The Potential of Semiconductor Diodes in High-Frequency Communications*

A. UHLIR, JR.†, ASSOCIATE MEMBER, IRE

Summary—**Graded *p-n* junctions that can be fabricated by solid-state diffusion are low-loss nonlinear capacitors at microwave frequencies. These diodes can be used to make low-noise amplifiers, amplifying frequency converters, harmonic and subharmonic generators, switches, limiters, and voltage-tuned passive circuits. Single junctions can control many watts of microwave power.**

Point-contact diodes are nonlinear resistors and as yet are unchallenged as microwave rectifiers. At lower frequencies, nonlinear-resistance action can be obtained in *p-n* junctions by introducing recombination centers.

A *p-i-n* diode is resistive at high frequencies. The value of the resistance depends upon the dc current. This variable resistance can be used as a broad-band microwave switch or attenuator. At low current densities, the *p-i-n* structure functions as a transmission line and so can serve as a support, protection, and connection for small-area *p-n* junctions made in the same single crystal of silicon.

I. Introduction

THE need for microwave crystal diodes stimulated early work on germanium and silicon. In particular, germanium of fair purity resulted and was an essential ingredient in the discovery of the transistor. Concepts created in the work on microwave diodes showed how to use the periodic table in further development of transistor materials.

It is now time for transistor technology to repay its debt to the microwave art. Single-crystal germanium and silicon are now available. New methods of making *p-n* junctions have been developed, such as solid-state diffusion of impurities. The *p-n* junction concept has been laid down and subjected to extensive (though by no means exhaustive) analysis. How can these developments be put to microwave use?

One answer to this question is that the *p-n* junction itself will be a valuable microwave circuit element. Its most spectacular use will probably be in low-noise microwave amplifiers. This and other uses of junction diodes are discussed. Also, an attempt is made to explain why junction diodes have not replaced point-contact diodes as microwave rectifiers and why the latter have not been greatly improved by transistor technology.

A major theme of this paper is that diodes differ qualitatively in their electrical characteristics. For a given circuit function, a particular type of electrical characteristic will generally be optimum. An introduction to three diode types will be given in Section II. Two types, the nonlinear capacitor and nonlinear resistor, can be described by simple equivalent circuits with frequency-independent elements; structures and physical mechanisms will be considered in Sections IV and V. Another type, the *p-i-n* structure discussed in Section VI is in a certain sense a variable resistor. Various circuit uses of diodes will be considered to determine which type of diode is most suitable. Rectification is discussed in Section VII. Frequency converters (Section VIII) may use nonlinear resistor or nonlinear capacitors; if the latter, amplification is possible. The control of microwave power by diode switches is discussed in Section IX. Sections X and XI deal with harmonic and subharmonic generation. The use of the nonlinear capacitor as a passive, electronically-variable capacitor is considered in Section XII.

II. Classification of Diodes

What is here called a nonlinear resistor is the ordinary conception of a rectifier, a diode that can convert ac power into dc power. The argument of the entire paper can be anticipated by suggesting that a rectifier may not be the right type of diode to use in frequency converters, switches, harmonic generators, and other circuits where rectification is not a specifically desired function.

For an ideal nonlinear resistor, the instantaneous current $i(t)$ is a function only of the instantaneous voltage $v(t)$. Practically, a device might be called a nonlinear resistor if, for the intended frequency range, its action can be analyzed approximately by assuming $i(t) = f(v(t))$. The anticipated relation between current and voltage is the familiar sort shown in Fig. 1. Generally, the polarity of the applied voltage makes a vast difference in the current. The polarity that gives a large current is called forward bias, the opposite, reverse bias.

An analogous description can be made of the nonlinear capacitor. The operation of such a device can be analyzed approximately by assuming that the instantaneous charge on the device is a single-valued nonlinear function of the instantaneous voltage applied to each terminal. The charge-voltage characteristic for a *p-n* junction is indicated in Fig. 2. The slope dQ/dV of this curve is the small-signal capacitance, which, like the charge Q, may be regarded as a function of the instantaneous voltage.

A time-varying capacitance can amplify. This principle has been utilized for a long time in vibrating-reed (or "dynamic capacitor") electrometers for amplifying low-frequency voltages from very high impedance sources. In Section VIII it is shown that analogous ef-

---

* Original manuscript received by the IRE, April 12, 1958. Based in part on investigations supported by the U. S. Army Signal Corps on contract DA 36-039 sc-73224.

† Bell Telephone Labs., Inc., Murray Hill, N. J.

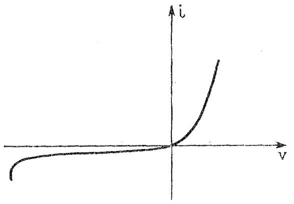


Fig. 1—Typical current-voltage characteristic of nonlinear resistor.

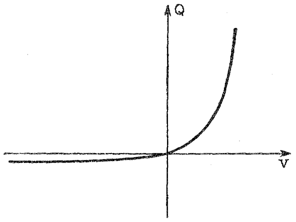


Fig. 2—Typical charge-voltage characteristic of nonlinear capacitor.

fects can be obtained at microwave frequencies by applying a time-varying voltage to a nonlinear capacitor, thus producing a time-varying capacitance. The direct analog of the vibrating-reed electrometer is the amplifying up-converter. Negative resistance amplification can also be obtained from time-varying capacitance and is the basis for a low-noise 6000-mc amplifier that is mentioned in Section VIII.

Amplification can similarly be obtained from nonlinear inductance. The name "varactor" has been proposed for any device whose operating principle is nonlinear reactance.[1] If this name gains currency, one would speak of nonlinear capacitor amplifiers, as well as amplifiers using saturable reactors, as varactor amplifiers. At present, the name "parametric amplifier" is often used. "Reactance amplifier" seems to be a much more descriptive term.

Another class of diodes seems to deserve the name "variable resistor." (It is regrettable that this term has been previously used as a synonym for nonlinear resistor.[2]) The structure is a layer of high-resistivity semiconductor with heavily-doped *p* and *n* regions on either side. A symbol for this structure is *p-i-n*. The shunt capacitance per unit area is remarkably small and the impedance at high frequencies is essentially resistive. The value of the resistance can be varied over a large range, being very high for reverse dc biases and very low when current flows through the diode in the forward direction. However, one cannot vary the resistance at a microwave rate, as is possible with certain nonlinear resistors (point-contact microwave diodes).

III. Impedance Measurements

The diode classifications outlined above define diodes by their two-terminal lumped-element characteristics. At high frequencies, ordinary components and transmission lines have dimensions that are appreciable compared to the wavelength, so that lumped-element concepts cannot be applied to them. However, the active regions of semiconductor diodes are generally small enough so that a lumped-element definition is tenable.

A suggestion by Waltz[3] makes possible microwave measurements of the actual junction or contact impedance of semiconductor diodes. This technique has been especially helpful in the development of *p-n* junction nonlinear capacitors, and this problem is solved in the following. The diode is mounted in some kind of a waveguide holder or crystal mount for coaxial line. The transmission line is connected to some impedance-measuring device adapted to transmission lines, such as a slotted line. The objective is to find the impedance of the junction itself. To do this, one must find the parameters that define the electrical transformation from the junction to the transmission line. The basis for the measurement is to make impedance *standards* that, like the junction, are small compared to the wavelength. They must be placed in exactly the same configuration as the *p-n* junction or point contact. The impedance standards may be, for example, small area pressure contacts to carbon of the type used in carbon resistors or pencil leads. (In evaluating gold-bonded germanium diodes, satisfactory "standard resistors" were made by bonding gold-gallium wire to *p* type germanium.) Two singular impedance standards are readily constructed, an open circuit and a short circuit.

The analysis of the data is not described here, except to note that it is particularly simple if the transformation between the diode and the transmission line can be regarded as lossless. Then the transformation at a single frequency may be represented by a length of line, a series reactance (or shunt susceptance), and an impedance transformation. The open circuit and short circuit determine length of line, and the product of the series reactance and the transformation ratio, so that relative junction impedances can be obtained without using a standard resistor.

If the diode is fabricated in some kind of a cartridge, one must use identical parts to fabricate the standard resistors. Care must be taken to insure that these resistors are inserted in the crystal mount in exactly the same way as the diode to be measured. The frequency limit of this technique depends upon the pains taken to insure equivalence of the electrical transformation when measurements are made on the standards and on the unknown. Dimensional tolerances should be held to within a few thousandths of a wavelength. Little difficulty is encountered in doing this at frequencies below 1000 mc.

IV. Nonlinear Capacitor Structures

The mechanism of *p-n* junction capacitance is described qualitatively, with mention of effects that can

---

[1] M. E. Hines, paper to be published.

[2] A. Uhlir, Jr., "Two-terminal *p-n* junction devices for frequency conversion and computation," Proc. IRE, vol. 44, pp. 1183–1191; September, 1956.

[3] M. C. Waltz, "A Microwave Resistor for Calibration Purposes," Bell Telephone Labs., Third Interim Rep. on Task 8 (Crystal Rectifiers), Signal Corps' Contract DA-36-039-sc-5589; April 15, 1955.

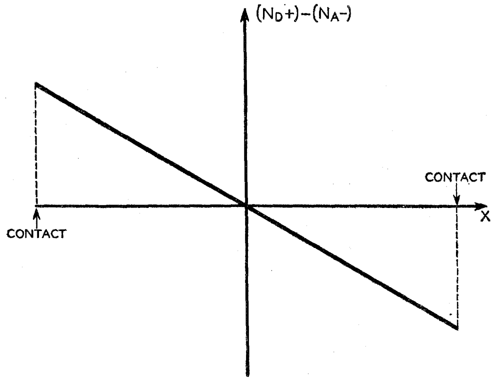


Fig. 3—Impurity distribution in linearly graded junction.

cause losses and lead to residual noise. Measurements on diffused silicon and welded-contact (gold-bonded) germanium diodes are given.

Generally speaking, semiconductor materials suitable for transistor fabrication exceed by a considerable margin the quality requirements for microwave nonlinear capacitors. *P-n* junctions are usually made by incorporating certain impurity atoms in the semiconductor lattice. These donor and acceptor atoms tend to ionize and become fixed positive and negative charges.

If the concentration of ionized donors is N_{D+} and the concentration of ionized acceptors is N_{A-}, the net charge density is $q(N_{D+}-N_{A-})$ and varies as shown in Fig. 3 for a linearly-graded junction, which will be used as an example. At zero bias the *p-n* junction is in equilibrium. For typical graded junctions, the concentration p of holes and n of electrons varies with distance as shown in Fig. 4(a). The place where the fixed charge density is zero is called the stoichiometric junction. It may be seen that in a small region around the stoichiometric junction, there are very few holes or electrons. This region is known variously as the depletion layer, exhaustion region, or space-charge region. The latter term refers to the fact that the net fixed charge is not neutralized by mobile carriers. Outside of the depletion layer the mobile carriers are present in almost exactly the right numbers to neutralize the the fixed charges. Evidently, holes are required to neutralize the negative fixed charges on the p side of the junction and electrons are required to neutralize the positive fixed charges on the n side of the junction.

