PHYSICAL LIMITATIONS ON FREQUENCY AND POWER PARAMETERS OF TRANSISTORS

By

E. O. Johnson

RCA Electronic Components and Devices,
Somerville, N. J.

Summary—*A simple analysis is used to show that the ultimate performance limits of a transistor are set by the product $Ev_s/2\pi$, where E is the semiconductor's dielectric breakdown voltage and v_s is its minority-carrier saturated drift velocity. This product, which has a value of about 2×10^{11} volts/second for silicon, emphasizes the fact that a semiconductor material has a maximum capability for energizing the electric charges that process a signal. If the operating frequency of the device is high, the frequency time period is short, and only a small amount of energy can be given to a charge carrier. Consequently, the power and power amplification must be relatively low. At low frequencies the inverse is true. In other words, device physics demands an inverse relation between frequency and power parameters that is independent of the thermal dissipation arguments commonly given to explain the trade-off between these parameters. The analysis leads to an effective means for making comparisons between existing devices.*

INTRODUCTION

IT SEEMS reasonable to suppose that an ultimate limit exists in the trade-off between the volt–ampere, amplification, and frequency capabilities of a transistor. This trade-off should somehow be related to material parameters. A trade-off relation linking these parameters would be useful if it could be derived in a manner that makes it independent of device design details, but yet applicable as a practical yardstick for evaluating device design refinement. J. M. Early,[1-3] and others[4-6] have considered this subject with more emphasis

---

[1] J. M. Early, "Maximum Rapidly-Switchable Power Density in Junction Triodes," *IRE Trans. on Electronic Devices*, Vol. ED-6, p. 322, July 1959.

[2] J. M. Early, "Speed in Semiconductor Devices," IRE National Convention, March 1962.

[3] J. M. Early, "Structure-Determined Gain-Band Product of Junction Triode Transistors," *Proc. IRE*, p. 1924, Dec. 1958.

[4] J. M. Goldey and R. M. Ryder, "Are Transistors Approaching Their Maximum Capabilities?" International Solid-State Circuits Conference, Feb. 1963.

[5] R. L. Pritchard, "Frequency Response of Grounded-Base and Grounded-Emitter Junction Transistors," AIEE Winter Meeting, Jan. 1954.

[6] L. J. Giacoletto, "Comparative High-Frequency Operation of Junction Transistors Made of Different Semiconductor Materials," *RCA Review*, Vol. 16, p. 34, March 1955.

on device design details than upon the general physical principles. One objective of the present paper is to derive general relations that show the performance limits of transistors independently of design details. Another objective is to demonstrate the use of these relationships in comparing existing devices and predicting the trend of future developments.

First, a basic voltage–frequency relation for semiconductor devices is discussed. This relationship is then extended to include space-charge constraints on device current. This combination leads directly to a relation that links maximum volt-amperes, device impedance level, and cutoff frequency. Finally, device power gain is brought into the picture through charge-control considerations. This leads to a relation that links power gain, dynamic range, and cross-modulation characteristics.

The mode of approach is to establish upper bounds on transistor performance by developing a highly idealized and simplified device model whose performance is not likely to be surpassed by that of any attainable design, no matter how optimized or cleverly conceived. This idealized performance is expressed in terms useful as a yardstick in comparing existing devices and highlighting their shortcomings. This mode of approach should complement the more usual one wherein performance improvements are extrapolated from the existing state of the art.

VOLTAGE–FREQUENCY RELATION

The charge-carrier transit-time cutoff frequency, f_T, of a charge-control type of device[7] such as a transistor is defined by the relation $f_T = (2\pi\tau)^{-1}$, where τ is the average time for a charge carrier moving at an average velocity, v, to traverse the emitter–collector distance, L. For a given value of L, minimum value of τ occurs when v is at its maximum possible value. For a semiconductor this maximum possible value is the saturated drift velocity v_s; for holes and electrons in common semiconductor materials[8] such as silicon and germanium v_s is approximately 6×10^6 cm/sec. This limiting velocity is reached at fields of the order of 10^4 volts/cm. With the carriers moving at velocity v_s, the transit time can be reduced even further by decreasing the distance L. The lower limit on L, however, is reached when the value of V/L, where V is the applied emitter–collector voltage, becomes equal to the dielectric breakdown field, E. This is approximately 10^5 volts/cm

---

[7] E. O. Johnson and A. Rose, "Simple General Analysis of Amplifier Devices with Emitter, Control, and Collector Functions," *Proc. IRE*, p. 407, March 1959.

[8] E. J. Ryder, "Mobility of Holes and Electrons in High Electric Fields," *Phys. Rev.*, Vol. 90, p. 766, June 1958.

in germanium and about twice this in silicon.[9] We are thus led to conclude that the best possible trade-off between the cutoff frequency, f_T, and the maximum allowable applied voltage, V_m, is given by

$$V_m f_T = \frac{E v_s}{2\pi} = \begin{cases} 2 \times 10^{11} \text{ volts/second for silicon} \\ 1 \times 10^{11} \text{ volts/second for germanium.} \end{cases} \quad (1)$$

This relationship[10] defines the upper limit on cutoff frequency, since the minimum value of V_m must have some value, say one volt, that is sufficiently greater than thermal voltage to ensure normal transistor collector action of the base–collector junction.

Practically attainable frequencies will be substantially less than the maximum possible frequencies indicated by Equation (1) because (1) the limiting velocity, v_s, will not be reached in all parts of the charge carrier path, (2) the electric field stress will not be geometrically uniform, and (3) the practical technology of small dimensions will have its own limitations. In practical circuit applications the maximum applied collector voltage is usually kept to a value below approximately one half of V_m to provide an adequate safety margin with respect to voltage transients and the instability effects caused by stray currents flowing to the base from the collector.

With respect to item (1), today's high-performance transistor designs are such that the collector depletion-layer thickness is two or three times the base thickness. Correspondingly, the velocity v_s is reached in approximately two-thirds to three-fourths of the carrier path. Thus, v_s is reduced by a factor of about 0.7. With respect to item (2), excluding semiconductor surface effects, the average electric field in the collector depletion layer ranges from one-half to about one-third of the maximum electric field, depending upon the impurity profile. Taking into account the relatively smaller field in the base region, the combined effect of items (1) and (2) is to reduce performance to a value that, at best, will be no greater than about one-fourth of that indicated by Equation (1). If the effect of the semiconductor surface on junction electric fields is taken into account, performance could be reduced by another factor of perhaps as much as two or three, particularly in higher voltage devices. However, by using the proper

---

[9] J. Shields, "Breakdown in Silicon on Junctions," *Jour. Electronics and Control*, Vol. 6, p. 130, Feb. 1959.

[10] First pointed out to author by W. M. Webster. Referred to by E. O. Johnson, "Whither the Tunnel Diode," AIEE-IRE Electron Devices Conference, Oct. 1962.

geometrical contour in the region where the junction intercepts the surface of the semiconductor, this effect can be reduced or eliminated.[11]

Table I lists V_m and f_T for a variety of available transistors made from both silicon and germanium, and spanning a wide range of frequency, current, voltage, and power capability. The other parameters listed in the table are considered later in the paper. The factual data in the table was obtained from handbooks and various commercial literature. This data, it should be noted, contains varying degrees of design and manufacturing conservatism. Figure 1 compares V_m and f_T values for these transistors with Equation (1).

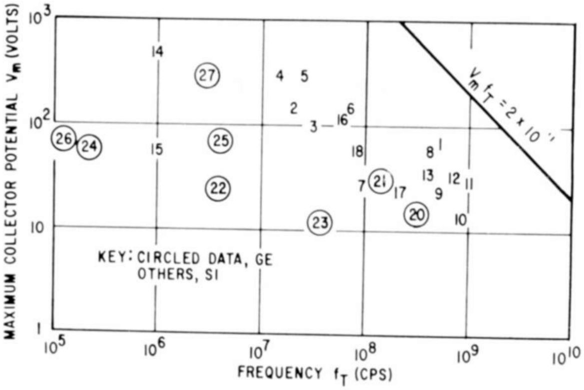


Fig. 1—Voltage–frequency relation for transistors (numbers refer to Table I).

As expected, the performance of all of the devices falls below the theoretical line. Also, as might be expected, the most recently developed devices tend to lie closest to this line. Silicon devices in almost every case are closer to the line than germanium devices. This probably stems from a lower value of E for germanium, and the fact that the germanium devices plotted in the figure are mostly of older designs. Grown-junction silicon transistors, which are the oldest devices of all and which are not shown in the figure, lie farthest away from the theoretical curve.

There is a pronounced tendency for performance to fall away from the theoretical curve in the lower frequency domain. This tendency arises from a combination of factors. Firstly, there is substantial difficulty in dealing with the lightly doped semiconductor material

---

[11] R. L. Davies and F. E. Gentry, "Control of Electric Field at the Surface of P-N Junctions," *IEEE Trans. on Electron Devices*, Vol. ED-11, p. 313, July 1964.

Table I—Device Data

				Measured Data				Calculated Data				
Device	Mat'l	Construction*	Function	Cutoff Frequency (cps)	V_m (volts)	I_m (amps)	C_0 (pf)	$P_m = V_mI_m$ (watts)	X_c at f_T (ohms)	I_mX_c (volts)	$(P_mX_c)^{1/2}$ (volts)	$(P_mX_c)^{1/2}$ f_T (volts/sec)
1. 2N3375	Si	OE	power rf	5×10^8	65	1.5	10	97.5	32	48	56	2.8×10^{10}
2. 2N3265	Si	DDE	power rf	2×10^7	150	25	900	3750	8.85	221	182	3.6×10^9
3. TA2400	Si	DDE	power rf	3×10^7	100	25	500	2500	11.6	290	170	5.1×10^9
4. TA2496	Si	EM	power sw	1.5×10^7	300	10	400	3000	26.5	265	282	4.2×10^9
5. TA2301	Si	EM	power audio	2.5×10^7	300	0.1	5	30	1270	127	195	5.0×10^9
6. TA2529	Si	EM	video driver	7×10^7	150	5×10^{-2}	2.5	7.5	910	46	82.6	5.8×10^9
7. 3N98	Si	MOS	FET	10^8	30	15×10^{-3}	2.5	0.45	1060	16	21.8	2.2×10^4
8. 2N2477	Si	DDEP	core sw	4×10^8	60	0.5	10	30.0	39.5	20	34.3	1.37×10^{10}
9. 2N2938	Si	DDEP	core sw	5×10^8	25	0.5	4	12.5	68.5	34	29.2	1.5×10^{10}
10. 2N2475	Si	DDEP	logic	8×10^8	15	0.5	3	7.5	66.5	33.3	22.4	1.8×10^{10}
11. 2N2857	Si	DDEP	rf amplifier	10^9	30	2×10^{-2}	1.8	0.6	88.5	1.8	7.3	7.3×10^9
12. 2N2708	Si	DDEP	rf amplifier	7×10^8	35	2×10^{-2}	1.5	.7	151	3	10.3	7.2×10^9
13. TA2600	Si	OE	power rf	3.5×10^8	36	5	25	180	18.2	91	57.2	2.0×10^{10}
14. 2N689	Si	D	thyristor (SCR)	10^6	500	16	50	8000	3200	51200	5060	5.06×10^9
15. 2N1513	Si	SD	power sw	10^6	60	8	200	480	800	6400	620	6.2×10^8
16. 2N2102	Si	TDP	general	6×10^7	120	1	15	120	177	177	145	8.7×10^9
17. 2N706	Si	DDP	logic	2×10^8	25	10^{-1}	6	2.5	133	13	18.2	3.6×10^9
18. 2N696	Si	DDP	power sw	8×10^7	60	0.5	20	30	100	50	54.8	4.3×10^9
19. 2N2016	Si	D	power	2.5×10^4	130	10	400	1300	15400	154000	4500	1.1×10^8
20. 2N960	Ge	EM	logic	3×10^8	15	10^{-1}	4	1.5	133	13	14.1	4.2×10^9
21. 2N1177	Ge	AD	rf amplifier	1.4×10^8	30	10^{-2}	2	0.3	570	5.7	13.1	1.8×10^9
22. 2N404	Ge	A	logic	4×10^6	25	10^{-1}	12	2.5	3310	331	91	3.6×10^8
23. 2N1301	Ge	AM	logic	3.5×10^7	13	10^{-1}	8	1.3	570	57	27.2	9.5×10^8
24. 2N301	Ge	A	power	2×10^5	60	10	300	600	2650	26500	1250	2.5×10^8
25. 2N2147	Ge	DA	power	4×10^6	75	5	400	375	100	500	194	7.8×10^8
26. 2N1358 (174)	Ge	A	power	10^5	70	15	400	1050	4000	60000	2050	2×10^8
27. TA1928	Ge	DA	power sw	3×10^6	320	10	300	3200	177	1800	820	2.5×10^9

* KEY:
OE—overlay epitaxial
DDE—double-diffused epitaxial
EM—epitaxial mesa
MOS—metal-oxide-semiconductor (field effect transistor)
DDEP—double-diffused epitaxial planar
D—diffused
SD—single-diffused
TDP—triple-diffused planar
DDP—double-diffused planar
AD—alloy-diffused
A—alloy
AM—alloy mesa
DA—drift-alloy

necessary for high collector-breakdown voltages. This stems both from the bulk material itself, and from the relatively greater surface effects experienced with lightly doped material. Secondly, the greatest amount of engineering effort, particularly developmental, has been applied in the high-frequency, relatively low-voltage domain. Thirdly, many low-frequency devices are not specifically designed to accommodate high voltages, since such voltages are not required in most typical applications.

To some degree the locations of points on Figure 1 are pessimistic, particularly for the Ge alloy devices. The values of their cutoff frequencies quoted in the literature are usually measured at collector potentials several-fold below V_m. If measured at V_m, the values of f_T would be somewhat greater than those listed in Table I. This effect is far less pronounced in the diffused-collector junction devices because their base–collector depletion region tends to expand into the collector region rather than into the base region as is the case with the alloy devices. Consequently, the effect on f_T is much less.

If device designers can extrapolate today's design perfection to higher frequencies, they will find an upper f_T limit at about 20 gigacycles. This can be seen from Figure 1 by linear extrapolation from point "1" or "11" parallel to the theoretical curve and down to the $V_m = 1$ volt line which is the nominal minimum for normal transistor operation. At 20 gigacycles the emitter–collector spacing must be of the order of 1000 Å.

The thyristor (SCR) (point "14") was included for comparison purposes that become clear in the next section. The developmental field-effect transistor (point "7") was included because it is a charge-control device and can be directly compared with the conventional bipolar transistor.[7]

CURRENT–FREQUENCY RELATIONS

The load current, I, through a charge-control device is basically defined by

$$I = \frac{Q}{\tau},$$

where Q is the total mobile charge in the emitter–collector region en route to the collector. If τ is considered as being a fixed quantity, then the ratio I/Q will also be a constant;

$$\frac{I}{Q} = \frac{1}{\tau} = \text{constant}, \qquad (3)$$

and I will increase linearly with Q. In a transistor the largest possible value of I will be attained when Q is approximately equal to the total fixed charge Q_f in the emitter–collector space. If Q were somehow made to exceed Q_f, the current I would become space-charge limited, as in a vacuum tube, and the controlling action of the base would be lost. This condition sets an absolute upper limit on I.

In practice this limit is not reached, because the mobile-charge density first exceeds the fixed-charge density in local regions of the emitter–collector path, for example, at the collector edge of base. This leads, as is well known, to the frequency-reducing phenomenon of base widening. It can also lead to an increase in the electric field in the collector depletion layer, thus reducing the maximum voltage that the device will tolerate before field breakdown occurs. The maximum rated current, I_m, of a device may be well below the maximum possible current for other reasons (e.g., power dissipation, second breakdown effects, and drop-off in current gain due to decreased emitter efficiency).

If we define the maximum current as the current I_m that causes the onset of significant base widening, Equation (3) can be written in terms of V_m and a capacitance approximately equal to the usually quoted value of the collector–base capacitance, C_0;

$$\frac{I_m}{C_0}=\frac{V_m}{\tau_C}, \tag{4}$$

where τ_C approximates the collector depletion-layer transit time. Note that the current/capacitance ratio is ideally invariant with collector area, doping level, and parallel circuit configurations of devices. For example, an increase in the doping level in the base will give a higher value of C_0, and this will allow a proportionally larger value of I_m. Equation (4) has some interesting consequences. First of all, if we replace V_m by its value from Equation (1) we see that

$$\frac{I_m}{C_0}=\frac{Ev_s}{2\pi f_T\tau_C}=Ev_s, \tag{5}$$

since $2\pi f_T\tau_C \cong 2\pi f_T\tau = 1$. The quantity I_m/C_0, a convenient measure of a device's speed in switching voltage, has a maximum attainable value that is equal to Ev_s.

Secondly, if we convert C_0 to reactive impedance X_c at the frequency f_T by using the relation $C_0 = (2\pi f_T X_c)^{-1}$, we obtain the relation

$$(I_m X_c) f_T = \frac{E v_s}{2\pi}. \tag{6}$$

This gives the most optimistic trade-off between current and frequency that one can expect from a transistor. A comparison of the devices listed in Table I with Equation (6) is shown in Figure 2. In marked contrast with Figure 1, we note that the data does not fall away from the theoretical curve in the low-frequency domain. Existing devices thus appear better at handling high currents than they are at handling high voltages. Devices having an emitter geometry that maximizes the emitter periphery–area ratio (points "1", "9", "10", and "13", for example) minimize current-crowding effects and so tend to be closer to the theoretical curve.

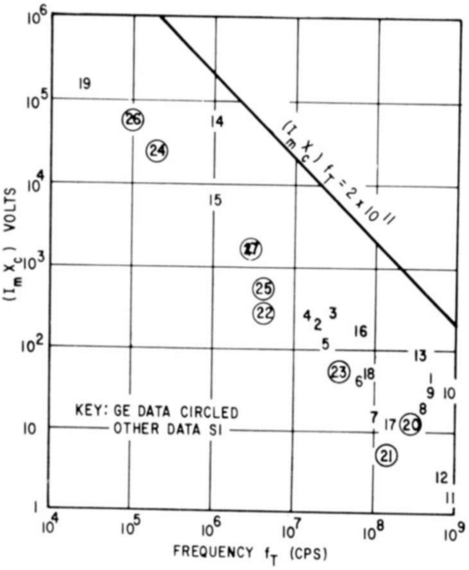


Fig. 2—Current–frequency relation for transistors (numbers refer to Table I).

Device current performance falls below the theoretical curve for reasons in addition to those listed above for the current itself. One reason is that because of current-crowding effects at the emitter periphery, the portions of the collector under the center of the emitter do not handle current. Portions of the collector that are offset from the emitter, such as those regions under base contracts, are also useless for handling load current. For these reasons the actual collector capacitance tends to be several-fold larger than is necessary for handling the rated value of collector current. Package and other stray capacitances, particularly with the high-frequency devices, also have the same deleterious effect.

As might be expected, the thyristor (SCR) device (point "14") is closer to the theoretical curve than any of the other devices. This device is not limited to a fixed I/C ratio because the maximum collector current can be almost arbitrarily large. Collector-current space-charge effects are canceled by the counter-current of carriers of opposite sign. The counter-flow of carriers of opposite sign creates a dense solid-state plasma that can carry currents not directly limited by the density of fixed charge in the device. Accordingly, a thyristor is inherently superior to a transistor in its volts/second switching capability. This superiority is further enhanced by the absence of current-crowding effects; base current is supplied from the collector rather than through the lateral base resistance as in a transistor. The price one pays for the inherent switching superiority of the thyristor is that current turn-off action becomes cumbersome and relatively slow. The f_T value noted for the thyristor in Table I was calculated from the turn-on time.

VOLT–AMPERE–FREQUENCY RELATIONS

The relation between volt–ampere product, impedance level, and frequency for a transistor is obtained by multiplying Equations (1) and (6). This combination gives

$$(P_m X_c)^{1/2} f_T = \frac{E v_s}{2\pi}, \tag{7}$$

wherein the volt–ampere product, $V_m I_m$, has been replaced by P_m. J. M. Early[1,2] has shown that the left-hand side of this relation can be derived by simple intuitive arguments. One interesting conclusion from Equation (7) is that for a given device impedance, the volt–ampere ability of a device must necessarily decrease as the device cutoff frequency, f_T, is increased.*

This decrease must occur, not because of the power dissipation arguments usually cited, but because a given semiconductor material has a limited volts/second capability for the charge carriers traversing it. For a device designed to operate at a low frequency, the time period available for energy transfer to the charge carriers is relatively long.

---

* First pointed out to the author by J. Olmstead and J. Avins of RCA.

Accordingly, energy transfer and power capability can be relatively large. For high-frequency devices the inverse is true. More specifically, the maximum electric field to which a charge carrier in the device can be subjected is the breakdown field, E. A charge carrier can traverse this field at a maximum velocity equal to v_s. The product Ev_s is the maximum rate at which a carrier can acquire volts of energy. If time is long, then the energy transfer can be large; conversely, if the time is short, the energy transfer is restricted.

The power-dissipation argument, sometimes inferred as a principle, holds that device size decreases with frequency capability and that this leads to decreased power dissipation capability and hence to a decreased ability to handle output power. As a principle this is fallacious. Equation (7) implies that, in principle, an arbitrarily large value of P_m can be achieved, for a given f_T, by connecting devices in parallel. The price paid for this is a proportionate decrease in the impedance X_c. In principle, these devices can be physically arranged in a distributed array such that the heat dissipation capability can be made almost arbitrarily high. Indeed, future practical power-output limits may well be established more by the problems of economics, circuitry, and device uniformity than by heat dissipation, per se.

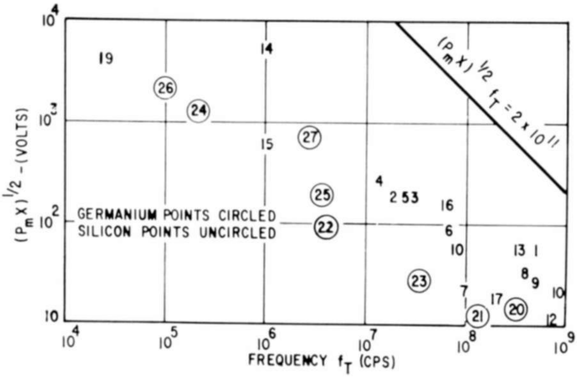


Fig. 3—Power–frequency relation for transistors (numbers refer to Table I).

Equation (7) points out that when the best power–frequency capability is desired for a transistor, the transistor should be designed to operate at relatively high current. This choice is a trade for a decrease in impedance level, but not for a decreased frequency capability as would be the case if the device were designed for low current and high voltage. Circuit requirements and practical device design problems limit how far one can proceed in the direction of high current.

Figure 3 compares the performance of the devices in Table I with

Equation (7). Again, as in Figure 1 and for the same reasons, performance tends to diverge more from the theoretical at the lower f_T values. Also, the superiority of silicon devices compared to germanium devices is strongly implied. The most recently developed designs for silicon power transistors, which use a large emitter periphery–area ratio, are closest to the theoretical curve for the reasons given previously.

RELATIONS INVOLVING AMPLIFICATION PROPERTIES

The maximum available power gain, G_p, of a charge-control device such as a transistor, is given by[7]

$$\begin{aligned} G_p &= G_I{}^2 \frac{Z_0}{Z_{\text{in}}} \\ &= \left(\frac{f_T}{f} \right)^2 \frac{Z_0}{Z_{\text{in}}}, \end{aligned} \tag{8}$$

where G_I is the current-amplification $(G_I = (2\pi f\tau)^{-1} = f_T/f)$, and Z_0 and Z_{in} are the output and input impedances, respectively. As noted here, G_p refers to the power gain that would be measured over the low-frequency pass band when the load is resistive. If we make the idealizing assumption that no electrode series resistances exist, Equation (8) can be approximated by

$$G_p = \left(\frac{f_T}{f} \right)^2 \frac{C_{\text{in}}}{C_0}, \tag{9}$$

where C_{in} is the input capacitance and C_0 is the output (base–collector) capacitance. Furthermore, if we assume that the emitter diffusion capacitance, C_d, dwarfs the emitter transition capacitance, then the maximum value of the input capacitance is given by

$$C_{\text{in}} = C_d \cong \frac{Q_m}{V_T} = \frac{I_m \tau_b}{V_T}, \tag{10}$$

where Q_m is the maximum total carrier charge en route to the collector, τ_b in the carrier base transit time, and V_T is the thermal voltage (kT/e volts). Other quantities have been previously defined. The output capacitance, C_0, is found from Equation (4):

$$C_0 = \frac{I_m \tau_c}{V_m}. \tag{11}$$

The maximum capacitance ratio is thus defined by Equations (10) and (11), and Equation (9) can be rewritten

$$G_p = \left(\frac{f_T}{f}\right)^2 \frac{V_m}{V_T} \frac{\tau_b}{\tau_c} \cong \left(\frac{f_T}{f}\right)^2 \frac{V_m}{V_T}. \tag{12}$$

The frequency ratio accounts for the input–output charge amplification of the device, and the voltage ratio, for the energy step-up per carrier. That is, in the ideal case an energy of V_T is required to place a charge on the control electrode, the base. As a consequence, each of the charges that flows to the collector ideally picks up an energy equal to V_m.

Equation (12) can be written in a form consistent with the other performance trade-off relations by using Equation (1);

$$(G_p V_T V_m)^{1/2} f = \frac{E v_s}{2\pi}. \tag{13}$$

Here G_p is evaluated at some operating frequency, f, sufficiently close to f_T for Equation (8) to be valid. The value of V_m corresponds to that used in Equation (1). Equation (13) emphasizes the fact that power amplification, like volt–ampere performance, basically depends upon the volts/second capability of the semiconductor material.

The actual performance of a transistor will fall far short of that predicted by Equation (13). First of all, the qualifications on practically attainable electric fields and average carrier velocities noted for Equation (1) will apply. This reduces the right-hand constant to an effective value equal to about one fourth, or less, of the theoretical value. Secondly, the practical working value of applied collector voltage will be about one half of V_m (see Table II). Thirdly, Equation (10) overestimates the value of C_d and, hence, G_p by perhaps as much as two times in a device operating at very high carrier densities. This disparity stems from an electric-field-enhanced carrier velocity in the base. Fourthly, this same electric field in the base doubles the input-energy requirement per carrier. Fifthly, another factor of at least two is added because of ohmic base resistance. The energy lost in the resistance of a series RC circuit is equal to the potential energy acquired by the charges on the capacitor when the charging time is more than a few-fold larger than the *RC* time constant. The energy loss in the base resistance can be very much greater than this when the driving signal's time constant is less the *RC* time constant of the circuit. Transition capacitance accentuates the problem because it increases the time constant of the transistor's input circuit. The effect of all these factors is to make the right-hand constant in Equation (13) at least thirty-fold less than the theoretical constant.

Table II

			Measured Data					Calculated Data
Device	Material	Construction*	Operating Freq. (f) (cps)	Power Gain (G_p) (db)	(ratio)	Voltage Operating	(volts) Max. (V_m)	$(V_m V_T G_p)^{1/2}$ (volts)
1. 2N1631	Ge	DA	1.5×10^6	47.7	6.0×10^4	12	34	232
2. 2N1425	Ge	DA	4.6×10^5	51.0	1.3×10^5	12	24	280
3. 2N1180	Ge	DA	1.0×10^7	35.0	3.2×10^3	12	30	15.5
4. 2N1638	Ge	DA	2.6×10^5	61.5	1.4×10^6	11	34	1100
5. 2N384	Ge	DA	5.0×10^7	21.0	1.3×10^2	12	40	11.4
6. 2N1066	Ge	DA	5.0×10^7	26.0	4.0×10^2	12	40	20
7. 2N1177	Ge	DA	1.0×10^8	14.0	2.5×10^1	12	30	4.3
8. 2N175	Ge	A	2.0×10^4	43.0	2.0×10^4	4	10	71
9. 2N139	Ge	A	4.6×10^5	37.0	5.0×10^3	9	16	45
10. 2N408	Ge	A	4.6×10^5	37.8	6.0×10^3	9	20	55
11. 2N2873	Ge	M	1.8×10^8	21.6	1.5×10^2	12	35	11.4
12. 2N2708	Si	DDEP	2.0×10^8	22.0	1.6×10^2	15	35	11.8
13. 2N2875	Si	DDEP	4.5×10^8	19.0	8.0×10^1	6	30	7.8
14. 3N98	Si	MOS	6.0×10^7	10.0	10	20	30	34.6

* KEY: **DA**—drift alloy
A—alloy
M—mesa
DDEP—double-diffused epitaxial planar
MOS—*M*etal *O*xide *S*emiconductor—field effect transistor
As noted in text, the voltage corresponding to V_T for the MOS is approximately the pinch-off voltage. Value used in calculation was 4.0 volts.

A plot of data in Table II is shown in Figure 4 along with the theoretical curve corresponding to Equation (13). The points closest to the theoretical curve correspond to the most recently developed devices and are roughly a hundred-fold removed. The next closest group of points, "1" through "7", correspond to the next most recently developed devices, the alloy drift types. The farthest removed group of points, "8", "9", and "10", correspond to the oldest device types, the alloy devices.

One reason for the tendency for the points to diverge increasingly from the theoretical curve at lower frequencies has to do with the previously discussed factors involving V_m, which apply to the results in Figure 1. Another reason is that circuit instability problems will tend to limit the amount of power gain sought by the device designer.

Making a speculative extrapolation to higher frequencies with Figure 4, as was done for the data in Figure 1, we see that at an operating frequency of 20 gigacycles the value of $(G_p V_T V_m)^{1/2}$ will be somewhat less than unity, and G_p will thus be limited to a maximum value of about ten.

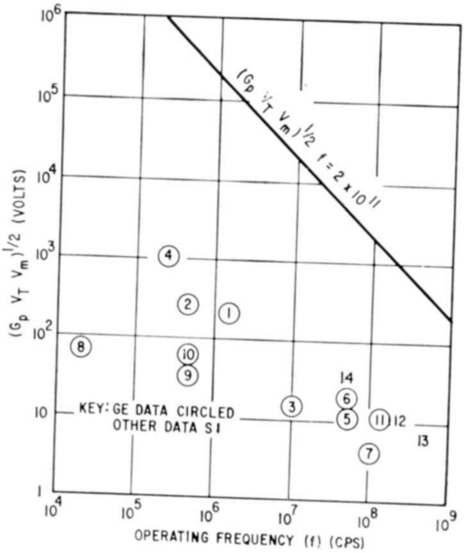


Fig. 4—Power-gain and frequency relation for transistors (numbers refer to Table II).

Equations (12) and (13) suggest that one must trade power gain for the energy required to place a charge on the control electrode.[7] For the bipolar transistor, this input energy has been noted as V_T volts; for the field-effect transistor, this energy can be associated with the pinch-off voltage, V_p, or the G_m/I ratio, and is ten to a hundred times greater than V_T. Both V_p and V_T define approximately the useful dynamic input voltage range of field effect and bipolar devices. Simple consideration shows that the larger this dynamic input-voltage range, the smaller the curvature components on the transfer characteristics tend to be. Consequently, the tendency for cross-modulation phenomena in the device will be smaller, particularly when the device operating point is swept along the transfer characteristic during reverse a-g-c action.*

---

* Varactor action in the collector circuit is not included in this argument.

One concludes that a bipolar transistor will tend to have more power gain at a given operating frequency than a field-effect transistor when both devices are constructed to have the same value of f_T. On the other hand, the field-effect transistor will have a greater dynamic range and, correspondingly, a greater ability to avoid cross-modulation interference effects.[12] Point "14" in Figure 4 shows that the field-effect transistor holds its own against the bipolar transistor when the dynamic range factor is taken into account.

CONCLUSIONS

A simple analysis of transistors shows that the product of the semiconductor breakdown field and the maximum carrier drift velocity, Ev_s, is the ultimate measure of a transistor's volt–ampere, power gain, and frequency performance. This product emphasizes that a given semiconductor material used in a transistor has a definite and fixed capability for imparting volts of energy per second to a charge carrier. For this reason the energy transfer to a charge carrier must necessarily decrease with frequency as must the volt–ampere and power-gain performance. However, if the designer is willing to pay the price of a decreased impedance level, a transistor or transistor array can, in principle, be made to give an almost arbitrarily high volt–ampere performance at any operating frequency within the frequency limits of the device.

For a given device cutoff frequency, f_T, there is a trade-off between power gain and dynamic range. The gain of a bipolar transistor will tend to be superior to that of the field-effect transistor, while the field-effect transistor will tend to have a superior dynamic range and ability to avoid cross-modulation interference.

The various trade-off relations can be plotted in a simple manner that compares transistors having a very wide range of parameter values.

A 60-MEGACYCLE VIDEO CHAIN FOR HIGH-DEFINITION TELEVISION SYSTEMS

BY

O. H. SCHADE, SR.

RCA Electronic Components and Devices
Harrison, N. J.

***Summary**—Direct observation of high-definition television images requires frame-scanning times of at least 1/15 second and video-frequency channels in the order of 30 to 60 megacycles wide. A theoretical analysis demonstrates that low-noise pictures can be obtained with camera tubes containing electron multipliers that furnish peak signal currents of the order of 15 microamperes. Reliable noise-free video amplifier systems that have constant amplitude response in a 60-megacycle bandwidth and that can deliver peak drive signals of 50 volts to a high-definition picture tube can be built with nuvistor tubes and conventional circuit components. The frequency-compensated camera preamplifier contains seven tubes, including a cascode input stage and a line driver. Linear-blanking insertion and black-level setting require two video amplifiers, a second line driver, and four pulse amplifiers. The picture-tube drive amplifier contains an input amplifier having adjustable gain, a cathode follower, and three nuvistor tetrodes operating in parallel. All stages are a-c coupled and self-adjusting so that chain reactions cannot occur. The system requires a 150-volt power supply.*

THEORY

HIGH-DEFINITION television systems generally use slow-speed scanning and frame times, T_f, in the order of seconds to transmit high-definition video signals over a moderately wide frequency band. A complete picture can thus be seen only on a second optical storage surface, generally a photographic film, and a second imaging system is required. Because of the considerable time delay between camera exposure and observation of the final image, adjustment and optimization of operating conditions can be a slow step-by-step process, and elaborate precautions are required to maintain adequate and long-time stability and calibration.

To eliminate the time delay and long-term stability requirements and obtain the advantages of immediate observation of cause and effect, it is expedient to scan only a fraction of the total image area at a higher frame repetition rate ($T_f = 1/15$ to $1/30$ second). Different areas of a larger image surface can then be studied by displacement of the small scanned area to different positions. The observations, however, are not conclusive when the unscanned and undeveloped area of the image alters the electric-field distribution surrounding the scanned area, as is often the case in camera tubes. Thus tests may be required with a fully scanned image surface and with a scanning-line density that leaves no unscanned areas in one frame period. The high scanning-line density required for high-definition (small beam diameter) tubes, however, results in a low limit of the observable spatial frequency response in the (horizontal) scanning-line direction unless the frequency band of the video signal amplifier is increased to high values. The electrical frequency band, Δf, in cycles per second required for a practical video system is given by

$$\Delta_f = \frac{2b\,(Xf_{s(x)})\,(Yf_{s(y)})}{T_f} \tag{1a}$$

where X and Y are the horizontal and vertical image dimensions in millimeters, $f_{s(x)}$ and $f_{s(y)}$ are the spatial cutoff frequencies of the system in cycles per millimeter in X and Y, b is the blanking factor and is equal to $T_f/T_{f(\text{unblanked})}$, and T_f is the frame time in seconds. The factor 2 appears in this equation because a minimum of two raster lines (or samples) are required in the Y direction to reproduce one cycle of the limiting spatial frequency $f_{s(y)}$.

It is customary in television engineering to state the resolution limits in terms of the total number ($N_{c(x)}$, $N_{c(y)}$) of television lines (half cycles) in a unit length equal to the vertical dimension (Y) of the image frame. If the following definitions are used:

$$N_{c(y)} = 2f_{s(y)}Y$$

$$N_{c(x)} = 2f_{s(x)}X\,(Y/X)$$

then Equation (1a) may be written

$$\Delta_f = \frac{0.5b\,(X/Y)\,N_{c(x)}N_{c(y)}}{T_f}. \tag{1b}$$

A bright flicker-free picture-tube image is obtained with a frame time of 1/30 second and 2-to-1 interlacing (60 fields per second). Satisfactory observations can also be made with T_f of 1/15 second and 2-to-1 interlaced fields, but the visual integration of the interlaced fields becomes inadequate at brightness levels exceeding approximately one foot-lambert for conventional picture-tube phosphors. Table I shows the electrical frequency band, Δf, for a number of conditions.

The usefulness of a wide-band channel is limited by the signal-to-noise ratio that can be obtained. The image-signal current is determined by exposure and charge storage in the camera-tube storage surface, and the amplitude of detail signals depends on the sine-wave response of the lens and of the camera tube and its electron multiplier. A *constant frequency response of the multiplier* of conventional camera tubes requires a high potential difference (500 to 600 volts) between the second dynode and the first dynode (grid No. 2). In addition, the multiplier focusing electrode and grid No. 3 must be negative with respect to grid No. 2 by approximately 50 volts to prevent a transit-time spread of the secondary electrons, which travel in semicircular paths about 3 centimeters long. The plane-parallel fields of the remaining dynodes provide uniform transit times at normal potential differences (20 to 200 volts). To eliminate phase-displaced image signals, which appear like reflections in an improperly terminated line, it is necessary to prevent secondary electrons generated by the beam at the field mesh near the storage surface from returning to the first dynode (grid No. 2). Therefore, the mesh electrode is made a few volts positive with respect to the drift-space potential (grid No. 4) or a suppressor electrode is inserted between grids No. 2 and No. 4 and is made negative with respect to grid No. 4.

Table I—Relation of Television Resolution Limits and Frequency Channel

T_f (seconds)	X/Y	$N_{c(y)}$	$N_{c(x)}$	b^*	Δf (Mc)
1/30	1	1000	1000	1.23	18.5
1/30	1	1000	2000	1.23	37.0
1/30	4/3	1000	1000	1.23	24.6
1/30	4/3	1000	2000	1.23	49.5
1/30	1	1500	1500	1.31	44.2
1/30	4/3	1500	1500	1.31	59.0
1/15	1	2000	2000	1.23	37.0
1/15	1	2000	3000	1.23	55.2
1/15	4/3	2000	2400	1.23	59.0

* Computed for a line blanking time of 5 microseconds and a 6 percent frame blanking time with $b = 1.06\ ((T_f/1.06N_y) + 5 \times 10^{-6})/(T_f/1.06N_y)$.

Figure 1 shows a schematic diagram of a wide-band (60-megacycle) preamplifier for a television system. The principal noise sources are the scanning beam and the first stage of the video amplifier. Amplifier noise becomes excessive for wide frequency bands when signal input currents are reduced to such small values that initial signal amplification by electron multiplier stages in the camera tube becomes necessary. The gain of this multiplication can be computed as follows.

Amplifier Noise

The signal-to-noise ratio of a cascode amplifier depends on its total input capacitance, C, and its equivalent noise resistance, R_{eq}. The r-m-s noise voltage, E_n, at the input of an electron-tube amplifier working into a "flat" frequency band Δf is given by

$$E_n = 1.3 \times 10^{-10} (R_{eq} \Delta f)^{1/2}.$$

When the input resistance, R, is shunted by a capacitance, C, a gain–frequency correcting network must follow the amplifier to restore a flat frequency response. This correction changes the "white" noise of the input stage to "peaked" noise, and the r-m-s noise voltage becomes

$$E_{n(c)} = \frac{E_n 2\pi \Delta f C R}{\sqrt{3}}.$$

The signal voltage after correction has the constant value

$$E_s = I_s R,$$

where I_s is the signal current. The signal to r-m-s noise ratio $[R]_a$ is then given by

$$\begin{aligned}[R]_a = \frac{E_s}{E_{n(c)}} &= \frac{I_s R \sqrt{3}}{2\pi \Delta f C R (1.3 \times 10^{-10}) (R_{eq} \Delta f)^{1/2}} \\ &= \frac{2.13 I_s \times 10^9}{C (R_{eq})^{1/2} (\Delta f)^{3/2}}. \qquad (2)\end{aligned}$$

The noise-spectrum components in the corrected channel are proportional to frequency (peaked noise). A signal to r-m-s noise ratio $[R]_a$ of 200 in a frequency band, Δf, of 60 megacycles can thus be obtained with a total input circuit capacitance, C, of 30 picofarads and equiva-

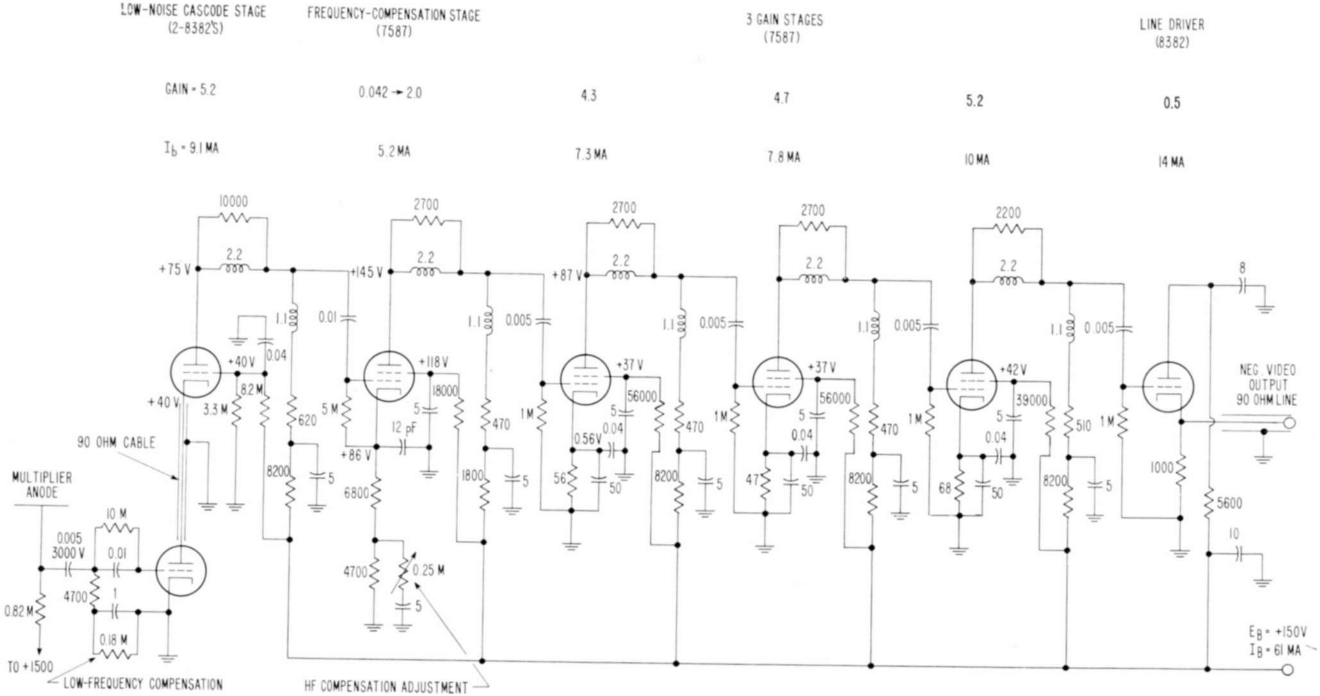


Fig. 1—Schematic diagram of frequency-compensated wide-band (60-megacycle) camera preamplifier using seven nuvistor tubes. (Except where indicated, all capacitance values are in microfarads, all inductance values in microhenries, and all resistance values in ohms.)

lent noise resistance, R_{eq}, of 235 ohms when the signal current from the multiplier, I_s, is approximately 20 microamperes. The nuvistor cascode amplifier shown in Figure 1 has an equivalent noise resistance of 235 ohms.

Signal and Noise

The signal to r-m-s noise ratio $[R]_0$ of the signal current, I_0, developed from the charge image in a camera tube can be computed basically from the following equation:

$$[R]_0 = (\bar{n}_0 \bar{a})^{1/2}$$

where $\bar{n}_0 = [Q_f/(X/Y)]/q_e =$ the total charge stored during the exposure time T_f in the unit area divided by the electron charge q_e; $\bar{a} = [N_{e(x)} N_{e(y)}]^{-1}$ is the equivalent sampling area computed from the equivalent noise pass bands $N_{e(x)}$ and $N_{e(y)}$. Substituting $Q_f = I_0 T_f/b$ and charge $q_e = 1.6 \times 10^{-19}$ ampere-second,

$$\bar{n}_0 = \frac{(6.25 \times 10^{18} I_0)\, T_f}{b\,(X/Y)}.$$

The value of $\bar{a}$ can be determined from Equation (1b):

$$\bar{a} = \frac{0.5 b\,(X/Y)}{T_f \Delta f_e}.$$

The value Δf_e is the noise-equivalent pass band computed from the product of the frequency-response functions of the storage surface, scanning beam, and video amplifier. It can be computed from Equation (1b) when the values $N_{c(x)}$ and $N_{c(y)}$ are replaced by the respective equivalent passband $N_{e(x)}$ and $N_{e(y)}$ of the system. The ratio $[R]_0$ in amperes per cycle is then given by

$$[R]_0 = \frac{3.12 I_0 \times 10^{18}}{(\Delta f_e)^{1/2}}. \qquad (3)$$

Beam Modulation and Total Noise

The scanning beam is modulated by electron absorption at the storage surface so that the return beam to the signal multiplier carries the signal modulation. The (mean) signal current, I_0, is thus equal to $m \times I_b$, where m is the modulation factor. Because both currents contain shot noise, their mean squared noise-current ratio is given by

$$\frac{\bar{i}_{n(b)}{}^2}{\bar{i}_{n(0)}{}^2} = \frac{\Delta f}{m \Delta f_e}$$

where $\bar{i}_{n(0)}{}^2$ is the mean squared noise of the signal current I_0 and Δf_e is its equivalent pass band. The signal to r-m-s noise ratio in the return beam $[R]_m$ is then given by

$$[R]_m = \frac{I_0}{(\bar{i}_{n(b)}{}^2 + \bar{i}_{n(0)}{}^2)^{1/2}} = \frac{I_0}{I_{n(0)} \left(1 + \dfrac{\Delta f}{m \Delta f_e}\right)^{1/2}}$$

or

$$[R]_m = \frac{[R]_0}{\left(1 + \dfrac{\Delta f}{m \Delta f_e}\right)^{1/2}}. \tag{4}$$

The beam modulation factor, m, obtainable for a complete discharge of a stored charge has a value less than unity because the mean return-beam value $(I_b - I_0) = I_0 \, [(1/m) - 1]$ must at least equal the negative peak values of the total noise modulation. This requirement can be written as follows:

$$I_0 = \left[\frac{1}{n} - 1\right] = p I_n \left[\frac{1}{m} + 1\right]^{1/2}.$$

Substitution of the ratio $I_0/I_n = [R]_0$ simplifies this expression:

$$\frac{1}{m} - 1 = \frac{p}{[R]_0} \left[\frac{1}{m} + 1\right]^{1/2}$$

where p is the ratio of peak noise amplitude to the r-m-s noise value. The value of p depends on the assumed probability of occurrence and is generally between 2.5 and 3.

The solution of the quadratic equation for m is

$$m = \frac{2a}{(2a+1) + (8a+1)^{1/2}}, \text{ with } a = \left(\frac{[R]_0}{p}\right)^2 \tag{5}$$

It follows that the maximum possible beam-current modulation for a complete discharge is a function of $[R]_0$. If $[R]_0 = 3$, a very low value, and if $\Delta f = \Delta f_e$, then when $p = 3$, $m = 0.333$ and $[R]_m = 1.5$.

When $p = 2.5$, $m = 0.4$ and $[R]_m = 1.6$. For a high value of $[R]_0$, the maximum modulation approaches unity and $[R]_m$ approaches $[R]_0/\sqrt{2}$.

In actual operation of camera tubes with electron multipliers, the beam current is adjusted by visual observation of highlights and the general noise level in the image. The beam is reduced to a value that just prevents "blocking" of highlights and that does not require a complete discharge beyond the mean charge level. As a result, the equivalent passband, Δf_e, is reduced to a lower value, $\Delta f_e'$, in the highlights; this change results in some loss of fine highlight detail but provides an improved signal-to-noise ratio $[R]_m$.

Example

The product of the sine-wave response of the storage surface and scanning beam measured on an experimental multiplier vidicon camera is approximately 55 per cent at $N_{c(x)} = N_{c(y)} = 1500$ lines. The equivalent pass bands are $N_{e(x)} = N_{e(y)} = 1050$ lines. When these values are substituted in Equation (1), the equivalent electrical pass band, Δf_e, is 28.8 megacycles for an aspect ratio $(X/Y) = 4/3$, $b = 1.31$, and the frame time $T_f = 1/30$ second. The bandwidth of the electrical system, Δf, calculated for $N_{c(x)} = N_{c(y)} = 1500$ is 59 megacycles.

When the measured highlight signal current, I_0, is 0.2 microampere and the modulation factor m is 0.4, the photon signal-to-noise ratio $[R]_0 = 147$ as determined from Equation (3), and the multiplier signal-to-noise ratio $[R]_m = 60$ from Equation (4).

The noise contribution from the cascode amplifier ($[R]_a = 200$) is therefore insignificant when the multiplier signal output current is 20 microamperes. The multiplier gain, g, has a moderate value equal to 20/0.2, or 100. Because multiplier gains in the order of 1000 are quite common, a 60-megacycle bandwidth is useful for observation of camera-tube performance at high-light signal currents as low as 0.02 microampere, for which $[R]_m = 20$. The threshold signal-to-noise ratio $[R]_m$ required for a continuously observed 3-bar test object[1] is in the order of 2.

DESIGN OF WIDE-BAND VIDEO AMPLIFIER SYSTEMS WITH NUVISTOR TUBES

A video amplifier system can be subdivided into three separate amplifiers:

(1) a linear preamplifier, located in the television camera, consisting of a low-noise input stage followed by a gain–frequency correction amplifier (shown in Figure 1) that compensates

---

[1] O. H. Schade, "An Evaluation of Photographic Image Quality and Resolving Power," *Jour. SMPTE*, Vol. 73, p. 81, Feb. 1964.

for the decreasing frequency response of the capacitive input impedance;

(2) a line amplifier (shown in Figure 5) for linear-blanking insertion and black-level setting; and

(3) a picture-tube drive amplifier (shown in Figure 8) that provides d-c restoration and video signals of the order of 50 volts.

Signal levels in the 50- to 90-ohm lines between these amplifiers should be of the order of 1 to 2 volts. To prevent ground currents, long lines should be grounded directly only at the input end. The over-all frequency response and gain to the output of the line amplifier should be constant to facilitate measurements with a wide-band oscilloscope connected to this line. The signal transfer curve of the picture-tube amplifier can be permitted to have a moderate white-signal compression (high-light gamma less than unity) to counteract expansion of the white signal by the transfer characteristic of the high-gamma picture tube.

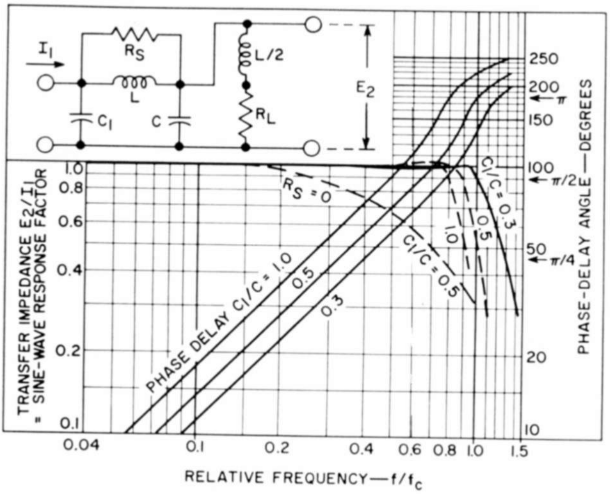


Fig. 2—Schematic diagram and characteristics curves for four-pole low-pass filter designed for constant-current input.

Interstage Coupling

The damped four-pole low-pass filter shown in Figure 2 is a modified pi-section filter designed for constant-current input. It is terminated at one end to provide a maximum gain–bandwidth product and low phase distortion. The capacitances C and C_1 represent the total plate- and grid-circuit shunt capacitances. The filter is reversible; C_1 should always be smaller than C. The series inductance, L, is shunted by a damping resistance, R_s, to provide minimum phase shift and a "flat" response. As shown in Table II, the value of R_s depends on the ratio C_1/C, which also determines the bandwidth factor, K, and the exact value of the terminating resistance, R_L. $K = \Delta f/f_c$, where Δf is the bandwidth over which the response is constant within the stated deviation and f_c is the theoretical cutoff frequency. The constants of the filter are determined by the following equations:

Table II

C_1/C	R_s/R_L	K	R_L/R	Deviation in amplitude
0.3	5.66	0.96	1.088	+1% to −2%
0.5	5.66	0.85	1.075	2.5%
1.0	18.9	0.76	1.0	3.0%

bandwidth for flat response: $$\Delta f = \frac{K}{RC\pi}, \tag{6a}$$

design resistance: $$R = \frac{K}{\Delta f C\pi} \tag{6b}$$
(for actual resistance R_L see Table II),

voltage gain of stage: $$g = \frac{Kg_m}{\Delta f C\pi} = g_m R_L,$$
inductance: $$L = R^2 C. \tag{6c}$$

For example, the capacitances C and C_1 of a coupling filter for a nuvistor tetrode can be estimated as follows:

tube input capacitance (cold) ..	6	pf
increase due to space charge ...	2.2	pf (for $g_m = 0.01$ mho)
circuit capacitance (socket, leads, coils, grid capacitor)	1.75	pf
total C	10.0	pf (approx.)
plate capacitance	1.5	pf
circuit capacitance (leads and coil)	1.2	pf
total C_1	2.7	pf

The ratio C_1/C is then about 0.27. Table II shows that R_s/R_L is 5.66 and K is 0.96 for a ratio of 0.3. For a frequency band Δf of 60 megacycles and a transconductance g_m of 10 millimhos, therefore, $R = 510$ ohms, $R_L = 550$ ohms, $R_s = 3000$ ohms, and $L = 2.6$ microhenries; the gain g is then 5.5.

When the driven tube has an unbypassed cathode resistance R_k, the nuvistor "hot" grid–cathode capacitance C_{gk} of approximately 6 picofarads is reduced by the gain reduction factor $1/(1 + g_{mk}R_k)$, where $g_{mk} = g_m(I_b + I_{c2})I_b$. For $g_{mk} = 12$ millimhos, for example, a cathode resistor of 90 ohms (used in cathode line drivers or for linearization) reduces the input capacitance by $1/(1 + 0.012 \times 90)$; i.e., $C_{gk} = 6 - 0.48 \times 6 = 3.1$ picofarad. The capacitance C is thus reduced to approximately 7 picofarads, and R_L is increased to 730 ohms, provided the shunt reactance of the cathode-to-ground capacitance is not significant. The cathode to heater and ground capacitance is approximately 2.5 picofarads, and the shunt reactance is 1060 ohms at a frequency of 60 megacycles. The voltage gain in the plate circuit of the stage is then given by Equation (5) divided by $(1 + g_{mk}R_K)$; i.e., $g = 7.3/1.9 = 3.85$. The capacitance values of an actual circuit may differ somewhat from estimated values and cause departures from a constant-amplitude response. The frequency response can be made constant by adjustment of the values of R_L and/or R_S, as illustrated in Figure 3. Errors in bandwidth require correction of the inductance values.

The gain–bandwidth product $g\Delta f$ of practical nuvistor stages using tetrodes that have a transconductance g_m of 10 millimhos can be calculated from Equations (6c) and (7) to be approximately $(0.9/\pi)$ gigacycles or 286 megacycles. The gain–bandwidth product of the picture-tube driver stage is increased considerably by operation of 3 tetrodes in parallel. With values of $C = 7.7$ picofarads (picture tube), $C_1 = 6$ picofarads (3 tetrode plate circuits), and $R_k = 27$ ohms per tube, the effective transconductance value is 39/1.46 millimhos and the gain–bandwidth product $g\Delta f$ is $(0.80 \times 27/7.5\pi) = 0.91$ gigacycle. The tube gain is degenerated by the 27-ohm cathode self-bias resistor to prevent white-signal compression at large signals. This resistor decreases the total input capacitance to 22 picofarads and permits an input filter resistance R_L of 220 ohms for a 60-megacycle bandwidth of the cathode-follower driver.

