References

[1] E. B. Christoffel, "Ueber die Fortpflanzung von Stössen durch elastische feste Körper," *Ann. di Matematica Milano*, vol. 8, pp. 193–243; 1877.
[2] I. Koga, "Thickness vibrations of piezoelectric oscillating crystals," *Physics*, vol. 3, pp. 70–80; August, 1932.
[3] R. Bechmann, "Über die Temperatur-Koeffizienten der Eigenschwingungen piezoelektrischer Quarzplatten und Stäbe," *Hochfrequenz. und Elektroak.*, vol. 44, pp. 145–160; November, 1934.
[4] J. Laval, "Élasticité des cristaux," *Compt. rend. Acad. Sci., (Paris)*, vol. 232, pp. 1947–1948; May 21, 1951.
——, "Sur l'élasticité du milieu cristallin, L'état solide," Congres Solvay, Stoops, Bruxelles, pp. 273–313; 1952.
[5] V. G. Zubov and M. M. Firsova, "On the measurement and new computation of the dynamic elastic constants of quartz," *Kristallografiya*, vol. 1, pp. 546–556; May, 1956. (In Russian.)
[6] R. Bechmann, "Elastic and piezoelectric constants of alpha-quartz," *Phys. Rev.*, vol. 110, pp. 1060–1061; June 1, 1958.
[7] R. D. Mindlin, Letter-type Rept., U. S. Signal Corps Contract DA 36-039 SC-87414; March 29, 1961. (Unpublished.)
[8] W. G. Cady, "Piezoelectricity," McGraw-Hill Book Co., New York, N. Y., London, England; 1946.
J. A. Schouten, "Tensor Analysis for Physicists," Clarendon Press, Oxford, England; 1951.
J. P. Musgrave, "The Propagation of Elastic Waves in Crystals and Other Anisotropic Media," *Repts. Progr. Phys.*, vol. 22, pp. 74–96, 1959, Physical Soc., London, England; 1959.
[9] K. Heegner, "Gekoppelte selbsterregte elektrische Kreise und Kristalloszillatoren," *Elektr. Nachrichtentech.*, vol. 15, pp. 359–368; December, 1938.
[10] "IRE standards on piezoelectric crystals, 1949," PROC. IRE, vol. 37, pp. 1378–1395; December, 1949.
[11] W. R. Ives, "A New Zero Temperature Coefficient Quartz Oscillator Plate," M.S. thesis, Colorado A & M College, Ft. Collins, Colo.; March, 1951.
[12] R. D. Mindlin and D. C. Gazis, "Strong Resonances of Rectangular AT-cut Quartz Plates," Rept., U. S. Signal Corps Contract DA 36-039 SC-87414; July, 1961.
[13] J. L. Saunders and D. L. Hammond, "Design and Development of Extended Temperature Range CR-(XA-21)/U Crystal Units," Wright Air Dev. Ctr., Wright Patterson AF Base, Ohio, Final Rept. AF 33(600)33889, September 16, 1956–June 26, 1959.
[14] R. Bechmann, "Frequency-temperature-angle characteristics of AT- and BT-type quartz oscillators in an extended temperature range," PROC. IRE, vol. 48, p. 1494; August, 1960.
[15] ——, "Influence of the order of overtone on the temperature coefficient of frequency of AT-type quartz resonators," PROC. IRE, vol. 43, pp. 1667–1668; November, 1955.
[16] W. P. Mason, "Zero temperature coefficient quartz-crystals for very high temperatures," *Bell Sys. Tech. J.*, vol. 30, pp. 366–380; April, 1951.
[17] I. Koga, M. Aruga and Y. Yoshinaka, "Theory of plane elastic waves in a piezoelectric crystalline medium and determination of elastic and piezoelectric constants of quartz," *Phys. Rev.*, vol. 109, pp. 1467–1473; March, 1958.
[18] W. P. Mason, "Low temperature coefficient quartz crystals," *Bell Sys. Tech. J.*, vol. 19, pp. 74–93; January, 1940.
[19] J. V. Atanasoff and P. J. Hart, "Dynamical determination of the elastic constants and their temperature coefficients for quartz," *Phys. Rev.*, vol. 59, pp. 85–96; January 1, 1941.
[20] R. Bechmann and S. Ayers, "The Shear Elastic Constants of Quartz and their Behaviour with Temperature," Post Office Res. Sta., Engrg. Dept. London, Dollis Hill, England, Res. Rept. No. 13524; December, 1951.
[21] E. Giebe and E. Blechschmidt, "Über Drillungsschwingungen von Quarzstäben und ihre Benützung für Frequenznormale," *Hochfrequenz. und Elektroak.*, vol. 56, pp. 65–87; March, 1940.
[22] R. Bechmann, "Thickness-shear mode quartz cut with small second- and third-order temperature coefficients of frequency (RT-cut)," PROC. IRE, vol. 49, p. 1454; September, 1961.
[23] P. A. Simpson, C. Barclay and F. P. Phelps, U. S. Patent No. 2,931,924; April 5, 1960.
[24] "IRE standards on piezoelectric crystals: determination of the elastic, piezoelectric, and dielectric constants—the electromechanical coupling factor, 1958," PROC. IRE, vol. 46, pp. 764–778; April, 1958.
[25] R. Bechmann, "Quarzoszillatoren und Resonatoren im Bereich von 50 bis 300 kHz.," *Hochfrequenz. und Elektroak.*, vol. 61, pp. 1–12; January, 1943.

Correspondence

Recombination Radiation Emitted by Gallium Arsenide*

When appropriately diffused GaAs diodes are biased in the forward direction they emit intense line radiation corresponding to gap transitions. Absolute measurements of the emitted radiation intensity indicate that at 77°K these diodes may be as high as 85 per cent efficient in the conversion of injected holes into photons of the gap energy. Data pertaining to the spectral distribution and speed of response of the emitted radiation is presented, and also the high conversion efficiency of the diode and its implications are discussed.

The diodes were fabricated from single-crystal *n*-type GaAs. A *p*-type layer was formed by diffusing in a sealed evacuated quartz tube zinc from a dilute solution of zinc in gallium. The wafers were lapped to 0.003 inch and diced. The die was then alloyed to a Au-Sn plated Mo tab [as shown in Fig. 1(a)] to form the base contact after which an InZn sphere was alloyed to the *p*-type layer in a position above the hole in the base tab to form the other ohmic contact. The diode was then etched to define a junction area of about 7.5×10^{-4} cm^2. The forward current-voltage characteristics of this diode at 298°K and 77°K are shown in Fig. 1. At 298°K the current varies as exp $(qV/2kT)$ up to about 0.1 a. At 77°K the current also varies as exp $(qV/2kT)$ between 10^{-3} and 3×10^{-2} a and below 10^{-10} a with an intermediate region in which it varies between exp $(qV/2kT)$ and exp $(qV/8kT)$.

Fig. 2 shows the relative intensity of the emitted photons as a function of photon energy. Two peaks in the radiation distribution are observed at room temperature. The photon energies at which the peaks occur agree approximately with those observed by Pankove,[1] but the relative intensities of the two peaks are diametrically opposed to those he reported. At room temperature our diodes emit an approximately equal number of photons in each line while at 77°K about 90 per cent of the emitted energy appears in the 1.44 ev band. From measurements of the absolute value of the emitted flux and its angular dependence, calculations which take into account the photon loss due to refraction and internal reflection at the GaAs surface indicate that at 300°K approximately 40 per cent of the holes injected at the *p-n* junction produce photons in the 1.33 ev band. A similar calculation based on the measurements at 77° would give 5 photons in the 1.44 ev band produced per injected hole. However, this result is in error because when the efficiency of producing 1.44 ev photons is close to unity, many photons which would normally not reach the measuring apparatus because of refraction and internal reflection are absorbed and re-emitted until they ultimately emerge from the crystal. By completely neglecting all refraction and internal reflection losses, a lower limit of 0.48 photon in the 1.44 ev band produced per injected hole can be established. On the other hand, if one as-

* Received May 25, 1962.

[1] J. I. Pankove and M. Massoulie, "Recombination Radiation in a Gallium Arsenide P-N Junction," The Electrochem. Soc., Electronics Div. (*Abstracts*), vol. 11, pp. 71–75, Spring Meeting, Los Angeles, Calif.; May 6–10, 1962.

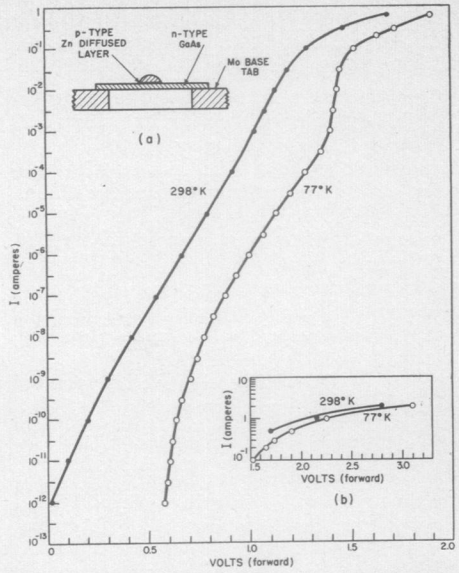


Fig. 1—Current-voltage characteristics for a zinc diffused GaAs diode. Pulse techniques were used for measurements above 10^{-1} a; all other measurements were dc. Higher current measurements are shown in inset (b) and a cross section of diode structure is shown in inset (a).

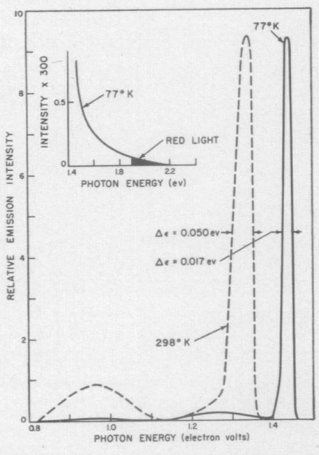


Fig. 2—Relative diode emission intensity as a function of photon energy at 298°K and 77°K. The emission intensity scale is relative and is not the same for both temperature plots. Inset shows the high energy tail at 77°K on an expanded scale.

sumes that the over-all quantum efficiency (photons per injected hole) is unity, one can determine from Fig. 2 that as an upper limit 0.85 photon in the 1.44 ev band are produced per injected hole.

Also shown in an inset in Fig. 2 is the high energy distribution of emitted photons. One sees that there is a small but finite tail which extends into the visible region of the spectrum. This is the source of the "red glow" reported by Mayburg and is quite visible at low temperatures. It is believed that this tail is due to recombination radiation of high energy electrons in the conduction band.

Measurements of emitted radiation at various injection currents above 30 a/cm² show that the radiation increases linearly with current up to current densities of 2.5×10^3 a/cm² at which point the radiation tends to saturate with increased current density. When the injection was performed with a current pulse with a decay time of about 2 μsec it was found that the emitted radiation had the same decay time. Measurements of the switching time for similar diodes (2×10^{-9} sec) indicate that the emission can be modulated in excess of 100 mc.

The above results indicate that it may be possible to fabricate diodes in which for every injected hole a photon of energy close to the band-gap is emitted from the diode. If such diodes are possible, they will, when forward-biased, extract heat from the surroundings. This is because the holes that are injected are from the high energy tail of the Fermi distribution, and the full band-gap energy is not necessary to inject them into the n-region. The heat extracted per unit time will be

$$-\frac{dQ}{dt} = I\left(\frac{h\nu}{q} - V\right)$$

where V is the applied voltage in volts and $h\nu/q$ is the average photon energy also in volts. It is evident from the above equation and Fig. 1 that diodes in which each injected carrier produces a photon of energy close to the band-gap can act as refrigerators.

The authors express their thanks to Dr. R. H. Rediker for many helpful discussions and suggestions, to F. M. Sullivan for the diode fabrication, and to D. H. Bates for assistance in taking the measurements.

R. J. Keyes
T. M. Quist
Lincoln Laboratory[2]
Mass. Inst. Tech.
Lexington, Mass.

[2] Operated with support by the U. S. Army, Navy and Air Force.

Microwave Photomixing of Optical Maser Outputs with a *PIN*-Junction Photodiode*

This paper reports the possibility of using a *PIN*-junction photodiode as a mixer in an optical superheterodyne system employing microwave-modulated light. With such a diode, we have detected signals from UHF through X band produced by photomixing[1,2] between axial-mode components of ruby lasers. The observations were made by exposing the laser beam onto a silicon *PIN* photodiode, so that the photocurrent included signals corresponding to differences between the simultaneous, discrete optical frequencies in the laser output. These heterodyned signals are then detected with high signal-to-noise ratio.

The experimental setup is like that employed by McMurtry and Siegman,[2] with the photodiode replacing the microwave phototube. The photodiode is a point-contact *PIN*-junction type, the P region consisting of a GaAu dot ~10 mils diameter deposited on an epitaxial I region ~10 μ thick (fabricated by Prof. J. F. Gibbons of Stanford). Reverse breakdown voltage is >150 volts. The ceramic-diode cartridge was partly cut away and the diode inserted in the center conductor of a coaxial line, with a small hole in the outer conductor to admit the laser light. One end of the diode holder was connected to the receiving system via coaxial cables and tuners; the other end was connected to a movable short. A break in the outer conductor, insulated by a Mylar sheet, allowed a dc bias voltage to be applied across the diode. Receivers with 10 Mc/sec IF bandwidth were used in the frequency ranges from UHF through X-band. The laser light output was pointed directly at the junction from approximately one foot away, without focusing. Back-bias voltages were generally between 50 v and 100 v.

The top trace of Fig. 1 shows typical microwave photo-beats from the diode as detected with a superheterodyne receiver tuned to the fundamental frequency spacing between adjacent axial optical resonances in the laser rod. The bottom trace shows the simultaneous spikes of photocurrent from the same diode. Observations like Fig. 1 have been made with several ruby rods of different lengths, at their fundamental frequencies and also at higher mode intervals. With a 122-mm long rod, we have detected beats at the first through sixteenth mode intervals, from 691 to 11,050 Mc/sec. A 39-mm long sapphire clad rod yielded beats up to the fifth mode interval, from 2135 to 10,675 Mc/sec. No outputs are observed except at these discrete frequencies.

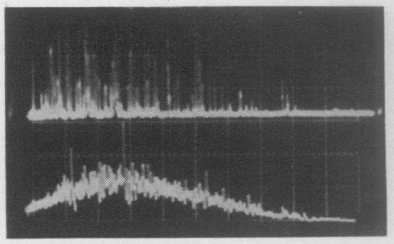


Fig. 1—Lower trace: photocurrent from PIN-junction photodiode with 50 v of back-biased voltage. Upper trace: simultaneous microwave output at 2135 Mc/sec. Sweep speed: 50 μsec/div. Sweep triggers at start of laser action.

The higher frequency outputs might possibly represent not higher mode intervals as claimed here, but rather harmonics of the fundamental frequency, generated by nonlinearities in the diode or receiver. To eliminate this possibility, we measured the harmonic generation in the photodiode itself,

* Received May 21, 1962; revised manuscript received June 11, 1962. This work was supported by the Signal Engineering Laboratories of the U. S. Army Signal Corps.

[1] A. T. Forrester, "Photoelectric mixing as a spectroscopic tool," *J. Opt. Soc. Am.*, vol. 51, pp. 253–259; March, 1961.

[2] B. J. McMurtry and A. E. Siegman, "Photomixing experiments with a ruby optical maser and a travelling-wave microwave phototube," *Appl. Optics*, vol. 1, pp. 51–53; January, 1962.

by inserting pulsed signals into the diode at the fundamental frequency and looking for harmonics. The conversion loss for harmonic generation was found to be large enough to eliminate this possibility.

In principle, the frequency at which the *PIN* photodiode can give substantial output is limited by the transit time through the intrinsic region of junction. For our diode, the expected value for the transit time is $\sim 10^{-10}$ sec, consistent with the experimental results presented here.

This method of optical heterodyning between simultaneous optical modes is a simple but powerful means for studying optical masers. The *PIN* photodiode serves as an optical-frequency mixer which is compact, simple, capable of fairly wide frequency range (when retuned at each frequency), and possessed of good infrared response. It requires, however, a separate amplifier or more sensitive receiver following it. By contrast, the microwave phototube offers the advantage of very wide instantaneous bandwidths (>3:1 in a single device), together with substantial internal amplification of the detected signal. For various reasons, the phototube may be more promising at higher beat frequencies, from 10 kMc/sec upwards.

These experiments were first stimulated by reports of Reisz's work[3] on similar diodes; we have more recently learned of similar work at frequencies up to 4200 Mc/sec by workers at the Philco Scientific Laboratories.[4]

H. Inaba†
A. E. Siegman
Stanford Electronics Labs.
Stanford University
Stanford, Calif.

[3] R. P. Riesz, "High-Speed Semiconductor Photodiodes," to be published.
[4] G. Lucovsky, M. E. Lasser, and R. B. Emmons, "Coherent Light Detection in Solid State Photodiodes," to be published.
† On leave of absence from Tohoku University, Sendai, Japan.

Silicon Field-Effect Transistor with Internal Epitaxial Channel*

Many techniques have been applied to the fabrication of field-effect devices since Shockley's detailed presentation of the principles.[1] These have included alloying, diffusion, etching, and mechanical cutting, for example.[2,3] But field-effect device channels must be regions of high sheet resistivity, typically about 4000 Ω/square, and with these techniques it is difficult to generate such layers reproducibly. Epitaxial techniques, however, have been able to meet the requirement in a highly satisfactory manner.

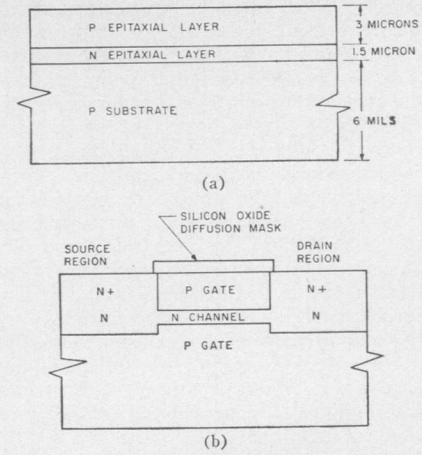


Fig. 1—(a) Epitaxial starting material. (b) Structure completed by means of single diffusion.

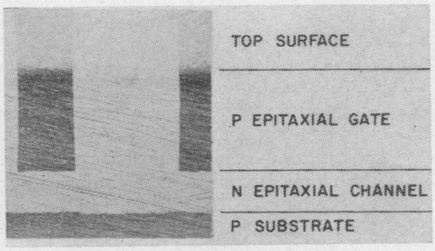


Fig. 2—Beveled and stained specimen showing final structure.

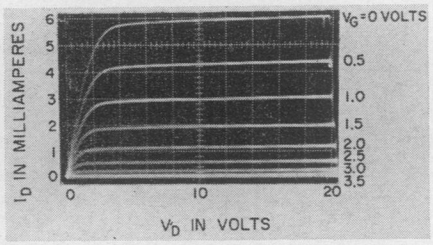


Fig. 3—Output characteristics of epitaxial field effect transistor.

We have employed a *p*-type silicon substrate having two epitaxial layers. A high sheet resistivity *n* layer which will serve as a low pinch-off voltage channel is first deposited onto the *p*-type substrate. A surface *p* layer is then deposited by switching dopant during the growing process. Thus a *pnp* structure is grown continuously. All resistivities are made equal, about 0.5 Ω cm. Typical dimensions are given in Fig. 1(a).

A conventional photoresist procedure is used to produce an oxide diffusion mask forming a compact, convoluted pattern on the surface. A phosphorus diffusion then defines the channel and forms the source and drain contact regions, as shown in Fig. 1(b). This represents a section through any portion of the pattern mentioned above; note that vertical dimensions are exaggerated for clarity, just as they are exaggerated in the beveled and stained section, Fig. 2. Diffusion depth is very noncritical. The depth must merely exceed 3 microns, the thickness of the *p* layer, and make contact with the *n* layer. Overshooting by a few microns has a trifling effect on channel length, which is of the order of 60 microns. Hence this approach to field-effect device fabrication assigns the necessary lateral geometry control to the diffusion operation, which is easy with conventional masking techniques, and it assigns sheet resistivity control to the epitaxial process, which is well adapted to meeting this requirement.

Excellent devices have been made in a passivated planar configuration. In addition, unpassivated mesa devices have exhibited good characteristics and a high degree of electrical stability. The present devices are being assembled in the TO-5 package. Fig. 3 shows an example of the output characteristics.

The same structure principles of lateral geometry control by oxide masked diffusion and sheet resistivity control by epitaxial deposition will of course be applicable to field-effect transistors with *p*-type channels.

Contacts can be applied to either or both of the gates, as desired. By connecting the gates electrically to the source, one obtains an efficient current limiter.[3]

G. C. Onodera
W. J. Corrigan
R. M. Warner, Jr.
Motorola
Semiconductor Products Div.
Phoenix, Ariz.

* Received May 17, 1962.
[1] W. Shockley, "A unipolar 'field-effect' transistor," Proc. IRE, vol. 40, pp. 1365–1376; November, 1952.
[2] G. C. Dacey and I. M. Ross, "The field-effect transistor," *Bell Sys. Tech. J.*, vol. 34, pp. 1149–1189; November, 1955.
[3] R. M. Warner, Jr., W. H. Jackson, E. I. Doucette, and H. A. Stone, Jr., "A semiconductor current limiter," Proc. IRE, vol. 47, pp. 44–56; January, 1959.

Optical Harmonic Frequency Ratio Measurements*

Generation of optical harmonics with high intensity ruby masers has been reported by several workers.[1-3] Tentative reports[4] of discrepancies in the expected ratio of 2:1 between the harmonic and fundamental frequencies as well as the possibility of observing phonon effects have led to the high resolution study discussed below.

The harmonics were generated by focussing the output of a cooled pulsed-ruby source, estimated at several joules, onto a KDP crystal, at 45° to $X-Y$ axes and normal to the Z axis. Since as a result of thermal tuning,[5] the output wavelength of the ruby may be different on consecutive pulses, it was essential to make simultaneous observations on both the fundamental and second harmonic. In order to accomplish this, the light in the forward direction was

* Received May 29, 1962. Work supported in part by a grant from the U. S. Air Force, monitored by the Air Force Office of Scientific Research.
[1] P. A. Franklin, A. E. Hill, C. W. Peters and G. Weinreich, "Generation of optical harmonics," *Phys. Rev. Lett.*, vol. 7, pp. 118–119; August, 1961.
[2] J. Giordmaine, "Mixing of light beams in crystals," *Phys. Rev. Lett.*, vol. 8, pp. 19–20; January, 1962.
[3] P. D. Maker, R. W. Terhune, M. Nisenoff and C. M. Savage, "Effects of dispersion and focussing on the production of optical harmonics," *Phys. Rev. Lett.*, vol. 8, pp. 21–22; January, 1962.
[4] H. S. Boyne, private communication.
[5] I. D. Abella and H. Z. Cummins, "Thermal tuning of ruby optical maser," *J. Appl. Phys.*, vol. 32, pp. 1177–1178; June, 1961.

passed through a collimating system, incident on a direct-view prism, then refocused onto the slit of a 3.4-meter Jarrell-Ash Ebert Spectrograph with a resolving power of about 250,000 in the eighth-order red. The direct-view prism displaces the red and ulraviolet images along the length of the slit by several millimeters. This serves to separate the images on the film, thus preventing superposition and possible confusion. With this arrangement it was possible to obtain single-flash images on Kodak 103-O plates in the eighth-order red and sixteenth-order ultraviolet. The separation of images was further aided by dispersion of air which caused a sensible relative displacement of the ultraviolet image to shorter wavelength. The focused spots were carefully centered on the entrance slit by photographing them with a separate camera and making the necessary alignment of external optics.

The plates were measured with a traveling microscope comparator, and corrections due to dispersion of air were applied.[6] Near threshold the expected 2:1 frequency ratio was found to hold to within the instrumental uncertainty of about 1.5 ppm (0.043 cm^{-1} at the second harmonic). At higher intensities the harmonic line appeared to be displaced toward the blue by as much as 10 ppm. This apparent shift is consistent with the effects to be expected from thermal tuning[5] during single high-power pulses. As a result of heating, the ruby output shifts toward the red during the pulse, producing a smeared image on the film. The red power output rises rapidly to a maximum, then decays slowly in time. The second harmonic output, proportional to the square of the input power, is produced most efficiently at the early part of the pulse before appreciable tuning occurs. The result of this is that the center of gravity of the ultraviolet line is displaced from the center of gravity of the red line. This effect is further complicated by the differential response of the 103-O emulsion to red and ultraviolet. One concludes from the threshold observations that to within the experimental uncertainty of 1.5 ppm the expected 2:1 ratio obtains.[7] The apparent shifts observed with high power are believed to be purely instrumental for the reasons given.

Mention should be made of the occasional appearance of fine structure in the ultraviolet images, consistent with the operation of several longitudinal modes in the ruby source.[8] The structure was not, however, visible in the red image in this experiment. We believe that the apparent absence of these lines results from the poor red contrast and sensitivity of the emulsion used which gives rise instead to a broadened image.

Some estimates of linewidths were made although the arrangement was not best suited for this. In a few clear cases close to threshold the width of the second harmonic line was essentially twice the fundamental width of 0.1 cm^{-1}. Phonon effects in the ultraviolet generating crystal which could cause line broadening are quite weak, and none were observed here.

The author is indebted to Prof. C. H. Townes and H. S. Boyne for useful discussions, and to Prof. R. Novick for his critical reading of the text.

I. D. Abella
Columbia Radiation Lab.
Columbia University
New York, N. Y.

[6] B. Edlen, "Dispersion of standard air," *J. Opt. Soc. Am.*, vol. 43, pp. 339–344; May, 1953.
[7] Similar results were found by H. S. Boyne and W. C. Martin, *J. Opt. Soc. Am.* (to be published), and J. Giordmaine, private communication.
[8] I. D. Abella and C. H. Townes, "Mode characteristics and coherence in optical ruby masers," *Nature*, vol. 192, pp. 957–959; December, 1961.

Highly Sensitive Static Magnetic Field Detector*

A simple and high sensitive scheme has been developed for the detection and measurement of small static magnetic fields. The scheme utilizes the diamagnetic properties of superconductors. The detector consists of a tin gate lying on top of, and insulated from, a lead ground plane. While the devices tested consisted of thin, vapor-deposited films, no doubt this device would be constructed of foils. The device is cooled to a temperature sufficiently low to allow the lead and tin to become superconducting. The approximate geometry of a simple device is shown in Fig. 1(a). The cross section shown in Fig. 1(b) has been greatly exaggerated for clarity.

The critical current of the tin gate is a marked function of the field applied to the sample. Fig. 2 shows a typical plot of the critical current of the tin as a function of angle, when the device is rotated in the stray field within a test dewar. The critical current varied from 205 ma to 8 ma. In some samples the magnification has been sufficient to produce an orientation in which the gate is held resistive.

Operation of the detector can be understood by noting that the field in the vicinity of the edge of an ellipsoid is approximately given by[1]

$$H_e = H_a\left(1 + \frac{a}{b}\right)$$

where

$H_e = H$ at the edge of the ellipsoid
$H_a = H$ applied normal to the major axis of the ellipsoid
a — length of major axis
b — length of minor axis.

This equation is valid for $H_e < H_c$, the critical field of the ellipsoid.

For the sample shown in Fig. 1, $a = 1$ cm, $b = 5000$ Å, $a/b = 20{,}000$. If, for example, the component of earth's field normal to the sample is 0.001 oersted, the edge field will be about 20 oersted. It is precisely this magnification that caused the orientation sensitivity shown in Fig. 2. The symmetry is expected since the edge field and the field due to current in the tin gate add in quadrature.

A slight orientation sensitivity observed in open field cryotrons has been attributed to this effect.[2] Here the control acts as the magnifier. The device is also sensitive to initial orientation in the dewar. The field trapped when the ground plane becomes superconducting manifests itself as an edge field.

This device is extremely useful for detection of small static magnetic fields. When a "field free" region is produced using Helmholtz coils, this detector is superb for detecting the null field. The device can also be used to accurately detect and map anomalies in the earth's magnetic field. In order to measure the field, it is only necessary to provide a null field and then accurately calibrate critical current of the tin gate vs applied field.

David J. Dumin
Applied Electronics Labs.
Stanford University
Stanford, Calif.
formerly at
IBM Labs.
Yorktown Heights, N. Y.

* Received June 5, 1962.
[1] See, for example, D. Shoenberg, "Superconductivity," Cambridge University Press, London, Eng.; pp. 164–171, 1952.
[2] A. Brenemann, private communication.

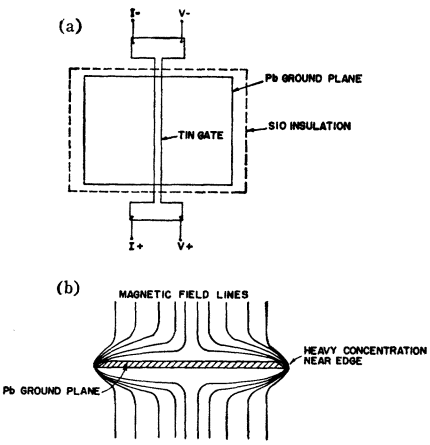


Fig. 1—(a) Geometry of a sample device. (b) Cross section showing concentration of flux lines.

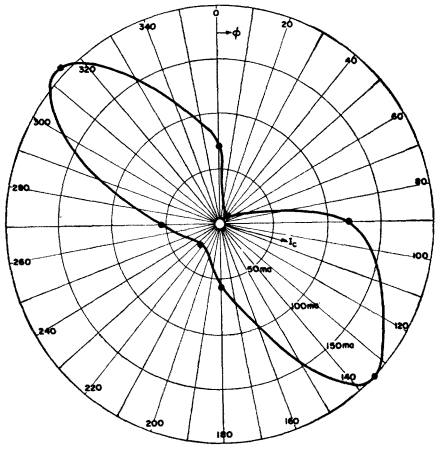


Fig. 2—Critical current of the tin gate vs angle in the dewar.

Light and Gravitation*

The frequency shift experienced by photons in falling through a gravitational field, as measured by Pound and Rebka,[1] can be explained in a simple and nonrelativistic way. If the effective weight of the photon is taken as $m=E/c^2=h\nu$, the energy gained in falling the distance l in the earth's field will be mgl and the momentum gained will be mgl/c. If m and c are the mass and velocity at the higher point, and m' and c' are the corresponding quantities at the lower point, then we have, from conservation of momentum, $m'c'=mc+mgl/c$, and from conservation of mass-energy, $m'c'^2=mc^2+mgl$. Solving these equations, we obtain $c'=c$ and $m'=m(1+gl/c^2)$. If λ and λ' are the wavelengths at the two points, and ν and ν' are the frequencies, we also have $\lambda'=\lambda(1-gl/c^2)$ and $\nu'=\nu(1+gl/c^2)$.

If we visualize a photon as consisting of a wave packet of length Δx containing N cycles, then $\Delta x=N\lambda$. From quantum mechanics the spatial extension of a wave packet is inversely proportional to its momentum. Hence $\Delta x'=N\lambda'$ with N, the number of cycles, remaining constant.

We now see rather clearly what happens to a photon as it gains energy in falling through a gravity field. The wave packet shrinks in both total length and wavelength, the velocity and number of cycles remaining constant. The decrease in wavelength at constant velocity accounts for the increase in frequency.

A further crucial test for any theory of light and gravitation is to predict the observed deflection of star images close to the sun during a solar eclipse.[2] In this case the photon is moving transverse to the direction of the gravity field; and if one calculates, as for example in Bohm,[3] the deflection of a photon, assuming it to be a particle of mass $m=E/c^2$ and velocity c, the resulting deflection will only be about half the observed value. However, the photon will have an extension in the direction of the field, and, as we have seen, the wavelength decreases in this direction. Thus the wavefronts of the photon will be tilted in the direction of the field, and this will cause a further deflection of the photon since it will follow the direction of its wavenormals. If this additional deflection is calculated, it turns out to be the missing half of the observed value.

Since we have seen that in the interaction of light and gravity the light velocity remains constant, it is tempting to revive an old idea, due originally to Faraday, that the gravity field is the medium for the propagation of electromagnetic radiation. Without inquiring into the question of how this comes about, let us make the assumption that the velocity of light is constant relative to the structure of a *local* gravity field and see how far this assumption will take us.

Planets, stars, galaxies, clusters of galaxies, all of the ascending hierarchy of assemblages of matter in the universe, possess local gravity fields. In gravitational terms, we may describe the universe in terms of relatively moving local gravity fields. We should then expect the phenomena of stellar aberration and Doppler shift to be explainable in terms of these relatively moving fields plus the assumption that the velocity of light is always equal to c relative to whatever local gravity field it is traversing.

Space in this communication does not permit a detailed description of how these accommodations take place, but it is possible to rough them in briefly. If we consider the Doppler shift from a moving stellar source, for example, we find the following. The star will have a local gravity field which is moving with respect to the galactic gravity field, for instance. Since the velocity of propagation must be the same in both fields, light originating on the star, on crossing the transition between the two fields, must have its wavelength increased or decreased depending on the relative motion of the two fields. If this light is finally observed on the earth, and the earth is moving with respect to the galactic field, another change in wavelength will occur on entering the earth's gravity field. The final wavelength observed will depend only on the relative motion of the star and earth.