Suppose the junction is biased slightly in the forward direction. This means applying to the contact on the p side a voltage that is positive with respect to the contact on the n side. This voltage will urge the hole and electron distributions to move toward each other, as shown in Fig. 4(b). For this motion to take place without leaving large unbalanced electric charges in the previously neutral p and n regions, it is necessary for additional holes and electrons to enter the semiconductor at the contacts. It is assumed that the contacts are of such a nature as to permit this. If a still larger forward voltage is applied, holes and electrons intermingle appreciably in a not-so-thoroughly-depleted layer and on either side of it [Fig. 4(c)]. Despite the intermingling, one can recover the charge if it is allowed to return in a time that is short compared to the time required for appreciable recombination. (Data are given below on a diffused silicon nonlinear capacitor to show that recombination is indeed negligible if the frequency is much higher than 1 mc.) When a substantial amount of the stored charge is thus represented by intermingled holes and electrons, one speaks of "carrier-storage capacitance."

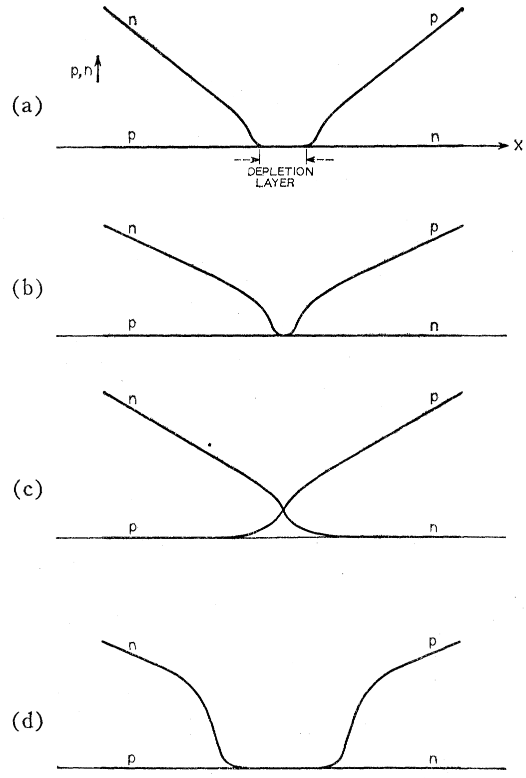


Fig. 4—Hole and electron concentrations in neighborhood of a graded *p-n* junction: (a) zero bias, (b) small forward, (c) large forward bias, (d) reverse bias.

When the junction is biased in reverse direction, the depletion layer widens as shown in Fig. 4(d). When the reverse voltage reaches the so-called "breakdown voltage," the junction begins to conduct copiously, usually because of generation of carriers in the depletion layer by a process known as avalanche multiplication. In the range between this reverse breakdown voltage and the slight forward voltage corresponding to Fig. 4(b), the charging and discharging of the junction takes place by the motion of hole and electron distributions toward and away from each other, without appreciable intermingling of holes and electrons. This type of capacitance is referred to as a depletion-layer capacitance. The thermally generated reverse saturation current of the junction in shunt with this capacitance produces full shot noise. The reverse current of clean silicon *p-n* junctions is so small that this shot noise is not important at the relatively low impedance levels of high-frequency circuits. Reverse currents of germanium *p-n* junctions cannot be so confidently neglected if the temperature is appreciably above room temperature.

The two loss mechanisms thus far mentioned—forward conductance, caused by recombination, and reverse leakage, caused by generation—produce shot noise that is practically independent of frequency. There is, in addition, the possibility of a frequency-dependent conductance arising from inability of the holes and electrons to redistribute themselves instantaneously in response to very rapid changes in applied voltage. This phenomenon has been called dispersion of the capacitance.[4] The dispersion of depletion-layer capacitance must occur at very high frequencies; it has never been observed experimentally. The fast response of the depletion-layer capacitance is attributed to the fact that the distances the holes and electrons must move are only small fractions of the depletion-layer width.[4] The dispersion of the storage capacitance has been one of the main topics in the theoretical and experimental study of *p-n* junctions.

The depletion-layer capacitance is probably the basis of the microwave diode amplifiers that will be described. One reason for this belief is a theoretical analysis of the effect of series resistance on the gain and noise of variable capacitance frequency converters.[5] For a capacitance varying sinusoidally with time, the theory shows that a six-to-one dynamic range of capacitance is all one might wish; a three-to-one range is almost as good. These modest dynamic ranges can be obtained with depletion-layer capacitance alone.

The depletion-layer capacitance of a linearly-graded junction is given as a function of voltage v by the approximate formula[4]

$$C \approx \frac{C_0}{\sqrt[3]{1-(v/\phi)}} \tag{1}$$

where ϕ is a constant that depends upon the impurity gradient but is about one-half volt for most silicon junctions. This formula is quite accurate for reverse biases (negative values of v) and holds reasonably well for moderate forward biases if ϕ is chosen empirically. A corresponding formula for abrupt junctions is

$$C \approx \frac{C_0}{\sqrt{1-(v/\phi)}}\,\cdot \tag{2}$$

It is sometimes suggested that the square-root relation (2) is "more nonlinear" than the cube-root relation, (1), and that the abrupt junction therefore is preferable to the graded junction. This argument is meaningless when not referred to a specific application and is probably false under most circumstances. The ease with which low series resistance can be obtained in graded junctions appears a decisive advantage in their favor.

Since the dynamic range of the capacitance is more than adequate, the parameters determining the gain and noise potentialities of the diode are, as shown in Fig. 5, the series resistance R_s and the minimum capacitance $C_{\min}$, the latter being the capacitance for reverse voltages just short of breakdown. Then, if an arbitrary impedance level is permitted, a single figure of merit will serve to describe the diode. This figure of merit may be written as a *cutoff frequency* f_c, defined by

$$f_c = \frac{1}{2\pi R_s C_{\min}}\,\cdot \tag{3}$$

Diffused silicon nonlinear capacitors with cutoff frequencies typically 60 to 120 kc, are being made from graded junctions with impurity gradients at the junction of 10^{23} to 10^{24} cm^{-4}.[6] The preferred impurity distribution for minimum series resistance is one in which the impurity gradient is everywhere as large as, or larger than, the impurity gradient at the stoichiometric junction.

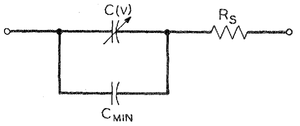


Fig. 5—High-frequency equivalent circuit of nonlinear capacitor.

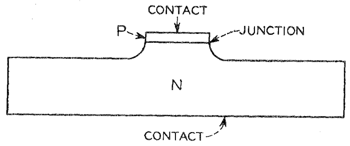


Fig. 6—*P-n* junction "mesa" diode.

The series resistance is inversely proportional to area and the capacitance is proportional to area, so the cutoff frequency is independent of area. In other words, the graded junction is a "planar" or "one-dimensional" structure. In this respect it differs from most high-frequency diodes, such as point-contact and welded-contact diodes, which must have small contact size for good high-frequency performance (because their series resistances vary inversely as the contact diameter). It is also a point of difference from the transistor triode, which requires two-dimensional flow of majority and minority carriers in the base layer.[7]

Apart from the series resistance, the diffused silicon diodes indeed appear to be voltage-dependent capacitors up to microwave (perhaps much higher) frequencies. The purely capacitive impedance should have no shot noise, and the shot noise of the reverse current is usually negligible. As a tentative hypothesis, it is suggested that the series resistance exhibits thermal noise. Then the cutoff frequencies that have been obtained lead one to expect microwave and UHF diode amplifiers to have noise figures lower than those now obtained with the best electron tubes. Preliminary evidence of low noise is given in Section VIII.

---

[4] W. Shockley, "The theory of *p-n* junctions in semiconductors and *p-n* junction transistors," *Bell Sys. Tech. J.*, vol. 28, pp. 435–489; July, 1949.

[5] D. Leenov, "Gain and noise figure of a variable capacitance up-converter," *Bell Sys. Tech. J.*, vol. 37; July, 1958.

[6] A. E. Bakanowski, N. G. Cranna, and A. Uhlir, Jr., "Diffused silicon and germanium nonlinear capacitors," presented at the IRE-AIEE Semiconductor Device Research Conference, Boulder, Colo., July, 1957.

[7] A. Uhlir, Jr., "Shot noise in *p-n* junction frequency converters," *Bell Sys. Tech. J.*, vol. 37; July, 1958.

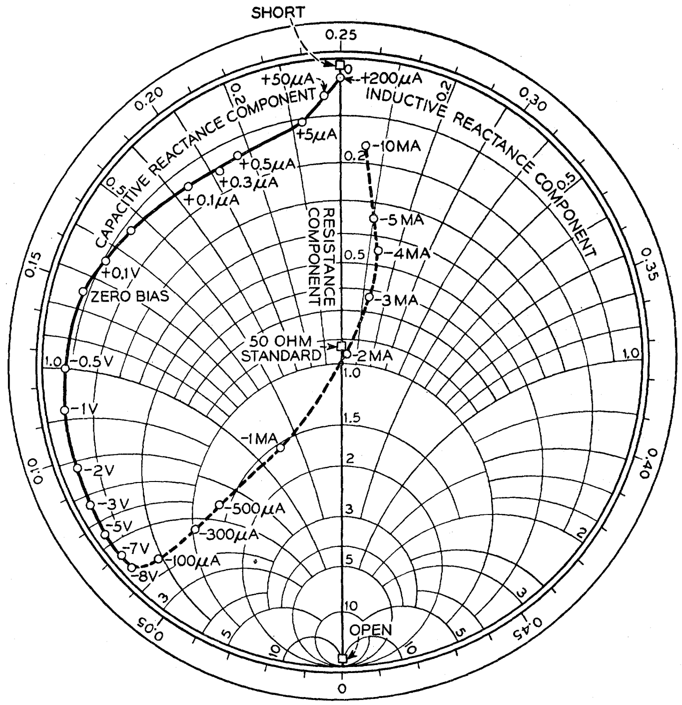


Fig. 7—Small-signal impedance of diffused silicon nonlinear capacitor at 1000 mc. Broken line is reverse breakdown region.