The Preamplifier

In the wide-band preamplifier circuit shown in Figure 1, the input tube of the cascode stage is mounted on a ground plate that supports the camera-tube socket. Its plate is d-c coupled over a short 90-ohm line to the second tube, which is located in the preamplifier chassis proper. The total input capacitance at the grid of the cascode amplifier is 25 to 30 picofarads. The parallel input resistance is assigned a value R of 4700 ohms. The multiplier signal current I_s of 15 microamperes develops an input signal voltage of 70 millivolts at low frequencies. The signal decreases to 49 millivolts at a frequency of 1.13 megacycles where $R = 1/\omega C$. Beyond this frequency the input voltage decreases in inverse proportion to frequency to 1.5 millivolts at 60 megacycles. The compensating inverse frequency characteristic is obtained by a single RC circuit (same product) in the cathode of the second stage. The input voltage E_{in} to this stage is the vector sum of the cathode voltage $E_k = g_{mk}/\omega C_k$ and the grid–cathode voltage, which are in quadrature at high frequencies. The corrected frequency-response factor, r_f, is given by

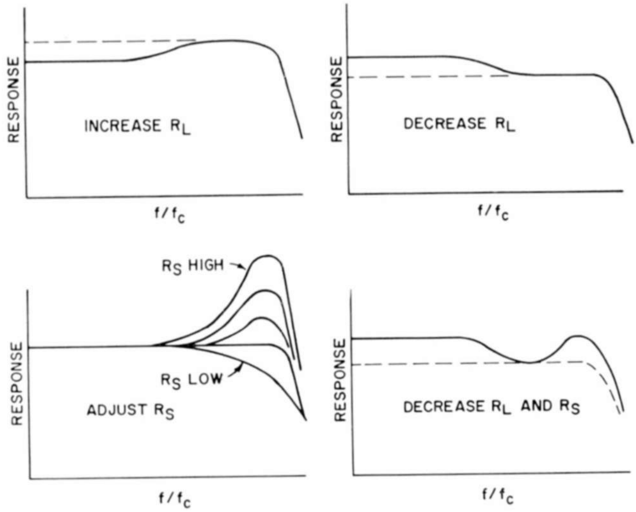


Fig. 3—Curves indicating adjustment of R_L and/or R_s to maintain constant frequency response.

$$r_f = \frac{\dfrac{g_{m(k)}}{\omega C_k}}{\left[\left(\dfrac{g_{mk}}{\omega C_k}\right)^2 + 1\right]^{1/2}}. \tag{7}$$

For $C_k = 15$ picofarads and $g_{mk} = 10$ millimhos, the response factor at 60 megacycles, $r_{60} = 0.87$; it increases to $r_{50} = 0.905$ and $r_{40} = 0.93$ at 50 and 40 megacycles, respectively. The small deficiency at higher frequencies is corrected in a later stage. The cascode input stage and the second stage operate with a small grid current (0.03 microampere) and use a grid-current self-bias resistor to obtain a high transconductance-to-plate-current ratio (g_m/I_b) and a low plate-voltage drop. Thus, adequate low-frequency compensation may be obtained for the grid-circuit time constant with a moderate B-supply voltage. The 90-ohm cable connecting the two cascode triodes increases the cathode–ground capacitance of the second triode to 21 picofarads. The corresponding reactance to ground of 126 ohms at 60 megacycles bypasses the normal cathode load resistance, which is the plate resistance ($r_p = 2500$ ohms) of the first triode. As a result, the effective plate impedance of the second triode decreases from its degenerated low-frequency value r_{p2} of approximately 80,000 ohms to about 6000 ohms at 60 megacycles. The resistance values of the coupling filter to the following stage must then be modified as shown in Figure 1. The voltage gain of the preamplifier, measured with a supply voltage of 140 volts, is 12 at 200 kilocycles and 600 at 60 megacycles; it provides an output signal of 1 volt in a 90-ohm line with a signal input current of 17 microamperes. Measured frequency-response functions are shown in Figure 4. The preamplifier requires a-c heater power of 6.36 watts and d-c power of 150 volts × 61 milliamperes, or 9.1 watts.

Blanking Insertion and Black-Level Setter

Figure 5 shows a schematic diagram of the circuit used to provide blanking insertion and black-level setting. The retrace time of the camera-tube beam must be shorter than that of the signal blanking time because the camera signal may contain white-signal "spikes" adjacent to its black-level signal that exceed the video signal by a factor of two. A large negative picture-tube blanking signal is therefore added to the video signal by tube (1a) to depress the white spikes and camera-blanking section below the desired blank level.

The black-signal portions of the camera signal are set to a constant negative potential by a delayed 2-microsecond clamp pulse (produced by tubes (1b) and (1c)) that causes a diode bridge to conduct and to reset the potential of the grid capacitor of tube (2) during the camera retrace time. This potential is adjusted by the black-level control to a value such that the unwanted depressed signal portions are clipped off at the desired black-level distance below the video signal by current cutoff in tube (2). The small "setup" of the black values in the video

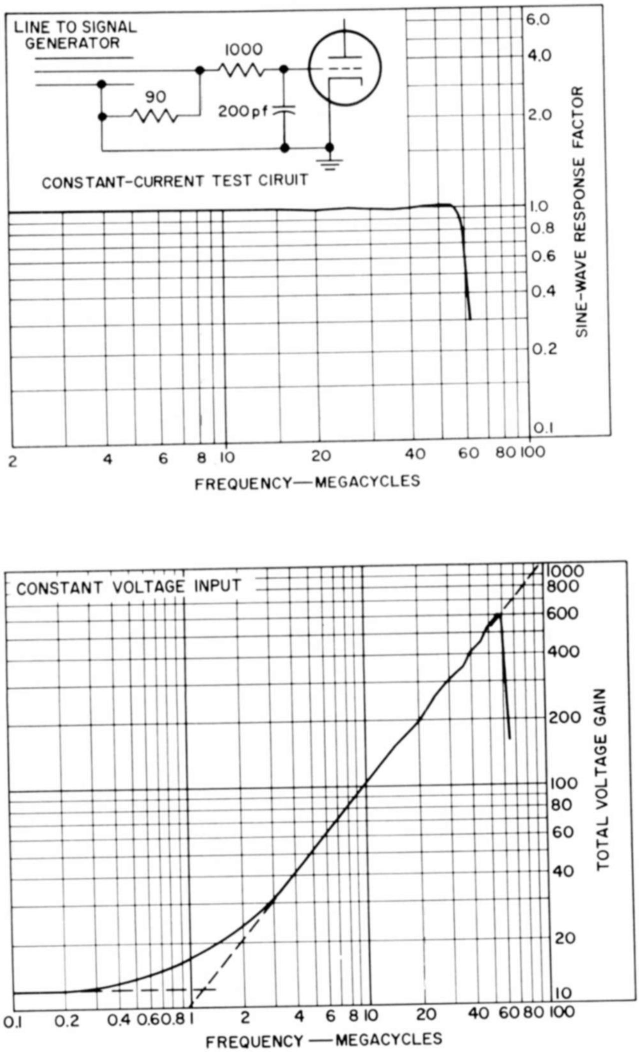


Fig. 4—Frequency-response curves for preamplifier shown in Figure 1.

signal with respect to this clipping level requires a linear characteristic down to current cutoff to avoid dark-signal compression. The cutoff curvature in tube (2) is, therefore, reduced by degeneration in a 100-ohm cathode resistance as shown in Figure 6. The remaining curvature is compensated by operation of tube (1) with an opposite curvature fixed in position by a d-c setting diode in its grid circuit. The over-all amplitude transfer curve is shown in Figure 7.

Picture-Tube Drive Amplifier

A schematic diagram of the picture-tube drive amplifier is shown in Figure 8. The triode gun of a 17-inch high-definition picture tube operated at 20 kilovolts requires a black-level bias, E_{cc0}, of approximately -230 volts and peak video signals in the order of 40 volts. The total grid-circuit capacitance of 7.75 picofarads determines the load resistance, R_L, of 560 ohms. The small output capacitance of the 7587 nuvistor tetrode permits parallel operation of three tubes. The zero-signal operating point is located on a grid-bias curve $E_{c1} = -0.5$ volt obtained by use of a 27-ohm self-bias resistor in each cathode circuit. The effect of the input signal voltage ($E_{in} = E_{c1} + E_k$) on the degenerated plate characteristic is shown in Figure 9. The operating point (0) and the d-c and a-c load lines per tube are indicated. The zero-signal plate current per tube is 14.5 milliamperes at a plate voltage, E_p, of 111 volts. The black level is restored at the picture-tube grid by a d-c restorer circuit containing two negative-peak rectifier diodes. A fraction of the rectified d-c output of one diode is connected in series with the full rectified voltage of the second diode to obtain peak rectification greater than unity to compensate for the black-level setup. The grid-drive voltage for the output tubes is developed across a 220-ohm resistor by a cathode-follower triode. The first stage in Figure 8 provides linearity compensation for the output and driver stages, as shown by the transfer curves in Figures 10 and 11. The signal-transfer curve of the 3-stage amplifier, shown in Figure 11, includes the d-c restoration; this curve was measured with blanked-level signals injected into the amplifier line input. Curve B shows the measured picture-tube screen brightness in foot-lamberts as a function of video line voltage.

The d-c power drain of the amplifier is 80 milliamperes at 150 volts. The current drain can be reduced to 35 milliamperes by class-B operation of the output stage. This approach requires a reversal of signal polarity in the output stage and cathode-drive operation of the picture tube. The circuit is d-c coupled from the cathode follower to the picture-tube cathode with d-c restoration at the control grid of the cathode follower, and the output tubes are operated with a zero-signal bias of -2 volts (2.5 milliamperes per tube). Screen-grid and plate voltages must be stabilized (Zener diodes are used), and considerably more negative feedback is required to avoid black-signal compression. The class-B circuit does not possess the stability and self-regulating features of the class A circuit, which can tolerate large supply-voltage variations because chain reactions cannot occur.

Component Layout and Precautions

The physical construction of nuvistor amplifiers is shown in Figures

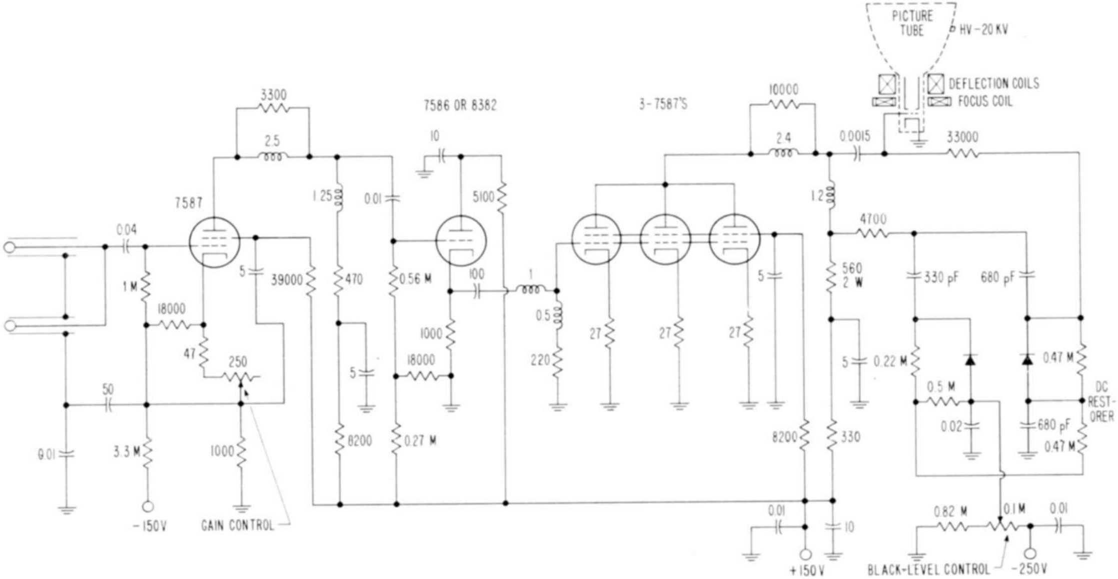


Fig. 8—Schematic diagram of 60-megacycle picture-tube drive amplifier using five nuvistors. (Except where indicated, all capacitance values are in microfarads, all inductance values in microhenries, and all resistance values in ohms.)

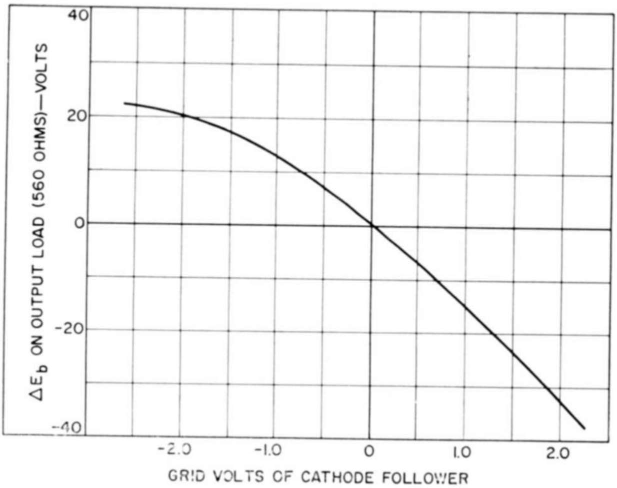


Fig. 9—Plate characteristics of 7587 nuvistor tetrode in circuit of Figure 8.

12 and 13. The inductances for interstage coupling are close-wound coils of No. 38 wire on 3/16 or 1/8 inch-diameter Lucite tubing; their values are computed as follows: $L = r^2n^2/(9r + 10l)$ microhenries, where r and l are the coil radius and length, respectively, in inches, and n is the number of turns. The radius, r, should be selected so that

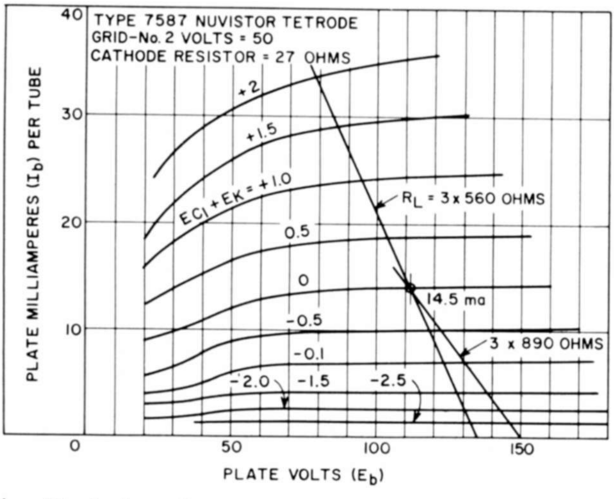


Fig. 10—Amplitude transfer curve of cathode-follower and driver stages in Figure 8.

l is approximately $2r$, and a length of wire approximately equal to the connecting lead length should be unwound from the coil to compensate for the lead inductance. Connecting leads should be as short as possible. The 150-volt miniature electrolytic bypass capacitors are stripped of their insulating case and mounted with brass pressure flanges directly on the chassis to eliminate the ground-lead inductance. These capacitors have a series resistance of approximately 0.4 ohm at 30 megacycles and are adequate for screen-grid and plate-circuit bypassing. Electrolytic cathode-bypass capacitors are shunted by 0.04-microfarad ceramic capacitors (smaller values would form tuned trap circuits with the lead inductances). One heater connection is grounded at each socket; the other is brought through a hole near the pin to the opposite side of the chassis and bypassed at the shielded feeder cable connection by a ceramic 0.01-microfarad capacitor. Power-supply lines or cables from remote power sources feeding individual amplifiers and associated circuits must be r-f bypassed at the amplifier feed points to prevent r-f leakage and regeneration over supply lines.

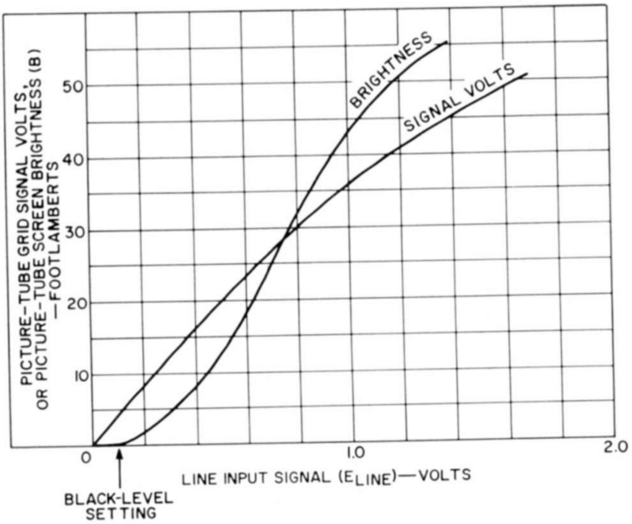


Fig. 11—Signal-transfer curve of three-stage amplifier shown in Figure 8, including d-c restoration.

The preamplifier in the television camera is enclosed in a metal compartment to prevent local capacitive r-f feedback over multiplier supply leads into its input. The first cascode tube is mounted close to the camera-tube socket on a wide ground plate to which all tube and

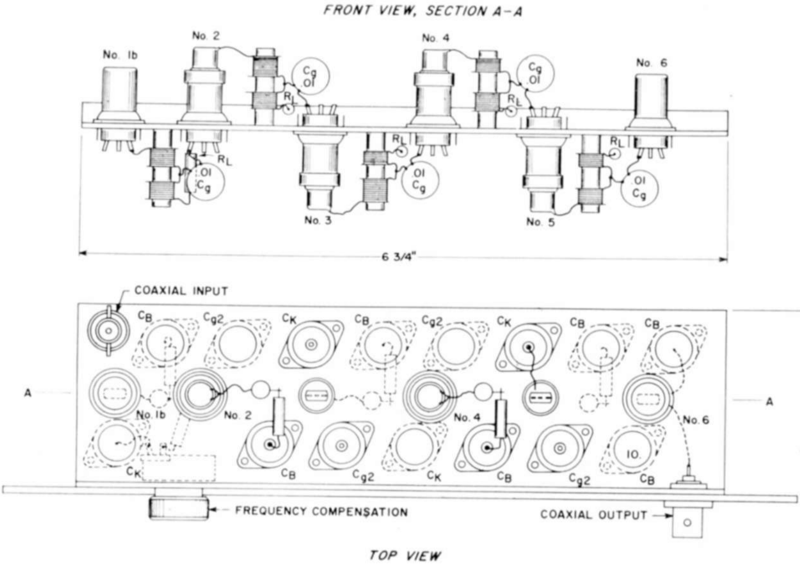


Fig. 12—Component layout of camera amplifier.

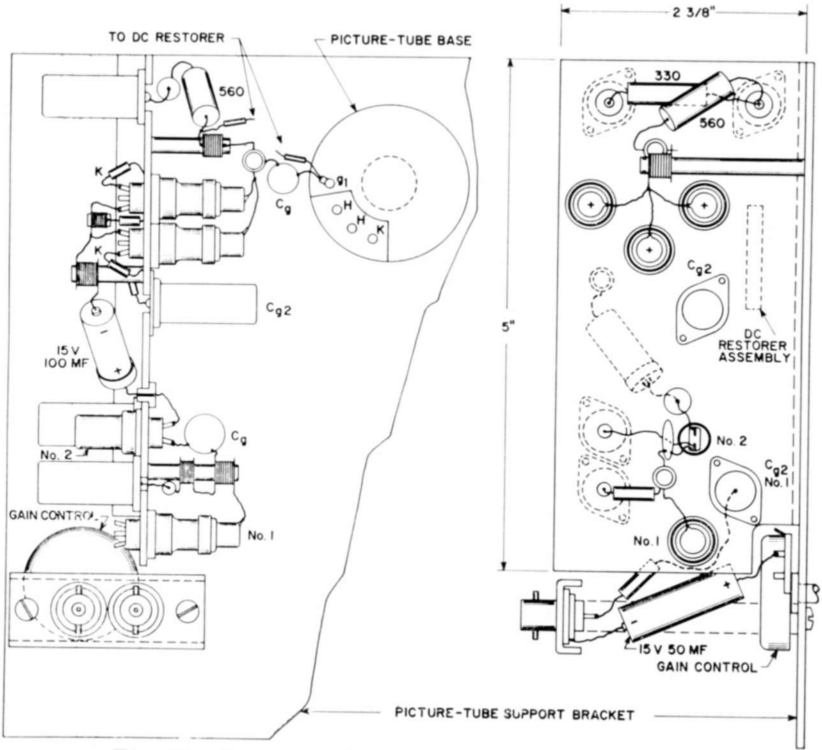


Fig. 13—Component layout of picture-tube amplifier.

multiplier electrodes terminating at the camera-tube socket are bypassed. Aside from the shield of the heater cable and the outer conductor of the short 90-ohm coaxial cable, the cascode input tube is thus grounded to the amplifier through the low-reactance ground plate that supports the box shield enclosing the camera-tube socket. This arrangement is necessary because normal shielded cables between the camera-tube socket and the resistor board containing bleeder circuits and low-frequency bypass capacitors become chokes at 60 megacycles.

The B-supply voltages of the amplifier are not critical because the circuits are self-adjusting. A larger decrease (to 100 volts, for example) merely causes a reduction in gain and obtainable power output that can be restored in most stages by use of smaller decoupling resistors.

CONCLUSIONS

It has been demonstrated that reliable low-noise wide-band video systems that have bandwidths of at least 60 megacycles and satisfy all the requirements of high-quality, high-definition television systems can be built with a small number of nuvistor tubes. The uncritical operating conditions, stability, and tolerance to high surge voltages and temporary overloads of nuvistor amplifiers are desirable properties for many applications. The gain–bandwidth product (for constant response) of a practical nuvistor tetrode stage has a value of 286 megacycles, and the gain/delay ratio is a maximum when the gain has the value 2.718. The optimum gain of a nuvistor stage is obtained with a flat-range bandwidth Δf of 105 megacycles and a minimum time delay per stage, $\beta/(2\pi\Delta f)$, of approximately $1/(2\Delta f)$ or 5.0 nanoseconds. The phase delay, β, of a ten-stage amplifier is thus only 90 degrees at a bandwidth of 10 megacycles. This performance demonstrates that stable negative-feedback amplifiers having a gain of 80 db and a bandwidth of 10 megacycles can be realized with nuvistor tubes.

PLASMA ACCELERATION BY ELECTRON CYCLOTRON RESONANCE*

BY

H. HENDEL, T. FAITH, AND E. C. HUTTER

RCA Astro-Electronics Division
Princeton, N. J.

***Summary**—Results are presented for a continuously operating electron cyclotron resonance plasma accelerator at a frequency of 2.45 kilomegacycles with absorbed power levels up to 0.5 kilowatt. Langmuir-probe measurements indicate that a space-charge field set up by energetic electrons (electron temperatures from 5 to 25 ev) accelerates plasma ions to energies of from 35 to 150 ev, as measured by a retarding-field ion-energy analyzer. The final ion energy shows a linear relationship to the electron temperature, in agreement with the theory. Strong electron confinement along magnetic field lines is observed in the discharge region. At powers up to 10 watts, approximately 30% of the absorbed microwave power is converted into kinetic energy of the ions when a magnetic bottle is formed at the rear of the accelerator. When a similar bottle is used with a power of 0.5 kilowatt, the efficiency falls to 4%. It is believed that this reduction is due to strong excitation in the magnetic bottle region, where electrons are connected by the magnetic field lines to the waveguide walls. The fact that low- and high-power efficiencies have been found comparable for other magnetic field shapes indicates that an accelerator designed to exclude field-line connections between wall and plasma should yield efficiencies of at least 30% at the higher power levels. The specific impulse, efficiency, and thrust attainable from a thruster based on electron cyclotron resonance are in the useful range for space-propulsion applications, especially in orbit-corrective missions. The absence of sensitive accelerating or electron-emitting electrodes promises long lifetime for such a thruster.*

INTRODUCTION

STUDIES carried out at RCA[1,2] and elsewhere[3,4] since 1962 have demonstrated that space-charge acceleration of ions under conditions of electron cyclotron resonance makes possible the effi-

---

* This work was supported by The AF Office of Scientific Research of the Office of Aerospace Research, under Contract No. AF49(638)-1342. The results reported here were presented at the Fourth Symposium on Advanced Propulsion Concepts, sponsored by The AF Office of Scientific Research and United Aircraft Corporation, Palo Alto, California, 26 April 1965.

[1] H. W. Hendel and T. T. Reboul, "Continuous Plasma Acceleration at Electron Cyclotron Resonance," AIAA preprint 63001.

[2] S. A. Ahmed and H. W. Hendel, "Space Charge Acceleration of Ions at Electron Cyclotron Resonance," AIAA preprint 64-24.

[3] D. B. Miller, "X-Band Magnetic Resonance Plasma Accelerator," AIAA preprint 64-702.

[4] R. Bardet, T. Consoli, and R. Geller, "Acceleration Continue de Plasma a Gradients de Champs Electromagnetique et Magnetique Statique. Resultats Experimentaux," *Phys. Rev. Ltrs.*, Vol. 10, p. 67, 15 May 1964.

cient transfer of microwave energy to directed plasma velocity energy. This method of plasma acceleration offers attractive possibilities for space-vehicle thrusters because erosion and sputtering problems are minimal. Devices of this type have indeed been run for more than 1500 hours, with no noticeable deterioration.

The general characteristics of space-charge acceleration of ions by electron cyclotron resonance have been described previously.[1,2] Efficiency values (ratio of beam kinetic power to r-f power absorbed by the discharge) of approximately 30% and exhaust velocities above 10^6 cm/sec were reported with a mercury propellant for power levels up to 10 watts.

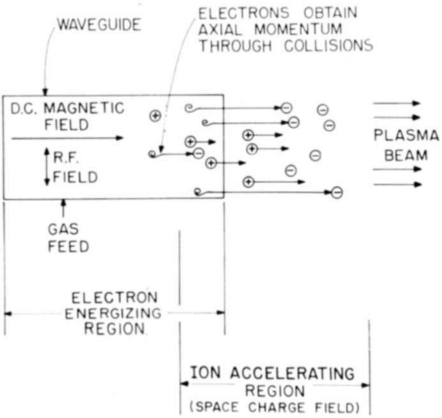


Fig. 1—Schematic diagram of electron cyclotron resonance plasma accelerator.

The work discussed here is primarily concerned with investigations of the basic physics of the acceleration process at power levels up to 500 watts. A study of the engineering use of satellite station-keeping thrusters of this type indicates that the inherent long life of such devices may make them especially suitable for use in synchronous satellites.

GENERAL BACKGROUND AND THEORY

Figure 1 is a schematic diagram of a simplified acceleration apparatus. Cold propellant gas is admitted into a section of S-band waveguide that is excited in a mode such that the electric vector of the r-f field is transverse to the axis of the waveguide and in the plane of the paper. A magnetic field, B, is applied in the axial direction of the waveguide. One end of the waveguide (left-hand side in Figure 1) is closed to prevent the escape of the propellant; the other end has an opening through which the accelerated plasma escapes.

The driving r-f frequency, f, and the magnetic field, B, within the discharge region (also called the electron-energizing region) are adjusted to be at the condition for electron cyclotron resonance $B_r = f(2\pi m/e)$. The gas pressure is adjusted so that the electron mean free path parallel to the magnetic field (as determined by electron–ion and electron–neutral collisions) is many times the Larmor radius of the electrons orbiting about a magnetic field line. Typical values for the conditions of the experiment are $f = 2.45$ gigacycles, $B = 870$ gauss, ν (electron collision frequency) $= 5 \times 10^7$ collisions per second, r (Larmor radius) $= 10^{-2}$ cm, and λ (mean free path parallel to B) $= 20$ cm.

Under these conditions, the gyrating electrons in the discharge region gain a relatively high average energy through r-f pumping,[5] their motion is eventually randomized by collisions with ions and neutrals, and they diffuse out rapidly from this region along the magnetic field lines. This diffusion process sets up a space-charge region that performs two major functions; (1) it accelerates the positive ions out of the waveguide, and (2) it reflects the lower-energy electrons back into the cavity. The higher-energy electrons penetrate this potential barrier and emerge with the ions to form a neutralized beam.

It has been shown in preivous work[1,2] that the following relation can be written for the discharge region:

$$\overline{W} = \frac{e^2 E_0^2}{4m\nu^2}, \qquad (1)$$

where $\overline{W}$ = mean electron energy gain between collisions,

E_0 = amplitude of transverse r-f electric field,

m = electron mass,

e = electron charge.

If values of $E_0 = 10$ volts/cm and $\nu = 5 \times 10^7$ collisions/sec are inserted into Equation (1), the mean energy gain, $\overline{W}$, will be 18 ev.

The ion-accelerating potential drop, V_d, across the self-generated space-charge region can be derived as follows. The ion current density, j_i, passing through the space-charge region is approximated by[6]

---

[5] B. Lax, W. P. Allis, and S. C. Brown, "The Effect of Magnetic Field on the Breakdown of Gases at Microwave Frequencies," *Jour. Appl. Phys.*, Vol. 21, p. 1297, Dec. 1950.

[6] D. Bohm, E. H. S. Burhop, and H. S. W. Massey, *Characteristics of Electrical Discharges in Magnetic Fields*, ed. by A. Guthrie and R. K. Wakerling, McGraw-Hill, 1949.

$$j_i = \frac{en}{2}\left(\frac{kT_e}{M}\right)^{1/2}. \tag{2}$$

The electron current density, j_e, penetrating to a point in the potential fall region V volts below discharge potential is given by

$$j_e = \frac{en}{2}\left(\frac{2kT_e}{\pi m}\right)^{1/2} \exp\left\{\frac{-eV}{kT_e}\right\}. \tag{3}$$

At the exit end of the space-charge region where $V = V_d$, j_e must equal j_i, and V_d is found to be

$$V_d = \frac{kT_e}{e} \ln\left(\frac{2M}{\pi m}\right). \tag{4}$$

where n = electron (or ion) density,
T_e = electron temperature in discharge region,
M = mass of ion (or neutral), and
k = Boltzmann's constant.

For mercury ions, $V_d \simeq 6kT_e/e$; i.e., the potential difference available for ion acceleration is about 6 times the electron thermal energy in the discharge region.

EXPERIMENTAL APPARATUS

A diagram of the experimental apparatus used is shown in Figure 2. Microwave power up to 2.0 kilowatts can be supplied at a frequency of 2.45 gigacycles and transmitted through a circulator to a mercury-plasma-filled region of the rectangular S-band waveguide. The function of the circulator is to prevent destruction of the magnetron due to sudden changes of the plasma load. The waveguide is generally terminated in a cutoff block with a cylindrical hole 2.4 cm in diameter, which allows the plasma to emerge but is beyond cutoff for the frequency used and of sufficient length (5 cm) to reflect r-f power incident on it.

A movable-stub tuner matches the impedance of the plasma load and thereby maximizes the r-f electric field intensity in the plasma. Incident and reflected power are monitored by a directional coupler, enabling absorbed power to be measured. A waveguide pressure window separates the input waveguide at atmospheric pressure from the plasma-filled waveguide at low pressure. Mercury vapor is supplied to the plasma acceleration region from a heated reservoir. To achieve magnetic-field configurations of varying extent and mirror ratios up to 2.4, three solenoids at variable distances are used. No external d-c fields are applied in the accelerator and no electron emitter is employed. Thus, the r-f drive both ionizes the gas and delivers energy to the electrons.

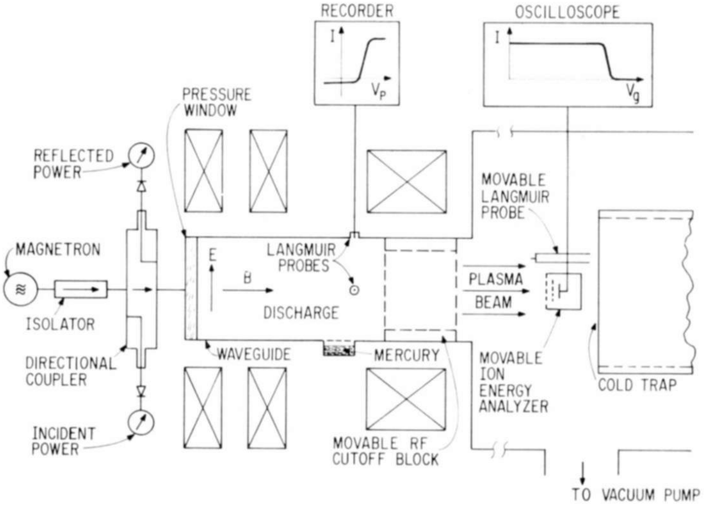


Fig. 2—Schematic diagram of apparatus.

The diagnostic equipment includes axially and radially movable Langmuir probes and retarding-field ion-energy analyzers. These analyzers, indicated schematically in Figure 2, consist of an ion-retarding grid and an ion-collecting disc housed in a Faraday cup. A sawtooth voltage (zero to +140 volts) is applied to the retarding grid; the ion collector is biased at −180 volts. Ion-collector current, I, versus retarding grid potential, V_g, is displayed on an oscilloscope (Figure 2). Some of these probes can be introduced through the cutoff block into the plasma-filled region of the waveguide; they can also be used at any point in the exhaust beam. Throughout the low-power experiments, sets of five fixed Langmuir and r-f probes were used in both the top and side walls of the plasma-filled waveguide. A scintillation counter with pulse-height analyzer monitored intensity and energy distribution of x-rays produced by run-away electrons observed under certain conditions.

The operating pressure in the expansion region was 2×10^{-5} torr or less.

EXPERIMENTAL RESULTS

To obtain as detailed an understanding as possible of the mechanism of the plasma acceleration and to verify in more detail the theoretical equations found to be approximately correct in the earlier work, many measurements have been made of plasma potential, floating potential, electron temperature, and plasma density. These measurements have been made in both the discharge and beam regions of the accelerator and at both low power levels (up to 10 watts) and high power levels (up to 500 watts). Measurements have also been made on beam ion current and energy distribution.

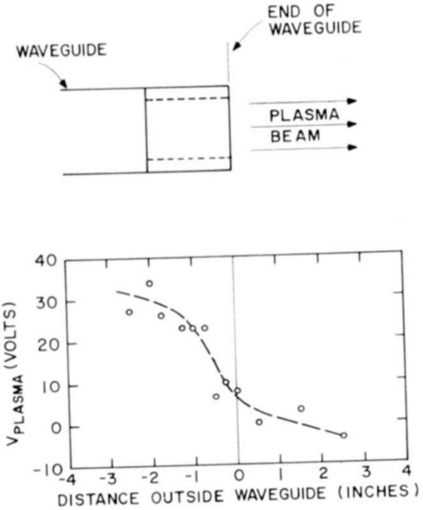


Fig. 3—Axial plasma potential distribution.

Low-Power Measurements

These measurements were carried out in c-w operation at absorbed r-f power levels up to 10 watts.

Plasma Potential Distributions

A movable cylindrical Langmuir probe 0.125 inch long and 0.040 inch across provided plasma potential, electron temperature, and density data along the axis of the discharge. As seen in Figure 3, the measured axial plasma potential in this experiment showed a total drop of 35 volts across the region traversed by the probe. The actual potential drop to a region further outside the waveguide will be considerably greater—about 50 volts. The fact that the probe measurements result in a slightly lower potential drop than was calculated is explained as due to the presence of the probes. This argument is strengthened by the better agreement of final ion energy (measured without the disturbing effects of probes in the space-charge region) and theory. Although most of the measured drop is shown as occurring at the cutoff block, experiments indicate that the block is not necessary to sustain a potential drop.

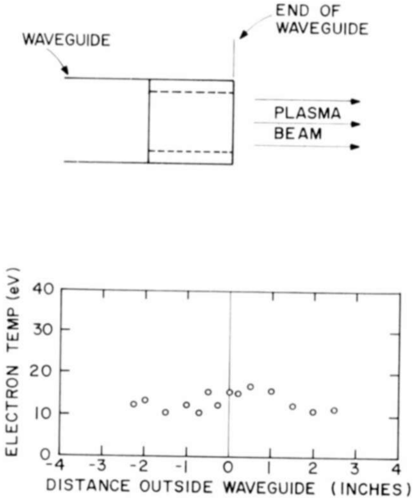


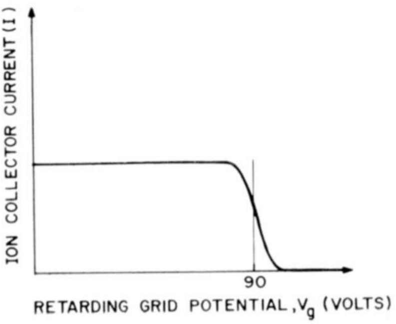
Fig. 4—Axial electron temperature distribution.

Electron Temperature

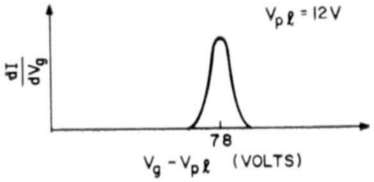
Electron velocity distributions were obtained from the portion of the Langmuir probe I_- versus V_p plot (I_- is the electron current to the probe, V_p the probe potential) slightly below plasma potential. Corrections for ion-probe current were made in the discharge region, where the ratio of electron to ion saturation currents was much smaller (~ 10) than it was in the beam region of lower magnetic field (~ 100). The linearity of the corrected $\ln(I_-)$ versus V_p curves in the region of interest showed that the electron velocity distribution was approximately Maxwellian. The measured electron temperatures, T_e, are shown in Figure 4. It can be seen from Figure 4 that T_e is nearly constant throughout the accelerator.

Beam Ion-Energy Distribution and Efficiency

The ion-energy analyzer outlined in Figure 2 was used to measure beam ion current and energy distribution at a cross section 10 to 14 inches "downstream" from the end of the waveguide. Five individual collectors covered radial positions from zero to 3 inches off axis; the ion flux and energy were fairly constant over this radial range. The analyzer trace (ion current to the negatively biased ion collector versus retarding-grid voltage) shown in Figure 5 (a) represents a typical result; the energy distribution is obtained by differentiating this trace and plotting against $V_g - V_{pl}$, i.e., the potential difference between ion-retarding grid and plasma, and is centered at 78 volts as shown in Figure 5 (b). This value is approximately 6 times the measured electron temperature as predicted in Equation (4).



(a) ION CURRENT vs RETARDING GRID POTENTIAL



(b) ION ENERGY DISTRIBUTION

Fig. 5—Ion energy distribution.

Figure 6 shows a comparison between electron temperatures at the walls of the waveguide and the beam ion energy. The expected linear relationship was measured. In making these runs, the electron temperature was changed by varying the input r-f power and propellant

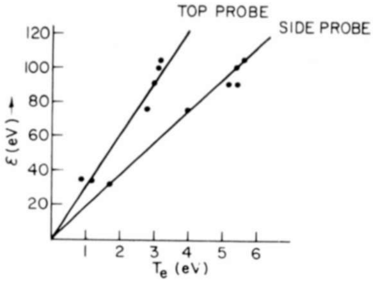


Fig. 6—Final ion energy versus electron temperature.

flow. Electron temperatures at side and top walls were, respectively, about 1/3 and 1/5 the electron temperature in the center of the discharge.

Engine efficiency (the ratio of beam kinetic power to r-f power absorbed by the discharge) was calculated from ion-energy-analyzer data and absorbed-power measurement for three magnetic-field configurations shown in Figure 7. In Figure 7(c), where the rear magnetic field strength was raised to form a magnetic bottle with a mirror ratio of 2.0, the maximum measured efficiency of 30% was obtained.

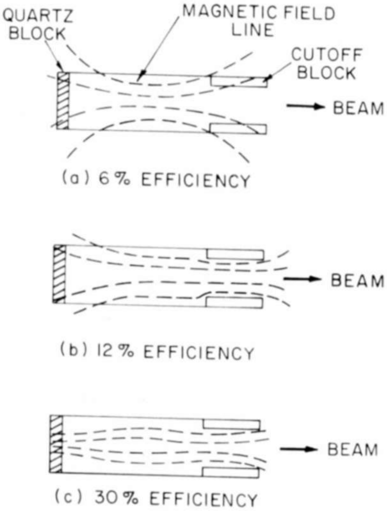


Fig. 7—Effect of magnetic-field configuration on efficiency.

High-Power Measurements

The most recent experimental work was carried out at absorbed r-f powers ranging up to 500 watts, again in c-w operation. Small plane Langmuir probes (0.001 to 0.010 inch in diameter) and a movable ion-energy analyzer having a 0.75-inch-diameter Faraday cup were used for these measurements. The pressure in the beam region was kept below 2×10^{-5} torr.

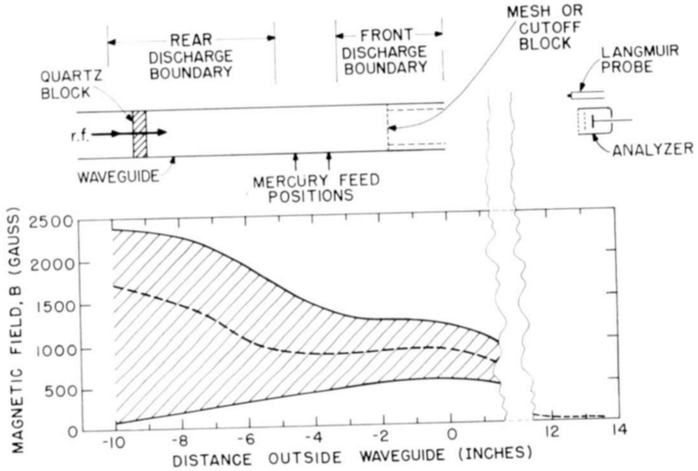


Fig. 8—Range of accelerator parameters in measurements of efficiency at high power.

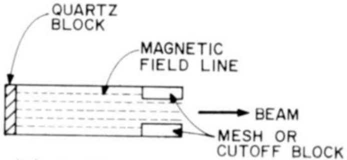
Efficiency Measurements

Figure 8 shows the range of accelerator parameters covered in the course of this work. The principal diagnostic tools were the Langmuir probes and the ion analyzer mentioned above and shown at the right of Figure 8. The discharge boundary adjacent to the beam was defined either by a grounded cutoff block, by a grounded mesh, or, less effectively, by the open end of the waveguide. A quartz block defined the rear-discharge boundary. Both boundaries were movable over the range indicated by the arrows at the top of the illustration. The range and distribution of magnetic field intensity are indicated in the shaded area of Figure 8; most of the experiments were conducted with a field distribution similar to that shown by the dotted line.

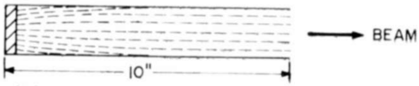
Some of the engine configurations tested are outlined in Figures 9(a)-(c). As can be seen, the efficiencies for (a) and (b) were very low and, although 11% efficiency was reached in arrangement (c), the

30% achieved at lower powers has not been reached at 500 watts in the experiments to date. Some of the reasons for these lower efficiencies are presented below.

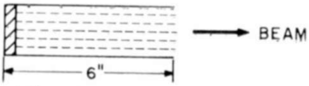
In the configurations shown in Figure 9(a) and (b), grounded metal walls intercepted a large number of magnetic field lines. In both configurations short-circuiting of electrons along the magnetic field lines to these walls contributed significantly to the reduction of engine efficiency. The effect shown in (a) is clear. Field lines from the resonant region are intercepted by the block or mesh. In (b), however, the only lines that intercept the waveguide walls are those connected to the bottle ($B > B_r$) region. It is possible to explain losses in this configuration by assuming either that considerable plasma excitation occurs due to nonlinear effects in the bottle region or that electrostatic fields in the bottle region enable electrons to cross the magnetic field lines. The latter condition appears less probable since good magnetic confinement occurs, as described later. The observation of 30% efficiency at lower powers with a field configuration similar to that of (b)



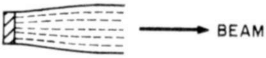
(a) 2-4% EFFICIENCY



(b) 4% EFFICIENCY



(c) 11% EFFICIENCY



(d) APPROXIMATION TO OPTIMIZED ENGINE

Fig. 9—Efficiency for various engine configurations.

is not inconsistent with the former explanation. At the lower powers the excitation region might have been confined to the resonant region, whereas at the higher power the area of excitation appears to have extended well into the bottle region, thus creating electrons that followed the field lines to the walls without transferring their energy to axial ion energy. The range of experiments at the two power levels indicates, however, the importance of shaping the accelerator walls to the magnetic-field contour.

Another factor reducing the efficiency in Figure 9(b) is the length of the chamber. The ion wall losses increase with the total wall area and thus with the waveguide length. Thus, an engine using this principle should be as short as possible. A configuration that is believed to represent a first approximation to an optimized engine is shown in Figure 9(d); it is anticipated that efficiencies up to the 30% previously observed at low power can be achieved in the region of 500 watts r-f absorbed power.

Discrimination between a Plasma Beam and a Stationary Plasma

It had been realized for some time that ion-velocity measurements obtained by an ion-energy analyzer could be in error if the measured ion velocity were due to the potential drop experienced by an ion while traversing a sheath around the analyzer rather than by the basic ion-accelerating action postulated. To check this point the following experiment was performed.

First, with the accelerator running under normal operating conditions, an ion-energy-analyzer trace at a given point in the beam region was recorded together with an I_- versus V_p plot from an adjacent Langmuir probe. Then the same measurements were made (at the same location) with the system at a higher pressure (3×10^{-4} torr), so that the acceleration effect did not take place, and only a stationary thermal plasma was present. The ion-energy-analyzer trace obtained was almost identical to that for the low-pressure condition where a directed plasma beam has been assumed to exist.

The data from these experiments are presented in Figures 10 and 11, with the results of the plasma-beam and thermal-plasma conditions shown on the left and right, respectively, of the illustrations. When plasma potential differences are not accounted for, i.e., when the derivative of the analyzer ion-current trace is plotted against V_g, the apparent ion-energy distributions are centered at 70 and 63 ev—nearly equal values. When, however, the plasma-potential measurement is accounted for, i.e., when the derivative is plotted against $V_g - V_{pl}$, the thermal-plasma case and the beam case can easily be differentiated. Plasma potentials obtained by standard procedures[7] are indicated in (c) and (d) of Figure 10. It is seen that while the plasma potential in the vicinity of the analyzer is near ground potential in the case of the plasma beam, plasma potential is highly positive for the thermal plasma. The correct energy distributions are shown in Figures 11(c) and (d). A thermal plasma is clearly indicated in Figure 11(d), in contrast to the obvious beam nature shown in Figure 11(c). Hence when the plasma potential correction is applied, plasma beams and thermal plasmas may be easily distinguished.

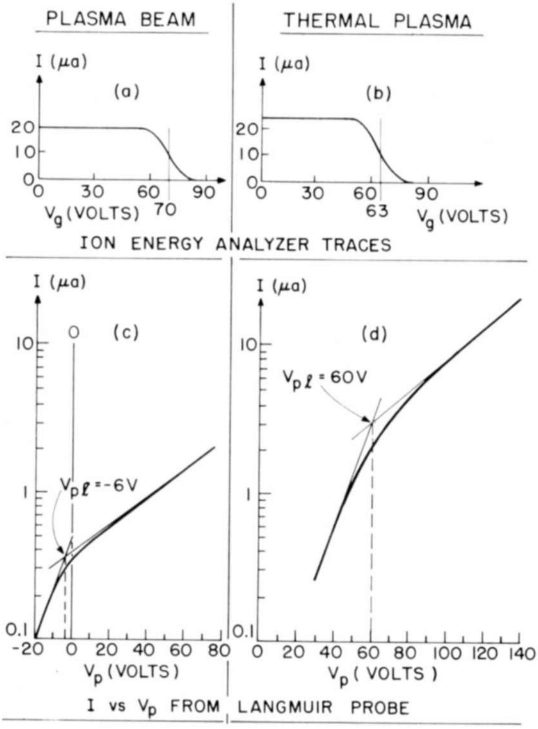


Fig. 10—Comparison of probe measurements from a plasma beam and a thermal plasma.

It should be emphasized that the misleading effects of an uncompensated difference between the potential of a measuring instrument and local plasma potential apply not only to the ion-energy analyzer,

---

[7] I. Langmuir and H. Mott-Smith, Jr., "Studies of Electric Discharges in Gases at Low Pressures, Part I," *General Electric Review*, Vol. 27, p. 449, July 1924.

but also to beam calorimeters, beam-thrust pendulums, and all similar plasma-beam detectors.

Supporting Experiments

A number of additional experiments not solely concerned with probing ion and electron properties within the beam were carried out and are described briefly below.

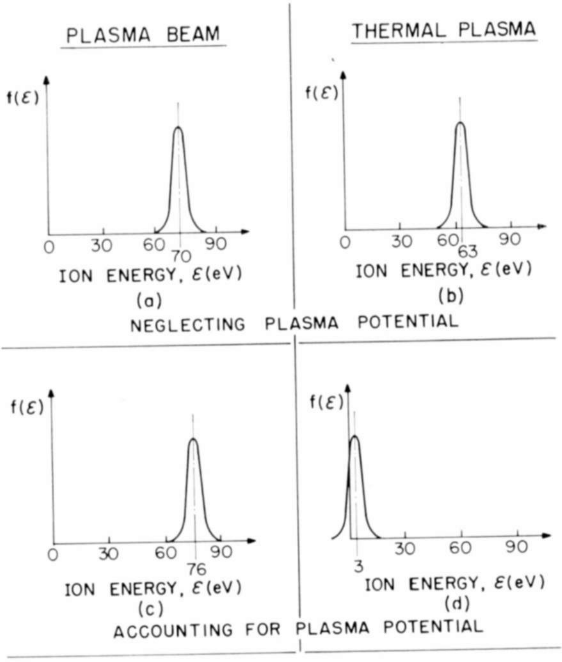


Fig. 11—Apparent and actual ion energy distributions.

High-Energy Electrons

As has been stated, only electrons with a velocity in excess of the mean thermal velocity in the discharge region can pass through the opposing space-charge electric field and issue along the beam of accelerated ions. Some information on the very-high-velocity electrons was obtained by scintillation-counter measurements of the emitted x-ray spectrum and comparison with Langmuir-probe electron-temperature measurements. Although the Langmuir probes indicated a Maxwellian distribution at 15 ev, high-energy electrons were measured at the same time with energy distributions around 20 kev and 70 kev. It was found that the presence of these "run-away" electrons could only be detected when the injected mercury vapor was pre-ionized. Such acceleration of electrons up to energies of a few million electron volts in electron cyclotron (fusion) heating experiments has been reported elsewhere.[8]

At a distance of 24 inches from the discharge, with a 0.12-inch-thick aluminum window, approximately 100 counts per second per cm^2 per kev were measured, with a maximum energy of 300 kev at 100 watts of absorbed microwave power. The density of these high-energy electrons in the plasma was found to be of the order of 10 electrons/cm^3, and the energy loss due to high-energy electrons is a small fraction of one watt.

Plasma Confinement along Magnetic Field Lines

Strong evidence of a high degree of magnetic confinement was found when an image of a mesh used at the front end of the accelerator appeared deposited upon the quartz block at the rear of the accelerator. The ratio of the size of the mesh image at the block to the size of the mesh itself checked very closely with the square root of the ratio of the magnetic field strength at the block to that at the mesh, as expected.

CONCLUSIONS FROM THE EXPERIMENTS

Experimental results for ion kinetic energy, electron temperature, and plasma-potential gradient are in good agreement with theoretical predictions based on an ion-accelerating mechanism employing a space-charge potential gradient set up by the energetic electrons. Continuous operation for many hundreds of hours has been achieved at absorbed power levels of up to 0.5 kilowatt.

Electron temperatures (up to 25 ev), ion energies (35 to 150 ev), and propulsive efficiencies (30% at low power and 11% at high power) have been measured. Apparently because of strong excitation in the magnetic bottle region at high absorbed power levels, electrons from this region, connected by the magnetic field lines to the wall, caused significant losses, and bottle configurations of the magnetic field yielded low efficiencies—about 4%. Experiments in a uniform magnetic field resulting in 11% efficiency indicate that efficiencies with a bottle field should be as high as previously measured (30%) if the waveguide is properly designed to exclude all magnetic field-line connection between wall and plasma.

---

[8] M. C. Becker, R. A. Dandl, H. O. Eason, Jr., A. C. Englund, R. J. Ken, and W. B. Ard, "An Investigation of Electron Heating at the Cyclotron Frequency," *Nuclear Fusion*, Suppl. Pt. 1, p. 345, 1962.

APPLICABILITY TO SPACE THRUST REQUIREMENTS

A thruster must be evaluated in relation to its potential mission. For this reason, studies have recently been made of the potential usefulness of an operating plasma engine based on the principles explained above and verified in the experimental program of the past year. In view of the still-present deep-space electric propulsion problem of obtaining an adequate long-life power supply with weight-to-power ratios lower than 10 kilograms per kilowatt, it is felt that the most likely near-term operational use of an electrical thruster will be for orbit-corrective missions with long required lifetimes. Boucher,[9] in a detailed comparison of various thruster systems for a station-keeping mission for a 1500-pound synchronous satellite, recently concluded that for an orbital lifetime in excess of approximately one year, high-specific-impulse electric-propulsion systems do offer distinct weight savings over low-specific-impulse devices (cold gas jets, chemical rockets, resistojets).

The principal system considerations are life expectancy and weight of the total propulsion plant. A thruster system using the principles of r-f electron cyclotron resonance will probably suffer a slight weight penalty when compared to an ion engine, for example, both because of its lower efficiency and because of the extra weight required for a d-c to r-f converter. If the total thruster system weight is small compared to that of the entire satellite system, however, this fact may be of less importance than the favorable life expectancy[10] of the thruster.

Table I summarizes typical characteristics that might be expected of a thruster system for a 3-year station-keeping mission for a 1500-pound synchronous satellite with the 24,500 lb-sec total impulse requirement considered by Boucher. Data are given for two possible propellants, sodium and mercury, and for two propellant utilization factors, 30% and 80%. No substantiated data have yet been taken on the propellant utilization of the r-f electron cyclotron plasma accelerator, but it can be noted that even the lower figure leads in the worst case to a total thruster system weight of 131 pounds, which is less than 10% of the vehicle weight. A comparable cold-gas system weighs approximately 300 lbs.

The lifetime expectations of this thruster are very good. The system has been run in the laboratory for over 1500 hours with no

---

[9] R. A. Boucher, "Electrical Propulsion for Control of Stationary Satellite," *Jour. Spacecraft and Rockets*, Vol. 1, p. 164, 1964.

[10] G. R. Seikel, D. B. Bowditch, and S. Domitz, "Application of Magnetic Expansion Plasma Thrusters to Satellite Station Keeping and Attitude Control Missions," AIAA preprint 64-677.

Table I—Anticipated Electron Cyclotron Resonance Thruster System for a 3-Year Station-Keeping Mission for a 1500 lb Synchronous Satellite; Total Impulse Requirement: 24,500 lb.sec.

Propellant	Na		Hg	
Thrust (lb)	$.46 \times 10^{-3}$		$.46 \times 10^{-3}$	
Beam power (watts)	26.5		9	
Specific Impulse (V_{exh}/g)	2950		1000	
Ion Energy (ev)	100		100	
Efficiency (%)	25		25	
Engine Power (watts)	106		36	
Solar Panel power (watts)	212		72	
Propellant utilization (%)	30	80	30	80
Thruster Component Weights (lbs)				
Thrusters (8)	16		16	
Propellant	28	11	82	31
Propellant tank	14	6	20	15
Power Supply Weights (lbs)				
Solar panel	7		3	
Power converter	15		10	
Total Thruster System Weights (lbs)	80	55	131	75
Life expectancy	Years			

visible deterioration. The electrodeless microwave discharge is ignited and maintained in the absence of any electron-emitting filament. Furthermore, since the ion-accelerating field is self-generated in the plasma by the energetic electrons, no ion-accelerating electrode is required. Some sputtering of the waveguide walls occurs, but this is a minor problem because the sputtering ions have relatively low energy (0 to 100 ev, depending on the discharge-plasma potential) and, therefore, a low sputtering yield. Also, waveguide dimensional tolerances are not stringent.

In conclusion, the electron cyclotron resonance accelerator described compares favorably with other station-keeping systems for those long-duration missions where thruster life is of particular importance.

RETARDING-FIELD ANALYZERS FOR THE MEASUREMENT OF AXIAL-VELOCITY DISTRIBUTIONS IN ELECTRON BEAMS

By

Martin Caulton

RCA Laboratories
Princeton, N. J.

Summary—*A study has been made of retarding-field analyzers for the measurement of axial-energy distributions in magnetically confined electron beams with energy spreads in the range of 1 to 50 volts. A retarding-field analyzer was developed incorporating a filter lens, in which the analyzing electrode does not intercept or collect the retarding current. This system has the advantages of ease of measurement, simplicity, and freedom from secondary-electron effects and surface charging. Lens and ion distortion effects were minor, and the potential depression due to space charge could be accounted for. Measurement capabilities were demonstrated on beams with a variety of velocity spreads.*

INTRODUCTION

IN THE course of a study of r-f wave propagation in magnetically confined electron beams with a large d-c velocity spread,[1-2] it became necessary to measure the axial velocity distribution of electrons in beams typical of microwave tubes. A retarding-field analyzer was developed that incorporated a filter lens in which the analyzing electrode does not intercept or collect the retarding current. This system proved to be suitable for the measurement of the axial energy distribution in the range from 1 to 50 volts of electrons in beams confined by a magnetic field.