In the case of stellar aberration we have to consider the behavior of a spherical wave crossing the boundary between two local gravity fields which are moving relative to each other in a direction at right angles to the direction of propagation. In this case a simple Huyghens construction will show that the pattern of waves in the second gravity field will be distorted in such a manner that the waves entering a telescope in the second field will be tilted and will appear to come from a source which is displaced from its true position by an amount depending on the relative motion of the two fields.

More interesting questions present themselves when we consider the rotating earth. To what extent does the structure of its gravity field that determines light velocity rotate with the earth? To determine this requires the measurement of an "ether drift" of 0.3 km/sec or less in temperate latitudes. The Michelson-Morley experiment and its later refinement by Joos,[4] in a 360° rotation of the apparatus, found no drift greater than about 1.0 km/sec, and it is clear that a precision of at least an order of magnitude better than this will be required to reach the actual rotational velocity. The Cedarholm-Townes ammonia maser experiment[5] which currently reports no drift greater than 0.015 km/sec is actually of no value in this measurement, because on rotation of the apparatus a considerable positive effect was observed which was ascribed to an unavoidable change in local magnetic conditions. The Michelson-Gale experiment[6] is consistent with the view that the earth rotates with respect to the propagation medium but since this experiment can be explained, though tortuously, by general relativity, a verification by a refined Michelson-Morley type of experiment is urgently desired. Techniques using microwaves and atomic frequency standards for the rotational experiment will be the subject of a future communication.

C. W. Carnahan
120 Fawn Lane
Menlo Park, Calif.

* Received December 28, 1961.

[1] R. V. Pound and G. A. Rebka, Jr., "Apparent weight of photons," *Phys. Rev. Lett.*, vol. 4, pp. 337–341; April 1, 1960.

[2] H. von Kluber, "The determination of Einstein's light-deflection in the gravitational field of the Sun," in "Vistas in Astronomy," Pergamon Press, New York, N. Y., vol. 3, pp. 47–77; 1960.

[3] D. Bohm, "Quantum Theory," Prentice-Hall, Inc., New York, N. Y., pp. 517–519; 1951.

[4] G. Joos, "Theoretical Physics," Hafner Publishing Co., New York, N. Y., 2nd ed., pp. 235–237; 1950.

[5] J. P. Cedarholm, G. F. Bland, B. L. Havens, and C. H. Townes, "New experimental test of special relativity," *Phys. Rev. Lett.*, vol. 1, p. 342; November, 1958.

[6] Joos, *op. cit.*, pp. 471–473.

Modes in Rectangular Guides Loaded with Magnetized Ferrite*

In previous works[1,2] the exact investigation of propagating modes of any order in a rectangular guide partially filled with a transversely magnetized lossless ferrite slab situated against one side wall has been carried out. The numerical calculations showed the existence of structures for which the first and second order modes have a unidirectional character.

Seidel and Fletcher published a paper[3] on an approximate solution of the same problem. They have considered the asymptotic case of vanishingly small waveguides. They found, besides the zero order ferrite-dielectric modes, three types of higher order modes, which they called FM (ferrite-metal), FAI (ferrite-air I) and FAII (ferrite-air II). The propagation constants of the FM and FAI modes are equal and opposite, and therefore from the Seidel and Fletcher analysis it appears that in the asymptotic case the structure cannot be unidirectional for modes of order higher than zero and for a certain range of the ferrite tensor permeability components.

This result seemed to be somehow in contrast with our numerical calculations since, although for guides of finite cross section, we found, for the same range of the ferrite tensor permeability components considered by Seidel and Fletcher, unidirectional higher order propagating modes.

We have therefore made the vanishingly small waveguide approximation in our characteristic equation [(11) of the above men-

* Received, February 5, 1962; revised manuscript received, February 23, 1962. The research reported in this document has been sponsored by the Electronics Research Directorate, AF Cambridge Res. Lab., AFRD (Air Research and Development Command) through its European Office, under Contract AF 61(052)-101.

[1] G. Barzilai and G. Gerosa, "Modes in rectangular guides filled with magnetized ferrite," *L'Onde électrique*, vol. 38, pp. 612–617, *Supplément Spécial, Congrès International Circuits et Antennes Hyperfréquences*, Paris; October 21–26, 1957.

[2] G. Barzilai and G. Gerosa, "Modes in rectangular guides partially filled with transversely magnetized ferrite," IRE Trans. on Antennas and Propagation, vol. AP-7, special supplement, pp. S471–S474; December, 1959.

[3] H. Seidel and R. C. Fletcher, "Gyromagnetic modes in waveguide partially loaded with ferrite," *Bell Sys. Tech. J.*, vol. 38, pp. 1427–1456; November, 1959.

tioned work[1]] to see if the bidirectional character of the higher order modes was peculiar to the asymptotic structure. However, from our characteristic equation we obtained the FM and FAII modes, but we did not find the FAI set. Therefore the asymptotic analysis of our characteristic equation was consistent with our previously obtained results, since the lack of the FAI modes makes possible structures with unidirectional higher order modes even in the asymptotic case.

We have investigated the reasons for the discrepancy between our asymptotic analysis and the one carried out by Seidel and Fletcher, and we have found that the Seidel and Fletcher FAI modes are trivial solutions with zero amplitude. The arguments supporting this conclusion are reported below.

Apart from the asymptotic approximation, the method of attacking the problem we have used and the one used by Seidel and Fletcher differ for the expression of the e.m. field in the vacuum region. We have expressed such a field as a superposition of a TE wave and a TM wave, while Seidel and Fletcher used two waves, one of which has the tangential electric field parallel at the vacuum-ferrite interface to the tangential electric field of one mode in the ferrite region, and the other has the tangential magnetic field parallel to the tangential magnetic field of the other mode in the ferrite region.

If we introduce into Maxwell's equations for a vacuum region a wave solution of the following form:

$$\begin{Bmatrix} E \\ H \end{Bmatrix} = \begin{Bmatrix} \sqrt{\mu_0/\epsilon_0}\, e \\ h \end{Bmatrix} B \cdot \exp\left[j(k_x x + k_y y + k_z z)\right] \quad (1)$$

(time dependence exp $j\omega t$ is assumed), we obtain a system of six linear algebraic homogeneous equations in the six components e_x, e_y, e_z, h_x, h_y, h_z of the two constant adimensional vectors $\mathbf{e}$ and $\boldsymbol{h}$. By setting equal to zero the determinant of the coefficients, we obtain (by normalizing the propagation constants with respect to $\omega\sqrt{\mu_0\epsilon_0}$):

$$k_x^2 + k_y^2 + k_z^2 = 1. \quad (2)$$

If (2) is satisfied, by computing all the 5-order minors of the 6-order determinant, it is found that they are all equal to zero, while there are 4-order minors different from zero; therefore, the 6-order determinant of the coefficients becomes a 4-rank determinant. This means that from Maxwell's equations we obtain only four independent equations for the six field components.

We can obtain a solution with only one arbitrary amplitude by adding to four of the six Maxwell's equations a fifth linear homogeneous relation among the six field components. This fifth linear relation must of course be independent from Maxwell's equations.

If we separately pick up two different linear homogeneous relations we can obtain two different wave solutions having the same form (1). However in order to obtain two independent waves it is necessary that four of the six Maxwell's equations and the two different additional linear homogeneous conditions are altogether six independent conditions. This is so if the 8×6 matrix, obtained by adding to the six Maxwell's equations the two additional linear homogeneous conditions for the two waves, is a 6-rank matrix, *i.e.*, at least one 6-order minor is different from zero. If, on the contrary, the rank of the above mentioned 8×6 matrix is less than 6, the two waves obtained are not independent.

In our work[1] we have chosen for the two waves used to express the e.m. field in the vacuum region the following additional linear equations (x axis = longitudinal axis of the guide):

$$e_x = 0 \quad \text{(TE wave)}; \quad (3)$$

$$h_x = 0 \quad \text{(TM wave)}, \quad (4)$$

while Seidel and Fletcher chose

$$e_{zfI}e_x - e_{xfI}e_z = 0 \quad \text{(parallelism between the tangential electric fields)}; \quad (5)$$

$$h_{zfII}h_x - h_{xfII}h_z = 0 \quad \text{(parallelism between the tangential magnetic fields)} \quad (6)$$

where the subscript f refers to the ferrite region.

Under conditions (3) and (4) the 6-order minors of the above mentioned 8×6 matrix are all equal to zero if

$$1 - k_x^2 = 0, \quad (7)$$

while under (5) and (6) this happens if:

$$k_x e_{zfI}\left[(k_x - 1/k_x)h_{zfII} - k_z h_{xfII}\right] - k_z e_{xfI}\left[k_x h_{zfII} - (k_z - 1/k_z)h_{xfII}\right] = 0. \quad (8)$$

Eqs. (7) and (8) represent, therefore, for the two different expressions of the e.m. field in the vacuum region, the no-independence condition for the two waves. It can be verified that if (7) is satisfied all components of the TE and the TM waves become proportional, reducing the two waves to TEM waves, and if (8) is satisfied all components of the two waves obtained under conditions (5) and (6) become proportional.

Consider now the boundary value problem considered by Seidel and Fletcher,[3] *i.e.*, a vacuum-ferrite interface. We can express the e.m. field in the vacuum and in the ferrite region as a superposition of two waves. If however we choose for the vacuum region two waves having all components proportional, it can be shown that the characteristic equation obtained from the boundary conditions is automatically satisfied, but the corresponding e.m. field is everywhere equal to zero.

Therefore the no-independence condition for the two waves used to express the e.m. field in the vacuum region always corresponds to a trivial solution for the boundary value problem considered by Seidel and Fletcher.[3] The same happens in the more complicated boundary value problem considered by us,[1] and in fact our characteristic equation also contains the trivial solution (7), which, however, we have neglected.

Under asymptotic approximation, condition (8) becomes:

$$k_x e_{zfI} - k_z e_{xfI} = 0 \quad (9)$$

or

$$k_x h_{zfII} - k_z h_{xfII} = 0. \quad (10)$$

Eq. (9) is identical to the characteristic equation for the FAI modes found by Seidel and Fletcher. Therefore the FAI modes are modes of zero amplitude and the relative characteristic equation is only the expression for the no-independence condition discussed above.

We note that if we compute the e.m. field corresponding to the characteristic equation for the FAI modes by using the asymptotic approximation we do not obtain zero, since the approximate expressions for the field components of the two waves in the vacuum region are not exactly proportional. This is perhaps the reason why Seidel and Fletcher, who started directly with the approximate expressions, did not realize that the FAI modes are solutions of zero amplitude.

From the preceding analysis it can be concluded that the FAI modes found by Seidel and Fletcher do not exist; therefore, even in the asymptotic case they have considered, undirectional higher order modes can be found.

The author wishes to express his thanks to Prof. G. Barzilai for encouragements and helpful suggestions and discussions in the course of this investigation.

GIORGIO GEROSA
Istituto di Elettronica dell' Università di Roma
Rome, Italy

Single-Sideband Suppressed-Carrier Modulation of Coherent Light Beams*

A method for producing single-sideband suppressed-carrier modulation at optical frequencies was recently devised and tested in our Laboratories with a coherent carrier obtained from a helium-neon gas laser. This type of modulation is of interest for improving signal-to-noise ratio and conserving bandwidth. It is also important in that it provides a solution to the problem of shifting the inherently fixed frequency of a laser beam to another frequency without introducing additional frequency components, thus making possible a variety of new communications and physics applications of the laser.

The principle of the modulator is that a rotating birefringent plate, or a system that simulates the action of such a plate, will act upon a circularly polarized light beam to produce a separable component shifted in frequency. The rotation is accomplished by the use of the Pockels electro-optic effect in a pair of potassium dihydrogen phosphate (KDP) crystals. This rotation can transform a left-circularly polarized light beam to produce a right-circularly polarized component at one of the modulation sideband frequencies. The displacement of this frequency is limited only to the maximum frequency at which the electro-optic effect can

* Received May 14, 1962; revised manuscript received May 21, 1962.

be produced, which has been shown to extend into the microwave range.[1]

A diagram of the single-sideband suppressed-carrier optical modulator (SSBSCOM), is shown in Fig. 1. Incoming light is passed through a combination of a plane polarizer and a quarter-wave birefringent plate that functions as a left-handed circular polarizer. This polarized light is passed through the two electro-optic crystals along their optic axes (or *c* directions) and then through a right-handed circular analyzer (the mirror image of the polarizer), which extracts the desired r.h. component and blocks the l.h. component. A modulating electric field is applied to each crystal in the c-axis direction by means of transparent electrodes. The effect of a rotating birefringent plate is obtained by having the *b* axes of the two crystals oriented 45 degrees apart, as shown, and applying the modulation voltages in phase quadrature. It can be shown that the output light consists of a l.h. component of one fixed amplitude and a r.h. component of a second fixed amplitude. The l.h. component rotates in synchronism with the input left-polarized light. However, the r.h. component executes one additional r.h. cycle of rotation during each modulation cycle, and therefore the frequency of this component is the sum of the input light frequency and the modulation frequency, *i.e.*, the upper sideband frequency.

The SSBSCOM has been tested at audio frequencies by using it to modulate light from a General Telephone & Electronics Laboratories gas laser. The crystals used were a pair of potassium dihydrogen phosphate bars 7 mm long and 10 mm square, with peak modulating voltages of about 2000 volts. The occurrence of SSBSC modulation was established by using the two-tone test that is employed for this same purpose at ordinary radio frequencies. Two modulating frequencies were applied, and the modulated light was detected with an RCA 7102 multiplier phototube. The modulation envelope was then displayed on an oscilloscope and examined for beat frequencies. Fig. 2 shows oscillograms obtained under four different conditions of operation. Tones of 1 and 1.5 kc were used, and the circuitry was arranged so that upper sideband output could be obtained for either tone alone [Figs. 2(a) and 2(b)] or for both tones together [Fig. 2(c)]; alternatively, upper sideband output could be obtained for one tone and lower sideband output for the other [Fig. 2(d)]. The upper traces are the waveforms of the voltage on the first crystal, the middle traces are the modulation envelopes, and the lower traces are base lines obtained with the light turned off. The envelopes of Figs. 2(a) and 2(b) are horizontal lines with a slight ripple at the modulation frequency, indicating extinction of the lower sideband and almost complete suppression of the carrier. The envelope of Fig. 2(c) displays the beat frequency of 0.5 kc between the two upper sideband frequencies. The envelope of Fig. 2(d) displays the beat frequency of 2.5 kc between the upper and lower sideband frequencies.

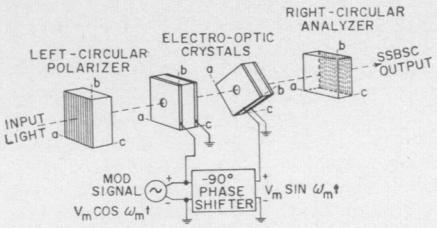


Fig. 1—Single-sideband suppressed-carrrier optical modulator (SSBSCOM).

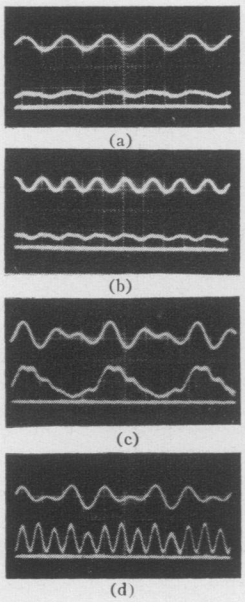


Fig. 2—Two-tone test of SSBSCOM. (a) 1-kc upper sideband output. (b) 1.5-kc upper sideband output. (c) 1-kc upper sideland output and 1.5-kc upper sideband output. (d) 1-kc upper sideband output and 1.5-kc lower sideband output.

Detailed mathematical analysis of the SSBSCOM reveals that the l.h. component at the output of the second crystal consists of a carrier of frequency f_c and a series of even-ordered double-sideband terms at frequencies $f_c \pm nf_m$, where f_m is the modulation frequency. The amplitudes of these l.h. frequency components are

$$L_n = \begin{cases} AJ_n(kV_m) & (\text{for } n \text{ even or zero}) \\ 0 & (\text{for } n \text{ add}) \end{cases} \quad (1)$$

where $k = 2\frac{1}{2}\pi n^3 r/\lambda$ n is the index of refraction, r is the appropriate electro-optic constant, λ is the light wavelength, A is the amplitude of the input light, V_m is the peak modulation voltage, and J_n are nth order Bessel functions of the first kind. All of these terms are eliminated by passing the light through the r.h. circular analyzer. The r.h. component at the output is a series of odd-ordered single-sideband terms at frequencies $f_c \pm nf_m$. The amplitudes of these terms are

$$R_n = \begin{cases} \sqrt{2}AJ_n(kV_m) & (2) \\ (\text{for } n = n = +1, -3, +5, -7, \cdots) \\ 0 \text{ (for all other values of } n) \end{cases}$$

The modulation transfer characteristic is R_1, which is nearly linear for rV_m up to about $\pi/2$ radian. At this value the incident light power has been divided as follows: 51 per cent carrier (suppressed), 45 per cent upper sideband fundamental, 3.9 per cent double sideband second harmonic (suppressed), and 0.1 per cent lower sideband third harmonic.

The writers acknowledge with thanks the help given by Dr. E. M. Conwell in this work.

C. F. Buhrer
V. J. Fowler
L. R. Bloom
General Telephone & Electronics Labs., Inc.
Bayside, N. Y.

Inductance from a Field-Effect Tetrode*

In their paper[1] on the field-effect tetrode Stone and Warner show that suitable biasing will cause this device to display transfer admittances of opposite sign and thus to behave as a gyrator. Of chief interest is the gyrator's ability to simulate inductances by inverting capacitive admittances. In practice attempts to find bias conditions which give a useable inductance combined with a high Q at a reasonable operating frequency usually prove abortive. It is suggested here that biasing the tetrode to produce negative self-admittance may well result in interesting combinations of these three parameters. The model, terminology and basic assumptions are similar to those of Stone and Warner.

The following criteria were applied to calculated results in order to determine which were of practical importance:

a) Results which involve working frequencies outside the range 1 kc to 100 Mc were rejected.
b) A range of 1 pf to 0.1 μf was imposed on both the load capacitance and the capacitance which would be needed to tune the inductance produced. (This is approximately the range for capacitors realizable by microminiature techniques used in the production of thin film circuits and solid circuits.)

The fact that negative self-admittance can be obtained for the field-effect tetrode suggests that it should be employed to cancel, or nearly cancel, real components of the input admittance of the gyrator-capacitor combination to yield inductance of very high Q. Two possible arrangements present themselves for this treatment:

First Quadrant, Case I where $y_{11} < 0$,
Third Quadrant, Case II where $y_{22} < 0$.

[1] I. P. Kaminow, "Microwave modulation of the electro-optic effect in KH_2PO_4," *Phys. Rev. Lett.*, vol. 6, p. 528; May 15, 1961.

* Received April 24, 1962; revised manuscript received May 7, 1962.

[1] H. A. Stone and R. M. Warner, "The field-effect tetrode," Proc. IRE, vol. 49, pp. 1170–1184; July, 1961.

The load capacitor is assumed to be of infinite Q to keep the analysis simple, *i.e.*, $y_c = j\omega C$.

First Quadrant, Case I

The input admittance of the combination is given by

$$Y_{IN} = y_{11} - y_{12}y_{21}/y_{22} + y_c \quad (1)$$

and the condition that the real part of Y_{IN} is zero, to give an infinite Q emerges as

$$\omega^2C^2 = y_{22}(y_{12}y_{21} - y_{11}y_{22}). \quad (2)$$

The imaginary part of Y_{IN} is

$$-1/j \frac{(y_{22}^2 + \omega^2C^2)}{\omega C y_{12}y_{21}}. \quad (3)$$

As $y_{12}y_{21} < 0$ this quantity is positive and equivalent to $1/j\omega L$, the admittance of an inductance L. For the condition imposed by (2) the product ωC is constant for any chosen G_1 so that the inductance is completely defined by

$$\frac{L}{C} = (y_{22}^2 + \omega^2C^2)/\omega^2C^2y_{12}y_{21}. \quad (4)$$

Choose bias conditions $V_1 = 1.2W_p$, $V_3 = W_p$ and $G_1 = 10^{-2}$ mho. These quantities determine the admittance parameters of the tetrode and hence fix ω^2C^2 and L/C for the combination. In Table I C_1 is the capacitance needed to tune the inductance produced. Results for $G_1 = 10^{-1}$ mho are also included.

TABLE I

	C(pf)	$f(Q=\infty)$	L	C_1(pf)
$G_1 = 10^{-2}$ mho	10^4	142 Kc	0.6 mh	2100
	10^3	1.42 Mc	60 μh	210
	10^2	14.2 Mc	6.0 μh	21
	10	142 Mc	0.6 μh	2.1
$G_1 = 10^{-1}$ mho	10^4	1.42 Mc	6.0 μh	2100
	10^3	14.2 Mc	0.6 μh	210
	10^2	142 Mc	60 mμh	21

The stability of the device under these bias conditions needs further investigation in applications to amplifiers, oscillators and mixers.

Third Quadrant, Case II

A similar analysis performed for this case with a choice of $V_1 = -W_p$, $V_3 = -1.2W_p$ and $G_1 = 10^{-2}$ mho allows Table II to be constructed.

TABLE II

C(pf)	$f(Q=\infty)$	L	C_1(pf)
10^4	422 kc	0.55 mh	260
10^3	4.22 Mc	55 μh	26
10^2	42.2 Mc	5.5 μh	2.6
10	422 Mc	0.55 μh	0.26

Thus it would seem that there is sufficient theoretical evidence to justify building some devices with small initial conductance and attempting to measure them while biased in this particular manner.

Gyrator Design

The limitations of pinch-off in both channels together with the requirements on initial conductance suggest the following design for a silicon gyrator.

A 5 Ω cm n-type channel should be combined with a 12 Ω cm p-type channel having a common channel thickness of 3 microns. The pinch-off voltage then becomes 14.5 v and the bias voltages suggested become 14.5 v and 17.5 v, respectively. A reverse bias of 3 v should not put the junction in danger of breakdown. For the n-type channel an initial conductance $G_1 = 2.5 \times 10^{-3}$ mho (400 Ω) would be obtained with an aspect ratio of 0.01 (see Fig. 1).

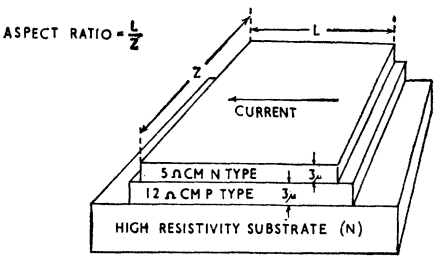


Fig. 1—Proposed construction for a gyrator.

Epitaxial techniques should be ideally suited for the fabrication of these otherwise delicate structures. The n-type floating substrate imposes no additional frequency limit as the capacitance of the junction it forms can be made small compared with that between the epitaxial layers.

The authors wish to thank the directors of the Plessey Company Ltd. for permission to publish this note and also the Ministry of Aviation for their support.

P. J. Etter
B. L. H. Wilson
The Plessey Company Ltd.
Caswell, Towchester
Northants, England

Generation and Radiation of Ultramicrowaves by Optical Mixing*

The presence of small nonlinearity in the dielectric constant of optically transparent materials has recently[1,2,3] been used for generation of harmonics and optical mixing of focused pulsed laser beams producing intense electric fields of the order of 10^5 volts/cm and above. The wavelength of the monochromatic light emitted by two lasers operating at different temperatures is separated a few angstroms apart with the frequency difference on the order of 500 to 2000 kMc. By using the nonlinearity of the medium it should be possible to generate this ultramicrowave beat frequency. The nonlinearity coefficient α in the dielectric constant expression $\epsilon = \bar{\epsilon}(1 + \alpha E)$ is usually quite small ($\sim 10^{-13}$ to 10^{-12} cm/volt) for optical harmonic mixing, but may be considerably larger for beat frequency generation, because ionic displacements play a larger role due to $f_1 - f_2$ being below the infrared absorption band of the material (between 3 to 10 teracycles/second for materials of interest).

The two input frequencies f_1 and f_2 (phase constants β_1, β_2, respectively) beating in the nonlinear dielectric medium cause a displacement current with phase constant $\beta_1 - \beta_2$ at the submillimeter beat frequency. This phase constant is, in general, different from the propagation constant

$$\left(\beta = \frac{\omega_1 - \omega_2}{c}\sqrt{\epsilon_{f_1 - f_2}}\right)$$

of waves at frequency $f_1 - f_2$. Due to the fact that the permittivity at ultramicrowaves is either larger or equal to that at optical wavelengths, *i.e.*, $\beta > \beta_1 - \beta_2$, radiation at submillimeter wavelengths from the displacement current at these frequencies is possible at an angle $\psi = \sin^{-1}(\beta_1 - \beta_2)/\beta$. In fact, the radiation pattern can be obtained from the theory of long-wire antennas,[4] and can be shown to have a polar coefficient of the form

$$\sin\theta \frac{\sin\left[\frac{\beta L}{2}(\sin\theta - \sin\psi)\right]}{\sin\theta - \sin\psi}.$$

Since a few centimeter lengths L of the medium provide a radiation path several hundred wavelengths long, extremely narrow radiated beams may be obtained at submillimeter frequencies (angular width of the major lobe $= (2\lambda/L)1/\cos\psi \sim 0.5$ to 1°). The power level at the maximum of the first sidelobe at $\psi \pm (3\lambda/L)1/\cos\psi$ is about 13.5 db below the principal maximum, independently of the actual value of L/λ as long as it is large. Since the nonlinear medium used is by definition transparent to the optical frequencies, use can be made of the multiple reflections of the input signals from the two end boundaries. In other words, the medium can be so dimensioned ($\beta_1 L = n_1\pi$, $\beta_2 L = n_2\pi$; where n_1 and n_2 are integers) and strongly reflecting end boundaries are provided to create a high Q resonator at the two incoming wavelengths. Under these conditions, the electric field strength E_{opt} at optical wavelengths is increased by a factor $1/(1-R)^{1/2}$, where R is the power reflection coefficient of the end boundaries at optical wavelengths (assumed equal for the two end boundaries, without any loss of generality). Since the displacement current at submillimeter frequencies is proportional to $(E_{opt})^2$ and the radiated power is equal to (displacement current)^2 × radiation resistance R_r, an improvement of $1/(1-R)^2$ may be achieved in the radiated power by this means.

* Received June 8, 1962.

[1] P. A. Franken, *et al.*, "Generation of optical harmonics," *Phys. Rev. Lett.*, vol. 7, pp. 118; August, 1961.
[2] M. Bass, *et al.*, "Optical mixing," *Phys. Rev. Lett.*, vol. 8, pp. 18; January, 1962.
[3] J. A. Giordmaine, "Mixing of light beams in crystals," *Phys. Rev. Lett.*, vol. 8, pp. 19; January, 1962.
[4] S. A. Schelkunoff, "Electromagnetic Waves," D. Van Nostrand Co., Inc., New York, N. Y.; 1948.

The total radiated power is given by

$$P_0 = \frac{1}{(1-R)^2}\,\frac{\eta}{2\pi}\left(2(\omega_1-\omega_2)\bar{\epsilon}\alpha P_{in}\frac{\eta}{\sin\psi}\right)^2 \cdot \int_0^{\pi/2} G^2(\theta)\sin\theta\, d\theta$$

where η is the characteristic impedance of the medium at frequency f_1-f_2, $\bar{\epsilon}$ the permittivity of the medium and $G(\theta)$ the polar coefficient of radiation for the above arrangement of multiple reflections.

$$G(\theta) = \sin\theta\left[\frac{\sin\left(\frac{\beta L}{2}(\sin\psi-\sin\theta)\right)}{\sin\psi-\sin\theta} + \frac{\sin\left(\frac{\beta L}{2}(\sin\psi+\sin\theta)\right)}{\sin\psi+\sin\theta}\right]$$

The dimensions of the medium are assumed much larger than the wavelength $\lambda=2\pi/\beta$ to neglect the induction field in the field expressions. In order to form an idea of the order of magnitude involved, let us assume the following parameters:

$f_1-f_2=10^{12}$ c/s
$P_{in}=6$ Kw(3 Joule, $\frac{1}{2}$ msec. pulse)
$\alpha=10^{-10}$ cm/volt
$R=0.99$
$R_r=200$ ohms.

The radiated power P_0 is calculated to be 45 mw. The value of α at these frequencies is not known at present and may be somewhat lower than the value assumed. The above presented scheme of nonlinear mixing provides a simultaneous generator and highly directional radiator at ultramicrowaves. The edge boundaries of the medium are shaped so as to allow the radiation out of the system.

The author is grateful to Dr. S. K. Ghandhi for his interest and encouragement during the course of this work.

Om P. Gandhi
Philco Scientific Laboratory
Blue Bell, Pa.

Tunnel Diode Loaded by a Shorted Transmission Line*

Nagumo and Shimura[1] have made a thorough analysis of the tunnel diode loaded by a shorted transmission line. Unfortunately, some readers may have trouble following this worthy paper in one place due to an equation that appears to have suffered a misprint and in other places due to apparent errors in some of the figures.

Since $\Delta t=T=2l/w$, the next to last equation in column one of page 1287 should read

$$\Delta t = T = 2l\sqrt{LC}.$$

The following equation is correct as it stands indicating that this error is only an unfortunate misprint.

More serious is the confusion created by some of the figures, of which Fig. 21 is a good example. In order to see that the path indicated is incorrect, let us repeat the analysis in more physical terms than used by Nagumo and Shimura. Consider, as an initial condition, that the transmission line in Fig. 1 is completely discharged. Then an equivalent circuit is that shown in Fig. A, where Z is purely resistive since the line is assumed lossless. The resulting operation point P_1 can be determined by drawing a load line corresponding to Z and E on the characteristic curve of the tunnel diode as shown in Fig. B. (The problem of bistability has been side-stepped for the time being by choosing a monostable load line for the first example.)

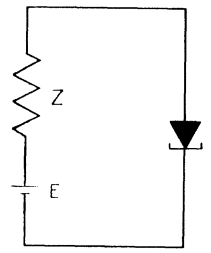


Fig. A—The equivalent circuit.

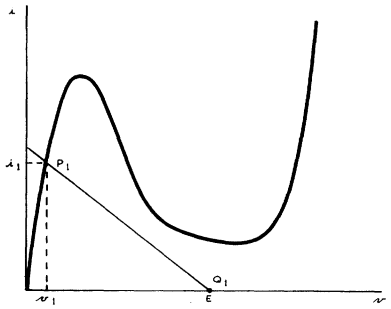


Fig. B—Determination of the first operation point, P_1.

Operation at P_1 is obtained by impressing a negative voltage (v_1-E) and a positive current (i_1) on the line. These polarities are consistent with a wave traveling away from the diode. When this wave reaches the shorted end of the line, it is reflected with inverted voltage and travels back to the diode. Now, in general, there will be a reflection at the diode, and a third wave will travel down the line toward the short. This third wave is determined by the fact that the sum of the three waves must be consistent in voltage and current with the characteristic of the diode. Considering at first only the first two waves plus E, the combined voltage will be E, and the combined current will be $i_1+(i_1-0)=2i_1$. In addition to this there will be the voltage and current of the third wave which have the ratio $-Z$. The second operating point P_2 can then be found by drawing a second load line through point Q_2 as shown in Fig. C. Now, as before, the third wave, having a positive voltage (v_2-E) and a negative current (i_2-2i_1), will be reflected at the shorted end of the line, reversing its voltage to become a fourth wave, and the fourth wave is reflected at the diode creating a fifth wave. The third operating point P_3 can be found by drawing a third load line through the point Q_3 whose coordinates are $v=E$, $i=i_2+(i_2-2i_1)=2i_2-2i_1$. The process may be continued indefinitely to trace the action of the circuit through as many cycles as may be desired.

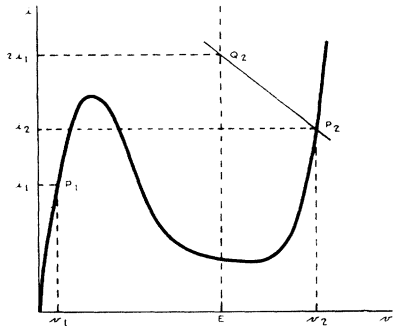


Fig. C—Determination of the second operation point, P_2.

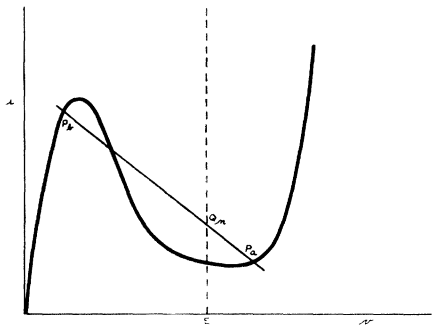


Fig. D—A load line showing bistability.