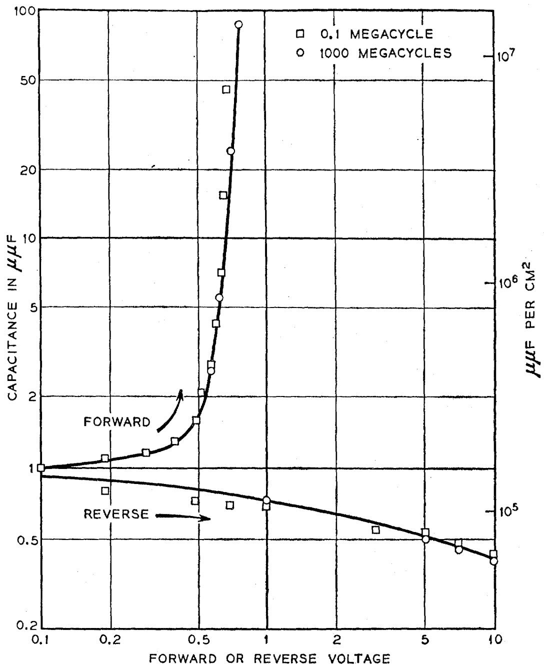


Fig. 8—Capacitance of a diffused silicon nonlinear capacitor as a function of voltage.

The diffused silicon "mesa-type" p-n junction structure is shown in Fig. 6. The 1000-mc small-signal measurements on a diode of this type are given in Smith chart form in Fig. 7. The solid curve represents the nonlinear capacitance bias range; the broken line shows the impedance when breakdown current is flowing. Some points corresponding to impedance standards are shown on the chart. The capacitance-voltage relation obtained from this chart is compared in Fig. 8 with the same relation determined by measurements at 100 kc. The resistance-reactance trajectory in rectangular coordinates (simply a transformation of the Smith chart) is shown in Fig. 9. The justification for the equivalent circuit of Fig. 5 is evident from Fig. 9 and the frequency-independence of capacitance shown in Fig. 8. Substantial reductions in series resistance have been realized since these graphs were prepared. Fig. 10 shows, as a function of bias, the crossover frequency at which the junction susceptance equals the junction conductance (measured at 100 kc). Evidently, the frequency-independent part of the conductance is negligible if the operating frequency is much higher than 1 mc.

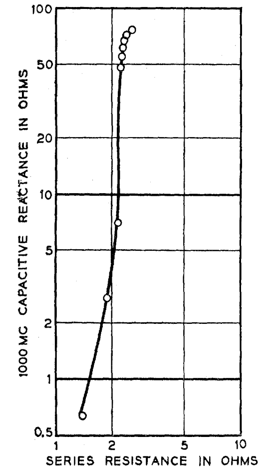


Fig. 9—Small-signal impedance of diffused silicon nonlinear capacitor, for various dc biases.

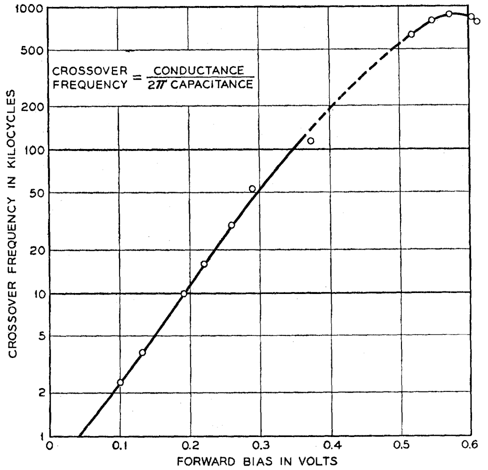


Fig. 10—Frequency at which conductance and susceptance are equal. Based on 100-kc measurements of 10^{-3} and 10^{-5} cm² areas of diffused silicon junction of gradient 5×10^{23} cm^{-4}.

Silicon p-n junctions are marketed by several concerns for use as electronically-variable capacitors in the VHF range. These devices do not have and do not require the high cutoff frequencies of the diodes intended for low-noise microwave amplifiers. Previously, experimental nonlinear capacitors made by alloying indium to n-germanium[8] had been tested in similar applications ($e.g.$, frequency control).

Historically, the first semiconductor diodes reported to give amplification were welded-contact germanium

---

[8] L. J. Giacolletto and J. O'Connell, "A variable-capacitance germanium junction diode for VHF," *RCA Rev.*, vol. 17, pp. 68–85; March, 1956.

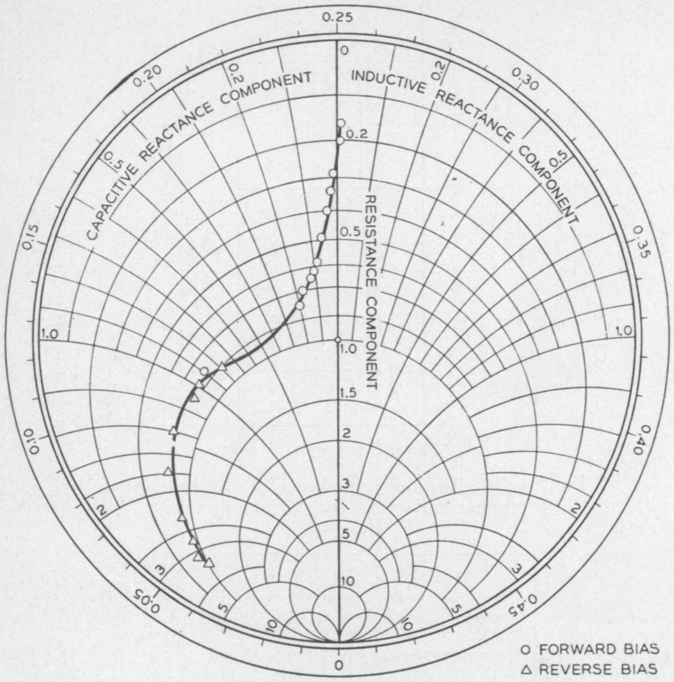


Fig. 11—Small-signal impedance of gold-bonded germanium diode at 9 kmc. Chart center is 60 ohms.

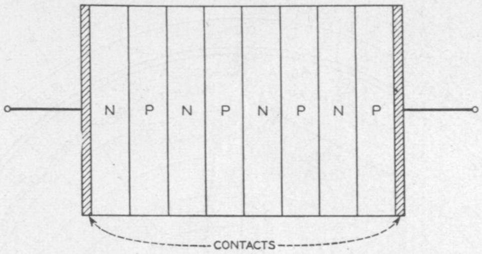


Fig. 12—Multiple-junction diode.

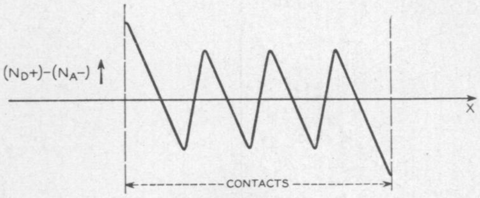


Fig. 13—Impurity distribution for nonlinear capacitance in a multiple-junction diode.

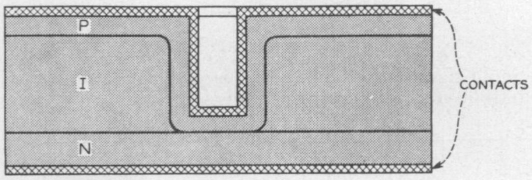


Fig. 14—Section through dimple diode made by diffusion in silicon.

diodes.[9] A modern version of this type of diode, a gold-bonded germanium diode, has been designed for use in a microwave relay transmitting modulator.[10] The cutoff frequency is 40 kmc, which is more than adequate for an amplifying modulator between 60 to 80 mc and 6000 mc. A Smith chart plot of the 9-kmc small-signal admittance of one of these diodes is given in Fig. 11.[11] For large forward currents, the resistance decreases and acquires a small inductive component; these effects are attributed to conductivity modulation of the series resistance.

While the bonded diodes are economical and practical, future developments will doubtless be predicated on the demonstrated success of the diffused junction nonlinear capacitor. The impedance level of a large-area *p-n* junction is inconveniently low for high-frequency circuits. A way of building up the impedance is to put a number of such junctions in series. The power-handling capability is increased, first, by the large junction areas and, second, by the multiplicity of junctions.

A series stack can be approximated in a single crystal of semiconductor without intervening metallic contacts (which might add resistance). Consider a single crystal with alternate layers of *n*- and *p*-type semiconductor, as shown in Fig. 12. One readily obtains nonlinear capacitance with the impurity charge distribution shown in Fig. 13. The *p-n* junctions in this structure are graded considerably more steeply than the *n-p* junctions. The steeper junctions have higher capacitance per unit area and may be regarded approximately as ac short-circuits. The equivalent circuit of the structure, then, is a series connection of the nonlinear capacitances of the more gradual junctions.

Another ramification is the production of a graded *p-n* junction embedded in a *p-i-n* diode, as shown in Fig. 14.[12] This dimple structure is resistant to atmospherically-induced changes in capacitance or breakdown voltage and can safely dissipate more power than equivalent mesa diodes. It is a way of contacting and handling very small *p-n* junctions.

In most diodes, reverse breakdown due to avalanche multiplication occurs at a number of localized discharges, each of which is called a microplasm. A large fraction of the dimple diodes break down at just one microplasm. When this microplasm turns on (starts to pass reverse current), a microwave transient is generated—an effect which appears to be the first observed conversion of dc power into microwave radiation by a *p-n* junction.[13]

A *p-i-n* diode of considerable linear extent should act as a transmission line. A sequence of *p-n* junctions in this kind of transmission line, as in Fig. 15, is a traveling-wave diode amplifier in a single piece of silicon.

---

[9] H. C. Torrey and C. A. Whitmer, "Crystal Rectifiers," McGraw-Hill Book Co., Inc., New York, N. Y., ch. 13; 1948.

[10] "Semiconductor diodes yield converter gain," *Bell Labs. Rec.*, vol. 35, p. 412; October, 1957.

[11] D. Leenov, private communication.

[12] N. G. Cranna and A. Uhlir, Jr., paper in preparation.

[13] J. L. Moll, A. Uhlir, Jr., and B. Senitzky, "Microwave transients from avalanching silicon diodes," to be published in PROC. IRE.

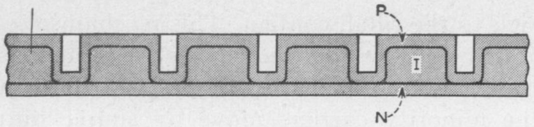


Fig. 15—Section through p-i-n transmission line with integral p-n junctions.

V. NONLINEAR RESISTOR STRUCTURES

Nonlinear resistors fall into two groups. P-n junctions can be used at frequencies up to several hundred megacycles. Point-contact diodes are used at higher frequencies, including millimeter waves.[14]

Recombination processes are necessary for nonlinear resistance action in p-n junctions. It is suggested that point-contact diodes are p-n junctions in a broad sense and also require recombination. The increasing variety of semiconductor materials being used for microwave point-contact diodes is noted. Finally, the possibility of of making nonlinear resistors that do not depend upon recombination is considered.

The forward current in a p-n junction is maintained by recombination of holes and electrons. One is concerned with the situation like that shown in Fig. 4(c) in which the hole and electron distributions overlap appreciably. For each electronic charge that flows in the external circuit in the forward direction, one hole and one electron must recombine. This recombination can take place in the neutral n or p-regions or in the depletion layer.