Many methods can be used to measure electron velocity and energy. Among these are (1) retarding electric fields, (2) magnetic velocity selectors using mass-spectrometric techniques, and (3) electric and magnetic deflection with a cathode-ray-tube display. The study of a magnetically confined electron beam imposes specific restrictions that rule out the use of several methods. Transition fields are difficult to measure with method (2) and this, together with size limitations, eliminated this method from consideration. In the presence of a strong

---

[1] M. Caulton, B. Hershenov, and F. Paschke, "Experimental Evidence of Landau Damping in Electron Beams," *Jour. Appl. Phys.*, Vol. 33, p. 800, March 1962.

[2] T. G. Mihran, "RF Current Behavior in Electron Beams with DC Velocity Spread," *Jour. Appl. Phys.*, Vol. 33, p. 1582, April 1962.

magnetic field, method (3) would not yield large deflections. In view of these and other considerations, efforts were concentrated on retarding-field analyzers.

Retarding-field measurements may be considered textbook problems, but phenomena and associated difficulties often described only in individual publications (not in one textbook) were encountered and had to be solved in turn. For this reason the considerations of the phenomenon and associated problems leading to the development of the filter lens are reviewed in Appendix I. Some general solutions are discussed in Appendix II. The assumptions made in the development of the filter lens are partially explained in the Appendices. The review may be helpful to workers interested in the problem of retarding-field measurements, even though they may not be concerned with the specific solution described here.

RETARDING-FIELD CONSIDERATIONS IN AXIAL VELOCITY DISTRIBUTION MEASUREMENTS

Boersch[3] describes precise measurements of electron energy distributions in high-energy electron beams (not confined by a magnetic field) using a modification of the basic retarding-field analyzer. In 1961 Simpson[4] published a study of retarding-field energy analyzers; his paper contains a bibliography of much of the work published prior to that date. In the present paper we are mainly concerned with one-dimensional axial velocity distributions, and some of Simpson's conclusions applying to total energy measurements are not applicable. The basic retarding-field analyzer is shown in Figure 1. The electron energy distribution can be determined by a measurement of the collected current, i, versus the retarding collector voltage, V_R (with respect to cathode).

For an arbitrary axial electron velocity distribution $f(u)$, the current i ideally collected by the analyzer is given by

$$i = e \int_{u_0}^{\infty} u f(u)\, du,$$

where u_0 is the lowest velocity that is collected by the analyzer. Ex-

---

[3] H. Boersch, "Experimentelle Bestimmung der Energieverteilung in Thermisch Ausgelösten Electronenstrahlen," *Zeit. fur. Physic,* Vol. 139, p. 115, Oct. 1954.

[4] J. A. Simpson, "Design of Retarding Field Energy Analyzers," *Rev. Sci. Instr.,* Vol. 32, p. 1283, Dec. 1961.

pressing u in terms of a voltage V (using $u = \sqrt{2(e/m)V}$), we have

$$i = \frac{e}{m} \int_{V_R}^{\infty} F(V)\,dV,$$

where $F(V)$ is the electron energy distribution and eV_R is the lowest potential that the electrons experience. For a given i versus V_R observation,

$$F(V) = -\frac{m}{e}\frac{\partial i}{\partial V_R}.$$

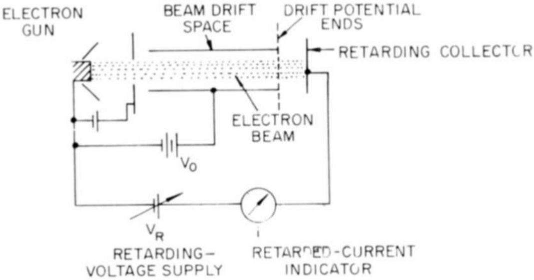


Fig. 1—Schematic of simple retarding-field analyzer.

Figure 2 shows this relation when $f(u)$ and/or $F(V)$ is a delta function. The rectangular distribution depicted in Figure 2(c) would ideally produce the constant slope i-V curve of Figure 2(d). If the distribution is a cutoff Maxwellian distribution as depicted in Figure 2(e), the resulting i-V curve would be (ideally) a straight line of slope $-e/kT$ on a semilog plot, as shown in Figure 2(f). The exponential nature of the i-V curve of the current from a thermal cathode has been observed over as much as 5 decades of current.[5] Note that a zero retarding voltage ($V_R = 0$) corresponds to the beam or drift voltage, V_0.

The square cutoff or the straight line semilog plots of Figure 2 are indeed ideal situations. More often, distortions of either of these shapes are observed. These distortions can be due to a variety of problems associated with retarding-field measurements.

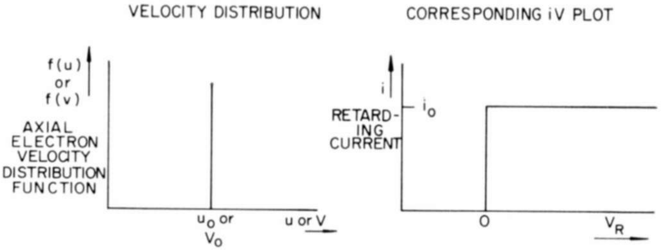
---

[5] H. Shelton, "Thermionic Emission from a Planar Tantalum Crystal," *Phys. Rev.*, Vol. 107, p. 1553, Sept. 15, 1957.

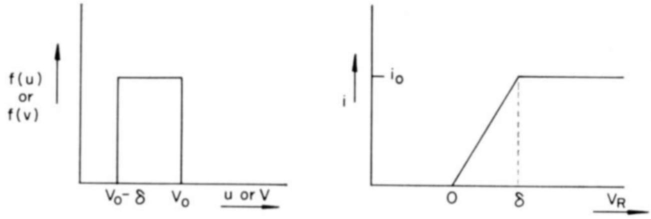
The Filter-Lens System

Filter-Lens Considerations

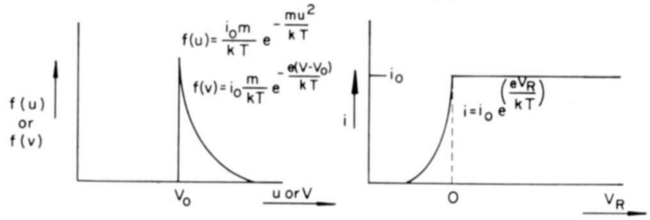
The development of a method to measure the axial velocity distribution of a magnetically confined electron beam with expected spreads of from one to fifty volts was the main objective of the study. In order to measure voltage spreads in this range, secondary electrons must be controlled, since they start to play an important role at these beam velocities. The basic retarding-field system of Figure 1 was not suitable for these measurements. Following the considerations described in Appendices I and II, a filter-lens system was developed to measure these voltage spreads.



(a) MONO-VELOCITY (b)



(c) RECTANGULAR (d)



(e) CUTOFF-MAXWELLIAN (f)

Fig. 2—Ideal velocity distributions and corresponding retarding-analyzer i-V plots.

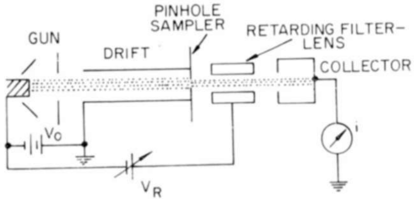


Fig. 3—Filter-lens system.

Figure 3 is a schematic representation of the filter lens. The retarding collector of Figure 1 is replaced by a hollow cylinder—the filter; this electrode retards but does not collect any electrons. The collector is a Faraday cage at high voltage that re-collects all secondary electrons. Lens distortions and secondary electrons at the collector do not change the *i-V* curves, since the constant-voltage collector measures only the current transmitted through the filter lens.

The filter-lens system suffers from two difficulties: (1) a lens effect can distort the voltage distribution, and (2) a space-charge potential depression is produced if the current density is high enough. Figure 4 illustrates the nonparallel equipotential lines and transverse fields that are present in the lens. If the magnetic field is large enough and the transmitted electrons are confined to a sufficiently small diameter, however, the electrons will cross equipotential lines that are predominantly planar, and lens effects will be negligible.

As shown in Figure 4, the voltage applied to the electrode only appears on the axis at a distance from the edge of the filter equal to

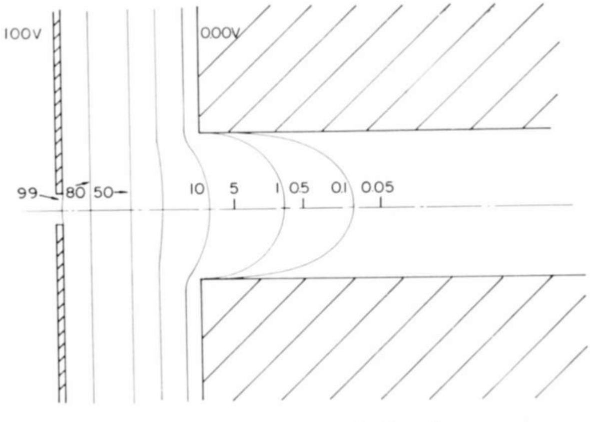


Fig. 4—Equipotential lines of filter-lens system.

the filter-electrode diameter. The equivalent planar diode spacing of this electrode system is greater than the actual electrode spacing. The length of the filter lens must be long enough for the electrode potential to appear on the axis. If the equivalent diode spacing is sufficiently great, an axial space-charge potential depression may occur (Appendices I, III). In addition, a potential depression can be formed across the beam if the space-charge is sufficiently increased.

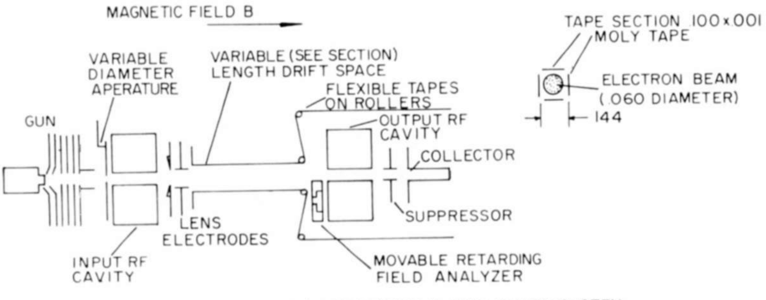
Lens effects can be reduced by making the lens-to-beam-diameter ratio large. This increases the possibility of both axial and transverse space-charge depressions. The lens-to-beam-diameter ratio must be determined from a consideration of maximum electron temperature excursions and the maximum current that is to be passed. It was empirically found that a lens-to-beam-diameter ratio of 2:1 would have little lens effect. However, the potential actually applied to the beam was the electrode potential plus a potential depression produced across the beam. The potential depression can be corrected for, as described later.

A serious problem in systems that are not bakeable is that of the charging of surfaces bombarded by electrons (Appendix I). The pinhole shown in Figure 3 picks up an insulating layer that is at a potential different from that applied to the electrode. This seriously disrupts measurements, as shown in Figure 14c. One way to remove this effect is to remove the pinhole as indicated in Figure 15(V). Without the pinhole the whole beam can be retarded by a filter lens. The lens can consist of any cylinder having a sufficiently large diameter and axial length so that the surface potential of the cylinder appears on its axis. The problem described previously applies. This system exhibits the phenomena of lens and ion (Appendix I) effects, and space-charge depression. The lens and ion effects, although detectable, are not too serious, as demonstrated later. The space-charge effects can be corrected for, as will be described.

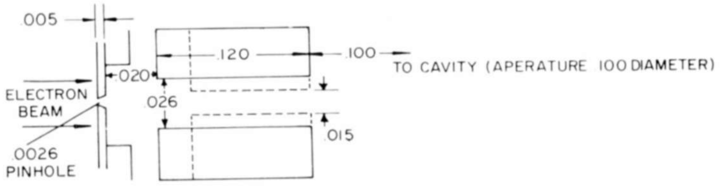
A solution for the pinhole charging that retains the beam-sampling properties of the pinhole is to electrically isolate the pinhole from the rest of the system as indicated in Figure 15(VI). The surface charging can be delayed by judicious control of the voltage of the electrons bombarding the pinhole. Surface charging may develop slowly, and if the primary electrons are of high energy, a conducting deposit may be formed. If electrons of low voltage are required to strike the pinhole, the pinhole voltage can be reduced while the experiment is proceeding and brought above 300 volts at other times. This technique proved very successful in this study.

Experimental System

Figure 5a is a schematic diagram of a demountable two-cavity klystron system in which the beam-velocity spread was analyzed in conjunction with a study of the electromechanical power propagated in the beam with various velocity spreads. The beam was confined by an axial magnetic field uniform throughout the length of the beam. The pinhole analyzer could be moved horizontally across the 0.060-inch-diameter beam, if desired, or removed from the beam. The second cavity, with a 0.100-inch-diameter aperature, served as the filter lens for the study of the whole-beam velocity spread (System V, Appendix II). Figure 5(b) shows the details of the pinhole analyzer, constructed in accordance with the considerations of the prevous section and Appendix I. It has short electrode spacings, compromise filter-to-beam-diameter ratio, and carbon-coated and electrically isolated pinholes.



(a) DEMOUNTABLE TWO CAVITY SYSTEM



(b) PINHOLE RETARDING FIELD ANALYZER DETAIL (DOTTED LINES SHOW ALTERNATE ANALYZER)

Fig. 5—Experimental apparatus.

The current–voltage retarding curves were displayed on an oscilloscope and recorded by photographs. Figure 6 indicates the power-supply system for the display. Two identical potentiometers were rotated synchronously to control the retarding voltage and the hori-

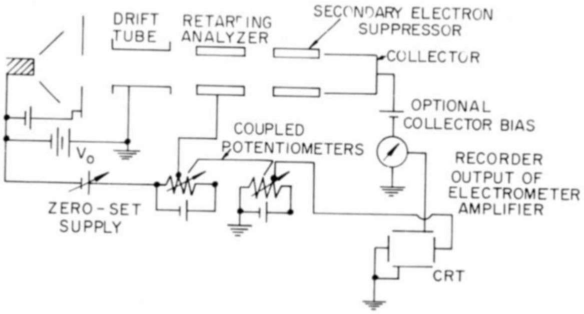


Fig. 6—Retarding-field display.

zontal-sweep voltage for the oscilloscope. The supplies swept the curves at 2 cps. A faster speed gave rise to a phase shift between the response read by the electrometer amplifier and the applied voltage; the phase shift then had to be corrected. Ion effects were more prominent at the low speeds, and the speed chosen is a compromise that depends on the individual problem.

RESULTS

Space-Charge Effects in Retarding-Field Systems

Figure 7 is the *i-V* retarding voltage curve of a 100-μa, 100-volt electron beam that, in the absence of any analyzer aberrations, should have displayed an ideal Maxwellian distribution cutoff. Figure 7 also shows the *i-V* curves of lower-current beams. The voltage at which cutoff begins is a significant function of beam current, which indicates that space-charge is causing a potential minimum. Examination of the geometry of the filter-lens system shown in Figures 4 and 5 suggests that space charge may build up to form an axial potential minimum due to the equivalent planar diode of the drift-space filter-lens system;

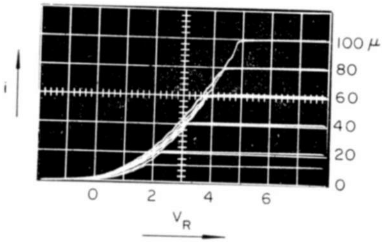


Fig. 7—Retarding *i-V* curves of 100-volt electron beam with minimum velocity spread.

such an axial potential variation would result in space-charge-limited transmission as in a thermonic diode. In addition, the filter lens is a cylindrical drift tube that, because of the transverse potential variation due to space-charge, can pass a current whose maximum limit depends on the drift potential and the ratio of beam diameter to tube diameter. The filter-lens system has both limitations, but that of the limiting-current drift tube is most important.

In Appendix III the results of Figure 7 are considered as caused by an axial potential depression. The data is viewed using Ferris' universal diode characteristics.[6] Ferris' analyses apply to a Maxwellian distribution of velocities and for electrons injected at zero potential. Although these conditions do not apply here, the analysis is carried out for the insight on how velocity spread may influence the space-charge potential depression. It is shown in Appendix III that the planar-diode limitation does not play a major role in our system.

When a beam is passed through any cylindrical drift tube at potential V_0, the potentials at the edge and at the center of the beam are reduced below V_0 as predicted by Poisson's equation. This reduction depends on the amount of current and the fraction of the tube that the beam fills. If enough current is passed, a virtual cathode will be formed, rearranging the current density and potential, thus limiting the amount of current that the tube will pass. Smith and Hartman[7] calculate the limiting current for beams partially filling the tube. For a 0.060-inch-diameter beam in a 0.100-inch tube (filter lens), they predict a maximum perveance of 14.6×10^{-6}. The data of Figure 7 showing the cutoff voltages for different injection currents have been plotted on two-thirds-power paper in Figure 8. The measured perveance is 12.7×10^{-6}. Other plots gave values ranging up to 15×10^{-6}. This is consistent with the theory. The measured contact potential of 0.9 volt also appears reasonable. In addition the very sharp initial cutoff of the current in Figure 7 conforms more closely to the discontinuous nature of the limiting current in a drift tube (the potential distribution shift when the critical current is reached). This discontinuity is sometimes very prominent under certain beam conditions. The perveances of the cutoff points for 0.040-inch- and 0.020-inch-diameter beams were determined (11.8×10^{-6} and 7.7×10^{-6}, respectively, with the same contact potential). The decrease is not as

---

[6] W. R. Ferris, "Some Characteristics of Diodes with Oxide-Coated Cathodes," *RCA Review*, Vol. X, p. 134, March 1949.

[7] L. P. Smith and P. L. Hartman, "The Formation and Maintenance of Electron and Ion Beams," *Jour. Appl. Phys.*, Vol. 11, p. 220, March 1940.

much as predicted from Smith and Hartman, but if the beam is not centered in the drift tube a difference is expected. The important consideration is that the change in perveance was not proportional to the beam-area change of 1/9, as it would be if the planar-diode limitation were the main consideration.

It is therefore concluded that the limiting current of the drift tube is the principal factor. The modification of the critical-current calculation to take into account a velocity distribution of electrons is not considered here, particularly since the exact conditions are determined by both the planar diode and the drift tube. Instead, an empirical study has been made, and a simple experimental determination of the space-charge correction is discussed next.

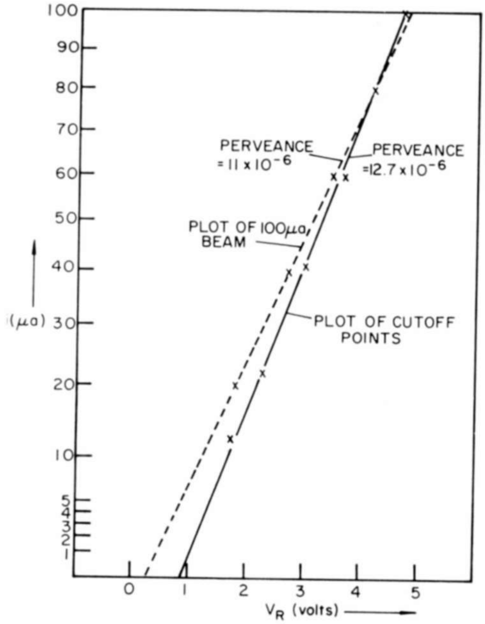


Fig. 8—Two-thirds-power plot of current versus retarding-field voltage (of Figure 7 data).

Experimental Space-Charge Correction

The filter-lens system produces a space-charge voltage differential that can mask the true velocity distribution, but the insight gained from the Ferris diode considerations implies that the differential may not appear with large velocity distributions. It was possible to produce in the system of Figure 5 a rectangular voltage distribution by velocity modulating the beam with r-f fields. This was done, and the resulting voltage spread was measured as a function of the r-f power, P_{in}, entering the input cavity. The peak r-f voltage, V_p, can be calculated from the expression

$$V_p = \sqrt{2R_0P_{\text{in}}}\,,$$

where R_0 is the equivalent shunt resistance of the cavity.

The ratio of the shunt impedance to the quality factor Q of the cavity was measured by plunger techniques as described by Ginzton.[8]

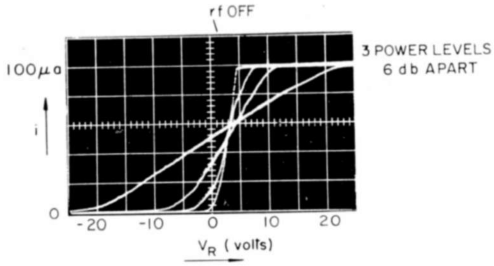


Fig. 9—*i*-*V* curves showing r-f modulation.

The Q of the cavity was independently measured by the impedance method.[9] The peak voltage calculated from these measurements for a given r-f input power was corrected for the beam–gap coupling coefficients[10] (radial and axial). The r-f peak voltage measured by the filter-lens system and the calculated voltage agree within 10%, which is on the order of the uncertainties of R_0/Q and Q measurements. It is concluded that the filter-lens system can accurately measure r-f voltage distributions.

The space-charge voltage differential can be corrected for by using the r-f modulation to produce a controlled voltage distribution in the beam. Figure 9 displays the effect of the r-f voltage swings on the $V^{3/2}$ current variation. Since $V_p = \sqrt{2R_0P_{\text{in}}}$, changing the power by a known amount should result in a predictable change in V_p (maintaining small-signal conditions). For this predictable change, knowledge of the exact value of V_p is not required, for if the measured V_p changes

---

[8] E. L. Ginzton, *Microwave Measurements*, p. 435, McGraw-Hill Book Co., Inc., N. Y., N. Y., 1957.

[9] See Ref. (8), p. 405.

[10] G. M. Branch, "Electron Beam Coupling in Interaction Gaps of Cylindrical Symmetry," *IEEE Trans. PGED*, Vol. ED-8, p. 193, May 1961.

as $P_{in}^{1/2}$, then V_p as measured is correct. In this way the power was systematically lowered until the voltage departed from a $P^{1/2}$ relationship. In Figure 9 the largest r-f power level produces a voltage cutoff of about 22 volts. Reducing the voltage modulation by 50% shows a cutoff of 11 volts, which indicates that the cutoffs are being read correctly. Reducing the voltage modulation by an additional 50% indicates a cutoff of 7 volts instead of the 5.5 volts expected. At this point, corrections due to space charge on the calculated V_p can be measured. This was done for the entire range of currents and beam sizes used in the experiments. In general it was found that if the applied voltage spread was approximately 1.5 times the space-charge voltage differential, V_D, at full beam current, the correction was negligible (V_D is the cutoff voltage in Figure 7).

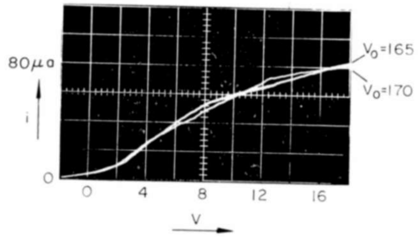


Fig. 10—*i-V* of large voltage distribution showing lens effects due to phase of scallop.

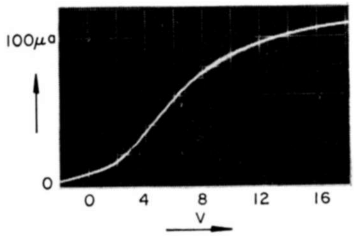
This correction is, of course, an approximation, since the superimposed r-f voltage distribution is rectangular while most of the distributions observed in practice have relatively sharp peaks and a velocity tail of slow electrons. The space-charge potential depression with such a tail is less than that of a rectangular distribution. The corrections observed with r-f are larger than that for a tail. The value of 1.5 V_D for no corrections is therefore conservative and on the high side.

Lens Effects

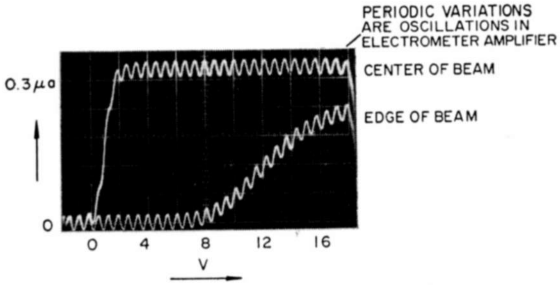
The whole-beam filter-lens system contributes lens effects (Appendix I). The analyzer-to-beam-diameter ratio is dictated by other considerations, and Figure 10 displays the lens effects on the *i-V* curves of a large voltage distribution. The irregularities move along the curve as the beam scallop entering the analyzer changes phase. This is evidenced only if the scallop wavelength is long (higher beam voltage and lower magnetic field). For short scallop lengths, such effects were not evident. Even with the irregularities, the voltage distribution does not appear to be appreciably distorted.

Electron Beams Influenced by Electric and Magnetic Fields

The large velocity spread indicated in Figures 10 and 11a was produced by an electrostatic lens (Figure 5) acting on electrons passing through aperatures having large potential differences. The outermost electrons travel through transverse electric fields, and the ensuing radial velocity is converted to rotational energy by the magnetic field. The resultant axial-velocity distribution of the beam varies with the initial transverse position of electrons inside the electrostatic lens. Figure 12 is the axial-voltage distribution of the *i-V* curves of Figure 11a. Most of the electrons are going slower than the electron beam voltage, V_0, and the faster electrons of the thermal spread are insignificant in number. This distribution is larger than those normally observed in magnetically confined electron beams, but it is more typical than the cutoff Maxwellian distribution, which is masked here.



(a) RETARDING iV CURVE OF WHOLE BEAM



(b) RETARDING iV CURVES OF PINHOLE SAMPLER OF BEAM OF (a)

Fig. 11—Retarding *i-V* curves produced by electric lens in magnetic field.

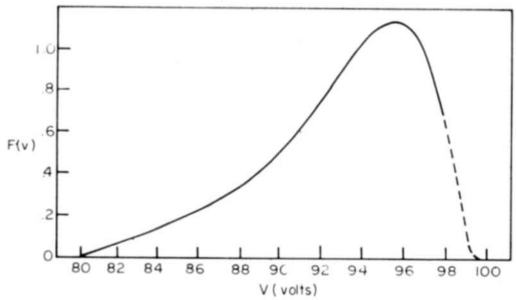


Fig. 12—Electron beam axial-voltage distribution produced by electric lens in a magnetic field.

Filter Lens and Pinhole Sampler

The electron beam can be analyzed by the isolated pinhole sampler described in Appendix II. Figure 11a displays the *i-V* characteristics of the whole beam with a large velocity spread; Figure 11b shows *i-V* waves with the pinhole sampler at the center and at the edge of the beam. The electrons at the edge of the beam are both slowed down (zero-current voltage shift) by 8 volts and spread out over 8 volts, which indicates good agreement with the whole beam spread of 16 volts seen in Figures 11a and 12. The center of a quiet beam seen by the pinhole analyzer has a sharper and lower-voltage cutoff than those shown in Figure 11b.

Ion Effects with Whole-Beam Analyzer

If the whole-beam system is operated with ions accumulating in the system, the electrons returning when the beam is cut off can lead to ion oscillations; these can impart additional energy to the primary electrons, yielding a greater amount of current at low and negative V_R. These ion effects appear if the V_R sweep is at a slow rate. Sweep rates of V_R as high as 20 cps were studied, and these showed negligible ion effects under most conditions. The presence of the effect is displayed in Figure 13. It is minimized in the lower curve by reducing the voltage of the beam and one of the electrodes (Appendix I).

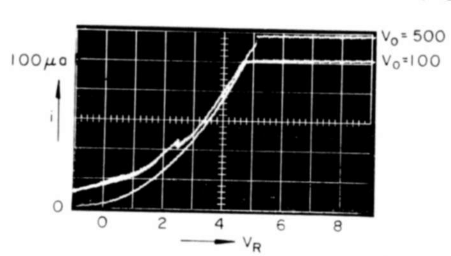


Fig. 13—Ion effects on *i-V* curves of cold beam.

CONCLUSIONS

For a variety of axial-velocity-distribution problems, there are many solutions and systems that will fit an experimenter's objectives. For the particular problem of velocity distributions corresponding to spreads from 1 to 50 volts, the filter-lens system, properly applied, offers many advantages. These are ease of measurement, simplicity of incorporation in any system, and freedom from secondaries and surface charging. The disadvantages are space-charge depression, lens effects, and ions if the whole beam is studied. Space-charge depression produced effects only for the smaller velocity spreads, and could be accounted for. Lens effects and ions do not overly distort the measurements.

It was demonstrated on beams with a variety of velocity spreads that the filter-lens system can accurately measure the axial energy distribution of magnetically confined electron beams.

ACKNOWLEDGMENTS

The substantial contributions of D. A. Daly, who participated in the early work developing the measuring systems, are gratefully acknowledged. The suggestions and criticism gained from conversations with A. L. Eichenbaum, L. S. Nergaard, P. T. Smith, and H. Sobol are very much appreciated. The author also wishes to acknowledge the work of O. Bekhor, who carried out many of the measurements described and the numerous members of the Microwave Research Laboratory Staff and Services who aided with these experiments.

APPENDIX I—SUMMARY OF RETARDING-FIELD MEASURING PROBLEMS AND SOLUTIONS

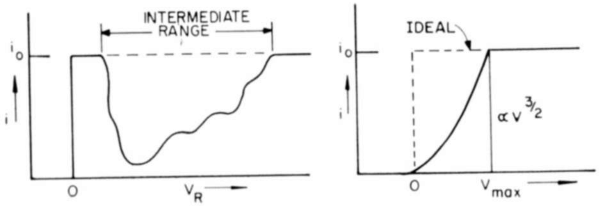
This Appendix describes often-observed phenomena leading to experimental problems. Table I lists the major problems associated with retarding-field measurements, with possible solutions to each problem. The third column refers to the systems of Figure 15, which are discussed later. The following comments apply to the numbered items of Table I.

(1) If secondary electrons are present the ideal square cutoff for intermediate collector voltages will be modified as shown in Figure 14a. Since most secondary electrons have low energy at collector voltages higher than depicted in Figure 14a, the total beam current will reach the collector. At low primary voltages (4 volts or less) the number of secondaries becomes negligible. If voltage spreads in the intermediate range are to be measured, secondaries will distort the

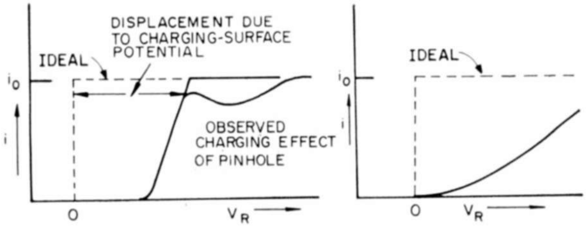
Table I—Retarding Field Measurement Considerations

Phenomenological Problems	Solutions	Systems to Correct (See Figure 15)
(1) Secondary electrons	(a) low collector voltage. (b) coat or construct with material of low atomic number. (c) Faraday-cage collector. (d) filter-lens system.	 II III
(2) Space-Charge Limitations	(a) short electrode separation. (b) low current density. (c) take potential depression empirically into account.	
(3) Lens Effects	(a) parallel-plate system. (b) beam diameter small compared with lens. (c) high magnetic field. (d) meshes or grids at high voltages. (e) if no magnetic field, then compensating lens possible.[4]	I IV
(4) Charging of Electrode Surfaces	(a) clean, bakeable system. (b) heat electrode above 250°C. (c) operate at high potentials (>300 volts). (d) construct system without current-intercepting surface. (e) alternate operation at low potential with maintenance at high potential. (f) a-c measurement.	 V VI
(5) Contact Potential	(a) measure. (b) all materials same, at same temperature, free of contamination.	
(6) Measurement Difficulties	(a) connection of electrometer properly, grounding. (b) shielding. (c) low-capacitance connections. (d) filter-lens systems.	 IV,V,VI
(7) Improper Sampling	(a) tubes restricting current area have small thickness compared with scalloping λ.	
(8) Ion Effects	(a) provide ion trap and/or sinks. (b) pulse system. (c) increase V_R sweep rate.	(System V can have ion effects)

curves. For voltage spreads greater than forty volts, the secondaries emitted are returned to the collector and have a negligible effect. As suggested in Table I, materials of low atomic number give fewer secondaries,[11] and coating the collector with a colloidal graphite solution is widely practiced. A Faraday-cage collector (depicted in System II of Figure 15) will also re-collect secondaries. However, in the presence of a magnetic field, the Faraday cage gives rise to other problems. The filter lens of System III, Figure 15, completely removes secondaries from consideration.



(a) SECONDARY EFFECTS (b) SPACE-CHARGE LIMITATIONS



(c) CHARGING EFFECTS (d) IMPROPER SAMPLING

Fig. 14—*i-V* curves corresponding to problems of Table I.

(2) A potential minimum or virtual cathode may form between the drift space and collector due to excessive space charge. The applied potential, V_R, will not be the reflecting voltage for such cases. The maximum current density, J_{max}, that can be passed between the drift plane and the retarding collector can be approximated by Child's equation;

$$J_{\text{max}} = 2.33 \times 10^{-6} \frac{V^{3/2}}{d^2},$$

---

[11] Spangenburg, *Vacuum Tubes*, pp. 49-57, McGraw-Hill Book Co., Inc., N. Y., N. Y., 1948.

where d is the effective separation between the electrodes. As suggested in Table I, shortening d or lowering the current density will alleviate this problem. Figure 14b illustrates the effects of space charge on the retardation of an ideal beam; the current i follows a $V^{3/2}$ relationship. As described in the text, space charge can be corrected for.

(3) Lens effects are produced by any analyzer distortion that causes electrons to shift direction. A combination of electric and magnetic fields can convert axial energy into rotational energy without changing the total kinetic-energy distribution. The study of axial velocity, rather than total energy, will show a distorted axial velocity distribution. The use of a parallel-plate system will remove transverse electric fields; a beam size small compared to electrode lens openings will minimize transverse-field effects on the beam. Very large axial magnetic fields confine the electrons in an axial direction. Meshes at high voltage can maintain a planar system, but the same mesh becomes a strong lens at low electron velocities.

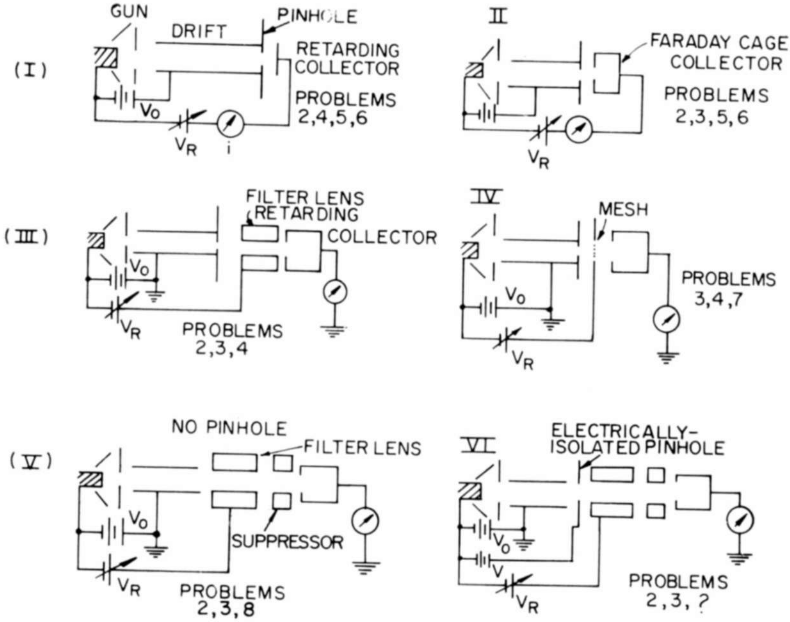


Fig. 15—Some retarding-field systems (problem numbers correspond to numbers in Table I).

(4) The charging of electrode surfaces can become severe in an unbaked system. If vacuum oils and greases are used there is always

a layer of their molecules on all surfaces, but this alone may not cause a charging problem. When electrons strike such a contaminated surface[12] the oil or grease can be broken down to a resistive or insulating layer. The surface of this layer can be at a potential different from that of the electrode. If the electron beam has sufficient energy, ions can be formed. Even at low pressures, surface charges can develop sizable potentials if ions are formed in an electric field. This problem may not exist in a clean, greaseless, and bakeable system. It is possible[13] to keep the resistance of the insulating films low by heating the electrode above 250°C and by maintaining the energy of the electrons striking the collector above 300 volts. The filter-lens system described in this paper keeps the retarding electrode out of the electron path. The retarding electrode is at a low voltage and a charging layer may be built up by the low-energy electrons bombarding it. It is possible to account for the resistive layer by a-c techniques, but an electrode can be kept from charging by keeping it at high potential except when the measurement is made. The potential of the charging layer will vary both with the energy and number of electrons; as a general rule,[11] if the potential is greater than the potential at which the secondary emission ratio is greater than unity, the electrode will charge to beam potential, while if the secondary emission ratio is less than unity, the electrode will approach zero potential. These factors can vary with position on the surface of the electrode, and Figure 14c illustrates the effects on observed *i-V* curves. The portion of the beam to be analyzed is often sampled by placing a movable pinhole across the beam. If the potential varies with transverse position due to surface charges on the sampler, lens effects such as portrayed in Figure 14c may result.

(5) Contact potential may only displace the cutoff. The contact potential must be measured if an absolute value of potential is required. If one of the surfaces has its work function changing while *i-V* curves are being taken, the measurements will be distorted.

(6) Measurement difficulties become apparent when currents are measured in the millimicroampere range. The major problem with the basic source-retarding collector of Figure 1 is that the retarding battery and the current meter are in the same series circuit. The electrometer must often be grounded, and its proper position in the circuit is dependent on which other electrodes are grounded. The resistance

---

[12] L. Mayer, "Photocontrol of Growth Rate of Thin Polymer Films Formed by Electron Bombardment," *Jour. Appl. Phys.*, Vol. 34, p. 2088, July 1963.

[13] K. M. Poole, "Electrode Contamination in Electron Optical Systems," *Proc. Phys. Soc.*, Vol. LXVI, Sec. B, p. 542, 1953.

of the battery retarding-voltage supply should be 10^{14} ohms or greater. Appreciable leakage to ground in the supply will often pass through the current meter, making low-current measurements difficult. In the course of the studies, a variable-voltage source was built, with every element checked for the introduction of leakage paths with an ultrasensitive microammeter. Proper shielding, incorporated to prevent stray pickup, may produce a large capacitance to ground, which limits the response time of the system. (At low current ranges the meter indicates all charging currents.)

The filter-lens system takes the collector meter out of the retarding supply circuit, making most of the precautions listed in the previous paragraph unnecessary. It is then quite simple to ground the collector through the current meter.

(7) It is often desirable to sample only a portion of the beam with a pinhole. Such a sampler can exclude electrons having a large transverse motion, particularly in the presence of a magnetic field.[4] The beam that is then studied will depend on the phase of its transverse rotation and the axial thickness of the pinhole.[14] The observed effects on the i-V curves with a pinhole that is too thick are indicated in Figure 14d. As V_R decreases, the cyclotron wavelength also decreases, causing more electrons to be collected on the tunnel in the pinhole electrode. To eliminate this problem it is necessary to keep the axial thickness of any sampling apparatus very much smaller than a scallop wavelength.

(8) If a significant number of ions are being produced, they can accumulate in low-voltage regions, particularly in the retarding analyzer. Ion oscillations can then impart extra energy to the electrons, leading to more electrons going faster than the thermal distribution predicts. System V of Figure 15 will produce ion effects in a low-pressure system under certain conditions. As pointed out, increasing the retarding-field sweep rate will minimize the effect of ions. If the faster sweep rate is not practical, the effects of ions can be minimized by providing a sink or trap for the ions. This can be accomplished by inserting a low-voltage electrode somewhere in the system, adjusting the voltage profile so that ions will drain out, or lowering the magnetic field to permit ions to spread to the walls. If the effects are severe, the system can be pulsed to allow the ions to drain away during off times.

APPENDIX II—RETARDING-FIELD SYSTEMS

In the course of the investigation, most of the problems described

---

[14] If the pinhole is thicker than a scallop wavelength it can be used to measure the distribution of transverse velocities in the beam.

in Appendix I were observed, analyzed, and solved using the principles listed in Table I. The basic system was modified for each problem, but each modification brought new problems. Figure 15 indicates six retarding-field systems that were studied; some of these are commonly used. With each system is a list of the problems of Table I that each may contain.

System I, similar to that of Figure 1, eliminates space-charge (2) by short electrode spacings, minimizes lens effects (3) by a small beam diameter, and eliminates sampling effects (7) by the use of a thin pinhole electrode. The other problems, however, cannot be completely eliminated. If large velocity spreads (greater than a few volts) are to be measured, and if the system is not extremely clean, surface charging will not permit application of known potentials.

System II can eliminate secondaries, but may introduce space-charge depression (effective diode spacing much larger than electrode separation) and severe lens effects; the retarding electrode of System II is a strong lens to electrons of low velocity.

System III contains the filter lens discussed in the text.

System IV is an alternative filter lens. The use of a mesh close to the drift space would appear to reduce lens effects and space-charge. However, low-energy electrons experience strong transverse forces due to the mesh wire and contribute to an even stronger lens effect. In addition, surface-charging causes the mesh openings to close up due to accumulated deposits, and after a relatively short operation time current transmission can fall by an order of magnitude.

Systems V and VI are discussed in the text.

APPENDIX III—LIMITATIONS OF A PLANAR DIODE

We consider here the implications of assuming that the potential minimum limiting the current is formed axially as in a planar diode. It will be deduced that the current limitations discussed do not follow those predicted by planar-diode considerations. However, it is instructive to follow the planar-diode limitations for a beam with a Maxwellian distribution of velocities.

Figure 16 illustrates the formation of a retarding potential minimum in a planar diode for electrons entering with velocities corresponding to an accelerating voltage V_0. In Figure 16a, when the injected current density reaches a critically high value I_1, the potential-versus-axial-distance curve exhibits a potential minimum at the anode or retarding plane. Figure 16b portrays the formation of a potential minimum, V_m, in front of the retarding plane; then V_m, not V_R, is the voltage that reflects electrons.

The Langmuir solutions for a space-charge-limited diode of electrons with Maxwellian velocity distributions have already been worked out. The space-charge retarding diode is similar, but the source electrons are injected at beam voltage and velocity, not at zero potential. Initially, let us ignore this difference to get some insight into the effect of a velocity distribution on potential minimum formation. Ferris[3] has presented the Langmuir solutions in a particularly simple and universal form. Figure 17 contains plots of normalized current ratios versus the reduced voltage $(\eta = eV/(kT))$ taken from Ferris. The area to the right of line AB is the emission-limited region; between AB and CD is the space-charge-limited region. To the left of CD is the Boltzmann, or retarding potential, region where a logarithmic response of I versus V would be observed. I_1 is the current at the intersection of the beam (or source) current I_s with CD; it is a critical current that places the potential minimum at the anode plane (Figure

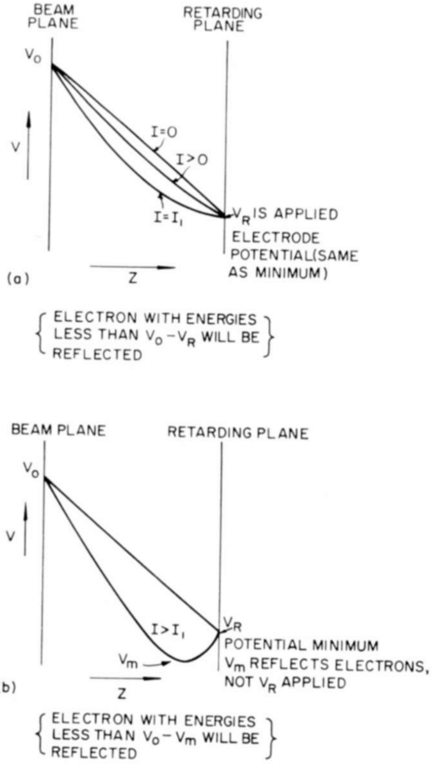


Fig. 16—Space-charge potential depression in diode.

16a). For currents and voltages less than this critical value, the *i-V* curves yield the true velocity distribution. If $I_s \gg I_1$, then I_1 will approach a value I_∞, (the horizontal asymptote of line CD). I_∞ would be the limiting value of I_1 if the beam current were infinite. It is the current flowing in the diode for infinite source current when the potential minimum is at the anode.

We are attempting to determine the effect of the known space-charge potential minimum on the observed velocity-spread measurements. Figure 17 illustrates this effect on the curves of a Maxwellian velocity distribution; distance $\Delta\eta$ in Figure 17 illustrates a correction at a particular value of current, with conditions of injected beam current $I_s = 5I_\infty$. If the parameters I_∞, I_s, and I_1 can be determined for the experimental system, corrections due to space charge can be calculated. Practical electron beams have velocity spreads that often mask the Maxwellian thermal velocity distribution, but an estimate of the effect can be made if it can be determined where on the universal curves the system is operating.

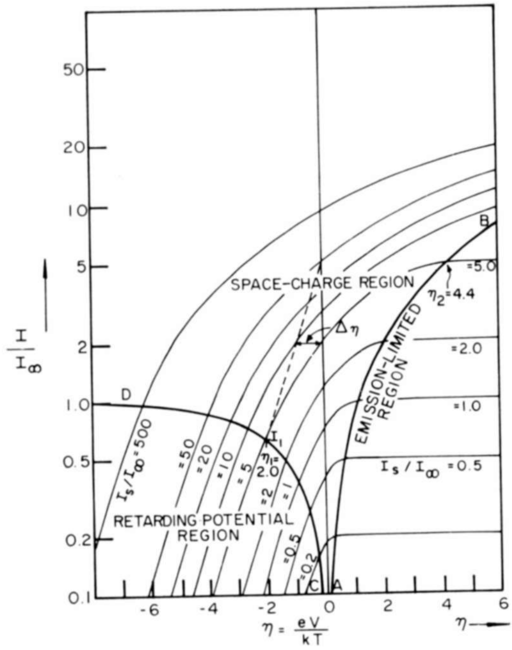


Fig. 17—Universal diode characteristics (from Ferris).

The data in Figure 7 can be used to calculate Ferris' parameters for the system of Figure 5. From Ferris,[6]

$$I_\infty = 0.1050\mathcal{P}\left(\frac{T}{1000}\right)^{3/2},$$

where $\mathcal{P}$ is the perveance $(I/V^{3/2})$ of the diode, and T is the absolute temperature. Figure 8 contains a two-thirds-power plot of the 100-μa current data represented in Figure 7; the slope indicates a perveance for the 100-μa beam of 11×10^{-6}. The $i = 0$ intercept of 0.25 volt appears to indicate the contact potential. The perveance would be less than this and the contact potential greater if the initial velocity-distribution effect on the diode conductance were taken into account.[15] If this correction is made, then $\mathcal{P} = 9.0 \times 10^{-6}$, $V_{\text{contact}} = +0.9$ volt. When we use the cathode temperature of 1600°K, we find $I_\infty = 2.0$ μa, and $I_s/I_\infty = 100/2.0 = 50$. Using this ratio to find the position on Ferris' curves, we obtain by calculation an emission cutoff of approximately 3.8 volts, and the start of the retarding region at approximately 0.6 volt. This appears to agree well with Figure 7 when the contact potential is taken into account $(V_{\text{cutoff}} = 3.8 + 0.9 = 4.7 \text{ volts})$.

Following Figure 17, for $I_s/I_\infty = 50$, $I_1 = 0.85\ I_\infty = 1.7$ μa (intersection of $I_s/I_\infty = 50$ with curve CD). This would signify that only at currents less than 1.7 μa would the i-V curves record the velocity distribution. To remove space-charge effects, I_s/I_∞ must decrease. If the velocity distribution is very large, representing a large equivalent temperature, T_{eq}, I_∞ will increase as $(T_{eq}/1000)^{3/2}$, and I_1/I_∞ on Figure 17 will approach unity. This signifies that the velocity distribution is so large that no correction of space-charge is necessary. Physically this is plausible because with a large velocity distribution (smaller number of electrons in each velocity class) the number of electrons turned back at a given voltage (thus removing its contribution to a potential minimum) would not be sufficient to cause a potential depression in the system.

This initial agreement of the data can be examined further. The estimated corrected perveance of 9×10^{-6} yields a diode spacing, d, as determined by

$$\mathcal{P} = 2.33 \times 10^{-6}\left(\frac{\text{beam area}}{d^2}\right).$$

Applying the known beam diameter of 0.060 inch, we calculate $d = 0.027$

---

[15] R. M. Matheson and L. S. Nergaard, "A Bridge for the Measurement of Cathode Impedance," *RCA Review*, Vol. XV, p. 485, Dec. 1954.

inch. The geometrical spacing is on the order of 0.100 inch. The discrepancy can be due to the presence of ions partially neutralizing the beam and shortening the effective diode spacing. Ion oscillations can also change the electron energy distribution and yield discrepancies. However, a large unexplained discrepancy is apparent.

Data similar to that of Figure 7 on beams with 0.040-inch and 0.020-inch diameters yielded perveance values that changed by an amount much less than is predicted from the perveance–beam-area variation. There is another discrepancy that indicates that inadequacy of the planar diode description. The consideration of limiting current injection points out that there may not be virtual cathode formation when the initial beam injection voltage is very high. The critical current[16] for virtual cathode formation can be expressed as

$$I_{\max} = \mathcal{P}(V_0^{1/2} + V_R^{1/2})^3.$$

With a beam voltage $V_0 = 100$ volts and a perveance estimated from known geometry, the beam current used is well below the maximum current so calculated. It was found that the data of Figure 7 did not change when the beam voltage was changed. It is concluded that the planar diode limitation does not play a major role in this system.

---

[16] R. Salzburg and A. V. Haeff, "Effects of Space Charge in the Grid-Anode Region of Vacuum Tubes," *RCA Review*, Vol. II, p. 336, Jan. 1938.

LOW-ENERGY-ELECTRON SCATTERING FROM PHOTOCONDUCTORS APPLIED TO A CAMERA TUBE*

By

A. D. Cope and W. F. Bruce

RCA Astro-Electronics Division
Princeton, N. J.

Summary—*Vidicons are extensively used for satellite-borne cameras because of their ruggedness, simplicity, and light weight. In many of these applications, information is read out at a relatively slow rate. Under these conditions, the performance of the tube can be substantially improved by using orthicon- or isocon-type scan. This paper describes work done in developing vidicon targets that give high performance in the isocon mode of operation. Tubes incorporating these new targets were found to have an illumination threshold two orders of magnitude lower than that of a standard vidicon, a signal-to-noise ratio comparable to that of an image orthicon, and a dynamic range ten times that of an image orthicon.*

Introduction

THE CONVERSION efficiency of television camera tubes is principally determined by two factors: (1) the quantum yield of the photosensitive surface and (2) the manner in which a video signal is derived from the charge image generated by the incident quanta. With respect to the first factor, both photoemissive and photoconductive layers are used as light-sensitive surfaces. Photoemissive surfaces have quantum efficiencies of 10 to 20% in the visible spectrum, while photoconductive surfaces exhibit quantum efficiencies greater than 50%. The high quantum yield of photoconducting films, particularly at long wavelengths, makes it desirable to investigate their use with the most efficient video-signal-generating techniques and thus to benefit from the combination of favorable characteristics.

With respect to the second factor, the image isocon, a variation of the familiar image orthicon, uses a unique method of read-out by a scanning electron beam to provide improved low-light performance. Isocon scan utilizes the distinguishable velocity spectrum of scanning-beam electrons that are scattered from the charge-storage surface to provide signal modulation. Electron optics designed to separate the

---

* This work was carried out under Air Force Contract No. AF33(657)-11581 under Task No. 415605 of Project No. 4156 administered under direction of the Air Force Avionics Laboratory Research and Technology Division.

return-beam electrons from the scattered electrons are required. Use of isocon scan in place of orthicon scan in an image isocon tube[1] has resulted in an increase by a factor of three in signal-to-noise ratio and expansion by a factor of ten in the brightness range within a single scene that can be accommodated without readjustment of the scanning beam.

Photoconductors that exhibit quantum conversion efficiencies between 50 and 100% at peak-response wavelengths are being used in vidicon tubes. For simplicity of operation, the vidicon signal is detected at the target. In this mode of operation, the limiting value of low-light signal is greater than the amplifier noise level. The low-light threshold can be reduced if the video signal is obtained from the modulated return beam after it is amplified by an electron multiplier. Such a photoconductive camera tube becomes competitive in sensitivity with the image orthicon. It also acquires the limited signal-to-noise and restricted dynamic range of the image orthicon. A more desirable combination of characteristics would be the superior photosensitivity of a photoconductor with the low-noise, high-gain readout of isocon scan. However, evaporated layers of photoconductors such as are used in vidicon camera tubes appeared, in preliminary tests, to exhibit low scattering coefficients.

To establish the feasibility of a photoconductive isocon tube, a study was made of the electron scatter coefficient of various photoconductive surfaces. Techniques that increased the scatter gain through surface treatment or overlayers were also studied. The investigations covered the following areas:

1. A variety of vacuum-evaporated photoconductor layers of the type suitable for vidicon application were evaluated for scattering properties. These studies were performed in a demountable camera-tube system.
2. Surface treatment of the photoconductor layers to provide improved scattering was investigated.
3. The effects upon the scattering of varying conductivity and voltage-modulation conditions were studied.
4. Prototype tube samples were built, and their operating characteristics were measured. Parameters were systematically varied in order to provide a comparison between the photoconductive isocon and other television camera tubes.

---

[1] A. D. Cope and H. Borkan, "Isocon Scan: A Low-Noise Wide Dynamic Range Camera Tube Scanning Technique," *Applied Optics*, Vol. 2, p. 253, March 1963.

ISOCON SCANNING

Basic Principle

Not all the beam electrons of low energy that are incident on a camera-tube storage target that is at a positive voltage with respect to the equilibrium potential established by the scanning beam will remain at the surface. A fraction will experience an elastic scattering that is nonspecular. The portion of the beam that has insufficient energy to reach the target is specularly reflected, as from a mirror. The spread in initial energy of scattered electrons is broad, whereas energy variation among reflected electrons is restricted to a small range, particularly in that component that is at right angles to the tube axis. Electron optics employed at the entrance to the electron multiplier allow only the scattered electrons to enter. When there is no stored signal at the target there should be no output signal. In the practical case, however, a small spurious output signal is present. With a stored signal at the target, an output current modulation in excess of 90% has been measured. Isocon scan is less degraded by the presence of excess beam current than other scanning means because of its greater signal-to-noise ratio.

Electron Optics

The electron optics that provide the capability for separating the scattered electrons from the reflected portion of the scanning beam were evolved during previous studies of the image isocon.[2,3] The significant electrode elements are shown in Figure 1. The collimated and focused beam is directed to the target with a combination of axial and radial velocity components that cause the reflected beam to be sufficiently off-axis to be intercepted by the separation electrode. By making the decelerating region between the mesh and the target small and the equipotentials flat surfaces parallel to the target plane, the reflected portion of the scanning beam from all parts of the target area being scanned retraces its path in returning through the deflection region. The reflected beam occupies a small region of the tube at the plane of the separation aperture. In isocon operation, the reflected electrons are directed off-axis to be incident upon the separation electrode.

The beam electrons that strike the target and are then elastically

---

[2] P. K. Weimer, "The Image Isocon," *RCA Review*, Vol. 10, p. 366, Sept. 1949.

[3] AF33(616)-5728 and AF33(616)-6497, WADD Technical Report No. 60-843, March 1961, A. D. Cope, C. C. Peterson, and H. Borkan.

scattered have a variety of initial velocity directions; this process is nonspecular. These scattered electrons reach the separator plane at a variety of radial positions, depending upon their initial radial-velocity component. Only those electrons with radial velocities matching the radial velocity of the reflected beam electrons are trapped by the separator; the rest of the scattered electrons pass through the separator aperture and enter the electron multiplier. Ideally, there should be no electrons entering the multiplier when there is no signal modulation, and, consequently, there should be zero output current. In practice, the minimum output current of the image isocon in the dark is 5 to 10% of the output current from a high-light condition. With the photoconductive isocon, the situation is more complex because of the photoconductor dark current. Although the output current can be near zero when there is no field applied to the photoconductor, indicating effective separation of the two classes of electrons, there is significant current when the normal operating potential is applied to the unilluminated photoconductor.

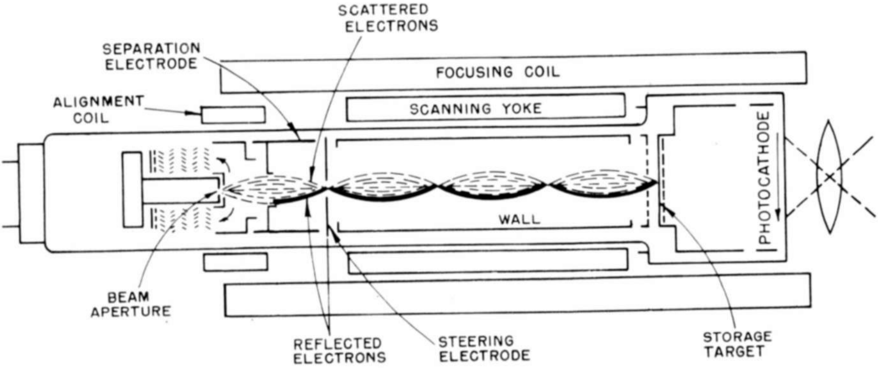


Fig. 1—Dual-operating-mode image-isocon–image-orthicon.