Before taking up the question of bistability, let us compare the results so far with those of Nagumo and Shimura. We find that we have drawn a set of load lines Q_nP_n having slopes $-1/Z$ in the v, i plane and with all the points Q_n lying on the line $v=E$. We can also draw in connectives P_nQ_{n+1} having slopes $+1/Z$, in which case we find that $Q_nP_nQ_{n+1}$ always forms a half diamond. If we take the trouble to normalize the variables in the manner of Nagumo and Shimura, we find that in the ξ, η plane the above lines have slopes of ± 1 (45°) and the points Q_n all lie on the η axis in agreement with Nagumo and Shimura.

Now consider a point Q_n and load line as shown in Fig. D. Do we take P_a or P_b to be P_n? So far we have considered the tunnel diode to be an ideal diode having no capacitance. Let us be more realistic and consider that there is a capacitance but that it is too small to make the rise times appreciable. Now if Q_n lies above the diode characteristic, as shown, the sum of the $2n-2$ waves considered thus far have more current flowing toward the diode than at that instant (before the reflection) is taken by the diode. The difference must flow into the capacitance, raising the voltage, and the circuit moves to point P_n. If Q_n had been below the diode

* Received September 18, 1961.
[1] J. Nagumo and M. Shimura, "Self-oscillation in a transmission line with a tunnel diode," Proc. IRE, vol. 49, pp. 1281–1291; August, 1961.

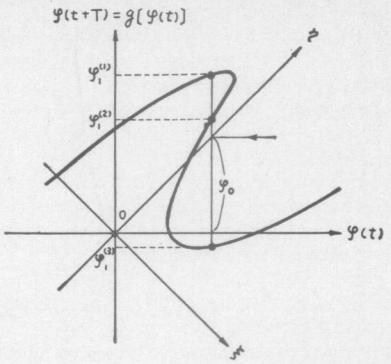


Fig. 1—The three roots of the equation $G(\Phi, \phi_0, 0) = 0$ are $\Phi = \phi_1^{(1)}$, $\phi_1^{(2)}$ and $\phi_1^{(3)}$.

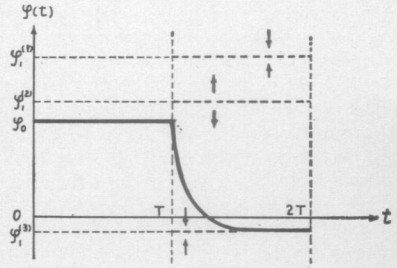


Fig. 2—$\phi(t+T)$ for $0 \leqq t \leqq T$ approaches to $\phi_1^{(1)}$ if $\phi_0 > \phi_1^{(2)}$, and to $\phi_1^{(3)}$ if $\phi_0 < \phi_1^{(3)}$.

characteristic, the circuit would have had to move in the direction of lower voltage coming to a point such as P_b. Thus the (geometric) reflection (of the construction lines) at Q_n is to the right if Q_n lies above the diode characteristic and to the left if Q_n is below the diode characteristic. Examining Fig. 21 we find that the reflection is incorrect at the next to bottom such point. Similar mistakes can be found in several other figures.

MALCOLM H. STEWARD
Westwood Division
Houston Fearless Corp.
Los Angeles, Calif.

Authors' Comment[2]

The first point which is indicated by Steward is apparently our careless mistake. The next to last equation in column one of page 1287 should read

$$\Delta t = T = 2l\sqrt{LC}$$

as pointed out by Steward.

Concerning the second point, which is more serious, Steward's conclusion seems to be incorrect. The reason is as follows.

Let us consider the differential-difference equation (38) in the case where the function g is triple valued. Let the three roots of the equation of Φ

$$G(\Phi, \phi_0, 0) = 0$$

[2] Received October 9, 1961.

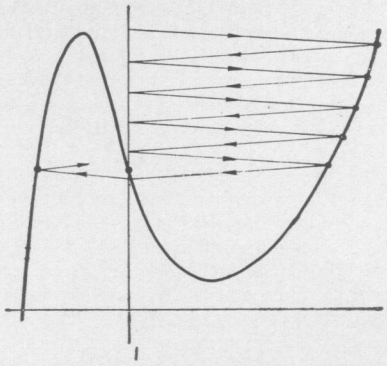


Fig. 3—If Steward's conclusion is correct, the operating points ought to behave as shown in this figure.

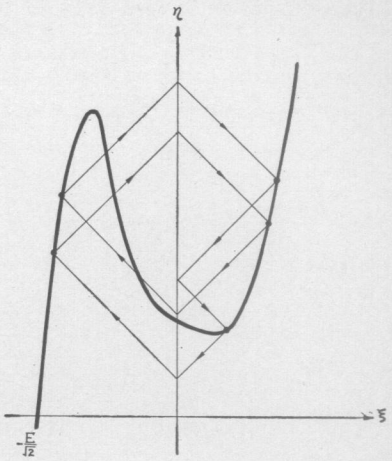
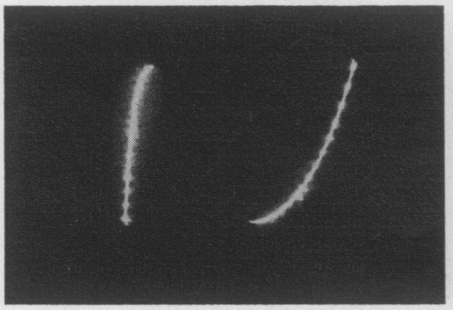
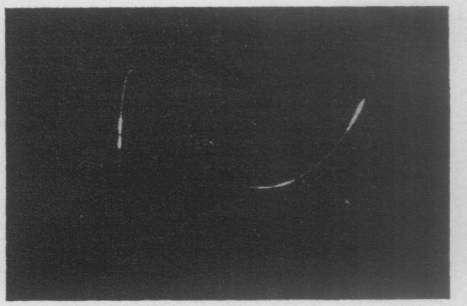


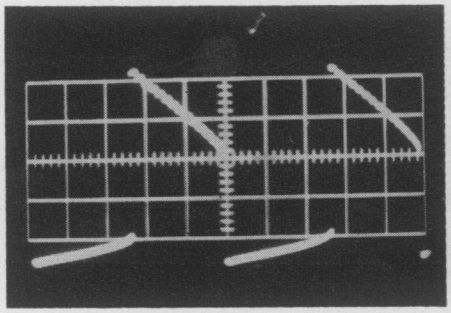
Fig. 5—Fig. 21 must be replaced by this figure.



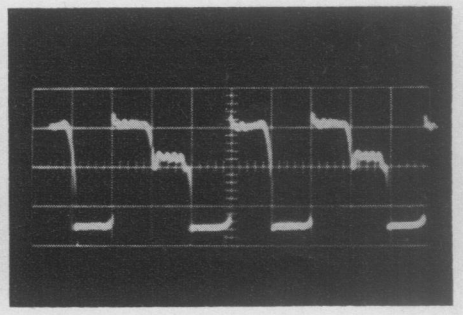
(a)



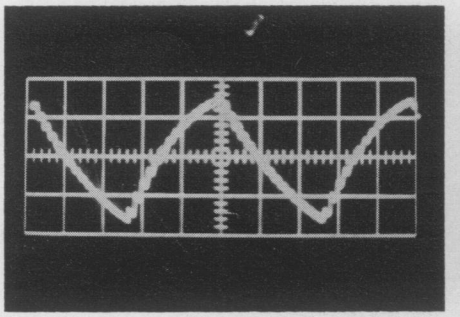
(a)



(b)

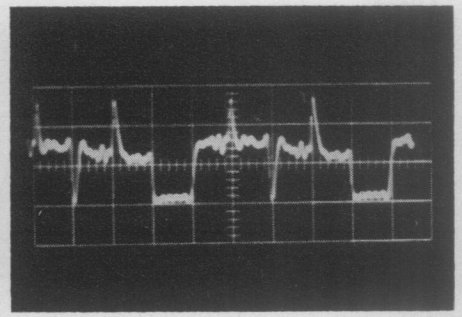


(b)



(c)

Fig. 4—The case $m=16$, $n=15$. $Z=1$ kΩ, $T/2=100$ μsec, $E=0.14$ volt. (a) The V-I characteristic. (b) The voltage waveform. 0.05 volt/division (vertical), 1.2 msec/division (horizontal). (c) The current waveform. $1m$ amp/division (vertical), 1.2 msec/division (horizontal).



(c)

Fig. 6—The case $m=2, n=3$ [(1:1)⇄(1:2)] $Z=220$ Ω, $T/2=450$ μsec, $E=0.28$ volt. (a) The V-I characteristic. (b) The voltage waveform. 0.1 volt/division (vertical), 900 μsec/division (horizontal). (c) The current waveform. $2m$ amp/division (vertical), 900 μsec/division (horizontal).

be $\Phi=\phi_1^{(1)}, \phi_1^{(2)}$ and $\phi_1^{(3)}$ ($\phi_1^{(1)}>\phi_1^{(2)}>\phi_1^{(3)}$), where ϕ_0 is a constant (see Fig. 1).

If the initial condition of the differential-difference equation (38) is assumed as

$$\phi_0(t) = \phi_0\text{: constant} \qquad \text{for } 0 \leqq t \leqq T,$$

(38) reduces to a differential equation of the first order:

$$C_0 \frac{d\phi(t+T)}{dt} = G(\phi(t+T), \phi_0, 0) \qquad \text{for } 0 \leqq t \leqq T. \quad \text{(a)}$$

Since $G(\phi_1^{(i)}, \phi_0, 0)=0$ for $i=1, 2$ and 3, this differential equation has three constant solutions:

$$\phi(t+T) = \phi_1^{(1)}, \phi_1^{(2)}, \phi_1^{(3)}, \quad \text{for } 0 \leqq t \leqq T.$$

From the fact that $G(\Phi, \phi_0, 0)<0$ for $\Phi>\phi_1^{(1)}, \phi_1^{(2)}>\Phi>\phi_1^{(3)}$ and $G(\Phi, \phi_0, 0)>0$ for $\phi_1^{(1)}>\Phi>\phi_1^{(2)}$, $\Phi<\phi_1^{(3)}$, we know that the solutions $\phi_1^{(1)}$ and $\phi_1^{(3)}$ are stable, while $\phi_1^{(2)}$ is unstable. Therefore, we can conclude that the solution of (a), $\phi(t+T)$ for $0 \leqq t \leqq T$, approaches to $\phi_0^{(1)}$ if $\phi_0>\phi_1^{(2)}$, and to $\phi_1^{(3)}$ if $\phi_0<\phi_1^{(2)}$. It is easily seen that this conclusion is equivalent to Steward's (see Fig. 2).

It must be pointed out, however, that this conclusion is based on the assumption that "$\phi_0(t)$ is constant for $0 \leqq t \leqq T$," which is implicitly assumed in Steward's physical discussion too. If $\phi_0(t)$ is not constant for $0 \leqq t \leqq T$, the above-mentioned conclusion does not necessarily hold. Even if the initial condition is assumed to be constant for the first period, we cannot make such an assumption for the next period because of the parallel capacitance C_0 of the tunnel diode. The effect of the term $C_0(d\phi(t)/dt)$ in (38) cannot easily be estimated as done in Steward's discussion.

To make this situation definite, let us consider the case where Z is large. If Steward's conclusion is assumed to be correct, the operating points ought to behave as shown in Fig. 3, which is obviously not true, as Fig. 4 of our experimental result shows.

However, Fig. 21 is our mistake by itself and we must replace it by Fig. 5 below. An experimental result which corresponds to it it shown in Fig. 6.

J. NAGUMO
M. SHIMURA
Dept. of Applied Physics
Faculty of Engineering
University of Tokyo
Tokyo, Japan

Mr. Steward's Reply[3]

In reply to the authors' comment I must maintain that my conclusions are not incorrect for the case assumed, which was that there is no capacitance across the tunnel diode. My wording was, " . . . there is a capacitance but that it is too small to make the rise times appreciable." In other words, the concept of capacitance was introduced only to the point of determining the choice between P_a and P_b. It was still to be considered that the voltages and currents were step functions. Nagumo and Shimura seem to alternately agree and disagree with my conclusions. The reason for that is that they are not always clear as to whether capacitance is being neglected or not, and the presence of capacitance in non-negligible amounts can make impressive differences in the results.

In their original paper (4) implicitly states that capacitance is to be neglected, and nothing in the text changes this assumption until the beginning of Section VI. All figures up to and including Fig. 29 are presumed by the reader to be based on the assumption of no capacitance, and on the basis of this assumption, many of them, including the revision shown in Fig. 5 of their comment, are incorrect. Toward the end of their comment, the authors reveal the cause of the confusion. Without ever stating the amount of capacitance involved, they have been drawing their figures so as to agree as far as possible with experimental results. The large number of discrepancies obtained indicate the strong effects of practical amounts of capacitance.

In Section VI of their original paper the authors state the equations governing the case in which capacitance is considered, but do not explain the effects of this capacitance on the results. I must confess I had not given this much thought either, until I read their comment. In order to get a qualitative feel for the importance of practical amounts of capacitance, let us study Fig. E below.

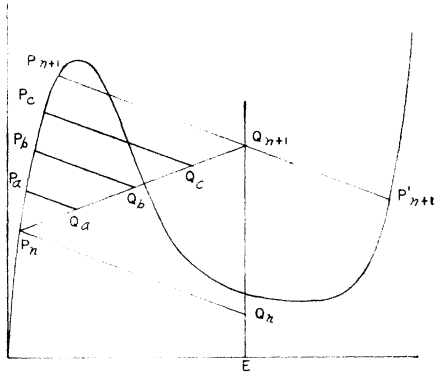


Fig. E—A moving loadline as when rise time is non-negligible.

Suppose we have arrived at P_n, and the next wave coming toward the diode takes us to Q_{n+1}. If the rise time of this wave is negligible, the next operation point must be P_{n+1}'. However, if, as a result of smoothing introduced in previous periods, the rise time is appreciable, we may think of the voltage and current applied to the diode as moving gradually from P_n through Q_a, Q_b, and Q_c to Q_{n+1}. During this time the diode capacitance will be charged by the difference between the diode and line currents. If the applied voltage and current jumped from P_n to Q_a and then stopped, the effect of the charging would be to move the operating point from Q_a along the road line Q_aP_a toward P_a. It can therefore be seen that the actual operation point will follow some curve from P_n to a point on Q_aP_a, to a point on Q_bP_b, to a point on Q_cP_c, to, finally, either P_{n+1} or P_{n+1}'. If the capacitance charges rapidly enough, compared to the rise time of the applied wave, to keep the operating point under the diode characteristic, the end point will be P_{n+1}. If not, the changing current will reverse part way through the process, and the end point will be P_{n+1}'. The motion of the operating point away from P_nQ_{n+1} generates a new wave on the transmission line traveling away from the diode to continue the process of oscillation. The rise time of this wave will correspond to the charging time of the diode capacitance.

If, as assumed above, the charging time is similar to the rise time (of the applied wave, not the new wave), it appears that the operation point will stay on the same branch of the diode characteristic until this is no longer possible because of passing the peak point or the valley point. Since both times are determined by the same capacitance, it appears that this might be a common condition and explains most of the figures shown by Nagumo and Shimura. On the other hand, if the inductance in series with the diode becomes appreciable, the result could be the other way around again. The situation can obviously get quite complicated. If a prediction of circuit performance is required, the best approach is to add the effect of inductance to (34) and (35) and then solve using the actual circuit values. These equations become

$$i(l,t) = f\left(E + v(l,t) - L\frac{\partial i(l,t)}{\partial t}\right) + C_0\frac{\partial v(l,t)}{\partial t}$$

$$\phi_1\left(t-\frac{T}{2}\right) + \phi_1\left(t+\frac{T}{2}\right) = Zf\left(E + \phi_1\left(t-\frac{T}{2}\right) - \phi_1\left(t+\frac{T}{2}\right) - L\phi_1'\left(t-\frac{T}{2}\right) - L\phi_1'\left(t+\frac{T}{2}\right)\right) + ZC_0\phi_1'\left(t-\frac{T}{2}\right) - ZC_0\phi_1'\left(t+\frac{T}{2}\right).$$

In conclusion, it appears that the figures shown by Nagumo and Shimura are probably correct in real life, but not for the simplified case assumed in most of their paper. In spite of this, their paper is very enlightening and draws many interesting conclusions of practical importance.

Authors' Reply[4]

In case the reactance of the tunnel diode is completely neglected, the circuit behavior is governed by and only by the difference equation (18) and the stability condition (25). If there exist many solutions which satisfy (18) and (25), therefore, it is impossible to decide which one of them is actually chosen to occur.

Nevertheless, only one circuit behavior is observed in actual fact. To give a theoretical explanation of this fact, it is necessary to consider the reactance of the tunnel diode. As the first step toward the attempt, we considered the parallel capacitance of the tunnel diode. We thus obtained the differen-

[3] Received December 4, 1961.

[4] Received January 8, 1962.

tial-difference equation (38) and defined the stability condition mentioned in our paper (Section VI). To our regret, we can not give complete discussion on this problem now.

Many figures displayed in our paper are drawn to show solutions of the difference equation (18) satisfying the stability condition (25), which coincide with our experimental results. It is expected, however, that the occurrence of these solutions displayed in our paper will be theoretically confirmed when the full treatment of the differential-difference equation (38) is accomplished.

Furthermore, it will be necessary in some cases to consider not only the parallel capacitance but also the series inductance of the tunnel diode, as pointed out by Steward.

Eventually, Steward's conclusion mentioned in his Reply does not seem to be in discord with our opinion. We express sincere thanks to Steward for his valuable discussions.

J. Nagumo
M. Shimura

A Total-Reflection Solid-State Optical-Maser Resonator*

Various schemes for solid-state optical maser oscillators have been proposed and are being tested in a number of laboratories.[1-6] Initially only pulsed operation was possible. More recently, attempts to set up continuously operating solid-state optical masers using special pumping techniques and cooling have been successful.[7,8]

In this communication a novel resonator and output coupling configuration for an optical maser is described. This configuration differs from others in current use in that it uses total internal reflection to reflect the light beam and frustrated total reflection for output coupling.

One possible configuration of the total reflection resonator is shown in Fig. 1. As the figure shows, the adjacent end walls of the resonator are at right angles to each other. If the resonator is placed in an optically rarer medium such that the ratio of the refractive indexes of the resonator medium and the external medium is greater than $\sqrt{2}$, a light beam in a direction parallel to the side walls will be totally reflected from the end walls and confined within the crystal. This eliminates the need for dissipative metallic or multilayer dielectric coatings. Power is abstracted from the resonator through frustrated total reflection, that is, by coupling to the rapidly decaying fields in the optically rarer medium associated with the phenomenon of internal reflection. This can be achieved by placing the hypotenuse of a right-angle prism parallel to and in close proximity of one of the totally reflecting end walls. The arrangement is shown schematically in Fig. 2. Frustrated total reflection will occur if the refractive indexes of the coupling prism and the coupling medium (relative to that of the resonator medium) are properly chosen. (It is readily found that for frustrated total reflection to occur, $n_2 < n_1 \sin \alpha_1 < n_3$, where n_1, n_2 and n_3 are the refractive indexes of the resonator, coupling medium and coupling prism, respectively, and α_1 is the angle of incidence.) The amount transmitted power depends on the angle of incidence α_1, on the ratio of the refractive indices n_2 and n_3 of the couping medium and the coupling prism, respectively, to the refractive index n_1 of the resonator, and on the thickness d of the coupling layer. It should be clear that the medium of the coupling layer need not be the same as the medium surrounding the other reflecting surfaces of the resonator.

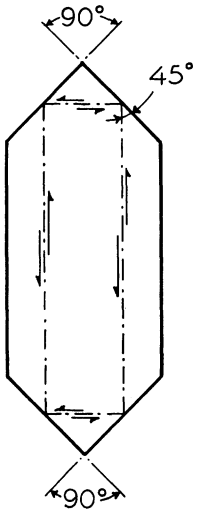


Fig. 1—The total reflection optical-maser resonator.

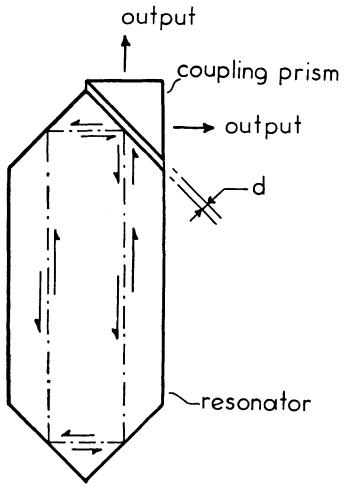


Fig. 2—The total reflection optical-maser resonator with coupling prism.

The threshold requirements for optical-maser oscillation are determined by the requirement that the amplification by stimulated emission be sufficient to compensate for the losses.[9] Thus, both the threshold concentration of active atoms in the host crystal and the pumping power required to maintain the threshold number of excited atoms are determined primarily by the losses. These consist of the scattering by the inhomogeneities in the rod material, the end-wall absorption, diffraction and scattering losses, and the useful output. In the total reflection resonator the end-wall losses are substantially eliminated. As a result the threshold pumping power is reduced and the available output is increased.

Another, perhaps more important, aspect of the reduced losses lies in the lower threshold concentration of active atoms. Since the light quanta emitted either by spontaneous or induced emission are necessarily smaller than the input pumping quanta, a substantial fraction of the pumping power is dissipated as heat within the host material. This dissipated heat is proportional to the concentration of active atoms. A lower concentration will thus result in a lower heat dissipation so that continuous operation without elaborate cooling and pumping techniques becomes feasible.

The total-reflection resonator exhibits another interesting property. All the other optical-maser resonator cavities use either metallic or multi-layer dielectric coating of the end walls to normally reflect the light beam. As a result a definite phase relationship exists between the two contradirectionally traveling wave systems. No such relationship is established by the boundary conditions existing in the total reflection resonator. It would therefore appear that this configuration can support two mutually independent contradirectionally traveling resonant wave systems. The only process that might couple these two wave systems is the emission process. It would therefore be of interest to determine if any correlation exists between these wave systems. Since power from the two contradirectionally traveling wave systems can be brought out separately from the two different faces of the coupling prism such an experiment can be readily performed.

L. Bergstein
W. Kahn
Polytechnic Institute of Brooklyn,
Brooklyn, N. Y.
C. Shulman
Polytechnic Institute of Brooklyn, and
The School of Technology,
The City University
New York, N. Y.

* Received June 22, 1962. The work reported herein was sponsored by the U. S. Army Signal Research and Development Labs., Office of Naval Research and the Air Force Office of Scientific Research, under Contract No. AF-18 (600)-1505.

[1] A. L. Schawlow and C. H. Townes, "Infrared and optical masers" *Phys. Rev.*, vol. 112, pp. 1940–1949; December, 1958.

[2] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, pp. 493–484; August, 1960.

[3] R. J. Collins, D. F. Nelson, A. L. Schawlow, W. Bond, C. G. B. Garret, and W. Kaiser, "Coherence, narrowing, directionality, and relaxation oscillations in the light emission from ruby," *Phys. Rev. Lett.*, vol. 5, pp. 303–305; October, 1960.

[4] G. E. Devlin, J. McKenna, A. D. May, and A. L. Schawlow, "Composite rod optical masers," *Appl. Optics*, vol. 1, pp. 11–15; January, 1962.

[5] L. F. Johnson, G. D. Boyd, and K. Nassau, "Optical maser characteristics of Tm^{+3} in $CaWO_4$," Proc. IRE (*Correspondence*), vol. 50, pp. 86–87; January, 1962 .

[6] L. F. Johnson, G. D. Boyd, and K. Nassau, "Optical maser characteristics of Ho^{+3} in $CaWO_4$," Proc. IRE (*Correspondence*), vol. 50, pp. 87–88; January, 1962.

[7] L. F. Johnson, G. D. Boyd, K. Nassau, and R. R. Soden, "Continuous operation of the $CaWO_4:Nd^{+3}$ optical maser," Proc. IRE (*Correspondence*), vol. 50, p. 213; February, 1962.

[8] D. F. Nelson and W. S. Boyle, "A continuously operating ruby optical maser," *Appl. Optics*, vol. 1, pp. 181–183; March, 1962.

[9] A. L. Schawlow, "Optical and infrared masers," *Solid State J.*, vol. 2, pp. 21–29; June, 1961.

Current-Voltage Characteristic of Tunnel Junctions*

Tunneling through thin insulating films, discovered some thirty years ago, has been studied extensively in connection with tunnel diodes.[1] Recently, it has been pointed out[2] that, in addition to tunneling, there are other modes of electron transport through thin films. We have been studying symmetrical current-voltage characteristics of metal-insulator-metal junctions, of the type illustrated in Fig. 1, which was made with an Al-Al_2O_2-Cu junction at room temperature and atmospheric pressure. The natural oxide coating of commercial Al foil was used in this case, but similar characteristics are observed with thicker oxide films formed by anodizing. The slope at the origin is that of a 1330-ohm resistor, and an empirical formula $I=(V+0.625V^3)/1330$ represents the curve, apart from the "hysteresis."

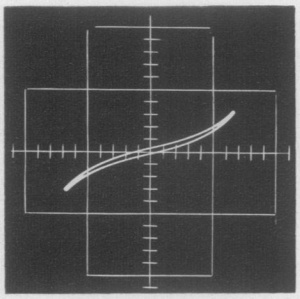


Fig. 1—Current-voltage characteristic of metal-insulator-metal tunneling. Maximum current, 1.59 ma, maximum voltage, 1.15; slope at origin indicates 1330 ohms; frequency, 10 kc.

There is no known theory for the hysteresis displayed. It is not attributable to capacitance in the usual sense, as shown by varying the frequency of the applied alternating voltage. We have observed curves enclosing more area, and curves with less are common. Open loops, with less included area, have been observed at 0.05 cps by Handy,[3] who formed junctions by air oxidation of evaporated Al films, on which various other metals were subsequently deposited by evaporation. No theory has been proposed to account for the hysteresis.

Our junctions are quite unstable, being formed by pressing the oxide-coated foil lightly against a flattened wire of Cu, Al or Au.

By tapping the support, an initially open junction is brought into the tunneling mode; an open or a short circuit may occur suddenly by further jarring or apparently spontaneously.

We are attempting to form more stable structures, to distinguish between different modes of transport, and to investigate the observed hysteresis.

H. P. Knauss
R. A. Breslow
Department of Physics
The University of Connecticut
Storrs, Conn.

* Received June 11, 1962. This work was supported in part by a grant from the Research Corporation.

[1] N. Holonyak, Jr., and I. A. Lesk, "Gallium arsenide tunnel diodes," Proc. IRE, vol. 48, pp. 1405–1409; August, 1960.

[2] P. R. Emtage and W. Tantraporn, "Schottky emission through thin insulating films," *Phys. Rev. Lett.*, vol. 8, pp. 267–268; April 1, 1962.

[3] Robert M. Handy, "Electrode effects on aluminum oxide tunnel junctions," *Phys. Rev.*, vol. 126, p. 1968; June, 1962. The authors are also indebted to Dr. Handy for a copy of his Ph.D. dissertation, Northwestern University, Evanston, Ill., 1961, unpublished.

A High-Speed Point Contact Photodiode*

Proposed optical communication systems require photodetectors capable of broad-band operation to frequencies in excess of 10^9 cps. A germanium formed point contact photodetector[1] that responds to the beat frequencies of the He-Ne gas maser up to 900 Mc has been constructed. The device should be capable of operation at frequencies up to 50 Gc.

This germanium photodetector was constructed using a P^+ body with an epitaxially grown π layer. The π layer is approximately 8 microns thick. A 2-mil diameter mesa was chemically etched. 0.6-mil arsenic doped gold foil was pointed and brought into contact with the π region in the center of the mesa. A short-current pulse was used to alloy the point material into the π layer. The alloyed region is approximately 4 microns in diameter and penetrates the π layer to a depth of 4 to 5 microns. The diode structure is shown in an expanded scale in Fig. 1.

Light is incident at the side of the π region as indicated in Fig. 1. The *V-I* characteristics of a typical diode are indicated in Fig. 2.

The operation of this type of photodetector was discussed by Gaertner.[2] The device is basically a *PIN* structure. With reverse bias the depletion layer of the junctions (*P-I* and *I-N*) spread out to occupy the entire *I* or in our case π region. This π region becomes a region of high and nearly constant electric field. Photons of sufficient energy to excite hole-electron pairs are incident on the π region. The generated carriers move with terminal velocity across the swept π region and are collected at the P^+ and N^+ regions. Since the carrier lifetimes in the P^+ and alloyed *N* region or N^+ region are quite small, typically 10^{-10} sec or less, these collected charges recombine quickly. The terminal velocity in the π region is approximately 10^7 cm/sec under a reverse bias condition of 3 volts or more.

An estimate of the ultimate performance can be obtained. The active π region (Fig. 1) is about 4 microns thick. With the terminal velocity of 10^7 cm/sec, photon excited carriers will transit this region in times less than 4×10^{-11} sec. The capacity can be estimated by considering the device as a parallel plate condenser. With a π layer of 4-micron thickness and a 4-micron diameter point region, the capacity is 4.5×10^{-4} pf (ϵ for germanium is 1.42×10^{-10} farad/m). The measured series resistance of a typical diode is about 10 ohms. The RC cutoff frequency will be on the order of 10^{13} cps. The transit time will thus limit the frequency response of the detected output to the 25- to 50-Gc region.

Tests were conducted with this diode mounted without encapsulation across the coaxial structure shown in Fig. 3. The output of the HE-NE gas maser at a wavelength of 1.153 microns was focused through the small hole in the side of the coaxial structure onto the π region in the vicinity of the point. A reverse bias of 3 to 4 volts was used. Both short-circuit and open-circuit tuning stubs

* Received June 15, 1962.

[1] R. P. Riesz of the Bell Telephone Labs. has submitted a report for publication to the *Journal of Applied Physics* detailing the construction and operations of a P-I-N junction photodetector.

[2] W. W. Gaertner, *Phys. Rev.*, vol. 116, no. 84, 1959.

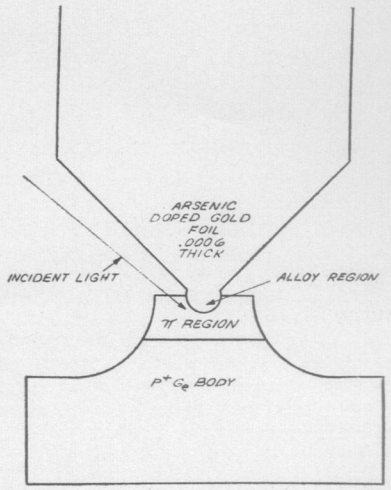


Fig. 1—Point contact photodiode structure.

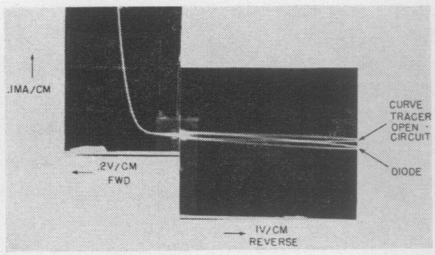


Fig. 2—Photodiode *I-V* characteristics.

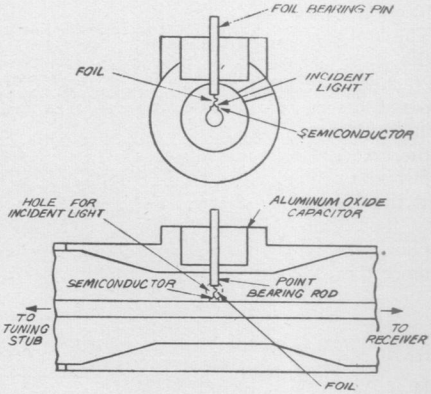


Fig. 3—Coaxial photodiode mount.

were placed on one coaxial terminal. The second terminal was connected to a conventional microwave receiver consisting of a tuned preselector and a single diode mixer followed by a 30-Mc IF amplifier.

The several modes present in the maser output were mixed in the diode to produce RF beat frequencies at 117 Mc. Beat frequencies as high as 936 Mc were detected. The output level of the 117-Mc beat frequency with proper adjustment of the RF tuning stub was comparable with that obtained from a 7102 photomultiplier. This high speed photomultiplier has a cathode response peaked near 8500 Å.

The author wishes to express his thanks to A. Kuper who supplied the germanium and to acknowledge the helpful discussion with R. P. Riesz of the Bell Telephone Laboratories.

L. U. Kibler
Bell Telephone Labs., Inc.
Holmdel, N. J.

A Method for Calibration of Laser Energy Output*

The development of pulsed high power ruby lasers has been hampered by the lack of accurate power level measurements. A calorimetric method has been reported on.[1] The following is a description of a method for calibrating the energy output of a laser; this method is especially applicable at low energy output, *i.e.*, near threshold.

The laser energy output is calibrated by attenuting the laser beam with neutral filters, directing it into a phototube (with S-1 spectral response), and integrating the current produced. This charge is directly dependent on the number of light quanta, and therefore the light energy.

The filters are calibrated by dividing the laser beam in two using a semireflective mirror (Fig. 1). The two beams are directed into separate phototube sensors. A dual-beam oscilloscope displays both signals; the display is then photographed, the camera shutter being appropriately synchronized with the firing of the laser. A series of photographs is taken, with various attenuating filters in one of the paths of the laser beam. The two displays are equalized by adjusting the oscilloscope preamplifier gains [Fig. 2(a)]. The ratio of these gains is then compared to the ratio obtained with no filters in the beams [Fig. 2(b)]. The resultant is the attenuation of each neutral filter. The filters are thus calibrated with a laser beam at the color at which they are to be used.