The graded p-n junction has one effect that tends to make it a difficult structure in which to obtain nonlinear-resistance action. The built-in field in the neutral p and n regions is such as to make any minority-carrier admittance capacitive rather than resistive. However, if nonlinear-resistance action can be obtained in spite of this effect, the graded junction will be very desirable because of the low series resistance relative to the depletion-layer capacitance.

To enhance recombination (a practice referred to as "ruining lifetime"), impurities may be added to the crystal to serve as catalysts for the recombination process; gold in silicon is an example.[15,16] Also, to lower lifetime, the geometrical arrangement of the lattice may be rendered imperfect by mechanical abuse at room temperature or elevated temperatures, or by irradiation with electrons, neutrons, etc. A sandblasted surface near the p-n junction is used in one experimental germanium diode.[17]

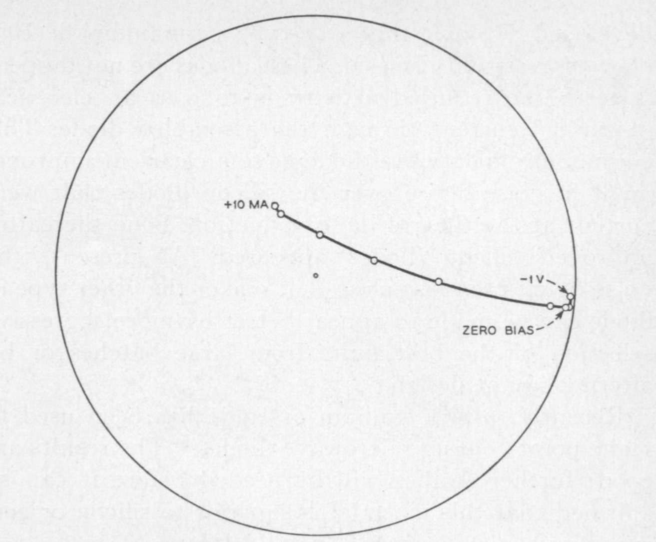


Fig. 16—Small-signal impedance of 1N23B silicon point-contact diode. Presented as 1000-mc reflection coefficient when mounted in commercial crystal mount at end of 50-ohm line.

Increased understanding of recombination processes may eventually lead to microwave p-n junction nonlinear resistors. In the meantime, point-contact diodes will serve as microwave nonlinear resistors, as shown by the measurements in Fig. 16. The physical structure of the point contact (to say nothing of the physical mechanism) is conjectural, for reasons that will be apparent from a brief description of processes used in fabrication.

Silicon point-contact diodes are made from p-type silicon doped to low resistivity (10^{-2} ohm-cm) with boron or aluminum. Lately, aluminum is preferred on the basis of empirical evidence that it gives better burnout resistance.[18] "Burnout" refers to any impairment of diode performance by electrical overload.

Point-contact diodes made in single-crystal silicon with controlled resistivity are more uniform than,[19] but are otherwise similar to, diodes made from polycrystalline silicon. High minority carrier lifetime in the starting material is not desired for obtaining nonlinear resistor action. Neither is it disadvantageous, because the effective lifetime at the contact is probably determined by mechanisms, speculated on below, for which the original lifetime is irrelevant.

The silicon is sliced and given a heat-treatment which may increase the resistivity of a thin surface layer. A high-resistivity surface layer would decrease the capacitance per unit area without proportionately increasing the series resistance and would assist in removing carriers from the contact according to one mechanism to be described. The diode is assembled by bringing a sharp tungsten point in contact with the surface. Good rectification is not obtained until the contact is mechanically disturbed, for example, by tapping the assembled unit with a small hammer.

Low-noise germanium point-contact diodes have been

---

[14] W. M. Sharpless, "Water type millimeter wave rectifiers," *Bell Sys. Tech. J.*, vol. 35, pp. 1385–1402; November, 1956.

[15] G. Bemski, "Recombination in semiconductors," this issue, p. 990.

[16] A. E. Bakanowski and J. H. Forster, paper in preparation.

[17] R. H. Rediker and D. E. Sawyer, "Very narrow base diode," PROC. IRE, vol. 45, pp. 944–953; July, 1957.

[18] E. J. Feldman, "Improved S-Band Crystal Diodes," Microwave Crystal Rectifier Symposium Record, Fort Monmouth, N. J., p. 196; February, 1956.

[19] J. H. Bollman, "Use of Single-Crystal Silicon in Microwave Varistors," Bell Telephone Labs., First Interim Rep. on Improved Crystal Rectifiers, Signal Corps' Contract DA36-039-sc-73224; May 15, 1957.

developed.[20] Single-crystal *n*-type germanium of 10^{-2} ohm-cm resistivity is used. These diodes are not tapped; instead, the required artistry is to pass an electrical "forming" current through the assembled diode. This germanium microwave diode gave a clear-cut improvement in noise figure over the silicon diodes that were current at the time of its introduction. Soon thereafter improved silicon diodes appeared. At present, the noise-figure race is so close that one or the other type of diode can be made to appear better by more aggressive selection of the best units from large batches or by altering circuit design.

Recently, *n*-type gallium arsenide has been used to make point-contact microwave diodes.[21] The results are good; further studies will be needed before it can be affirmed that this material is superior to silicon or germanium.

Point-contact diodes for use as microwave rectifiers are made by much the same techniques as the diodes for superheterodyne use. However, the requirement of high impedance, to be explained in Section VII, generally leads the designer to use smaller contact areas in diodes for rectifier use.

The point contact may be regarded as a kind of *p-n* junction if one realizes that the surface of a semiconductor is likely to have an electric charge.[22] The magnitude and even the sign of this charge depends upon the surface treatment, but under given conditions may be fairly constant. The surface charge may be regarded as fixed in comparison to the mobile carriers. If the surface charge is opposite in sign to the fixed charge due to impurities in the bulk of the semiconductor, a kind of *p-n* junction results. It seems beyond doubt that such a surface *p-n* junction exists at the emitter of a point-contact transistor made on *n*-type germanium, and at the emitter and collector contacts of a surface-barrier transistor. For transistor action to occur in either of these devices, it is necessary that the surface charge be quite strong. Then, in a layer of semiconductor just beneath the surface, there will be a high concentration of carriers neutralizing the surface charge. Forward current flows by injection of these carriers into the bulk material, where they are minority carriers. Nonlinear resistance action requires some mechanism that prevents the return of these minority carriers. It is questionable that the initial bulk minority-carrier lifetime is short enough to accomplish this automatically, even in polycrystalline semiconductors.

One possible mechanism for nonlinear resistance in a point-contact diode with carrier injection is simply the difficulty the injected carriers may have in finding their way back to the small contact. This mechanism cannot explain why rectification frequencies are as high as they are, if one assumes (as in ordinary *p-n* junction theory) that the minority carriers move by simple diffusion. However, the flow of forward current is accompanied by an electric field that hastens the departure of minority carriers from the neighborhood of the contact. For this field to be adequate, the resistivity very near the contact must be higher than the bulk resistivities that are used. Heat-treatment or forming may very well produce such a resistivity alteration.

Carrier injection into the bulk semiconductor may not be important in microwave point-contact diodes, which are made on bulk materials of much lower resistivity than preferred for transistors. It is possible that most of the forward current is carried by carriers moving from the bulk material into the surface layer, there to recombine.

The author has constructed a theory of shot noise in *p-n* junction frequency converters.[7] It should also be applicable to point-contact nonlinear resistors whether the carriers are emitted from surface to bulk or vice versa. The theory shows the importance of local-oscillator waveform in determining the noise figure of nonlinear resistor superheterodyne circuits.

Future progress in nonlinear resistors may employ structures that use collection, rather than recombination, to remove carriers. For example, *n-i-n* and *n-p-n* diodes are nonlinear resistors up to frequencies which compare with the transit time for an electron to go from one *n* region to the other. If symmetrical, such structures could not function as passive rectifiers, but should be usable frequency converters.

Another proposal is the "drift diode," with the impurity distribution shown in Fig. 17.[23] The impurity gradient near the junction gives rise to an electric field, even under equilibrium conditions, that is in such a direction as to discourage the return of injected carriers.

VI. *P-I-N* DIODES

When the first *p-n* junctions became available, everyone who studied them was impressed by their superb low-frequency rectification characteristics, compared to the previously available point-contact diode. But some device engineers were not content and proposed the *p-i-n* structure,[24,25] shown in Fig. 18(a). The symbol *I* stands for intrinsic or high-resistivity semiconductor. The intrinsic layer gives a very much larger breakdown voltage than can be obtained in simple *p-n* junctions. The somewhat more surprising feature, which makes the structure an excellent power rectifier, is that in forward bias the intrinsic region is filled with injected car-

---

[20] G. C. Messenger and C. T. McCoy, "Theory and operation of crystal diodes as mixers," PROC. IRE, vol. 45, pp. 1269–1283; September, 1957.

[21] D. A. Jenny, "A gallium arsenide microwave diode," PROC. IRE, vol. 46, pp. 717–722; April, 1958.

[22] J. Bardeen and W. H. Brittain, "Physical principles involved in transistor action," *Phys. Rev.*, vol. 75, pp. 1208–1223; April 15, 1949.

[23] C. H. Knowles, "Characteristics of the Drift *P-N* Junction," presented at the IRE-AIEE Semiconductor Device Research Conference, Purdue, Ind., 1956.

[24] R. N. Hall, "Power rectifiers and transistors," PROC. IRE, vol. 40, pp. 1512–1518; November, 1952.

[25] M. B. Prince, "Diffused *p-n* junction silicon rectifiers," *Bell Sys. Tech. J.*, vol. 35, pp. 661–684; May, 1956.

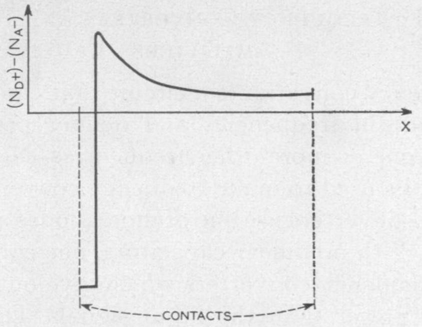
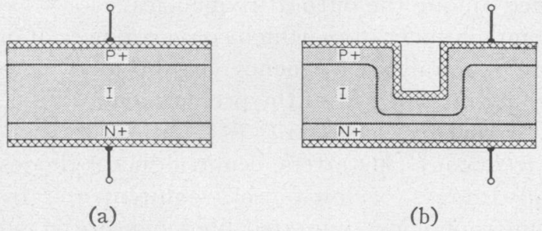


Fig. 17—Impurity distribution in drift diode.