The necessary radial component of velocity is added to the collimated scanning beam by a combination of the alignment field (an adjustable magnetic field directed at right angles to the tube axis) and the steering electrodes, which deflect the beam normal to the tube axis. The alignment field acts only upon the outgoing beam, whereas the steering field acts upon both the outgoing and the return beam. Thus, they can be adjusted independently, thereby permitting optimum removal of reflected electrons without loss of a significant fraction of the scattered electrons.

This particular isocon design also permits operation in the orthicon scanning mode, because the return current from an axially aligned beam can be made to pass through the separation aperture unimpeded. Thus, an accurate comparison can be made of orthicon and isocon scan in the same tube using identical target modulation. Also, in the process of tube setup a continuous transition from orthicon to isocon scan occurs that can be followed in the video output display. Under these conditions, the adjustment for optimum separation over the entire scanned area can be determined more easily than with alternative isocon designs.

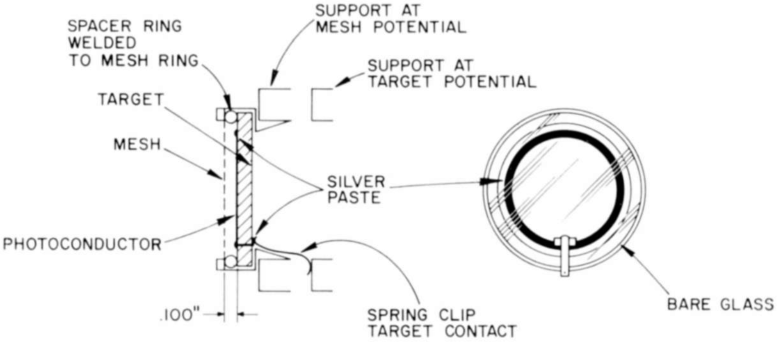


Fig. 2—Photoconductor-isocon target assembly.

Isocon scan requires a uniform decelerating field with a high potential gradient in the region near the scanned surface of the target. The decelerating field is provided by the mesh on the scanned side of the target. The function of this mesh is to eliminate the varying radial forces that the beam experiences in different parts of the scanning raster if the decelerating field is not uniform over the entire area. By reducing the scan in the return beam, the mesh ensures that the separation will be uniform when scanning the edges as well as the center of the target. Orthicon tubes with a field mesh are available; the advantages have been demonstrated in a more uniform center-to-edge performance in the video output. For either method of scan, there is one inherent disadvantage in having the mesh in the path of the scanning beam. Half of the beam is intercepted before it reaches the target, and half of the return electrons are intercepted. Also, the mesh contributes scattered electrons that do not represent signal modulation.

A photoconductive isocon was constructed by replacing the storage target and photocathode of the structure shown in Figure 1 by the assembly shown in Figure 2.

The target assembly for a photoconductive isocon consists of a photoconductor layer approximately 5 microns thick deposited on a transparent signal-plate electrode deposited on a glass support. In addition, a mounting for the target structure is required in order to provide close uniform spacing between the decelerator mesh and the photoconductor. The mesh is electrically connected to the support cup, as shown in Figure 2. The glass target disc for the evaporated photoconductor layer provides insulation between the two elements. The substrate electrode that covers the central area of the glass support disc is carried through an opening to a contact area on the reverse side. Contact is made by a spring strip when this assembly is inserted into the target support cup. This insertion takes place after the faceplate is sealed to the large stem assembly and before the gun seal is made.

The general configuration of the tube is that of the three-inch image orthicon. The image section is foreshortened to bring the faceplate to within an inch of the photoconductor layer, permitting a fifty-millimeter lens to reach focus without striking the faceplate. Although it is possible to sense the video signal generated by the scanning beam at the signal plate, smaller signal amplitudes can be detected by either of the return-beam/multiplier techniques. For example, operation of a vidicon in the conventional mode (signal detection at the target) was compared with operation in the isocon mode. With just enough illumination to give a minimum detectable signal in the conventional mode, the voltage modulation was 250 times greater at the output of the multiplier in the isocon mode.

MEASUREMENT OF PHOTOCONDUCTOR SCATTERING

At the start of the present study, only a few fragments of information were available concerning the behavior of various photoconductors as low-energy electron-scattering devices. These indicated that the scatter coefficient (the ratio of scattered current to incident current) is lower than that obtained with the bombardment of insulator surfaces such as glass or MgO used as storage targets in the image isocon.

Vacuum evaporated samples of a photoconductor on an electroded glass support were prepared and inserted in an ion-pumped demountable camera-tube system. This vacuum system proved to be a successful innovation, since it permitted extensive examination of samples without contamination from oil or other condensable vapors that would obscure the surface phenomena unique to a particular sample.

A schematic diagram of the demountable pumping system employed for these tests is shown in Figure 3. This ion-pumped system proved to be an outstanding improvement over previous demountables employing oil-diffusion pumping. Low-energy scattering, which is extremely sensitive to surface contamination, showed stable performance over several days. The electron multiplier gain also remained stable over many days if the system was not opened to the atmosphere. This had never been achieved with the oil-pumped system.

When scanning a light-modulated storage target, there is no direct means of determining the current incident on the target. For example, an insulating photoconductive target surface with no field applied will be driven to its most negative value when scanned continuously and allowed to come to equilibrium. The spread of energies in the beam is such that 25% or less of the beam electrons have an axial energy that will permit them to reach a surface only a few volts above its equilibrium potential. This fraction drops to less than 5% for the last few tenths of a volt above the equilibrium. When bombarding a surface with electrons having energies below the first crossover potential, it has been assumed that the sum of the electron current out of the target, I_T, and the scattered current collected at the first dynode, I_{dy_1}, represents the useful part of the current that is incident on the target. The scattering coefficient is

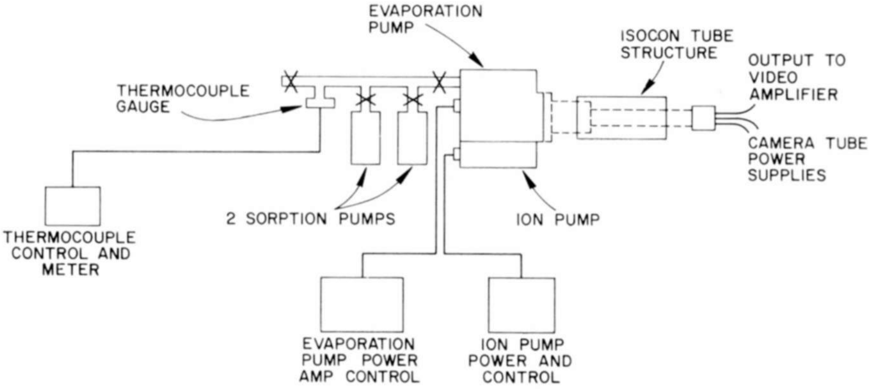


Fig. 3—Oil-free demountable pumping station.

$$K_s = \frac{I_{dy_1}}{I_{dy_1} + I_T}.$$

The scatter coefficient shown in Figure 4 as a function of target current, with target potential as a parameter, was measured for a magnesium-oxide-target image isocon, such as shown in Figure 1. In this tube design, the target potential can be accurately established by flooding the tube with light so that the low-capacitance target is fully charged and is at the potential of the mesh. The mesh acts as the collector of the secondaries excited by bombardment of the photocathode side of the target.

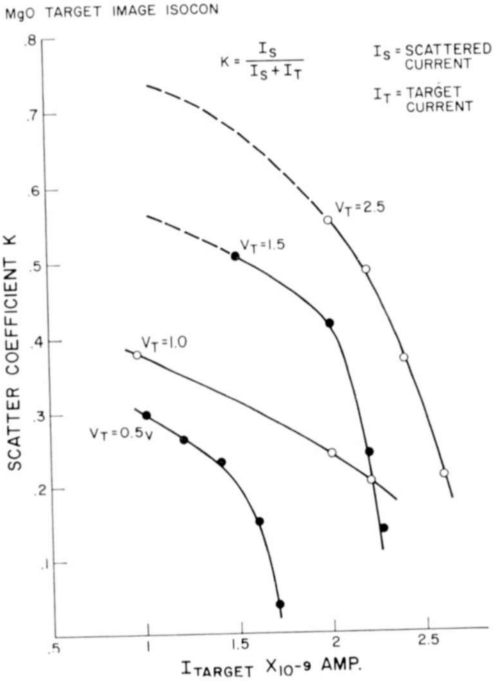


Fig. 4—Scatter coefficient as a function of target current.

At target currents below 1.7×10^{-9} ampere, the increase in scatter coefficient with target potential is supralinear. At a fixed target potential (at the start of the scan), the decrease in scatter coefficient with increasing target current is the result of a complexity introduced by the scanning process. If the fixed amount of charge per unit area that accumulates between scans is removed at a more rapid rate, the potential of the scanned surface decreases more rapidly, and the average potential will be lower.

When a photoconductor target is scanned with high beam current, the scatter coefficient in the illuminated areas is different from that in the dark areas because the potential rise at the scanned surface is greater in the regions with increased conductivity. To overcome this effect, it is desirable to measure the scattering at an incident beam current that is too low to alter the target potential. This operating condition is more easily achieved with photoconductor targets, which have capacitances several orders of magnitude greater than the image isocon target.

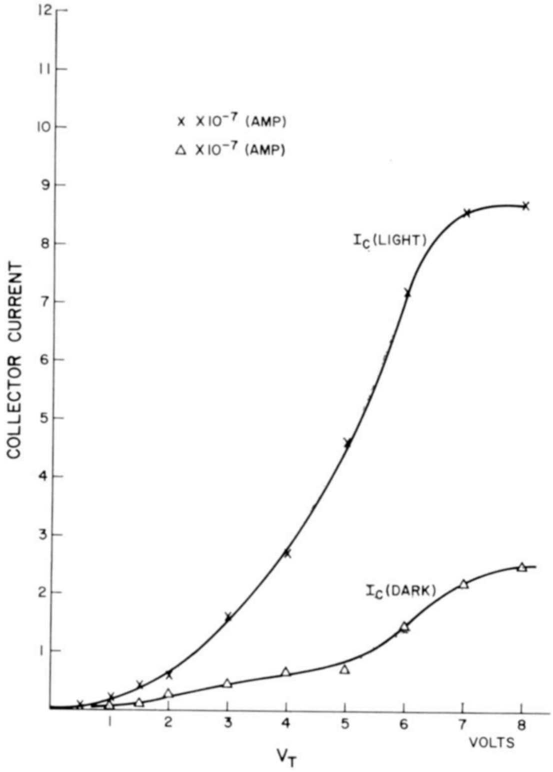


Fig. 5—Scattered current from a photoconductor layer.

The related scattered and target currents from a photoconductor layer measured while scanning with a small incident beam current are shown in Figures 5 and 6, respectively. Scatter coefficient measurements of the various photoconductor samples were taken with high illumination flooding the target and with the minimum usable incident beam current.

PHOTOCONDUCTOR MATERIALS

The results obtained from a variety of photoconductors, some with and some without a high-secondary-emission overlayer, indicated one common property related to the scatter coefficient. The potential of the surface being probed by the beam electrons has a strong effect upon both the energy with which the incident electrons strike the surface and the fraction of incident electrons that will be scattered. It is likely that these two effects are related, since the penetration of the bombarding electron into the surface it strikes will be greater with increasing energy.

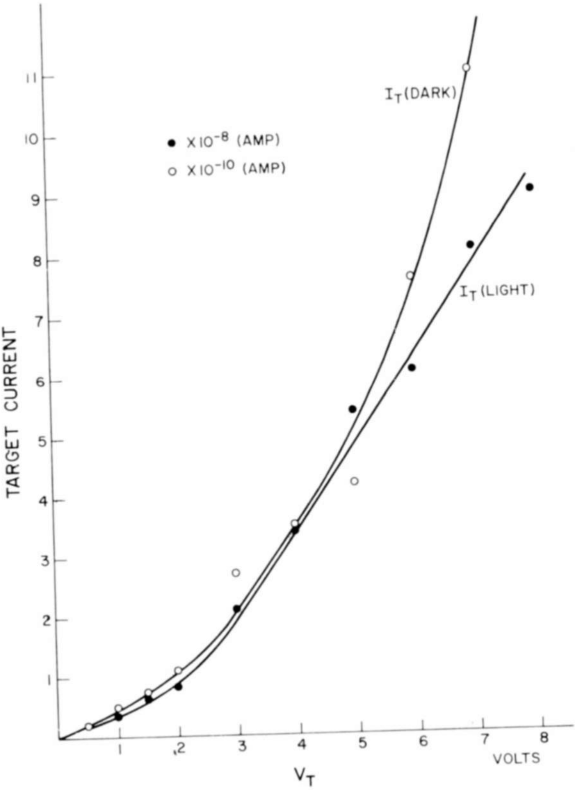


Fig. 6—Target current through a photoconductor layer.

In the present application, a relatively small range of incident electron energies is of interest. For low-illumination applications, only fractions of a volt of modulation result from the light excitation of the photoconductor. Higher potentials at the scanned surface can be obtained by increasing the field across the photoconductor, since this permits more current to flow. In general, the photo-current increases linearly with the applied field, whereas the dark current increases as the square of the applied field. Since the dark conductivity causes a rise in the surface potential of the target upon which the photomodulation is added, excessive dark current tends to reduce the contrast in the signal and make the detection of small photo-excited variations more difficult.

Typically, the photoconductive isocon output signal rises as the target potential is increased, reaching a peak at about eight to ten volts. Above this potential the signal drops or becomes nonuniform over the area, because large voltage modulations make the separation of scattered electrons more difficult. As the potential of the target electrode is increased to twenty volts or greater, there is a tendency for the video signal to increase rapidly. It appears, however, that this results from the creation of secondary electrons, as evidenced by the target-current reversing polarity. If the video signal were being taken from the target, there would be a clear-cut difference between the low-bombarding-energy operation and the high-velocity case involving secondary electrons. With the signal being detected in the electron stream returning from the target region of the tube, the scattered and secondary electrons overlap in energy and initial velocity, so that a smooth transition from one operating mode to the other takes place. The greatest practical difference, as far as the photoconductive isocon is concerned, is that operation with the incident electrons above the first crossover potential to generate secondaries of relatively low initial energy tends to result in very poor erasure of the charges stored at the target. Moving objects in the scene are obscured. Operating in the range where the target current is essentially zero, the readout is nondestructive and a continuous display of one scene can be achieved. Such a capability could be useful if a prolonged viewing of a single field were desired.

Sb_2S_3 with excess Sb, which exhibits a higher photosensitivity than stoichiometric Sb_2S_3, proved to be too conducting for a good isocon target. Arsenic compounds, such as As_2S_3 and As_2Se_3, either alone or in combination, showed a tendency for burn-in of an image as well as excessive lag. The combination of As_2S_3 and Sb_2S_3 proved interesting but not reproducible when later samples were prepared in tubes. Selenium with an overlayer of Sb_2S_3 is a useful combination since the dark current of Se is extremely low. At present, the use of this material is excluded from practical applications, however, because the inherent instability of amorphous selenium limits the life of such a target. This target did serve, however, to prove the effectiveness of markedly reducing the dark conductivity. The last group of Sb_2S_3 samples examined employed an insulating substrate whose purpose is to reduce the dark current.

As shown in Table I, the general trend is for the scatter coefficient to increase with increasing target potential when the target conductivity is low. At a fixed target potential, the scatter coefficient remains essentially constant as the incident beam current is varied over a range of ten to one. The apparent value of the scattering coefficient drops to a low level only when the target conductivity is so low that the target potential cannot be maintained at a constant value during scanning with the smallest useful beam current.

A variety of materials were applied as surface layers to enhance the scattering coefficient. Insulating materials with a high secondary-emission ratio were chosen, since there appears to be a reasonable correlation between the two phenomena. Barium fluoride is effective but somewhat high in resistivity, thus requiring very thin layers if a steady current of reasonable magnitude is to be drawn through the target. Magnesium fluoride with its lower resistivity is more easily managed and is the best of the materials examined for this application. The material is not too refractory for easy evaporation and does not decompose.

One means for ensuring a high scattering coefficient with a sufficient field across the photoconductor to give a high photosensitivity is to reduce the dark current. A method for accomplishing this is to use a thin insulating barrier layer between the substrate electrode and the photoconductor. This technique was borrowed from a previous development[4] and was found to be quite effective in this application; Tl_2O, GeO, Al_2O_3, and Sb_2O_3 were all found to give improved results as dark-current suppressors with Sb_2S_3. This layer must be thin (approximately 0.1 micron) and uniform if it is not to introduce nonuniformities in the target conductivity over the scanned area.

TUBE PERFORMANCE EVALUATION

The performance of several photoconductive isocon tubes are given in the following section. As shown in Table II, the scatter coefficient remains almost constant over the useful operating range when the appropriate beam required for the particular illumination is employed.

---

[4] AF33(657)-8843, "Applied Research on Photoconductive Photo-Tape," AL-TDR-64-18, pp. 26-29, issued March 1964.

Table I—Scatter Coefficient of Various Photoconductor Materials

Target No.	Photoconductor	Over-layer	Target Voltage	Scatter Coefficient		Comments
				Dark	Light	
4	Sb_2S_3(Sb)	——	10	0.45	0.1	Good picture
			13	0.43	0.14	
			15	0.43	0.15	
			18	0.45	0.12	
			20	0.44	0.22	
			23	0.42	0.22	
			25	0.42	0.20	
12	$As_2S_3 + As_2Se_3$ (1:2)	——	6	—	0.30	Image burn-in
			10	—	0.45	
			13	0.69	0.50	
			15	0.70	0.53	
			20	0.74	0.45	
			25	0.58	0.31	
13	$As_2S_3 + As_2Se_3$ (1:1)	——	8	—	0.67	Image burn-in
			11	0.94	0.69	
			15	0.82	0.58	
			20	0.78	0.51	
			25	0.55	0.43	
14	$As_2S_3 + Sb_2S_3$ (1:1)	——	5	—	0.36	Image burn-in
			8	—	0.30	
			12	0.77	0.30	
			15	0.67	0.31	
			20	0.83	0.29	
			25	0.90	0.83	
15	Se	Sb_2S_3	8	—	0.5	Low dark current, good contrast
			10	—	0.4	
			13	—	0.4	
			15	—	0.38	
			20	—	0.4	
16	Se + Sb_2S_3	MgF_2	8	—	0.84	Good picture
			10	—	0.66	
			13	—	0.64	
			15	—	0.63	
			18	—	0.64	
			21	—	0.67	
17	Sb_2S_3 over Tl_2O	——	3	—	0.07	
			5	0.13	0.09	
			8	0.06	0.10	
			10	0.12	0.14	
18	Sb_2S_3 over Ge	——	3	0.02	0.09	Fair picture
			6	0.15	0.11	
			9	0.12	0.12	

Table I (continued)—Scatter Coefficient of Various Photoconductor Materials

Target No.	Photoconductor	Over-layer	Target Voltage	Scatter Coefficient		Comments
				Dark	Light	
19	Sb_2S_3 over Sb_2O_3	——	3	0.35	0.21	Fair picture
			5	0.29	0.24	
			7	0.25	0.23	
			9	0.24	0.23	
			11	0.25	0.24	
20	Sb_2Sb_3	——	3	0.5	0.29	Fair picture
			5	0.32	0.28	
			7	0.31	0.24	
			9	0.23	0.24	
			11	0.25	0.26	
21	Sb_2S_3 over Tl_2O	——	3	—	0.20	Fair picture
			5	0.38	0.24	
			7	0.41	0.28	
			9	0.40	0.25	
			11	0.35	0.25	

Table II—Scatter Coefficients of Operating Tubes

Tube No. 7		
Illumination on tube face (foot-candles)	K_{Dark}	K_{Light}
0.189	0.56	0.61
0.095	0.56	0.59
0.038	0.59	0.60
0.035	0.59	0.61
0.021	0.59	0.61
0.011	0.60	0.61
Tube No. 10		
0.17	0.46	0.78
0.133	0.78	0.67
0.08	0.78	0.97
0.08	0.89	0.79
0.035	1.0	0.77
0.021	0.93	0.97
0.011	0.63	0.83
0.0055	0.67	0.61

Scatter Coefficient and Signal Modulation

The contribution from the different current sources whose sum is the tube output is shown schematically for orthicon scan in Figure 7 and for isocon scan in Figure 8.

For the orthicon return-beam operation, the maximum output current occurs when the beam is totally reflected at the target. The output current modulation, μ, of significance for assessing the expected signal-to-noise performance is

$$\mu = \frac{\Delta I_L}{I_c} = \frac{I_d - I_L}{I_c},$$

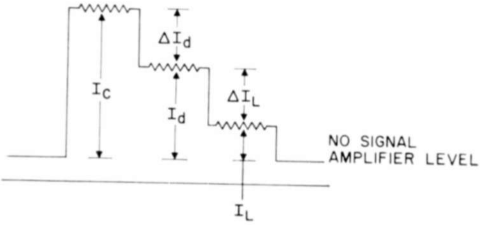


Fig. 7—Components of the return beam current when employing orthicon scan.

where ΔI_L is signal modulation from light-excited current;

I_c is total collected current with no field applied to the target;

I_d is dark signal current with target field applied, no illumination; and

I_L is light signal current with target field applied and illuminated.

For isocon scan there are two modulation factors of interest. First, with no field applied to the target, the output current, I_c, should be only a few per cent of the high-light output current. This figure is significant in estimating the efficiency of scattered electron separation. A large value of I_c indicates poor separation of scattered from reflected electrons or the presence of spurious current leaking into the electron multiplier. This modulation factor is

$$\mu_T = \frac{\Delta I_T}{I_L} = \frac{I_L - I_c}{I_L},$$

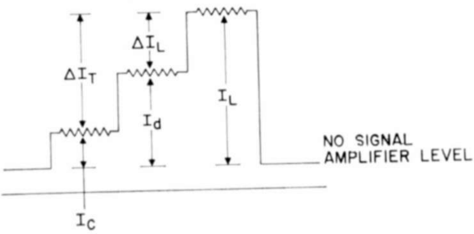


Fig. 8—Components of the return beam current when employing isocon scan.

where ΔI_T is the current from the illuminated target.

The second modulation factor is concerned with light-excited current and is of practical significance in estimating the signal-to-noise performance;

$$\mu_p = \frac{\Delta I_L}{I_L} = \frac{I_L - I_d}{I_L}.$$

Thus, I_d should be made as small as possible to permit the highest modulation and the best signal-to-noise performance. Measured values for one tube are shown in Figure 9.

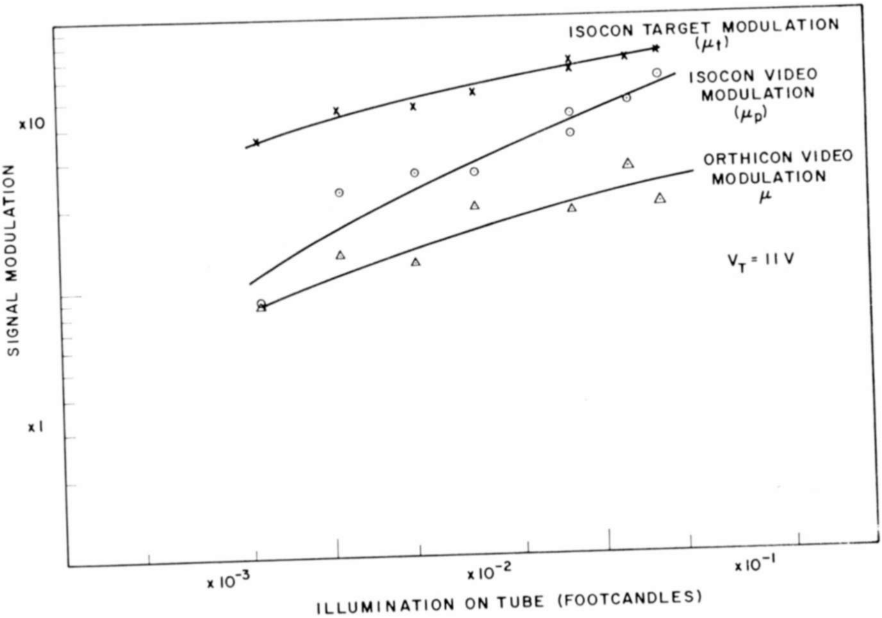


Fig. 9—Signal modulation of photoconductor isocon–orthicon readout.

Signal-to-Noise Ratio and Dynamic Range

Figure 10 shows signal-to-noise ratio as a function of illumination for both orthicon and isocon scan, each with the same applied potential and illumination. The beam current was adjusted to its optimum value at each illumination level. Here again there is an improvement when isocon scan is used.

An illumination range of 1000 to 1 can be handled with a single setting of the scanning-beam current when the isocon signal is being detected. When orthicon return-beam scan is used, the illumination range is greater in the case of the photoconductive isocon than for the photoemissive image isocon because the low gamma (0.6 to 0.7) of the photoconductor layer compresses the range of potential modulation developed at the scanned side of the target.

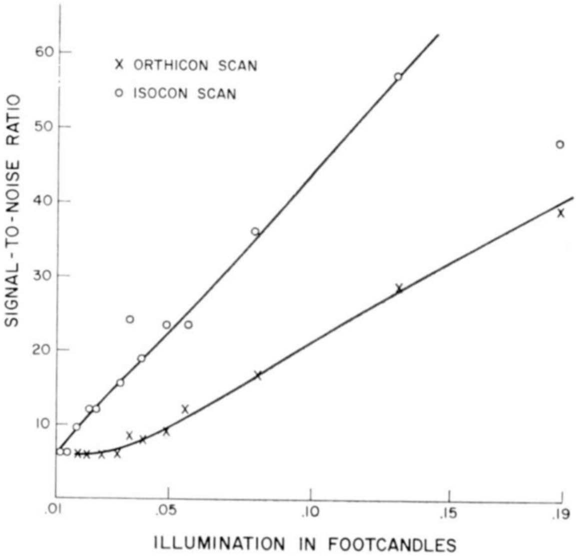


Fig. 10—Signal-to-noise ratio as a function of illumination for isocon and orthicon scan of a photoconductor.

These observations with the tubes are consistent with similar measurements obtained earlier in the demountable setup. Excess beam current, which is the source of a rapid increase in orthicon noise, does not degrade the isocon signal. The noise in this instance is characteristic of the scattered emission.

Resolution and Lag

Good resolving properties are favored by the photoconductive target geometry and the large potential modulations of the photoconductive target. Isocon resolving power is equivalent to orthicon resolution, except at extremely low illumination, where the lower isocon noise gives superior results.

Figure 11 shows the output signal from both orthicon and isocon scan at a variety of illumination levels. The test pattern consisted of a series of black and white bars. The difference in noise is readily seen.

These tubes show moderate lag in either mode of operation resulting from a combination of the photoconductive decay time constant and target capacitance. Scanning a smaller target area materially decreases the lag.

Sensitivity

Vidicon tubes with visible light sensitivity of the 200- to 400-microampere-per-lumen range require face-plate illuminations of 10^{-1} foot-candle or greater. A photoconductive isocon with similar photosensitivity will give a useful signal in the 10^{-3} foot-candle range. This operating range, between 10^{-3} and 10^{-1} foot-candle, bridges the gap between the image orthicon and the conventional vidicon. This is the same operating range encompassed by the combination of intensifier and vidicon or by tubes using high-gain targets operating by secondary-emission conductivity.

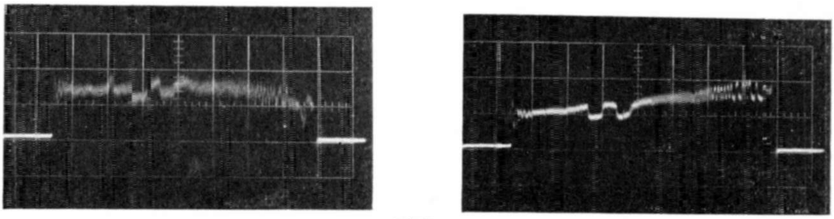
SUMMARY AND CONCLUSIONS

The effective scattering coefficient from a photoconductor target must be greater than 0.5 if the isocon-scan signal amplitude is to be equal to or greater than the orthicon-scan signal amplitude. This can be achieved with certain target materials of low dark conductivity and with low applied potentials. The target-operating-potential range must be carefully restricted to limit the maximum potential rise at the scanned surface. The noise accompanying the signal is less with isocon scan than with orthicon scan when the signal amplitudes are equivalent. Isocon scan gives a higher signal-to-noise ratio than orthicon scan with both photoemissive and photoconductive tubes. The improvement is greater for the case of the photoemissive tube.

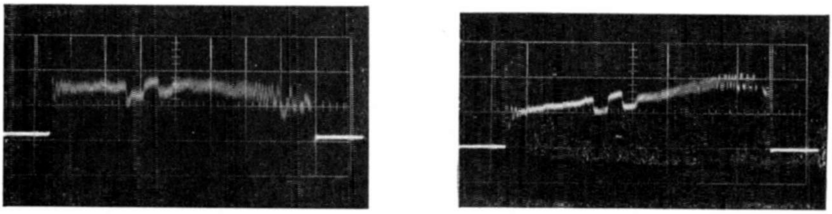
The low gamma of the photoconductor assists orthicon scan in handling wide variations of scene brightness. Despite this, isocon scan continued to out-perform orthicon scan when there is a large dynamic range in the scene.

The decrease in isocon signal output when the optimum potential of the scanned surface is exceeded restricts the operating potential of the target. Orthicon operation can be enhanced by increasing the target potential and accepting the added dark current.

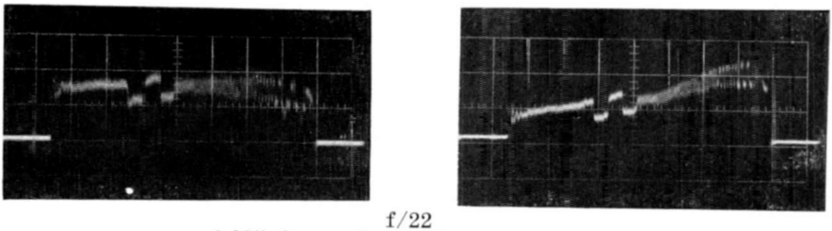
ORTHICON SCAN ISOCON SCAN



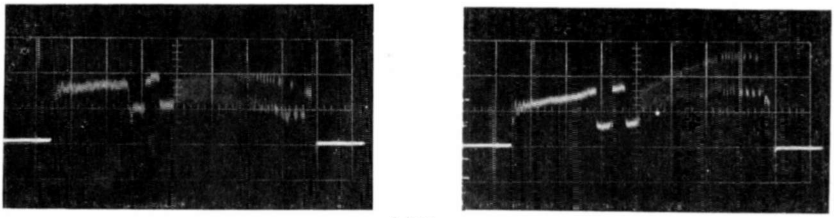
f/45
0.011 ft.-candle incident illumination



f/32
0.021 ft.-candle incident illumination



f/22
0.035 ft.-candle incident illumination



f/16
0.080 ft.-candle incident illumination

Fig. 11—Oscillograms of photoconductor isocon–orthicon signal output at various illumination levels. (Single-line scan across a pattern of vertical bars—100% contrast.)

A sensitive low-dark-current photoconductor with high scattering of low-energy electrons will give superior low-light performance when the signal is detected by isocon-scan techniques. Overlayers have been shown to be effective in supplying enhanced scattering. Insulating barrier layers have provided reduction in dark current.

The photoconductive isocon has demonstrated its capability to provide a hundred-fold reduction in the low-light sensitivity of a vidicon with a signal-to-noise ratio twice that of the return-beam orthicon scan. A 1.5 times amplification of the isocon output signal over the orthicon return-beam signal is typical.

The development of a photoconductive isocon has been shown to be feasible within the capability of present materials. The optimum target becomes complex when control of dark conductivity by means of a barrier layer, and a highly scattering overlayer is required. With the development of low-dark-current photoconductors, this complexity would be diminished.

In instances where a slow scan rate for readout is to be employed, as is demanded by the need to restrict the transmission bandwidth (a common situation in satellite and space-probe applications), the photoconductive isocon warrants further development to achieve its optimum potential. The advantages are

1. The target is constructed on a rigid support and is not readily damaged by shock and vibration.
2. The low-light performance exceeds that which can be achieved by other photoconductor tubes.
3. The noise in the output signal is not critically dependent upon beam-current adjustment.
4. The lag characteristic of low-light operation of present photoconductors is not a limiting factor when single-frame exposures to the light image at widely spaced intervals are to be employed.

Sensitive photoconductive targets with high scattering coefficients, when combined with isocon scan, will result in a tube that will be as sensitive as the image orthicon in applications where smear from the combination of photoconductive and capacitance lag is not critical.

ACKNOWLEDGMENTS

The authors wish to express their thanks for the assistance of numerous associates who made material contribution to this study; E. C. Hutter and S. Gray for their guidance, C. J. Busanovich for design and preparation of the photoconductor targets, and H. C. McCandless for construction of demountable tube structures.

STRESS ANALYSIS OF BUTT-TYPE CERAMIC-TO-METAL SEALS

By

S. W. Kessler

RCA Electronic Components and Devices
Lancaster, Pa.

Summary—*Equations are derived defining the stresses at any location in the metal or ceramic of a butt-type ceramic-to-metal seal. Simplifications are made for rapidly determining the approximate stresses at the jointure. The problems encountered in making large high-temperature seals are discussed in light of the stress analysis.*

General Derivation

HIGH-TEMPERATURE ceramic-to-metal seals are used in vacuum devices such as thermionic energy converters, and must be strong enough to resist the stresses imposed by thermal cycling at operating temperatures as high as 800°C. This paper describes a stress analysis made to optimize the design of these seals and thereby minimize residual stresses.

The analysis is limited to butt-type seals of the type shown in Figure 1, in which the diameter-to-wall-thickness ratio is greater than 5. It is assumed that the metal members are infinitely long cylinders, and that the ceramic is a short cylinder. For an infinitely long cylinder, β, a factor depending on the rigidity of the cylinder, must be equal to or greater than 2.6. The value of β is given by

$$\beta^4 = \frac{Eh}{4a^2D}, \tag{1a}$$

where D is the modulus of rigidity,

$$D = \frac{Eh^3}{12(1-\mu^2)},$$

E is the modulus of elasticity, h is the wall thickness, a is the mean radius of the seal, and μ is Poisson's ratio.

If Poisson's ratio is 1/3, which is approximately true for most materials, then

$$\beta^4 = \frac{8}{3a^2h^2} \quad \text{and} \quad D = \frac{3Eh^3}{32}. \tag{1b}$$

If all the stresses are elastic, the radial displacements of the metal and ceramic cylinders are equal to the differential thermal expansion of the members at the seal.[1] This relationship may be expressed algebraically as follows:

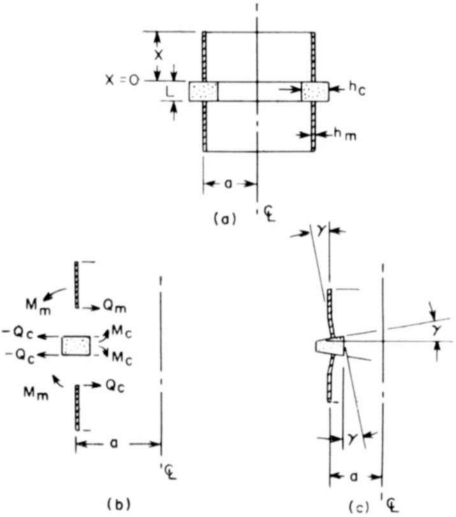


Fig. 1—The forces and deflections on a butt seal of molybdenum to alumina.

$$|w_m| + |w_c| = a(\alpha_m - \alpha_c)\Delta T, \tag{2a}$$

where w is the radial displacement, α is the coefficient of expansion, and ΔT is the temperature change over which the deflections occur. The subscripts m and c are used to denote metal and ceramic members, respectively.

If the expressions for $|w_m|$ and $|w_c|$ given by Timoshenko[2] are substituted in Equation (2a) and a positive sign is assigned for a deflection toward the axis of the cylinder, the displacements are given by

---

[1] T. D. Riney and J. W. Elek, "Thermoelastic Stresses in Balanced and Unbalanced Seals," *Bell Syst. Tech. Jour.*, p. 1519, Vol. 41, Sept. 1962.

[2] Equations derived by S. Timoshenko, *Theory of Plates and Shells*, McGraw-Hill Book Co., N. Y., 1940, were used in this analysis.

$$|w_m| + |w_c| = \frac{1}{2\beta_m{}^3 D_m} [\beta_m M_m \Omega(\beta_m x) + Q_m \theta(\beta_m x)] - \frac{2\beta_c a^2}{E_c h_c} [Q_c X_1(\beta_c L) + \beta_c M_c X_2(\beta_c L)] = a(\alpha_m - \alpha_c)\Delta T, \text{ at } x = 0. \quad (2b)$$

M's are the bending moment per linear inch of circumference, Q's are the shear forces per linear inch of circumference, and x and L are the distances along the length of the metal cylinder and the ceramic cylinder, respectively, and

$$\Omega(\beta_m x) = e^{-\beta_m x} (\cos \beta_m x - \sin \beta_m x),$$

$$\theta(\beta_m x) = e^{-\beta_m x} \cos \beta_m x,$$

$$X_1(\beta_c L) = \frac{\cosh \beta_c L + \cos \beta_c L}{\sinh \beta_c L + \sin \beta_c L},$$

$$X_2(\beta_c L) = \frac{\sinh \beta_c L - \sin \beta_c L}{\sinh \beta_c L + \sin \beta_c L}.$$

If the metal and ceramic cylinders are at right angles at their jointure and if the strain is elastic, the slopes of the deflections are equal,[2] as shown in Figure 1(c). By substitution of the first derivatives of the deflections, this condition may be expressed algebraically;

$$\gamma = \left(\frac{dw_m}{dx}\right)_{x=0} = -\left(\frac{dw_c}{dl}\right)_{x=0 \text{ or } l=L/2}, \quad (3a)$$

and

$$\frac{1}{2\beta_m{}^2 D_m} [2\beta_m M_m \theta(\beta_m x) + Q_m \phi(\beta_m x)] = -\frac{2\beta_c{}^2 a^2}{E_c h_c} [Q_c X_2(\beta_c L) + 2\beta_c M_c X_3(\beta_c L)]$$

where
$$X_3(\beta_c L) = \frac{\cosh \beta_c L - \cos \beta_c L}{\sinh \beta_c L + \sin \beta_c L}, \quad (3b)$$

and $\phi(\beta_m x) = e^{-\beta_m x}(\cos \beta_m x + \sin \beta_m x)$.

Figure 1(b) indicates that the shear force per linear inch, Q, and the bending moment, M, are the same for the ceramic and metal members, i.e.,

$$Q_m = Q_c = Q \text{ and } M_m = M_c = M. \tag{4}$$

Therefore, if the terms in Equations (2b) and (3b) are rearranged and Q and M are substituted for Q_c, Q_m, M_c, and M_m, the following simultaneous equations are obtained:

$$M\left[-\frac{2\beta_c{}^2a^2X_2(\beta_cL)}{E_ch_c}+\frac{\Omega(\beta_mx)}{2\beta_m{}^2D_m}\right]+Q\left[\frac{\theta(\beta_mx)}{2\beta_m{}^3D_m}-\frac{2\beta_ca^2X_1(\beta_cL)}{E_ch_c}\right]$$
$$= a(\alpha_m-\alpha_c)\Delta T, \tag{5}$$

and

$$M\left[\frac{\theta(\beta_mx)}{\beta_mD_m}+\frac{4\beta_c{}^3a^2X_3(\beta_cL)}{E_ch_c}\right]=Q\left[-\frac{2\beta_c{}^2a^2X_2(\beta_cL)}{E_ch_c}-\frac{\phi(\beta_mx)}{2\beta_m{}^2D_m}\right]. \tag{6}$$

When the dimensions and the Young moduli for the materials used in a butt seal are inserted into these equations, the shear force per linear inch of seal and the bending moments, in pounds, can be calculated.

APPROXIMATE SOLUTION

For most seals, the wall thickness of the ceramic is much greater than that of the metal; consequently, the terms involving the ceramic in Equations (5) and (6) are small in comparison to the metal terms. Therefore, these equations may be rewritten at $x = 0$ as

$$\frac{M}{2\beta_m{}^2D_m}+\frac{Q}{2\beta_m{}^3D_m}=a(\alpha_m-\alpha_c)\Delta T, \tag{7}$$

and

$$\frac{M}{\beta_mD_m}=-\frac{Q}{2\beta_m{}^2D_m}. \tag{8}$$

These equations can then be solved for Q and M:

$$Q = 4\beta_m{}^3D_m\, a(\alpha_m-\alpha_c)\Delta T, \tag{9a}$$

and

$$M = -2\beta^2D_m\, a(\alpha_m-\alpha_c)\Delta T. \tag{10a}$$

Q and M can be expressed simply in terms of the dimensions by substitution of the equations defining β and D into Equations (9a) and (10a);

$$Q = 0.782 h^{3/2} a^{-1/2} E_m (\alpha_m - \alpha_c) \Delta T, \tag{9b}$$

$$M = 0.726 h^2 E_m (\alpha_m - \alpha_c) \Delta T. \tag{10b}$$

STRESSES ALONG THE PRINCIPAL AXES

Figure 2 illustrates the principal axes along which the major stress components act (i.e., shear S_s, hoop S_h, and bending S_b stresses). Both the shear and hoop stresses are a maximum at $x = 0$; the bending stress is a maximum at $l = 0$. The maximum bending stress is considered later.

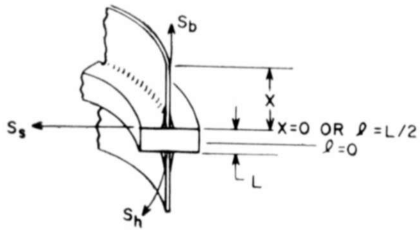


Fig. 2—Principal axes and major stress components.

Shear Stress—The maximum shear stress, S_s, is determined by dividing Equation (9b) by the wall thickness. Because the wall thickness changes from h_m to h_c at the jointure, the severest case can be defined by use of the wall thickness of the metal;

$$S_s = 0.782 h^{1/2} a^{-1/2} E (\alpha_m - \alpha_c) \Delta T. \tag{11}$$

Total Shear Force—The total shear force, F, is determined by multiplying Equation (9b) by the mean circumference;

$$F = 1.564 \pi h^{3/2} a^{1/2} E (\alpha_m - \alpha_c) \Delta T. \tag{12}$$

Hoop Stress—By means of Timoshenko's expressions (Reference (2); p. 391), the hoop forces, N_ρ, are

$$N_\rho = \frac{Ehw}{a}. \tag{13a}$$

If the deflection of the ceramic is ignored, as in the approximate solution, the radial displacement is given by

$$w_m = \frac{1}{2\beta^3 D} (\beta M + Q). \tag{14}$$

By substitution of the deflection of the metal (w_m) and the definition of β for w and β in Equation (13a), the hoop stress can be expressed in the terms of the seal dimensions as

$$N_\rho = 4Eh(\alpha_m - \alpha_c)\Delta T. \tag{13b}$$

When this expression is divided by the wall thickness of the metal for the severest case, the equation for hoop stress becomes

$$S_h = 4E(\alpha_m - \alpha_c)\Delta T. \tag{15}$$

Bending Stress—Timoshenko (Reference (1), p. 54) has defined the bending stress as

$$S_b = \frac{6M}{h^2}. \tag{16a}$$

If M is expressed as in Equation (10b), the bending stress is

$$S_b = 4.37E(\alpha_m - \alpha_c)\Delta T. \tag{16b}$$

It is interesting to note that S_b and S_h are nearly equal at $x = 0$; however, the bending stresses in the ceramic increase to a maximum at the middle of the ceramic, while the hoop stress may decrease.

STRESSES WITHIN THE CERAMIC

The stresses considered thus far have been limited to the jointure between the ceramic and the metal. However, seal-assembly failure is frequently caused by cracks in the interior of the ceramic.

The general equation for the deflection, w, of a short cylinder loaded by a uniform radial shear and radial moment at the end is

$$w = C_1 \sin\beta l \sinh\beta l + C_2 \sin\beta l \cosh\beta l + C_3 \cos\beta l \sinh\beta l + C_4 \cos\beta l \cosh\beta l. \tag{17}$$

For convenience, the origin is shifted to the middle of the cylinder, as shown in Figure 2. With such a shift of origin, the jointure is at $l = L/2$, and Equation (17) is an even function of l; therefore, C_2 and C_3 are zero, and the second and third terms drop out. The constants C_1 and C_4 are then selected to satisfy the end conditions,

$$M_{l=L/2} = D\left(\frac{d^2w}{dl^2}\right)_{l=L/2} = M \tag{18}$$

$$Q_{l=L/2} = -D\left(\frac{d^3w}{dl^3}\right)_{l=L/2} = Q \tag{19}$$

where Q and M are determined by solution of Equations (5) and (6). Equation (17) can then be rewritten and differentiated;

$$w = C_1 \sin\beta l \sinh\beta l + C_4 \cos\beta l \cosh\beta l \tag{20}$$

$$\frac{dw}{dl} = C_1\beta(\sin\beta l\cosh\beta l + \cos\beta l\sinh\beta l) + C_4\beta(\cos\beta l\sinh\beta l - \sin\beta l\cosh\beta l), \tag{21}$$

$$\frac{d^2w}{dl} = 2C_1\beta^2(\cos\beta l\cosh\beta l) - 2C_4\beta^2(\sin\beta l\sinh\beta l), \tag{22}$$

$$\frac{d^3w}{dl} = 2C_1\beta^3(\cos\beta l\sinh\beta l - \sin\beta l\cosh\beta l) - 2C_4\beta^3(\sin\beta l\cosh\beta l + \cos\beta l\sinh\beta l). \tag{23}$$

If the second and third derivatives, Equations (22) and (23), are substituted into the end conditions, and if, for the sake of simplicity, α is set equal to $\beta l/2$, the following equations are obtained:

$$M_{l=L/2} = -2C_1D\beta^2(\cos\alpha\cosh\alpha) + 2C_4D\beta^2(\sin\alpha\sinh\alpha), \tag{24}$$

and

$$Q_{l=L/2} = -2C_1D\beta^3(\cos\alpha\sinh\alpha - \sin\alpha\cosh\alpha) + 2C_4D\beta^3(\sin\alpha\cosh\alpha + \cos\alpha\sinh\alpha). \tag{25}$$

The constants C_1 and C_4 in Equations (24) and (25) are given by

$$ {}_1 = \frac{1}{2D\beta^2}\left\{\frac{Q\sin\alpha\sinh\alpha}{\beta[\cos\alpha\cosh\alpha(\sin\alpha\cosh\alpha+\cos\alpha\sinh\alpha)-\sin\alpha\sinh\alpha(\cos\alpha\sinh\alpha-\sin\alpha\cosh\alpha)]}\right. $$
$$ \left. -\frac{M(\sin\alpha\cosh\alpha+\cos\alpha\sinh\alpha)}{\cos\alpha\cosh\alpha(\sin\alpha\cosh\alpha+\cos\alpha\sinh\alpha)-\sin\alpha\sinh\alpha(\cos\alpha\sinh\alpha-\sin\alpha\cosh\alpha)}\right\} \qquad (26) $$

nd

$$ {}_4 = \frac{1}{2D\beta^2}\left\{\frac{Q\cos\alpha\cosh\alpha}{\beta[\cos\alpha\cosh\alpha(\sin\alpha\cosh\alpha+\cos\alpha\sinh\alpha)-\sin\alpha\sinh\alpha(\cos\alpha\sinh\alpha-\sin\alpha\cosh\alpha)]}\right. $$
$$ \left. -\frac{M(\cos\alpha\sinh\alpha-\sin\alpha\cosh\alpha)}{\cos\alpha\cosh\alpha(\sin\alpha\cosh\alpha+\cos\alpha\sinh\alpha)-\sin\alpha\sinh\alpha(\cos\alpha\sinh\alpha-\sin\alpha\cosh\alpha)}\right\} \qquad (27) $$

These expressions for C_1 and C_4 can then be substituted in Equations (24) and (25) to determine the stresses anywhere within the ceramic.

The point of greatest interest within the ceramic in a butt seal is at $l = 0$. If $l = 0$ is substituted in Equations (24) and (25), the following results are obtained:

$$ \begin{aligned} \cos\alpha\cosh\alpha &= 1 \\ \sin\alpha\sinh\alpha &= 0 \\ \cos\alpha\sinh\alpha &= 0 \\ \sin\alpha\cosh\alpha &= 0 \end{aligned} \qquad (28) $$

The bending moment, M, at the center is then given by

$$ M_{l=0} = \frac{\left(\dfrac{Q_{l=L/2}}{\beta}\right)\left(\sin\dfrac{\beta L}{2}\sinh\dfrac{\beta L}{2}\right) + M_{l=L/2}\left(\sin\dfrac{\beta L}{2}\cosh\dfrac{\beta L}{2} + \cos\dfrac{\beta L}{2}\sinh\dfrac{\beta L}{2}\right)}{\dfrac{(\sin\beta L + \sinh\beta L)}{2}}. \qquad (29) $$

The shear force per linear inch, Q, at the center is given by

$$ Q_{l=0} = 0. \qquad (30) $$

RESULTS

Solutions to Equations (9b), (10b), (11), and (12) are plotted in Figures 3, 4, and 5 for three molybdenum-to-alumina seals of different diameters that were brazed at 1235°C. The diameters chosen are ¾, 1½, and 2 inches. Several conclusions can be drawn from the curves and the equations: (1) the shear stresses increase with wall thickness of the metal, but decrease with increasing diameters; (2) the bending

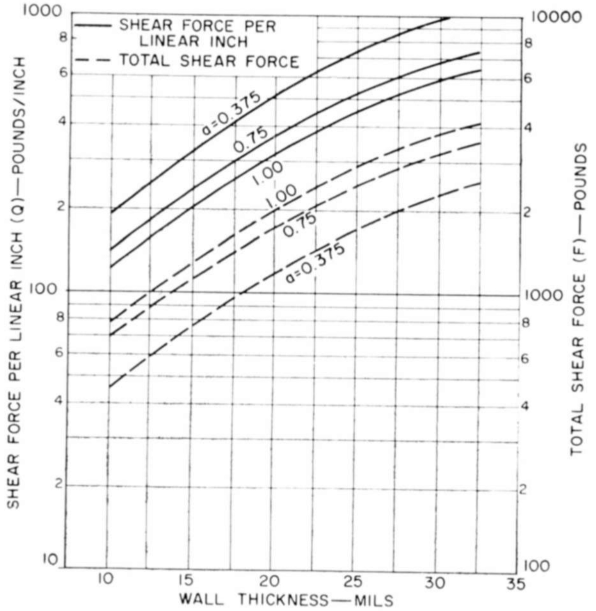


Fig. 3—Solutions to Equation (9b), $Q = 0.782h^{3/2}a^{-1/2}E_m(\alpha_m - \alpha_c)\Delta T$, and Equation (12), $F = 1.564\pi h^{3/2}a^{1/2}E_m(\alpha_m - \alpha_c)\Delta T$.

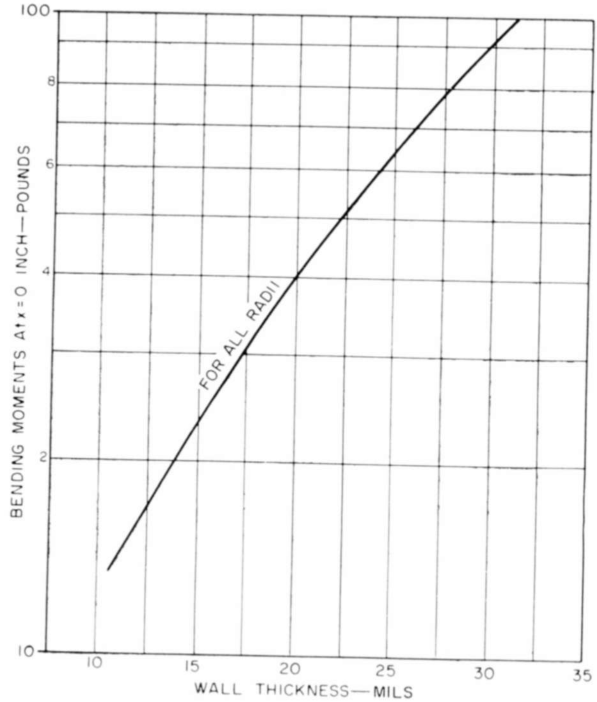


Fig. 4—Solution to Equation (10b), $M = 0.726h^2E_m(\alpha_m - \alpha_c)\Delta T$.

moments increase rapidly with wall thickness, but are independent of diameter; and (3) the hoop and bending stresses are solely dependent upon choice of materials, and are independent of the seal dimensions.

Solutions to Equations (5), (6), and (29) for ceramics of different lengths are illustrated in Figure 6 for a seal with a diameter of 1½ inches and a ceramic wall thickness of 0.065 inch. The ceramic is brazed to a molybdenum cylinder at 1235°C. As the length of the ceramic is increased from 0.033 to 0.200 inch, the stresses at the jointure decrease very rapidly. Increasing the length of the ceramic beyond 0.200 inch does not reduce the stresses at the jointure appreciably, although a minimum occurs at 0.300 inch. For ceramics shorter than 0.750 inch, the bending moments at the middle, $l = 0$, are greater than the bending moments at the end, $l = L/2$. There is a short initial drop in the bending moments at the middle of the ceramic for ceramics of 0.066 to 0.100 inch in length, but there is no further significant reduction in the bending moments in the middle until a ceramic length of 0.500 inch is assumed. The shear stresses at the ends decrease in the same manner as the bending moments, but at the middle of the ceramic, shear stress is zero for all lengths, as shown previously by substitution of $l = 0$ in Equation (25).

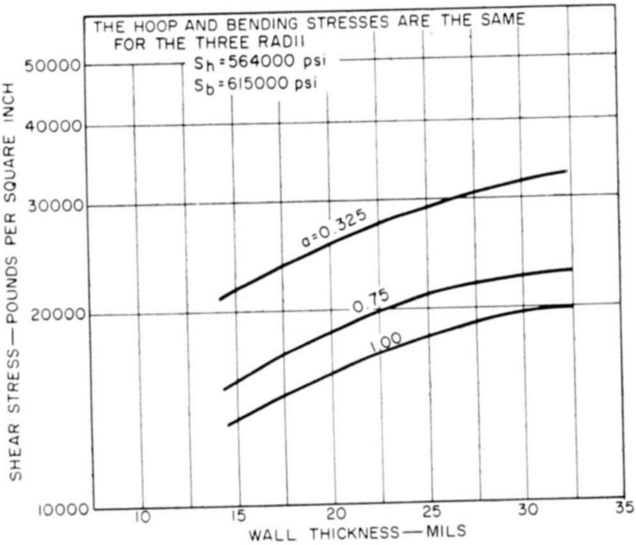


Fig. 5—Solution to Equation (11), $S_s = 0.782 h^{1/2} a^{-1/2} E_m (\alpha_m - \alpha_c) \Delta T$.

DISCUSSION

The results summarized above are contrary to general conclusions drawn in the past, which state that as the total mismatch is increased the shear stresses become greater and cause a seal to fail. The derived equations describe a less severe stress condition as the diameter is increased, in spite of increasing shear forces. The shear forces increase as the square root of the radius, as indicated in Equation (12). This fact helps to explain some of the difficulties encountered in making larger seals.

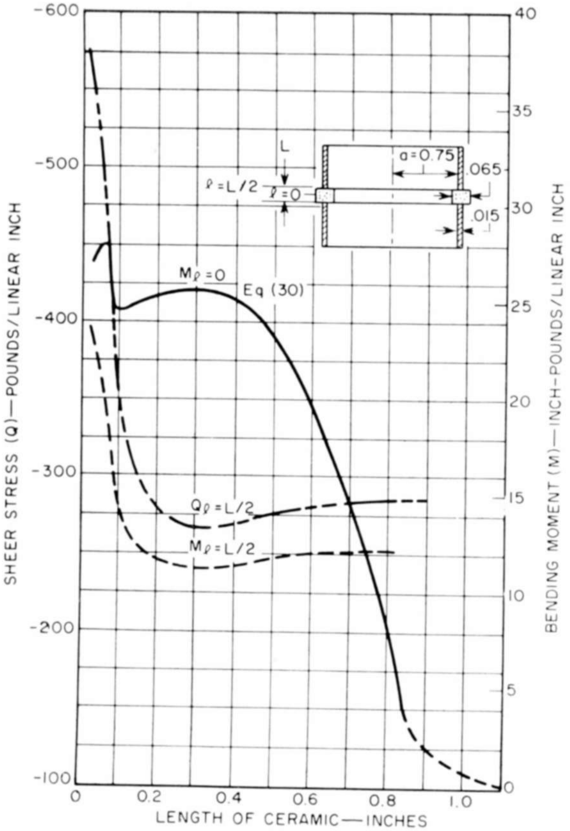


Fig. 6—Shear stress and bending moments for ceramics of different lengths.