The laser beam is allowed to enter one phototube with intensity peaks yielding more than 200 μa while the reflected beam is attenuated to less than 10 μa. The output signals have the same complex shape, indicating linearity of phototube and associated circuitry.

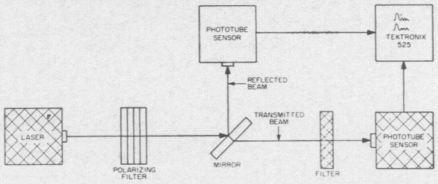
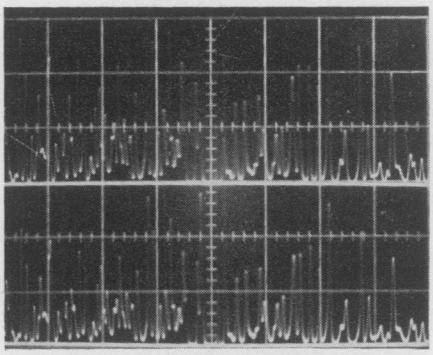
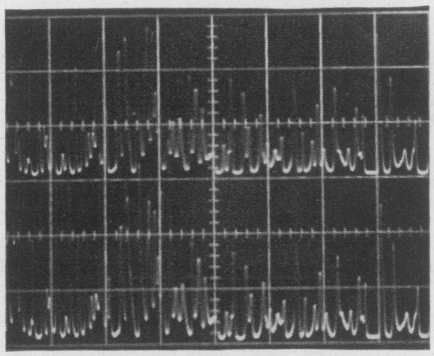


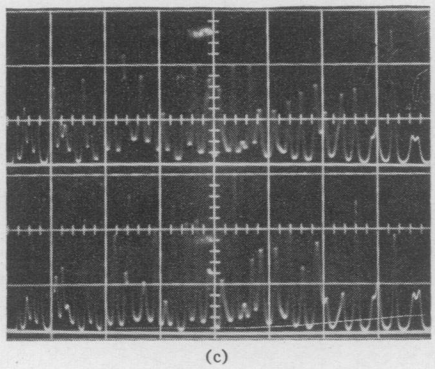
Fig. 1—Filter calibration. Laser output measurement uses only cross-hatched items.



(a)



(b)



(c)

Fig. 2—Reflected and transmitted laser beams, 20 μsec/cm. (a) Polarized; upper, transmitted, 100 mv/cm; lower, reflected, 50 mv/cm. (b) Polarized; upper, transmitted through B filter, 50 mv/cm; lower, reflected, 100 mv/cm. (c) Unpolarized; upper, relected, 50 mv/cm; lower, transmitted, 50 mv/cm.

Fig. 2(c) indicates the output of the laser beam beyond the mirror. The detailed waveforms of each pulse appear different. Investigations indicate that this is due to the fact that each pulse is differently polarized. The beam-splitting mirror favors certain polarizations for transmission and reflection. In order to achieve identical signals in both channels, it was found necessary to remove this polarization effect. By inserting a polarizing filter[2] after the laser, linearly polarized light is available to the mirror. The ratio of intensities of the two beams is determined by subtracting the signals in a differential amplifier. By varying the angle of the polaroid and mirror, the ratio of intensity of the split signals is adjusted, the difference signal being nulled out. Apparently the polarization of the laser beam is dependent on the output intensity since the nulled-out condition changes with input energy.

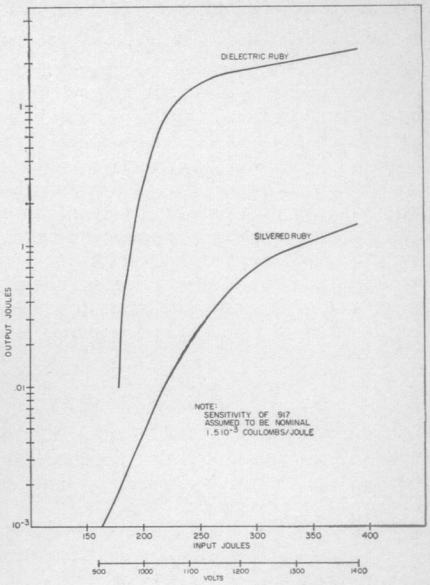


Fig. 3—Output energy curves.

The sensor is a RCA 917 phototube whose load impedance is 4.0 K ohms. The signal is amplified, integrated by a capacitor, followed by a peak detector.

The calibration of the integrator is dependent on the sensitivity of the photo cathode surface. The integrator design is such that a pulse of 0.60 v amplitude and 120 μsec length is integrated to 4 v. A current of 0.15 ma through 4.0 K will yield 0.60 v. This current for 120×10^{-6} sec determines a charge flow of 1.8×10^{-8} coulombs. The sensitivity of the S-1 cathode is approximately 1.5×10^{-3} coulombs/joule;[3] thus 1.2×10^{-5} joules yields a 4-v integrated output.

Fig. 3 indicates the results of calibration of an early Raytheon laser design, Model LH-1, using two different ruby rods.

Alvin L. Glick
Raytheon Company
Wayland, Mass.

* Received May 11, 1962; revised manuscript received, May 21, 1962.

[1] S. Koozekanani, P. P. Debye, A. Krutchkoff and M. Cifran, "Measurements of the laser output," Proc. IRE (*Correspondence*), vol. 50, p. 207; February, 1962.

[2] Four "Polaroid" filters in line, No. KN-36.

[3] The sensitivity range for the 917 is 0.90 to 3.0 μa/w at 7000 A°, a typical value being 1.5 μa/w. Correspondence from R. C. Park, RCA Electron Tube Div.

For absolute calibration of phototube see R. W. Engstrom, "Absolute spectral response characteristics of photosensitive devices," *RCA Rev.*, vol. 21, p. 184; June, 1960.

On a Random Failure Mechanism*

When Lusser[1] was Reliability Coordinator at Redstone Arsenal, he suggested that random failures of components were caused by random stresses. In the course of an unsuccessful search (unsuccessful in that the hypothesis of random stresses was found insufficient and unnecessary) for a mathematical connection between random failures and random stresses, the following results were found which may be of interest.

Assume that a stress X varies randomly with time with unvarying statistics. Then the fraction of time the stress spends in a stress interval dx is

$$f_X(x)dx = \frac{dt}{T}. \tag{1}$$

where T is an observation interval and $f_X(x)$ is the probability frequency function of the stress distribution.

Assume a simplified linear stress-failure mechanism wherein a parameter p degrades as a function of only the applied stress,

$$\frac{dp}{dt} = h(X).$$

In a time T the degradation Δp is

$$\Delta p = \int_0^T h(X)dt. \tag{2}$$

Substitution of (1) into (2) yields

$$\Delta p = T\int_{-\infty}^{\infty} h(x)f_X(x)dx = kT \tag{3}$$

where k is a constant defined by the integral. This result is essentially independent of the random nature of the stress contrary to Lusser's suggestion. If X is a constant stress of value S,

$$f_X(x) = \delta(x - S) \tag{4}$$

where δ is the Dirac delta function. Substitution of (4) into (3) yields

$$\Delta p = Th(S) \tag{5}$$

which describes the classical stress-failure curves which have been published, for example, for circuit breakers. Eq. (5) enables one to determine $h(X)$ experimentally.

Returning to (3), assume that a component fails when $\Delta p > \phi$ and that k and ϕ are distributed in a component population with frequency functions $f_k(x)$ and $f_\phi(x)$. We have by a change of variable,

$$P(x < kT \leq x + dx) = \frac{1}{T} f_k\left(\frac{x}{T}\right)dx.$$

If k and ϕ are independent random variables, their joint distribution is given by

$$P(x < kT \leq x + dx,\ y < \phi \leq y + dy) = \frac{1}{T} f_k\left(\frac{x}{T}\right) f_\phi(y)dxdy$$

Hence,

$$P(\Delta p < \phi) = \frac{1}{T}\int_{x=y}^{x=\infty}\int_{y=-\infty}^{\infty} f_k\left(\frac{x}{T}\right) f_\phi(y)dxdy.$$

Let $z = x - y$. Then,

$$P(\Delta p > \phi) = \frac{1}{T}\int_0^T\int_{-\infty}^{\infty} f_k\left(\frac{x}{T}\right) f_\phi(x - z)dxdz. \tag{6}$$

P in (6) can be interpreted as the fraction of components in the population that have failed at time T. Its derivative is the failure rate at that time.

Let us evaluate (6) for the case when T is small or when k has a unique value $\bar{k}$.

Then,

$$\frac{1}{T} f_k\left(\frac{x}{T}\right) = \delta(x - \bar{k}T). \tag{7}$$

Substitution of (7) into (6) yields, after integration,

$$P(\Delta p > \phi) = \int_{-\infty}^{\bar{k}T} f_\phi(u)du.$$

Differentiation yields the failure rate r at time T,

$$r = \frac{dP}{dT} = f_\phi(\bar{k}T). \tag{8}$$

The shape of a typical stress-failure distribution for components produced without any kind of testing or screening is shown in Fig. 1. Such distributions have been seen, for example, in the case of germanium diodes where the stress was mechanical acceleration. The "tail" extending down to very low values of stress seems to be characteristic of such distributions.

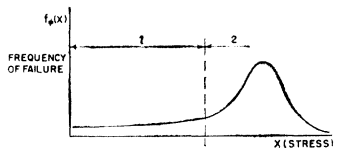


Fig. 1—Typical stress-faliure distribution.

By (8), the failure rate as a function of time may be obtained by replacing the abscissa x in Fig. 1 by $\bar{k}T$. Such a time history of failure rates has the characteristics of the classical curve wherein there is a period 1 in which the failure rate is approximately constant ("random" failures) followed by a rising failure rate period 2 ascribed to wear-out failures.

This suggests the following procedure for reducing (if not eliminating) "random" failures:

1) Determine the stresses that cause component failures.
2) Determine the mechanism by means of which the stress causes the component failures.
3) Modify the component manufacturing process and/or the environmental conditions in which the component is used so that the desired reliability can be achieved.
4) Test (100 per cent) to weed out all components having excessive degradation rates k and inadequate stress margins ϕ.

Essentially, such procedures seem to have been developed by various people (e.g., Remington Rand[2,3]) resulting in very low random failure rates. These procedures appear to have been developed on the basis of experience rather than from any mathematical insight. Incidentally, an examination of Reid and Raymond[3] suggests that the electronic component failures in the Athena computer are a highly probable result of maintenance procedures rather than a fault of the components.

HARRY J. GRAY
The Moore School of Electrical Engineering
University of Pennsylvania
Philadelphia, Pa.

* Received by the IRE, December 20, 1961.

[1] Robert M. Lusser (22 pamphlets on the reliability of missiles), Army Rocket and Guided Missile Agency, Redstone Arsenal, Ala.

[2] N. S. Prywes, H. Lukoff, and J. Schwarz, "UNIVAC-LARC high-speed circuitry: case history in circuit optimization," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 428–429, 437–438; September, 1961.

[3] L. W. Reid and G. A. Raymond, "The Athena computer, a reliability report," *Proc. Eastern Joint Computer Conf.*, Philadelphia, Pa. AIEE Publication T114, pp. 20–24; December 3–5, 1958; July, 1959.

WWV and WWVH Standard Frequency and Time Transmissions*

The frequencies of the National Bureau of Standards radio stations WWV and WWVH are kept in agreement with respect to each other and have been maintained as constant as possible since December 1, 1957, with respect to an improved United States Frequency Standard (USFS).[1] The corrections reported here were arrived at by means of improved measurement methods based on transmissions from the NBS stations WWVB (60 kc) and WWVL (20 kc). The values given in the table are 5-day running averages of the daily 24-hour values for the period beginning at 1800 UT of each day listed.

The time signals of WWV and WWVH are also kept in agreement with each other. Since these signals are locked to the frequency of the transmissions, a continuous departure from UT2 may occur. Corrections are determined and published by the U. S. Naval Observatory. The time signals are maintained in close agreement with UT2 by properly offsetting the broadcast frequency from the USFS at the beginning of each year when necessary. This new system was commenced on January 1, 1960.

Subsequent changes were as follows:

FREQUENCY OFFSET, WITH REFERENCE TO THE USFS

January 1, 1960, —150 parts in 10^{10}
January 1, 1962, —130 parts in 10^{10}

TIME ADJUSTMENTS, WITH REFERENCE TO THE TIME SCALE UT2

December 16, 1959, retardation, 20 milliseconds
January 1, 1961, retardation, 5 milliseconds
August 1, 1961, advancement, 50 milliseconds

* Received June 21, 1962.

[1] Refer to "National Standards of Time and Frequency in the United States," PROC. IRE, vol. 48, pp. 105–106; January, 1960.

Adjustments were made at 0000 UT on the foregoing dates; an advancement means that the signals were adjusted to occur at an earlier time than before.

WWV Frequency
With Respect to U. S. Frequency Standard

1962	Parts in 10^{10}†
May 1	−130.1
2	−130.0
3	−129.9
4	−129.8
5	−129.7
6	−129.6
7	−129.5
8	−129.4
9‡	−129.4
10	−129.9
11	−129.8
12	−129.7
13	−129.7
14	−129.7
15	−129.6
16	−129.5
17‡	−129.5
18	−130.3
19	−130.2
20	−130.3
21	−130.4
22	−130.4
23	−130.3
24	−130.4
25	−130.4
26	−130.4
27	−130.4
28	−130.4
29	−130.4
30	−130.3
31	−130.3

† A minus sign indicates that the broadcast frequency was below nominal. The uncertainty associated with these values is $\pm 5\times10^{-11}$.
‡ WWV frequency adjusted or interrupted as follows:
May 9, -0.5×10^{-10} adjustment, 1900 UT
May 17, -0.7×10^{-10} adjustment, 1900 UT

National Bureau of Standards
Boulder, Colo.

Electromagnetic Scattering from Radially Inhomogeneous Spheres*

Although the analysis of electromagnetic wave propagation through planar media having a varying index of refraction along the direction of propagation has received considerable attention in both the classic and more recent literature, comparatively little work has been directed at the similar problem of propagation through media which are spherically symmetric and have a varying refractive index in the radial direction. The chief work on this latter problem is that of Tai,[1] who derived the equations in the radial variable for the TE (m-type) and TM (e-type) modes. Using these equations as a basis, a method of obtaining the scattering from radially inhomogeneous spheres has been formulated for high-speed computer calculation.

* Received by the IRE, December 11, 1961. The research reported here was sponsored in part by Aeronautical Systems Div., AF Systems Command, U. S. Air Force, Wright-Patterson AFB, Dayton, Ohio, under Contract No. AF 33 (616)-8039.

[1] C. T. Tai, "The electromagnetic theory of the spherical Luneberg lens," *Appl. Sci. Res.*, sec. B, vol. 7, pp. 113–130; 1959.

The far electric fields scattered from the sphere of Fig. 1 are given by

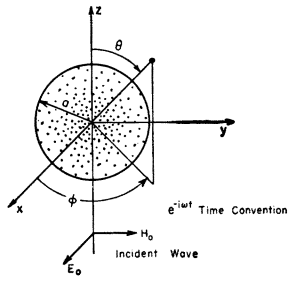


Fig. 1.

$$E_\theta^s = \frac{-iE_0e^{ikr}}{kr}\cos\phi\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\cdot\left[a_n^s\frac{Pn^1(\cos\theta)}{\sin\theta}+b_n^s\frac{\partial Pn^1(\cos\theta)}{\partial\theta}\right], \quad (1)$$

$$E_\phi^s = \frac{+iE_0e^{ikr}}{kr}\sin\phi\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\cdot\left[a_n^s\frac{\partial Pn^1(\cos\theta)}{\partial\theta}+b_n^s\frac{Pn^1(\cos\theta)}{\sin\theta}\right], \quad (2)$$

where a_n^s and b_n^s are, respectively, the magnetic and electric scattering coefficients of the sphere under investigation and $k=\omega\sqrt{\mu_0\epsilon_0}$ is the wavenumber in the ambient medium. The scattering coefficients are calculated from

$$a_n^s = -\frac{kaj_n(ka)-z_n^{(m)}(ka)[kaj_n(ka)]'}{kah_n^{(1)}(ka)-z_n^{(m)}(ka)[kah_n^{(1)}(ka)]'}, \quad (3)$$

$$b_n^s = -\frac{kaj_n(ka)-y_n^{(e)}(ka)[kaj_n(ka)]'}{kah_n^{(1)}(ka)-y_n^{(e)}(ka)[kah_n^{(1)}(ka)]'}, \quad (4)$$

where the symbol (′) means d/dkr and the quantities $z_n^{(m)}(kr)$ and $y_n^{(e)}(kr)$ satisfy the following Riccati equations:

$$\frac{dz_n^{(m)}}{dkr} = \tau(r)+\left[\kappa(r)-\frac{n(n+1)}{(kr)^2\tau(r)}\right][z_n^{(m)}]^2 \quad (5)$$

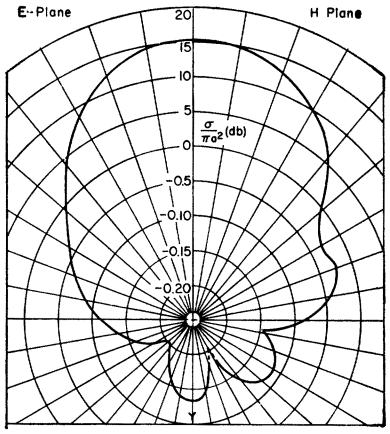
$$\frac{dy_n^{(e)}}{dkr} = \kappa(r)+\left[\tau(r)-\frac{n(n+1)}{(kr)^2\kappa(r)}\right][y_n^{(e)}]^2, \quad (6)$$

when $\mu(r)=\tau(r)\mu_0$ and $\epsilon(r)=\kappa(r)\epsilon_0$ are the constitutive parameters of the sphere. The quantity $-iz_n^{(m)}$ is defined to be the TE modal surface impedance at radius r, normalized to the intrinsic impedance of the ambient medium, and the quantity $-iy_n^{(e)}$ is defined to be the TM modal surface admittance at radius r, normalized to the intrinsic admittance of the ambient medium. An IBM 704 computer has been programmed to numerically integrate (5) and (6) under appropriate end conditions for arbitrary complex functions $\kappa(r)$ and $\tau(r)$. As examples, normalized E- and H-plane scattering patterns, given by

$$\sigma_E = \lim_{r\to\infty} 4\pi r^2\frac{|E_\theta^s|_{\phi=0^\circ}^2}{|E_0|^2}, \quad (7)$$

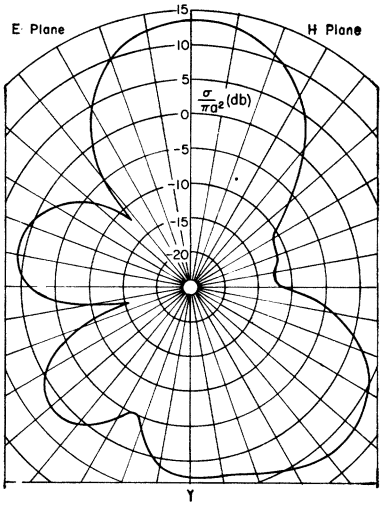
$$\sigma_H = \lim_{r\to\infty} 4\pi r^2\frac{|E_\phi^s|_{\phi=90^\circ}^2}{|E_0|^2}, \quad (8)$$

have been calculated using this program with two real permittivity distributions, known in the literature as the Luneberg lens and the Eaton-Lippmann lens. These distributions and corresponding patterns appear in Figs. 2 and 3. Further details are presented in an Antenna Laboratory memorandum.[2]



Bistatic Scattering
Cross-Sections Of
A Luneberg Lens
ka = 5
$\kappa = 2-\left(\frac{kr}{ka}\right)^2$
$\tau = 1$

Fig. 2.



Bistatic Scattering
Cross-Sections Of An
Eaton-Lippmann Lens
$\kappa = \begin{cases} \frac{2-kr/5}{kr/5} & 0.5 < kr \le 5 \\ \infty & 0 \le kr \le 0.5 \end{cases}$
$\tau = 1$

Fig. 3.

The work of D. Call and L. Du in programming the IBM 704 computer is acknowledged with gratitude.

R. J. Garbacz
Antenna Lab.
Dept. of Electrical Engineering
Ohio State University
Columbus, Ohio

[2] R. J. Garbacz, "Electromagnetic Scattering by Radially Inhomogeneous Spheres," Antenna Lab., Ohio State University, Columbus, Rept. No. 1223-3 (in process).

Space-Charge-Layer Width and Capacitance of Symmetrical Step Junctions*

The behavior of semiconductor p-n junction under large forward bias has received much attention in the recent years. Due to the effect of conductivity modulation and/or generation-recombination current in the space-charge region, it was found that the current varies as $\exp(qV/2kT)$ in the high forward-bias region, where V is the applied voltage less the ohmic drop in the neutral regions.[1] However, an inductive effect, which was observed at some high voltages, is not as yet satisfactorily explained. In junction capacitance calculation, it is usually assumed that the difference in quasi-Fermi levels across the junction is constant and is equal to the applied voltage, leading to a capacitance which increases with applied voltage, and is in contrast to the observed finite maximum capacitance at some high forward bias. It is the purpose of this correspondence to show that, for step junctions, the inductance effect is associated with drop of quasi-Fermi levels in the neutral regions in which the corresponding carriers are majority carriers.

For step junctions at thermal equilibrium Adirovich, *et al.*,[2] showed that the Poisson equation can always be put into a closed integral form. If a transition-layer width is defined as the distance between points, between which the electrostatic potential difference equals 95 per cent of the built-in voltage between p- and n-regions, it was found that the transition-layer width thus obtained agrees with the result from space-charge approximation. The agreement is within 2 per cent for bulk impurity densities ranging from 10^{13} to 10^{19}. Fig. 1 shows the normalized potential distribution across a symmetrical step junction with various impurity doping on both sides of a symmetrical junction.

To calculate the capacitance of the junction under bias, one has to solve the normalized Poisson equation

$$\frac{d^2u}{dy^2} = e^{(u_p-u_n)/2}\sinh\left(u-\frac{u_p+u_n}{2}\right) - \frac{a}{2}, \qquad y>0, \tag{1}$$

where $y=x/L_D$ is the distance in units of Debye length, u, u_p, and u_n are, respectively, the electrostatic potential, quasi-Fermi levels for holes and electrons, multiplied by q/kT, and a is the doping density divided by the intrinsic carrier density. Eq. (1) can be solved only if the variations in u_p and u_n are known. We shall therefore try to find an approximate solution by defining a transition region, bounded by two planes at

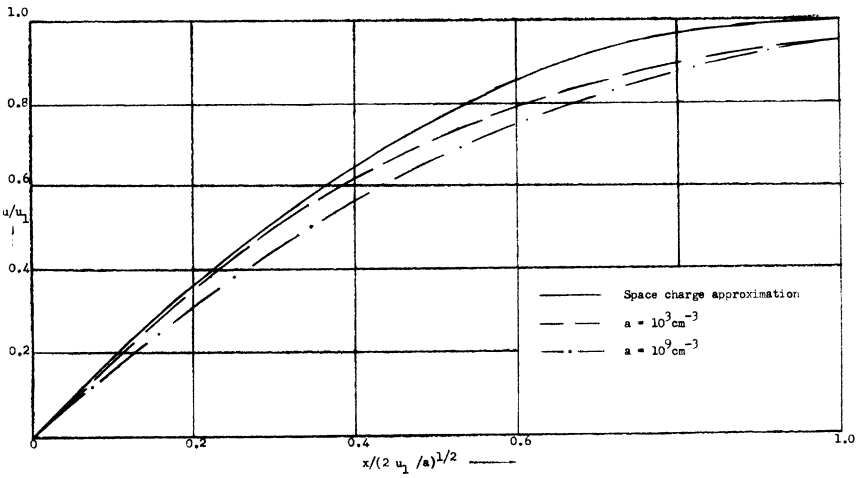


Fig. 1—Potential distribution in one side of the space-charge region. u_1 is the electrostatic potential in the neutral region.

$y=\pm w$, by the following conditions:

$$u''(y) = u'(y) = 0, \quad y > w; \qquad u(0) = 0. \tag{2}$$

We shall assume that the quasi-Fermi levels remain approximately constant throughout the transition region and, therefore, that u' and u'' are identical with u_0' and u_0'' in the transition region, where $u_0=u-(u_p+u_n)/2$. This approximation is evidently valid under a forward-bias condition and a moderate reverse-bias condition.[1] Inasmuch as the ohmic drops are negligible, u_p-u_n across the junction is equal to the applied voltage. As the forward bias is increased, however, the ohmic drops are no larger negligible; as a result, only a fraction of the applied voltage appears across the junction and u_p-u_n becomes less than v, where $v=qV/kT$. Assuming predominance of minority-diffusion current, these drops can be approximately expressed as $\delta u_p=\alpha e^{v/2}$ with similar expression for δu_n. Here α is a parameter characteristic of a particular junction device and is a measure of the effective resistivity of, and effective diffusion length of minority carriers in, the respective neutral regions. An approximate form of (u_p-u_n) across the junction at high forward bias is, therefore,

$$u_p - u_n = 2\alpha e^{v/2}. \tag{3}$$

It is seen that α is usually small, of the order of e^{-v_D}, there v_D is the built-in voltage. Clearly (3) is not valid in the extreme case in which v is much higher than the built-in voltage so that the electrostatic potential across the junction diminishes to zero and the current varies linearly with the applied voltage. With the above, (1) can be integrated once to read

$$\frac{du_0}{dy} = \sqrt{2}\,e^{v/2-\alpha e^{v/2}}\left[u_1\sinh u_1 - \cosh u_1 + \cosh u - u\sinh u_1\right]^{1/2}, \tag{4}$$

where u_1 is the value of u_0 evaluated at edge of the transition layer.

Introducing a normalizing capacitance $C_0=q^2n_iL_D/kT(=256\ \mu\mu f/cm^2$ for Si), the normalized capacitance

$$c = \frac{d}{dv}\int_{-l}^{+l} p\,dy,$$

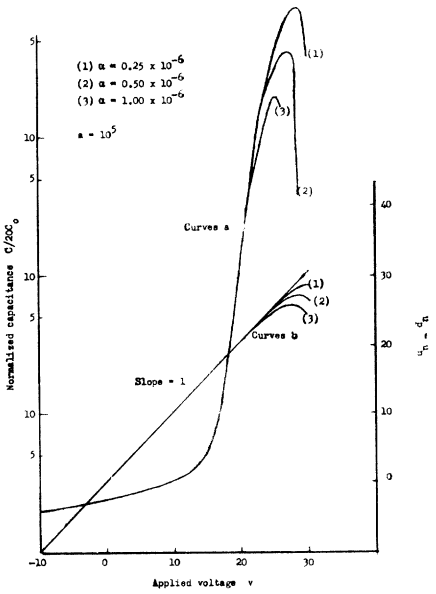


Fig. 2—(a) Capacitance of a symmetrical step junction. (b) Drop of quasi-Fermi levels across a symmetrical step junction.

can be written as the sum of three capacitances: 1) diffusion capacitance c_d; 2) space-charge capacitance c_t; 3) transition-region-carrier capacitance c_s, where

$$c_d = \int_{-l}^{-w}\frac{dn}{dv}\,dy + \int_{w}^{l}\frac{dp}{dv}\,dy,$$

$$c_s = -2\frac{du_0'(0)}{dv}, \qquad c_t = 2\int_0^w \frac{dp}{dv}\,dy.$$

The diffusion capacitance can be obtained by assuming that minority-carrier densities decay exponentially in the neutral regions with appropriate diffusion lengths, whereas the other two can be put into closed integral form using (4). A numerical evaluation of the capacitances was obtained and the result is shown in Fig. 2, together with the variations of u_p-u_n across the junction. Observe that the capacitance increases to a maximum at some voltage beyond v_D; at still higher voltage, the capacitance decreases and would tend to zero in the limit as the junction de-

* Received December 22, 1961; revised manuscript received January 5, 1962.

[1] C. T. Sah, R. N. Noyce, and W. Shockley, "Carrier generation and recombination in p-n junctions and p-n junction characteristics," PROC. IRE, vol. 45, pp. 1228–1243; September, 1957.

[2] E. I. Adirovich, Iu. S. Riabinkin, and K. V. Temko, "Equilibrium distribution of potential, field and concentration of current carriers in fused junctions," *Soviet J. Tech. Phys.*, vol. 3, pp. 49–59; January, 1958. (Translation.)

vice under consideration becomes a resistive element. Note also that the deviation of u_p-u_n from v is small even at very high bias, whereas the resulting decrease in capacitance is large. This perhaps explains why the usual assumption of $u_p-u_n=v$ across the junction has not been able to detect the obvious inductance effect in p-n junctions. Since the impedance associated with majority-current flow increases with frequency and bulk resistivity, it appears that the voltage at which the inductive effect occurs is lowered with increase in either bulk resistivity or the frequency of the applied voltage, in general agreement with the result of Sah for linearly graded junctions.[3]

The author wishes to acknowledge the help of H. Golde in programming for numerical solution.

C. C. WANG
Dept. of Electrical Engineering
University of Washington
Seattle, Wash.

[3] C. T. Sah, "Effects of electrons and holes on the transition layer characteristics of linearly graded p-n junctions," PROC. IRE, vol. 49, pp. 603–618; March, 1961.

Power Carried by the Helitron Waves on an E-Type Filamentary Electron Beam*

Waves on an E-type filamentary electron beam in a transverse-field circuit have been pointed out by Pantell,[1] who found the existence of the fast and slow space charge waves, but did not analyze the kinetic power flow of each wave. It is the purpose of this paper to show that the kinetic power flow on the E-type filamentary electron beam can be expressed in a normal mode form as well. These waves are named "fast helitron waves" and "slow helitron waves," respectively. It is found that similar expressions for the power of each wave component to those of O- and M-type electron beams can be obtained by this analysis.

In a drift space of the E-type beam as shown in Fig. 1, the following equations are obtained from the force equation for electrons which are traveling to the z-direction with dc velocity u_0, under the conditions of electrostatic focusing. It is assumed that the electron traces the path of a helix and all electrons form a right circular hollow cylinder which is thin in the radial dimension.

$$\frac{d^2x}{dt^2}=\eta V_0\frac{x}{x^2+y^2}$$
$$\frac{d^2y}{dt^2}=\eta V_0\frac{y}{x^2+y^2}$$
$$\frac{dz}{dt}=\sqrt{\eta V_0}\tan\phi \qquad (1)$$

where the definitions of quantities are given as follows:

$\eta=1.759\times10^{11}$ c/kg,
$V_0=(V_c-V_s)/\ln(s/c)$, $V_c>V_s$,
V_s = voltage between outer cylinder and ground,
V_c = voltage between inner cylinder and ground,
ϕ = electron pitch angle.

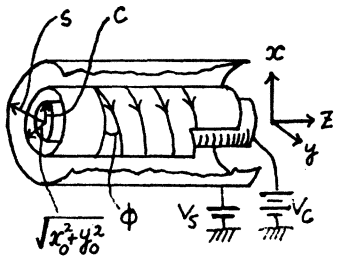


Fig. 1—Electrons form a right circular hollow beam that is thin in the radical direction and the electron traces the path of a helix between coaxial cylinders.

Under a small signal theory, the motion of the electron is given as a sum of a dc quantity and an RF disturbance.

$$x=x_0+x_1,$$
$$y=y_0+y_1,$$
$$z=z_0. \qquad (2)$$

The subscript 0 refers to the dc quantity and the subscript 1 refers to the RF motion.

Using (2), (1) becomes

$$\frac{dv_{1x}}{dt}+2\omega_h^2x_1=0,$$
$$\frac{dv_{1y}}{dt}+2\omega_h^2y_1=0,$$
$$\frac{dx_1}{dt}=v_{1x},$$
$$\frac{dy_1}{dt}=v_{1y} \qquad (3)$$

where ω_h is named "helitron angular frequency,"

$$\omega_h=\sqrt{\eta V_0/(x_0^2+y_0^2)}.$$

Now, we introduce the effective transverse current[2] K, as well as Chu's kinetic voltage U.

$$2U=U_x+jU_y,$$
$$2K=K_x+jK_y,$$
$$U_x=-u_0v_{1x}/\eta,$$
$$U_y=-u_0v_{1y}/\eta,$$
$$K_x=j\omega\sigma_0x_1,$$
$$K_y=j\omega\sigma_0y_1, \qquad (4)$$

where σ_0 is the total charge of the electron beam per transverse unit area. The subscript x denotes an x component and the subscript y denotes a y component. From (3) and (4), the following space charge wave equations are obtained:

$$\left(\frac{d}{dz}+j\beta_e\right)U=-j\frac{2\omega_h^2}{\omega q\sigma_0}K,$$
$$\left(\frac{d}{dz}+j\beta_e\right)K=-j\frac{\omega q\sigma_0}{u_0^2}U. \qquad (5)$$

These expressions are very similar to those for the longitudinal beam[3] (one dimensional).