(a) (b)

Fig. 18—(a) *P-i-n* diode. (b) Dimple structure of *p-i-n* diode with small effective area. In example, dimple is 5 mils diameter, thickness at bottom of dimple is about 3 mils.

riers and no longer has a high resistivity. Therefore, the forward drop at large currents is moderate. While this structure does its intended job of rectifying power frequencies such as 60 cycles per second or 400 cycles per second, it becomes a poor rectifier at frequencies as low as a megacycle (depending upon the thickness of the intrinsic region).

As long as thinking about microwave diodes revolved about the rectifier, there was little inclination for anyone to place in microwave circuits a device that could not even rectify one megacycle. Moreover, the low-frequency capacitance of these diodes was of the order of 20 mmf and up; what could be more absurd than to put such a device in a high-frequency circuit?

But, let us consider the *p-i-n* diode shown in Fig 18(b). To be sure, the dimensions of this diode are slightly smaller than those customarily used in the smallest power rectifiers, but they are still of the same order of magnitude and are enormous compared to the dimensions of the active area of point-contact microwave rectifiers; the zero-bias capacitance at 100 kc is 13.8 mmf. The raw data of a 1000 mc measurement of this diode is shown in the Smith chart of Fig. 19.[26] The parameter that is varied is the dc bias. Also shown are the measurements made on a short circuit and an open circuit constructed in the same diode cartridge and mounted in the same crystal holder. At zero bias and reverse biases the impedance is extremely high compared to the chart-center impedance of 45 ohms, while at moderately large forward currents the impedance of the diode is very small compared to 45 ohms. The effective shunt capacitance in reverse bias is only 0.3 mmf.

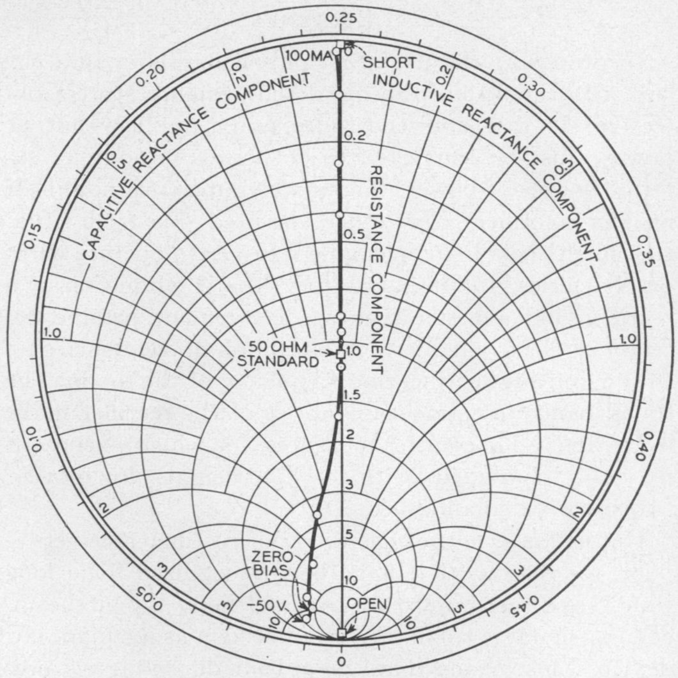


Fig. 19—Small-signal 1000-mc impedance of *p-i-n* dimple diode of Fig. 18(b).

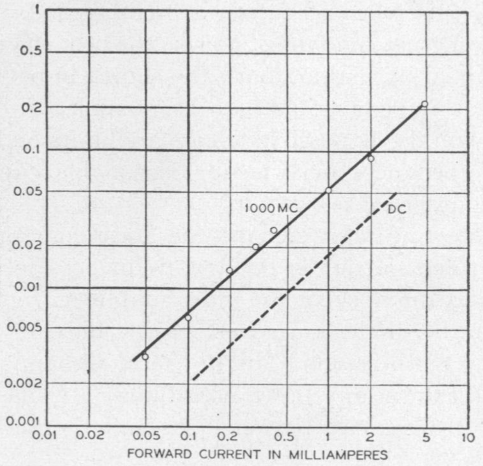


Fig. 20—Conductance of *p-i-n* dimple diode.

The value of the conductance is plotted as a function of the dc current in Fig. 20. The dashed line in this figure gives the approximate value of the dc conductance at corresponding dc currents. At high frequencies, one observes only the conductivity-modulated resistance of the intrinsic region. Accordingly, the high-frequency conductance is larger than the dc conductance.

Evidently the *p-i-n* diode can be used as an electronically-variable attenuator of microwave frequencies. The variable resistivity of an intrinsic region can also be utilized in distributed structures to provide variable attenuation;[27] the limit in this direction would be a *p-i-n* transmission line with variable attenuation.

[26] Fabricated by N. G. Cranna; measured by D. E. Iglesias.

[27] E. M. Gyorgy and G. L. Pearson, private communication.

VII. RECTIFIERS

In proposing the use of a nonlinear resistor in a circuit, one should keep in mind the ultimate system objective. If this objective necessarily implies that ac power must be converted to dc power without an external source of power, then an asymmetrical nonlinear resistor (rectifier) must indeed be used. One such situation might be a light, portable receiver to operate without batteries. In the laboratory, the combination of a rectifier and a meter makes a convenient passive detector of electromagnetic radiation of all frequencies up to the limiting rectification frequency of the diode. The broad-band detection capabilities of the rectifier make it attractive for counter-measures, for which purpose it may be combined with a lightweight, low-power, transistor video amplifier.

Unfortunately, the sensitivity of rectifier receivers is poor, for reasons that have been recognized for a long time,[28] and just briefly are outlined here. When the incoming signal is small, the rectifier acts as a square law device. This means that the output dc voltage is proportional to the square of the RF voltage. Accordingly, the efficiency of rectification is proportional to RF power and decreases when the available power decreases. Just when efficiency is most needed, it is lacking. For a given amount of power, the best efficiency can be obtained by making both the source impedance and the diode impedance as high as possible, to obtain as large a voltage as possible. There are limits on how high these impedances may be for reasonable circuits and diode contact areas.

In detecting faint signals, the noise competing with the rectified output is the low-frequency noise of the diode and the noise of the video amplifier. In the usual receiver the diode is used at zero dc bias; it would be contrary to the second law of thermodynamics for the diode to exhibit any but thermal noise. (A biased diode can have any noise between one-half thermal and infinity.[29])

Theoretically, improved sensitivity can be obtained by lowering the temperature, for two reasons. The important reason is the improvement in efficiency of rectification, for a given input power. This effect may be predicted from the theoretical rectifier characteristic

$$i = i_s\{ e^{qv/kT} - 1) \qquad (4)$$

which shows that the nonlinearity improves as temperature is lowered. The other reason is the reduction of the thermal noise of the diode, but this reduction is of limited advantage unless the noise in subsequent amplifiers can be correspondingly reduced. Specially designed diodes would be required for very low temperature use.

VIII. FREQUENCY CONVERTERS AND DIODE AMPLIFIERS

A frequency converter is a circuit that can accept signals at certain frequencies and deliver proportionate signals at one or more other frequencies. Nonlinear devices may be used to make frequency converters. Many frequency converters use one or more diodes as nonlinear elements. With nonlinear capacitors, one can make amplifying frequency converters which give output signals of greater power than the input signals. Diode amplifiers can be derived as special amplifying frequency converters that have the input frequency or frequencies included among the output frequencies.

Harmonic generators, which convert power of one frequency to a multiple frequency, are discussed in Section X and are excluded from the present connotation of the term "frequency converter."

As frequency converters, semiconductor diodes have the advantages of low-power requirements, freedom from microphonics, unlimited life when protected from overloads, and low cost. Moreover, the nonlinear capacitor is a low-noise or medium-power high-frequency amplifying device, with all of the above advantages and vastly improved resistance to electrical burnout, compared to point-contact diodes.

A circuit element whose value varies with time is a frequency converter. A picturesque representation of such an element might be a rotary variable capacitor driven by a large motor. The capacitance—the ratio of charge to voltage—is then a periodic function of time. Another type of such element would be a rheostat whose slider was made to move periodically by some mechanical drive. Then the resistance— the ratio of voltage to current—is a periodic function of time. Such conceptual objects would be linear frequency converters. Their efficiencies and impedances would not depend upon the magnitude of the impressed electrical signals (except that the electrostatic forces between the capacitor plates should not be large enough to react upon the motor drive).

A nonlinear capacitor or a nonlinear resistor can be made to imitate the mechanically driven frequency converters. These imitations may be used at microwave frequencies. The scheme is to apply a relatively large periodic voltage to the nonlinear element. Then it is found that small signals applied at frequencies other than the fundamental and harmonics of the large voltage are converted in frequency just as if the device were a time-varying resistance or capacitance. The small-signal response of such a frequency converter is linear, as can be shown very simply. Suppose that $v(t)$ is written

$$v(t) = v_l(t) + v_s(t) \qquad (5)$$

where v_l is the large-signal part, periodic or not, and v_s

---

[28] Torrey and Whitmer, *op. cit.*, ch. 11.

[29] Torrey and Whitmer, *op. cit.*, ch. 6.

can be regarded as arbitrarily small. Then $f(v)$ can be expanded in a Taylor's series;

$$f(v(t)) = f(v_l(t)) + f'(v_l(t))v_s(t). \tag{6}$$

For the nonlinear resistor, $i = f(v)$ and it is natural to define a small-signal conductance $G = di/dv = f'(v)$. Then, from (6), one has

$$i(t) = i_l(t) + i_s(t) \tag{7}$$

where

$$i_l(t) = f(v_l(t)) \tag{8}$$

and

$$i_s(t) = G(t)v_s(t). \tag{9}$$

The last equation is the one of interest, because it shows that small variations in current are linearly related to small variations in voltage. In similar fashion one may define, for the nonlinear capacitor, $C = dQ/dv$ and obtain, for the small-signal charge Q_s,

$$Q_s(t) = C(t)v_s(t). \tag{10}$$

The above relations in the time domain show the underlying simplicity of the mathematical approach used to linearize the problem. The detailed operation of frequency converters can best be analyzed in the frequency domain. The procedures used in such an analysis will be outlined. For the nonlinear resistor, suppose that $G(t)$ is periodic with a fundamental frequency b. This condition could prevail in a mechanically-varied resistor, but, of course, here we are most interested in periodic $G(t)$ resulting from the application of a periodic large-signal beating-oscillator voltage $v_l(t)$ to a nonlinear resistor type of diode. One can write

$$G(t) = \sum_{n=-\infty}^{+\infty} G_n e^{2\pi jnbt}. \tag{11}$$

If the applied voltage is given by

$$v(t) = \sum_{s\pm} \sum_{m=-\infty}^{\infty} v(mb + s)e^{2\pi j(mb+s)t}, \tag{12}$$

the resulting current is

$$\begin{aligned} i(t) &= G(t)v(t) \\ &= \sum_{s\pm} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} G_n v(mb + s)e^{2\pi j((m+n)b+s)t} \end{aligned} \tag{13}$$

and may be seen to contain the same frequency components as the impressed voltage, that is, one can deal with a closed set of frequencies. The impressed voltages and currents are related by a conversion matrix. Thus

$$i(mb + s) = \sum_n Y_{mn}v(nb + s \tag{14}$$

where the conversion matrix is

$$Y_{mn} = G_{m-n}. \tag{15}$$

Exactly the same steps may be followed for the time-varying capacitance, with the result

$$Q(mb + s) = \sum_n C_{m-n}v(nb + s). \tag{16}$$

It is customary to treat current rather than charge as a variable in circuit analysis. For any frequency ν,

$$i(\nu) = 2\pi j\nu Q(\nu). \tag{17}$$

Thus

$$i(mb + s) = 2\pi j(mb + s) \sum_n C_{m-n}v(nb + s) \tag{18}$$

so that the conversion matrix is given by

$$Y_{mn} = 2\pi j(mb + s)Cm\text{-}n. \tag{19}$$

A frequency converter is made up of the periodically varied element, described by a conversion matrix, plus impedances terminating all of the frequencies $mb+s$. Having chosen these terminations, one may, in principle, determine the performance of the frequency converter by linear network analysis.