It has been observed that most difficulties in making high-temperature seals for thermionic energy converters are not related to the characteristic failures in power-tube seals. There is always excellent solder flow and filleting, and the seals have a leak rate of less than 10^{-7} cm^3/sec. When a cross section is made of the seal, cracks cannot always be found in the ceramic; if cracks are formed, they are relatively close to the metalizing and do not extend all the way through the body. Sometimes separations are observed at the grain boundaries in the fillet, as shown in Figure 7. This latter characteristic is particularly noticeable in brazes of 60% nickel and 40% palladium.

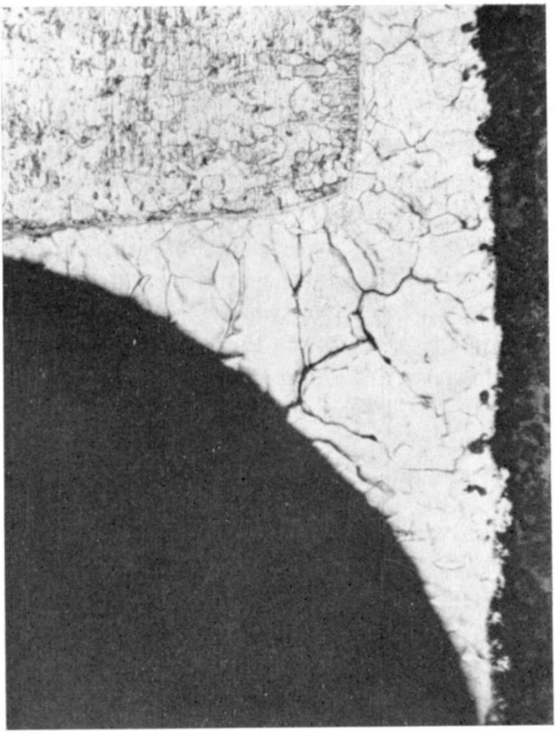


Fig. 7—Nickel-palladium braze between molybdenum and an alumina ceramic showing voids or cracks in the brazing materials (200×).

Low-Melting-Point Solders

The solders used for making the high-temperature brazes were eutectics; their composition is very nearly 50-50 atomic per cent. A characteristic of a simple atomic ratio is a tendency toward long-range order; that is, atoms arrange themselves in the lattice so that their nearest neighbors are unlike themselves. Such structures have lower ductility than a structure with random arrangements of the atoms. If the solder lacks ductility, it is not able to relieve the stresses caused by thermal mismatch of materials.

Equi-Cohesive Temperature

Solders also may lack ductility near their melting point because they are above their equi-cohesive temperature.* Above the equi-cohesive temperature, strain occurs by grain rotation, which causes void formation along the grain boundaries as shown in Figure 7.

Distribution of Stress

It was assumed in the derivation that the stresses on a seal are uniform. This condition may not be true in actual practice, and the larger the diameters the greater may be the nonuniform stresses. As the diameter of a seal is increased, circumferential temperature gradients and variations in the solder gap are greater. A circumferential temperature gradient that occurs while the solder is solidifying causes all the shear forces to act through the metal, which is last to solidify. The solder gap may vary if the mating parts of the seal are not flat. Flatness faults can result from poor machining of parts or from warpage of the ceramic during metalizing. If the solder gap is large, the braze material usually fills it and increases its thickness in this region. If there is a variation in the solder gap, the solder thickness can no longer be ignored (as was done in deriving the theory) because most brazing alloys have thermal expansions considerably greater than either member being joined. Therefore, circumferential temperature gradients and variations in the solder gap create an imbalance force that acts on an assembly while it is cooling from the brazing temperature or during any subsequent thermal cycles.

Shear Force During Brazing

There may be insufficient weight on a seal above the equi-cohesive temperature to prevent the metal member from sliding relative to the ceramic. Figure 8 is a vector diagram of the forces acting on a seal during solidification of the braze joint. The shear force to the left is the force calculated by use of Equation (12); the downward force is

---

* The term equi-cohesive temperature is an old metallurgical term that is seldom used in modern technology; however, the term best describes a state that exists just below the melting point of a metal where the grain boundaries are weaker than the bulk of the grain. At temperatures very close to the melting point, the metal will deform as though the grain boundaries were molten.[3]

[3] N. J. Grant, "Stress Rupture Testing," p. 41, *High-Temperature Properties of Metals*, American Society for Metals, 1950.

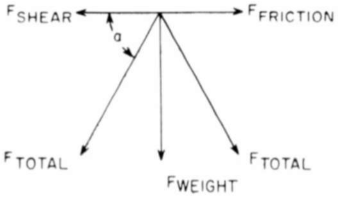


Fig. 8—Forces acting on a seal during solidification of the brazed joint.

an empirical force resulting from the weight applied to the seal during brazing. Countering the shear force is a frictional force that, if not sufficiently great, allows the grain boundaries to rupture as the ceramic and metal cylinder slide as a result of their different thermal expansions.

Throughout the previous discussion, it was assumed that the brazing alloy was near its melting point, because such an assumption fits the observations made in analyzing seal failures. However, the strength of the ceramic should not be overlooked. It is known that the silica-free ceramic bodies used in making the high-temperature seals are relatively weak at room temperature (35,000 to 40,000 psi), but little is known of the effect of increasing temperature on the strength of alumina ceramic. From the meager data available, it is estimated that most alumina bodies are only one-half as strong at 800°C as they are at room temperature. In addition, little is known of the creep properties of ceramics at very high temperatures (greater than 1000°C), where the grains are bonded by a plastic, glassy phase.

ACKNOWLEDGMENTS

I. E. Martin's assistance throughout the derivation of the equations presented in this paper is gratefully acknowledged.

MULTIPLE-INPUT LARGE-SIGNAL MIXER ANALYSIS

By

PETER TORRIONE AND SHUI YUAN

RCA Communications Systems Division
New York, N. Y.

Summary—*Theoretical analyses on resistive mixers generally assume (1) that the input to the mixer consists of only two voltages, that of the local oscillator and one signal; and (2) either that both signal and local oscillator are small but comparable in amplitude (Taylor Series method), or that the signal is small compared to the local oscillator (small-signal analysis).*

The method of intermodulation analysis described in this paper is a large-signal analysis that includes the case of any number of inputs, i.e., the local oscillator, signal voltages, and undesired or jamming inputs. The most important advantage of this analysis is that it yields a completely closed-form representation for the mixer responses by means of the generalized power relations for resistive mixers with multiple inputs. This is in contrast to the Taylor Series representation of the diode characteristic, which gives approximate solutions for small signals and local oscillator and requires a great many terms (more than 10) if the voltages across the diode are large. When the Bessel representation of the diode's spurious response is used, it is only necessary to decide the responses of interest and the frequencies at which power is being dissipated. From these, the equivalent circuit may be determined and the spurious outputs calculated.

The spurious responses generated by the mixer with one signal and local oscillator present can be minimized by judicious choice of local oscillator and output frequencies. In the more general case of multiple inputs, the distortion and spurious responses produced cannot be eliminated by choice of frequencies or filtering.

The importance of the multiple-input analysis is that it is able to predict the magnitudes of intermodulation and cross modulation for large input voltages.

INTRODUCTION

R. H. PANTELL[1] AND C. H. PAGE[2] developed some general power relationships for a nonlinear-resistance frequency converter under large local-oscillator voltages and small-signal voltages. Pantell's analysis is limited to cases where the diode con-

---

[1] Richard H. Pantell, "General Power Relationships for Positive and Negative Nonlinear Resistive Elements," *Proc. IRE*, pp. 1910-1913, Vol. 46, Dec. 1958.

[2] C. H. Page, "Frequency Conversion With Positive Nonlinear Resistors," *Journal of Research of the National Bureau of Standards*, Vol. 56, No. 4, p. 179, April 1956.

ductance varies only as a function of the local-oscillator voltage and is independent of signal level. Also, the only frequencies of interest are d-c, the local-oscillator and signal frequencies, and the intermediate (difference) frequency output.

In the following analysis it is possible to take into account any number of inputs signals using no small-signal approximations. In this way, it is possible to determine the distortion at the desired signal frequency caused by any number of undesired inputs. For simplicity the analysis is derived using three inputs (two signal inputs and the local oscillator). The analysis is extended by analogy to any number of inputs.

GENERAL POWER RELATIONSHIPS

The voltages and currents across a nonlinear element can be expressed as the sums and differences of the various harmonics of the input frequencies. For the case of three input frequencies, the Fourier series for the voltages and currents are

$$v = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} V_{lmn} \exp\{j(lx + my + nz)\} \tag{1}$$

$$i = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} I_{lmn} \exp\{j(lx + my + nz)\} \tag{1a}$$

where $x = \omega_0 t =$ the local-oscillator frequency,

$y = \omega_1 t =$ the desired signal frequency,

$z = \omega_2 t =$ the extraneous signal frequency.

Since i and v represent real functions of time,

$$\begin{aligned} V_{lmn} &= V^*_{-l-m-n}, \\ I_{lmn} &= I^*_{-l-m-n}, \end{aligned} \tag{2}$$

where the asterisk denotes the complex conjugate.

The average real power of a frequency $(l\omega_0 + m\omega_1 + n\omega_2)$ is

$$\begin{aligned} W_{lmn} &= 2 \text{ Re } [V_{lmn} I^*_{lmn}] \\ &= [V_{lmn} I^*_{lmn} + V^*_{lmn} I_{lmn}]. \end{aligned} \tag{3}$$

The coefficients of Equations (1) and (1a) can be found by a standard three-dimensional Fourier analysis.

$$V_{lmn} = \frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} v \exp\{-j(lx + my + nz)\}\, dx\, dy\, dz, \quad (4a)$$

and

$$I_{lmn} = \frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} i \exp\{-j(lx + my + nz)\}\, dx\, dy\, dz, \quad (4b)$$

where v is the voltage across the diode as a function of time and i is the current variation in the diode with time.

From the above relations an equation involving $V_{lmn}I^*_{lmn}$ can be written;

$$\sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} - l^2 V_{lmn} I^*_{lmn}$$

$$= \frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} v \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} - l^2 I^*_{lmn} \exp\{-j(lx + my + nz)\}\, dx\, dy\, dz.$$

$$(5)$$

The above expression can be simplified by substituting

$$\frac{\partial^2 i}{\partial x^2} = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} - l^2 V^*_{lmn} \exp\{-j(lx + my + nz)\}.$$

Equation (5) now becomes

$$\sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} - l^2 V_{lmn} I^*_{lmn} = \frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} v \left(\frac{\partial^2 i}{\partial x^2}\right) dx\, dy\, dz. \quad (6)$$

The integral on the right is integrated by parts. These integrals become

$$\frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi} \left(v \frac{di}{dx} \bigg|_0^{2\pi} - \int_0^{2\pi} \frac{\partial v}{\partial x} \frac{di}{dx}\, dx \right) dy\, dz.$$

Since v and i are periodic functions of x, y, and z,

$$v \frac{di}{dx} \Bigg|_{0}^{2\pi} = 0$$

and the right hand side of Equation (5) becomes

$$\left(\frac{1}{2\pi}\right)^3 \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{di}{dv}\left(\frac{dv}{dx}\right)^2 dx\, dy\, dz = h_l. \tag{7}$$

From Equations (3) and (7),

$$\sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} l^2 V_{lmn} I^*_{lmn} = \sum_{l=0}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} l^2 [V_{lmn} I^*_{lmn} + I_{lmn} V^*_{lmn}]$$

$$= \sum_{l=0}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} l^2 W_{lmn} = h_l = \frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{\partial i}{\partial v}\left(\frac{\partial v}{\partial x}\right)^2 dx\, dy\, dz. \tag{8a}$$

Similar relationships for h_m and h_n can be developed:

$$\sum_{l=-\infty}^{\infty} \sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} m^2 W_{lmn} = h_m = \frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{\partial i}{\partial v}\left(\frac{\partial v}{\partial y}\right)^2 dx\, dy\, dz, \tag{8b}$$

and

$$\sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=0}^{\infty} n^2 W_{lmn} = h_n = \frac{1}{8\pi^3} \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{\partial i}{\partial v}\left(\frac{\partial v}{\partial z}\right)^2 dx\, dy\, dz. \tag{8c}$$

By definition a positive nonlinear resistor is an element with a current–voltage characteristic such that $di/dv = g(v) \geq 0$ for all v, where $g(v)$ is the conductance. From Equations (8),

$$\left.\begin{aligned} h_l &= \sum_{l=0}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} l^2 W_{lmn} \geq 0 \\ h_m &= \sum_{l=-\infty}^{\infty} \sum_{m=0}^{\infty} \sum_{m=-\infty}^{\infty} m^2 W_{lmn} \geq 0 \\ h_n &= \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=0}^{\infty} n^2 W_{lmn} \geq 0 \end{aligned}\right\} \qquad (9)$$

as derived by Pantell.[1]

The conductance variation with time can be expressed as a three-dimensional Fourier series in the three independent input frequencies;

$$g(t) = \frac{\partial i}{\partial v} = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} G_{uvw} \exp\{j(ux + vy + wz)\}. \qquad (10)$$

Also, by differentiating Equation (1) with respect to x and squaring,

$$\left(\frac{\partial v}{\partial x}\right)^2 = \sum_{lmnpqr=-\infty}^{\infty} -(lp) V_{lmn} V_{pqr} \exp\{j(lx + my + nz\} \exp\{j(px + qy + nz)\}. \qquad (11)$$

Substituting Equations (10) and (11) into Equation (8),

$$h_l = \frac{1}{8\pi^3} \int_0^{2\pi}\!\!\int_0^{2\pi}\!\!\int_0^{2\pi} \sum_{lmnpqr=-\infty}^{\infty} \sum_{uvw=-\infty}^{\infty} -(lp) G_{uvw} V_{lmn} V_{pqr}$$

$$\exp\{j(l+p+u)\,x + (m+q+v)\,y + (n+r+w)\,z\}\,dx\,dy\,dz. \qquad (12)$$

Because the exponential function is orthogonal over the period 2π, the integrals exist only when

$$\left.\begin{aligned} l + p + u &= 0 \\ m + q + v &= 0 \\ n + r + w &= 0 \end{aligned}\right\} \qquad (13)$$

Equation (12) thus simplifies to

$$h_l = \sum_{lmnpqr=-\infty}^{\infty} -(lp) V_{lmn} V_{pqr} G_{(-l-p)(-m-q)(-n-r)}. \qquad (14a)$$

Similarly,

$$h_m = \sum_{lmnpqr=-\infty}^{\infty} -(mq)\, V_{lmn} V_{pqr} G_{(-l-p)(-m-q)(-n-r)}, \tag{14b}$$

and

$$h_n = \sum_{lmnpqr=-\infty}^{\infty} -(nr)\, V_{lmn} V_{pqr} G_{(-l-p)(-m-q)(-n-r)}. \tag{14c}$$

Equations (14) are the generalized power relations for the case of three inputs (two signals and the local oscillator) to the mixer. When there are n inputs ($n-1$ signals and the local oscillator), Equations (14) become

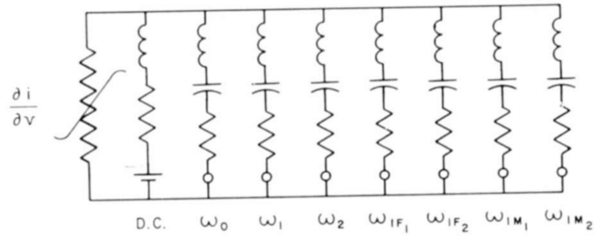
$$h_{l_1} = \sum_{l=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \cdots \sum_{l_n=-\infty}^{\infty} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \cdots \sum_{p_n=-\infty}^{\infty} -(l_1 p_1)\, V_{l_1 l_2 \ldots l_n} V_{p_1 p_2 \ldots p_n} G_{(-l_1-p_1)(-l_2-p_2)\ldots(-l_n-p_n)} \tag{15a}$$

and thus

$$h_{l_m} = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \cdots \sum_{l_n=-\infty}^{\infty} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \cdots \sum_{p_n=-\infty}^{\infty} -(l_m p_m)\, V_{l_1 l_2 \ldots l_n} V_{p_1 p_2 \ldots p_n} G_{(-l_1-p_1)(-l_2-p_2)\ldots(-l_n-p_n)}. \tag{15b}$$

SPECIFIC EXPRESSIONS FOR A PRACTICAL MIXER WITH THREE INPUTS

When there is more than one signal at the input of the mixer, the desired output is distorted by the intermodulation and cross-modulation products. For the case of two inputs, the mixer can be represented by the following equivalent circuit



where ω_0 = radian frequency of local oscillator

ω_1 = radian frequency of signal 1

$$\omega_2 = \text{radian frequency of signal 2}$$

$$\omega_{IF1} = \text{radian frequency of } \omega_1 - \omega_0$$

$$\omega_{IF2} = \text{radian frequency of } \omega_2 - \omega_0$$

$$\omega_{IM1} = \text{radian frequency of } 2\omega_1 - \omega_2 - \omega_0$$

$$\omega_{IM2} = \text{radian frequency of } 2\omega_2 - \omega_1 - \omega_0.$$

The general power relations for this mixer are expressed by Equations (14). When the mixer is represented by the equivalent circuit, the power is dissipated only at those frequencies that have an appreciable admittance. The general power relations can be simplified, and using the fact that for resistive devices

$$V_{lmn} = V_{-l-m-n} = V^*_{-l-m-n}$$

and

$$I_{lmn} = I_{-l-m-n} = I^*_{-l-m-n}$$

Equations (8b) and (14b) give

$$\begin{aligned}
&4W_{-12-1} + W_{-110} + W_{010} + W_{11-2} \\
&\quad = 4V_{-12-1}\,[-(2V_{-12-1}G_{2-42} + V_{-110}G_{2-31} + V_{010}G_{1-31} + V_{11-2}G_{0-33}) \\
&\qquad + (2V_{1-21}G_{000} + V_{1-10}G_{0-11} + V_{0-10}G_{1-11} + V_{1-12}G_{2-1-1})\,] \\
&\quad + 2V_{-110}\,[-(2V_{-12-1}G_{2-31} + V_{-110}G_{2-20} + V_{010}G_{1-20} + V_{11-2}G_{0-22}) \\
&\qquad + (2V_{1-21}G_{01-1} + V_{1-10}G_{000} + V_{0-10}G_{100} + V_{-1-12}G_{20-2})\,] \\
&\quad + 2V_{010}\,[-(2V_{-12-1}G_{1-31} + V_{-110}G_{1-20} + V_{010}G_{0-20} + V_{11-2}G_{-1-22}) \\
&\quad + (2V_{1-21}G_{-11-1} + V_{1-10}G_{-100} + V_{0-10}G_{000} + V_{-1-12}G_{10-2})\,] \\
&\quad + 2V_{11-2}\,[-(2V_{-12-1}G_{0-33} + V_{-110}G_{0-22} + V_{010}G_{-1-22} + V_{11-2}G_{-2-24}) \\
&\qquad + (2V_{1-21}G_{-211} + V_{1-10}G_{-202} + V_{0-10}G_{-102} + V_{-1-12}G_{000})\,].
\end{aligned} \tag{16}$$

Examining Equation (16), it is seen that each term on the right hand side of the equation is equal to a corresponding term on the left hand side of the equation. For example,

$$\begin{aligned}
&4W_{-12-1} = 4V_{-12-1}[-2V_{-12-1}G_{2-42} + V_{-110}G_{2-31} + V_{010}G_{1-31} + V_{11-2}G_{0-33}) \\
&\quad + (2V_{1-21}G_{000} + V_{1-10}G_{0-11} + V_{010}G_{1-11} + V_{-1-12}G_{2-1-1})\,].
\end{aligned} \tag{17}$$

Since, from Equations (3) and (16),

$$W_{lmn} = 2V_{lmn}I_{lmn}, \tag{18}$$

Equation (17) becomes

$$\begin{aligned} I_{-12-1} &= \frac{V_{010}}{2}(G_{1-11} - G_{1-31}) + \frac{V_{-110}}{2}(G_{01-1} - G_{2-31}) \\ &+ V_{1-21}(G_{000} - G_{2-42}) + \frac{V_{11-2}}{2}(G_{2-1-1} - G_{0-33}). \end{aligned} \tag{19a}$$

Similarly

$$\begin{aligned} I_{-110} &= V_{010}(G_{100} - G_{1-20}) + V_{-110}(G_{000} - G_{2-20}) \\ &+ 2V_{1-21}(G_{01-1} - G_{2-31}) + V_{11-2}(G_{20-2} - G_{0-22}) \end{aligned} \tag{19b}$$

$$\begin{aligned} I_{010} &= V_{010}(G_{000} - G_{0-20}) + V_{-110}(G_{-100} - G_{1-20}) \\ &+ 2V_{1-21}(G_{-11-1} - G_{1-31}) + V_{11-2}(G_{10-2} - G_{-1-22}) \end{aligned} \tag{19c}$$

and

$$\begin{aligned} I_{11-2} &= V_{010}(G_{-102} - G_{-1-22}) + V_{-110}(G_{-202} - G_{0-22}) \\ &+ 2V_{1-21}(G_{-211} - G_{0-33}) + V_{11-2}(G_{000} - G_{-2-24}). \end{aligned} \tag{19d}$$

Thus if the external circuit constraints are specified, the circuit operating conditions and its parameters are completely determined when the various conductances are calculated.

EXPRESSIONS FOR CONDUCTANCES

The conductance of a nonlinear resistance diode as expressed by Equation (10) is, of course, a function of the current–voltage characteristics of the diode. For an exponential diode,

$$i = I_{\text{SAT}}(e^{\alpha v} - 1)$$

$$\frac{\partial i}{\partial t} = g(t) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} G_{uvw} \exp\{j(ux + vy + wz)\}. \tag{20}$$

For large-signal analysis, conductance variation is a function of the local oscillator voltage and the two signal voltages. Equation (20) is therefore,

$$\frac{\partial i}{\partial v} = \alpha I_{\mathrm{SAT}} \exp\{\alpha (V_0 \cos x + V_1 \cos y + V_2 \cos z)\}$$

$$= \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} G_{uvw} \exp\{j(ux + vy + wz)\}$$

$$= \sum_{u=-\infty}^{\infty} G_u \exp\{jux\} \sum_{v=-\infty}^{\infty} G_v \exp\{jvy\} \sum_{w=-\infty}^{\infty} G_w \exp\{jwz\} \qquad (21)$$

where $G_{uvw} = G_u G_v G_w$. G_u, G_v and G_w can be solved by Fourier series analysis;

$$g(t) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} G_{uvw} \exp\{j(ux + vy + wz)\}$$

$$= \alpha I_{\mathrm{SAT}} \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} I_u(\alpha V_0)\, I_v(\alpha V_1)\, \mathrm{I}_w(\alpha V_2) \exp\{j(ux + vy + wz)\}. \qquad (22)$$

Hence, the G's in Equations (16) are determined as functions of the voltages across the nonlinear element of the diode. For example,

$$G_{120} = [I_1(\alpha V_0)\, I_2(\alpha V_1)\, I_0(\alpha V_2)]\alpha I_{\mathrm{SAT}} \qquad (23)$$

where $I_u(\alpha V)$ is the modified Bessel's function of the first kind, u^{th} order, and argument (αV).

CONCLUSION

Although this analysis is limited to three inputs (two signals and local oscillator), similar expressions can be derived for n inputs. The generalized power relations for n inputs are expressed by Equations (15). These power relations provide a means of calculating the output power at any frequency of interest. Thus, if the intermodulation ratio is to be determined for the case of three inputs, W_{121} and W_{110} must be calculated;

$$\frac{\mathrm{IM}}{\mathrm{IF}}=\frac{W_{121}}{W_{110}}.$$

In general W_{lmn} is determined by the following procedure.

(1) Obtain I_{SAT} and α of the diode by measuring its current–voltage characteristic (as in Vol. 15, Radiation Laboratory series).

(2) Arbitrary chose V_{100}, V_{010} and V_{001} as the local oscillator and the two signal voltages, respectively, across the nonlinear element of the diode.

(3) Calculate G_{uvw} according to Equation (22).

(4) Calculate W_{lmn} according to Equations (16) and (17).

(5) Calculate I_{lmn} according to Equations (18) and (19).

(6) Circuit operating conditions such as total input power, power loss, and impedance mismatch can be determined at each frequency completely with specified external circuit constraints. Note that before the power at the frequencies of interest can be calculated, the number of voltages at various frequencies that appear across the nonlinear resistance must be known so that the exact equivalent circuit can be obtained. Since the analysis starts out with voltage Equation (1), it is therefore accurate only if all voltages are taken into consideration. The advantage of this analysis is that it is applicable to any mixer operating condition, i.e., small-signal, large-signal, and for any number of inputs.

GENERATION OF MICROWAVE POWER BY PARAMETRIC FREQUENCY MULTIPLICATION IN A SINGLE TRANSISTOR

By

Martin Caulton,* Harold Sobol,* and Robert L. Ernst†

Summary—*A UHF overlay transistor (2N3375) has been used to produce watts of power at microwave frequencies with good efficiency. By simultaneously operating the transistor as a power amplifier and the base–collector junction capacitance as a varactor frequency multiplier, substantial power has been delivered well above the transistor cutoff frequency. As an amplifier–frequency-doubler the transistor in a common-base circuit has yielded 3.5 watts of r-f power at 1000 mc, 1.8 watts at 1500 mc as an amplifier–tripler, and 0.3 watt at 2000 mc as an amplifier–quadrupler. For these results the input was driven with 1 watt of r-f power at 500 mc. As an oscillator–multiplier the transistor has delivered 3.3 watts at 920 mc and 1.5 watts at 1500 mc. These results are with the transistor in the conditionally stable common-base circuit. In a more stable common-emitter configuration, amplifier–doubler operation yielded powers at 1000 mc as high as 2.3 watts.*

A simplified analysis of the basic multiplication mechanism is presented, and the experiments in which the microwave power is produced are described. An equivalent circuit of the varactor section of the transistor was developed and the parameters characterizing the multiplication mechanism have been measured; this allows the deduction of features of the transistor that make its varactor action feasible. Predicted results based on the mechanisms described and the measured parameters agree well with experiment. Collector efficiencies as high as 40% have been observed. This application of the transistor is a realization of the integration of two previously separate devices in a single unit with an overall performance equal to that of the individual units.

Introduction

A SINGLE transistor has been used to produce watts of power at microwave frequencies with good efficiency. The 2N3375, a silicon n-p-n UHF power transistor, fabricated with the overlay structure[1] has produced this power at frequencies well above the $f_{\max}$ of the transistor. Power output at frequencies above $f_{\max}$ was obtained by simultaneously operating the transistor as a power amplifier and as a varactor multiplier.

---

* RCA Laboratories, Princeton, N. J.

† RCA Communications Systems Division, New York, N. Y.

[1] D. R. Carley, "A 5 Watt 500 MC Transistor," paper presented at PGED meeting, Washington, D.C., Nov. 1963.

Parametric amplification in transistors was reported in the literature several years ago,[2-4] but no appreciable power above the cutoff frequency was realized. Pulfer and Lindsay[5] reported the observation of simultaneous transistor amplification and varactor multiplication to produce one watt of power at UHF frequencies. Early in 1964 several workers[†] using the 2N3375 transistor as a multiplier reported appreciable power above $f_{\max}$. Subsequently, H. Lee[‡] using the transistor as a doubler in a common-base circuit obtained 3 watts of power at 1000 mc. The present paper[*] is a study of the amplification–parametric-multiplication mechanism and an investigation into the possibilities of using the transistor above the cutoff frequency in this manner. Experiments are described in which the microwave power is produced in both common-base and common-emitter circuits.

Mechanisms that can be used to amplify and multiply r-f power are illustrated in Figure 1. Figure 1(a) illustrates a power amplifier operating in a class-C mode; this nonlinear device can produce a current pulse rich in harmonics of the input-signal frequency. Only the desired frequency harmonic, $n\omega$, is loaded; the load admittance, Y, together with the output capacitance, C_0, presents a real component, $1/R_L$, to the amplifier at frequency $n\omega$. The other frequencies are unloaded. The output power at the loaded harmonic frequency, $n\omega$, is proportional to the square of the amplitude of the $n\omega$ component of amplifier output current. We will refer to the device of Figure 1(a) as the class-C multiplier. Figure 1(b) is a class-C power amplifier, but with a voltage-dependent capacitance $C(V)$, or varactor, across its output terminals. As in Figure 1(a) the desired frequency harmonic, $n\omega$, is loaded, but the other frequencies are either unloaded or have idlers ($Y(k\omega) = -jk\omega C_0$) provided for them. If the capacitance $C(V)$ is a strong function of V, the output harmonic power of Figure 1(b)

---

[2] R. Zuleeg and V. W. Vodicka, "Parametric Amplification Properties in Transistors," *Proc. IRE*, Vol. 48, p. 1785, Oct. 1960.

[3] V. W. Vodicka and R. Zuleeg, "Transistor Operation Beyond Cutoff Frequency," *Electronics*, p. 56, 26 Aug. 1960.

[4] J. Lindmayer and C. Wrigley, "On Parametric Amplification in Transistors," *Proc. IRE*, Vol. 49, p. 1335, Aug. 1961.

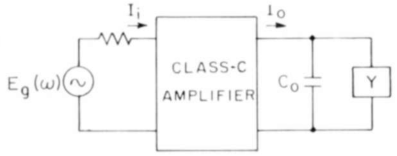
[5] J. K. Pulfer and A. E. Lindsay, "Simultaneous Amplification and Parametric Frequency Multiplication in VHF Power Transistors," *Proc. IEEE*, Vol. 52, p. 212, Feb. 1964.

[†] D. Snider, Lincoln Laboratories, communication in early 1964 to J. Hilibrand, RCA.

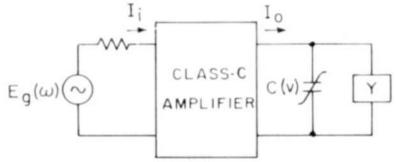
[‡] H. Lee, RCA Electronic Components and Devices, Somerville, N. J., private communication.

[*] This work was first reported at the IEEE Electron Devices Meeting in Washington, D.C., on October 31, 1964 as a paper entitled "A Transistor Microwave Power Source".

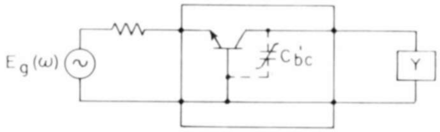
depends predominantly on the amplitude of the fundamental current, rather than on the harmonic current. Since the transistor–varactor combination requires gain only at the fundamental frequency, it is possible to obtain conversion gain at a harmonic frequency greater than the $f_{\max}$ of the transistor, provided that the varactor cutoff frequency is considerably in excess of the desired harmonic.



(a) CLASS-C MULTIPLIER



(b) AMPLIFIER AND VARACTOR MULTIPLIER



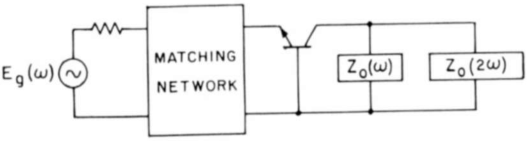
(c) COMBINATION TRANSISTOR – VARACTOR PARAMETRIC MULTIPLIER (TPM)

Fig. 1—Frequency-multiplication mechanisms.

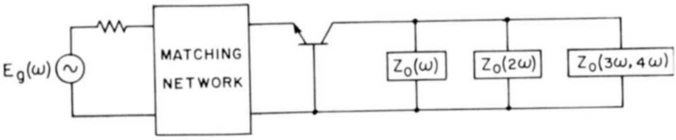
Figure 1(c) illustrates the mechanism responsible for the generation of the microwave powers described in this paper. The nonlinear capacitance or varactor diode shown separately in Figure 1(b) is eliminated and in its stead the collector–base capacitance, $C_{b'c}$, serves as the varactor. The device illustrated will be referred to as the transistor parametric multiplier (TPM). The transistor may be considered in two sections, one capable of gain at the fundamental frequency and a second section in which the capacitance $C_{b'c}$ serves as a varactor. Thus, even though the transistor is not capable of amplification at a harmonic frequency above the cutoff, conversion gain may be possible at the harmonic frequency as long as the transistor is capable of gain at the fundamental frequency. Similar power and loading considerations apply as with a separate varactor following a transistor amplifier.

EXPERIMENTS

The 2N3375 has been operated as a transistor-amplifier–parametric-multiplier (TPM) in both the common-base and common-emitter configurations. This section describes the circuits used and the powers and efficiencies obtained.



(a) DOUBLER CIRCUIT



(b) TRIPLER OR QUADRUPLER CIRCUIT

Fig. 2—Common-base multiplier circuits.

Common-Base Configuration

The common-base configuration with the appropriate a-c circuits is shown in Figure 2. The input circuit properly matches the generator to drive the transistor at the fundamental frequency. The output circuit of the doubler consists of (1) an idler that presents a reactance $j/(\omega C_0)$ across the base–collector capacitance at the fundamental frequency ω, where C_0 is the time-averaged value of this junction capacitance, and (2) a load that presents an impedance $j/(2\omega C_0) + R_L$ at the doubled frequency 2ω. Both double- and triple-stub tuners or lumped-constant π or T networks were used for the input circuit. Often, lumped-constant networks (as described by R. Minton[6]) together with stub tuners proved to be somewhat better than stub tuners alone. The reactive load that provided an idler for the fundamental frequency consisted of lumped elements at the transistor pins. The harmonic load circuit employed a triple-stub tuner for matching, or a tunable cavity with variable coupling and a termination.

---

[6] R. Minton, "Design of Large-Signal VHF Transistor Power Amplifiers," *RCA Applications Note SMA-36*, RCA Electronic Components and Devices, Harrison, N.J., July 1964.

The tripler or quadrupler circuit shown in Figure 2(b) employed an additional idling circuit at the doubled frequency and the proper resonance and loading at either 3ω or 4ω. Multiple-resonant circuits, such as stub tuners, were quite difficult to adjust, but traps to filter out unwanted frequencies before the load circuit made tuning easier.

Table I lists the powers, P_{om}, and collector efficiencies, $\eta_c = P_0/(I_c V_c)$, measured with the common-base circuit used as doubler, tripler, and quadrupler TPM's. Also included are the power output and efficiency for a simple class-C amplifier operating at 500 mc. In all cases the transistor was driven with 1 watt of power at 500 mc. The powers at the different frequencies were measured using a band-pass filter to allow only the desired frequency to reach the power meter. It is noted that collector efficiencies of over 40% were obtained for doubling to 1000 mc with a nominal collector-to-base voltage of 28 volts. With higher voltages, the efficiency fell, but powers of 3.6 watts at 1000 mc and 1.8 watts at 1500 mc were achieved.

Table I—Common-Base Multiplier Performance ($V_{bc} = 0$)

Circuit	f_{out} (mc)	P_{in} (500 mc) (watts)	P_{om} (watts)	V_{CB} (volts)	η_c (%)
Straight-through	500	1	4.8	28	44
Amplifier	500	1	6.7	42	42
Doubler	1000	1	2.6	28	41
(×2)	1000	1	3.6	42	30
	920	0	3.3	31	41
Tripler	1500	1	1.3	28	22
(×3)	1500	1	1.8	32	20
	1500	0	1.5	30	14
Quadrupler	2000	1	0.3	28	4
(×4)	2000	1	0.3	28	4

The common-base circuit is only conditionally stable, and under some tuning conditions the circuit went into oscillation. Appreciable powers were obtained under such conditions, namely 3.3 watts at 920 mc and 1.5 watts at 1500 mc (zero watts input in Table I). Under the 920-mc oscillation condition some radiation was noted at 460 mc, the fundamental frequency. There was also evidence of radiation at 1.5 times the fundamental frequency.

We did not achieve correspondingly large powers while operating the devices as a quadrupler TPM, but it is expected that the device is capable of yielding greater powers with a better output circuit. The circuitry employed in all of the above experiments was not sophisticated. Traps to filter out unwanted frequencies from each succeeding circuit element would aid in achieving powers closer to the maximum possible. In addition, the base–emitter junction was operated with zero d-c bias voltage at all times. Further optimization may be possible by setting a better operating point.

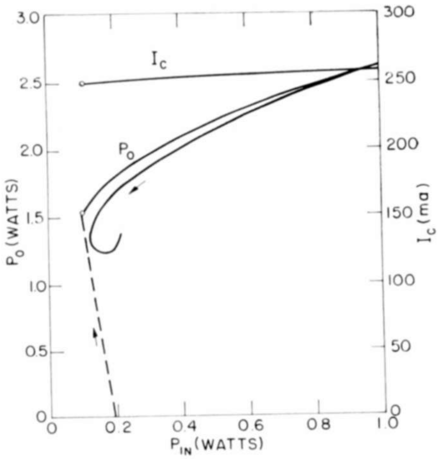


Fig. 3—Power output, P_o, and collector current, I_C, versus power input for common-base doubler.

The output power as a function of input power for the common-base transistor doubler is shown in Figure 3. The collector current as a function of input power is also shown on the same graph. The hysteresis and the nonlinear behavior of the output power indicate that the device may be behaving as a locked oscillator at the fundamental frequency. However, it was possible to obtain a nearly linear input–output characteristic under some tuning conditions giving lower power outputs. The "snapping in" of the collector current as shown in Figure 3 appeared to occur under most frequency-doubling conditions. Figure 4 presents the output versus input power curves and collector-current behavior for the tripler. Here the results indicate a stable class-C amplifier and multiplier in that both the power output and collector current vary smoothly with input power. Figure 5 shows output–input and collector-current characteristics for the device operating as a

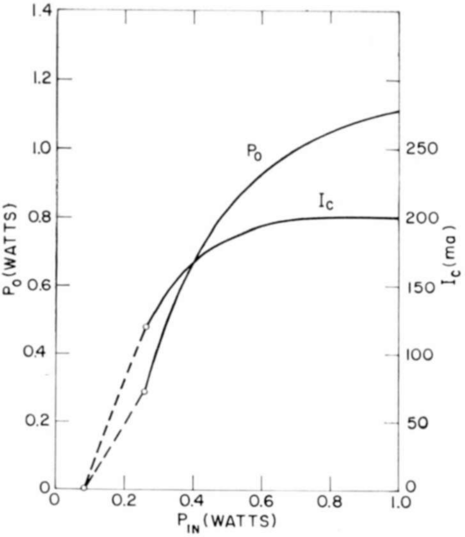


Fig. 4—Power output, P_o, and collector current, I_C, versus power input for common-base tripler.

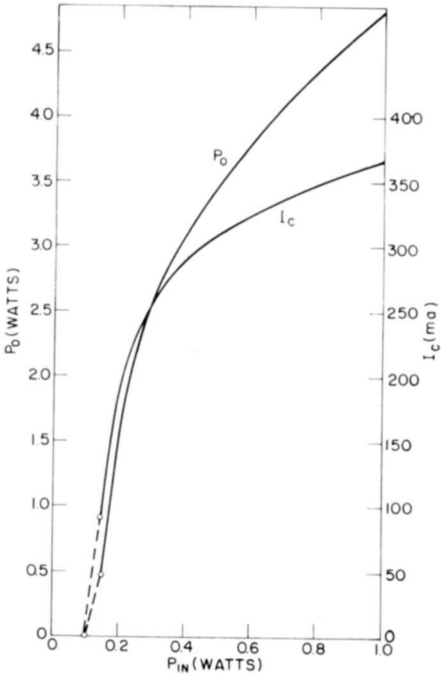


Fig. 5—Power output, P_o, and collector current, I_C, versus power input for common-base amplifier.

simple class-C amplifier at 500 mc. The amplifier and the amplifier–tripler both appeared to be fairly stable operating modes. The power curves of the quadrupler showed large hysteresis, indicative of either a locked oscillator or mode switching in varactors as described by McDade.[7,8]

Common-Emitter Configuration

Because of the conditional stability of the common-base operation, some effort was devoted to the use of the transistor in the more stable common-emitter configuration. The position of the base–collector capacitance $C_{b'c}$ in a common-emitter circuit is indicated in Figure 6.

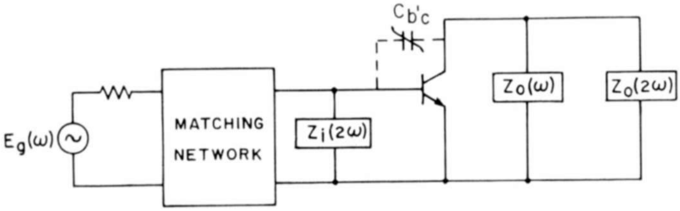


Fig. 6—Common-emitter doubler circuit.

A harmonic by-pass circuit is required to provide a low-impedance path that shunts the harmonic current around the base–emitter junction. This by-pass circuit, which is discussed later, can be in the form of a series *LC* circuit tuned at 1000 mc for the doubler. The rest of the circuit is in accordance with the considerations already described for the common-base configuration.

The power and efficiencies obtained on two transistors (common-emitter configuration) are indicated in Table II. Measured results for operation both as a simple class-C amplifier and as an amplifier doubler are listed. While the collector efficiency (29%) and the power output at 1000 mc (2.3 watts) are both lower than that of the common-base circuit, the common-emitter configuration is much more stable than that of the common-base circuit. Figure 7 shows the output power and collector current as a function of the input power for the circuit. Figure 8 shows the power and current behavior for a 500-mc common-emitter class-C amplifier. The smooth variation of output power and

---

[7] J. C. McDade, "Varactor Microwave Frequency Multipliers," ASTIA report AD282191, 1962.

[8] J. C. McDade, "Jump Phenomena in Varactor Diode Circuits," ASTIA report AD269919, 1962.

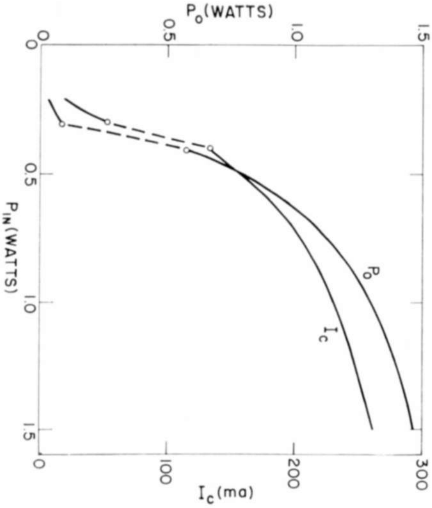


Fig. 7—Power output, P_o, and collector current, I_C, versus power input for common-emitter doubler.

collector current with input power in both Figures 7 and 8 indicate stable operation for both the class-C amplifier and the doubler TPM. The comparison of amplifier and doubler efficiencies is discussed later.

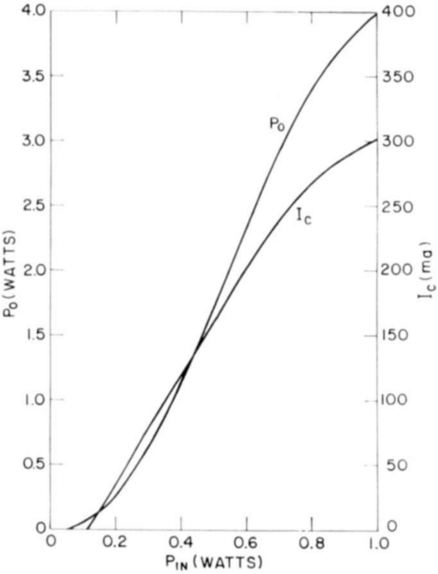


Fig. 8—Power output, P_o, and collector current, I_C, versus power input, P_{in}, for the common-emitter amplifier.

ANALYSIS OF THE IDEAL TRANSISTOR PARAMETRIC MULTIPLIER

Before attempting to correlate the experiments with theory, a simplified analysis of the TPM will be given and its performance compared to that of the class-C multiplier. Since the mechanism of multiplication in the TPM is similar to that of a conventional varactor multiplier, the details of the analysis are omitted, and only the highlights are presented. A description of conventional varactor multiplier theory may be found in the literature.[9]

Table II—Common-Emitter Multiplier Performance
($V_{be} = 0$, $P_{in} = 1$ watt at 500 mc)

Transistor	Circuit	V_{CE} (volts)	P_{om} (watts)	η_c (%)
#1	CE × 1	32	3.0	36
		40	3.8	32
#2	CE × 1	32	4.0	40
#1	CE × 2	32	1.3	18
		40	1.6	17
#2	CE × 2	32	2.3	29

An ideal TPM will be considered. Losses, parasitic reactances, and saturation effects are neglected in the simplified analysis. These effects, as well as the important parameters for selecting a transistor for a TPM, are discussed later. The ideal TPM obeys the Manley–Rowe equations and has 100% conversion efficiency, where the conversion efficiency is defined as the ratio of the multiplied power output to the available fundamental power output after amplification. An ideal class-C multiplier that can generate a delta function current pulse will also have 100% conversion efficiency. The analysis presented below is for a common-base doubler. It also applies to the common-emitter doubler when the proper input circuitry is used. The analysis may be easily extended to higher-order multipliers.

---

[9] P. Penfield and R. P. Rafuse, *Varactor Applications*, MIT Press, Cambridge, Massachusetts, 1962.

Power gain is achieved in the fundamental gain section of the TPM by well-known transistor mechanisms. The common-emitter circuit has both current gain and voltage gain while the common-base circuit has only voltage gain. The frequency multiplication is primarily a function of the collector–base junction and the output circuit.

A modified version of the collector section of a lossless Early[10] equivalent circuit of the common-base amplifier is shown in Figure 9. An ideal current generator drives the variable capacitance section of the TPM. The collector circuits are assumed to have lossless components. Resonance at the fundamental frequency is achieved by an idler circuit consisting of an inductance L_1 in series with a capacitance C_1, and the effective capacitance of the collector-to-base junction. In a similar fashion the series combination of the inductance L_2 and the capacitance C_2 resonates with the effective capacitance of $C_{b'c}$ at the second-harmonic frequency. The effective load to which power is delivered at the second harmonic is represented by the resistance R_2. Filters are included in both the load and idler branches to allow only the second-harmonic and fundamental currents, respectively, to flow in these branches. The idler is not used in the case of the class-C multiplier.

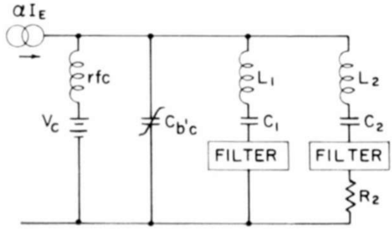


Fig. 9—Equivalent collector circuit for ideal TPM.

The current generator produces a current pulse rich in harmonics that may be represented as

$$\alpha I_E = \sum_{-\infty}^{\infty} I_n \exp\{jn\omega t\}, \tag{1}$$

where ω is the fundamental frequency. The base–collector junction

---

[10] R. L. Pritchard, et al, "Transistor Internal Parameters for Small-Signal Representation," *Proc. IRE*, Vol. 49, p. 736, April 1961.

capacitance is $C_{b'c}$. The junction is assumed to be abrupt and therefore

$$C_{b'c} = C_{\min} \left(\frac{\phi + V_B}{\phi + V} \right)^{1/2}, \tag{2}$$

where ϕ is the contact potential, V_B is the breakdown potential, $C_{\min}$ is the capacitance at breakdown potential, and V is the applied potential.

The voltage V developed across the base–collector junction also can be expanded in a Fourier series:

$$V = \sum_{-\infty}^{\infty} V_n \exp\{jn\omega t\} = \sum_{-\infty}^{\infty} V_{n0} \exp\{j(\theta_n + n\omega t)\} \tag{3}$$

where V_{n0} is the magnitude of the n^{th} harmonic and θ_n is the phase.

When a small-signal circuit analysis (a-c voltage small compared to d-c reverse-bias on junction) neglecting harmonics higher than the second is used, the following equations may be derived relating the Fourier components of Equations (1) and (3):

$$I_1 = -j \frac{Q_o V_2 V_1^*}{R_2 8 V_o}, \tag{4}$$

$$I_2 = \frac{Q_o^2}{R_2} \frac{1 + jQ_o}{1 + Q_o^2} V_2 - j \frac{Q_o V_1^2}{R_2 8 V_o}, \tag{5}$$

where $Q_o = 2\omega C_o R_2$ and $V_o = V_c + \phi$. V_c is the reverse bias applied to the base–collector junction, and V_1 and V_2 are the voltage amplitudes of the fundamental and first-harmonic frequencies, respectively.

C_o is the effective capacitance of $C_{b'c}$ and is given by

$$C_o = C_{\min} \left(\frac{\phi + V_B}{\phi + V_c} \right)^{1/2}. \tag{6}$$

The admittance at the fundamental frequency presented to the current generator section is, from Equation (4),

$$\frac{I_1}{V_1} = -j \frac{Q_o V_2 \exp\{-2j\theta_1\}}{8 R_2 V_o}. \tag{7}$$

This admittance is real if

$$2\theta_1 - \theta_2 = -\frac{\pi}{2}. \tag{8}$$

If we assume that this phase condition can be obtained through proper tuning, the normalized amplitude, X, of the second-harmonic voltage across the capacitor is then determined from the cubic equation

$$X^3 - |a_2|^2 \frac{1+Q_o^2}{Q_o^2} X^2 - 64 |a_1|^4 \frac{1+Q_o^2}{Q_o^2} = 0 \tag{9}$$

where

$$|a_1| = \left| \frac{I_1 R_2}{Q_o V_o} \right|,$$

$$|a_2| = \left| \frac{I_2 R_2}{Q_o V_o} \right|,$$

$$X = \left(\frac{V_{20}}{V_o} \right)^2. \tag{20}$$

The parameters a_1 and a_2 are normalized fundamental and second-harmonic currents produced in the current generator section of the transistor.

A pure varactor solution is obtained by setting $|a_2| = 0$ while a class-C multiplier solution is obtained by setting $|a_1| = 0$.

The harmonic power output is given by

$$P_{o2} = \frac{Q_o^2}{2R_2} \frac{V_{20}^2}{1+Q_o^2}. \tag{10}$$

The variation of X as a function of a_2/a_1, for some typical parameters is plotted in Figure 10. It can be concluded that the voltage developed across the junction is nearly that of a pure varactor. Thus, the ratio of the harmonic power output of the class-C multiplier to that of the TPM is

$$\frac{X_{\text{class-C}}}{X_{\text{TPM}}} \approx \frac{1}{4} \frac{|a_2|^2}{|a_1|^2} |a_1|^{2/3} \left(\frac{1+Q_o^2}{Q_o^2} \right)^{2/3}. \tag{11}$$

For large Q_o ($\gg 1$)

$$\frac{X_{\text{class-C}}}{X_{\text{TPM}}} \rightarrow \frac{1}{4} \frac{|I_2|^2}{|I_1|^2} |I_1|^{2/3} \left(\frac{R_2}{V_o} \right)^{2/3} \frac{1}{Q_o^{2/3}}, \tag{12}$$

while for small Q_o ($\ll 1$)

$$\frac{X_{\text{class-C}}}{X_{\text{TPM}}} \to \frac{1}{4} \frac{|I_2|^2}{|I_1|^2} |I_1|^{2/3} \left(\frac{R_2}{V_o} \right)^{2/3} \frac{1}{Q_o^2}. \tag{13}$$

A TPM operating with a fundamental frequency close to $f_{\max}$ of the transistor will have a second harmonic produced by the pulse generator that is less than the fundamental produced by the generator. Therefore, for large Q_o, the power output of the TPM will be considerably greater than the output of the class-C multiplier. When Q_o is small the second-harmonic output of the class-C multiplier may be close to that of the TPM. Large Q_o implies large C_o and heavy "pumping" in the TPM, but a great deal of shunting of harmonic current in the class-C multiplier. Small Q_o and consequently small C_o implies less pumping and less shunting, so the output of the class-C multiplier may be comparable to that of the TPM.

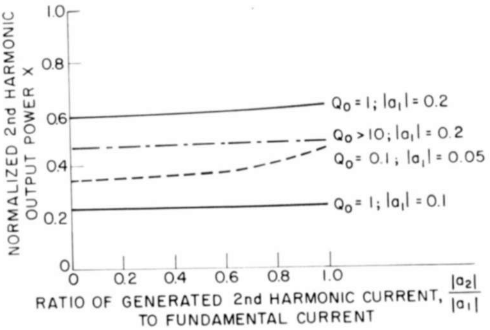


Fig. 10—Normalized power output of TPM doubler as a function of the ratio of generated second harmonic to fundamental currents.

Viewed from another angle, for a fixed C_o and R_2, the ratio of power output of the TPM to that of the class-C multiplier *increases* as the frequency is increased. The optimum fundamental frequency is probably close to f_t. Operating at higher frequencies will result in low fundamental gain and reduced multiplied output power.

The analysis presented above is limited to small-signal conditions and cannot be expected to accurately predict the large-signal results obtained earlier. However, the analysis is useful in providing insight into the nature of the operation of a TPM.

The resistance presented to the pulse generator at the fundamental frequency can be found from Equations (7), (8), and (9), and is

$$R_o(\omega) = 2\left(\frac{V_o}{I_1R_2}\right)^{2/3} \frac{Q_o}{(1+Q_o{}^2)^{1/6}\,\omega C_o}. \tag{14}$$

In summary, the important conclusion is that the ideal TPM can be treated as a transistor driving a load given by Equation (14), and the multiplication process can be treated purely as a varactor process.

Thus far we have made no mention of losses. The primary loss to be considered with a conventional junction transistor used as a TPM is the base spreading resistance $r_{bb'}$, just as it is the primary loss effect in a power amplifier. The effect of the base spreading resistance is to introduce a varactor cutoff frequency for the TPM given by

$$f_c = \frac{1}{2\pi C_{\min}(r_{bb'}+r_s)}. \tag{15}$$

The series resistance, r_s, present in the collector circuit is also included in the denominator of Equation (15). The effect of losses is to reduce the multiplication below the 100% predicted by the Manley–Rowe equations.

A requirement for good multiplication efficiency, using conventional transistors at high frequencies, is therefore that $r_{bb'}$ must be extremely low. Some of the modern high-frequency high-power transistors, however, are constructed in such a way that not only is $r_{bb'}$ extremely small, but also part of $r_{bb'}$ is by-passed and has little influence on the varactor cutoff frequency. This is discussed in the next section.

TRANSISTOR PARAMETRIC MULTIPLIER: EQUIVALENT CIRCUITS, CHARACTERISTICS, AND CAPABILITIES

In order to evaluate the frequency-multiplication capabilities of a particular TPM, it is necessary to find the parameters of the device that determine its varactor performance. A convenient way of doing this is to use an equivalent circuit. A knowledge of the physical parameters that contribute to the equivalent circuit elements make it possible to specify the geometrical and electrical properties of the transistor that determine the varactor characteristics. In this section an equivalent circuit is developed that represents the 2N3375 at the frequencies and conditions of use. The unknown equivalent-circuit elements are found by impedance measurements, and from these the varactor parameters are calculated.

A conventional junction transistor can be represented at high frequencies by either the Early or hybrid-π equivalent circuits. As mentioned in the previous section, the base spreading resistance and collector series resistance degrade the varactor performance. On the other hand, a planar transistor using the overlay structure does not have the same limitations. The reason for this is seen from a consideration of the geometry of the overlay transistor, a cutaway view of which is shown in Figure 11. There is a large base–collector junction area in addition to the active area under the emitter. The equivalent circuits mentioned above are not suitable to represent this transistor. The base–collector junction area not under the emitter results in a capacitance that must be accounted for in the equivalent circuit. The equivalent circuit of Figure 12, a modification of the model for the planar transistor described by Wolf,[11] is a useful starting point to determine the circuit elements pertinent to varactor action in the 2N3375. The elements shown are for zero emitter current. The effective resistance and junction capacitance between the base and collector leads constitute the equivalent varactor section of the transistor. The inner base–collector capacitance, $C_i(V)$, and the outer base–collector capacitance, $C_o(V)$, of Figure 12 are determined by the emitter area and total base area, respectively. These are in the ratio of approximately 1:10 for the 2N3375. Thus $C_o(V) \cong 10C_i(V)$, and similarly the inner resistance, r_i, and outer resistance, r_o, are related by $r_i \cong 10r_o$.