Next, we define the helitron normal mode $a_\pm$ using U and K:

$$a_\pm=\frac{1}{4\sqrt{Z_0}}(U\pm Z_0K) \qquad (6)$$

where z_0 is the characteristic beam impedance for the small-signal quantities of the helitron waves,

$$Z_0=\frac{2V_0}{|I_0|}\frac{\sqrt{2}\beta_h}{\beta_e} \qquad (7)$$

where I_0 is the dc beam current ($|I_0|=-\sigma_0u_0$), $\beta_h=\omega_h/u_0$ and $\beta_e=\omega/u_0$.

Using (6) and (7), (5) then becomes

$$\left(\frac{d}{dz}+j(\beta_e\mp\sqrt{2}\beta_h)\right)a_\pm=0. \qquad (8)$$

These are the normal mode forms of the helitron waves. Accordingly, for an arbitrary excitation in terms of the mode amplitudes at the input plane $z=0$, the solution of (8) are,

$$a_{\pm(z)}=a_{\pm(0)}e^{-j(\beta_e\mp\sqrt{2}\beta_h)z}. \qquad (9)$$

It is easy to see that from (9) the a_+ mode, the fast helitron wave, travels with a phase velocity slightly greater than the dc velocity, whereas the a_- mode, the slow helitron wave, travels with a phase velocity slightly less than u_0. By $\omega=\sqrt{2}\omega_h$, it is possible to have an infinite phase velocity associated with fast helitron waves.

From (9), K and u may be expressed as follows:

$$U=2\ \mathrm{Re}\{\sqrt{Z_0}[a_{+(z)}+a_{-(z)}]e^{j\omega t}\},$$
$$K=2\ \mathrm{Re}\left\{\frac{1}{\sqrt{Z_0}}[a_{+(z)}-a_{-(z)}]e^{j\omega t}\right\} \qquad (10)$$

The power carried by helitron waves can be given by the generalized Chu's power theorem as follows:

$$P=\tfrac{1}{2}\ \mathrm{Re}\,(UK^*)=2(|a_+|^2-|a_-|^2). \qquad (11)$$

In order to show that P is independent z, we differentiate (10) with z and (8) and their complex conjugate substitute into this expression. It then follows immediately that $dP/dz=0$.

In conclusion, the power flow on the E-type filamentary electron beam in the drift space discussed here is given as (10) and it is seen that the fast helitron wave carries positive power, whereas the slow helitron wave carries negative power.

ICHIRO SAKURABA
Faculty of Engineering
Hokkaido University
Hokkaido, Japan

* Received by the IRE, December 11, 1961.

[1] R. H. Pantell, "Electrostatic electron beam coupler," IRE TRANS. ON ELECTRON DEVICES, vol. ED-8, pp. 39–43: January, 1961.

[2] S. Saito, "Power carried by the cyclotron waves and the synchronous waves on a filamentary electron beam," PROC. IRE, vol. 49, pp. 969–970; May, 1961.

[3] W. H. Louisell, "Coupled Mode and Parametric Electronics," John Wiley and Sons, Inc., New York, N. Y.; 1960.

A Technique for Equalizing Parabolic Group Delay*

It can be shown[1] that the phase shift in a band-pass filter that is a minimum phase-shift network is

$$\theta = -\sum_{b=1}^{n} \tan^{-1}\frac{X - X_{bi}}{X_{br}} \tag{1}$$

when X = normalized frequency variable, X_{br} and X_{bi} are real and imaginary components of the pole positions (*i.e.*, poles of filter transfer function), n = number of resonant circuits.

If group delay is defined as the derivative of the phase shift with respect to frequency

$$d = \frac{1}{\pi\Delta f}\sum_{b=1}^{n}\frac{B_b}{1 + (A_b + B_bX)^2} \tag{2}$$

where

$$A_b = \frac{1}{X_{br}} \quad \text{and} \quad B_b = -\frac{X_{bi}}{X_{br}}$$

and d = absolute group delay of the filter, Δf = 3-db bandwidth of the filter. At the band center

$$(x = 0)$$

then

$$d_0 = \frac{1}{\pi\Delta f}\sum_{b=1}^{n}\frac{B_b}{1 + A_b^2}. \tag{3}$$

Letting Δd = the differential group delay at any normalized frequency X, then

$$\Delta d = d - d_0. \tag{4}$$

Parabolic group delay which occurs in the pass band centers of maximally flat band-pass filters can be equalized using the conventional all-pass network which provides compensating group delay and no amplitude discrimination. All-pass equalizers usually are lattice or bridged-T networks which are often difficult to physically realize. This paper considers a technique for equalization that does not use all-pass networks as equalizers.

For a maximally flat band-pass filter with *two or more* resonant circuits it can be shown that

$$\Delta d_f \cong \frac{CX_f^2}{\pi\Delta f_f} \qquad (0 < X_f < 0.4), \tag{5}$$

where

Δd_f = differential group delay of filter
X_f = normalized frequency variable of filter
f = frequency
f_0 = center (*i.e.*, resonant) frequency of filter
Δf_f = 3-db bandwidth of filter.

For narrow bandwidth filters

$$\left(\frac{\Delta f_f}{f_0} \leqq 2\%\right)$$

$$X_f \cong \frac{2|f - f_0|}{\Delta f_f} \tag{6}$$

* Received December 12, 1961.
[1] M. Dishal, "Dissipative band-pass filters," PROC. IRE, vol. 37, pp. 1058–1060; September, 1949.

then

$$\Delta d_f = \frac{4C|f - f_0|^2}{\pi(\Delta f_f)^3}. \tag{7}$$

This is a parabolic differential group-delay function that is concave upward. For a single-tuned circuit (*i.e.*, maximally flat filter with one resonator)

$$\Delta d_e \cong \frac{-X_e^2}{\pi\Delta f_e} \tag{8}$$

where

Δd_e = differential group delay of equalizer
X_e = normalized frequency variable of equalizer
f = frequency
f_0 = center (*i.e.*, resonant) frequency of equalizer
Δf_e = 3-db bandwidth of equalizer.

For narrow bandwidth equalizers

$$\left(\frac{\Delta f_e}{f_0} \leqq 2\%\right)$$

$$X_e \cong \frac{2|f - f_0|}{\Delta f_e} \tag{9}$$

then

$$\Delta d_e = \frac{-4|f - f_0|^2}{\pi(\Delta f_e)^3}. \tag{10}$$

This is a parabolic differential group-delay function that is concave downward. For equalization

$$\Delta d_f + \Delta d_e = 0 \quad \text{and} \quad \Delta d_f = -\Delta d_e$$

$$\frac{4C|f - f_0|^2}{\pi(\Delta f_f)^3} = \frac{4|f - f_0|^2}{\pi(\Delta f_e)^3} \tag{11}$$

$$\therefore\ (\Delta f_e)^3 = \frac{1}{C}(\Delta f_f)^3$$

$$\Delta f_e = C^{-1/3}\Delta f_f. \tag{12}$$

For maximally flat band-pass filters, (2) becomes

$$d = \frac{1}{\pi\Delta f}\sum_{r=1}^{n}\frac{B_r}{1 + (A_r + B_rX)^2} \tag{13}$$

where

$$A_r = \csc\left(\frac{2r - 1}{2n}\right)\pi$$

and

$$B_r = -\operatorname{ctn}\left(\frac{2r - 1}{2n}\right)\pi$$

$$d_0 = \frac{1}{\pi\Delta f}\sum_{n=1}^{n}\frac{B_r}{1 + A_r^2} = \frac{D_0}{\pi\Delta f}. \tag{14}$$

Using a parabolic approximation for differential group delay

$$\frac{CX^2}{\pi\Delta f} \cong d - d_0. \tag{15}$$

The constant C is selected at a value of $X = 0.2$. A table of values of C and D_0 for n from two to eight resonators follows:

n	C	D_0
2	1.414	1.414
3	1	2.000
4	1.083	2.613
5	1.236	3.236
6	1.414	3.864
7	1.604	4.494
8	1.799	5.126

This technique should be applicable to RF and IF amplifier chains in which single-tuned and double-tuned circuits can be readily used in alternate coupling networks.

R. M. KURZROK
RCA Surface Commun. Sys. Lab.
New York, N. Y.

A Useful Technique for Treatment of Certain Statistical Sampled-Data Control Systems*

In any statistical treatment of control systems the first- or second-order statistic, namely the average value and mean-square value, has special significance. Two theorems appeared in Ku and Wolf,[1] and Wolf,[2] which are given below.

Theorem 1: Given a random process $g(t)$ with known statistical quantities and a deterministic operator L such that if $G(q)$ is the resulting transform-random process according to the relation $G(q) = L\{g(t)\}$ where q is the transform-random variable corresponding to the original random t, then $\langle G(q)\rangle_G = L[\langle g(t)\rangle_g]$ where $\langle\ \rangle_G$ is the ensemble average with respect to G and $\langle\ \rangle_g$ is the ensemble average with respect to g and these are commutative with respect to each other and the deterministic operator L.

Theorem 2: Given a random process $g(t)$ and its transform $G(q)$, the mean nth ensemble average of the given process is $\langle g^n(t)\rangle_g = L^{-1}[\langle C^{n-1}\{G(q)\}\rangle_G]$ where $C^k\{\ \}$ is the convolution transform of the kth order.

Second-order statistic is specifically obtained for a continuous case using Laplace transform.[2] This seems to be quite useful. This technique can now be developed for sampled-data case using z transforms. Closed-form solution can be immediately obtained.

In this short note we will use above theorems to obtain specific results for the sampled-data case. Let the system consist

* Received by the IRE, December 28, 1961; revised manuscript received, January 15, 1962.
[1] Ku, Y. H. and Wolf, A. A., "Transform ensemble method for the analysis of linear and non-linear systems with random inputs," *Proc. Natl. Electronics Conf.*, Chicago, Ill., 1960, Natl. Electronics Conf. Inc., Chicago, Ill., vol. 15, pp. 441–455; 1960.
[2] Wolf, A. A., "On poles and zeros of a random process in linear and nonlinear systems," *Proc. Natl. Electronics Conf.*, Chicago, Ill., Natl. Electronics Conf. Inc., Chicago, Ill., vol. 15; 1960.

of a sampler, hold circuit $H(s)$ and linear element $G(s)$. Let the input be $x(t)$ and the output be $y(t)$ (sampled output $y(nT)$). For this system we can write using z transform $Y(z)=GH(z)X(z)$ which gives

$$y(nT) = z^{-1}\{GH(z)X(z)\}$$

or using Theorem 1, we get

$$\langle y(nT)\rangle = z^{-1}\{GH(z)\cdot\langle X(z)\rangle_X\}.$$

Hence, this gives us a way of getting the mean value of the output at sampling instants, knowing the mean value of the input using z transforms.

In statistical work we are usually concerned with the mean-square values; then let us try to find mean-square value of output first at sampling instants.

Using Theorem 2 we can write

$$\langle y^2(nT)\rangle_y = z^{-1}[\langle C\{Y(z)\}\rangle_y,$$

where $C\{Y(z)\}$ is the convolution of $Y(z)$ *i.e.*,[3]

$$C[Y(z)] = Y(z) * Y(z) = \frac{1}{2\pi j}\int_\Gamma p^{-1}Y(p)\,Y\left(\frac{z}{p}\right)dp.$$

In our case $Y(z)=GH(z)X(z)$

$$\therefore [CY(z)] = \frac{1}{2\pi j}\int_\Gamma p^{-1}GH(p)X(p)GH\left(\frac{z}{p}\right)X\left(\frac{z}{p}\right)dp$$

$$= \frac{1}{2\pi j}\int_\Gamma p^{-1}GH(p)GH\left(\frac{z}{p}\right)X(p)X\left(\frac{z}{p}\right)dp$$

or

$$\langle C[Y(z)]\rangle_Y = \frac{1}{2\pi j}\int_\Gamma p^{-1}GH(p)GH\left(\frac{z}{p}\right)\cdot\left\langle X(p)X\left(\frac{z}{p}\right)\right\rangle_X dp$$

Hence, we can write that

$$\langle y^2(nT)\rangle_y = z^{-1}\left[\frac{1}{2\pi j}\int_\Gamma p^{-1}GH(p)GH\left(\frac{z}{p}\right)\cdot\left\langle X(p)X\left(\frac{z}{p}\right)\right\rangle_X dp\right]. \quad (1)$$

Now we must define what we mean by $\langle X(p)X(z/p)\rangle_X$. Consider $\langle x(t)x(t+\tau)\rangle_x$. It can be easily shown

$$L_{(t)}^{(s)}L_{(\tau)}^{(k)}[x(t)x(t+\tau)] = X(k)X(s-k),$$

where

$L_{(t)}^{(s)}$ is defined as the Laplace transform w.r.t.t with new variable s.

and

$L_{(\tau)}^{(k)}$ is defined as the Laplace transform w.r.t.τ with new variable k.

Now

$$z_{(k)}^{(p)}z_{(s)}^{(z)}[X(k)X(s-k)] = X(p)X\left(\frac{z}{p}\right),$$

where

$z_{(k)}^{(p)}$ is defined as the z transform w.r.t.k with new variable p,

and

$z_{(s)}^{(z)}$ is defined as the z transform w.r.t.s with new variable z.

By Theorem 1 we can write

$$z^{(p)}z^{(z)}[\langle x(t)x(t+\tau)\rangle_x] = \left\langle X(p)X\left(\frac{z}{p}\right)\right\rangle_X.$$

Hence we can find $\langle X(p)X(z/p)\rangle_X$ as follows:

1) Find $\langle x(t)x(t+\tau)\rangle_x$, which by definition is $\phi_{xx}(\tau)$, the autocorrelation function of the input.
2) Then find
$$L_{(t)}^{(s)}L_{(\tau)}^{(k)}\langle x(t)x(t+\tau)\rangle_x,$$
which will be of the form
$$\langle X(k)X(s-k)\rangle_X.$$
3) Then find
$$z_{(k)}^{(p)}z_{(s)}^{(z)}\langle X(k)X(s-k)\rangle_X,$$
which will give
$$\left\langle X(p)X\left(\frac{z}{p}\right)\right\rangle_X.$$

Substituting in integral (1), we have the complex integral which can be easily evaluated by complex integration,[3] giving us the mean-square value at sampling instants in a closed form. Closed-loop system causes no more difficulty.

If we are interested in obtaining the mean-square value in general and not only at sampling instants, we consider the modified z-transform approach.

For the system considered above, we can write $Y(z, m)=GH(z, m)X(z)$ which can be written by Theorem 1 as

$$\langle y(nT, m)\rangle_y = Z_m^{-1}[GH(z, m)\langle X(z)\rangle_X]$$

where

$$0 < m < 1.$$

Now

$$\overline{y(t)} = \int_0^1 \langle y(nt, m)\rangle_y dm.$$

In order to get the mean-square value at the output we again apply Theorem 2. Applying the same reasoning as above, we can write $\langle y^2(nT, m)\rangle_y$

$$= Z_m^{-1}\left[\frac{1}{2\pi j}\int_\Gamma p^{-1}GH(p, m)GH\left(\frac{z}{p}, m\right)\cdot\left\langle X(p)X\left(\frac{z}{p}\right)\right\rangle_X dp\right] \quad (2)$$

where $\langle X(p)X(z/p)\rangle_X$ has already been defined above. The integral can be evaluated by usual techniques.[3] Now

$$\overline{y^2(t)} = \tau\int_0^1 \langle y^2(nT, m)\rangle dm.$$

Then (1) and (2) give us the closed-form mean-square values at the output for sampling instants and in general, respectively, when the input-correlation function is known. Other methods to get these results are known and are equivalent to this,[3] but it seems that for the study of sampled-data systems this point of view proves to be quite handy. The major advantage of this technique is that we can apply Theorem 2 continuously to get higher moments without any greater difficulty and this helps us to construct the probability distribution at the output because this probability density function can be expressed in terms of moments of the output.

S. C. Gupta
Dept. of Electrical Engineering
University of California
Berkeley, Calif.

[3] E. I. Jury, Class Notes in "Applications of z-Transform Theory," University of California, Berkeley, Fall, 1961 (to be published).

Duration and Spacing of Sferic Pulses*

Distributions of durations and the spacing between sferic pulses were obtained from a limited sample of sferics recorded simultaneously at two stations. About 9000 sferic waveforms from the 35-mm film records were analyzed to provide preliminary information for a receiver development project. The duration of a sferic was arbitrarily defined as the length of time (after the sweep was triggered) that the amplitude exceeded 10 mv/m. A total oscilloscope sweep of 500 μsec was used for these recordings, and when the amplitude still exceeded 10 mv/m at the end of the trace, the duration was tabulated as >500 μsec. The frequency response of the recording equipment was approximately 1 to 150 kc. The triggering circuits had a 60 per cent bandwidth centered at 10 kc and were calibrated to the peak-to-peak amplitude of an equivalent CW signal which would just produce a trigger. The triggering sensitivity was set at various levels during the recording periods.

Fig. 1 shows the duration distributions for a period on May 27, 1960 when a large thunderstorm was located from 10 to 30 miles from the omnidirectional station at Goodland, Kansas.[1] The Brighton, Colorado station (near Denver) employed an Ephi directional system[2] which permitted the selection of sferics from pre-selected azimuths, in this case a 6° sector centered over Goodland (102° azimuth). During the recording interval the storm near Goodland was also the storm nearest Brighton (distance 160–180 miles), although there were clouds beyond Goodland which could also have produced some sferics. The Goodland station was operated at a high triggering level (324 mv/m) to avoid saturation or overlapping of waveforms on the film resulting from the sferics of the nearby storm.

Observations at Leoti, Kansas and Brighton, Colorado on May 19, 1960 indi-

* Received December 29, 1961; revised manuscript received, January 29, 1962.

[1] Time references are in Mountain Standard Time (105th meridian).

[2] G. Heney, R. F. Linfield and T. L. Davis, "The Ephi system for VLF direction finding," *J. Res. NBS*, vol. 65C, pp. 42–49; January–March, 1961.

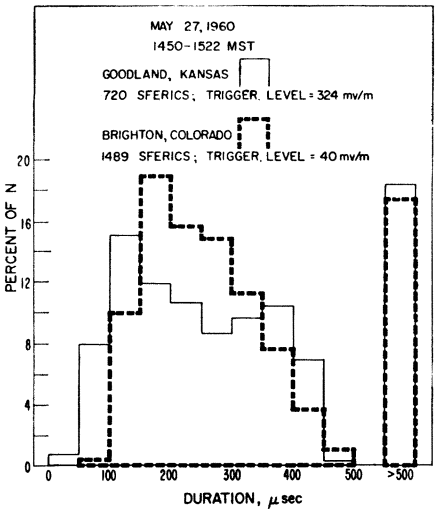


Fig. 1—Comparative duration distributions (N = total sferics).

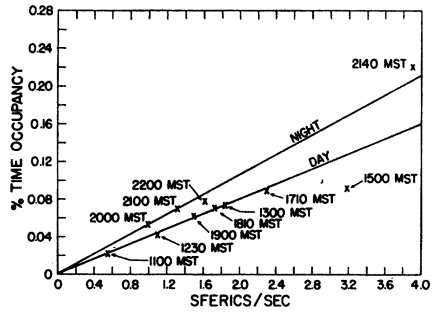


Fig. 2—Percentage of sample interval occupied by sferics—Leoti, Kansas, May 19, 1960.

cated a considerable shift in the duration peak as weather and propagation factors varied during the course of the day. Sferic durations were longer at night or when the source was at a great distance, and there were indications of a shift toward shorter durations coincident with tornado occurrences.[3]

The percentage of the sample interval occupied by sferics was calculated for a series of observations made at Leoti with the same omnidirectional equipment later used at Goodland. Sferics tabulated as ">500 μsec" were arbitrarily assigned a duration of 750 μsec for this part of the study. The total time occupancy is shown as a function of sferic rate in Fig. 2 for both day and night conditions. An extension of the two curves to higher sferic rates would indicate sferic time occupancy of about 4 to 5 per cent at 100 sferics/sec, and 12 to 15 per cent at a rate of 300 sferics/sec. However, this assumption of a linear relationship between rate and time occupancy may not be justified, and direct observation at the higher rates is needed. A change in the gating sensitivity (triggering level) will of course change the relative time occupancy.

Observations of the rate-vs-triggering level relationship were made at Brighton

[3] C. A. Samson and R. F. Linfield, "Sferic observations of the severe weather on May 19, 1961," *J. Geophys. Res.*, February, 1962.

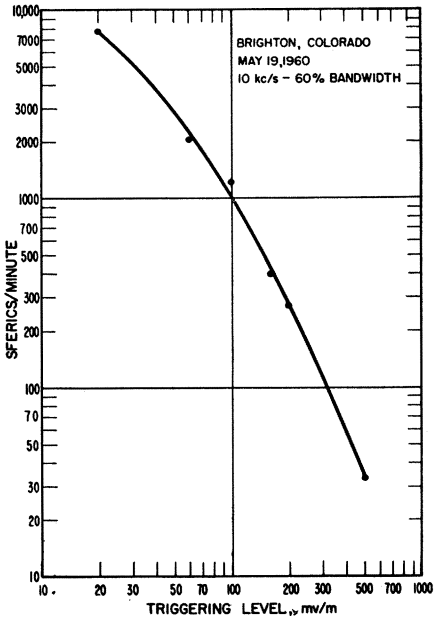


Fig. 3—Rate-vs-triggering level relationship.

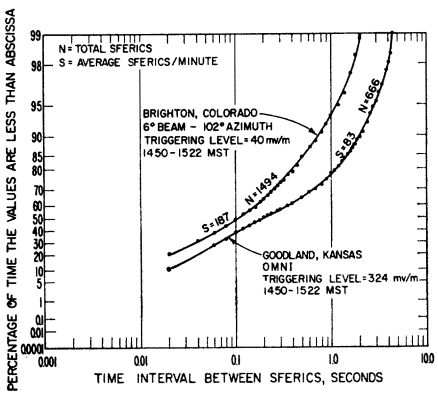


Fig. 4—Distribution of intervals between sferics, May 27, 1960.

over a period of about 9 hours on May 19, 1960, using an omnidirectional antenna. The average curve for this relationship is shown in Fig. 3. There was widespread storm activity in the midwest on this date, but no storms within several hundred miles of Brighton. Observations of this type indicate that some limiting value of sferic rate is approached as the sensitivity of the receiver is increased. The highest rate actually observed on May 19 was at about 2200 MST, when 12,880 sferics were recorded in one minute at a triggering level of 20 mv/m. The average rate during this interval was 215 sferics/sec, at a time when the major storm activity was 500–700 miles distant.

Fig. 4 shows the distribution of the time interval between sferics for the data used in Fig. 1. Other data indicate that the interval distribution varies as the sferic rate changes, regardless of variations in storm location.

R. F. Linfield
DECO Electronics, Inc.
Boulder, Colo.
C. A. Samson
Central Radio Propagation Lab.
National Bureau of Standards
Boulder, Colo.

RMS Currents in Variable-Width Pulse (VWP) Power Supply Circuits*

The use of "chopping techniques" employing variable-width pulse (VWP) methods for controlling and regulating dc power is becoming increasingly common in electronic applications. An important parameter in a VWP circuit is the rms current flowing. Herein is described a simple method for determining rms currents by means of measuring certain dc currents only.

In Fig. 1 is shown a common embodiment of the basic VWP circuit, employing a variable-width pulse for output-voltage control and regulation. The dc input E_1 is greater than the required dc output E_2. Power transistor Q_1 is turned on and off at a fixed frequency corresponding to a period T. The time t that the transistor is turned on is varied to give pulses of variable widths. Similar action may be obtained with SCR's, but turn-on turn-off circuits are more complicated than with transistors.

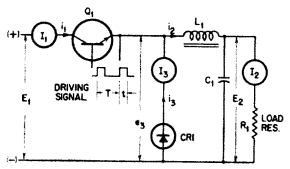


Fig. 1—Basic VWP circuit.

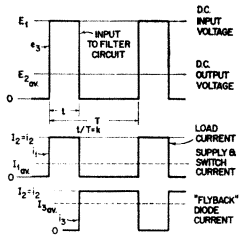


Fig. 2—Voltage and current waveforms of VWP circuit.

Fig. 2 shows major voltage waveforms and the corresponding current waveforms. It is assumed that L_1 of Fig. 1 is large, the load current ripple is small, and essentially constant current flows through L_1. Accordingly,

$$i_2 = I_2 = I_{2\,\mathrm{av}} = I_{2\,\mathrm{peak}}. \quad (1)$$

Under these conditions, the average dc value of the voltage wave e_3 appears across the output capacitor C_1. Therefore, the output voltage is equal to the average of the variable-width rectangular wave shown in Fig. 2. Hence,

$$E_2 = E_{2\,\mathrm{av}} = kE_1, \quad (2)$$

where

$$k = t/T. \quad (3)$$

By varying k, the output voltage can be controlled and regulated.

* Received January 8, 1962; revised manuscript received March 2, 1962.

Normally, instrumentation is not readily available for accurate rms measurements over wide ranges of pulsed currents, operating frequencies, etc; further, when the instruments are available, they are usually of the thermocouple type, and sensitive to damage if overloaded.

For rectangular pulses of current, as shown in Fig. 2, rms measurements need NOT be made; dc average current measurements will suffice. The rms current in Q_1 is given by

$$I_1 = \sqrt{I_1\,\text{av} \times I_2\,\text{av}}, \tag{4}$$

i.e., the value of the duty cycle also need NOT be measured.

This is proved as follows:

From Fig. 2 it is apparent that

$$I_1\,\text{av} = kI_2. \tag{5}$$

However, the rms value of a rectangular wave is given by

$$\text{Irms} = I = \sqrt{k}\, I\,\text{peak}. \tag{6}$$

But from (5) it is determined that

$$k = \frac{I_1\,\text{av}}{I_2}. \tag{7}$$

Using (1), and substituting (7) into (6) gives:

$$I_1 = \sqrt{\frac{I_1\,\text{av}}{I_2\,\text{av}}} \times I_2\,\text{av}$$

$$= \sqrt{I_1\,\text{av} \times I_2\,\text{av}}, \text{ Q.E.D.} \tag{8}$$

The average value of the current in CR1 is given by

$$I_3\,\text{av} = I_2(1 - k). \tag{9}$$

This *average value* is determined from Kirchhoff's Law, that the algebraic sum of all currents at a point must be equal to 0.

The rms current is similarly determined, and is given by, using (1) and (8),

$$I_3 = \sqrt{I_2^2 - I_1^2} = \sqrt{I_2^2 - I_1\,\text{av} \times I_2\,\text{av}}$$

$$= \sqrt{I_2(I_2 - I_1\,\text{av})}. \tag{10}$$

In higher power circuits, and in circuits requiring isolation between the output and input and/or a higher output voltage than input voltage, a transformer must be employed, as shown in Fig. 3. If N is the ratio of secondary voltage to primary voltage, then the following relationships can be shown to exist:

$$I_1\,\text{av} = kNI_2. \tag{11}$$

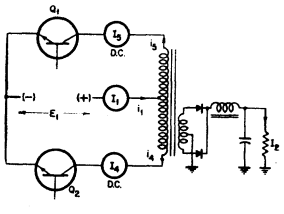


Fig. 3—VWP circuit with output transformer.

For the rms current from the dc source

$$I_1 = \sqrt{I_1\,\text{av} \times I_2 \times N}. \tag{12}$$

For the rms switching devices' currents, assuming balanced conditions,

$$I_4 = I_5 = \sqrt{\frac{I_1\,\text{av} \times I_2 \times N}{2}}. \tag{13}$$

V. Wouk
E. Poss
Electronic Energy Conversion Corp.
New York, N. Y.

Gain of Multisampler Systems*

Two recent notes[1,2] have presented methods for obtaining the output of multisampler systems. I believe the method presented here is simpler than the preceding two since it consists of only one step—writing the answer.

Given the system block diagram, first imagine that all samplers are closed (or that the system is continuous). Since block diagrams are only trivially different from signal flow graphs, Mason's technique may be applied directly to the block diagram. Each path is traversed in the normal manner set down by Mason. The multisampled character of the system is accounted for by writing an asterisk each time a sampler is traversed. If desired, a viniculum can be placed above all terms between this asterisk and the one preceding it to indicate that the sampler operates on the product of these terms. In writing a loop gain, the starting point must be the output of a sampler.

This simple technique is applicable under the following conditions. If interest is centered on the system's loop subgraph, two conditions are of interest in evaluating the gain of these isolated "subsystems." 1) The "subsystem" contains no loops which possess continuous (unsampled) paths from the output of a sampler in a feedback path to the output node of the "subsystem." 2) There exist such paths, but the samplers in question operate on the output of the "subsystem."

If condition 1) applies, the discrete and continuous output of the system can be written as explained above. The last example given by Lendaris[2] satisfies this requirement and can therefore be solved by inspection. If condition 2) applies, only the discrete system output can be found in this manner. However the continuous output can still be determined by inspection using Tou's results as follows:

First open all samplers which violate condition 1). Apply Mason's rules to determine the output response due to input signals. To this result add a term of the form GB^*/A^* where G is the sum of the continuous gains from the output of each opened sampler to the "subsystem" output, B^* is the sum of the Z-transforms of the signals appearing at these open samplers due to the various inputs, and A^* is the sampled determinant found from the loop gains involving the open samplers. (It is assumed that all samplers in the feedback path have a common node.)

For a system having many loops with all of the above conditions violated, the technique proposed in this note would provide a graph reduction process for arriving at a solution.

George C. Reis
Dept. of Electrical Engineering
Drexel Institute of Technology
Philadelphia, Pa.

* Received February 20, 1962.

[1] J. T. Tou, "A simplified technique for the determination of output transforms of multiloop, multisampler, variable-rate discrete-data systems," Proc. IRE (*Correspondence*), vol. 49, pp. 646–647; March, 1961.
[2] G. C. Lendaris, "Input-output relationships for multisampled loop systems," Proc. IRE (*Correspondence*), vol. 49, p. 1709; November, 1961.

A Theorem for the Construction of Root Loci*

Theorem: Given any two polynomials in s with real coefficients: $D(s)$ and $N(s)$ and a constant K. Let $z_1z_2 \cdots z_n$ denote all the roots of $N(s)$, and $p_1p_2 \cdots p_m$ denote all the roots of $D(s)+KN(s)$. The degree of $D(s)$ is higher than that of $N(s)$, $m>n$. Then

$$\prod_{i=1}^{i=m} (z_j - p_i) = C_j \qquad j = 1, 2 \cdots n \tag{1}$$

where $C_1C_2 \cdots C_n$ are constants independent of K.

Proof: Let A_0 denote the coefficient of s^m in $D(s)$. Then

$$D(s) + KN(s) = A_0 \prod_{i=1}^{i=m} (s - p_i). \tag{2}$$

Substituting z_j for s in the above equation gives

$$D(z_j) + KN(z_j) = A_0 \prod_{i=1}^{i=m} (z_j - p_i). \tag{3}$$

Because $N(z_j)=0$ by definition of z_j, (3) is reduced to (1) with $C_j = D(z_j)/A_0$.

The practical significance of the above theorem is that using (1) and other known rules of root-loci plot, a complete set of poles for a given gain can usually be determined directly from the root loci without making extensive algebraic calculations.

A special case of application is to evaluate the vector separation of a closed-loop pole-zero pair. Control systems usually have zeros closer to the origin than the "control poles," and invariably these zeros draw closed-loop poles to their vicinity to form pole-zero pairs. As the latter poles are in the near zone of origin, they are liable to introduce long error tails in the transient response of the closed-loop system. The magnitude of an error tail is directly proportional to the separation of the zero from the pole which gives rise to it.

Another special application is in the synthesis of optimum systems satisfying a quadratic criterion (least mean-square error with limited mean-square control effort, etc.). It

* Received by the IRE, January 15, 1962.

has been shown that the optimum closed-loop system function is related to the transfer function of the controlled plant by a root-square locus plot (root-locus plot in the ω^2 plane).[1] The above theorem helps to reduce the most difficult part of the synthesis procedure (finding the optimum closed-loop system function) to a completely graphical basis.

S. S. L. CHANG
New York University
New York, N. Y.

[1] S. S. L. Chang, "Synthesis of Optimum Control Systems," McGraw-Hill Book Co., Inc., New York, N. Y.; ch. 2; 1961.

Noise Figure of Moody and Wacker's Broad-Band Tunnel Diode Amplifier*

It is the aim of this note to determine the noise figure of Moody and Wacker's broad-band tunnel diode amplifier.[1] An n-stage amplifier is shown in Fig. 1. A tunnel diode of negative resistance $-R_n$ is connected between each two adjacent filter sections. The characteristic impedance of the ith filter section is chosen such that[2]

$$\frac{1}{R_{0i}} = \frac{1}{R_{01}} + \frac{(i-1)}{R_n} \quad (i = 1 \cdots n+1). \quad (1)$$

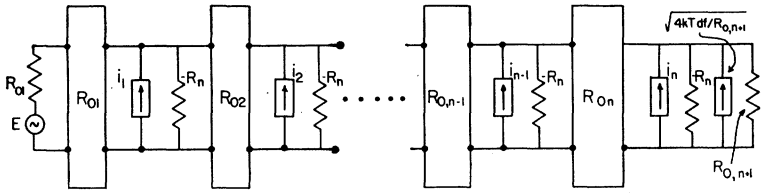


Fig. 1.