If any other set of frequencies $mb+s'$ is considered, another conversion matrix can be constructed. The linear problem for this new set of frequencies may be solved entirely separately from the first set of frequencies. Thus, a signal containing a spectrum of frequencies can be resolved into frequency components; separate analyses are conducted for each closed set of frequencies generated by the frequency components of the original signal.

When diodes that are neither nonlinear resistors nor nonlinear capacitors are used in frequency converters, they can still be described by conversion matrices for sets of frequencies $mb+s$. The elements of the conversion matrix are complex, in general. For example, a variety of conversion matrices arise from considering *p-n* junction action as nonlinear injection of minority carriers, followed by diffusion, drift, and recombination.[2]

The general conversion matrix involves an infinite number of frequencies, which usually makes an exact circuit analysis difficult. Most analyses neglect all but two or three frequencies, with no justification except the reasonableness of the results. An exception is an analysis of gain and noise in up-converter amplifiers,[5] in which a nonlinear capacitor is assumed having the equivalent circuit of Fig. 5. A way of treating the infinite number of frequencies is given which is logically consistent with the presence of series resistance.

The problem now is to obtain some general insights from the conversion matrices and any other considerations that can be applied. By the conversion matrix analysis of particular situations, one finds that amplifying frequency converters are possible with time-varying capacitance, but one soon suspects that a time-varying resistance cannot amplify. The truth of the latter surmise is easily demonstrated by the following argument. A mechanically-varied rheostat cannot have an electrical output that exceeds the electrical input. Therefore, a linear-for-small-signals frequency converter made with a nonlinear resistor cannot amplify, because it has exactly the same conversion matrix as a hypothetical mechanically-varied rheostat, as long as nonlinear resistors with negative resistance ($f'(v)<0$) are excluded.

The correspondence between the rheostat and the nonlinear resistor must not be carried too far. The rheostat can be an amplifier (of the mechanical signal). Also, noise in a rheostat is obviously thermal noise, while the nonlinear resistor mechanisms suggested above imply shot noise. Thus, a given $G(t)$ waveform has definite signal transmission properties but may produce different amounts of noise, depending upon the physical mechanism.

To obtain insight into the properties of linear frequency converters utilizing nonlinear capacitors, the analysis thus far is used only to suggest what signal frequencies should be considered. In the search for general principles, one turns naturally to conservation of energy within the (almost) lossless nonlinear capacitor. But this principle by itself is quite empty, for any reasonable amount of power will be cheerfully supplied from the beating oscillator (often called the "pump") if so doing produces the desired signal transfer.

What is needed is a "second law," like the second law of thermodynamics or the principle of conservation of momentum in dynamics. The required second law is given by Manley and Rowe in a general analysis of nonlinear reactances.[30] The results obtained for several simple and important types of frequency converters will be discussed.

Diagrams are invaluable for discussions of this kind, but no particular representations have been universally adopted. One obstacle to universality is that it is mathematically most convenient to use negative frequencies, so that Fourier series can be written, as in (7), with complex exponentials. Therefore, the transition to positive frequencies will be made explicit and finally a simplified diagram will be suggested.

Fig. 21 is a pictorial representation of the Fourier components of a periodic function such as the beating-oscillator voltage. Positive and negative harmonic frequencies are used. The complex amplitude of each component is the vector in the x-y plane perpendicular to the frequency axis at the corresponding frequency mb. Since the voltage is a real function of time, $v(-mb)$ is the complex conjugate of $v(mb)$.

Graphs like Fig. 21 are more complicated than necessary for a general discussion. In Fig. 22(a), the spectrum of Fig. 21 is schematized by replacing the two-dimensional vector by a vertical line whose length might be the magnitude of the vector; what matters is that the presence of a line denotes the existence of a vector. One could represent the vectors $v(mb+s)$ according to the same scheme. However, it is helpful to consider a hypothetical signal spectrum $mb+s$, $mb+s'$, $mb+s''$, $\cdots$ founded upon a perhaps infinite set of frequencies s, s', s'', $\cdots$, as indicated in Fig. 22(b). In the equations [*e.g.*,(8)], one also encounters the frequencies $mb-s$, shown in Fig. 22(c). Because the signal voltage is real, Fig. 22(c) does not represent any information not already represented in Fig. 22(b). Therefore, in calculations with conversion matrices, it is sufficient and advisable to consider either Fig. 22(b) or 22(c), but not both.

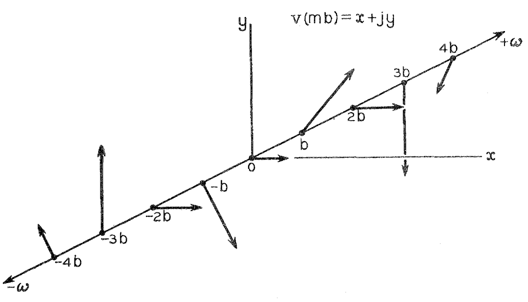


Fig. 21—A representation of the Fourier components of the local-oscillator voltage.

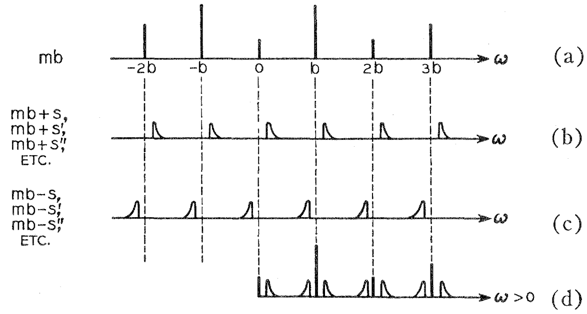


Fig. 22—Frequency components in a linear frequency converter: (a) local oscillator, (b) small-signal spectrum, (c) conjugate small-signal spectrum, (d) combined spectrum for positive frequencies.

In circuit design, as opposed to circuit analysis, there is much to be said for thinking only of positive frequencies. Then one combines Fig. 22(a) through 22(c), discarding negative frequencies, to obtain Fig. 22(d). Note that some of the signal spectra are inverted when measured in positive frequencies; these correspond to negative frequencies in Fig. 22(b) and are called "lower sidebands."

The possibility of using the sidebands around 2b, 3b, etc., is occasionally useful; the practice is known as harmonic mixing. However, the most important frequency converter applications use the three signal bands shown in Fig. 23(a). A diagram equivalent to Fig. 23(a) is shown in Fig. 23(b); the distinction between inverting and noninverting signal bands is indicated by small vertical arrows.

The first circuit example is the upper-sideband converter shown in Fig. 24(a). It is assumed that filters

---

[30] J. M. Manley and H. E. Rowe, "Some general properties of nonlinear elements—Part I. General energy relations," PROC. IRE, vol. 44, pp. 904–913; July, 1956.

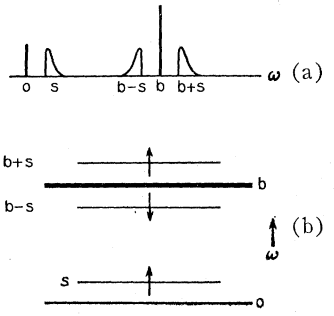


Fig. 23—(a) Frequencies involved in nonharmonic frequency conversion. (b) Diagram corresponding to (a); small arrows distinguish inverting and noninverting signal bands.

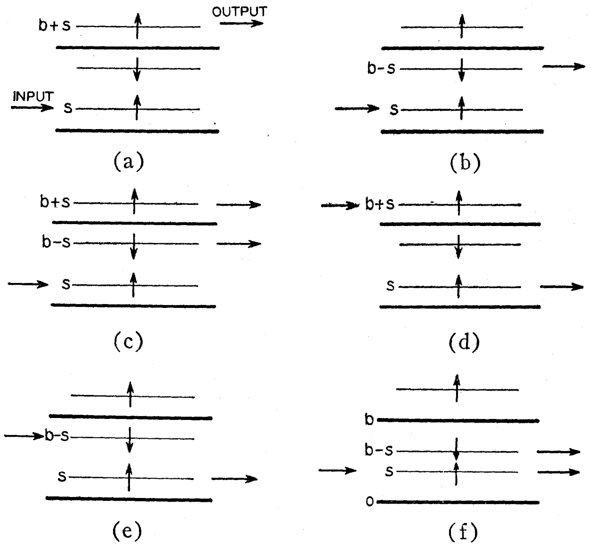


Fig. 24—Frequency relations in various types of frequency converters. (a) Upper sideband up-converter. (b) Lower sideband up-converter. (c) Double sideband up-converter. (d) Upper sideband down-converter. (e) Lower sideband down-converter. (f) Amplifier.

are used, if necessary, to prevent power leaving or entering the nonlinear capacitor at the lower sideband. Manley and Rowe[30] show that the power gain from input to output is exactly equal to the ratio $(b+s)/s$, and that the circuit is stable. The reduction of gain by series resistance has been analyzed;[5] the results can be expressed in terms of the ratios of the signal frequencies to the cutoff frequency. Some gain and bandwidth calculations on nonlinear capacitance up-converters have been discussed in the literature.[31] Another analysis puts forth the general principle that the 3-db bandwidth is about 40 per cent of the input frequency s, for upper-sideband operation.[32]

A lower-sideband up-converter is diagrammed in Fig. 24(b). Here, the power gain is $-(b-s)/s$. What does a negative power gain mean? As Manley and Rowe ex-explain, it means that unlimited amplification is possible, but instability is also possible: for certain terminations oscillations will occur and deliver power to the terminations at both $b-s$ and s. The greater the amplification, the less the bandwidth. There is no direct theoretical noise penalty associated with this "regenerative gain," since an ideal nonlinear capacitor is not a source of noise, no matter how operated.