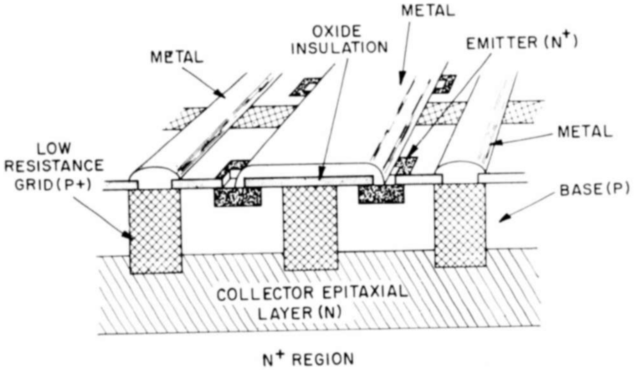


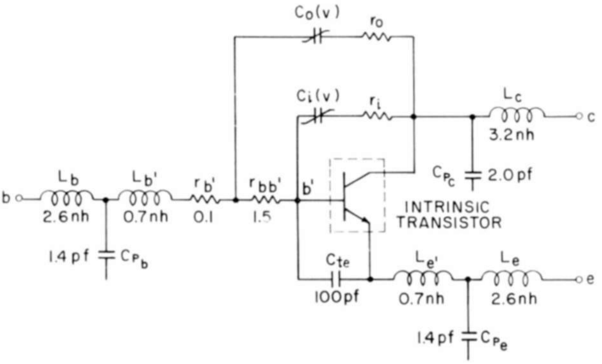
Fig. 11—Cutaway view of 2N3375.

---

[11] H. Wolf, "Recent Advances in High-Frequency Power Transistors," *Solid State Design*, Vol. 4, p. 21, June 1963.

Thus to a first-order approximation, the equivalent varactor elements are r_o and $C_o(V)$. $C_o(V)$ and r_o are fairly independent of emitter current. It is to be noted that although the resistances $r_{bb'}$ and r_i have little influence on the varactor cutoff frequency, they play significant roles in the transistor action at the fundamental frequency.

The values of the other elements of the equivalent circuit are given in Figure 12. The parasitic inductances* $L_{b,e,c}$ and case capacitances $C_{P_{b,e,c}}$ were determined by removing the transistor pellet from the case and making measurements of the input and output short-circuit admittances on an admittance bridge. Measurements were carried out at 500 mc and 1000 mc with the case connected both in common-emitter and common-base configurations. The values of the remaining elements were supplied by the manufacturer or derived from measurements and/or calculations based on the transistor properties.†



$L_{b,e,c}$ = parasitic inductances external to transistor pellet
$L_{b',e'}$ = parasitic inductances internal to transistor pellet
$r_{bb'}$ = base spreading resistance
$C_{P_{b,e,c}}$ = case, package capacitances to ground
C_{te} = intrinsic base-to-emitter capacitance
$C_i(V), r_i$ = inner intrinsic base-collector junction capacitance, series resistance
$C_o(V), r_o$ = outer base-collector junction capacitance, series resistance (external to intrinsic base region).

Fig. 12—Equivalent circuit of 2N3375 transistor ($I_e = 0$).

The pertinent elements of Figure 12 with zero emitter current are

---

* The external parasitic inductances indicated in Figure 12 are those measured with the transistor in the bridge mount. There is evidence that these are somewhat larger than what would be present with the 2N3375 in an actual circuit. However, these are the values that must be accounted for in the calculations based on the measurements described here.

† Some of the estimated values do not alter our calculations appreciably.

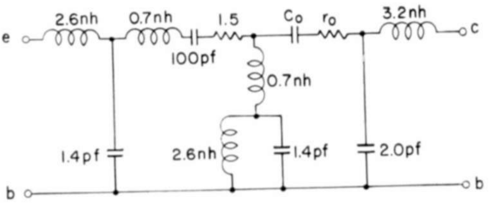


Fig. 13—Simplified common-base equivalent circuit (reduced from Figure 12, $I_e = 0$).

shown in Figure 13 connected in a common-base circuit. In order to determine the values of r_o and $C_o(V)$, measurements were made of Y_{22}, the short-circuit output admittance, with zero emitter current, I_e. These measurements were performed at 500 mc over a range of collector–base bias voltages, V_{bc}. The frequency of 500 mc was chosen in order to evaluate the equivalent-circuit parameters at the fundamental frequency of operation. From the values of Y_{22} and the values of the elements of Figure 13, r_o and C_o were calculated. They are tabulated in Table III as a function of bias voltage. r_o and C_o were also

Table III—r_o and C_o for Open- and Short-Circuit Condition on Two Transistors as Functions of Bias Voltage

Transistor	I				II	
Measurement	h_o		Y_{22}		Y_{22}	
V_{CB} (volts)	r_o (ohms)	C_o (pf)	r_o (ohms)	C_o (pf)	r_o (ohms)	C_o (pf)
0	2.1	31.0	3.1	28.6		
10	1.8	9.1	1.8	8.4	1.5	8.7
20	1.8	6.4	1.6	6.2	1.8	6.7
30	2.0	5.3	2.0	5.2	1.4	5.6
40	1.8	4.5	1.8	4.5	1.8	4.9
50	1.8	4.1	1.8	4.0	2.0	4.5

calculated in a similar fashion from measurements of h_o, the open-circuit output admittance. Some of these results are also shown in Table III. They agree with the values determined from the Y_{22} measurements. In addition, similar measurements made with an open- and short-circuit common-emitter configuration also yields confirming values of r_o and C_o. A check at 1000 mc, although the effects of the parasitic reactances now become very large, also gives good agreement with the previous results. It is therefore concluded that the equivalent circuit of Figure 12, which contains the basic varactor elements, is an adequate representation, and the parameters determined from this circuit representation are reasonably accurate.

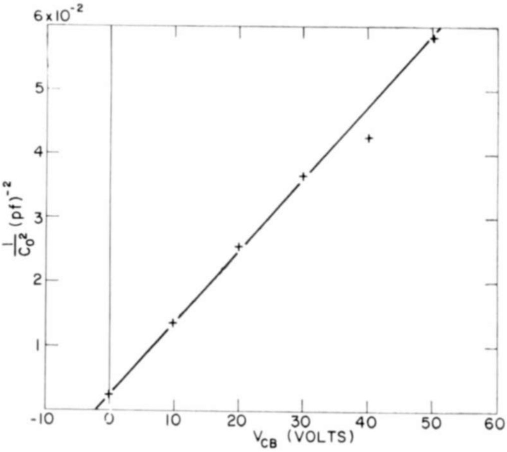


Fig. 14—Plot of $1/C_o^2$ versus V_{CB}.

A plot of $1/C_o^2$ as a function of bias voltage is shown in Figure 14. We conclude that the collector–base junction is abrupt. One can use the fairly constant value of r_o and the capacitance (extrapolated to its minimum value at a breakdown voltage of 65 volts) to find a varactor cutoff frequency, f_c. Calculated values range from 23 to 25 gc.* Measurements of the varactor cutoff frequency using different techniques yield values in the same range.† The values found on one transistor

---

* Earlier values of f_c reported by the authors (see footnote p. 287) were about one half the above. The lower values were in error due to neglect of losses in the admittance bridge.

† Measurement made by A. P. Anderson, RCA Victor Co. Ltd., Montreal, Canada.

by open- and short-circuit measurements were 24.5 and 24.8 gc. Another 2N3375 type transistor with a slightly different construction yielded a value of 23.4 gc from short-circuit measurements. If a breakdown voltage of 75 volts is assumed, a value of f_c as high as 27 gc is found. The average internal series resistance determined from most measurements varied from 1.6 to 1.8 ohms. This resistance presumably would include r_o and any part of $r_{bb'}$ that is present in the a-c path from base to collector leads. The value of r_o is consistent with estimates based on the transistor construction.‡

It is important to point out again that this reasonably high value of f_c for the TPM is due to the transistor construction, whereby the contribution to the varactor parameters is $C_o(V)$ and r_o of Figure 12, rather than $C_i(V)$ and $r_i + r_{bb'}$. This statement applies basically to the planar overlay transistor construction. In a transistor where the major collector–base junction is in the active region of the transistor, the characteristics of the varactor would be determined by $r_i + r_{bb'}$ and $C_i(V)$, rather than by r_o and C_o. The varactor cutoff frequency would then be drastically lowered. Such transistors would not be expected to yield good multiplier performance at microwave frequencies.

It was mentioned earlier that the common-emitter TPM requires a by-pass circuit at the harmonic frequencies between the base lead and ground. Reference to Figure 12 shows that there is a fairly high impedance between the varactor $C_o(V)$ and the emitter lead that cannot be tuned out by a simple matching network in the common-emitter configuration. The harmonic by-pass circuit circumvents this difficulty.

MULTIPLICATION PERFORMANCE AND EFFICIENCY

With the varactor parameters specified, one can compare the performance of the collector–base varactor and transistor amplifier with the theoretical multiplication efficiency and power predicted from the Penfield–Rafuse curves.[9] Table IV contains data selected from Tables I and II that can be used to compare performance. Column 3 is the maximum output power obtained for each condition listed in columns 1 and 2; the input r-f power for all operations listed was 1 watt at 500 mc.

The power output obtained as an amplifier (×1) permits a calculation of the conversion efficiency. This is based on the model that the observed amplifier power, P_{o1}, drives the collector–base capacitance,

---

‡ D. R. Carley, RCA Electronic Components and Devices, Somerville, N. J.

Table IV—Multiplication Performance—Measured and Predicted

1 Transistor Condition	2 Circuit	3 P_{om} (watts)	4 η_v (%)	5 η_{pr}(%) ($f_c = 25$ gc)	6 I_{dc} (ma)	7 η_{cm} (%)	8 η_{cpr} (%)	9 η_{cm}/η_{c1} (%)
#1	CB(×1)	4.8			390	44		
$V_{CB} = 28$ volts	CB(×2)	2.6	54	68	226	41	51	93
	CB(×3)	1.3	27	52	211	22	41	50
#1	CB(×1)	6.7			380	42		
$V_{CB} = 42$ volts	CB(×2)	3.6	54	68	285	30	38	72
#2	CE(×1)	3.0			260	36		
$V_{CE} = 32$ volts	CE(×2)	1.3	43	68	226	18	28	50
#2	CE(×1)	3.8			296	32		
$V_{CE} = 40$ volts	CE(×2)	1.6	42	68	235	17	27	53

$\eta_v = P_{om}/P_{o1} =$ calculated multiplication efficiency

$\eta_{pr} =$ predicted multiplication efficiency from Penfield–Rafuse

$I_{dc} =$ measured collector current

$\eta_{cm} =$ measured collector efficiency

$\eta_{cpr} = (\eta_{c1}/\eta_{pr})(P_{dc1}/P_{dcm}) =$ estimated collector efficiency

$\eta_{cm}/\eta_{c1} =$ separate varactor efficiency

$C_{b'c}$, as a varactor multiplier, producing the harmonic power P_{om}. The calculated multiplication efficiency, $\eta_v = P_{om}/P_{o1}$, may be compared with the maximum efficiency, η_{pr}, predicted by Penfield–Rafuse. These (η_{pr}) have been evaluated, neglecting idler losses, for a varactor frequency cutoff of 25 gc. Columns 4 and 5 list the results. The measured conversion efficiencies are somewhat less than predicted, as might be expected.

The collector efficiency, η_{cm}, at a fundamental ($m = 1$) or harmonic ($m > 1$) is defined as the ratio of the r-f power out, P_{om}, to the d-c power supplied to the circuit, P_{dc};

$$\eta_{cm} = \frac{P_{om}}{P_{dc}}. \tag{16}$$

If we assume that the fundamental circuit losses remain constant with multiplication and that the previous model is still applicable, P_{om} is given by the product of the fundamental driving power, P_{o1}, and the predicted varactor efficiency, η_{pr}. Then

$$\eta_{cm} = \eta_{c1}\eta_{pr}\frac{P_{dc_1}}{P_{dc_m}}. \tag{17}$$

The measured and estimated efficiencies are presented in columns 7 and 8. In general the agreement is good. As before, the measured results are somewhat less than predicted.

It is to be noted that the collector current always decreases under multiplication conditions. This tends to improve the collector efficiency, as can be seen from Equation (17). More experimentation will be required to determine the cause of this current fall. Possibilities include a change in the conduction angle or a change in the shape of the current pulse, either of which may be produced by a change in bias or the presence of harmonic currents in the inactive base. An increase in bias, V_{be}, generally improves the collector efficiency of a class-C amplifier, but this usually requires a higher input drive level to obtain a given output power. In all of the values quoted the input power was maintained at 1 watt. Different tuning conditions for amplifying and multiplying may also have a profound effect on the collector current. The effect of tuning on collector current was observed, and found to be greater for the multiplier than for the amplifier.

The unexpected lowering of the collector current in a multiplier gives relatively high efficiencies for the TPM. However, it must be pointed out that the effects of emitter–base bias have not yet been investigated. Column 9 presents the ratio of the multiplier collector efficiency, η_{cm}, to the amplifier collector efficiency, η_{c1}. This ratio is the efficiency that would be required of an independent varactor driven by the 2N3375 amplifier to obtain the same multiplier power output. The TPM, with its integrated varactor, gives a higher overall efficiency, especially with the common-base circuit, than would be expected from a transistor followed by one of the available high-power varactor diodes.

The determination of the varactor cutoff permits an examination of the predicted optimum operating voltage with actual operation. For varactors operating at a fundamental frequency less than 5% of the cutoff frequency, the operating bias predicted by Penfield–Rafuse for either maximum power or maximum efficiency is between 34 and 36% of the breakdown voltage. For the specified breakdown of approximately 70 volts, the optimum operating voltage for the varactor is around 25 volts. Our best overall efficiencies were obtained at bias voltages close to this value. The highest fundamental power available from the 2N3375 used as a class-C amplifier occurs with a collector bias voltage in the range of 40 to 50 volts. Therefore, it appears the bias requirements for maximum fundamental power and maximum multiplication power differ. This agrees with observations. For some transistors the highest harmonic power was achieved at bias voltages greater than 40 volts. Generally these transistors produced the highest harmonic powers. The maximum multiplication power for others occurred at lower voltages, below that for maximum fundamental power gain, and the harmonic power for these was correspondingly lower. The achievement of high multiplication power in a TPM is probably associated with the use of transistors having a high breakdown voltage. For example, a transistor with a breakdown voltage of 100 volts (as has been measured on some 2N3375's) should be operated with a bias of 35 volts for maximum multiplication efficiency. This is closer to the collector bias for high power gain and would exhibit high harmonic power.[†]

Operation of a Transistor Beyond the Normal $f_{\max}$

TPM operation is one method of extending the usable frequency range of certain transistors beyond $f_{\max}$ (the maximum oscillation frequency range of a uniform base transistor), where $f_{\max}$ is defined by

$$f_{\max} = \sqrt{\frac{f_a}{8\pi r_{bb'} C_{cb}}} \tag{18}$$

[†] Increasing the collector breakdown voltage will not always increase the multiplication efficiency. The collector resistance is usually increased to raise the breakdown voltage and this tends to lower the varactor cutoff frequency.

and f_a is the alpha cutoff frequency.

Figure 15 shows the power gain versus frequency measured on a 2N3375 as a common-emitter amplifier (×1). The measured gain falls off beyond 500 mc faster than 6 db per octave (dotted curve). We can use the predicted Penfield–Rafuse conversion efficiency for a varactor having $f_c = 25$ gc to estimate the multiplied power. Figure 15 also shows the multiplied power versus frequency based on using this transistor as a doubler (×2), tripler (×3), and quadrupler (×4). The transistor's usable frequency limit has effectively been transformed from a $1/f^2$ fall to a $1/f$ fall. The frequency range for which the power gain is greater than unity has been more than doubled.

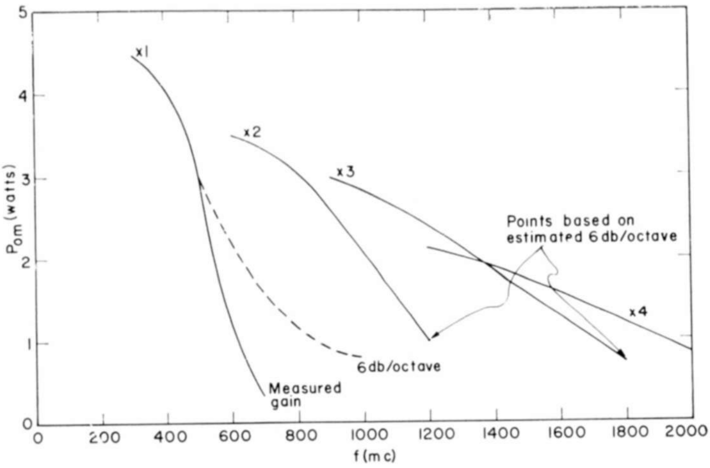


Fig. 15—Power output, P_{om}, as function of frequency for various amplifying and multiplying conditions (common-emitter circuit, power input = 1.0 watt, $V_{CE} = 30$ volts).

For the transistor response shown it appears that doubling is most effective at lower frequencies, while tripling or quadrupling would yield higher powers for higher output frequencies. The powers shown would be over twice as great in a common–base circuit with higher collector–base voltages. The achievement of power at 1000 mc would appear to be best performed by a TPM tripler, which was not done in the 1000 mc powers reported in this paper. In addition, tripling would produce more stable common-base operation as pointed out earlier. It may be concluded that the highest power performance of the 2N3375 at 1000 mc remains to be demonstrated.

Shockley[12] and others[13,14] have discussed another method of obtaining useful power from a transistor at frequencies in excess of the limit defined in Equation (18). This alternative method is to operate in the transit-time mode. In this mode, a negative resistance can be produced between the emitter and collector terminals over certain frequency bands where the base lead is open circuited and the $\text{Arg}(1-\alpha)$ is negative, where α is the usual collector–emitter current ratio. Thus the transit-time mode is represented as a two-terminal device and the r-f current circulates in the loop from emitter to collector to load. The lowest frequency band in which the negative resistance can exist is roughly defined by the limits

$$\frac{1}{2} < f\tau < 1 \tag{19}$$

where τ is the base transit time. Experimental observation of this mode has been reported[14] at micro- and milliwatt power levels using drift transistors. The large excess phase shift of a drift transistor makes it particularly suited to transit-time operation.

The TPM mode in planar transistors differs from the transit-time mode in that the high-frequency current in a TPM must flow through the base lead to obtain efficient, high-power operation. In the transit-time mode, on the other hand, no r-f current flows through the base lead. A practical circuit may, in some isolated cases, generate powers above f_{max} by a combination of the two modes. However, in the circuit configurations described in this paper the transit-time mode would not be strongly excited and it is concluded that the dominant process producing the substantial microwave power observed is due to TPM operation.

CONCLUSIONS

A single transistor has been used as a transistor parametric multiplier (TPM) to produce watts of power at microwave frequencies. Both amplifying and oscillating TPM's have been constructed with

---

[12] W. Shockley, "Negative Resistance Arising from Transit-Time in Semiconductor Diodes," *Bell Sys. Tech. Jour.*, Vol. 33, p. 799, July 1954.

[13] J. F. Gibbons, "An Analysis of the Modes of Operation of a Simple Transistor Oscillator," *Proc. IRE*, Vol. 49, No. 9, p. 1383, Sept. 1961.

[14] R. Zuleeg and V. W. Vodicka, "Microwave Operation of Drift Transistors in Transit-Time Mode," *IRE Trans. on CT*, Vol. CT-8, No. 4, p. 426. Dec. 1961.

both common-base and common-emitter configurations using simple circuitry. Overall efficiencies have been obtained which appear to be as good as or better than the combination of a transistor and separate varactor diode.

An analysis indicates that the ratio of power available from a TPM to that available from an ordinary class-C multiplier improves as the fundamental frequency is increased. The TPM is expected to have its best operation compared to that of a class-C multiplier when operating at a fundamental frequency close to f_t.

A transistor using the planar overlay type construction, such as the 2N3375, is required for successful operation as a TPM. This construction leads to a lower resistance in series with the effective collector–base junction capacitance, and the resulting varactor has a relatively high cutoff frequency. The equivalent circuit presented for the overlay transistor adequately represents the varactor section of the TPM. Using this circuit the varactor parameters are calculated from admittance measurements. Calculations of predicted operating efficiencies based on this equivalent circuit are consistent with measurements. Further work is required to explain the lowering of collector current during multiplier operation, and the resulting enhancement of collector efficiencies. The operation of the 2N3375 as a TPM effectively modifies the $1/f^2$ loss of power in the frequency range beyond f_t to that of a $1/f$ loss.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to J. Hilibrand for calling our attention to the phenomena of power output above the $f_{\max}$ of a transistor and contributing to the course of the work with many helpful discussions. We also wish to thank others from RCA Electronic Components and Devices, Somerville, New Jersey, who participated in illuminating discussions of the problem—in particular, D. R. Carley and R. Minton. Also, the authors are grateful to A. P. Anderson of RCA Victor Co., Ltd., Montreal, Canada, who aided in calling our attention to some of the problems associated with varactor cutoff measurements.

RCA Technical Papers[†]

First Quarter, 1965

Any request for copies of papers listed herein should be addressed to the publication to which credited.

"Color TV—Today and Tomorrow," E. W. Engstrom, *Radio-Electronics* (January) 1965

"Effects of Growth Parameters on Dislocations in CaF_2," M. S. Abrahams and P. G. Herkart, *Jour. Appl. Phys.* (January) 1965

"Electron Mobility Studies in Surface Space-Charge Layers in Vapor-Deposited CdS Films," A. Waxman, V. E. Henrich, F. V. Shallcross, H. Borkan, and P. K. Weimer, *Jour. Appl. Phys.* (January) 1965

"An Experimental Photo-Tape Frame-Camera System," W. J. Poch, *Jour. S.M.P.T.E.* (January) 1965

"High-Temperature Specific Heats of Ge, Si, and Ge-Si Alloys," D. Gerlich, B. Abeles, and R. E. Miller, *Jour. Appl. Phys.* (January) 1965

"Influence of Source Symmetry on Cerenkov Radiation," L. W. Zelby and J. Vollmer, *Jour. Appl. Phys.* (Communications) (January) 1965

"A New Isolator Using a Solid-State Plasma Waveguide," M. Toda, *Trans. IEEE PTGMTT* (Correspondence) January 1965

"Paralleled Microwave Circuits — Phase-Synchronized and Quasi-Statically Coupled," B. Hershenov, *Proc. IEEE* (Correspondence) (January) 1965

"Point-to-Point Wave Propagation Through an Intermediate Layer of Random Anisotropic Irregularities: Phase and Amplitude Correlation Functions," D. A. de Wolf, *Trans. IEEE PTGAP* (January) 1965

"The Promise and Challenge of the Computer," D. Sarnoff, *Signal* (January) 1965

"A Technique for the Generation of Highly Linear FM Pulse Radar Signals," P. Z. Peebles, Jr. and G. H. Stevens, *Trans. IEEE PTGMIL* (January) 1965

"Ambipolar Diffusion of Free Carriers in Insulating CdS Crystals," P. Mark, *Phys. Rev.* (4 January) 1965

"Time- and Position-Dependent Superconductivity," R. H. Parmenter, *Phys. Rev.* (4 January) 1965

"Nonlinear Reactance of Superconducting Films," J. Gittleman, B. Rosenblum, T. E. Seidel, and W. Wicklund, *Phys. Rev.* (18 January) 1965

"An Application of Decision Tables as the Source Language for Automatic Testing," B. H. Scheff, *Trans. IEEE PTGAS* (February) 1965

"Electro-optic TV Communications System," W. J. Hannan, J. Bordogna, T. E. Penn, and T. E. Walsh, *Proc. IEEE* (Correspondence) (February) 1965

"Evaluation of New Photomultipliers for Scintillation Counting," H. R. Krall, *Trans. IEEE PTGNS* (February) 1965

[†] Report all corrections to *RCA Review*, RCA Laboratories, Princeton, New Jersey.

"An Improved Detection Technique for Random-Access Discrete Address (RADA) Communications, R. C. Sommer, *Proc. IEEE* (Correspondence) (February) 1965

"Improved Performance of $GaAs_{1-x}P_x$ Laser Diodes," J. J. Tietjen and S. A. Ochs, *Proc. IEEE* (Correspondence) (February) .. 1965

"The Insulated Gate Tunnel Junction Triode," S. R. Hofstein and Coauthor, *Trans. IEEE PTGED* (February) 1965

"Measurements of Interaction in an Assembly of γ-Iron Oxide Particles," E. Della Torre, *Jour. Appl. Phys.* (February) 1965

"Precision Measurement of the Microwave Dielectric Constant of Liquids," E. Fatuzzo and P. R. Mason, *Jour. Appl. Phys.* (February) 1965

"Selecting a UHF Antenna," H. E. Gihring, *Broadcast News* (February) 1965

"State-of-the-Art Report on Electrical Energy Sources," P. Rappaport, *Electronic Industries* (February) 1965

"Television Camera System for the Command Module of the Apollo Spacecraft," J. L. Lowrance and P. M. Zucchino, *Jour. S.M.P.T.E.* (February) 1965

"A Visual Instrumentation System for a Lunar Orbiter," F. J. Bingley, *Jour. S.M.P.T.E.* (February) 1965

"Effect of Collisions on the Magnetoconductivity Tensor of a Quantum Plasma," S. Tosima, J. J. Quinn, and M. A. Lampert, *Phys. Rev.* (1 February) 1965

"Nearest-Neighbor Splitting of the Luminescence Levels of ZnS_xSe_{1-x}," W. H. Fonger, *Phys. Rev.* (1 February) 1965

"Theory of Giant Quantum Oscillations in Ultrasonic Attenuation in a Longitudinal Magnetic Field," J. J. Quinn, *Phys. Rev.* (1 February) 1965

"Achieving Long Focal Length Television Lenses Electronically," F. Himelfarb, *Jour. S.M.P.T.E.* (March) 1965

"Carrier Mobility and Current Saturation in the MOS Transistor," S. R. Hofstein and Coauthor, *Trans. IEEE PTGED* (March) 1965

"The Case for NTSC in Europe," G. H. Brown, *Radio-Electronics* (March) 1965

"Cavity Sensing of Cryoelectric Memory Planes," L. L. Burns and A. R. Sass, *Jour. Appl. Phys.*, Part 2 (March) 1965

"Continuous Room-Temperature $Nd^{3+}: CaMoO_4$ Laser," R. C. Duncan, *Jour. Appl. Phys.*, Part 1 (Communications) (March) .. 1965

"Cooperative Jahn-Teller Distortions in the Systems $ZnMn_2O_4$-Cu_2GeO_4 and Zn-Mn_2O_4-$ZnCuGeO_4$," M. Robbins and P. K. Baltzer, *Jour. Appl. Phys.*, Part 2 (March) 1965

"Current-Pumped Abrupt-Junction Varactor Power Frequency Converters," B. S. Perlman, *Trans. IEEE PTGMTT* (March) .. 1965

"The Defect Structure of Grown Silicon Dioxide Films," A. G. Revesz, *Trans. IEEE PTGED* (March) 1965

"The 'Del' Operator in Curvilinear Coordinates," R. B. Schilling, *Amer. Jour. Phys.* (Notes and Discussion) (March) 1965

"Efficient High Power Mobile Transmitter," F. A. Barton, *Trans. IEEE PTGVC* (March) 1965

"Hydrogen-Induced Surface Space-Charge Regions in Oxide-Protected Silicon," J. Olmstead, J. Scott, and P. Kuznetzoff, *Trans. IEEE PTGED* (March) 1965

"Improved Maser Performance through Pump Modulation," R. D. Ray, *Proc. I.E.E.E.* (Correspondence) (March) 1965

"Medicine and Biology Challenge to the Engineering Profession," V. K. Zworykin, *Proc. I.E.E.E.* (March) 1965

"Mobile Network of the Singapore Component of the Royal Malaysia Police," G. Sartore and Coauthor, *Trans. IEEE PTGVC* (March) 1965

"Multichannel VHF for Marine Communications," R. T. Brankley, *Trans. IEEE PTGVC* (March) 1965

"Progress in Optical Computer Research," W. F. Kosonocky and Coauthor, *IEEE Spectrum* (March) 1965

"The Sign of the Trigonal Field Splitting of β-Site Ions in Spinel," S. B. Berger, *Jour. Appl. Phys.*, Part 2 (March) 1965

"Simplified Calculation of Transient Response," M. S. Corrington, *Proc. IEEE* (March) 1965

"The Spectrotor, a Selective Infrared Detector of High Sensitivity," H. Pullan and R. Swarbrick, *Rev. Sci. Instr.* (March) 1965

"A TACAN-Compatible Accurate DME for Short Ranges," C. J. Hirsch, *Trans. IEEE PTGANE* (March) 1965

"Temperature Dependence of n-Type MOS Transistors," F. P. Heiman and H. S. Miiller, *Trans. IEEE PTGED* (March) .. 1965

"Use of Transient Measurements in Metal-Insulator-Semiconductor Structures to Determine Semiconductor Doping Densities," A. Waxman, *Proc. IEEE* (Correspondence) (March) 1965

"Optical Properties and Band Structure of Wurtzite-Type Crystals and Rutile," M. Cardona and G. Harbeke, *Phys. Rev.* (1 March) 1965

"Energy Levels of Dy^{2+} in the Cubic Hosts of CaF$_2$, SrF$_2$, and BaF$_2$," Z. J. Kiss, *Phys. Rev.* (15 March) 1965

"Zeeman Effect Studies of the $^5I_7$-$^5I_8$ Transition in CaF$_2$:Dy^{2+}, Z. J. Kiss, C. H. Anderson, and R. Orbach, *Phys. Rev.* (15 March) 1965

"Advances in the Facsimile Art During 1964," W. H. Bliss, *IEEE International Convention Record, Part 1, Communications 1; Wire and Data Communication* 1965

"The AN/TRC-97—New Tactical Multichannel Microwave Communication System," W. J. Connor and E. J. Sass, *IEEE International Convention Record, Part 1, Communications 1; Wire and Data Communication* 1965

"Pictorial Transmission with HIDM," M. R. Winkler, *IEEE International Convention Record, Part 1, Communications 1; Wire and Data Communication* 1965

AUTHORS

WILLIAM F. BRUCE received a B.S. degree in Physics from Bucknell University in 1961. While at Bucknell, he was employed as an assistant in the electrical measurements laboratory. He has completed two years of graduate study at Purdue University toward the MS degree in Physics. During the summer of 1960, he was employed by the Lansdale Division of the Philco Corp. as a student engineer, and was concerned with the design of transistor manufacturing equipment. During the summer of 1962, he was employed by the Physical Research Group of RCA Astro-Electronics Division, where he made tests on high resolution vidicons. He rejoined RCA in June 1963, and has been concerned with measurements on television pickup tubes, including isocon low-energy electron scattering, storage vidicons, and high-resolution scanning. Mr. Bruce is a member of Sigma Pi Sigma.

MARTIN CAULTON received the Bachelor's degree in 1950, the M.S. in 1952, and the Ph.D. degree in 1954, all in Physics, from Rensselaer Polytechnic Institute. He was an instructor in Physics from 1950 to 1953 while at Rensselaer, and did part of his doctoral work at the Brookhaven National Laboratories, in high-energy nuclear physics, where he was research associate from 1953 to 1954. From 1954 to 1955 he was a Fullbright scholar at the Imperial College of Science and Technology in London. From 1955 to 1958 he was a member of the Technical Staff at Bell Telephone Laboratories working in research and development on low-noise microwave tubes. In 1958 he became assistant professor of physics at Union College, Schenectady, New York. Since joining RCA Laboratories in 1960, he has been engaged in work on microwave power tubes, multivelocity flow problems in electron beams, and microwave solid-state devices. Dr. Caulton is co-author of a textbook on physical electronics and is also adjunct professor of electrical engineering at Drexel Institute of Technology.

A. Danforth Cope graduated from Colgate University in 1938 and continued in graduate study at Yale University (Physics). He joined the Radio Corporation of America in 1941 and worked at Harrison, N. J. and Lancaster, Pa. Since 1949, he has been at the David Sarnoff Research Center, Princeton, N. J. During most of this period his area of interest has been phototube and camera tube design and performance problems. From 1949 to 1961, Mr. Cope worked on photoconductor materials suitable for application to camera tubes. At present he is associated with the RCA Astro-Electronics Division, Physical Research Laboratory, carrying out research in adapting scanning photosensors to the specific needs of space instruments. This has included such diverse problems as soft x-ray solar imaging, long integration television systems applicable to astronomy, and meteorological satellite sensors.

Mr. Cope is a member of Phi Beta Kappa, Sigma Xi, and the Optical Society of America.

Robert L. Ernst received the BEE degree from Manhattan College in 1961 and is now working toward the MSEE degree from the Polytechnic Institute of Brooklyn. In 1961, Mr. Ernst joined the Radio Systems Division of the Western Union Telegraph Company as a microwave engineer and was active in design and development of microwave components. In 1963 he joined the RCA Communications Systems Division in New York, N. Y. He has been active in a program of developing interference reduction techniques for microwave receivers. He has also been involved in the study of electronically tunable microwave bandpass filters utilizing YIG crystals and varactor diodes as well as the study of ferroelectric capacitors for low distortion microwave receiver applications. Mr. Ernst is a member of the Institute of Electrical and Electronic Engineers.

Thomas J. Faith, Jr., received a Bachelor of Engineering degree in 1959, and an M.S. in Physics in 1961, both from Stevens Institute of Technology. He was employed by the Stevens Physics Department as a Research Assistant from 1959 to 1961, where he worked in the fields of high energy physics and low temperature physics. Since June 1961, when he joined the technical staff of the RCA Astro-Electronics Division, Princeton, New Jersey, he has been engaged in experimental work on plasma accelerators. He is a member of the AIAA.

Hans W. Hendel received the B.S. degree in Physics in 1946, the M.A. in 1948, and the Ph.D. (Dr. rer nat.) in 1953 from the Technische Hochschule in Munich, Germany. During 1951, Dr. Hendel held a fellowship at M.I.T., Cambridge, and from 1953 to 1957, he developed optical systems at the Agfa Camera Works in Munich, Germany. From 1957 to 1961, as a research physicist at the USA SRDL, Fort Monmouth, New Jersey, he worked on plasma heating and confinement. As Leader of the Plasma Physics and Propulsion Group at RCA Astro-Electronics Division, Dr. Hendel is doing research on mechanisms applicable to space propulsion. During 1964, Dr. Hendel was in charge of Air Force sponsored work on electron cyclotron resonance plasma acceleration. Since 1963, his group is performing, among other projects, work on pumping methods for high power gas lasers and on tuned laser spectroscopy. Dr. Hendel is a member of the American Physical Society and the AIAA.

E. C. Hutter studied physics at the University of Virginia, receiving the B.S. (Cum Laude) in 1939, the M.A. in 1941, and the Ph.D. in 1943. From 1943 to 1947, he was a research physicist at RCA Laboratories, Princeton, New Jersey. In 1947, he became a research associate at the Institute of Textile Technology at Charlottesville, Virginia. In 1949, he joined the Acetate Research Laboratories of the E. I. duPont de Nemours Company. He returned to the RCA Laboratories in 1951 to work on Project Typhoon and on other general flight simulation projects. Dr. Hutter was placed in charge of the studies relating to environmental problems and power supply for the ARS (satellite) project and continued in this capacity for the succeeding Project JANUS, the forerunner of the TIRCS project. Upon the formation of the Astro-Electronics Division in 1958, he was appointed Manager of Physical Research. Among other projects, his group is performing research on plasma acceleration and on dielectric tape development for satellite TV camera applications. Dr. Hutter is a member of the Institute of Electrical and Electronics Engineers, the American Physical Society, the American Institute of Aeronautics and Astronautics, and Sigma Xi.

E. O. Johnson received the BSEE from Pratt Institute in Brooklyn in 1948 and pursued graduate studies in physics and electrical engineering at Princeton University and at the Swiss Federal Institute of Technology in Zurich. He worked in the fields of plasma, solid-state-surface and general-device physics at RCA Laboratories, Princeton, from 1950 to 1959. Currently, he is the Manager of Engineering for the RCA Electronic Components and Devices operation based at Harrison, New Jersey. His responsibility covers vacuum tubes and many different types of solid state devices.

S. W. Kessler, Jr., received the B.S. degree in industrial engineering from the University of Pittsburgh in 1949. While engaged in graduate work at Johns Hopkins University from 1949 to 1953, Mr. Kessler worked as a Research Assistant, investigating the kinetics of the solidification of metals. He joined RCA in 1953 and was assigned to the Chemical and Physical Laboratory at Lancaster, Pennsylvania. He has worked in various fields including silicon rectifier design and thermionic energy conversion. Recently, he has been responsible for the development of special metal vapor arc lamps for use in industrial test equipment. Mr. Kessler is a member of American Society for Metals and American Institute of Metallurgical Engineers.

Otto H. Schade was born and educated in Germany; he came to the United States in 1926. He joined RCA Electronic Components and Devices, Harrison, New Jersey, in 1931. Since 1938, he has specialized in television circuits, camera tubes, and picture tubes. From 1944 to 1957 he worked on a unified general method of image analysis and specification, including practical methods for measuring the "aperature" effect (square-wave and sine-wave response function) and fluctuation levels (noise) of optical, photographic, and electronic image-system components and the eye. He has had the responsibility for the thermal and electrical design of nuvistor tubes. More recently, he has developed an accurate method for calculating the resolving power of television and photographic imaging systems to assist in the evaluation of high-definition television systems, and a new electron optic providing minimum aberrations and uniform focus in television camera tubes with larger (50 × 50 mm) image surfaces. Dr. Schade has received numerous honors, including the Modern Pioneers Award of the National Association of Manufacturers (1940), the Morris Liebmann Memorial Prize of the Institute of Radio Engineers (1950) and a Fellowship (1951) from the Institute of Radio Engineers. In 1951 he was made a Fellow, and also was the first recipient of the David Sarnoff Gold Medal Award, of the Society of Motion Picture and Television Engineers. In June 1953, he was invested with the honorary degree of Doctor of Engineering by Rensselaer Polytechnic Institute. In 1960, he received the Progress Medal Award of the Society of Motion Picture and Television Engineers for his outstanding technical contribution in the engineering phases of the motion picture and television industries.

HAROLD SOBOL received his BSEE from CCNY in 1952, his MSEE and Ph.D. in E.E. from the University of Michigan in 1958 and 1959, respectively. From 1952 to 1955 Dr. Sobol worked on radar and missile guidance problems at the Willow Run Laboratories of the University of Michigan. He had a Sperry Fellowship, 1955-1956, for graduate study in Electron Physics. From 1956 to 1960 he worked on microwave tube problems, including high-power traveling-wave tubes, in the Electronic Physics Laboratory of the University of Michigan. In 1960, he joined the IBM Research Center at Yorktown Heights, N. Y., where he worked on circuit and thermal problems in high-speed superconducting computer devices. In May 1962, he joined the Microwave Research Laboratory of RCA Laboratories at Princeton, New Jersey, to work on high-power tubes. In August 1963, he became head of the Microwave-Power Generation Group. Recently he has worked on microwave solid-state devices and integrated circuits.

Dr. Sobol is a member of the Institute of Electrical and Electronics Engineers, Sigma Xi, Eta Kappa Nu, and Tau Beta Pi.

PETER S. TORRIONE graduated from CCNY in January 1963 with a BEE degree. Following his graduation, he joined the RCA Communications Systems Division in New York City as a specialized trainee, and subsequently joined the microwave group. He has performed theoretical analyses and experimental measurements on X-band resistive frequency converters and has done work on L-band vacuum-tube VCOs for application in altimeters. In addition, he has done development work on broadband transistor amplifiers in the VHF and UHF bands, miniature filters, and the design of bandpass and band-reject filters, in waveguides. He has also worked on high-order, high-efficiency, step-recovery frequency multipliers. Mr. Torrione is a member of the IEEE and the PTGMTT.

S. YUAN received the BSEE from the University of California in 1952 and the MSEE from Columbia University in 1956. From 1953-1955 he was with the RCA Communications Systems Division in New York City, working on design and development of television broadcasting components. From 1957-1958 he was employed as a Senior Engineer with the Electronics Division of Curtiss-Wright Corp., where he was engaged in their system development program. From 1958-1963 he was associated with the Electronics Research Laboratory of Columbia University as a Senior Research Engineer. During that period he made some basic contributions to the theory and design of highly-efficient frequency multiplier employing variable reactances. In 1963 he rejoined RCA as a Senior Member Technical Staff in the Solid-State RF Techniques Group. He is currently involved in the interference reduction and frequency multiplier programs. Mr. Yuan is a member of Tau Beta Pi and Eta Kappa Mu.

RCA REVIEW

BOARD OF EDITORS

Chairman
R. S. HOLMES
RCA Laboratories

E. I. ANDERSON
Home Instruments Division

A. A. BARCO
RCA Laboratories

E. D. BECKEN
RCA Communications, Inc.

G. H. BROWN
Radio Corporation of America

A. L. CONRAD
RCA Service Company

E. W. ENGSTROM
Radio Corporation of America

A. N. GOLDSMITH
Honorary Vice President, RCA

J. HILLIER
RCA Laboratories

E. C. HUGHES
Electronic Components and Devices

E. O. JOHNSON
Electronic Components and Devices

E. A. LAPORT
Radio Corporation of America

H. W. LEVERENZ
RCA Laboratories

G. F. MAEDEL
RCA Institutes, Inc.

W. C. MORRISON
Broadcast and Communications Products Division

L. S. NERGAARD
RCA Laboratories

H. F. OLSON
RCA Laboratories

J. A. RAJCHMAN
RCA Laboratories

D. F. SCHMIT
Radio Corporation of America

L. A. SHOTLIFF
RCA International Division

C. P. SMITH
RCA Laboratories

W. M. WEBSTER
RCA Laboratories

Secretary
C. C. FOSTER
RCA Laboratories

REPUBLICATION AND TRANSLATION

Original papers published herein may be referenced or abstracted without further authorization provided proper notation concerning authors and source is included. All rights of republication, including translation into foreign languages, are reserved by RCA Review. Requests for republication and translation privileges should be addressed to *The Manager*.

PUSH-PULL SATURATED CORE TUNNEL DIODE INVERTERS*

By

R. Feryszka and P. Gardner

RCA Electronic Components and Devices
Somerville, N. J.

Summary—*Previous work on tunnel diode inverters for use with direct energy conversion sources such as thermionic generators indicated that the push–pull saturated square-hysteresis-loop (Marzolf) tunnel-diode inverter held the most promise for achieving high efficiency and power output with good power-to-weight ratio.*

A theory of the operation of this type of inverter has been developed. This theory predicts the performance of the inverter for varying input voltage and load from the I-V *characteristics of the tunnel diodes and transformer parameters, and shows that the diode figure of merit previously believed to be the limiting efficiency of such inverters can be exceeded by suitable design of the inverter.*

Second-order effects such as nonperfect squareness of the core-material hysteresis loop are also discussed, and their effect on inverter performance indicated.

Performance parameters of this type of inverter with outputs to 220 watts are presented, and unique operation features are discussed. Design and fabrication techniques are also briefly discussed.

Introduction

TUNNEL-DIODE inverter circuits and their application to low-voltage power sources have been explored by H. F. Storm and D. P. Shattuck,[1] F. G. Gevert, S. Wang, and R. U. Broze,[2] J. M.

---

* The work presented in this paper was sponsored by the United States Air Force, Air Force Aero Propulsion Laboratory, Aerospace Power Division, Wright Patterson Air Force Base, and was directed by H. Briggs, Captain USAF, project officer.

[1] H. F. Storm and D. P. Shattuck, "Tunnel Diode D-C Power Converter," *Trans. A.I.E.E.*, Vol. 80, p. 347, July 1961.

[2] F. G. Gevert, S. Wang, and R. U. Broze, "Heat to A.C. Conversion Using Thermoelectric Generators and Tunnel Diodes," *Solid-State Electronics*, Vol. 3, p. 100, Sept. 1961.

Marzolf,[3,4] A. C. Scott,[5] D. J. Hanrahan,[6] and Carlson and Gardner.[7] Although these authors emphasized the feasibility of tunnel-diode inverters for such power sources as thermoelectric generators, thermionic generators, sea-water batteries, and fuel cells, one common conclusion was reached—the maximum obtainable efficiency of a tunnel-diode inverter is limited by the figure of merit of the diode, η_0 (diode conversion efficiency);

$$\eta_0 = \frac{I_p - I_v}{I_p + I_v} \frac{V_v - V_p}{V_v + V_p}, \tag{1}$$

where I_p = peak current, I_v = valley current, V_p = peak voltage, and V_v = valley voltage.

Although the maximum efficiency predicted by Hanrahan was higher for a nonsaturated core than for a Marzolf inverter, the latter was selected for this study for the following reasons:

(a) for high power levels the difference in efficiency, which is $I_m/\Delta I$, where I_m is the magnetizing current, is in the range of 1% or less and is insignificant,

(b) core weight and hence core losses are much lower for a saturated core,

(c) the output of a Marzolf inverter is a high-quality a-c square wave at any desired voltage and requires much less filtering if the output must be converted to d-c.

(d) the Marzolf inverter offers a relatively constant frequency a-c output, while a nonsaturated core changes the frequency in a range of approximately 3 to 1.

Detailed studies were performed on saturated-core push–pull inverters to determine their mode of operation and to optimize their design, particularly at high powers. A detailed theory of the operation of these inverters has been developed, including second-order effects such as core material squareness factor and operating frequency, and equa-

---

[3] J. M. Marzolf, "Tunnel Diode Static Inverter," U.S.N.R.L. Report No. 5706, Oct. 25, 1961.

[4] J. M. Marzolf, "Sea Water Battery Plus Tunnel Diode Converter as a Power Source," U.S.N.R.L. Report No. 5961, May 1963.

[5] A. C. Scott, "Design and Oscillation Application of High Current Tunnel Diodes," Doctorial Thesis MIT, Aug. 1961.

[6] D. J. Hanrahan, "Analysis of Tunnel-Diode Converter Performance," U.S.N.R.L. Report No. 5722, Dec. 1, 1961.

[7] F. M. Carlson and P. D. Gardner, "Design and Application of High Current Tunnel Diodes to DC-AC Inverters," *Proc. Power Sources Conf.*, 1963.

tions have been derived that permit the design and predict the performance of such inverters.

ANALYSIS OF OPERATION

In the analysis of the operation of saturated-core push–pull (Marzolf) tunnel-diode inverters, the following assumptions were made:

(1) The total d-c resistance of the primary circuit, which includes the power-source internal resistance, transformer primary resistance, and diode series resistance, must be less than the negative resistance of the tunnel diode, i.e., the tunnel diode must be biased with a monostable d-c load line in its negative-resistance region.

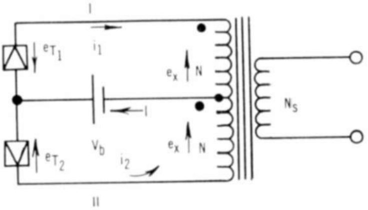


Fig. 1—Push-pull saturated-core tunnel-diode inverter circuit for no-load conditions.

(2) The a-c conditions necessary for the tunnel diodes to operate as relaxation oscillators must be met, i.e., the product of the negative resistance, circuit resistance, and diode capacitance must be less than the circuit inductance. This criteria is very easy to meet.

(3) The transformer core material has a perfectly rectangular hysteresis loop.

(4) All losses, including core losses and lead losses are neglected.

(5) When the effects of loading the inverter are considered, the load is purely resistive.

Open-Circuit Condition

Figure 1 shows the circuit of a push-pull saturated core inverter. It consists of a transformer that has a core of square hysteresis loop material, with two tunnel diodes connected in series across the transformer primary. The power supply is connected between the primary center tap and the common ends of the diodes.

When the circuit is energized, a d-c current equal to the valley current, I_v, of the diodes flows in both primary loops. Since the windings are in the same direction and the currents flow in opposite directions, no magnetizing occurs.

Assuming the current in loop I starts increasing at time t, with the core still unsaturated, a voltage equilibrium exists;

$$V_b - e_{T_1} - e_x = 0. \qquad (2)$$

When the current reaches the value $i = I_v + I_m$, the core is saturated, and $e_x = N(d\phi/dt) \times 10^{-8} = 0$, since no further change in flux is possible. Since e_x is now zero, the tunnel-diode voltage must increase rapidly until the peak point is reached and the diode switches to the forward voltage. Since the core is still saturated, e_{T_1} and i drop rapidly until $i_1 = I_v + I_m$. At that point the core begins to unsaturate and a voltage e_x appears across the primary. The equilibrium now becomes

$$V_b + e_x - e_{T_1} = 0. \qquad (3)$$

The current continues to decrease slowly until the valley point is reached. At this point e_x is again zero, and the diode switches to the low-voltage state.

It has been assumed that during a half cycle, e_x is constant. If Equations (2) and (3) are to hold, then e_{T_1} and e_{T_2} must also be constant. These assumptions are only valid when I_m is small compared to I_p. In practice, e_{T_2} varies over a wide range for small changes in current, leading to distorted waveshapes.

The region where the diode's operating point changes slowly between I_v and $I_v + I_m$ is defined as the "dwell" region. The switching cycle described above is illustrated in Figure 2.

In a push–pull circuit, both diodes operate simultaneously, i.e., when the first diode is in a low-voltage state (at I_v) with i_1 increasing, the second diode is at a high-voltage state (at $I_v + I_m$) with i_2 decreasing.

The time relationships for i_1, i_2, ϕ_1, ϕ_2, and I are shown in Figure 3. The currents, i_1 and i_2, are shifted by $T/2$ (Figure 3a) where T is the period of oscillation. The total current supplied by the power supply is constant and equal to $I_m + 2I_v$ (Figure 3b). The fluxes ϕ_1 and ϕ_2 are in phase with their appropriate magnetizing currents i_1 and i_2 (Figure 3c).

Although the fluxes are out of phase by $T/2$, the induced voltages, e_x, in both primary halves are in phase, since $d\phi/dt$ is positive for the first loop and negative for the second loop (Figure 3d).

It should be noted that this analysis is valid for only one specific

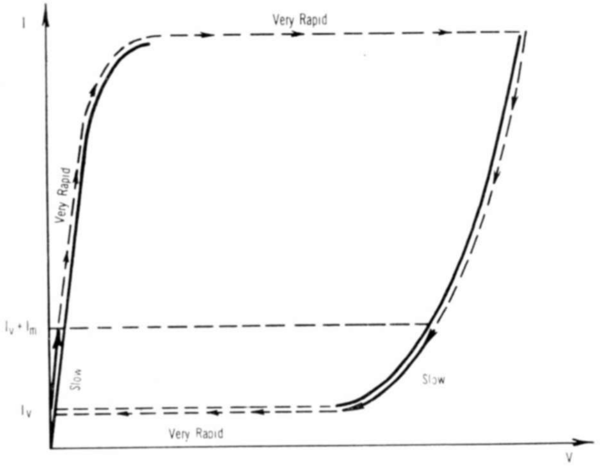


Fig. 2—Tunnel-diode characteristic showing switching path and dwell regions for unloaded condition.

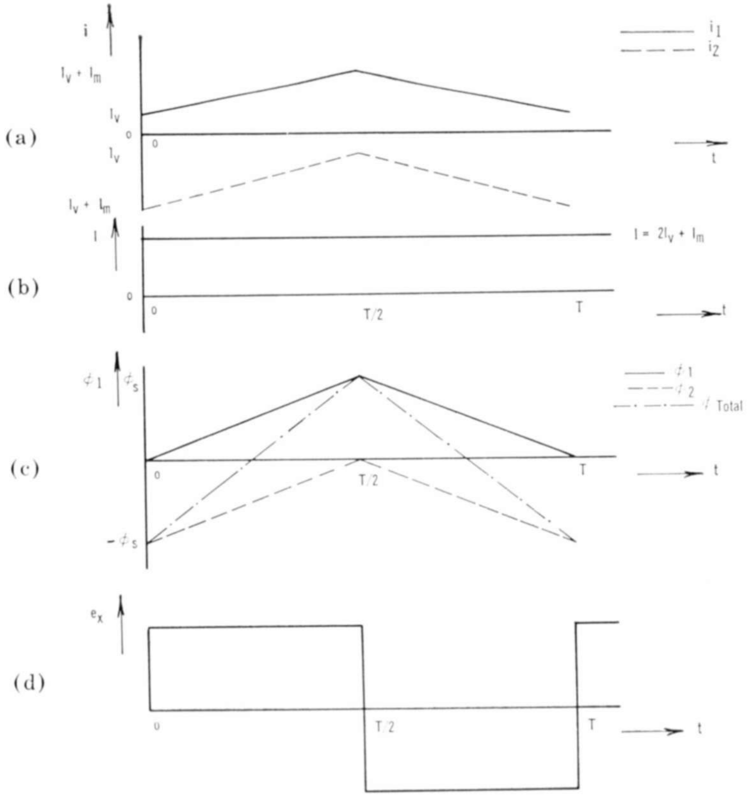


Fig. 3—Time relations for i_1, i_2, I, ϕ_1, ϕ_2, ϕ_{total}, and e_x under no-load conditions.

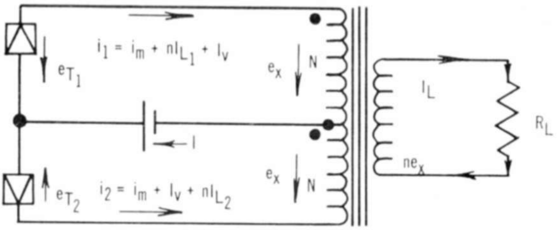


Fig. 4—Push-pull saturated-core tunnel-diode inverter circuit for loaded condition.

bias voltage, V_b, when the diodes dwell between I_v and $I_v + I_m$ (for open circuit), and the input current drawn from the power supply is equal to $2I_v + I_m$. This condition is defined as threshold operation.

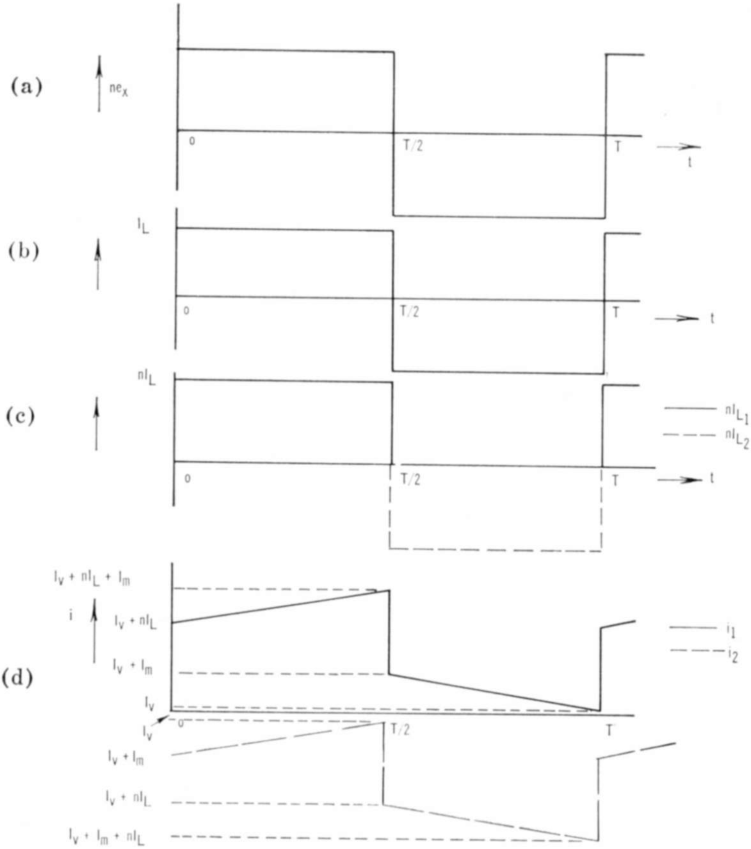


Fig. 5—Time relations between e_x, I_L, i_1, and i_2 for loaded conditions.

Loaded Condition

Assume that the inverter is supplying a current nI_L to a load. Figure 4 shows the directions of current flow and the voltages in the circuit.

Since a purely resistive load is being considered, the load current nI_L is in phase with the existing voltage, e_x, as shown in Figures 5a and 5b. Thus, the load current is carried in loop I during the first half cycle, as shown in Figure 5c. The total current is $nI_L + I_m + I_v$ in loop I and $I_m + I_v$ in loop II during the first half cycle, and $nI_L + I_m + I_v$ in loop II and $I_m + I_v$ in loop I during the second half cycle, as shown in Figure 5d.

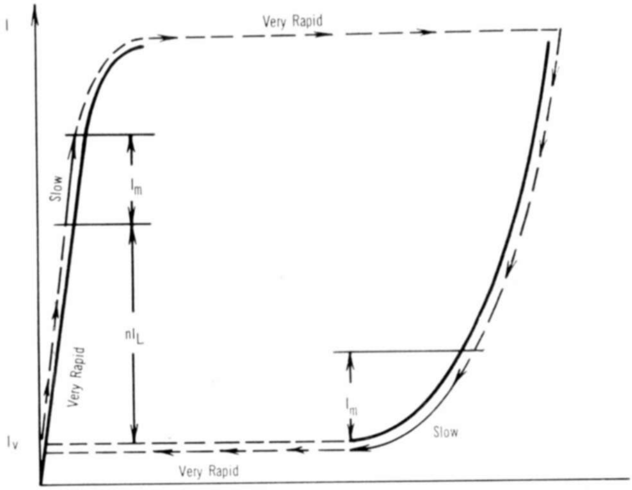


Fig. 6—Tunnel-diode characteristic showing switching path and dwell regions for loaded condition.