The mean-square noise output due to the source is

$$kTdfR_{01}. \quad (2)$$

Let the noise of the diodes be represented by current generators $i_1 \cdots i_n$. If all diodes are identical, have a current I_d, and show full shot noise, then

$$\overline{i_1^2} = \overline{i_2^2} = \cdots = \overline{i_n^2} = 2eI_d df. \quad (3)$$

The ith diode looks into an impedance consisting of the resistances R_{0i}, $-R_n$ and $R_{0,i+1}$ connected in parallel and hence it sees the total resistance $\frac{1}{2}R_{0i}(i=1, \cdots, n)$. Each diode gives instantaneously a voltage $\frac{1}{2}i_iR_{0i}$ at its location in the filter chain and hence a voltage wave $\frac{1}{2}i_iR_{0i}$ flows from the left to the right. There is also a voltage wave flowing from the right to the left and that wave is partially reflected at all the interconnections between the ith diode and the source.

The reflection coefficient at the ith interconnection is

$$\frac{(1/R_{0,i+1}) - (1/R_{0,i} - 1/R_n)}{(1/R_{0,i+1}) + (1/R_{0,i} - 1/R_n)} = \frac{R_{0i}}{R_n} \quad (4)$$

and hence a wave $\frac{1}{2}i_iR_{0i}$ generated at the ith interconnection gives rise to a reflected wave $\frac{1}{2}i_iR_{0i}(R_{0,i-1}/R_n)$ at the previous interconnection and to a wave $\frac{1}{2}i_iR_{0i}(1+R_{0,i-1}/R_n)$ continuing to the left. The total wave coming back to the load, if ϕ is the phase angle due to the travel back and forth along one filter section, is therefore

$$\begin{aligned}
&\frac{1}{2}i_iR_{0,i} + \frac{1}{2}i_iR_{0i}\left(\frac{R_{0,i-1}}{R_n}\right)\cos\phi \\
&+ \frac{1}{2}i_iR_{0i}\left(1+\frac{R_{0,i-1}}{R_n}\right)\left(\frac{R_{0,i-2}}{R_n}\right)\cos 2\phi + \cdots \\
&\quad \frac{1}{2}i_nR_{0i}\left(1+\frac{R_{0,i-1}}{R_n}\right)\left(1+\frac{R_{0,i-2}}{R_n}\right)\cdots \\
&\cdot\left(1+\frac{R_{02}}{R_n}\right)\frac{R_{01}}{R_n}\cos(i-1)\phi \\
&= \frac{1}{2}i_iR_n\left[\frac{R_{0i}}{R_n} + \frac{R_{0i}}{R_n}\frac{R_{0,i-1}}{R_n}\cos\phi\right. \\
&\quad + \frac{R_{0,i-1}}{R_n}\frac{R_{0,i-2}}{R_n}\cos 2\phi + \cdots \\
&\quad \left. + \frac{R_{02}}{R_n}\frac{R_{01}}{R_n}\cos(i-1)\phi\right] \quad (5)
\end{aligned}$$

since $(1+R_{0i}/R_n) = R_{0i}/R_{0,i+1}$.

Hence the noise figure F is:

$$\begin{aligned}
F = 1 + \frac{e}{2kT}I_dR_n\left(\frac{R_n}{R_{01}}\right)\left[\left(\frac{R_{01}}{R_n}\right)^2\right. \\
+ \left(\frac{R_{02}}{R_n} + \frac{R_{02}}{R_n}\frac{R_{01}}{R_n}\cos\phi\right)^2 + \cdots \\
+ \left(\frac{R_{0n}}{R_n} + \frac{R_{0n}}{R_n}\frac{R_{0,n-1}}{R_n}\cos\phi + \cdots\right. \\
\left.\left. + \frac{R_{02}}{R_n}\frac{R_{01}}{R_n}\cos(n-1)\phi\right)^2\right] \quad (6)
\end{aligned}$$

If the noise contribution of the load resistance $R_{0,n+1}$ is taken into account, one obtains a noise figure

$$\begin{aligned}
F_L = F + \frac{T_L}{T}\frac{R_n^2}{R_{0,n+1}R_{01}} \\
\cdot\left[\frac{R_{0n}}{R_n} + \frac{R_{0n}}{R_n}\frac{R_{0,n-1}}{R_n}\cos\phi + \cdots\right. \\
\left. + \frac{R_{02}}{R_n}\frac{R_{01}}{R_n}\cos(n+1)\phi\right]^2 \quad (7)
\end{aligned}$$

where T_L is the noise temperature of the load.

At low frequencies $\phi = 0$, and F_L may be written

$$\begin{aligned}
F_L &= (F_L)_{1.f.} \\
&= 1 + Mn\frac{R_{01}}{R_n} + \frac{T_L}{T}\left(1 + n\frac{R_{01}}{R_n}\right) \quad (7a)
\end{aligned}$$

where $M = eI_dR_n/(2kT)$ is the noise measure of the tunnel diode.

The low-frequency noise figure thus increases linearly with n and is quite large. The high-frequency noise figure is somewhat better, because of the destructive interference of waves coming from the same noise source.

A. VAN DER ZIEL
Elec. Engrg. Dept.
University of Minnesota
Minneapolis, Minn.

Parametric Machines*

It seems that Stockman has rediscovered the reluctance motor[1] and its dual, the electrostatic motor.[2] These elementary machine types are used to introduce the subject of electromechanical energy conversion in at least one modern text book.[3]

Insofar as an electromechanical machine or transducer can be viewed from its electrical terminals as composed of circuit parameters R, L, and C, every case of electromechanical energy conversion may be explained by the time variation of L or C. (Though in the case of the homopolar machine, a stretch of the imagination is necessary.) Variation of the particular parameter is due to motion of one portion of the device with respect to another. Variation of a reactive (energy storage) element is required because the mechanical and electrical systems are coupled by an electric or magnetic field, the energy of which is the medium of exchange. Thus, all electromechanical machinery may be termed "parametric." This circuit viewpoint is not the most fundamental, but it is most convenient for the sake of analysis.

D. J. HANRAHAN
Energy Conversion Branch
U. S. Naval Res. Lab.
Washington, D. C.

Author's Comment[4]

As will be evident from the following, I have not "rediscovered" the reluctance and electrostatic motors, and I shall be glad to explain the difference between my parametric time-constant motors and the motor types described in the Fitzgerald and Kingsley textbook.[3] Actually, my motors have

* Received by the IRE, December 22, 1962; revised manuscript received, January 8, 1962.
[1] N. F. Moody and A. G. Wacker, "Broad-band tunnel diode amplifier," PROC. IRE, vol. 49, p. 835; April, 1961.
[2] This means that the line is matched at all interconnections for waves going from the left to the right.

* Received January 8, 1962.
[1] H. E. Stockman, "Parametric oscillatory and rotary motion," PROC. IRE (*Correspondence*), vol. 48, pp. 1157–1158; June, 1960.
[2] H. E. Stockman, "Parametric variable-capacitor motor," *ibid.*, vol. 49, pp. 970–971; May, 1961.
[3] A. E. Fitzgerald and C. Kingsley, Jr., "Electric Machinery," McGraw-Hill Book Co., Inc., New York, N. Y.; 1952.
[4] Received January 19, 1962.

very little in common with the old motor types as far as basic theory goes. This becomes apparent from a study of the instantaneous torque equation for the old motor,

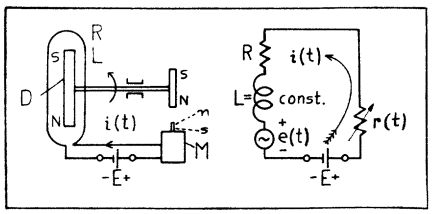
$$T = -\text{const.}\,\phi^2\frac{d\mathcal{R}}{d\alpha}, \qquad (1)$$

where ϕ is the instantaneous flux, $\mathcal{R}$ the reluctance, and α associated with rotor angular deviation. T switches sign with the derivative, so that synchronous operation with $\omega_{\text{lines}} = \omega_{\text{rotor}}$ is the only possibility in the basic case. No average torque is produced at any other speed (except for harmonic effects). Eq. (1) also shows that the old motor lacks dc torque and therefore is not self-starting. The maximum average torque expression

$$T_{\max} = \text{const.}\,\phi^2_{\max}(\mathcal{R}_Q - \mathcal{R}_P) \qquad (2)$$

spells out the unfortunate limitation of the old motor that if $T_{\max}$ is exceeded by the torque requirement of the mechanical loading, loss of motor action and shutdown results. My time-constant motor behaves quite differently since it is an *ASYNCHRONOUS* motor, not intended for operation in a synchronous state. Accordingly, the reluctance rate of change is largely independent of ω_{line}, and if the lines frequency were to be suddenly increased from 60 to 70 cps, as an experiment of thought, the old motor would fall out of synch, while my motor merely would slow down, and then continue at constant speed. Similarly, mechanical overload makes my motor slow down, while in accordance with (2), it makes the old motor fall out of synch. Further, my motor is self-starting, except for one peculiar position of the rotor; a defect that can be remedied. These striking differences are logical in view of the fact that the quoted textbook, on page 89, claims that it is discussing a synchronous machine of reluctance type, while my motor is an asynchronous machine of time-constant type. In essence, the time-constant motor is a brushless dc motor, using an ac carrier supply as dc substitute in about the same fashion as a magnetic amplifier. It extracts its driving torque from the double-valued force-function established by the time constant; a phenomena totally different from that causing the torque in the old motor. Any suspicion that my motor is a rediscovery of an old motor is unfounded, at least as far as a comparison with the old reluctance motor goes. The same is true in a comparison of my electric field version asynchronous motor with the electric field synchronous motor described in the quoted textbook, which correctly claims that the analysis of its electrostatic machine parallels that of its electromagnetic machine.

Within the ramifications of his note, Hanrahan states without any verification that every case of electromagnetic energy conversion may be explained by the time variation of L and C, that variation of a reactive element is required, and that all electromechanical machinery may be termed "parametric." That these opinions are unjustified is indicated by the fact that if a fundamental law were to be formulated, it would be just the opposite of what is claimed by Hanrahan: "Parametric action is *NOT* a necessary requirement for energy conversion." This is amply demonstrated by a commercially available motor, shown in principle with its equivalent two-terminal network in Fig. 1.[5] This motor is parametric in a quantity not even included as a possibility in Hanrahan's statement, namely R. It consists of a spinning magnet NS, which via a smaller magnet ns activates the variational resistance device M, which functions like a carbon mike. The produced ac drives via the inductor L the nonsalient pole magnetic disk D. There is no intended or essential change in any storage element, but still the device converts electrical energy to mechanical energy, and quite successfully, too. If the variational resistance device M is replaced by a brush and a contact element, a nonparametric motor results, demonstrating the most fundamental fact that we can do without parametric action in energy conversion, if we so wish.



(a) (b)

Fig. 1.

HARRY E. STOCKMAN
Lowell Tech. Inst.
Lowell, Mass.

[5] Made by SINE-SER Co., Waltham, Mass.

Dependence of Parametric Element Nonlinearity on Tuning Circuit*

In a recent letter on the behavior of the nonlinearity in a parametric device employing a capacitor, Helgesson[1] states that the difference produced by series and parallel tuning affects only the second and higher order nonlinear terms and leaves the first unchanged. This is correct only for low-level pumping in which the diode is swept through a small portion of its characteristic. Helgesson's Eq. (5), which compares the proportional reactance sweep in the series tuned circuit with the proportional susceptance sweep in the parallel tuned circuit, assumes an approximate linear relation between charge and voltage which is invalid for high-level pumping. It will be shown below that under these conditions considerable advantage can be obtained by using a series tuned high-impedance terminated circuit.

The standard method of analyzing each of the dual arrangements is identical but for the interchange of δq, the incremental charge on the nonlinear capacitor, and δv, the incremental voltage, and starts from a power series of δq in terms of δv for the parallel tuning and δv in terms of δq for the series tuning. These series can be conveniently compared if they are expressed in terms of the dimensionless quantities $\delta q/q_0$ and $\delta v/v_0$, the charge q_0 and the voltage v_0 corresponding to the operating point and being measured with respect to the point of infinite ac capacitance.

Assuming the static characteristic to be of the form

$$q = kv^{(1-n)}, \qquad (1)$$

one obtains

$$\delta q/q_0 = (1-n)\{\delta v/v_0 + (n/2!)(\delta v/v_0)^2 + \cdots\} \qquad (2)$$

$$\delta v/v_0 = (1-n)\left\{dq/q_0 - \frac{n}{2!(1-n)}(\delta q/q_0)^2 + \cdots\right\} \qquad (3)$$

for the series and parallel tuning respectively. With these variables the second expansion has the greater nonlinearity. In each case the maximum gain-bandwidth gB of the amplifier is given by the ratio of the second to the first term. Thus

$$gB = (n/2)(\delta v/v_0) \qquad (4)$$

for parallel tuning, and

$$gB = \frac{n}{2(1-n)}\frac{(\delta q)}{(q_0)} \qquad (5)$$

for series tuning.

Hence, provided $\delta q_{\max}/q_0$, the maximum fractional charge excursion, is not less than $(1-n)\delta v_{\max}/v_0$ the series tuning will be preferable. This can be shown to be true as follows.

A typical voltage charge characteristic is shown in Fig. 1, upper and lower limits of satisfactory operating voltage, v_2 and v_1, being indicated. These might correspond to the points of reverse breakdown and forward conduction respectively. The whole range between v_1 and v_2 may be utilized if the parallel tuned amplifier is biased to a voltage $(v_1+v_2)/2$ and the series tuned amplifier to a charge of $(q_1+q_2)/2$. The corresponding maximum voltage and charge excursions are given by

$$\delta v_{\max}/v_0 = (1 - v_1/v_2)/(1 + v_1/v_2) \qquad (6)$$

$$\delta q_{\max}/q_0 = (1 - q_1/q_2)/(1 + q_1/q_2) = \{1 - (v_1/v_2)^{1-n}\}/\{1 + (v_1/v_2)^{1-n}\}. \qquad (7)$$

Investigation of these shows that as v_1/v_2 varies from unity (small allowable capacitance variation) to zero (maximum allowable capacitance variation) $(\delta q_{\max}/q_0)/(\delta v_{\max}/v_0)$ varies from $1-n$ to unity. Hence from (4) and (5), the ratio of the gain bandwidth of the two amplifier arrange-

* Received January 15, 1962.

[1] A. L. Helgesson, "Nonsymmetrical properties of nonlinear elements in low- and high-impedance circuits," PROC. IRE (*Correspondence*), vol. 49, p. 1569; October, 1961.

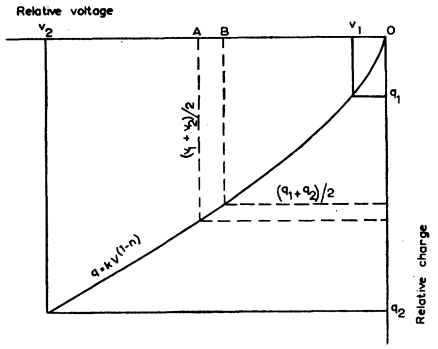


Fig. 1—Voltage-charge characteristic of nonlinear capacitance showing limits of operation and biasing points.

ments varies from unity for small pumping to $1-n$ for large pumping always in favor of the series tuned layout.

G. H. B. Thompson
Standard Telecommun. Labs., Inc.
Harlow, Essex, England

Author's Reply[2]

Thompson has raised a very good point, with which the author fully agrees. Series and parallel tuning are compared for equal bias conditions,[1] but as Thompson shows, optimum pumping is achieved in each case for different bias points. For $0<n<1$, the optimum bias point for series tuning is always in a region of higher nonlinearity, resulting in larger gain-bandwidth product.

It is interesting to compare the maximum pumping obtainable from abrupt and graded junction varactors in series and parallel resonant circuits. In parallel tuning, $\Delta C_B/C_B$ is determined from (4) of Helgesson[1] with $|v(t)|=\phi-V_B$, and $\Delta S_B/S_B$ for series tuning is computed from this by using Thompson's improvement factor of $1/(1-n)$.

	$n=\frac{1}{2}$	$n=\frac{1}{3}$
Parallel Tuning $(\Delta C_B/C_B)_{\max}$	$\frac{1}{4}$	$\frac{1}{6}$
Series Tuning $(\Delta S_B/S_B)_{\max}$	$\frac{1}{2}$	$\frac{1}{4}$

Of the four combinations, the series tuned circuit with an abrupt junction varactor has the highest gain-bandwidth capability in addition to the elimination of the higher order perturbations mentioned previously. Series tuning has the additional advantage that in the frequency regions of greatest interest at present, the source of loss in the varactor is series resistance, and the second harmonics and other mixing products are prevented from dissipating power in this resistance because of the high impedance termination at these frequencies. For parallel tuning, these frequencies cannot be completely short circuited, and they develop some voltage across the capacitance and dissipate some loss in the series resistance.

There is, however, one minor disadvantage of series tuning that has not been mentioned to date. When the charge voltage expansions are derived with the dc terms included,[3] it can be seen that average elastance is a function of the magnitudes of the ac charges. This has two effects in the circuit. The first is that the optimum bias voltage is not determined, as might be incorrectly inferred from point B of Thompson's diagram, by taking the voltage corresponding to the average charge. Because of the shift in bias charge due to the pumping, the average charge is not uniquely determined by the bias voltage. In Thompson's notation, if $V_1=0$ and $n=\frac{1}{2}$, point A is $V_2/2$ and point B is $V_2/4$. It can be shown that the optimum bias voltage for series tuning is $\frac{3}{8}V_2$.

The second effect produced by the average-charge shift is that the average elastance is a function of the signal amplitude. This produces a slight phase distortion which may be important in systems that require phase coherence. However, this perturbation can be shown to be about a factor of 3 less than that produced by the third-order term for parallel tuning, as calculated from (6) of Helgesson.[1] [Eqs. (6) and (7) of this reference are in error. The right-hand sides of both should be multiplied by ω_i.]

In summary, the case for series tuning appears to be well established.

Alan L. Helgesson
Sylvania Elec. Products, Inc.
Waltham, Mass.

Varactor Charge-Voltage Expansions for Large Pumping Conditions*

In a previous correspondence,[1] charge-voltage expansions for a varactor were derived for small ac variations about a fixed bias point. The purpose of this correspondence is to extend these results to the large pumping case by deriving the complete expansions with the inclusion of the dc terms. It will be shown that (2) of Helgesson is valid under large-pumping conditions, but that (3) must be modified. From these new expansions, the dependence of the average charge and the average elastance on the magnitudes of the ac variations is calculated.

The expression for the derivative of the total charge as a function of the total voltage,

$$\frac{dQ}{dV}=C_0\left[1-\frac{V}{\phi}\right]^{-n}, \tag{1}$$

is integrated, and the boundary condition that Q is zero when V is zero is imposed, giving

$$Q=\frac{C_0\phi^n}{1-n}\left[\phi^{1-n}-(\phi-V)^{1-n}\right]. \tag{2}$$

This is the fundamental equation relating charge and voltage on the varactor.

Charge Expansion

We are now separate dc and ac components and define

$$\begin{aligned} Q(t) &= Q_{\text{ave}}+q(t) \\ V(t) &= V_{\text{ave}}+v(t) \end{aligned} \tag{3}$$

where $q(t)$ and $v(t)$ have zero mean. For the first expansion, we will consider the case in which the circuitry external to the varactor restricts the average voltage to be equal to that of a bias supply

$$V_{\text{ave}}=V_B, \tag{4}$$

and allows $v(t)$ to consist of only a small finite number of sinusoidal components

$$v(t)=\sum_{k=1}^{K}|V_k|\cos(\omega_k t+\theta_k). \tag{5}$$

For this case we write (2) in the form

$$Q_{\text{ave}}+q(t)=\frac{C_0\phi^n(\phi-V_B)^{1-n}}{1-n}\cdot\left[\left(\frac{\phi}{\phi-V_B}\right)^{1-n}-\left(1-\frac{v(t)}{\phi-V_B}\right)^{1-n}\right] \tag{6}$$

and expand into the series

$$\begin{aligned} Q_{ave}+q(t) &= \frac{C_0\phi^n}{1-n}\left[\phi^{1-n}-(\phi-V_B)^{1-n}\right] \\ &+C_B\left[v(t)+\frac{nv^2(t)}{2!(\phi-V_B)}+\frac{n(n+1)v^3(t)}{3!(\phi-V_B)^2}\right. \\ &\left.+\frac{n(n+1)(n+2)v^4(t)}{4!(\phi-V_B)^3}+\cdots\right] \\ &=Q_B+C_Bv(t)+\mathcal{C}_a v^2(t)+\mathcal{C}_b v^3(t)+\cdots \end{aligned} \tag{7}$$

where the average capacitance as determined by the bias voltage is

$$C_B=\frac{C_0\phi^n}{(\phi-V_B)^n}. \tag{8}$$

Inspection of (7) shows that average charge is not completely determined by the bias voltage, but has a component dependent upon the magnitides of the ac voltages. Writing

$$Q_{\text{ave}}=Q_B+Q_{\text{ac}} \tag{9}$$

the first term in (7) gives the charge due to the bias voltage

$$Q_B=\frac{C_0\phi^n}{1-n}\left[\phi^{1-n}-(\phi-V_B)^{1-n}\right] \tag{10}$$

and the contribution from the ac voltages is

$$Q_{\text{ac}}=C_B\left[\frac{n}{2!(\phi-V_B)}\frac{1}{2}\sum_{k=1}^{K}V_k^2+\frac{n(n+1)(n+2)}{4!(\phi-V_B)^3}\frac{3}{8}\sum_{k=1}^{K}V_k^4+\cdots\right]. \tag{11}$$

In this case the fact that the average charge changes with the magnitudes of the ac voltages is only of academic interest, since all constants multiplying powers of $v(t)$ in (7) are determined by the bias voltage, and the circuit properties of the device are independent of the average charge. Eq. (2) of Helgesson[1] is therefore valid for large pumping, except that it contains the additional average charge Q_{ac}.

[2] Received January 30, 1962.
[3] A. L. Helgesson, "Varactor charge-voltage expansions for large pumping conditions," this issue, same page.

* Received January 30, 1962.
[1] A. L. Helgesson, "Nonsymmetrical properties of nonlinear elements in low- and high-impedance circuits," PROC. IRE (*Correspondence*), vol. 49, pp. 1569–1570; October, 1961.

Voltage Expansion

For the second expansion we consider the case in which the ac charge variation is restricted to consist of a small finite number of sinusoidal components.

$$q(t) = \sum_{k=1}^{K} |Q_k| \cos(\omega_k t + \theta_k). \quad (12)$$

It would be most desirable if we could simultaneously impose a restriction on the average charge, but unfortunately, dc charge sources are not available. Instead, biasing of the device must be done with a voltage source and the constraint on the average value is again given by (4).

Solving (2) for V in terms of Q gives

$$V = \phi - \left[\phi^{1-n} - \frac{Q(1-n)}{C_0\phi^n}\right]^{1/(1-n)}. \quad (13)$$

There are several expansions that can be made in this mixed system of constraints. The one that leads to the most convenient form for the circuit equations is obtained by writing (13) in the form

$$V_B + v(t) = \phi - \left[\left(\phi^{1-n} - \frac{Q_B(1-n)}{C_0\phi^n}\right) - \frac{[Q_{ac} + q(t)](1-n)}{C_0\phi^n}\right]^{1/(1-n)}, \quad (14)$$

recognizing from (10) that

$$\phi^{1-n} - \frac{Q_B(1-n)}{C_0\phi^n} = (\phi - V_B)^{1-n}, \quad (15)$$

and expanding as

$$V_B + v(t) = V_B + \frac{Q_{ac} + q(t)}{C_B} - \frac{n[Q_{ac} + q(t)]^2}{2!(\phi - V_B)C_B^2} - \frac{(1-2n)n[Q_{ac} + q(t)]^3}{3!(\phi - V_B)^2C_B^3} - \cdots$$
$$= V_B + S_B[Q_{ac} + q(t)] - S_a[Q_{ac} + q(t)]^2 - S_b[Q_{ac} + q(t)]^3 - \cdots. \quad (16)$$

This is the complete form for (3) of Helgesson.[1] The solution for Q_{ac} may be made by noting that $v(t)$ has zero average value so that

$$\frac{Q_{ac}}{C_B} - \frac{n\left[Q_{ac}^2 + \frac{1}{2}\sum_{k=1}^{K}|Q_k|^2\right]}{2!(\phi - V_B)C_B^2} - \frac{(1-2n)nQ_{ac}^3}{3!(\phi - V_B)^2C_B^3} - \cdots = 0. \quad (17)$$

In general, the solution of this equation is difficult because it is of high order in Q_{ac}. However, the case of greatest practical interest is fortunately that for which $n = \frac{1}{2}$, for which the equation reduces to a quadratic. The average charge contributed by the ac variations is then

$$Q_{ac} = 2(\phi - V_B)C_B \cdot \left[1 - \sqrt{1 - \frac{1}{8(\phi - V_B)^2C_B^2}\sum_{k=1}^{K}|Q_k|^2}\right]. \quad (18)$$

For an example of the effect of Q_{ac}, let us consider a three frequency device in which

$$q(t) = \frac{Q_1e^{j\omega_1 t} + Q_1^*e^{-j\omega_1 t}}{2} + \frac{Q_2e^{j\omega_2 t} + Q_2^*e^{-j\omega_2 t}}{2} + \frac{Q_3e^{j\omega_3 t} + Q_3^*e^{-j\omega_3 t}}{2} \quad (19)$$

and

$$\omega_1 + \omega_2 = \omega_3. \quad (20)$$

The voltage component produced at $+\omega_1$, for example, is from (16)

$$V_1 = \frac{Q_1}{C_B} - \frac{Q_{ac}Q_1}{2(\phi - V_B)C_B^2} - \frac{Q_2^*Q_3}{4(\phi - V_B)C_B^2}$$
$$= \frac{Q_1}{C_B}\sqrt{1 - \frac{|Q_1|^2 + |Q_2|^2 + |Q_3|^2}{8(\phi - V_B)^2C_B^2}} - \frac{Q_2^*Q_3}{4(\phi - V_B)C_B^2}. \quad (21)$$

The magnitudes of the ac charges change the average elastance of the varactor in each of the three circuits to the value

$$S_{ave} = S_B\sqrt{1 - \frac{|Q_1|^2 + |Q_2|^2 + |Q_3|^2}{8(\phi - V_B)^2C_B^2}}. \quad (22)$$

The term producing power conversion is unaffected and is determined only by the bias voltage. The elastance shift due to the fixed pump magnitude may be accounted for by proper design, but a small perturbation dependent upon the magnitudes of signal and idler charges will always exist.

To compute the maximum pumping obtainable, we assume that the pump charge is large in comparison to that of the signal and idler, and that the varactor is driven to the point of forward conduction. Setting $V_B + v(t) = \phi$ and $q(t) = |Q_p|_{max}$ in (16) we solve for the pump charge with the help of (18). For $n = \frac{1}{2}$ the result is

$$|Q_p|_{max} = \sqrt{\frac{8}{3}}(\phi - V_B)C_B. \quad (23)$$

The relative elastance change for maximum pumping is

$$\frac{\Delta S_{ave}}{S_{ave}} = \frac{S_a|Q_p|_{max}}{S_{ave}} = \frac{S_B\sqrt{\frac{1}{6}}}{S_B\sqrt{\frac{2}{3}}} = \frac{1}{2}, \quad (24)$$

which is twice the maximum $\Delta C_B/C_B$ obtainable in a parallel-tuned configuration.

If the varactor is not pumped to its maximum value, but is to be driven between a specified upper voltage V_u and a lower voltage V_l, the pumping charge and the proper bias voltage may again be computed from (16) and (18). The general solution is algebraically cumbersome, but if we assume $\phi \approx 0$ the bias voltage is given by the simpler form

$$V_B = \frac{3[V_l + V_u] - 2\sqrt{V_lV_u}}{8}. \quad (25)$$

In addition to presenting the complete forms of the charge-voltage expansions, the important conclusions to be drawn from this analysis are that in the charge controlled configuration, the average charge and average elastance are functions of the ac charge magnitudes, and that more percentage pumping may be obtained than in the voltage controlled circuit.

Alan L. Helgesson
Sylvania Elec. Products Inc.
Waltham, Mass.

A Broad-Band Ku-Crystal Diode Switch*

A broad-band diode switch which can be made to operate over the 12.0 to 17.0 Gc frequency range with a minimum bandwidth of 2 Gc is described in this note. This bandwidth is described here to be an isolation greater than 20 db and insertion loss less than or equal to 1.5 db. The results shown here were taken with a D4151A diode, although from the data it appears that other germanium point contact diodes in suitable packages would work equally well. The D4151A was purposely designed to have the lead inductance and package capacitance one half the value normally found in commercial glass-package germanium diodes. A pill-type package should work equally well in this band.

Because the IN283 glass diode worked extremely well in waveguide at X band, we initially tried extending the switching action to Ku band with this type diode. Although we obtained definite switching action, it was not too successful. The difference between the isolation and insertion states was only on the order of 5 db above 13 Gc with insertion losses of 4 to 6 db. Obviously this is not a good device for systems applications. Since the IN283 worked extremely well from 5 Gc to 12.5 Gc, we thought that since the lead inductance and package capacitance were lowered by a factor of two, if the switching states were somehow a function of these parameters, combined with the junction capacitance, switching might be made to occur at higher frequencies with the D4151A diode used as the switching elements.

Studies made at this laboratory on waveguide semiconductor switches from 5 Gc to 12 Gc indicated quite strongly that by reactively loading the diodes, better switching results are obtained. The reason for this appears to be that the diode and the reactive structures act together as a band-pass filter element where the reactive element is sized for the particular frequency range under consideration.

In this frequency range, a single diode element, suitably loaded, gives a 0.9 Gc bandwidth. However, two diode elements suitably loaded and spaced one quarter wavelength apart give substantially better results. The bandwidth is much larger and the switching ratio is improved over the band. Switching ratio is defined as the difference between the isolation and insertion losses in db. The isolation is due to the reactive structure and diode presenting a large reactive impedance which results in the energy being reflected back toward the source. Normally this reflected energy is 1 db down from the incident power, where the 1 db is dissipated across the resistances within the diode.

Fig. 1 shows the results obtained from a one-diode reactive loaded switch. Here the peak occurs at 13.9 Gc, although this peak can be shifted up and down in frequency by suitably changing the reactive loading on the diode. Fig. 2 shows the characteristics of a two-diode switch which is

* Received by the IRE, January 5, 1962.

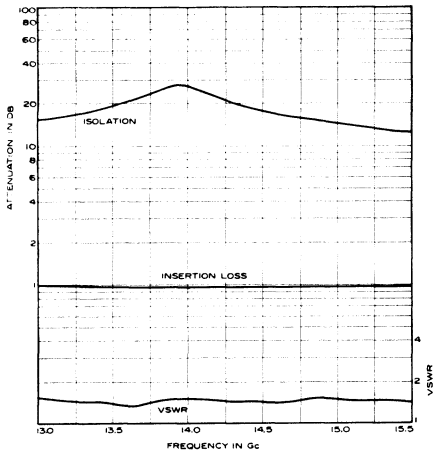


Fig. 1—Characteristics of a single-diode Ku-band switch.

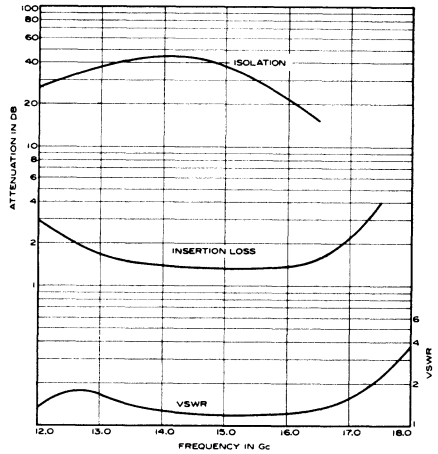


Fig. 2—Characteristics of a two-diode Ku-band switch.

improved over the one-diode switch. The first apparent factor is the increase in the bandwidth which is larger by a factor greater than two. The isolation also increased greatly while the insertion loss increased by a factor somewhat less than two. The insertion losses shown in Figs. 1 and 2 can be greatly reduced because our biasing arrangements are not ideal and some energy is being dissipated in the dc input. For the Ku-band switch reported here, the reactive loading took the form of capacitive elements at the diode position. This loading will vary appreciably with frequency at X and C band for a gold-bonded germanium IN283, the loading becomes inductive.

The isolation condition shown in Figs. 1 and 2 occurred with the diode biased in the reversed direction at -50 v. At -10-v bias, the isolation is nearly 90 per cent of the value recorded at -50 v. In the insertion direction, the loss is inversely proportional to the forward current; the larger the forward current, the smaller the insertion loss. It is the ability of the diode to handle current that limits the insertion loss since the forward resistance of the diode is the one source of dissipative losses in the diode.

We have not made power tests on this switch, but we expect the switch to propagate an average power of 1 w and a peak power of 150 w at 001 duty cycles. These estimates are derived from an examination of the diode's dc characteristics. These, however, need to be verified. But the feasibility of the switch for low-power receiver applications is obvious where relatively large bandwidths are required. We also feel that crystal switches can be extended to higher frequency ranges if careful attention is paid to package design and reactive loading.