Hines[1] has pointed out a remarkable difference between double-sideband up-converters [Fig. 24(c)] using nonlinear capacitors and those using nonlinear resistors. The Manley-Rowe relations show that unlimited gain is possible for the nonlinear capacitors. Moreover, for a given input power, the more power delivered to the upper sideband, the more delivered to the lower sideband. Contrast this to the miserly economy of a nonlinear-resistor up-converter, wherein, to obtain maximum output at one sideband, power must be reflected at the other sideband.

An upper-sideband down-converter, shown in Fig. 24(d), produces a stable power loss of $s/(b+s)$ when a nonlinear capacitor is used, so nonlinear resistors are preferred in this type of circuit. The lower-sideband down-converter shown in Fig. 24(e) can give unlimited gain, but is not a generally satisfactory circuit when large frequency ratios are involved. The basic reason is the following situation. For given terminations, the gain in such a down-converter is less by $(s/(b-s))^2$ than the gain in the reverse direction; *i.e.*, with the circuit operating as a lower-sideband up-converter [Fig. 24(b)].

One kind of nonlinear-capacitor amplifier is diagrammed in Fig. 24(f). Unlimited gain is possible; the theoretical amplitude-gain-bandwidth product is approximately equal to the frequency being amplified.[32] As shown, the output is at two frequencies, the original and one generated by frequency conversion. The new frequency can be discarded, but it is usually wise to send it on to the next stage. The most obvious reason for preserving the new signal is that it and the amplified signal at the original frequency together determine the original information regardless of possible fluctuations in the local oscillator. In addition, a single stage of amplification by a lumped nonlinear capacitor is approximately bilateral. To obtain unidirectional gain, one might hope to arrange nonlinear capacitors in sequence in a transmission line (or in a single piece of silicon, as shown in Fig. 15), with a directional phasing of the local-oscillator supply to the successive diodes. This scheme fails if only one signal frequency band is transmitted from diode to diode, for then each diode merely presents to the signal a negative impedance that is independent of the local-oscillator phase.

Practically, of course, it is much easier to transmit both signal bands than to stop one of them. A UHF traveling-wave amplifier using four nonlinear capacitors has been built that gives low noise and unidirectional gain.[33] A few milliwatts of pump power are sufficient for this circuit.

In Fig. 24(f), no frequency values are indicated but the spacing of the lines suggests that s and $b-s$ are

---

[31] C. F. Edwards, "Frequency conversion by means of a nonlinear admittance," *Bell Sys. Tech. J.*, vol. 35, pp. 1403–1416; November, 1956.

[32] A. E. Bakanowski, "The Nonlinear Capacitor as a Mixer," Bell Telephone Labs., Second Interim Rep. on Task 8, (crystal rectifiers), Signal Corps Contract No. DA-36-039-36-5589; January 15, 1955.

[33] R. S. Engelbrecht, private communication.

approximately equal to each other. This particular frequency relation has the advantage that single-tuned circuits can be used to tune both s and $b-s$. In the process of amplifying s it is not only possible but also necessary that an amplified signal emerge at an inverting frequency such as $b-s$. Therefore $b-s$ must be terminated with an "idler" impedance capable of absorbing power; *i.e.*, having a resistive component. Depending upon its physical nature, the idler termination may be a source of noise. In a receiver, one might use the antenna as an idler. If the effective source temperature is low, idler noise will be low; if not, a low-noise receiver is of little value. Another possibility is to make the idler frequency much higher than the signal frequency; then the idler noise is deamplified by the frequency ratios $(b-s)/s$.[34] Idler noise can be obviated if systems considerations permit simultaneous reception of signals at s and $b-s$. On this basis, a 3-db noise figure, with 35 db of gain, has been measured for a 6-kmc amplifier using a diffused silicon nonlinear capacitor and a few hundred milliwatts of 12-kmc pump power.[35] For signal-side-band use, one would have a noise figure of 4.5 db to 6 db, depending on whether the idler termination is noiseless or exhibits room-temperature thermal noise.

This experimental result can surely be improved by further development, but it is already considerably better than a microwave superheterodyne receiver using point-contact nonlinear resistors, which might give an over-all noise figure of 6.5 db in converting 6000 mc to 30 mc. It should be noted, of course, that point-contact noise figures have been bettered at certain frequencies with traveling-wave electron-beam tubes; *e.g.*, less than 4 db at 3000 mc.

Nonlinear resistors, either point-contact diodes or some future junction devices, will continue to have some uses. They are generally to be preferred when a high frequency must be converted to a much lower frequency. In such down-converters, nonlinear capacitors are difficult to stabilize. An important use for down-converters is in frequency standards; one can obtain a signal whose frequency is the difference between some standard high-frequency oscillator (such as an ammonia maser) and a high frequency which is to be compared with the standard. As long as a device can be fairly described as a nonlinear resistor, its behavior is independent of frequency. Therefore, nonlinear resistors could be exceedingly broad-band devices. This potentiality has not been realized, because the usual point-contact nonlinear resistor contains a cat's-whisker with sufficient inductance to limit the bandwidth.

Up-converters accept a signal at a given frequency and have an output at a very much higher frequency. They are typically found in transmitting modulators. In up-converters, it is possible to use *p-n* junction nonlinear capacitors to get gain and nevertheless have unconditionally stable circuits. In addition to amplification, the *p-n* junction has the advantage of good power-handling capability. Bonded germanium diodes and diffused silicon diodes have been found satisfactory in transmitting modulators.[10]

Amplifying up-converters may also be used in low-noise receivers. Up-converters employing nonlinear capacitors have demonstrated lower noise figures than are commonly obtained with electron tubes. For example, an up-converter from 460 mc to 9375 mc has been built with 9 db of gain and a 2-db noise figure.[36]

Junction diodes can be used in frequency shifters to generate carrier frequencies for the several channels of a microwave relay system.[10] Having generated one stable carrier frequency by harmonic generation from a low-frequency oscillator, one can obtain a neighboring carrier frequency by mixing the first with an accurate low frequency. The reliability and power-handling ability of *p-n* junction nonlinear capacitors favors their use in this application.

IX. CONTROL OF MICROWAVE POWER

Microwave circuit elements whose impedances can be altered enable one to vary the transmission of microwave power or its distribution to branches of the circuit. It may not matter if the impedance change is resistive or reactive. If a reactive change is satisfactory, the nonlinear capacitor has a power-handling advantage over point-contact nonlinear resistors, for two reasons. One is the larger area of junction nonlinear capacitors. The other is the fact that a good nonlinear capacitor absorbs but little of the power that it controls. There is no known limit on the speed of switching by either nonlinear resistors or nonlinear capacitors. This statement may seem extraordinary in comparison, for example, with the present limitations of ferrite switches, but its truth is evident if one considers that a switch with either type of diode is just a kind of frequency converter.

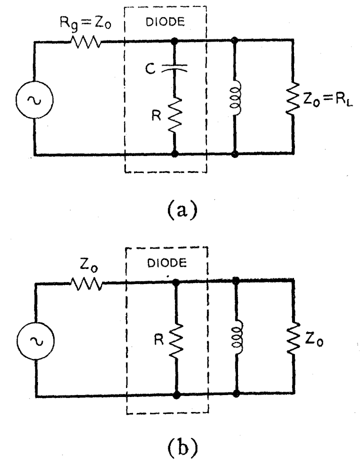
If a relatively slow (*e.g.*, 1 mc) variation of the microwave power is all that is required, the *p-i-n* "variable resistor" diode permits the construction of broad-band switches and attenuators. The *p-i-n* diode has the advantage of even larger size than the *p-n* junction nonlinear capacitor suitable for the same frequency (excepting multiple-junction or series-parallel nonlinear capacitors). Despite large sizes possible with the *p-i-n* structure, one must be careful when it is used with large powers, because, included among its wide range of resistive values, there is usually one resistance that matches well enough to absorb a large fraction of the available power. When the available microwave power exceeds the allowable dissipation of the *p-i-n* diode, one must make sure that the diode does not exhibit the matching impedance except for a short time during a switching operation.

---

[34] S. Bloom and K. K. N. Chang, "Theory of parameteric amplification using nonlinear reactances," *RCA Rev.*, vol. 18, pp. 578–593; December, 1957.

[35] G. F. Herrman, M. Uenohara, and A. Uhlir, Jr., "Noise figure measurements on two types of variable reactance amplifiers using semiconductor diodes," this issue, p. 1301.

[36] G. T. Knapp of the New York Bell Telephone Co. collaborated with the author in obtaining these measurements.



(a)

(b)

Fig. 25—Equivalent circuits of a diode transmission-line switch. (a) Reverse bias (low-loss condition), (b) forward bias (high-loss condition).

Let us consider what can be done in switching microwave power by a simple circuit in which a nonlinear capacitor is shunted across a transmission line of characteristic impedance Z_0. It will be assumed that an inductor is shunted across the line to tune for maximum transmission when the diode exhibits its minimum capacitance, as in Fig. 25(a). When the diode is biased appreciably in the forward direction, the capacitance becomes very large and the effective impedance of the diode is just its series resistance, as shown in Fig. 25(b).

The insertion loss of a shunt element of admittance Y is

$$\text{Insertion loss} = (1 + \tfrac{1}{2} Z_0 Y)^2. \tag{20}$$

At the band center, the insertion loss under reverse bias is

$$\text{Insertion loss, reverse bias} = \left\{ 1 + \frac{Z_0}{2R_s} \cdot \frac{1}{1 + (f_c/f)^2} \right\}^2, \tag{21}$$

where f_c is the cutoff frequency defined by (3). For forward bias, the effect of the tuning inductor is usually negligible and the insertion loss is

$$\text{Insertion loss, forward bias} \approx \left\{ 1 + \frac{Z_0}{2R_s} \right\}^2. \tag{22}$$

The fraction of the incident power dissipated in the series resistance is

$$\frac{P_{\text{dissipated in diode}}}{P_{\text{available}}} = \frac{Z_0/R}{1 + (f_c/f)^2} \div \text{insertion loss} \tag{23}$$

for reverse and

$$\frac{P_{\text{dissipated in diode}}}{P_{\text{available}}} = \frac{Z_0}{R} \div \text{insertion loss} \tag{24}$$

for forward bias.

For an example, a calculation is given for 6-kmc switching with a diode of 100-kmc cutoff frequency. A reverse bias insertion loss of 0.5 db permits a forward bias insertion loss of 25 db. In reverse bias, 11 per cent of the incident power would be dissipated in the diode; in forward bias, 10 per cent. The loaded Q under reverse bias is less than unity, so the switch is quite broad band.

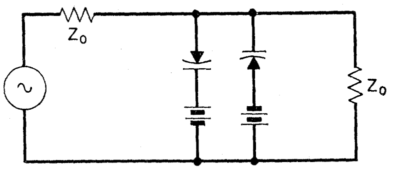


Fig. 26—A protective limiter circuit using two nonlinear capacitors.