The switching path around the tunnel-diode characteristic is shown in Figure 6, and corresponds to the current waveforms shown in Figure 5d. At $t = 0$, $i_1 = I_v + nI_L$. The current i_1 then increases slowly until $t = T/2$, when $i_1 = nI_L + I_m + I_v$. The core is now saturated, and the diode switches very rapidly to a current $i_1 = I_v + I_m$ in the forward region. During the second half cycle, i_1 decreases slowly to $i_1 = I_v$ at $t = T$. For the second diode, at $t = 0$, $i_2 = I_m + I_v$ and decreases to $i_2 = I_v$ at $t = T/2$. At this point, the second diode must provide the load current, and thus it switches to $i_2 = I_v + nI_L$. The current i_2 increases slowly to $i_2 = I_v + nI_L + I_m$, and the cycle is repeated.

Thus, during a cycle, the d-c current I supplied by the power supply is $i_1 + i_2$.

$$I = nI_L + 2I_v + I_m. \tag{4}$$

Thus, for a given load, the current drawn by the inverter is constant.

It can be seen from Figure 6 that the maximum load current obtainable while still maintaining core saturation is

$$nI_{L(\max)} = I_p - I_v - I_m, \tag{5}$$

and the current drawn from the power supply is

$$\begin{aligned} I_{\max} &= I_p - I_v - I_m + 2I_v + I_m \\ &= I_p + I_v. \end{aligned} \tag{6}$$

Again, this analysis holds only for a particular bias voltage.

Having described the basic mode of operation of the inverter under no-load and loaded conditions, a more general analysis of the operation of this type of inverter can now be carried out.

ANALYSIS OF PUSH–PULL SATURATED CORE INVERTER OPERATION

General Case with Inverter Loaded

a. Bias Voltage

It has been assumed that the bias voltage is such that the dwell and equilibria conditions previously described hold. However, in practice the input voltage will vary, and thus it is important to consider the effect of varying bias voltage.

During half a cycle, the following equilibria must hold

$$\text{1st diode:} \qquad V_b - e_{T_1} - e_x = 0,$$

$$\text{2nd diode:} \qquad V_b - e_{T_2} + e_x = 0.$$

We make the assumption that I_m is small and that e_x is constant. Therefore,

$$V_b = \frac{1}{2}\,(e_{T_2} + e_{T_1}). \tag{7}$$

It can be seen from Equation (7) that for a given input voltage, V_b, dwell points must be established to satisfy the equation, and that there is a minimum bias voltage for operation. As $e_{T_1(\min)} \cong 0$ and $e_{T_2(\min)} = V_v$, the minimum input voltage possible is $V_{v/2}$.

However, the minimum dwell points are, for the loaded inverter,

$$\text{1st diode:} \quad nI_L + I_v \rightarrow nI_L + I_v + I_m, \tag{8}$$

$$\text{2nd diode:} \quad I_v + I_m \rightarrow I_v, \tag{9}$$

and

$$I_{IN} = nI_L + 2I_v + I_m. \tag{10}$$

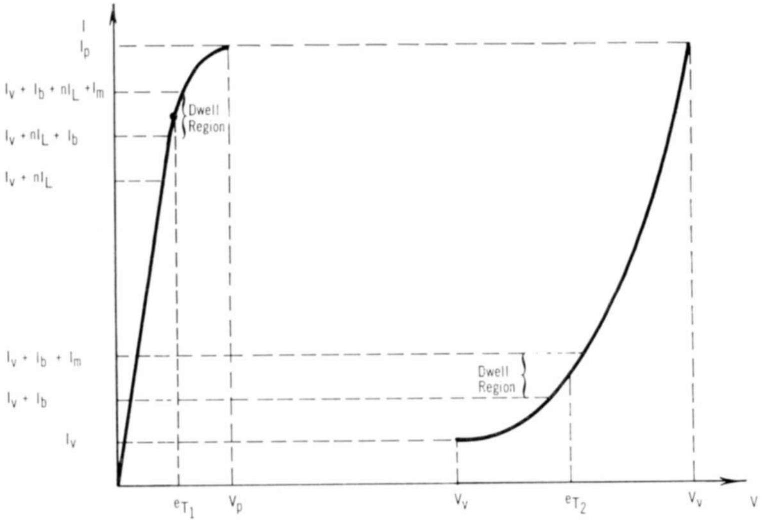


Fig. 7—Tunnel-diode characteristic showing dwell region currents and voltages for loaded condition.

This condition, where minimum current is drawn for a given load, is defined as "threshold" operation, and will be satisfied only for a particular input voltage, V_b. As the diode voltages e_{T_1} and e_{T_2} increase with increasing current, this value of V_b for threshold will also be the minimum input voltage at which the inverter will operate under given conditions of load. Also, the value of V_b for threshold operation will be greater at full load than at zero load.

Thus, in the general case, the dwell points for half a cycle will be, as shown in Figure 7,

$$\text{1st diode:} \quad nI_L + I_v + I_b \rightarrow nI_L + I_v + I_b + I_m,$$

$$\text{2nd diode:} \quad I_v + I_b + I_m \rightarrow I_v + I_b, \tag{11}$$

where I_b is some additional d-c current depending on the bias voltage, V_b. These dwell points are reversed in the second half cycle.

In order to compute the required input voltage, V_b, an expression for the *I-V* characteristic of the diodes is required. In general,

$$V_b = \frac{1}{2I_m}\left[\int_{I_v+I_b}^{I_v+I_b+I_m} e_2(i_2)\,di + \int_{I_v+I_b+nI_L}^{I_v+I_b+nI_L+I_m} e_1(i_1)\,di\right]. \tag{12}$$

To determine the voltages e_1 and e_2, $i_1 = f_1(e_1)$ and $i_2 = f_2(e_2)$ are required. For the tunneling region of a tunnel diode, Karlovsky[8] derived the expression

$$I = A''' \, (e_1 + e_2 - g_V)^2. \tag{13}$$

As e_1 is only required from 0 to V_p, Equation (13) suggests that a simple square function may be used.

The forward region of the tunnel diode is essentially the normal diode injection region (for good I_p/I_v ratio diodes, because the excess current component of I_v is then small), and thus follows the normal exponential function. Again, as e_2 is only required over a small range, a square law seems reasonable.

The relations for e_{T_1} and e_{T_2} given by Equations (14) and (15) are reasonable approximations to the *I-V* characteristics of both Ge and GaAs tunnel diodes over the required ranges as shown in Figures 8 and 9.

$$e_{T_1} = V_p\left[1 - \sqrt{1 - \frac{i}{I_p}}\right], \tag{14}$$

$$e_{T_2} = V_v + (V_f - V_v)\left[\sqrt{\frac{i - I_v}{\Delta I}} + \frac{1}{3\Delta I}\sqrt{(i - I_v)(I_p - i)}\right], \tag{15}$$

where $\Delta I = I_p - I_v$ and V_f = forward voltage. Note that Equation (15) gives slightly lower voltages for the forward region, leading to pessimistic results in later calculations.

Substituting these expressions in Equation (12) and integrating

---

[8] J. Karlovsky, "Simple Method for Calculating the Tunneling Current of an Esaki Diode," *Phys. Rev.*, Vol. 127, p. 419, July 15, 1962.

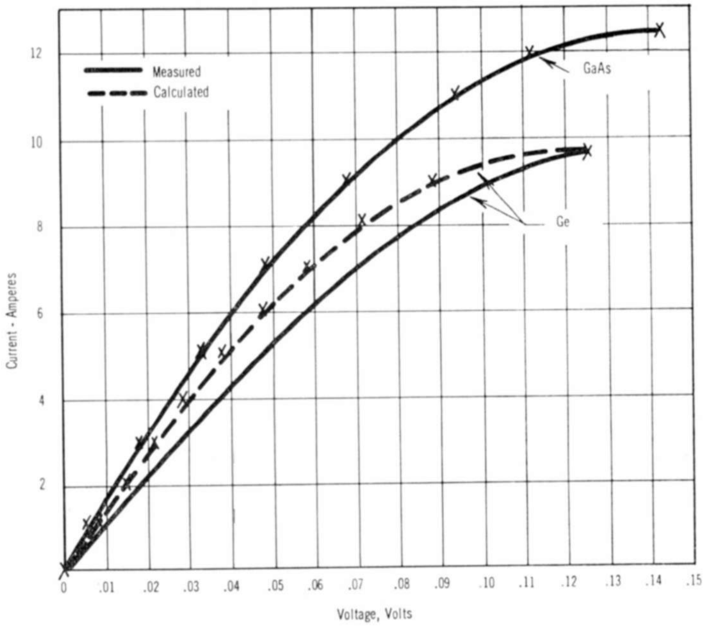


Fig. 8—Comparison between measured and calculated I-V characteristics of the tunneling region of gallium arsenide and germanium tunnel diodes.

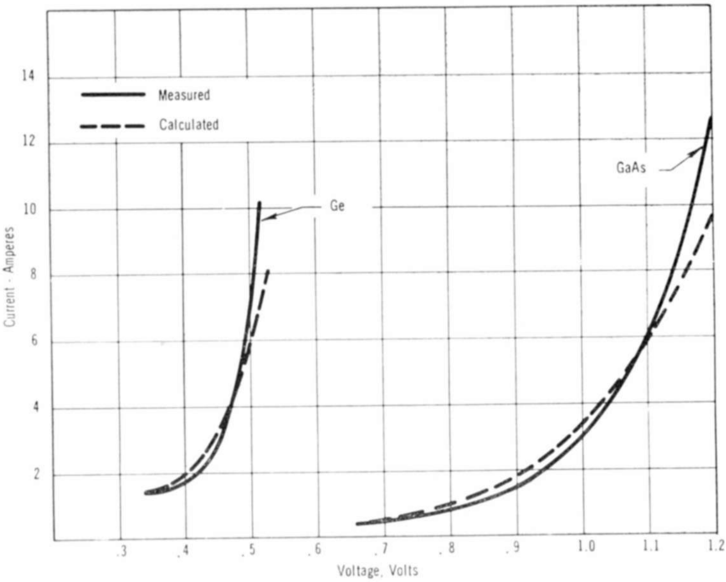


Fig. 9—Comparison between measured and calculated I-V characteristics of the forward region of gallium arsenide and germanium tunnel diodes.

gives the input voltage for given dwell points. However, to simplify the results, an approximation will be made by taking the mean diode voltages at the extremes of the dwell region (see Figure 7), i.e., taking the diode voltages at the currents given in Equation (11) and averaging.

Hence,

$$V_b = \frac{1}{4}\left\{ 2(V_v + V_p) + \frac{V_f - V_v}{\sqrt{\Delta I}}\left[\sqrt{I_b + I_m} + \sqrt{I_b}\right] + \right.$$

$$\frac{V_f - V_v}{3\Delta I}\left[\sqrt{(I_b + I_m)(\Delta I - I_b - I_m)} + \sqrt{(I_b)(\Delta I - I_b)}\right] -$$

$$\left. \frac{V_p}{\sqrt{I_p}}\left[\sqrt{\Delta I - nI_L - I_b} + \sqrt{\Delta I - nI_L - I_b - I_m}\right]\right\}. \quad (16)$$

Equation (16) gives the input voltage, V_b, required for operation at a given load and given dwell points. With the exception of I_m and I_b, all the terms in Equation (16) can be obtained from the diode characteristics.

The magnetizing current, I_m, required to saturate the transformer core may be calculated from the core parameters and the number of primary turns

$$I_m = \frac{B_s}{\mu_s} \frac{1}{0.4\pi N}, \quad (17)$$

where B_s = saturation flux density in gauss,

μ_s = permeability at saturation in gauss per oersted,

l = mean magnetic path length in centimeters,

N = one half the number of primary turns.

b. *Output Voltage*

Having established the required input voltage, the output voltage can be calculated using Equation (2). The exciting voltage $e_x = V_b - e_{T_1}$. Hence, the r-m-s voltage across each half of the transformer primary, V_T, is given by

$$V_T = V_b - e_{T_1}, \quad (18)$$

since e_x is a square wave. Substituting in Equation (14) gives

$$V_T = V_b - V_p + \frac{V_p}{2\sqrt{I_p}} \left(\sqrt{\Delta I - nI_L - I_b} + \sqrt{\Delta I - nI_L - I_b - I_m} \right) \tag{19}$$

Equation (19) gives the r-m-s voltage across each half of the primary as a function of the load current. Using Equations (16) and (19), this voltage may be computed for any load and any input voltage.

Having calculated the r-m-s voltage across each half of the primary, the output voltage, operating frequency, power output and inversion efficiency of the inverter may be calculated.

The r-m-s output voltage is the product of the r-m-s voltage across each half of the primary and the ratio, n, of the secondary turns to half the primary turns. Therefore

$$V_{\text{out}} = nV_T. \tag{20}$$

c. *Operating Frequency*

It is known that

$$N \frac{d\phi}{dt} \times 10^{-8} = e_{\text{rms}}.$$

As the core is driven from $-\phi_s$ to $+\phi_s$ during half a cycle (see Figure 3c),

$$N \frac{2\phi_s}{T/2} \times 10^{-8} = V_T,$$

where T is the period of oscillation. Therefore,

$$f = \frac{V_T}{4\phi_s N} \times 10^{8} \text{ cycles per second.}$$

Since $\phi = B_s A$, the operating frequency is given by

$$f = \frac{V_T}{4ANB_s} \times 10^{8} \text{ cycles per second.} \tag{21}$$

d. *Output Power*

The output power is the product of the output current and voltage

$$P_{\text{out}} = nV_T I_L. \tag{22}$$

e. *Inversion Efficiency*

The inversion efficiency, η, is given by

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}}.$$

The input current is the sum of the diode currents during a full cycle. Therefore,

$$I_{\text{in}} = nI_L + I_m + 2I_v + 2I_b, \tag{23}$$

$$P_{\text{in}} = V_b \, (nI_L + I_m + 2I_v + 2I_b), \tag{24}$$

and

$$P_{\text{out}} = V_T n I_L. \tag{25}$$

Therefore

$$\eta = \frac{V_T n I_L}{V_b (nI_L + I_m + 2I_v + 2I_b)} \times 100 \text{ percent.} \tag{26}$$

A summary of the equations predicting the performance of saturated square hysteresis loop core, push–pull tunnel-diode inverters is given in Table I.

Special Cases

a. *Threshold Operation*

As can be seen from Equation (16), the input voltage required for threshold operation (i.e., $I_b = 0$) increases from zero load to full load. Since the output voltage from the power source will decrease with increased load, it is most important to consider V_b at full load.

If the input voltage is lower than that required for threshold operation, then one of two possibilities occur. One is that the inverter will cease to oscillate in the push–pull mode, reverting to inphase oscillation with zero output. A second possibility is that the inverter may operate in some other mode than that of push–pull saturated core inverters.

b. *Maximum Load Current*

The maximum load current that can be drawn from the inverter is

$$nI_{L(\max)} = \Delta I - I_m - I_b, \tag{27}$$

and hence it varies with bias voltage, decreasing as the bias voltage is increased. The output voltage, nV_T, increases with increasing bias voltage. Hence, the power output should peak at some input voltage where the increase in nV_T is greater than the reduction in $nI_{L(\max)}$. The peak efficiency, however, will occur at some combination of load and bias voltage not necessarily the same as that for maximum power output.

Table I—Analytic Expressions Predicting the Performance of Saturated Square Hysteresis Loop Core, Push-Pull Tunnel Diode DC to AC Inverter

---

Bias Voltage

$$V_b = \frac{1}{4}\left\{ 2(V_v + V_p) + \frac{V_f - V_v}{\Delta I}\left[\sqrt{I_b + I_m} + \sqrt{I_b}\right] \right. + \frac{V_f - V_v}{3\Delta I}\left[\sqrt{(I_b + I_m)(\Delta I - I_b - I_m)} + \sqrt{I_b(\Delta I + I_b)}\right] \left. - \frac{V_p}{I_p}\left[\sqrt{\Delta I - nI_L - I_b} + \sqrt{\Delta I - nI_L - I_b - I_m}\right]\right\} \tag{16}$$

Output Voltage

$$V_T = V_b - V_p + \frac{V_p}{2\sqrt{I_p}}\left(\sqrt{\Delta I - nI_L - I_b} + \sqrt{\Delta I - nI_L - I_b - I_m}\right) \tag{19}$$

Frequency

$$f = \frac{V_T}{4ANB_s} \times 10^8 \tag{21}$$

Input Current

$$I_{\text{in}} = nI_L + I_m + 2I_v + 2I_b \tag{23}$$

Input Power

$$P_{\text{in}} = V_b[nI_L + I_m + 2I_v + 2I_b] \tag{24}$$

Output Power

$$P_{\text{out}} = nV_TI_L \tag{25}$$

Inversion Efficiency

$$\eta = \frac{nV_TI_L}{V_b[nI_L + I_m + 2I_v + 2I_b]} \times 100 \tag{26}$$

---

c. *Overload Condition*

If the maximum load current defined by Equation (27) is exceeded, the diode supplying the load current will switch to the high state. This is because the diode current can only increase to I_p, at which time $di/dt = 0$, and e_x becomes zero, so that the voltage equilibrium is disturbed and the diodes switch.

Thus, under overload conditions, the switching of the diodes is no longer controlled by the saturation characteristics of the transformer, and normal inverter operation ceases. The d-c input current will reach a maximum value of $I_p + I_v$, and cannot exceed this level if the diodes are oscillating. Hence, in the tunnel-diode inverter, no external overload or short-circuit protection is required.

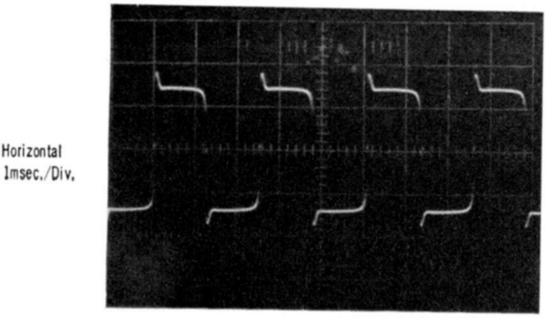


Fig. 10—Output waveform of a saturated-core push-pull tunnel-diode inverter—no-load condition.

TUNNEL-DIODE INVERTER PERFORMANCE

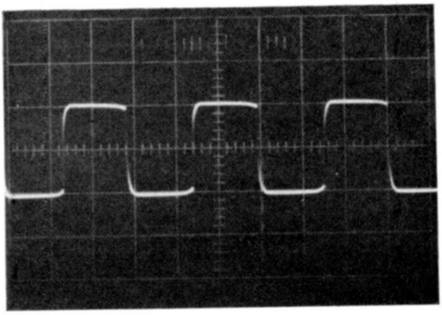
Several tunnel diode inverters ranging in output from 3 watts to 200 watts were fabricated and tested. The low-power inverters were operated from a low impedance, low-voltage regulated power supply, while the high-power units were tested with a homopolar generator as the power source. The power output was determined by measuring the r-m-s voltage across a known load resistance, and computing the power using the relation $P_{\text{out}} = V_{\text{rms}}{}^2/R_L$. Input current and voltage were measured using calibrated low-resistance shunts and voltmeters, and operating frequency was measured with a frequency counter.

Output Wave Shape

From Figures 3 and 5, it can be seen that the predicted waveform, under all load conditions, is a square wave. Figure 10 is an oscillogram of the waveshape of an unloaded inverter taken at the secondary terminals. Notice that it is a good square wave, but that there is a spike on the leading edge. This spike is the overvoltage due to the tunnel-diode current rising rapidly from the dwell region ($I_v + I_m$ at no load) to I_p before switching to the forward region of its characteristic.

It would, therefore, be expected that as the inverter is loaded, the amplitude of the spike would be reduced, since the dwell region approaches I_p, and that the spike would disappear at full load conditions.

Figure 11 shows an oscillogram of an inverter under full load. Notice that the wave shape is still square, and that there is no spike on the leading edge.



Horizontal
1msec./Div.

Fig. 11—Output waveform of a saturated-core push-pull tunnel-diode inverter—full-load condition.

Inversion Efficiency, Operating Frequency, and Input Current

Using Equations (16), (19), (21), and (22), inversion efficiency, operating frequency, power output and input current were computed for transformer No. 5B using 12-ampere GaAs diodes.

Figure 12 compares the variation of inversion efficiency, operating frequency, and input current, all computed for a given input voltage. Measured values are taken at the same input voltage and at power levels from 0 to 3 watts.

Comparison between the computed and measured performance shows good agreement. The curves are in excellent agreement, indicating the validity of the analysis. The discrepancy in operating frequency is easily within the errors in the specifications for saturation flux density and cross-sectional area of the core; furthermore, since the number of primary turns is small (3 per half), it is difficult to be precise. For example, a value of $B_s = 6800$ gauss was assumed, while the manufacturer's data shows that B_s may vary between 6600 and 8200 gauss.

It should be noted that in the computations, the transformer losses are not taken into account. It would, therefore, be expected that the computed efficiency would be somewhat greater than the measured efficiency. However, test results (see Figure 12) indicate that at full load the efficiency is somewhat greater than predicted. This can be explained as follows. As the load increases, the bias current, I_b, for a given bias voltage decreases, and as a result the high-voltage-state diode dwells closer to the valley region. As can be seen in Figure 9, the diode characteristic in this region is flatter than that calculated from Equation (15) and hence E_{T_2}, and thus V_T, are greater than the calculated values, leading to greater output power.

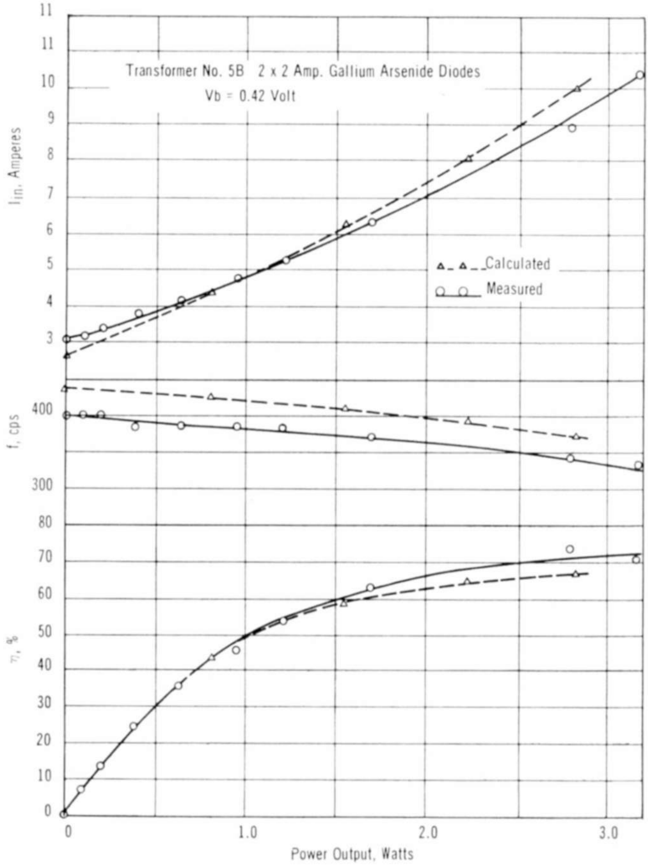


Fig. 12—Comparison between calculated and measured variation of efficiency, frequency and input current with load.

Figures 13(a) and 13(b) compare the variation of inversion efficiency, operating frequency and input current, computed at two input voltages, with measured values taken at the same two voltages at a power level of 40 watts.

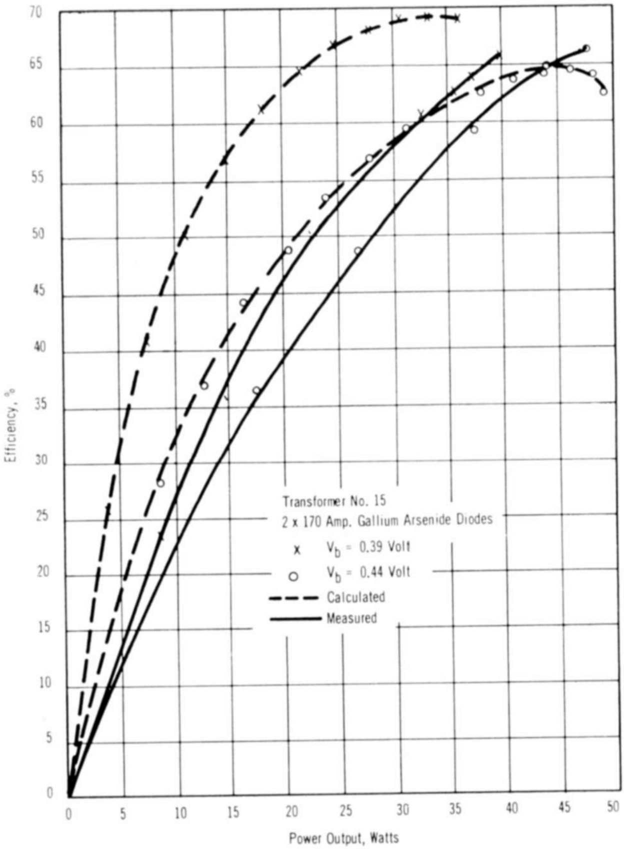


Fig. 13(a)—Comparison between calculated and measured variation of efficiency.

Notice that at this higher power level, the discrepancy between computed and measured efficiency and input current are quite large at partial load conditions, but at full load good agreement exists. It should also be noted that at the higher input voltages, the discrepancies are smaller. The difference between measured and calculated frequency can easily be explained as before.

Table II lists the performance of tunnel inverters with power outputs of 3 watts to 200 watts. In all cases efficiencies greater than the figure of merit were achieved. It should be noted that pairs of diodes with the same figures of merit will not necessarily have the same efficiency in a circuit because of differences in the forward part of the I-V characteristic, which does not affect the figure of merit.

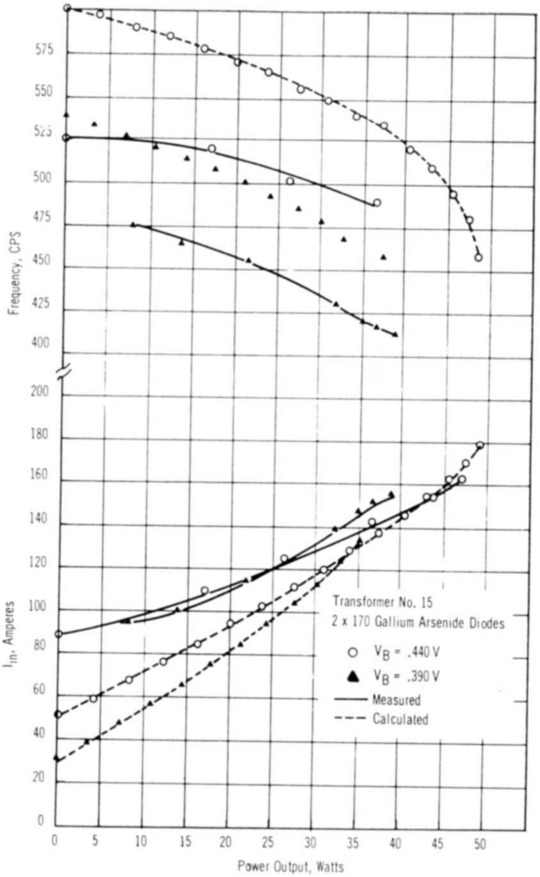


Fig. 13(b)—Comparison between calculated and measured variation of frequency and input current.

The discrepancy between measured and calculated efficiency and input current are due to second-order effects not significant at low power levels.

Second-Order Effects

As noted in the previous section, the major discrepancy between predicted and measured inverter performance was observed when high-power inverters were operated at partial load. Under these conditions, the efficiency was much lower than predicted.

Investigation showed that the low efficiency was due to the input current being much greater than would be predicted at these partial-load conditions, the expected values being $I_m + 2I_v + 2I_b$ at open circuit and $I_m + 2I_v + 2I_b + nI_L$ when the inverter is loaded.

Further investigation revealed that under these no-load or partial-load conditions, the input current depended not only on I_m, I_v, and I_b, but also varied with the operating frequency, squareness of the core material hysteresis loop, circuit resistance, diode peak current, and magnetizing current. These effects are discussed in more detail below.

Table II—Comparison between Measured Efficiency and Diode Figure of Merit for Inverters of Various Power Levels

Diodes	Figure of Merit %	Power Output Watts	Measured Efficiency %	Frequency cps
12 amp	60	3	78	400
25 amp	61	6	79	400
170 amp	59	40	70	400
350 amp	55	80	68	400
600 amp	55	140	68	400
1000 amp	50	220	—	400

a. Effect of Hysteresis Squareness and Circuit Resistance

In the derivation of the theoretical expressions for efficiency, it was assumed that the total circuit resistance was zero and that the hysteresis loop of the core material used was perfectly square, i.e., $B_r/B_{(\max)} = 1$ where B_r is the remanence value of B. In practice, however, neither of these assumptions are correct, and at high power levels their effects are significant.

If the total circuit resistance, R_s, is not zero, but is smaller than the negative resistance of the tunnel diodes (a necessary condition for operation), then the voltage equilibrium equations become

$$\text{low-voltage-state diode:} \qquad V_b - e_{T_1} - e_x - i_1 R_s = 0,$$

$$\text{high-voltage-state diode:} \qquad V_b - e_{T_2} + e_x - i_2 R_s = 0.$$

When the core is saturated, $i_2 = I_v + I_b$, and since R_s is small, $i_2 R_s$

may be neglected. However, during switching i_1 changes from $I_v + I_b + I_m$ (assuming no load) to I_p. Also, if $B_r/B_{\max} < 1$, $nd\phi/dt > 0$ after saturation, and a small voltage $e'_x (e'_x < e_x)$ will be developed as the low-voltage-state diode current increases toward I_p. The voltage equilibria are now:

$$\text{low-voltage-state diode:} \quad V_b - e_{T_1} - e'_x - i_1 R_s = 0,$$

$$\text{high-voltage-state diode:} \quad V_b - e_{T_2} + e_x = 0.$$

As e_{T_2} is much larger than V_b, and e'_x is small, the high-voltage-state diode will switch to a low-voltage state. However the low-voltage-state diode will be delayed from switching, with the current increasing slowly to I_p, since e_x increases as dt decreases. Since $V_T = (e_{T_2} - e_{T_1})/2$, during this time $V_T = 0$, as e_{T_1} is small. Thus, poor squareness and high resistance prevent the diodes from switching simultaneously, leading to an input current greater than predicted.

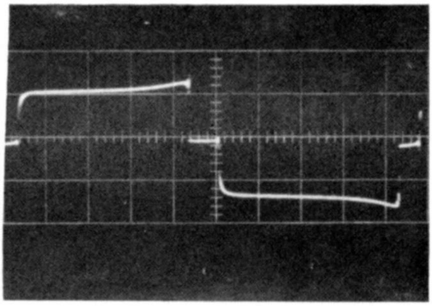


Fig. 14—Output wave shape showing step caused by poor squareness and high resistance.

As the squareness is improved, $d\phi/dt$ is reduced, and hence dt can decrease, i.e., the delay is lessened.

Figure 14 is an oscillogram of the output wave shape of the 40-watt inverter operating at partial load. The step can be clearly seen. At low power levels, these effects are very small. Figure 15 shows the variation of input current with input voltage under no-load conditions for a three-watt inverter. It can be seen that the measured and predicted values agree quite closely.

b. *Effect of Operating Frequency*

The delay caused by the poor squareness of the core-material hysteresis loop and circuit resistance is constant for a given core material and primary circuit. Under these conditions, the operating frequency is determined by the core area alone. Thus, the lower the operating frequency, i.e., the greater the period of oscillation, the smaller the effect of the switching delay.

Figure 16 shows the variation of the input current with frequency and input voltage. It will be noted that the input current decreases with decreasing frequency to a minimum value at a given input voltage. This minimum occurs where $I_b = 0$.

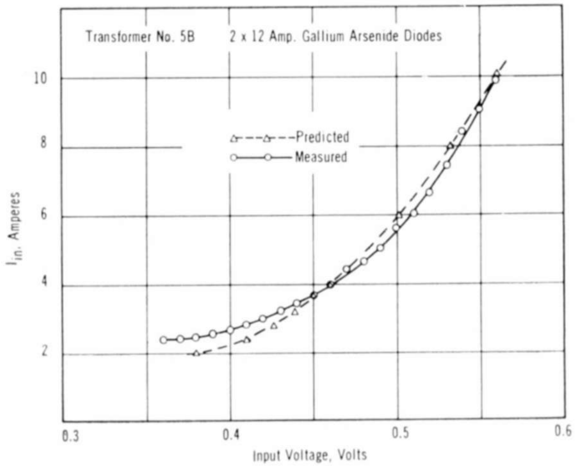


Fig. 15—Comparison between calculated and measured variation of input current with input voltage for open circuit.

c. Effect of I_p

For a given transformer (I_m, R_s, frequency, and squareness fixed) increasing peak current will increase the delay, and thus reduce the partial-load efficiency of the inverter.

d. Effect of I_m

For given core material and operating frequency, increasing I_m reduces the delay, since the current change in the diode is reduced.

Figure 17 shows the effect of increasing the magnetizing current, I_m. Three transformers were designed having I_m in the ratios 1:2:3, but the same operating frequency. It can be seen that for increasing I_m, the input current at no-load conditions decreases.

e. *Effect of Input Voltage*

As can be seen from Figure 13(b), as the input voltage is increased, the discrepancy between measured and calculated values of input currents are reduced. This is due to the increase of I_b with increasing input voltage, reducing the current swing, and hence the switching delay, in the diodes.

OPERATING FEATURES

Overload and Short-Circuit Operation

Figure 18 shows the effect of overload on an 80-watt inverter. Figure 18a shows the output wave form at full load, Figure 18b shows the wave form with slight overload, and Figure 18c shows the wave form for heavy overload. It will be noted that the output voltage decreases with overload, and the operating frequency increases.

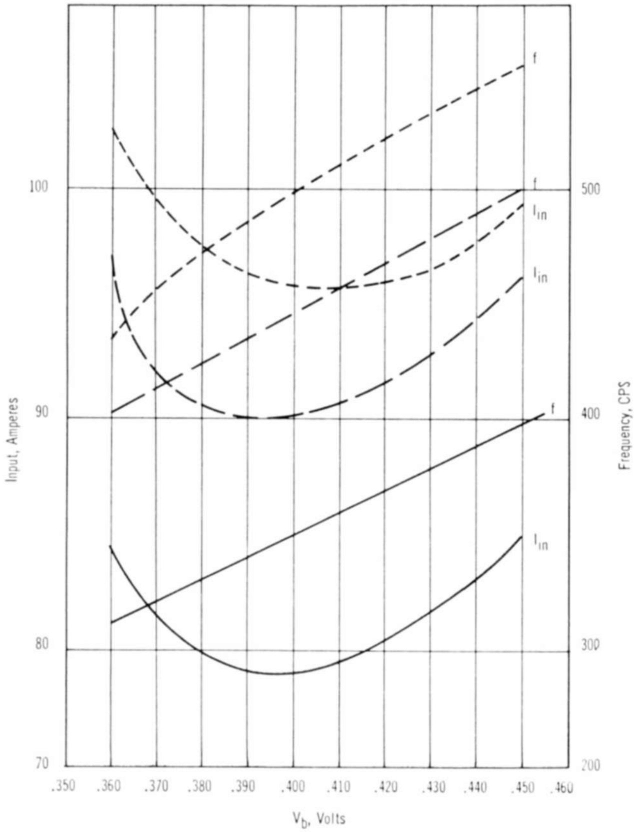


Fig. 16—Variation of input current with frequency and input voltage—no-load condition.

Tests were carried out to determine the effect of a d-c motor load on the performance of tunnel diode inverters. Initially, the motor presents a short circuit to the inverter, with the load decreasing as the motor speed increases to its normal value. Thus, a 100-watt motor initially overloads a 100-watt inverter.

However, the overload characteristics of the inverter are of value in overcoming this problem, as the inverter provides some output even at the overload condition (see Figure 18c). This output starts the motor (assuming it is unloaded), and as the motor speed increases, it unloads the inverter, the output of the inverter increases, causing the motor speed to increase and so on until the motor is at normal speed.

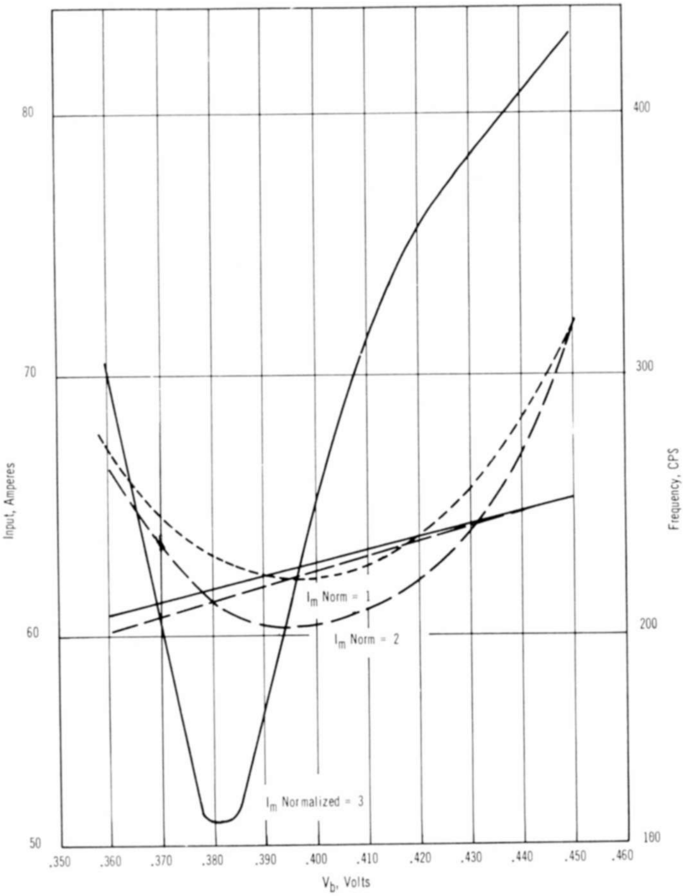
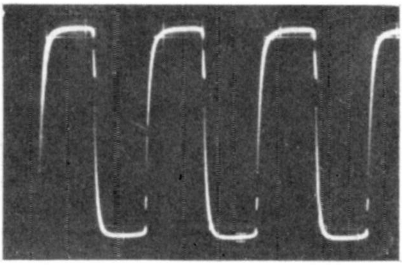


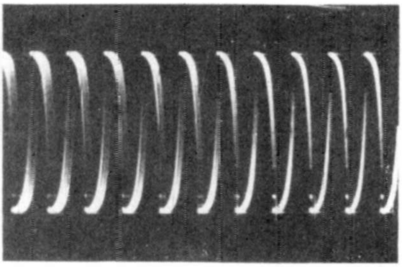
Fig. 17—Variation of input current with I_M and input voltage with frequency constant—no-load condition.

This effect is illustrated in Figure 19. Figure 19a shows the output just after switching on the motor, and Figure 19b shows the output with the motor at full speed.

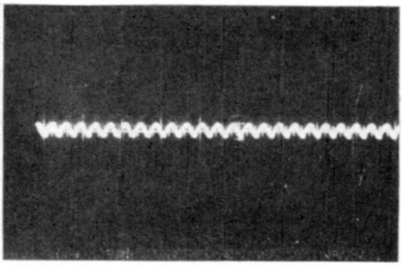
This built-in short-circuit protection has the advantage of isolating the load from the supply, and by eliminating the need for external short circuit protection, reduces the weight and complexity of the circuit. It also allows the inverter to be started under overload or short-circuit condition, with normal operation resuming after the overload is removed.



(a)
Full Load



(b)
Slight
Overload

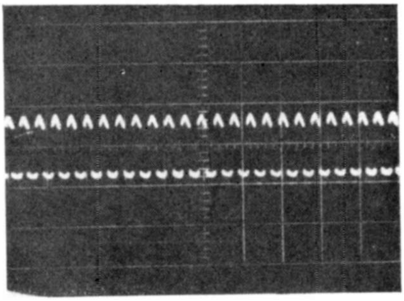


(c)
Heavy
Overload

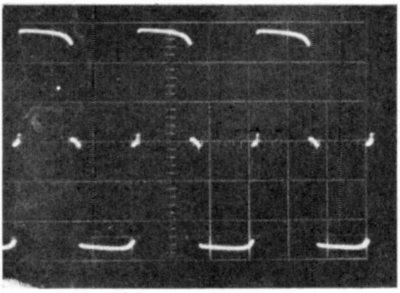
Fig. 18—Tunnel-diode inverter overload characteristic.

Operation at Elevated Temperatures

The major restriction on the operation of present tunnel-diode inverters at elevated temperatures is the solder materials used in the diode fabrication. The dopant material and solders have relatively low melting points, limiting operation to about 100°C. These limitations may be overcome by replacing these solders with materials having higher-temperature capability. However, such materials must be carefully investigated to obtain good electrical characteristics.



(a) Output waveform—motor switched on.



(b) Output waveshape—motor running at normal speed.

Fig. 19—Output waveshape of inverter showing effect of d-c motor starting load.

An analytic study was performed to determine the variation of efficiency with temperature, assuming the solution of the solder material limitation, and using the normal temperature coefficients of peak, valley, and forward voltages for tunnel diodes[9] (—0.1 mv/°C, —0.8 mv/°C and —1.1 mv/°C, respectively), and the usual variation of valley current with temperature.

---

[9] R. M. Minton and R. Glicksman, "Theoretical and Experimental Analysis of Germanium Tunnel Diode Characteristics," *Solid-State Electronics*, Vol. 7, p. 491, July 1964.

Figure 20 shows the variation of efficiency with temperature computed for diodes having a figure of merit of 69%. It will be noted that the efficiency remains at a high value over a wide range of temperatures, and is 66% at 200°C.

As the major effect of elevated temperatures is to lower the diode voltages, use of materials such as GaAs-GaP alloys with higher band gaps than GaAs could improve high-temperature performance, since the initial voltages would be higher.

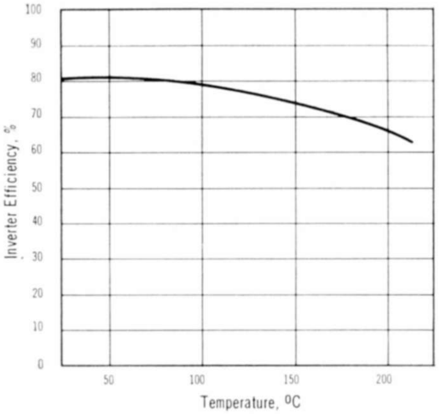


Fig. 20—Variation of inverter efficiency with operating temperature.

Operation with Thermionic Generator as a Power Source

The operation of tunnel-diode inverters with thermionic generators has previously been demonstrated[7] with an output of 20 watts being achieved. Further tests have demonstrated an output of 60 watts at 110 volts, again demonstrating the feasibility of obtaining useful power from thermionic-generator–tunnel-diode inverter systems.

Control of Frequency and Phase of Operation

The operating frequency of these inverters,

$$f = \frac{V_T}{4ANB_S} \times 10^8 \text{ cycles per second,} \qquad (21)$$

changes with load as the diode dwell points, and hence V_T, changes with load. For applications where a constant operating frequency independent of load is required, or where both frequency and phase of several inverters must have a given relationship, such as in a three-phase inverter, some form of control is required. As the tunnel diode is a two-terminal device, any control must be external to the diode.

Marzolf[10] described a circuit for using tunnel diodes in series, which was triggered by an external pulse generator. This idea was applied to the saturated-core, push–pull inverter to provide frequency and phase control of the inverter.

Although this circuit operates satisfactorily, it has the disadvantage of consuming relatively large powers and is critical in design. Therefore a simple circuit called the "Overload" circuit was developed.

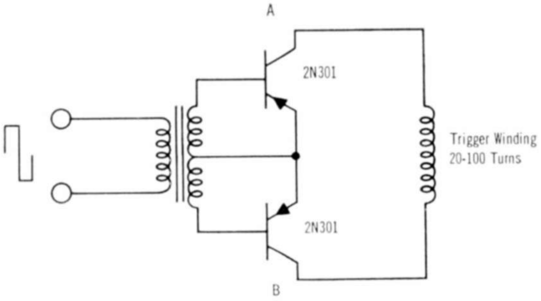


Fig. 21—Overload triggering circuit.

If an inverter is loaded such that the primary load current is greater than $\Delta I - I_b - I_m$, the transformer core can not be saturated, and a different mode of operation results. In this case, when V_p is reached, the diode must switch to the high-voltage state, even though the core is still unsaturated, as there can be no further current change, causing e_x to collapse and disturbing the voltage equilibria.

This effect is utilized for triggering by employing the simple circuit shown in Figure 21. Assume that the inverter is operating normally, and that at some time $T_{\text{trig}} = 0$, when point A is negative with respect to B, a trigger pulse is applied so that the base of transistor A is driven negative.

Under these conditions, transistor A conducts, and because transistor B is forward biased, this transistor also conducts. Thus, the trigger winding is short circuited, overloading the inverter and causing the tunnel diodes to switch to their opposite states. At this time, transistor A is forward biased, and remains conducting. However, transistor B is reverse biased, and has a positive base bias, and is turned off, unloading the trigger winding. Normal operation continues

---

[10] J. M. Marzolf, "Application of Tunnel Diode Static Inverter to Higher Input Voltages," U.S.N.R.L. Report No. 5802, June 1961.

until the trigger pulse polarity is reversed, at time $t = T_{\text{trig}/2}$ causing transistor B to conduct. Again the trigger winding is short circuited, and the tunnel diodes reverse their states.

Thus, the inverter will operate at a frequency determined by the trigger frequency, provided the trigger frequency is greater than that of the normal operating frequency of the inverter.

If n_T is the number of turns on the trigger winding, then the peak current $\hat{I}_T$ flowing through the transistors is

$$\hat{I}_T = \frac{(\Delta I - I_b - I_m - nI_L)N}{n_T}. \tag{28}$$

As the inverter is loaded, the peak triggering current decreases, and the efficiency loss is reduced.

This circuit is extremely simple, and consumes very little power. Since the short-circuit condition of the trigger winding is very short, little power is dissipated in the transistors, eliminating the need for heat sinks, while maintaining high efficiency. Also note that the overload circuit requires no d-c power supply.

Figure 22 is a block diagram of a three-phase tunnel-diode inverter, using overload triggering. Such a circuit was built and operated, using separate 1-watt inverters.

DESIGN AND FABRICATION CONSIDERATIONS

Core Material Parameters

A study of the core materials listed in Table III established the following guidelines for the proper design of a push–pull tunnel-diode inverter with square-hysteresis-loop saturated cores.

1) For 5- to 10-watt inverters, where the magnetizing current value is significant compared to the diode peak current, square Permalloy 80 or equivalent is satisfactory. Test results indicate good agreement with the predicted performance from zero load to full load. Although this material has a relatively low squareness, the wave shape was very close to a square wave.

2) For inverters with higher power outputs, the choice of core material depends on the character of the load. For an inverter that is intended to operate at full load only, the problem of hysteresis-loop squareness is not critical. In this case, a square Permalloy 80 core or equivalent can be used. Usually cores with higher flux densities have higher squareness. In that case

Table III—Summary of Important Parameters of Core Materials Investigated

Material Type	Manufacturer	Flux Density (kilogauss)	Permeability at Sat., at 400 cps (gauss/oersted)	Squareness (%)
Square Permalloy 80	Magnetics Inc.	6.6-8.2	~25,000	80 up
48 Alloy	Magnetics Inc.	11.5-14.0	~25,000	80-92
Square mu 49	Magnetic Metals Co.	14.0-15.5	—	92 up
Ferramic "H"	Indiana General	~3.4	—	~55
Ferramic "T-1"	Indiana General	~4.4	4,000	—

the increase in I_m may be compensated for by lower values of bias current I_b or by moving the high-voltage-state dwell region upward, which increases the output voltage. Such cases should be analytically evaluated. In cases where the load varies from partial load to full load, only cores with high squareness should

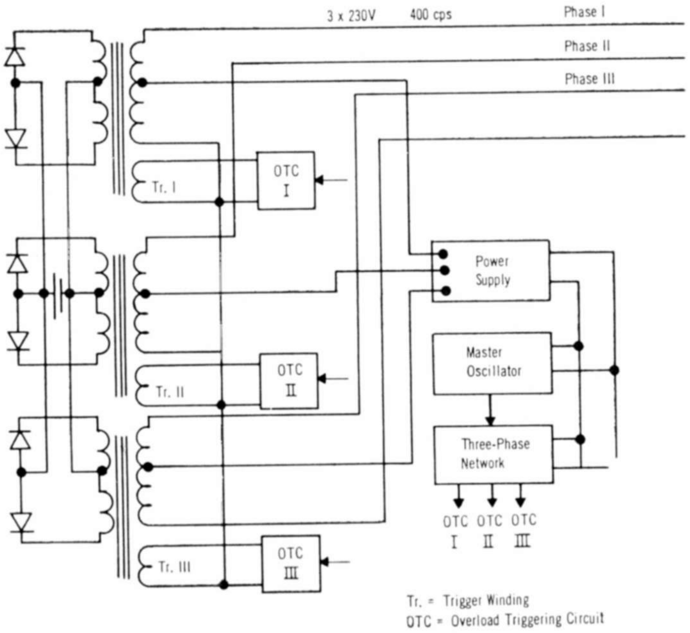


Fig. 22—Three-phase tunnel-diode push-pull inverter block diagram.

be considered. This includes the case when a low operating frequency is required.

3) Ferrite cores, due to nonsquareness, low saturation flux density, and very low saturation permeability, are not suitable.

Shape of Core Material

In initial transformer studies, ring-type laminations were used because they provide a gapless core, and thus have the best saturation characteristics and the lowest losses. However, toroidal cores have the limitation of requiring either the use of sectional primary windings or flexible windings. As the d-c resistance of the primary must be kept in the low micro-ohm range for high-power inverters, toroidal cores become unattractive from fabrication considerations.

Thus, studies were made with transformers made with DU-type laminations, as this shape most nearly approached the gapless structure. Tests indicated performance similar to toroidal cores, and because of the ease of fabrication of primary windings, this shape was used for all high-power inverters.

Inverter Fabrication

The use of DU-type laminations allowed easy fabrication of very-low-resistance primary windings. The technique used was to mill a spiral cut in a thick-walled tube of appropriate dimensions, and then to braze the center tap and diode connections to the spiral. Figure 23 is a photograph of a completed primary winding.

The transformer is completed by assembling the laminations through both the primary and secondary windings, and then clamping the stack together. Figure 24 shows two views of a completed inverter with the tunnel diodes attached. The connections are made by bolting the transformer to the diodes and the diodes to a baseplate by insulated thru-bolts, so as not to short the diodes.

This type of construction allows simple fabrication and assembly with maximum flexibility in changing core material, secondaries, and diodes.

Computer Program

The hand computation of the performance of an inverter can be seen to be very time consuming because of the inter-relation of Equations (15) and (18).

Therefore, a computer program was written so that complete performance calculations could be performed quickly. The values of the diode and transformer parameters and the range of supply voltage and

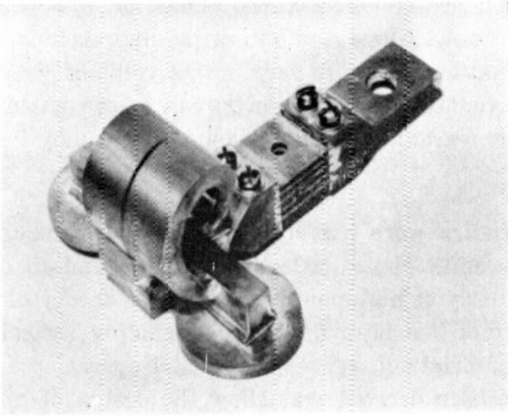


Fig. 23—Primary winding of a high-power inverter (200 watt).

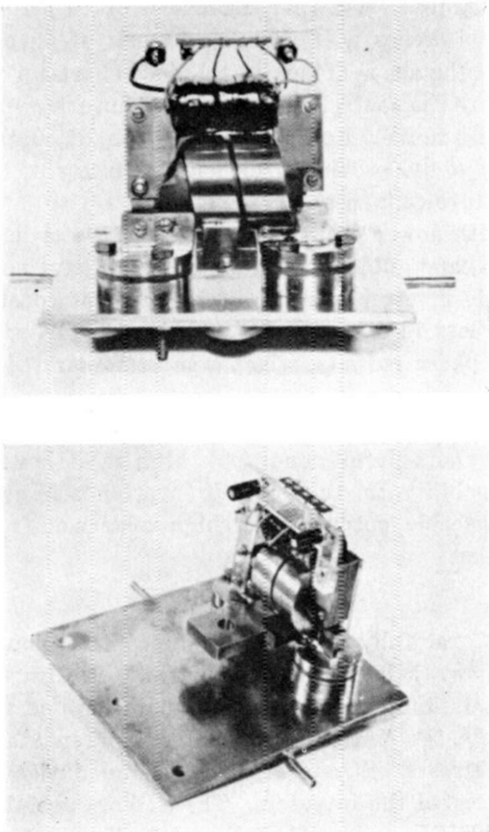


Fig. 24—Front and back views of a complete high-power inverter (200 watt).

load in question are supplied to the computer, and the performance expressions are solved using an interative approach to the V_b and I_b relation of Equation (15). By varying the value of V_b, a complete set of performance curves for this inverter can be calculated, or by changing transformer parameters, an optimum design may be achieved.

CONCLUSIONS

Detailed studies were performed on saturated-core push–pull inverters to determine their mode of operation and to optimize their design, particularly at high powers. A detailed theory of the operation of these inverters has been developed, including second-order effects such as core-material squareness factor and operating frequency, and equations have been derived that allow the design of and predict the performance of such inverters. These studies have shown that for saturated-core push–pull inverters, the previously accepted theory of operation is inadequate, and that efficiencies of 80 percent are attainable using tunnel diodes with figures of merit of 60 percent. It has been shown that the shape of the tunnel-diode characteristic and transformer design are the controlling elements in inverter efficiency. Thus, improvements in tunnel diodes from the use of materials such as GaAs-GaP alloy would provide efficiencies in excess of 80 percent and higher temperature capability.

Inverters with power outputs of up to 200 watts have been fabricated and tested, and guidelines for their design have been established. The tunnel diode inverter can also be operated at constant frequency with negligible loss in efficiency by external triggering. This feature also allows polyphase outputs, which was demonstrated with a three-phase inverter.

Because of its simplicity, high power to weight ratio, tolerance of overload and short-circuit conditions, high radiation resistance, and temperature capability, the tunnel-diode inverter is an attractive device for obtaining useable outputs with high efficiency from thermionic generators.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of R. Glicksman for his many helpful suggestions, and R. Buzzard, who programmed the equations for computer computation and who performed many of the tests. P. A. Hoss and L. Murray grew materials for the tunnel diodes, Mrs. E. Strouse fabricated the diodes, and P. Britt fabricated and tested the inverters. The authors are also indebted to B. J. Wilson, USNRL, and to M. Bell and A. Hom, who developed the homopolar generator used in testing the high-power inverters.

A SOLID-TO-SOLID DIFFUSION TECHNIQUE

BY

J. SCOTT AND J. OLMSTEAD

RCA Electronic Components and Devices
Somerville, N. J.

Summary—*This paper describes a solid-to-solid diffusion technique that permits precise control over a large range of surface concentrations. The addition of a suitable dopant to an alkoxysilane results in a doped oxide film that is used as a diffusion source. The effects on the diffused-layer characteristics of impurity concentrations in the oxide, diffusion temperature, diffusion time, and diffusion ambient are presented. Methods of obtaining silicon dioxide layers by the pyrolytic decomposition of an alkoxysilane are also discussed.*

Application of this technique permits the simultaneous diffusion of n-type and p-type impurities without interaction, and still affords individual control of surface concentration of all areas. As a result, the number of high-temperature heat cycles required is reduced for many purposes, and greater flexibility in design is obtained. Because of the unique placement of the source, novel photolithographic techniques are possible.

This new diffusion technique provides ease of control and reproducibility. A comparison with conventional diffusion techniques is presented.

INTRODUCTION

THE CONTINUING increase in the number of diffusion techniques used to produce p-n junctions demonstrates the general dissatisfaction with the techniques available for the varying needs of the industry.[1] Examination shows that all these techniques are attempts to improve one of two basic systems. The first of these systems, generally referred to as the two-zone or carrier-gas process, has been described in detail by Frosch and Derick[2] (this work describes the diffusion system, together with the masking quality of silicon dioxide). The diffusion system consists of a furnace that contains two controllable heat zones. The impurity source is placed in the first zone, and the wafers to be diffused are placed in the second zone. A slightly oxidizing gas is then passed through the furnace and transports the vapor from the impurity in the first hot zone to the wafers in the second

---

[1] C. S. Fuller and C. J. Frosch, *Transistor Technology*, Vol. 3, Van Nostrand, New York, 1958; F. M. Smits, "Formation of Junction Structures by Solid-State Diffusion," *Proc. IRE*, Vol. 46, p. 1049, June 1958.