The authors would like to acknowledge the assistance of M. Groll of Sylvania, Waltham, Mass., for supplying many diodes.

K. W. Beck
J. J. Rowley
Sylvania Electric Products, Inc.
Microwave Device Division
Mountain View, Calif.

Radiation from a Magnetic Line Dipole Source of Finite Width*

The problem of radiation from a magnetic line dipole source lying in a lossy plane is of great interest since, in addition to radiating components, a surface wave is also launched. What is of importance in this problem is the fact that if the line dipole source is of finite width the surface-wave component can be either reduced to zero, or it can be maximized.

Assume a line dipole source lying in the $y=0$ plane, extending from $z=-\infty$ to $z=+\infty$. For $y>0$, let the medium be characterized by μ_0, ϵ_0, and the medium for $y\leq 0$ by μ, ϵ such that the wave number at the surface $y=0$ is given by $k_s=i\omega\epsilon_0(R_s+iX_s)=2\pi/\lambda_s$. The solution of this problem, normalized for a point source, is given by[1]

$$u(x,y) = -i\pi\left[H_0^{(2)}(kr) - k_s e^{-k_s y}\int_{-\infty}^{y} e^{k_s\eta}H_0^{(2)}(k\rho)d\eta + \frac{2k_s\exp(-k_s y)\exp(-i\sqrt{k_s^2+k_s^2}\,|x|)}{(k^2+k_s^2)^{1/2}}\right] \quad (1)$$

for harmonic time dependence $e^{i\omega t}$; $u(x,y)=H_z$, and $\partial/\partial z$ was assumed zero; $k^2=\omega^2\mu_0\epsilon_0=(2\pi/\lambda_0)^2$; $H_0^{(2)}(kr)$ are Hankel functions of second kind, and order zero and $r^2=x^2+y^2$, and $\rho^2=x^2+\eta^2$.

The first and second term of (1) represent the source and secondary contribution to the cylindrically radiated field; the last term represents the surface wave. Let us, for the moment, consider this last term alone, and extend the solution to a number of line dipole sources distributed from $x=-d/2$ to $x=d/2$. Assuming that there is no interaction between the sources, and that the sources are in phase, the total contribution to the surface-wave field will be given by

$$U_s = \sum_n u_{sn}[(x-nx_0),\, y], \quad (2)$$

where x_0 is the distance between neighboring line dipole sources. If there are N sources distributed uniformly over a distance d, (2) can be written as

$$U_s = \sum_{n=-N/2}^{N/2} u_{sn}\left[\left(x-\frac{nd}{N}\right),\, y\right] \rightarrow \int_{-N/2}^{N/2} u_{sn}\left[\left(x-\frac{nd}{N}\right),\right] y\,dn \quad (3)$$

for large N.

Substitution of the expression for u_{sn}, from (1)

$$u_{sn} = A\exp\left[-i\sqrt{k^2+k_s^2}\left|x-\frac{nd}{N}\right|\right], \quad (4)$$

where

$$A = -\frac{2\pi i k_s}{(k^2+k_s^2)^{1/2}}\exp(-k_s y),$$

and integration between the indicated limits results in

$$U_s\big|_{|x|>\frac{d}{2}} = NA\frac{\sin B}{B}\exp(-i\sqrt{k^2+k_s^2}\,|x|) \quad (5)$$

$$U_s\big|_{|x|\leq\frac{d}{2}} = -\frac{NA}{B}\left[1-\cos\left(\frac{2B}{d}x\right)\exp(-iB)\right] \quad (6)$$

where $2B=d(k^2+k_s^2)^{1/2}$. In the limit as $B\rightarrow 0$, (6) reduces to zero, and (5) to the surface wave term due to one line dipole source at (0, 0).

It is evident that for $x>d/2$ the surface-wave term will vanish if $B=\pm m\pi$, $m=1, 2, \cdots$, in which case k_s must be pure real or pure imaginary. There is no value of B for which (5) can be equal to zero in the case of complex k_s. Assuming, then, that k_s is either pure real or pure imaginary,

$$d = \pm\frac{2m\pi}{(k^2\pm k_s^2)^{1/2}} = \pm\frac{m\lambda_0\lambda_s}{(\lambda_s^2\pm\lambda_0^2)^{1/2}} \quad (7)$$

for elimination of the surface-wave term for all values of $x>d/2$. It follows from (1) and (7) that for k_s pure imaginary we actually have a "leaky" surface wave,[2] whereas for k_s pure real we have a "true" surface wave. This must be so, since for $k_s^2<0$ real values of d require $\lambda_s>\lambda_0$, *i.e.*, the phase velocity along the surface will exceed c, the speed of light.

When (7) is satisfied, we obtain from (6)

* Received by the IRE, January 10, 1962.

[1] S. N. Karp, F. Karal, "Surface Waves on a Right Angled Wedge," AF Cambridge Res. Ctr., Bedford, Mass., Tech. Rept. No. AFCRC-58-368; August, 1958.

[2] F. J. Zucker, "The guiding and radiation of surface waves," *Proc. Symp. on Modern Advances in Microwave Techniques*, Polytechnic Inst. of Brooklyn, Brooklyn, N. Y.; November 8–10, 1954.

$$U_s|_{|x|} \leq \frac{d}{2} = -\frac{2NA}{m\pi}\cos^2\left(\frac{m\pi}{d}x\right) \quad (8)$$

representing a standing wave with nodes at $x = \pm d/2m$.

Somewhat similar results are obtained if the individual sources are not fed in phase. If a phase shift

$$\beta_n = \beta_0\left(\frac{d}{2} + \frac{nd}{N}\right)$$

is assumed from source to source

$$U_s|_{|x|} > \frac{d}{2}$$

$$= NA\frac{\sin D}{D}\exp\left[-i\sqrt{k^2+k_s^2}\,|x| + \frac{\beta_0 d}{2}\right] \quad (9)$$

with $2D = d(\sqrt{k^2+k_s^2}+\beta_0)$. As in the preceding case, the surface wave for $x > d/2$ will disappear for $D = \pm m\pi$, $m = 1, 2, \cdots$.

In both cases, the surface-wave term will be a maximum when $d \to 0$. Consequently, it can be concluded that for maximum efficiency in launching a surface wave a source of infinitesimal thickness is required. Furthermore, since in this case k_s must be pure real, a purely inductive surface impedance is required, in agreement with conditions for support of surface waves derived by Barlow and Cullen.[3] When the surface impedance is complex, or pure real, a "leaky" surface wave is obtained. The surface-wave term can be eliminated only if the surface impedance is not complex.

LEON W. ZELBY
The Moore School of Electrical Engineering
University of Pennsylvania
Philadelphia, Pa.

[3] H. E. M. Barlow, A. L. Cullen, "Surface waves," *Proc. IEE*, vol. 100, pt. 3, pp. 329–347; November, 1953.

Sensitivity of a Magnetoresistance Wattmeter*

As reported previously,[1,2] it was proved that the magnetoresistance effect in a semiconductor could be used as a wattmeter. Here, the theory has been developed in connection with a Hall-effect wattmeter.

The resistance of a rectangular semiconductor specimen of length l, width w, thickness t under a high magnetic field B as shown in Fig. 1(a) is approximately expressed by

$$R(B) = R(0)\frac{\rho(B)}{\rho(0)}\left\{1 + \frac{w}{l}\left(\mu B - \frac{4}{\pi}\ln 2\right)\right\}, \quad (1)$$

* Received by the IRE, January 2, 1962.

[1] S. Kataoka, "Magneto-resistance multiplier with higher gain," PROC. IRE, to be published.

[2] S. Kataoka and S. Kobayashi, "The application of the magnetoresistance effect in an intermetallic semiconductor to the measurement of electric power (Part I. Fundamental)," *Bull. Electrotech. Lab.*, vol. 25, pp. 827–834; November, 1961.

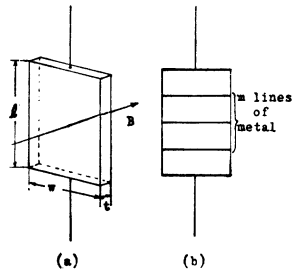


Fig. 1—Semiconductor element. (a) Without metal line. (b) With m metal lines.

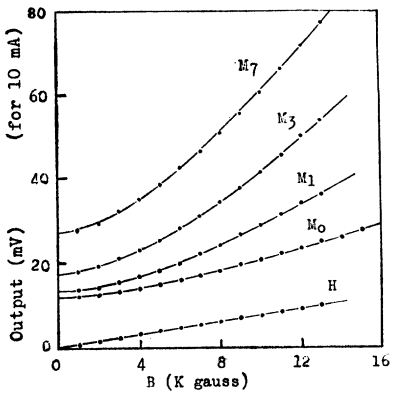


Fig. 2—Galvanomagnetic characteristics of an InAs crystal. H: Hall effect. Mm: Magnetoresistance with m lines of metal.

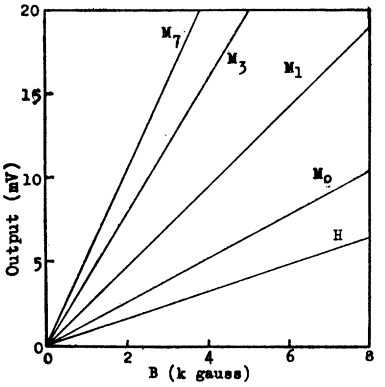


Fig. 3—Output characteristics of a magnetoresistance wattmeter. H: Hall effect. Mm: Magnetoresistance with m lines of metal.

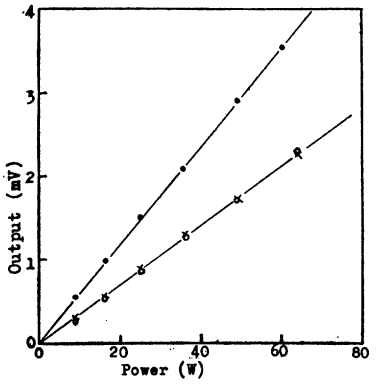


Fig. 4—Output characteristics of a Halltron wattmeter; ○ = Hall output, × = magnetoresistance output (Hall terminals opened), ● = magnetoresistance output (Hall terminals shorted).

where

$R(B)$ = resistance in a magnetic flux density B,
$R(0)$ = resistance without magnetic field,
$\rho(B)$ = resistivity in a magnetic flux density B,
$\rho(0)$ = resistivity without magnetic field,
μ = mobility of charge carriers.

The first term $\rho(B)/\rho(0)$ increases with B^2 and reaches a saturation at high value of B. The second term $\{1 + w/l(\mu B - 4/\pi \ln 2)\}$ is obtained by analyzing an electric field distortion employing the Schwarz-Christoffel conformal transformation.[3]

Thus, $R(B)$ increases linearly with B at high magnetic field and a voltage drop variation across the semiconductor is proportional to a product of a current I passed through it and a magnetic field variation B', if the semiconductor is biased magnetically at a relatively high magnetic field B_0. At ac, the time-average voltage of the semiconductor is proportional to the active product of B' and I, that is $\bar{v} = KB'I\cos\phi$, to be used as a wattmeter. Here, K, a proportionality constant, represents a sensitivity of the device and can be found by differentiating (1), in considering $R(0) = \rho(0)l/wt$ and $\rho\mu = \mathfrak{R}_H$.

$$K = \frac{dR(B)}{dB} = R(O)\frac{\rho(B)}{\rho(0)}\frac{w}{l}\mu = \frac{\mathfrak{R}_H}{t}, \quad (2)$$

where $\mathfrak{R}_H$ is a Hall coefficient of the semiconductor.

It is quite interesting to note that the sensitivity of a magnetoresistance wattmeter is exactly the same as a Hall-effect wattmeter in so far as the semiconductor element remains the same

Next, let us consider a case where m lines of metal are attached onto the semiconductor plate as shown in Fig. 1(b). The resistance of such an element is

$$Rm(B) = Rm(0)\frac{\rho(B)}{\rho(0)} \cdot \left\{1 + \frac{(m+1)w}{l}\left(\mu B - \frac{4}{\pi}\ln 2\right)\right\}, \quad (3)$$

and the sensitivity of a wattmeter

$$Km = \frac{dRm(B)}{dB} = \frac{(m+1)\mathfrak{R}_H}{t}. \quad (4)$$

This relation leads to a conclusion that the output would be increased by a factor of $(m+1)$.

Fig. 2 shows experimental results of the magnetoresistance and the Hall-effect characteristics of an InAs polycrystal specimen of 8×4×0.1 mm and Fig. 3 the deduced output characteristics of both kinds of wattmeter. Fig. 4 shows the results on output characteristics of a Halltron wattmeter[1,2] in both ways of use.

[3] V. H. J. Lipmann and F. Kuhrt, "Der Geometrieeinfluss auf den transversalen magnetischen Widerstandseffekt bei rechteckförmigen Halbleiterplatten," *Z. Naturforsch.*, Band 13a, pp. 462–474; June, 1958.

All these experimental results seem to verify the above theoretical considerations. The considerable deviation from the theory for large m may be due to the facts that the metallic line has a width of 0.2 mm and that these were soldered on only one side of the specimen.

SHOEI KATAOKA
Electrotechnical Lab.
Nagata-cho, Chiyoda-ku
Tokyo, Japan

A Tunnel-Diode Wide-Band Frequency Doubling Circuit*

A recent note[1] describes an elegant wide-band frequency doubling circuit which makes use of the fact that the volt-ampere characteristic of a tunnel diode closely approximates a parabola in the region of the peak point. The circuit as described is of limited usefulness because of sensitivity to variations in ambient temperature. The operating point of the tunnel diode is set by the base biasing network of the transistor (Fig. 1). However, the tunnel diode is in the emitter circuit so its bias will change if the emitter-base voltage of the transistor changes. For a silicon transistor V_{eb} has a temperature dependence of approximately -2 mv/°C. Since the peak point of the diode occurs at approximately 55 mv, this shifts the bias of the tunnel diode by -3 per cent/°C. With a signal of 25 mv peak, this shift is approximately 8 per cent of peak swing per °C. Therefore a 1°C temperature change will cause alternate peaks of the output wave to differ in amplitude by as much as 16 per cent (see Fig. 2).

By using forward biased silicon diodes in the bias network to compensate for the temperature dependence of V_{ce}, considerable reduction in temperature sensitivity can be achieved. The biasing diodes should, of course, be in close thermal contact with the transistor. This lowers the input impedance, but not to such an extent as to cause any difficulty in driving the circuit.

Further stabilization can be realized by biasing at the valley point of the tunnel diode. While the tunnel diode characteristic alone is not a good approximation of a parabola at this point, the transfer characteristic of the transistor-tunnel diode combination is (Fig. 3). The gain of the squaring portion of the circuit is lower, but since the quiescent current is lower by a larger factor, slightly more voltage gain is available from the circuit as a whole. An improvement of about a factor of six (=valley point voltage/peak point voltage) in temperature stability results.

The resulting circuit (Fig. 4), while still not suitable for wide temperature ranges, can be used in many applications where temperature changes are reasonably small and the bias can be checked and adjusted occasionally if necessary. It has been operated for eight hours at a time in a non air-conditioned laboratory without noticeable bias drift after a short warm-up period.

J. H. BURBO
Instrumentation Lab.
Mass. Inst. Tech.
Concord, Mass.

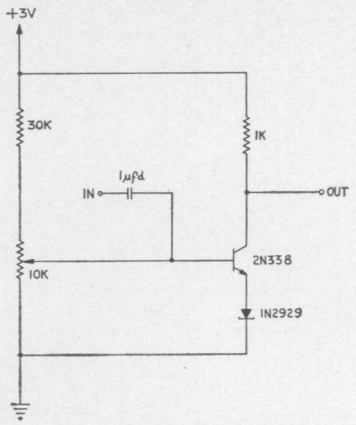


Fig. 1—Neu's frequency doubling circuit.

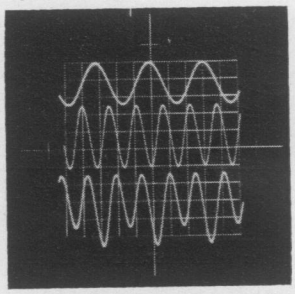


Fig. 2—Top: 800 cps input. Center: 1600 cps output, bias correct. Bottom: 1600 cps output, bias 5 mv offset. Vertical sensitivity: 50 mv/div.

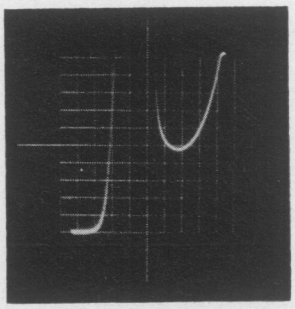


Fig. 3—Transfer characteristic of transistor-tunnel diode combination (peak point is off scale). Horizontal (input): 75 mv/div. Vertical (output): 200 mv/div.

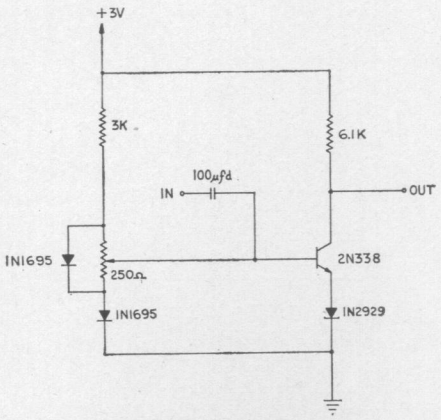


Fig. 4—Improved frequency doubler.

Graphical Expressions of Synchronous Conditions in the Transverse-Type Electron Beam Parametric Amplifier*

There are several types of pumping in the transverse-type electron beam parametric amplifier, each of which has a different synchronous condition. This condition is explained physically by deriving it from a geometrical sum of the signal and idler waves.

The case of amplification by the coupling of two fast cyclotron waves is shown in Fig. 1, where ω_s, ω_i and ω_p represent signal, idler and pump frequency, respectively. In both waves the electron rotates by $\omega_c t$ in time t while it proceeds axially at $v_0 t$, where ω_c is cyclotron frequency and v_0 is the axial mean velocity of electron. From a simple geometrical consideration the following phase velocities are obtained:

$$v_s = v_0/1 - \frac{\omega_c}{\omega_s}, \qquad v_i = v_0/1 - \frac{\omega_c}{\omega_i}$$

where the subscripts s and i refer to signal and idler. Summing up the displacement of these waves, a new wave pattern is obtained, where the relation $\omega_s+\omega_i=\omega_p$ holds at the initial plane and the rate of rotation observed from a moving electron becomes $2\omega_c t$. From the new pattern the following relation is given:

$$\omega_p\left(1 - \frac{v_0}{v_p}\right) = 2\omega_c$$

This is nothing but a synchronous condition, needed for the pump wave, so that the Doppler frequency is equal to two times cyclotron frequency.

Fig. 2 shows the case of amplification by the coupling of two synchronous waves in quadrupole amplifier. In this case the phase velocity of pump wave remains equal to the electron velocity. Then the synchronous condition is given by

$$\omega_p\left(1 - \frac{v_0}{v_p}\right) = 0,$$

showing that Doppler frequency equals zero.

Fig. 3 shows the case where amplification occurs by the coupling of the fast cyclotron and slow synchronous waves in the axially

* Received by the IRE, January 3, 1962.
[1] F. D. Neu, "A tunnel-diode wide-band frequency doubling circuit," PROC. IRE, vol. 49, pp. 1963–1964; December, 1961.

* Received January 3, 1962; revised manuscript received, February 23, 1962.

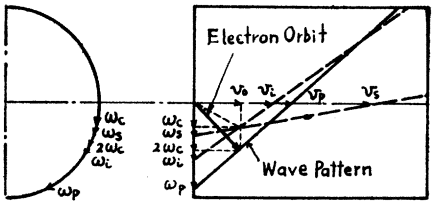


Fig. 1—Amplification by the coupling of fast cyclotron waves (quadrupole pump).

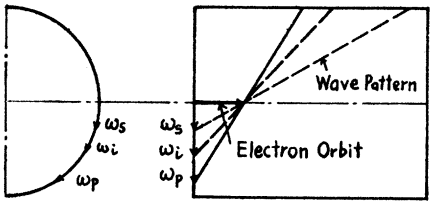


Fig. 2—Amplification by the coupling of synchronous waves (quadrupole pump).

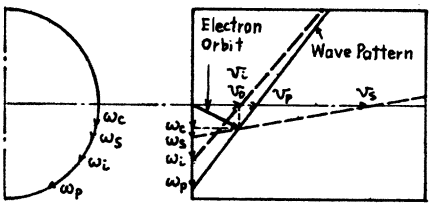


Fig. 3—Amplification by the coupling of fast and slow synchronous waves (axially symmetric field pump).

symmetric field pump. By means of a similar process the synchronous condition is found as follows:

$$\omega_p\left(1-\frac{v_0}{v_p}\right)=\omega_c.$$

Here, the Doppler frequency equals the cyclotron frequency.

In conclusion it may be stated that the number of $2\omega_c$, 0 and ω_c, included in the above equations, corresponds to the number of cyclotron waves related to the amplification process.

K. KAKIZAKI
Central Research Laboratory
Tokyo Shibaura Electric Co., Ltd.
Kawasaki, Japan

Symmetrical DC Converter Using 6-a Tunnel Diodes*

A possibly useful application of the tunnel diode is in a dc converter circuit to raise the output level of a low voltage source such as a thermoelectric generator. The symmetrical converter circuit shown in Fig. 1 is

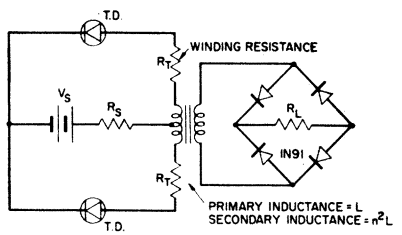


Fig. 1—Symmetrical tunnel-diode converter circuit.

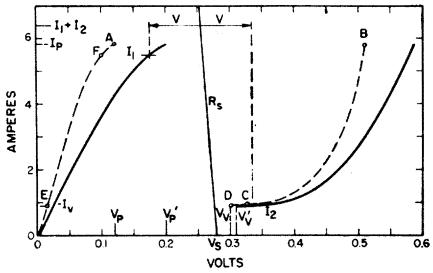


Fig. 2—Dashed line: $v-i$ characteristic of tunnel diode. Solid line: $v-i$ characteristic of tunnel diode in series with transformer primary resistance R_T.

particularly interesting because it allows a relatively steady current to be drawn from the source.[1] This results in a higher efficiency (ratio of output power to *available* power at the input) than is possible with a single diode circuit.[2,3]

The operation of this circuit has been analyzed and examined experimentally for the antisymmetric mode of oscillation. The voltage-current trajectory for one diode can be traced on Fig. 2. Starting at point A the diode switches *very* quickly (in the constant current switching time) to B and quickly (in the order of the time to charge the transformer leakage inductance) to C. It then relaxes *slowly* (in the order of L/R time constants) to D, very quickly to E, quickly to F, and slowly back to A. The locations of points C and F depend upon the loading. For no load F coincides with E and C with B; while under maximum loading (minimum R_L) F coincides with A and C with D.

Under maximum loading, therefore, the diodes effectively switch back and forth between their peak and valley points. The power output is

$$P_{\text{out}}=\tfrac{1}{2}(V_v'-V_p')(I_p-I_v) \qquad (1)$$

where $V_v'=V_v+R_TI_v$ and $V_p'=V_p+R_TI_p$ as indicated in Fig. 2. The efficiency is

$$\epsilon\approx\frac{4m}{(1+m)^2}\left(\frac{V_v'-V_p'}{V_v'+V_p'}\right)\left(\frac{I_p-I_v}{I_p+I_v}\right), \qquad (2)$$

where $m=R_c/R_s$ and $R_c=\frac{1}{2}(V_v'+V_p')/(I_p+I_v)$ is the resistance which the circuit presents to the source. The source resistance R_s must be less than $\frac{1}{2}(V_v'-V_p')/(I_p-I_v)$ to insure stable biasing of the circuit. Therefore

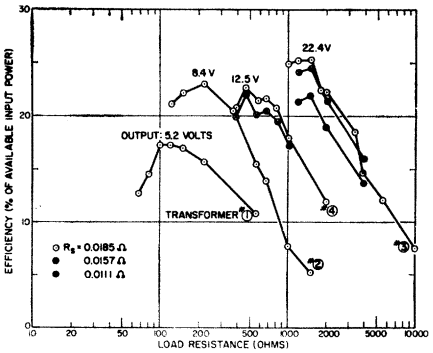


Fig. 3—Efficiency vs load resistance and source resistance for various transformers. Transformer #1—$R_T=0.125$ ohm, #2—0.004 ohm, #3—0.0018 ohm, #4—0.0042 ohms.

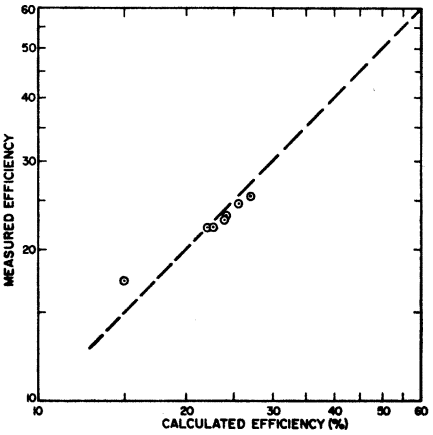


Fig. 4—Comparison of measured and calculated efficiencies.

$$m>\left(\frac{V_v'+V_p'}{V_v'-V_p'}\right)\left(\frac{I_p-I_v}{I_p+I_v}\right). \qquad (3)$$

Thus for maximum efficiency it is desirable to have V_v'/V_p' and I_p/I_v as large as possible and $V_v'/V_p'\geq I_p/I_v$.

Efficiency has been measured in a circuit using two "solution grown" germanium diodes with matched v-i characteristics as shown in Fig. 2.[4] Efficiency vs load resistance is plotted for several values of source resistance and for various transformers in Fig. 3. In Fig. 4 maximum efficiency is compared with that calculated from (2). Output power is between 300 and 400 mw.

If the load resistance is made less than

$$R_L<2n^2\frac{V_v'-V_p'}{I_p-I_v}, \qquad (4)$$

the diodes relax quickly from B to D and from E to A. The transformer flux does not have time to become equal in the windings and the antisymmetric mode is no longer favored by the winding polarity. A symmetric mode then obtains which is characterized by a much higher frequency and which does not couple to the load.

A. C. SCOTT
Dept. of Elec. Engrg.
Massachusetts Institute of Technology,
Cambridge, Mass.

* Received January 31, 1962. This note is based on a thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Science in the Department of Electrical Engineering at the Massachusetts Institute of Technology, Cambridge, on August 21, 1961.

[1] This circuit was suggested to the author by Dr. Paul Pittman of the Westinghouse Res. Labs., Pittsburgh, Pa.

[2] H. F. Storm and D. P. Shattuck, "Tunnel Diode D-C Power Converter," presented at AIEE Winter General Meeting, New York, N. Y.; January 29–February 3, 1961.

[3] S. Wang, "Converter efficiency and power output of a tunnel diode relaxation oscillator," PROC. IRE, vol. 49, pp. 1219–1220; July, 1961.

[4] The author is indebted to the staff of the Bell Telephone Labs., Murray Hill, N. J., for the opportunity to fabricate these diodes during the summer of 1960.

Magneto-Ionic Duct Propagation Time (Whistler-Mode) vs Geomagnetic Latitude at 4 KC*

When the transmitted frequency f is well below the plasma frequency f_p in ionospheric propagation, the extraordinary ray direction is approximately that of the earth's magnetic field lines. At increasing geomagnetic latitudes of origination, the propagation time from the originating point to the conjugate point should increase with the path length. Correspondingly, the group velocity is lessened in varying degree with height in the ionosphere, and for different lengths of time dependent upon the ray path length in the ionosphere. Thus, it seems reasonable to assume that on a graph of the propagation time vs geomagnetic origination latitude, some portion of the curve may be quite flat or perhaps have a negative slope. A simplified analysis follows to show that this situation is quite possible, if the ray follows the field line closely.

The assumptions employed are listed below:

1) Energy travels in the direction of the earth's magnetic field.
2) Circular polarization with the right-hand sense for propagation from south to north is assumed (these first assumptions are completely equivalent to assuming a longitudinal extraordinary mode.
3) The earth's magnetic field is represented by an earth-centered magnetic dipole.
4) The collisional frequency μ is zero.

The group velocity v_g is given by[1]

$$v_g = \frac{c}{n + f dn/df} = \frac{2c(f_H - f)^{3/2}[f^2(f_H - f) + f f_p^2]^{1/2}}{2f^3 - 4f^2 f_H + 2f f_H^2 + f_p^2 f_H}$$

where

$$n^2 = 1 + \frac{f_p^2}{f(f_H - f)},$$

f_p = plasma frequency = $9\sqrt{N} \times 10^3$ c/s,

f_H = gyro frequency = $2.8 \times 10^5 B$ c/s,

n = the refractive index,

N = the electron concentration electrons/cm^3,

B = the magnetic field flux density, gauss,

f = wave frequency, cps.

The conventional first- and second-order approximations are given by

$$v_g \doteq 2c\frac{f^{1/2}(f_H - f)^{3/2}}{f_H f_p}, \qquad f_p^2 \gg f f_H$$

$$v_g \doteq \frac{2c(f f_H)^{1/2}}{f_p}, \qquad f_p^2 \gg f f_H,\ f \ll f_H.$$

The propagation time T is to be evaluated graphically from

$$T = \int_{\text{path}} \frac{ds}{v_g} \triangleq \frac{D}{\sqrt{f}}$$

which defines D, the dispersion. The magnetic path length s is

$$s = \frac{R_0}{2}\left\{(1 + 3\sin^2\lambda)^{1/2}\sin\lambda + \frac{1}{\sqrt{3}}\ln\left[\sin\lambda + \frac{(1 + 3\sin^2\lambda)^{1/2}}{\sqrt{3}}\right]\right\}_{\lambda_1}^{\lambda_0}$$

where

λ_0 = the originating latitude, degrees,

λ_1 = the latitude to which path length is desired,

R_0 = the maximum radius of the magnetic dipole path, in units of earth radii, $\sec^2\lambda_0$.

The height h above the earth's surface is

$$h = (R_0\cos\lambda - 1)a$$

where a is the earth's radius. Assuming symmetry of the dipole field above and below the geomagnetic equator, only one-half the path need be considered for integration purposes.

Due to evident approximations in the foregoing, computations were run by hand at a single frequency of 4 kc. The results are given in Fig. 1 for a one-way path. In Table I the one-way dispersion as a function of originating latitude is given.

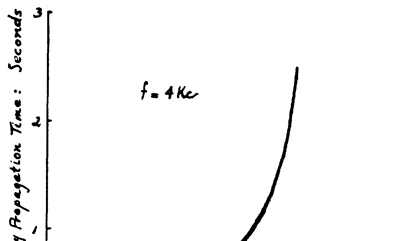


Fig. 1—Propagation time vs originating geomagnetic latitude.

TABLE I
DISPERSION VS ORIGINATING LATITUDE

λ_0 (degrees)	$D = Tf^{1/2}$ (seconds)$^{1/2}$
10	24.5
20	46.3
30	32.2
40	32.2
50	47.7
60	159.8

Granted that the foregoing is an approximate analysis, the phenomenon demonstrated by Fig. 1 is very interesting. A more elaborate investigation is certainly warranted at other frequencies and for other parameter models. The models used in this analysis were those of a dipole magnetic field for the earth and an electron density profile compiled from Penn State data, NRL rocket data, USSR data (rocket, February 21, 1958), USSR satellite data, and "whistler" estimates.

CHARLES V. GREENMAN
RCA
Camden, N. J.

* Received January 29, 1962.

[1] G. R. Ellis, "On the propagation of whistling atmospherics," *J. Atmospheric and Terrestrial Phys.*, vol. 8, pp. 338–344; June, 1956.

A Two-Term Analytical Approximation of Tunnel-Diode Static Characteristics*

In the articles previously published on tunnel-diode circuitry design, either a polynomial expression[1] or sectional straight-line approximations are used to represent the tunnel-diode static current-voltage characteristics. This paper will present a simple two-term exponential approximation. The accuracy of this approximation compared with actual tunnel-diode characteristics will also be discussed.

The static I-V characteristic of a tunnel

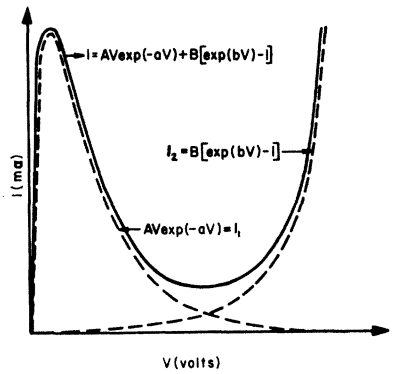


Fig. 1—Tunnel-diode static I-V characteristics and the two current components.

diode is shown as the solid curve in Fig. 1. It is found that a plot of the natural logarithm of (I/V) vs the applied voltage gives a straight line in the negative resistance region. This suggests that the total tunnel-diode current may be expressed as[2]

$$I = I_1 + I_2 = AV\exp(-aV) + B[\exp(bV) - 1]. \quad (1)$$

The first term is the tunnel current and the second term is the ordinary diode current. They are shown as the dotted curves in Fig. 1. The constants a, A, b, and B are determined as follows:

For voltages less than the valley voltage, the second term of (1) can be neglected giving

$$I_1 = AV\exp(-aV) \quad (2)$$

which can be written as

$$\ln(I_1/V) = \ln A - aV. \quad (3)$$

In the straight line plot of $\ln(I_1/V)$ vs the applied voltage, the slope is equal to $(-a)$, and the zero voltage intercept is equal to $\ln A$, as shown in Fig. 2.