Still better performance is possible at lower frequencies. For the same diode at 400 mc, a reverse bias loss of 0.2 db would permit, theoretically, a forward bias loss of 58 db and a loaded Q less than 2. The corresponding dissipated powers would be 2.6 per cent and 0.25 per cent of the incident power.

The reverse bias calculations must be modified unfavorably when the signal voltage is large, because the capacitance is nonlinear and the breakdown voltage must not be exceeded. However, the power that can be controlled will still be much larger than the allowable power dissipation of the diode(s).

On the assumption that thermal conduction through the semiconductor is the determining process the allowable dissipation P per diode can be estimated from

$$\begin{aligned} P &\approx 1.5\kappa d \Delta T && \text{for mesa diodes} \\ P &\approx 4\kappa d \Delta T && \text{for dimple diodes} \end{aligned} \tag{25}$$

where d is the junction diameter, κ is the thermal conductivity, and ΔT may conservatively be taken as 100°C for silicon diodes in contact with a heat sink near room temperature. For example, these relations give 1 watt for a 0.003-inch diameter mesa diode and 5 watts for a 0.005-inch diameter dimple diode, in good accord with experience.

The limiter circuit shown in Fig. 26 is closely related to the switch. Again, the diodes' capacitance can be tuned for good low-level transmission. High-level signals are clipped. The bias batteries may sometimes be eliminated. One application of such a limiter is in maintaining a fairly constant output amplitude from a variable-frequency oscillator; the accuracy might be improved by a feedback circuit that adjusts the bias voltage in accordance with a power monitor. The limiter can also be used as a protective circuit. Since most of the incident high-level power is reflected, a limiter of this kind can protect against powers that exceed by many times the allowable dissipation of the diodes.

X. Harmonic Generation

A nonlinear device generates harmonics when excited by a single-frequency generator. The nonlinear resistance of point-contact rectifiers has long been used for this purpose. In this section, theoretical and experimental evidence is given in support of the idea that nonlinear capacitors should be much better harmonic generators. Harmonic generation seems to be an expedient way of obtaining millimeter waves. Another application

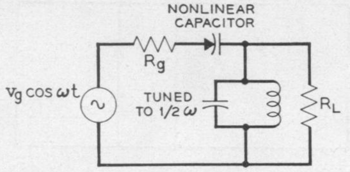


Fig. 27—Subharmonic generator.

is in frequency standards where a low-frequency crystal-controlled oscillator output is multiplied to obtain standard microwave frequencies. In either of these applications the most serious limitation is the weak output obtainable with present crystal rectifiers.

It is expected that much higher efficiency in harmonic generation can be obtained with nonlinear capacitors. Ideally, such a diode cannot convert any of the incident power into dc power, nor can it dissipate any of this power. Hence, if it is possible to put an ideal lossless filter between the generator and the diode, passing only the fundamental, and to have another lossless filter that permits only the desired harmonic to leave by way of the output, then it should be possible to get nearly perfect efficiency in harmonic generation.[30] Since *p-n* junction nonlinear capacitors can be made with much better power-handling capability than point-contact diodes, as well as better efficiency, they should indeed make superior harmonic generators.

Some experiments have been very encouraging even though they made use of miscellaneous filters and tuning elements that happened to be available. Diffused silicon diodes of the mesa type were used. For example, 58 mw of available power at 430 mc was doubled to give 17 mw of 860 mc, a "conversion loss" of 5 db. In tripling, 30 mw of 330 mc gave 5 mw of 990 mc, or a 7-db conversion loss. Three hundred milliwatts of 1200 mc gave 0.4 mw of 8400 mc (29 db down to 7th harmonic).

XI. SUBHARMONIC GENERATION (FREQUENCY DIVISION)

In frequency measurement and the establishment of frequency and time standards, it would be valuable to be able to divide a given frequency by an integer. From the study of differential equations, such as the Matthieu equation, it is known that a time-varying capacitance may be used to divide a frequency by 2; that is, to generate the one-half harmonic.[37]

This conclusion is also suggested by the small-signal analysis of frequency converters. In the "amplifier" scheme diagrammed in Fig. 24(f), it is possible to generate spontaneous oscillations at any pair of frequencies s and $b-s$. If the signal frequency s is just equal to one-half of the beating oscillator frequency b, then the lower-sideband frequency $b-s$ is equal to s. Hence, the two signal frequencies can both be tuned by a single resonant circuit. This fact makes it particularly easy to generate oscillations at $b/2$. Indeed, it is found experimentally that there is a distinct tendency to lock in at the one-half harmonic when a singly-tuned resonant circuit is approximately tuned to this frequency.

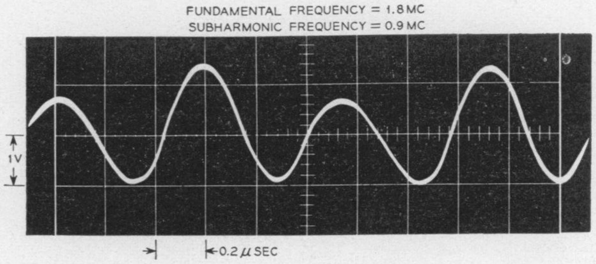


Fig. 28—Waveform observed across tank circuit of Fig. 27.

An experimental test at low frequency was made in the circuit shown in Fig. 27, where an alloy germanium transistor was used as a low-frequency nonlinear capacitor by tying together the emitter and collector. The resulting oscillogram is shown in Fig. 28.[38] Because the resonant circuit was not infinitely sharp, an appreciable amount of the fundamental frequency appears at the output. It is clear from this figure that an elementary description of one-half harmonic generation is to say that "alternate cycles are different." Similar waveforms have been observed on a traveling-wave oscilloscope with 450-mc input. Generation of the one-half harmonic of 12 kmc is easily observed in the apparatus used as a 6-kmc negative resistance amplifier.

For a fixed fundamental oscillator, there are two possible phases of subharmonic oscillation. Ordinarily, the phase is determined by chance. However, it has been suggested that the two phase possibilities could be used in computers to represent binary digits; for this application, it is necessary to devise methods of establishing the desired phase.[39]

XII. VOLTAGE-TUNED CIRCUITS

Direct applications for the voltage-dependent capacitance of *p-n* junctions are to be found in varying the tuning of oscillators, amplifiers, and filters.[8,40] The capacitance presented to *small* high-frequency signals depends only on the dc bias voltage applied to the junction.

The high-frequency Q of a nonlinear capacitor is determined by the series resistance and is a maximum when the capacitance is a minimum; *i.e.*, at the maxi-

---

[37] L. Brillouin, "Wave Propagation in Periodic Structures," McGraw-Hill Book Co., Inc., New York, N. Y., sec. 45; 1946.

[38] A. Uhlir, Jr., "Possible Uses of Nonlinear Capacitor Diodes," Bell Telephone Labs., Eighth Interim Rep. on Task 8, Signal Corp. Contract No. DA-36-039-S65589; July 15, 1956.

[39] J. von Neuman, U. S. Patent No. 2,815,488; December, 1957.

[40] W. Y. Pan and O. Romanis, "Automatic Frequency Control of Television Receivers Using Junction Diodes," in "Transistors I," RCA Labs., Princeton, N. J., pp. 598–608; 1956.

mum reverse voltage. This maximum Q is given by f_c/f, where f_c is the cutoff frequency defined in (1) and f is the frequency of use. At 1 kmc, a diffused silicon junction with an f_c of 100 kmc would have a Q of 100 at -7 volts bias. It would have two and one-half times as much capacitance at zero bias and accordingly would have there a Q of 40. Use at slight forward bias is not excluded, but when appreciable forward current flows, shunt conductance is added to the series resistance losses.

Since the dc current is very small throughout the generally useful voltage range, the dc power required to maintain the desired bias is exceedingly small: 10^{-9} watts, for example, with silicon junctions. Accordingly, low-power feedback circuits can be used to tune an oscillator in response to an error signal.

The high-frequency voltage must be small compared to the relative change of capacitance with voltage, if the effective capacitance is to be independent of signal level (and if the diode is not to become active and perhaps break into oscillation). It is sometimes possible to put the diode at a low-impedance point of a tuned circuit. Otherwise, several diodes can be connected in series (they may be in parallel at dc to keep the control voltage small) or the multiple junction structure of Fig. 13 can be used.

For reverse biases, the capacitance of graded silicon junctions is remarkably independent of temperature, as shown in Fig. 29.[12] Also, for reverse biases, the relative variation of capacitance with voltage is not sensitive to variations in the impurity gradient. These facts make control of capacitance to within 1 per cent seem feasible. Present production techniques do not permit satisfactory yields of single junctions selected to such close tolerances. However, it is easy to combine two or more junctions, after measurement, to obtain a desired capacitance. "Channels" caused by surface charges can produce reversible changes in junction capacitance. These changes are ordinarily negligible at high frequencies, because the series resistance of the channel is usually high. For applications requiring exceptional long-term stability of capacitance, the dimple diode of Fig. 14 affords protection against surface effects.

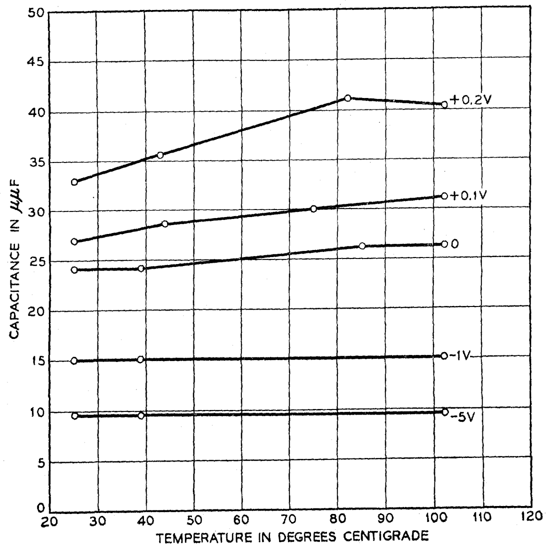


Fig. 29—Capacitance of graded junction in silicon, as a function of temperature and bias. Gradient is 10^{23} cm^{-4}.

XIII. Conclusions

Low-noise UHF and microwave amplification can be obtained at room temperature with the nonlinear capacitance of *p-n* junction diodes of special but simple design. Nonlinear capacitors are also capable of performing many other circuit functions, with the exception of rectification. *P-n* junction diodes are small, reliable, and reproducible. They can control substantial amounts of high-frequency power while consuming very little control power.

XIV. Acknowledgment

The author wishes to thank his colleagues for permission to quote their results in advance of publication. Also, R. M. Ryder has made many specific and helpful suggestions.

But for the patience of the Bell Telephone Laboratories and the U.S. Army Signal Corps, nonlinear capacitors would not have been refined to the point where low-noise microwave amplification could be proved.