For large values of forward bias, the first term of (1) can be neglected and since $\exp(bV) \gg 1$, (1) reduces to

$$I_2 \approx B\exp(bV), \quad (4)$$

* Received January 31, 1962; revised manuscript received, February 15, 1962.

[1] M. Schuller and W. W. Gartner, "Large-signal circuit theory for negative-resistance diodes, in particular tunnel diodes," PROC. IRE, vol. 49, pp. 1268–1278; August, 1961.

[2] A. Ferendeci, "A Study of Tunnel Diode Characteristics," M.S. thesis, Case Institute of Technology, Cleveland, Ohio; 1961.

which can be written as

$$\ln I_2 = \ln B + bV. \quad (5)$$

The plot of $\ln I_2$ vs the applied voltage should give a straight line with a slope equal to b and the zero voltage intercept equal to $\ln B$, as is shown in Fig. 2.

If (2) and (4) are plotted, then the sum of the two curves should approximate the actual tunnel-diode curve. This approximation has been applied to germanium, silicon and Ga-As tunnel diodes with results better than ±10 per cent accuracy over the complete curve. A typical germanium tunnel diode (1N-2941), is used to illustrate the results of this two-term approximation as shown in Fig. 3. The actual diode curves are obtained from a photograph taken from a curve tracer oscilloscope. From the graphs of $\ln (I/V)$ vs V and $\ln I$ vs V, the constants a, b, A, and B are calculated giving the approximation as:

$$I = 0.044V \exp(-16.8V) + 5.4 \times 10^{-7}[\exp(15.4V) - 1] \text{(amperes)}. \quad (6)$$

Using (6) the currents for corresponding voltages are calculated and the results are plotted on the same graph paper with the actual tunnel-diode curves for comparison. The current difference between the actual and the calculated curves is normalized with respect to the peak current and plotted on the lower portion of the graphs. It is seen that the approximation is accurate to ±6 per cent or better.

For engineering purposes the four constants may be determined by measuring four sets of current and voltage values at four pilot points to be fitted accurately. Two of the pilot points should be in the tunnel current region and the other points in the ordinary diode current region. By substituting these I-V values into (3) or (5), the constants may be found. The peak point corresponds to the first $dI/dV = 0$ point; therefore, if the peak point is selected to be one of the pilot points, then the constant $a = (1/V_p)$.

It is not surprising to find that for the simple approximations, the constants a and b are not multiples of (e/kT), especially the ordinary diode constant b. This is possibly due to the valley current which effects the ordinary diode current considerably within the regions of interest.

A. Ferendeci
W. H. Ko
Engrg. Div.
Case Inst. Tech.
Cleveland, Ohio

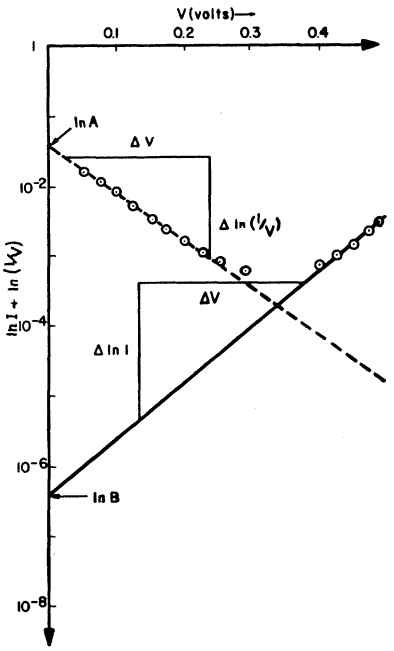


Fig. 2—ln (I/V) vs V and ln (I) vs V curves.

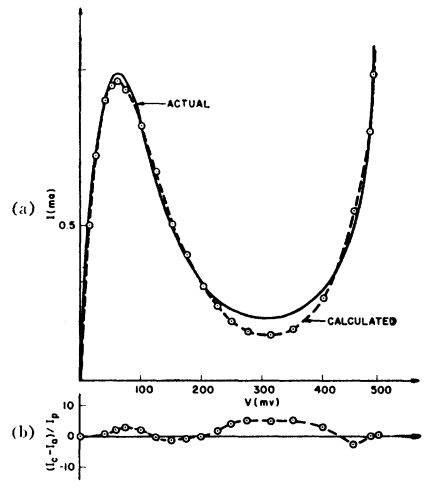


Fig. 3—1N-2941 Germanium tunnel diode. (a) Actual and calculated characteristics. (b) Per cent error between actual and calculated characteristics.

Nonlinear Distortion in Tunnel-Diode Amplifiers*

An important performance characteristic of tunnel-diode amplifiers is dynamic range. The upper limit imposed upon the dynamic range of a tunnel-diode amplifier is usually due to nonlinear distortion. In this paper, an approximate theoretical analysis of the nonlinear distortion within tunnel-diode amplifiers will be presented.

The current-vs-voltage characteristic of a typical tunnel diode is shown in Fig. 1. In the negative resistance region this i-e curve can be approximated by an exponential function:

$$i = Ae^{ae} \quad (1)$$

where

i = tunnel-diode current,
e = tunnel-diode voltage,
A and a are constants.

Letting $X = e - e_b$ = tunnel-diode ac voltage, where e_b = voltage at operating point of the tunnel-diode amplifier

$$i = Ae^{a(x+e_b)} \cong Ae^{ae_b}\left[1 + ax + \frac{a^2}{2!}x^2 + \frac{a^3}{3!}x^2 + \cdots\right]. \quad (2)$$

At the operating point

$X = 0$ and $e = e_b$, then

$$a = \frac{\left.\frac{di}{de}\right|_{e=e_b}}{i}, \quad (3)$$

$$A = ie_b^{-ae_b} \quad (4)$$

Constants a and A can be evaluated using (3) and (4) for an operating point within the negative resistance region of a particular tunnel diode.

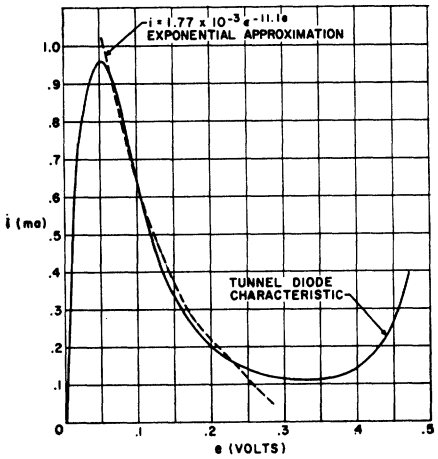


Fig. 1—Typical current-vs-voltage characteristic of tunnel diode and exponential approximation in the negative resistance region.

To determine the second- and third-order harmonic and intermodulation distortion the first four terms of (2) are employed:

$$i = Ae^{ae_b}\left[1 + ax + \frac{a^2}{2}x^2 + \frac{a^3}{6}x^3\right]. \quad (5)$$

Assuming two equilevel signals of peak voltage E and angular frequencies ω_1 and ω_2,

$$x = E\cos\omega_1 t + E\cos\omega_2 t. \quad (6)$$

Substituting (6) into (5) and expanding, it can be shown that

$$\begin{aligned} i = Ae^{ae_b}\Big\{&\left(1 + \frac{a^2E^2}{2}\right) \\ &+ (aE + \tfrac{3}{8}a^3E^3)(\cos\omega_1 t + \cos\omega_2 t) \\ &+ \frac{a^2E^2}{4}(\cos 2\omega_1 t + \cos 2\omega_2 t) \\ &+ \frac{a^2E^2}{2}\cos(\omega_1 + \omega_2)t + \frac{a^2E^2}{2}\cos(\omega_1 - \omega_2) \\ &+ \frac{a^3E^3}{24}(\cos 3\omega_1 t + \cos 3\omega_2 t) \\ &+ \frac{a^3E^3}{8}[\cos(2\omega_1 + \omega_2)t + \cos(2\omega_2 + \omega_1)t \\ &+ \cos(2\omega_1 - \omega_2)t + \cos(2\omega_2 - \omega_1)t]\Big\}. \quad (7) \end{aligned}$$

* Received January 26, 1962.

Then

$$\frac{\text{Signal}}{\text{2nd Harmonic}} = \frac{4(1+\frac{3}{8}a^2E^2)}{aE} \quad (8)$$

$$\frac{\text{Signal}}{\text{2nd Order Intermodulation}} = \frac{2(1+\frac{3}{8}a^2E^2)}{aE} \quad (9)$$

$$\frac{\text{Signal}}{\text{3rd Harmonic}} = \frac{24(1+\frac{3}{8}a^2E^2)}{a^2E^2} \quad (10)$$

$$\frac{\text{Signal}}{\text{3rd Order Intermodulation}} = \frac{8(1+\frac{3}{8}a^2E^2)}{a^2E^2}. \quad (11)$$

The ratios of signal-to-harmonic distortion and signal-to-intermodulation distortion are given in (8)–(11). Although these ratios appear to be independent of bias voltage, they are functions of the constant a which is evaluated for a particular value of bias voltage. At different operating points, a different value of a would generally yield a better fit of the exponential approximating function.

This theoretical analysis is of limited accuracy. Higher-order curvatures have been neglected and possible mismatches seen by the tunnel diode at the distortion frequencies have not been considered. The exponential approximation obviously breaks down for ac voltage magnitudes that are large enough to swing into the tunnel-diode positive resistance region. Nevertheless the exponential approximating function provides simple equations that can be used to obtain "ballpark" estimates of the various signal-to-distortion ratios.

Because of the different sources of error previously mentioned, it is seldom profitable to employ the more exact representation of the tunnel-diode characteristics and experimental techniques are more expedient.

R. M. Kurzrok
A. Newton
RCA Surface Commun. Sys. Lab.
New York, N. Y.

Capacitance and Charge Coefficients for Varactor Diodes*

It has been shown desirable to know the Fourier coefficients of the charge and capacity of a varactor diode driven by a single frequency voltage.[1,2] These coefficients have been previously calculated either in approximate series form or in terms of untabulated functions. It is possible to express them in an elegant and simple functional form.

In the back-biased diode, the capacity as a function of voltage is taken to be

$$C = C_0\left(1 - \frac{V}{\phi_0}\right)^{\nu}. \quad (1)$$

We impose

$$V = V_0 + 2V_1 \cos \omega t. \quad (2)$$

Hence

$$C = C_0\left\{\left(1 - \frac{V_0}{\phi_0}\right) - \frac{2V_1}{\phi_0}\cos \omega t\right\}^{\nu} \quad (3)$$

$$C = C_d\left\{1 - \frac{2V_1}{\phi_0 - V_0}\cos \omega t\right\}^{\nu}, \quad (4)$$

where

$$C_d = C_0\left(1 - \frac{V_0}{\phi_0}\right)^{\nu}, \quad (5)$$

C_d is the capacity of the varactor at voltage V_0.

We want the complex Fourier coefficient

$$C_n = \frac{1}{\pi}\int_0^{\pi} C_d\left\{1 - \frac{2V_1}{\phi_0 - V_0}\cos \omega t\right\}^{\nu} \cdot \cos n\omega t d(\omega t) \quad (6)$$

$$C_n = C_d\frac{(-1)^n}{\pi}\int_0^{\pi}(1 + \alpha \cos \theta)^{\nu} \cos n\theta d\theta, \quad (7)$$

where we have set

$$\theta = (\omega t + \pi) \quad (8)$$

and

$$\alpha = \frac{2V_1}{\phi_0 - V_0}. \quad (9)$$

The integral (7) may now be related to a known integral, known as the Laplace integral.

From Erdelyi:[3]

$$\frac{1}{\pi}\int_0^{\pi}(z + \sqrt{z^2 - 1}\cos \theta)^{\nu} \cos n\theta \, d\theta = \frac{\Gamma(\nu + 1)}{\Gamma(\nu + n + 1)} P_{\nu}^{n}(z), \quad (10)$$

where $P_{\nu}^{n}(z)$ is the associated Legendre function.

Multiplying and dividing by z^{ν}, we have

$$\frac{1}{\pi}\int_0^{\pi}\left(1 + \frac{\sqrt{z^2 - 1}}{z}\cos \theta\right)^{\nu} z^{\nu} \cos n\theta d\theta = \frac{z^{\nu}}{\pi}\int_0^{\pi}(1 + \alpha \cos \theta)^{\nu} \cos n\theta d\theta, \quad (11)$$

with

$$\alpha = \frac{\sqrt{z^2 - 1}}{z} \quad \text{or} \quad z = \frac{1}{\sqrt{1 - \alpha^2}}. \quad (12)$$

Thus

$$C_n = C_d\frac{(-1)^n}{\pi}\int_0^{\pi}(1 + \alpha \cos \theta)^{\nu} \cos n\theta d\theta = C_d(-1)^n\frac{\Gamma(\nu + 1)}{\Gamma(\nu + n + 1)} z^{-\nu} P_{\nu}^{n}(z). \quad (13)$$

$P_{\nu}^{n}(z)$ is tabulated[4] for non-integral values of ν, and we may evaluate any Fourier coefficient of capacity simply, and with no approximations at any drive level.

The same procedure is used in the calculation of the Fourier coefficients of the charge Q. We find Q as a function of V by integrating (1)

$$Q = \frac{-C_0\phi_0}{\nu + 1}\left(1 - \frac{V}{\phi_0}\right)^{\nu+1}. \quad (14)$$

The Fourier integral is of the same form as before

$$Q_n = -Q_d\frac{(-1)^n}{\pi}\int_0^{\pi}(1 + \alpha \cos \theta)^{\nu+1} \cdot \cos n\theta d\theta, \quad (15)$$

where

$$Q_d = \frac{C_0\phi_0}{(\nu + 1)}\left(1 - \frac{V_0}{\phi_0}\right)^{\nu+1} = \frac{C_d(\phi_0 - V_0)}{(\nu + 1)}. \quad (16)$$

Using (10) we have

$$Q_n = Q_d(-1)^{n+1}\frac{\Gamma(\nu+2)}{\Gamma(\nu+n+2)} z^{-(\nu+1)} P_{\nu+1}^{n}(z) \quad (17)$$

$$C_n = C_d(-1)^n\frac{\Gamma(\nu + 1)}{\Gamma(\nu + n + 1)} z^{-\nu} P_{\nu}^{n}(z), \quad (13)$$

where

$$z = \frac{1}{\sqrt{1 - \alpha^2}}. \quad (12)$$

Eqs. (12) and (13) are repeated here to summarize the results.

D. B. Leeson
Dept. Elec. Engrg.
Stanford Univ.
Stanford, Calif.

Invariant Stability Parameters*

In a recent paper,[1] it is shown that there exists an invariant stability factor k defined by

$$k = \frac{2\rho_{11}\rho_{22} - \text{Re}(\gamma_{12}\gamma_{21})}{|\gamma_{12}\gamma_{21}|} \quad (1)$$

where the γ may be any of the conventional z, y, h, g twoport matrix parameters, and $\rho_{11} = \text{Re}(\gamma_{11})$, etc. The quantity k is invariant under arbitrary lossless terminations, under interchange of input and output, and under "immittance substitution," a transformation group involving the arbitrary interchanging of impedance and admittance formulations at both ports (replacing any set of the z, y, h, g parameters by any other). These transformations are associated with stability,[1] and provided $\rho_{11}, \rho_{22} \geq 0$, k is a unique measure of the degree of conditional ($-1 \leq k < 1$) or unconditional ($k > 1$) stability of the twoport. This means that k is a measure of what would happen in the worst possible case (as regards stability) of arbitrary passive terminations.

There is, however, a need for a stability parameter which will indicate the degree of

* Received January 22, 1962.

[1] S. Sensiper, and R. D. Weglein, "Capacitance and charge coefficients for parametric diode devices," Proc. IRE, vol. 48, pp. 1482–1483; August, 1960.

[2] ——, "Capacitance coefficients for varactor diodes," Proc. IRE (*Correspondence*), vol. 49, p. 810; April, 1961.

[3] A. Erdelyi, "Note on Heine's integral representation of associated Legendre functions," *Philos. Mag.*, ser. 7, vol. 32, pp. 351–352; October, 1941.

[4] A. Lowan, "Tables of Associated Legendre Functions," Natl. Bur. Standards, Columbia Univ. Press, New York, N. Y., pp. 230 ff.; 1945.

* Received January 23, 1962.

[1] J. M. Rollett, "Stability and power gain invariants of linear two-ports," IRE Trans. on Circuit Theory, vol. CT-9, pp. 29–32; March, 1962.

stability of a twoport in a particular situation of interest, and not only in an "idealized" worst possible case. Such a parameter (if it exists) must take account of source (Γ_1) and load (Γ_2) immittances, and will be invariant under interchange of input and output (including terminations) and under immittance substitution.

Instability occurs when the total immittance at either port is zero, *i.e.*, when

$$\Gamma_1 + \Gamma_{in} = \frac{(\Gamma_1 + \gamma_{11})(\Gamma_2 + \gamma_{22}) - \gamma_{12}\gamma_{21}}{(\Gamma_2 + \gamma_{22})} = 0 \quad (2)$$

and similarly for $(\Gamma_2 + \Gamma_{out})$. The zeros of this expression[2] are the zeros of $(\Gamma_1 + \gamma_{11})(\Gamma_2 + \gamma_{22}) - \gamma_{12}\gamma_{21}$, provided that the source immittance is passive (and similarly for the load immittance), and provided that the characteristic frequencies of the twoport with infinite immittances are left half-plane. Consequently,

$$(\Gamma_1 + \gamma_{11})(\Gamma_2 + \gamma_{22}) - \gamma_{12}\gamma_{21} = 0 \quad (3)$$

is called the characteristic equation[3,4] of the system.

Now the characteristic function of (3) is invariant under interchange of input and output, but not under immittance substitution. A search for suitably invariant functions of the characteristic function has unearthed a quantity

$$\chi = \frac{(\Gamma_1 + \gamma_{11})(\Gamma_2 + \gamma_{22}) - \gamma_{12}\gamma_{21}}{\sqrt{(\Gamma_1\Gamma_2\gamma_{12}\gamma_{21})}} \quad (4)$$

which is "semi"-invariant under immittance substitution, *i.e.*, invariant except for an ambiguity in phase of $m\pi/2$ (m integral). This ambiguity may be removed by forming such functions as $|\chi|$, χ^4 or $\chi \exp jm\pi/2$, etc.

The properties of χ may be briefly summarized.

1) χ is invariant under interchange of input and output and semi-invariant under immittance substitution.
2) The zeros of χ are the zeros of the characteristic function of (3), with the provisos mentioned above.
3) The generalized signal gain (*i.e.*, ratio of an output current/voltage to a signal current/voltage) is given by

$$A = \frac{(-1)^n}{\chi}\sqrt{\left(\frac{\gamma_{21}}{\gamma_{12}}\right)}\sqrt{\left(\frac{\Gamma_2'}{\Gamma_1'}\right)} \quad (5)$$

where the immittance representations of Γ_1', Γ_2' are chosen (independently of the choice of matrix parameters γ) according to the particular ratio required, *i.e.*, impedance for voltage, admittance for current; n is taken as 0 if Γ_2, Γ_2' are equal, and 1 if they are reciprocal, to preserve conventional phase relations.

Thus, for example, voltage gain is given by

$$A_v = \frac{v_2}{v_g} = \frac{(-1)^n}{\chi}\sqrt{\left(\frac{\gamma_{21}}{\gamma_{12}}\right)}\sqrt{\left(\frac{Z_2}{Z_1}\right)} \quad (6)$$

in general, or in particular

$$A_v = \frac{-y_{21}Z_2}{(Y_1 + y_{11})(Y_2 + y_{22}) - y_{12}y_{21}}; \quad (7)$$

while transadmittance is given by

$$Y_m = \frac{i_2}{v_g} = \frac{(-1)^n}{\chi}\sqrt{\left(\frac{\gamma_{21}}{\gamma_{12}}\right)}\sqrt{\left(\frac{Y_2}{Z_1}\right)}. \quad (8)$$

The quantity[5] $\sqrt{(\gamma_{21}/\gamma_{12})}$ is also semi-invariant under immittance substitution.

4) The transducer gain G_T is given by

$$G_T = \frac{4}{|\chi|^2}\cdot\left|\frac{\gamma_{21}}{\gamma_{12}}\right|\cdot\frac{P_1P_2}{|\Gamma_1\Gamma_2|} \quad (9)$$

where $P_1 = \mathrm{Re}\,(\Gamma_1)$, etc.; $P_1P_2/|\Gamma_1\Gamma_2|$ is invariant under immittance substitution.

The usefulness of χ lies in the fact that it is an invariant measure of stability, or in effect, an invariant return difference.[4,6] The size and shape of the plot of χ as a function of frequency in the complex plane is unique, apart from rotations of $m\pi/2$. Thus the shape of the locus can be examined, as in the Nyquist test, to provide information about the zeros;[3,6] while the distance of χ from the critical point, the origin, is proportional to the stability of the system at real frequencies. In simple cases, a knowledge of the magnitude of χ provides sufficient information for the circuit designer, in the form of a plot of $|\chi|$ or $|\chi|^2$ against frequency.

The stability parameter χ or its magnitude may be measured by making use of the properties given in paragraphs 3) and 4) above. Thus

5) If the reverse generalized signal gain, found by interchanging input and output, is denoted by A^r then

$$\chi = 1/\sqrt{(A \cdot A^r)} \text{ or } |\chi|^2 = 1/|A \cdot A^r|. \quad (10)$$

6) If the reverse transducer gain is denoted by $G_T{}^r$ then

$$|\chi|^2 = \frac{4}{\sqrt{(G_T \cdot G_T{}^r)}}\cdot\frac{P_1P_2}{|\Gamma_1\Gamma_2|}. \quad (11)$$

Since $P_1P_2/|\Gamma_1\Gamma_2|$ is invariant, there is no need to include it as a factor. If power measurements are made, it is more convenient to plot a reduced stability parameter $|\chi'|^2$ where[7]

$$|\chi'|^2 = \frac{1}{\sqrt{(G_T \cdot G_T{}^r)}} = \frac{|(\Gamma_1 + \gamma_{11})(\Gamma_2 + \gamma_{22}) - \gamma_{12}\gamma_{21}|^2}{4P_1P_2|\gamma_{12}\gamma_{21}|}. \quad (12)$$

The product of the transducer gain G_T and the stability parameter $|\chi'|^2$ is the invariant maximum stable power gain,[5,1] $|\gamma_{21}/\gamma_{12}|$, of the twoport. Thus $1/|\chi'|^2$ is the efficiency of the reciprocal part of the twoport, so that $|\chi'|^2$ may be called the "reciprocal attenuation";[1] the transducer gain is then given by dividing the maximum stable power gain by the reciprocal attenuation. By analogy, χ may be called the (complex) reciprocal signal attenuation. It is the main suggestion of this letter that these invariant reciprocal attenuation parameters are closely connected with stability.

J. M. Rollett
British Telecommun. Res., Ltd.
Taplow Court, Taplow
Nr. Maidenhead, Berks.
England

Demodulators*

In the detection of a signal ensemble which is a sample of white Gaussian noise, an approximation to the optimum receiver consists of a full-wave square law demodulator followed by an integrator, or summer.[1,2] If a full-wave linear demodulator is used in place of the "optimum" demodulator, the degree of suboptimality is of interest.

Using the notation of Peterson, Birdsall, and Fox,[1] the parameter d is seen to be a convenient measure of the reliability of detection.

$$d = \frac{2E}{N_0} = \frac{(M_{SN} - M_N)^2}{\sigma_N^2}. \quad (1)$$

where

E = signal energy in a record of length T
N_0 = noise spectral density
M_{SN} = mean value at demodulator output (signal plus noise)
M_N = mean value of demodulator output (noise alone)
σ_N^2 = variance of noise.
σ_{SN}^2 = variance of signal plus noise.

The performance of the receiver, regardless of the decision criteria used, is conveniently represented by the Receiver Operating Characteristic (ROC) curves; the parameter d serves to identify individual members of such a family of curves. A convenient meth-

[2] The condition that the real part of the total port immittance be positive with arbitrary lossless terminations is that the over-all stability factor of Rollett[1] be greater than unity.

[3] F. B. Llewellyn, "Some fundamental properties of transmission systems," Proc. IRE, vol. 40, pp. 271–283; March, 1952.

[4] J. G. Linvill and J. F. Gibbons, "Transistors and Active Circuits," McGraw-Hill Book Co., Inc., New York, N. Y.; 1961.

[5] C. G. Aurell, "Representation of the general linear four-terminal network and some of its properties," *Ericsson Technics*, vol. 11, pp. 155–179; 1955.

[6] H. W. Bode, "Network Analysis and Feedback Amplifier Design," D. Van Nostrand Co., Inc., Princeton, N. J., pp. 151, 166; 1945.

[7] L. G. Gripps and J. A. G. Slatter, "Amplifier gain and stability," *J. Brit. IRE*, vol. 22, p. 417; November, 1961.

[8] M. A. Karp, "Power gain and stability," IRE Trans. on Circuit Theory (*Correspondence*), vol. CT-4, pp. 339–340; December, 1957.

* Received January 22, 1962.

[1] W. W. Peterson, T. G. Birdsall, and W. C. Fox, "The theory of signal detectability," IRE Trans. on Information Theory, vol. IT-4, pp. 171–212; September, 1954.

[2] J. J. Bussgang and W. L. Mudgett, "A note of caution on square-law approximation to an optimum detector," IRE Trans. on Information Theory (*Correspondence*), vol. IT-10, p. 504; September, 1960.

od of computing the effect of the suboptimum (linear) demodulator is to form the ratio

$$\mu = \frac{d_{\text{square-law}}}{d_{\text{linear}}} = \frac{\left[\frac{\sigma_{SN}^2 - \sigma_N^2}{\sqrt{2}\sigma_N^2}\right]^2}{\left[\frac{\sqrt{\frac{2}{\pi}}(\sigma_{SN} - \sigma_N)}{\sqrt{1 - \frac{2}{\pi}}\sigma_N}\right]^2}$$

$$= \frac{(\pi - 2)}{4}\left[1 + 2\left(\frac{\sigma_{SN}}{\sigma_N}\right) + \frac{\sigma_{SN}^2}{\sigma_N^2}\right]. \quad (2)$$

In the small signal case $\sigma_N \simeq \sigma_{SN}$, and $\mu = 1.14$. Since E is proportional to the length of record T, the factor μ indicates that, for equivalent receiver performance, the record length T must be increased by 14 per cent when the suboptimum linear demodulator is used.

J. KNUDSON
Western Dev. Labs.
Philco Corp.
Palo Alto, Calif.

The Dielectric Constant of a Semiconductor as Related to the Intrinsic Activation Energy*

A general equation which shows the relationship, for a semiconductor, between impurity activation energy, dielectric constant and the effective mass is shown as

$$E = \frac{E_h}{\epsilon_s^2}\left(\frac{m_\phi}{m_0}\right) \quad (1)$$

Here E is the activation energy of the impurity, m_ϕ is the effective mass, and ϵ_s is the dielectric constant of the semiconductor. E_h is the first ionization potential of hydrogen, and m_0 is the rest mass of an electron. Eq. (1) gives the basic relationship between the effective mass and the dielectric constant.

The energy gap or intrinsic activation energy is related to the square of the optical dielectric constant. This concept was first given by Moss in 1952.[1] The equation showing the connection between the two parameters is given by:

$$E_g\epsilon^2 = \text{constant} \quad (2)$$

where E_g is the intrinsic activation energy between the valence band and the conduction band. ϵ is the optical dielectric constant.

A relationship between the impurity activation energy and the carrier effective mass may be found from the theory of Brillouin zones of one dimension. If we consider the effective mass as determined by one dimensional K-space, Brillouin zone theory gives an equation of the form that is shown in the following:[2]

$$m_\phi = \frac{h^2}{4\pi^2}\left(1 \Big/ \frac{d^2E}{dK^2}\right). \quad (3)$$

Substituting the equation for the effective mass (1) and rearranging terms to form a differential equation, we have for this new expression

$$\frac{d^2E}{dK^2} = \frac{E_h h^2}{4\pi^2\epsilon^2 m_0 E}. \quad (4)$$

Multiplying both sides of (4) by $2dE$ and integrating will give

$$\left(\frac{dE}{dK}\right)^2 = \frac{E_h h^2}{2\pi^2\epsilon^2 m_0}\ln E. \quad (5)$$

This equation is equivalent to that obtained by Moss from his investigation of photoconductivity. The approximate value of (ln E) is found by plotting a curve of the ratio E_g/E_h vs $1/\epsilon^2$. Here (ln E) is found from the curve made up of the parameters for several semiconductors. This curve and the value of its slope, ln E, are shown in Fig. 1. The slope of this curve was found to be equal to a value of 11.7. Therefore, the product of the first ionization potential of hydrogen and the slope of the curve (11.7) is equal to the constant given in (2). The approximate value of this constant was found to be 159 ev. The energy gap or intrinsic activation energy is approximated by

$$E_g = \frac{159 \text{ ev}}{\epsilon^2}. \quad (10)$$

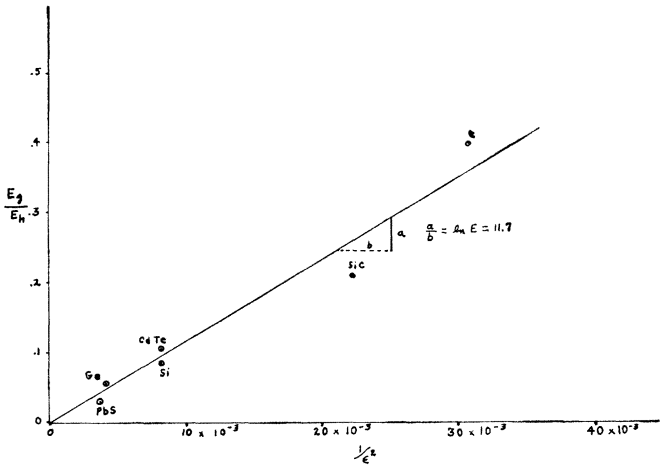


Fig. 1.

This equation relates the energy of an electron or hole to the wave number. An electron's wave energy in one dimensional K-space is shown by

$$E = \frac{h^2K^2}{8\pi^2 m_\phi}. \quad (6)$$

The derivative of this equation for energy with respect to the wave number, K, results in

$$\frac{dE}{dK} = \frac{h^2K}{4\pi^2 m_\phi}. \quad (7)$$

If this derivative is substituted into (5), then the energy relationship becomes:

$$\frac{h^2K^2m_0}{8\pi^2 m_\phi^2} = \frac{E_h}{\epsilon^2}\ln E. \quad (8)$$

Here $K = 1, 2, 3, 4, \cdots, n$.

The left-hand side of (8) will represent the intrinsic activation energy for the allowed wave number K. Eq. (8) now becomes

$$E_g = \frac{E_h}{\epsilon^2}\ln E. \quad T = \text{constant in °Kelvin.} \quad (9)$$

Moss has indicated that indium antimonide and indium arsenide would not follow the relationship of (2). Calculated values of the dielectric constant ϵ_c, as given in Table I, show this to be correct for the above equation. Table I has values of measured energy gap or intrinsic activation energy E_g and dielectric constant ϵ_m taken from Moss.[1] ϵ_c is the calculated value of the dielectric constant from (10).

TABLE I

			$T = 300°K$
Element	E_g	E_m	E_c
C (diamond)	5.40 ev	5.7	5.4
Si	1.12 ev	11.7	11.8
Ge	0.65 ev	16.0	15.6
Sn (Grey)	0.09 ev	?	42.0
PbS	0.40 ev	16.8	19.9
CdTe	1.43 ev	11.0	10.5
SiC	2.80 ev	6.7	7.5
PbTe	0.32 ev	28.6	22.3
PbSe	0.26 ev	21.0	24.7
CdS	1.30 ev	11.6	11.1
InSb	0.18 ev	15.7	29.7
InAs	0.33 ev	12.2	21.9
InP	1.26 ev	9.6	11.0
GaAs	1.35 ev	11.5	10.9
GaSb	0.70 ev	14.5	15.0
GaP	2.24 ev	8.4	8.4
AlSb	1.60 ev	10.2	9.9
AlAs	2.16 ev	?	8.6

C. F. COLE, JR.
Box 233
Clear, Alaska

* Received January 29, 1962; revised manuscript received, February 13, 1962.

[1] T. S. Moss, "The Optical Properties of Semiconductors," Academic Press Inc., New York, N. Y., pp. 48–49; 1959.

[2] C. Kittel, "Introduction to Solid State Physics," John Wiley and Sons, Inc., New York, N. Y., 2nd ed., p. 289; 1956.