[2] C. J. Frosch and L. Derick, "Surface Protection and Selective Masking During Diffusion in Silicon," *Jour. Electrochem. Soc.*, Vol. 104, p. 547, Sept. 1957.

hot zone. The first hot zone controls the vapor pressure of the source and can be eliminated if the source used is sufficiently volatile at room temperature.

The second approach, which has been described by L. A. D'Asaro,[3] utilizes a closed box containing the wafers and the impurity source. The impurity is introduced from a solution of impurity oxide in silicon dioxide. When the box is heated to diffusion temperature, it fills with the vapors of the impurity. The silicon wafers then act as a "sink" for these impurities.

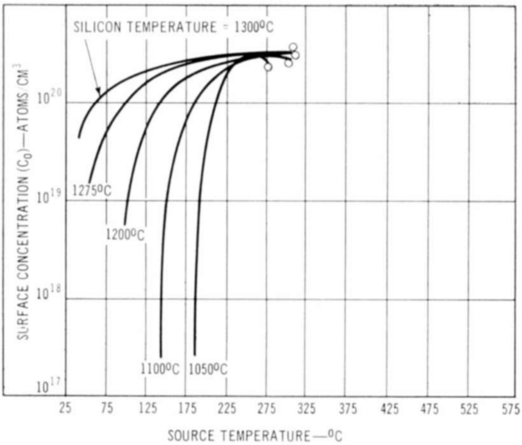


Fig. 1—Surface concentration as a function of source temperature for P_2O_5 two-zone diffusion system.

Both of these methods and their many variations have been applied with reasonable success for surface concentrations at or near the solid solubility of the impurity in the silicon at the diffusion temperature. For surface concentrations much below this value, however, these methods do not give reproducible results and become less controllable.

Figure 1 illustrates the problems encountered in attempting to achieve low surface concentrations. This figure shows surface concentration as a function of source temperature for a typical P_2O_5 two-zone diffusion system. It can be seen that control of source temperature is extremely difficult if a surface concentration less than about 10^{20} atoms per cubic centimeter is required.

The box method becomes troublesome because the whole system cannot approach a satisfactory near-equilibrium condition. (This is

---

[3] L. A. D'Asaro, "Diffusion and Oxide Masking in Silicon by the Box Method," *Solid-State Electronics*, Pergamon Press, 1960, Vol. 1, No. 1, pp. 3-12.

also true, but to a lesser extent, for the two-zone system.) In theory, the surface concentration is determined by the vapor pressure of the impurity source. The vapor pressure, in turn, is controlled by the temperature and, in the case of the box method, by dilution of the source. In practice, the walls of the box complicate the situation by becoming a secondary source, because the quartz box is a sink for the impurity. Near equilibrium can only be approached as the box becomes saturated with the impurity. Unfortunately, as the box approaches saturation, the surface concentration approaches the maximum solid solubility of the impurity in silicon at the diffusion temperature, i.e., the source effectively switches from the dilute mixture to the saturated walls of the container. Thus, the box method is acceptable only for high concentrations.

Because of these major difficulties, deposition and redistribution methods have been adopted to achieve values of surface concentration below the solid solubility. First, a short deposition step utilizing one of the above techniques is employed to saturate the surface with dopant. A drive-in step without the presence of the source is then used to redistribute the impurities and achieve the final depth and surface concentration. However, the fact that the redistribution is accomplished without the presence of a source changes the type of distribution from the complementary-error function to a modified Gaussian type. Furthermore, the control of depth and of surface concentration become interdependent, so that two steps are required and the process is thus less easily controlled. It is necessary to control two times and two temperatures, as well as the ambient, during drive-in. The latter is necessary because oxide is being regrown during the drive-in and a portion of the original impurity introduced in the deposition step is being consumed in the oxide.

THE SOLID-TO-SOLID DIFFUSION TECHNIQUE

In an ideal arrangement for solid-state diffusion, the impurity partial pressure should remain constant and independent of diffusion time. Furthermore, the partial pressure should not be a strong function of temperature. Transport should be accomplished in such a manner that all areas of the wafer surface would be subject to equal impurity-flow densities. A close approximation of such an ideal arrangement would be to place an inert solid source containing the desired impurity in intimate contact with the silicon and diffuse from this solid at the same temperature.

The solid-to-solid diffusion process described employs an adaptation of the technique described by E. L. Jordan[4] for the deposition of silicon dioxide films. Jordan describes the deposition of silicon dioxide by the pyrolytic decomposition of an alkoxysilane such as ethyl silicate. In the present process, simultaneous pyrolysis of a metal–organic compound containing the desired impurity is used to deposit an oxide containing this impurity on a silicon wafer.[5] Subsequent heating of the coated wafer results in diffusion of the impurity into the silicon. The surface concentration is then determined primarily by the concentration (impurity level) and diffusion coefficient of the impurity in the oxide. A single high-temperature diffusion step is employed to obtain this junction.

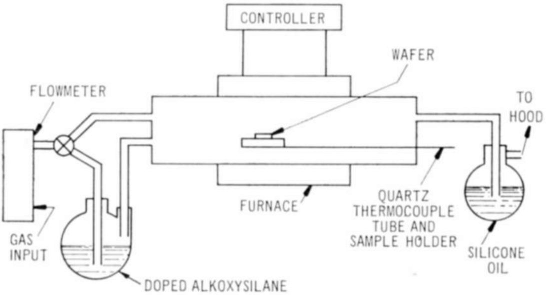


Fig. 2—Apparatus used for deposition of silicon dioxide films.

The concentration of impurity in the doped oxide depends on the partial pressure of the metal–organic compound during the deposition step. If the dopant is dissolved in the alkoxysilane, the partial pressure P_A of the dopant for an ideal system is given by

$$P_A = X_a P_a^0$$

where P_a^0 is the vapor pressure of the pure substance and X_a is the mole fraction of the dopant A. In a nonideal solution, the interaction of solute and solvent causes deviations from this relationship. In any case, there is a unique dependence of vapor pressure on the mole fraction, i.e., the proportion in the mixture determines the surface concentration.

Experimental Method

Figure 2 shows the apparatus used to accomplish the depositions. The carrier gas is saturated by passage through the bubbler containing

---

[4] E. L. Jordan, "A Diffusion Mask for Germanium," *Jour. Electrochem. Soc.*, Vol. 108, p. 478, May 1961.

[5] J. Scott and J. Olmstead, U. S. Patent 3,200,019.

the mixture of organo-oxysilane and the organic dopant. Pyrolysis of the gas in the hot zone of the furnace leads to the deposition of an impurity-containing silicon dioxide on the wafer. Because of the relatively low temperatures (approximately 750°C) and short times (approximately 10 minutes) involved, there is little or no diffusion into the silicon during the deposition step.

The diffusion is accomplished as described above and as shown in Figure 3. In step 1, the silicon wafer is cleaned and placed in the furnace. In step 2, the "doped" oxide is deposited. In step 3, the wafer is heated at the diffusion temperature for the desired time to obtain the required depth.

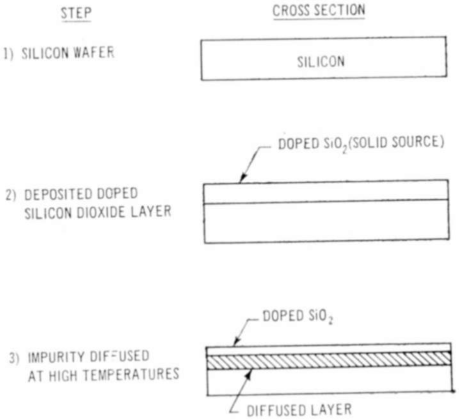


Fig. 3—Steps used in solid-to-solid diffusion process.

With this technique a wide variety of surface concentrations can be obtained easily once variables such as surface concentration as a function of impurity percentage in the bubbler have been established.

EXPERIMENTAL DATA

The following data, which were generated for boron diffusion using trimethyl borate as the organic impurity, demonstrate the control available with this technique. Figure 4 shows the relationship between diffusion depth and the square root of time. This curve demonstrates the same time dependence observed with existing techniques.

Figure 5 shows surface concentration at various diffusion temperatures as a function of the percentage of trimethyl borate in the oxysilane solution for a 750°C, 10-minute deposition. It can be seen that the surface concentration can be controlled easily over two orders of

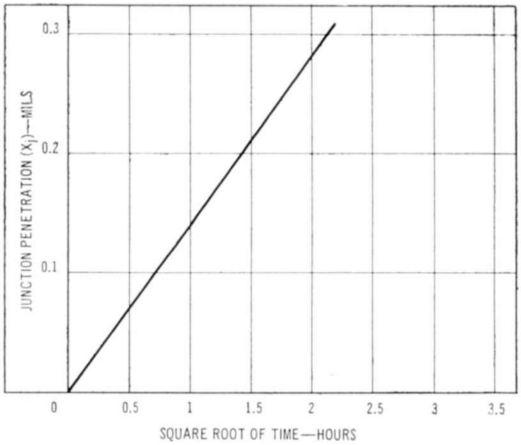


Fig. 4—Diffusion depth as a function of the square root of time for boron diffusion using trimethyl borate.

magnitude. Figure 6 shows the same type of data for a low-vapor-pressure organic tripropyl borate. The lower vapor pressure provides the advantage of a more constant mixture that does not deplete as easily.

Figure 7 shows surface concentration as a function of diffusion time. It can be seen that depletion occurs more rapidly with a diffusion source obtained by a one-minute deposition than with the thicker layer

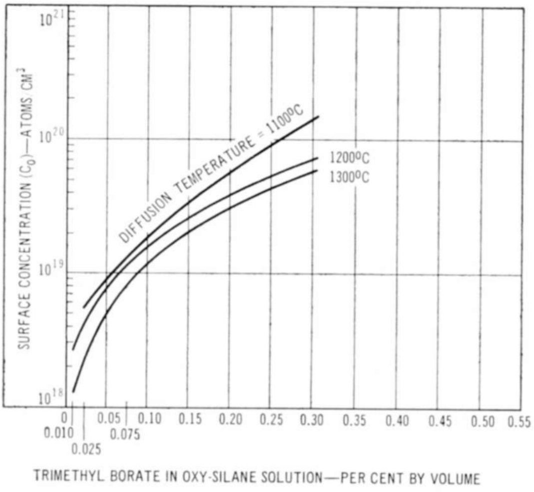


Fig. 5—Surface concentration as a function of solution concentration for three diffusion temperatures (750°C, 10-minute deposition).

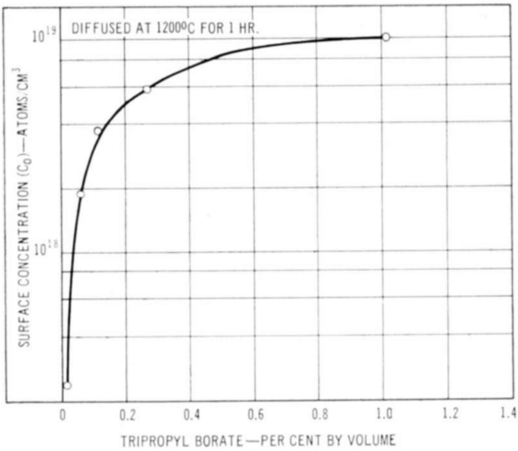


Fig. 6—Surface concentration as a function of solution concentration for low-vapor-pressure organic tripropyl borate.

of oxide obtained from a ten-minute deposition. The thicker oxide is a good approximation to an infinite source for the diffusion times indicated. The data shown in this figure and in Figure 4 support the assumption of a complementary-error-function type of impurity distribution. The curves of surface concentration as a function of time for three different diffusion temperatures in Figure 8 also demonstrate that the oxide source depletes very slowly with time and for most diffusions may be considered an infinite source.

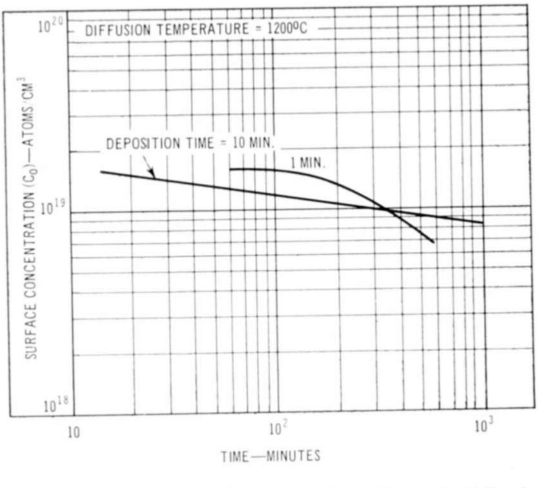


Fig. 7—Surface concentration as a function of diffusion time.

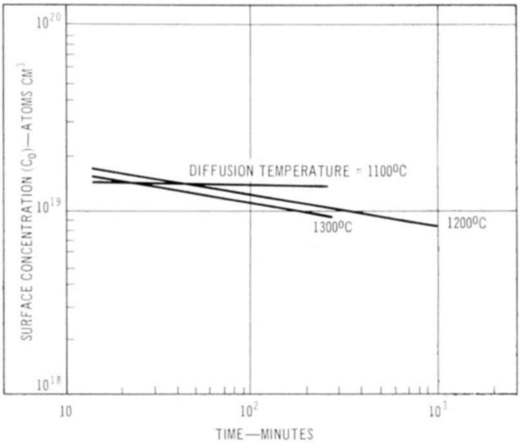


Fig. 8—Surface concentration as a function of diffusion time for three diffusion temperatures.

The middle curve of Figure 8 is replotted in Figure 9 for two types of diffusion ambient gas. The ambient gas can be used to obtain some further control of surface concentration. The differences in the diffusion process for different gases can be explained by the chemistry of the reduction of the impurity oxide. Most reported systems employ oxides of the impurities as the transporting species. These oxides must then be reduced to supply the silicon with the elemental impurity. This reduction occurs somewhere at the Si-SiO$_2$ interface.

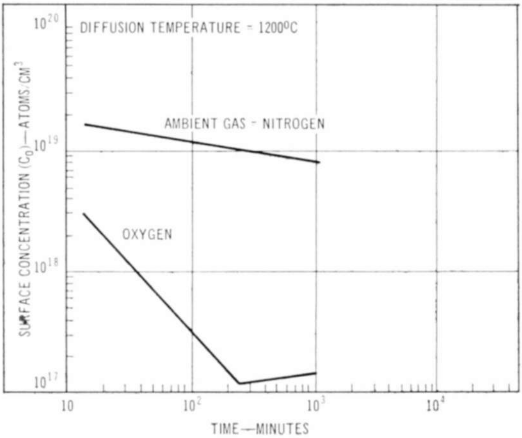


Fig. 9—Surface concentration as a function of diffusion time for two types of ambient gas.

If R represents a chemical impurity and RO its oxide, the reduction process proceeds as $2RO + Si \rightleftharpoons SiO_2 + 2R$. The dissociation reaction, $2RO \rightleftharpoons 2R + O_2$, is ruled out by a consideration of free-energy change for the reaction. While the first reaction is proceeding, the silicon also oxidizes by the reaction $Si + O_2 \rightleftharpoons SiO_2$.

If the oxygen concentration is decreased, the third reaction shifts to the left. The increase in concentration of silicon then shifts the first reaction to the right and thereby increases the concentration of the elemental impurity available for diffusion. In addition, as the concentration of oxygen in the ambient increases, oxide formation occurs at the interface of the doped oxide and the silicon, and the concentration at the surface is further reduced. Therefore, the ambient may have control over both surface concentrations, C_0, and the concentration profiles obtained. Figure 10 illustrates the control of surface-conductivity parameters and reproducibility obtained utilizing this technique.

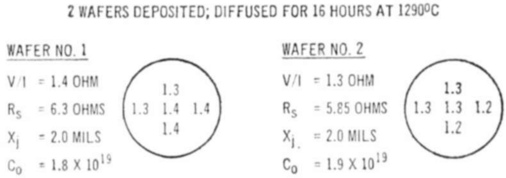


Fig. 10—Results obtained on two wafers prepared by new solid-to-solid diffusion technique.

A system for n-type diffusions can be obtained in a similar manner by use of trimethyl phosphate in place of the trimethyl borate used for the p-type diffusion. As in the case of boron, the surface concentration is not a strong function of deposition time. A standard deposition time of 10 minutes and a temperature of 750°C were chosen for the following data.

Figure 11 shows surface concentration as a function of the percentage of trimethyl phosphate at temperatures of 1100, 1200, and 1300°C. These curves, which represent drive-in times of both 1 and 4 hours, show some dependence on time of diffusion. Figure 12 shows a slight dependence of surface concentration on diffusion temperature for various dilutions. Figure 13 shows that, as in the case of boron, an adjustment of the ambient gas produces a change in the surface concentration obtained.

The diffusion technique has one other major advantage in addition to the control of surface concentration and reproducibility. Placing

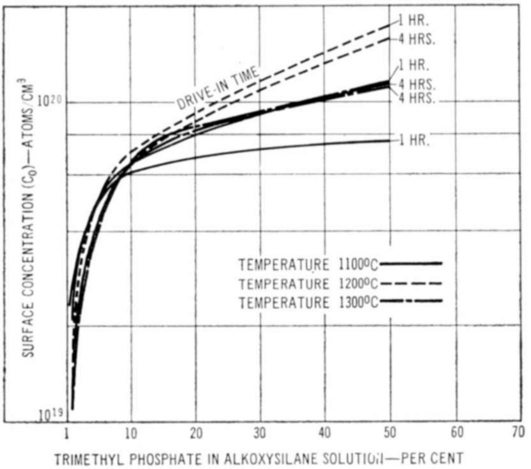


Fig. 11—Surface concentration as a function of solution concentration for n-type diffusions at three different temperatures.

the diffusion source directly on the surface of the semiconductor permits great versatility in selective solid-state diffusion. The versatility of the new system is best shown by a comparison with conventional diffusion methods. Figure 14 outlines the conventional sequence: growth of a silicon dioxide mask, etching of openings by photolithographic techniques, and then diffusion of impurities through the holes to form the p-n junction.

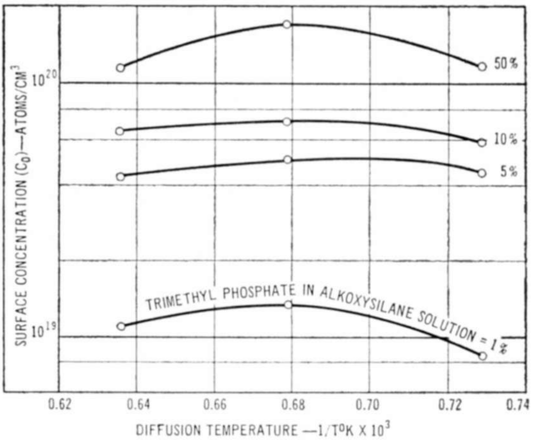


Fig. 12—Surface concentration as a function of diffusion temperature for various solution concentrations.

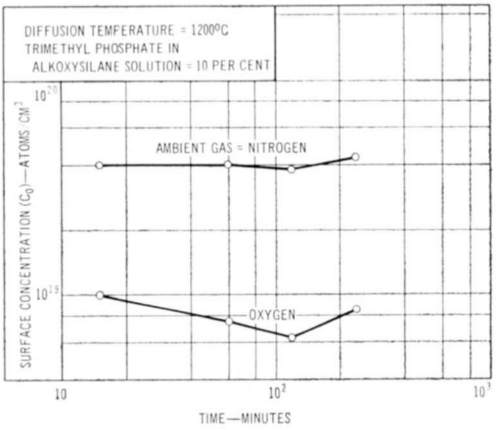


Fig. 13—Surface concentration as a function of diffusion time for two types of ambient gas.

In contrast, the use of a solid source provides a choice of sequences, as shown in Figure 15. The operation can be performed conventionally, or the procedure can be reversed so that the "solid source" is deposited over the wafer and then removed by photolithographic techniques from the areas where diffusion is not desired before the wafer is exposed to diffusion temperatures.

Because of this new flexibility, simultaneous diffusions of both p- and n-type impurities are possible. Simultaneous diffusions of p, p$^+$, n, n$^+$, and various other combinations have been carried out. Combination of this technique with existing diffusion methods makes possible

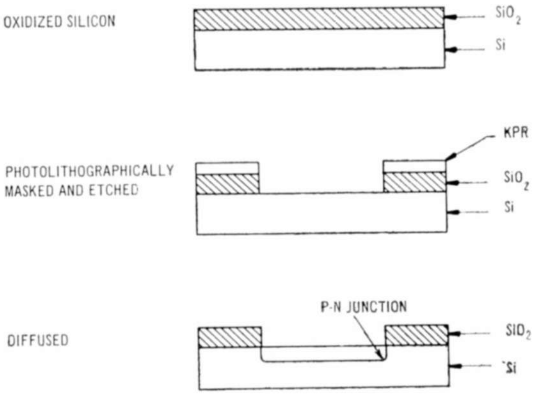


Fig. 14—Steps in conventional diffusion sequence.

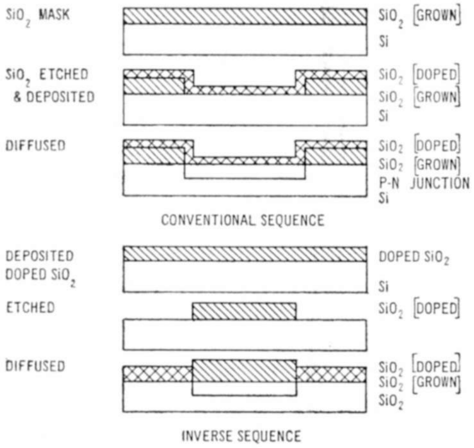


Fig. 15—Steps in solid-to-solid diffusion technique.

the fabrication of many complex structures that could not previously be made without the use of numerous high-temperature diffusions.

ACKNOWLEDGMENT

The authors wish to acknowledge the helpful suggestions and additional data supplied by W. J. Greig, J. A. Nava, and L. Krassner.

A CASSEGRAINIAN FEED FOR WIDE-BAND SATELLITE COMMUNICATIONS

BY

P. FOLDES, S. KOMLOS, N. K. CHITRE, R. SCHWERDTFEGER, AND T. SZIRTES

RCA Victor Company, Ltd.,
Montreal, Canada

Summary—*Canada's first satellite communication ground station is located at Mill Village, Nova Scotia. The antenna is an 85-foot paraboloid with a new type of Cassegrainian feed system. The antenna covers a wide frequency range and has adjustable polarization, high efficiency, low noise temperature, and large power-handling capability. Automatic tracking in both direction and polarization is possible. The beam width of the feed is about the same in nearly all of the receiving frequency band (3700–4200 mc) and the transmitting band (5925–6425 mc). The antenna uses a multi-mode multi-horn source aperture that illuminates the hyperboloid–paraboloid reflector system with an optimum field distribution for both sum and difference modes. The low loss (about 0.25 db in the complete duplexer multi-mode coupler mode-filter system) and high quality of overall match ($VSWR \cong 1.25$) is achieved by the use of high precision ridged waveguide elements. The main feature of the antenna is the precisely controlled primary pattern, which assures very low spillover radiation around the hyperboloid. As a consequence, the overall antenna efficiency is about −2 db in the receiver band, and about −3 db in the transmitter band.*

INTRODUCTION

IN RECENT years, considerable effort has gone into the study of ground antennas for satellite communications systems. The basic problem is the provision of an antenna with autotrack capability and high gain/noise-temperature ratio at low cost. Since operation with several different satellite systems may be desirable, the antenna should be capable of operation over a wide frequency band and with various polarizations.

Analysis of existing and planned wide-band communication satellite systems indicates that the ground antenna should have a gain/system-noise-temperature ratio of at least 40 decibels for elevation angles of 5° or more (system noise is here expressed in decibels relative to 1°K). Because of inherent limitations on system noise temperature in the 4-gigacycle band, where the down-link frequency of presently planned satellite systems lies, an antenna gain of nearly 59 decibels for the received signal is required. Several approaches to the problem have been described in the literature.

The installation at Andover, Maine uses a horn reflector.[1,2] In other installations, the electrical advantages of the horn reflector have been combined with the structural simplicity of a symmetrical Cassegrainian arrangement.[3-5] In these installations, the horn reflector serves as the feed system, thereby assuring low spillover radiation. The possibility of still further structural simplification was shown by P. D. Potter[6] and P. A. Jensen,[7] who utilized higher order wave-guide modes in a relatively small source aperture in order to control the secondary pattern and obtain high gain/noise-temperature ratios. Foldes and Komlos[8] showed that optimum performance can be extended into difference modes and into a wide frequency band by using a multimode-multihorn source, resulting in an arrangement somewhat similar to Ricardi's 12-horn system.[9]

The present paper is a continuation of the work described in Reference (8). The design of the antenna was based on the following requirements:

(a) optimization of the performance-to-cost ratio for the overall antenna complex, where performance is characterized by the ratio of antenna gain to noise temperature;

(b) use of a minimum number of microwave component types;

(c) provision of large flexibility in operational modes, and coverage of wide frequency band in order to accommodate the feed for the possible change of frequency requirements;

(d) assurance of low circuit loss and difference mode minimum depth by the use of highly symmetrical waveguide cross sections;

(e) good aperture field distribution for the main reflector in the

---

[1] J. S. Cook and R. Lowell, "The Autotrack System," *Bell System Tech. Jour.*, Vol. 42, Part 2, p. 1283, July 1963.

[2] A. J. Giger, S. Pardee, Jr., and P. R. Wickliffe, Jr., "The Ground Transmitter and Receiver," *Bell System Tech. Jour.*, Vol. 42, Part 1, p. 1063, July 1963.

[3] C. E. Profera and A. F. Sciambi, "A High Efficiency Low-Noise Antenna Feed System," *IEEE PTGAP International Symposium on Space Communication*, Boulder, Colo., p. 83, July 1963.

[4] G. B. Von Trentini, "Erregersysteme für Cassegrain—Antennen," *Frequenz*, p. 491 (Sonderausgabe Bd 17/1963).

[5] S. P. Morgan, "Some Examples of Generalized Cassegrainian and Gregorian Antennas," *Trans. IEEE PTGAP*, Vol. AP-12, p. 685, Nov. 1964.

[6] P. D. Potter, "A New Horn Antenna with Suppressed Sidelobes and Equal Beamwidths," *Microwave Jour.*, Vol. 6, p. 71, June 1963.

[7] P. A. Jensen, "A Low-Noise Multimode Cassegrain Monopulse Feed with Polarization Diversity," *NEREM Record*, p. 94, 1963.

[8] P. Foldes and S. G. Komlos, "A New Multimode Monopulse Feed," *NEREM Record*, p. 100, 1963.

[9] L. J. Ricardi and L. Niro, "Design of a Twelve-Horn Monopulse Feed," *IRE International Convention Record*, Part 1, Vol. 9, p. 49, 1961.

receiver frequency band (3700–4200 mc) while maintaining at least –3 db efficiency in the transmitter (5925–6425 mc) frequency band (this requirement gives more importance to the receiving operation, which is the basic system limitation in any satellite communication ground station);

(f) provision of freedom for the field alignment of the feed relative to the main reflector;

(g) provision of both circular and orientable linear polarization (in changing the plane of polarization, movement or rotation of large elements is to be avoided).

DESCRIPTION OF OVERALL SYSTEM

Requirement (a) above virtually dictates the use of some form of Cassegrainian system. In this type of radio optics, the radiating source first illuminates a primary reflector, which in turn reflects the wave toward a secondary reflector. In the simplest form of Cassegrainian system, the radiating source aperture is small and the subreflector is in its far field. Then the subreflector is a symmetrical hyperboloid and the secondary reflector is a symmetrical paraboloid. Mechanical symmetry results in the simplest construction and minimum weight, and therefore was used as the basis of the antenna design. An 85-foot-diameter solid paraboloid with a focal distance of 36 feet and r-m-s surface tolerance (half-path-length error) of 0.040 inch is used for the main reflector. This type of paraboloid is relatively easily fabricated and has known mechanical characteristics. It represents a good compromise among weight/aperture-area ratio, rigidity, and weight of feed-support structure.

To illuminate such a relatively flat paraboloid with low spillover requires an illuminating aperture of 10 or more wavelengths. Therefore, an 8.5-foot-diameter hyperboloid subreflector with a focal distance of 18 feet is used. Such a subreflector is large enough to virtually eliminate spillover radiation around the edge of the paraboloid; the spillover radiation problem is thus transferred to the region around the edge of the subreflector.

The control of this spillover radiation and, more generally, the control of the illumination provided by the source system for the hyperboloid, is the central problem of any Cassegrainian antenna and is discussed later in more detail. The control of the illumination is dependent on proper field distribution in the source aperture. The purpose of the source system is to provide this distribution at its aperture where the guided wave (in the source) is launched into a freely propagating wave.

Figure 1 is a simplified block diagram of the radiating source system when it is operated in a linearly polarized mode that may be oriented in any plane. In principle, this system is relatively simple and, with the exception of the mode filter, does not contain any radically new elements. The transmitter (Tx) terminal (WR-159) accepts high power (up to 10 kilowatts CW) in the 5925 to 6425 megacycle band.

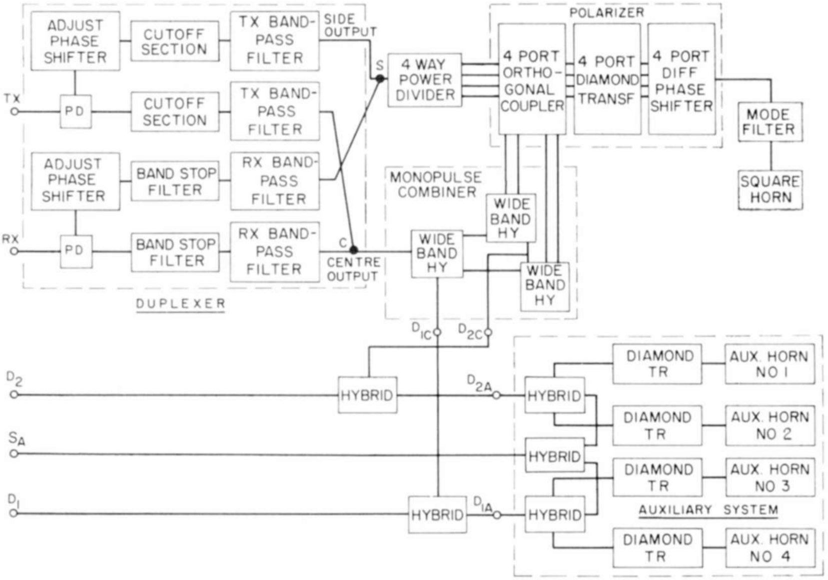


Fig. 1—Block diagram of feed system.

This power is divided into two equal parts. Each half of the power travels through a cutoff section and a band-pass filter that provide more than 125 db attenuation below 4200 megacycles and have less than 0.13 db attenuation for the Tx frequencies. In addition to these filters, the arm connected to the side output of the source system has an adjustable phase shifter using a short-slot hybrid with noncontacting plungers. The outputs of the Tx band-pass filters are connected to the receiver (Rx) bandpass filters at the junctions (S) and (C), which form the duplexing points, with negligible junction effects. Beyond these junction points, both the center and side-arm powers of the transmitter are divided into four equal parts by the use of identical wide-band hybrids. (The four-way power division is necessary to introduce monopulse tracking capability for the 3700 to 4200 megacycle receiving band). After these manipulations, the center and side-arm powers are recombined in a four-port orthogonal coupler that has four square-waveguide outputs. Each of these cross sections support the TE_{10} and TE_{01} modes corresponding to the center and side outputs of the system. Each of these modes is decomposed in the following symmetrical diamond transformers and converted into a left and right circularly polarized wave by the 90° differential phase shifter that terminates the four-port circuit. Up to the output of this phase shifter, the system can be considered as a wideband four-horn system that radiates linear polarization where the plane of polarization can be oriented by the adjustable phase shifter. If the power divider following the Tx terminal is left out and the Tx terminal is directly connected to the center or side arm of the duplexer, then the system becomes left or right circularly polarized.

The waveguides between the four-port orthogonal coupler and terminals C and S have rectangular, double-ridged cross sections. These assure low loss in the Rx frequency band and the suppression of higher order modes at Tx frequencies. The wide-band folded hybrids have a Gaussian side-wall taper that results in the unique feature of better than —38 db reflection coefficient for the sum terminal and —46 db isolation for the difference terminal. The difference-mode terminal is matched only for Rx frequencies.

The waveguide cross sections between the orthogonal coupler and differential phase shifter (polarizer assembly) are square with ridge loading on each side. The orthogonal coupler is a wide-band version of a component described elsewhere.[10] The differential phase shifter is a waveguide section loaded by 11 irises that have a binomially tapered penetration. All the above-mentioned components have individual reflection coefficients of —37 db or better in the Rx and Tx frequency bands.

The four-port polarizer launches four individual TE_{10} modes into the mode filter, which is an oversized square-waveguide section with various loading elements. The mode filter has two basic purposes: (1) to match the four-port circuit to the multimode horn and (2) to produce a symmetrically tapered field distribution in the aperture of this horn. It should be emphasized that the matching conditions must be simultaneously fulfilled for the communication (sum) and tracking (difference) modes in their corresponding frequency bands. Furthermore, the matching elements must not destroy the axial ratio and main to cross-polarized difference-mode ratio. Additionally, the pattern-

---

[10] S. G. Komlos, P. Foldes, and K. Jasinski, "Feed System for Clockwise and Counterclockwise Circular Polarization," *Trans. IRE PGAP*, Vol. AP-9, p. 577, Nov. 1961.

shaping elements must work in such a way that the width of the pattern in the Tx band is about equal to that of the Rx band, although the Tx band is 60% higher in frequency. With the exception of some basic theoretical considerations regarding the cross section and length of the mode-filter section, the details of this element were determined by cut and try procedures. The end result was an oversized square-waveguide cross section capable of supporting the TE_{10}, $TE_{11} + TM_{11}$, TE_{20}, $TE_{21} + TM_{21}$, and TE_{30} modes. This waveguide, approximately 2λ long, has a series of crosses at the input end for matching the sum and difference modes and a series of posts at the output end for pattern shaping. The mode filter is followed by a square horn that has about a 5λ aperture size in the transmitting band.

The operation of the system for the received frequencies is basically identical to that for the Tx operation, except that the direction of propagation is opposite and that band-stop filters are used in the receive arm of the duplexer instead of the cutoff waveguide sections of the transmit arm.

The difference-mode operation of the system is somewhat unconventional, mainly because it combines two independent monopulse systems. One monopulse system is realized in the center multimode system and is obtained through the three wide-band hybrids of the monopulse combiner. This system is quite effective, even though the difference-mode aperture is only 1/2 the optimum value.[11] This limitation is removed by the addition of the auxiliary system, which is a standard four-horn monopulse system consisting of four circularly polarized rectangular horns located symmetrically around the multimode square center horn. The horns are diagonally fed, and circular polarization is achieved by the selection of horn length and aperture aspect ratio. The auxiliary and center systems together assure an optimum difference mode, i.e., they provide a spillover efficiency almost as good as for the communication mode, and thus provide good protection against ground reflections at low elevation angles. The combined difference modes (D_1 and D_2) also have the advantage of lower cross coupling than that which exists in the center monopulse system alone, and nearly constant difference-mode slopes with varying attitudes of incoming linear polarization.

Figure 2 is a photograph of the complete radiating source system prior to its installation into the feed-positioning mechanism and feed-housing cone. Figure 3 shows the hyperboloid subreflector positioning mechanism.

---

[11] P. W. Hannan and P. A. Loth, "A Monopulse Antenna Having Independent Optimization of the Sum and Difference Modes," *IRE International Convention Record,* Part 1, Vol. 9, p. 57, 1961.

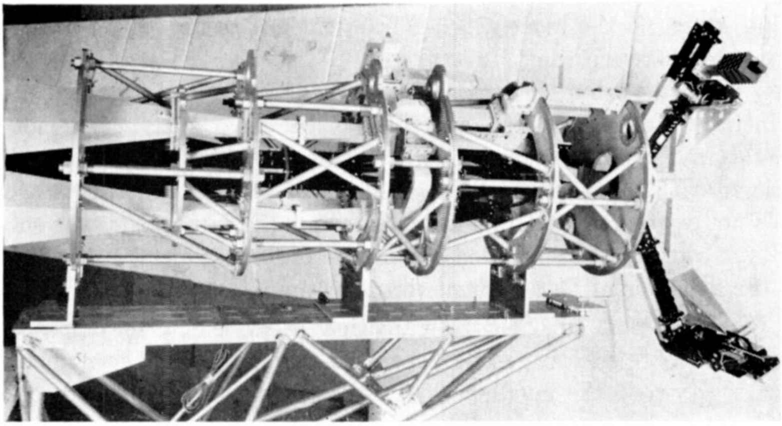


Fig. 2—Photograph of complete feed system.

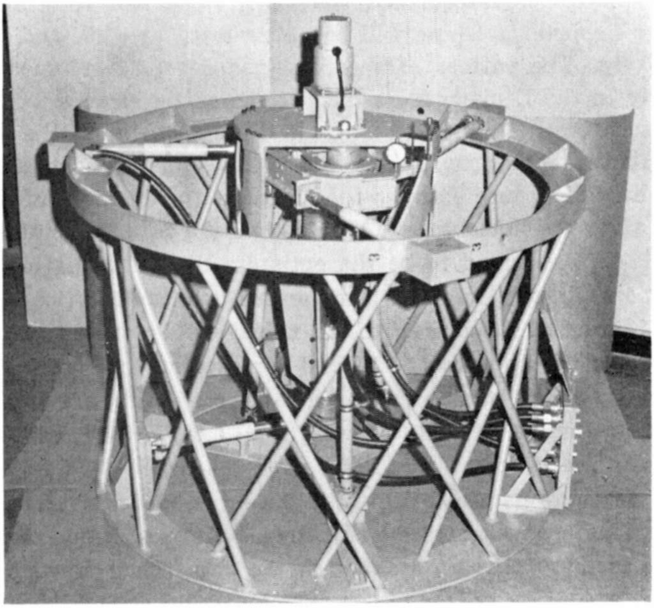


Fig. 3—Photograph of hyperboloid adjusting and mounting mechanism.

MODE-FILTER DESIGN

Electromagnetically, the most interesting component in the radiating source is the mode filter. Ideally it is a device that provides a transition between the four incident TE_{10} modes and the desired complex mode distribution in the aperture of the multimode horn. In principle, the field in the aperture of the multimode horn can be synthesized by the proper selection of the phase between the incident elementary TE_{10} modes and the geometry of the mode filter. Although it is not suggested that this complicated synthesis problem was solved or even completely understood during the present project, a description of the behavior of this element can be given.

The four-port to one-port transition in its basic, uncompensated form is a very effective mode generator, supplying a large number of modes. Even if the single-port (oversized) square-waveguide section is in cutoff for the $m + n \geqq 3$ modes, some of these modes reach the horn, because a relatively short mode filter is required for wideband operation. This attenuation cannot be increased by the reduction of mode-filter cross section because the impedance-matching problem becomes unmanageable.

For sum-mode performance, the TE_{10}, TE_{21}, and TE_{30} modes are the most important. These will be abbreviated as α, δ, and ξ modes, respectively. The pattern of the combined $\alpha + (A \angle \phi)\, \delta$ mode differs primarily in the E plane from the pattern of the pure α mode, while the pattern of the $\alpha + (B \angle \phi)\, \xi$ mode differs mostly in the H plane. In the above expressions, A and B are the amplitude ratios, while ϕ is the phase difference between the modes in the aperture of the horn. Table I gives the sidelobe level, κ, in db and the 9-db width of the main beam (9 db was found to be the optimum edge illumination for the hyperboloid) in relative units for various amplitude ratios and mode phase differences. According to the table, an E-plane beam width to H-plane beam width ratio of 0.72 and E-plane side-lobe level of 13.2 db (with $A = B = 0$) can be improved to a beam-width ratio of 0.9 and E-plane side-lobe level of 19 db with $A = 0.6$, $B = 0.2$, and $20° \leqq \phi \leqq 60°$. As the propagation velocities of the δ and ξ modes are very close, it seems to be possible to keep these two modes nearly in phase. Then, the major problem is the adjustment of A and B and the phasing of the α mode relative to δ and ξ in the operational frequency band.

The adjustment of A, B, and ϕ is relatively easy experimentally when the frequency band is narrow (about 6 to 8%) and side-lobe level requirement is only 20 db. Double-frequency-band operation is somewhat more difficult, but still manageable when the VSWR requirement is not very stringent. However, it is quite difficult to obtain high-quality patterns and match in double the frequency band.

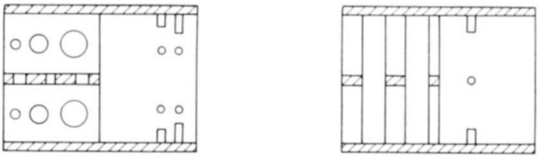
Figure 4 indicates the geometry of some of the most successful mode-filter configurations. Each of these arrangements has special features that are advantageous for certain applications. The arrangement shown in Figure 4(a) results in very high quality patterns in about 10% frequency band by the fine adjustment of the coupling holes between the four individual waveguides. However, the matching of this coupling system requires additional posts or irises.

Table I—Pattern Characteristics of the $\alpha + (A\angle\phi)\delta$ and $\alpha + (B\angle\phi)\xi$ Modes

$\alpha + (A\angle\phi)\delta$	E plane							
$A \backslash \phi^\circ$	0		20		40		60	
	κ (decibels)	θ_{9db}	κ (decibels)	θ_{9db}	κ (decibels)	θ_{9db}	κ (decibels)	θ_{9db}
0	13.2	.72						
.2	17.3	.78	16.8	.78	15.6	.78	14.5	.76
.4	23.5	.84	21.2	.82	17.2	.84	14.0	.79
.6	28.6	.90	25.5	.92	21.6	.92	19.0	.89
.8	37.5	1.06	37.5	1.06	21.2	1.08	18.0	1.20

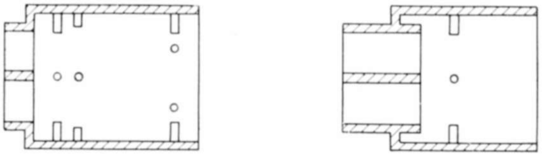
$\alpha + (B\angle\phi)\xi$	H plane							
$B \backslash \phi^\circ$	0		20		40		60	
	κ (decibels)	θ_{9db}	κ (decibels)	θ_{9db}	κ (decibels)	θ_{9db}	κ (decibels)	θ_{9db}
0	23.0	.99						
.2	27.0	1.04	27.5	1.02	25.6	1.00	24.0	1.00
.4	36.0	1.10	35.5	1.08	33.5	1.04	32.0	1.02
.6	40	1.18	38.6	1.18	34.0	1.14	31.1	1.10
.8	40	1.40	38.5	1.25	33.0	1.23	30.2	1.16

The cross arrangement, Figure 4(b), is quite good for double-frequency-band operation, and has the advantage that cross separation controls the match and pattern of the Tx band, while the posts conveniently shape the Rx pattern. The stepped transition of the arrangement in Figure 4(c) is a convenient way to balance the δ and ξ modes, and thus improve impedance behavior at the high-frequency end. It also has been a good matching characteristic for the difference modes. Finally, in the arrangement in Figure 4(d) (which was used in the present installation) a choke is connected parallel with the 4-port to 1-port transition as a means of obtaining a very high quality wideband match.



a
TRANSITION WITH
COUPLING HOLES

b
TRANSITION WITH
COUPLING CROSSES



c
TRANSITION WITH
STEPS AND POSTS

d
TRANSITION WITH
CHOKE

Fig. 4—Various forms of mode filters.

Figure 5 shows the average 9-db beamwidth (θ_{9dbav}), average first side-lobe level (κ_{av}) and beam-width ratio $\theta_{9dbE}/\theta_{9dbH}$ with the arrangement in Figure 4(b) in double-frequency-band operation. For comparison, the figure shows the same characteristics for an identical-size square horn excited by pure TE_{10} mode. It can be seen, that the multimode horn has less elipticity in the main-beam cross section and lower side-lobe level. Both of these features improve the gain/noise-temperature of the overall antenna system.

The posts indicated in some of the mode-filter configurations are very convenient elements in shaping the sum-mode pattern, but they tend to increase the cross-polarized component in the difference modes. On that basis their penetration was restricted to assure at least 10-db cross coupling between the two difference channels.

HIGH-POWER LOW-LOSS DUPLEXER

Another special area of the feed system where some new components **had** to be developed is the duplexer. The primary function of the duplexer is to permit the use of a single antenna for simultaneous transmission at +70 dbm cw in the 6175 to 6425 megacycle band (F_{tb}) and reception of a —100 dbm signal in the 3960 to 4200 megacycle band (F_{rb}). The secondary function is to provide the selection of right or left circular polarization, or a linear polarization with any orientation.

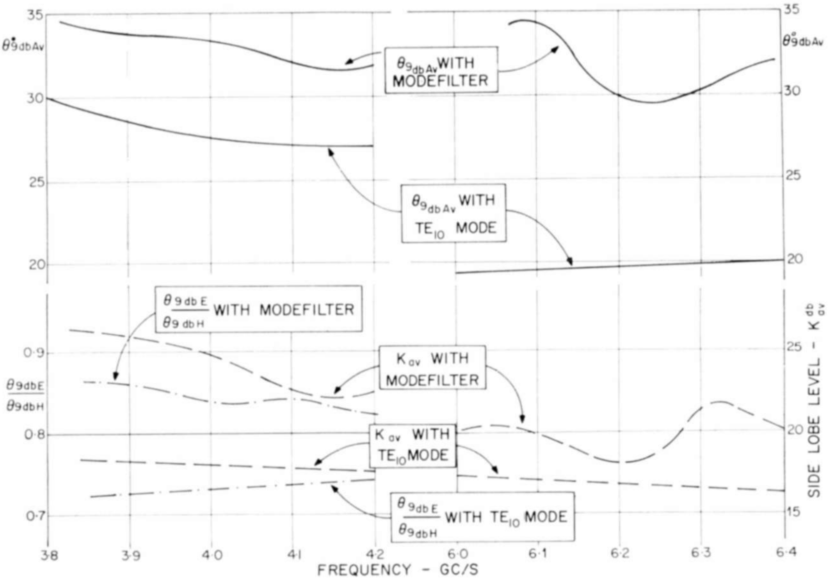


Fig. 5—Main characteristics of the source pattern for TE$_{10}$ mode and for modes achieved by mode filter.

The indicated frequency bands are determined by the duplexer filters and could be shifted by substituting differently tuned filters. Figure 6 shows the assembly drawing of the duplexer, where the symbols Rx, Tx, and A indicate the receiver, transmitter, and antenna terminals, respectively. The receiver arm contains a band-pass filter for the receiver channel and a band-stop filter to provide extra isolation at the transmit frequency. Both of these filters are designed in a single ridged waveguide (SR175), and a transformer is provided to match this arm to the standard WR229 waveguide. The transmitter arm consists of a band-pass filter for the transmitter channel and a cutoff section to enhance the isolation in the receiver channel.

Since an extremely low noise temperature is required in the receiver band, the receive arm must provide a very low loss path for the received signals so as to limit the duplexer's contribution to the system noise temperature. On the other hand, due to the extremely large difference (>170 db) in the levels of the transmitted and received signals, the duplexer must provide isolation between the Tx and Rx terminals greater than 185 db at the transmitter frequency and 100 db elsewhere, so as to limit the deterioration of receiver noise caused by interfering signals. Under these conditions the design must result in minimum noise for the receiver, whether it comes from thermal or intermodulation sources.

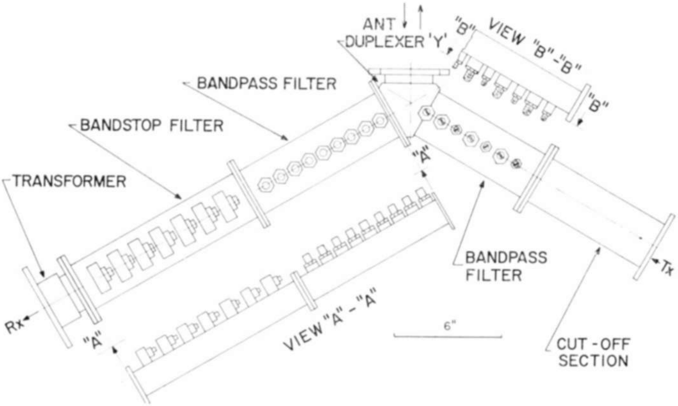


Fig. 6—Assembly drawing of duplexer.

Electromagnetically, the design must provide the largest possible slope for the transfer curve between the Rx and Tx terminals and between the highest receive and lowest transmit frequencies. A standard band-pass filter by itself cannot be expected to provide the largest ratio of isolation to pass-band loss, since it tends to have fairly large losses in the pass band and spurious responses at about 1.5 times the pass-band frequency. A band-stop filter alone would not be able to handle the high transmitter power (voltage limitation in the first cavity) and would provide isolation over only a narrow band. The optimum arrangement was found to be a band-pass filter providing more than 85 db isolation throughout the 5925 to 6425 megacycle transmit band and a band-stop filter to enhance this by another 100 db over a 70-megacycle band centered on the transmitter frequency. The total isolation of the filter combination is more than 100 db over the whole 500-megacycle transmit frequency band. For application where additional isolation is provided between the low-noise front and the mixer of the main receiver, this isolation is adequate for the whole transmitter band. When such isolation is not provided, the band-stop filter must be changed whenever the transmit frequency is changed.

The total receiver channel loss in the duplexer is kept below 0.14 db, the major contributor being the band-pass filter (≈ 0.09 db) with the band-stop filter (≈ 0.04 db), the transformer, and the Y junction contributing the remainder. This represents a noise-temperature contribution of about 9.5°K. When the band-stop filter is removed, the noise-temperature contribution drops to 6.5°K.

In the transmitter arm of the duplexer, 125 db isolation is provided between the transmitter and receiver for the 4-gigacycle spurious signals originating in the high-power amplifier. The loss in the transmitter band is kept to a minimum (0.1 db) to prevent overheating of the duplexer due to absorption of the transmitter power.

A band-pass filter was included in the transmitter arm because it provides the possibility of tuning out reflections in the duplexer Y, more isolation at 4 gigacycles for a given pass-band loss, smaller physical size, and a certain amount of protection against the modulation products on the high-frequency side of the transmitter frequency band. It also provides an extremely well-defined short circuit at the receiver frequency. However, to provide 125 db isolation, a six- or seven-cavity design would have been needed; the compromise arrived at to reduce the complexity consists of a three-cavity band-pass filter contributing about 65 db isolation and a cutoff section adding another 60 db.

Duplexer Y

The duplexer Y is a symmetrical shunt junction that accepts a double-ridged waveguide (DR175) on the antenna port and has built-in transformers to single ridged waveguide (SR175) on the receiver arm and to rectangular waveguide (WR175) on the transmitter arm. The Tx band-pass filter is soldered to the Y, while the Rx band-pass filter is attached by a flange.

Transmitter Band-Pass Filter

The Tx band-pass filter is a three-cavity direct-coupled design in which TM_{01}-mode cylindrical cavities are used to minimize transmission loss in the transmit band and reflection loss in the receive band. The cross section of the filter is made up of three intersecting circles; the dimensions of the diameters and the spacings of the centers of the circles are such as to give proper response. The edges at the aperture have a 0.062-inch radius to give good high-voltage breakdown properties. Tuning screws, used only for fine adjustment of the match, have very small penetrations, thus increasing the voltage-handling capability to nearly that of the undisturbed waveguide. In fact, power handling is limited by the heat production, and practical experience shows that no cooling fins are required for CW transmitter power output up to 10 kilowatts.

The Cutoff Section

The cutoff section is a six-inch length of special-cross-section rectangular waveguide (WR120) with transformers added on both ends. The width of the guide is chosen to maximize the ratio of isolation in the receive band to insertion loss in the transmit band in the TE_{10} mode. The height of the guide is the maximum permissible that will still prevent TE_{01}-mode propagation below 6425 megacycles. Two-step transformers adapt to WR175 on the Tx band-pass filter side and to WR159 on the output side. The transformers and cutoff waveguide are built in one piece.

The Receiver Band-Pass Filter

This unit is of the direct-coupled design in a single-ridged waveguide (SR175). The waveguide is optimized to provide minimum loss in the receive band without propagating the TE_{20} mode in the transmit band. The coupling posts are capacitive in the pass band and provide high isolation in the transmit band, where they are almost series resonant. The elements have no sharp corners, thus reducing the possibility of arcing at high power. Because the attenuation at transmit frequencies is more than 15 db per cavity, only the first cavity at the Y junction handles appreciable power.

Band-Stop Filter

The design consists of seven external cavities coupled to an SR175 waveguide through slots on the top wall, with the length of the slots controlling the stop-band width. The equal-cavity design gives a frequency response characterized by the Tchebychev functions of the second kind, i.e., maximum isolation for a given number of sections. The coupling lengths between cavities are approximately $3/4\ \lambda_g$ at f_T and are adjusted to give a reflection minimum in the receive band. Final matching in the receive band is achieved by small adjustments of tuning screws below each slot.

The maximum isolation of these filters is more than 135 db, too great to permit tuning and measurement on any conventional microwave bench. A special high-isolation swept-frequency setup has been developed for this purpose to plot isolations up to 135 db.

Twin Duplexer

In the actual twin-duplexer system, two of the above-described duplexers are connected in parallel on two arms of the Rx band and Tx band power-divider hybrids. In this arrangement the arm connected to the side system of the feed contains the adjustable phase shifter, while a phase-compensating section is inserted into the other arm. When the phase shifter is in the center position, the phase-compensating section assures 0° (vertical) polarization at midfrequency. In this phase-shifter position, the polarization attitude may differ by ±3° from the vertical at the limits of the frequency band. When the phase shifter is at its end positions, the polarization is ±90° at midfrequency and is within 6° relative to this at band limits. Attitude error can be reduced by a more refined phase-compensating section, but this was not required for the present application. The attitude error was taken into account by calibration accurate to ±0.2°.

Table II—Cone Structure Characteristics

Overall Length	218 inches
Diameter at Base	96.1 inches
Diameter at Top	51.4 inches
Included Angle	11° 42′ 28″
Weight of Skin	1520 lbs.
Total Weight (including equipment)	4500 lbs.
Maximum Lateral Deflection at the Tip	0.025 inch
Lowest Natural Frequency Under Full Load (calculated)	50 cps

MECHANICAL CONSIDERATIONS

The feed system is made up of three somewhat independent subsystems, namely the microwave circuit, the feed cone, and the hyperboloid.

The microwave circuit consists of the radiating source and the duplexer. These units are supported by independent frames and can be readily separated for maintenance.

The main function of the cone, which is fabricated of 3/16-inch welded aluminum sheets, is to house and support the feed frame assembly and its positioning mechanism. The cone also houses parametric amplifiers, monopulse receivers, tunnel-diode amplifiers, downconverters, preamplifiers, noise-measuring equipment, and the like. In addition it can accommodate two men for maintenance or testing of equipment. A 1600 cfm blower changes the air in the cone every 10 minutes. Physical characteristics of the cone are shown in Table II.

The feed-horn assembly is so mounted that it may be adjusted ±2.5 inches in a direction perpendicular to its axis, ±7 inches parallel to its axis, and the axis may be tilted ±3°. These adjustments are used to ensure that the axis and phase center of the horn assembly coincide with the axis and focus of the paraboloidal dish, respectively.

The second major unit in the Cassegrainian feed is the subreflector assembly. The hyperboloid reflector has a five-degree freedom of movement. It can be moved along and rotated about two mutually perpendicular axes, both parallel to the aperture plane, and it can also be moved along the axis of the subreflector. The principle of operation can be described by the use of Figure 7 where the point A represents the axis of the subreflector, which is perpendicular to the plane of the drawing, S. By suitable adjustment of the bar lengths, K_1, K_2, and K_3, any desired lateral and angular position of the subreflector may be obtained. Figure 8 is a sketch of the most important parts of the mechanism. A reversible electric motor provides the axial motion of the hyperboloid.

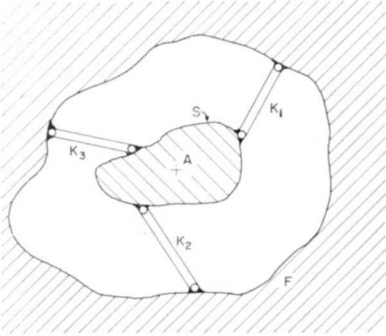


Fig. 7—Principle of hyperboloid adjusting mechanism.

MEASURING TECHNIQUES

The successful development of a feed system, which is relatively complicated in fine details, depends to a very large degree on the measuring techniques used. It is beyond the scope of the present paper to give a full description of these techniques, but some of the more interesting ones are mentioned briefly.

Reflection Coefficient

The swept-frequency method is used for the measurement of reflection coefficient of the wide-band components. The necessarily wide-band hybrid used as a directional coupler is a slightly modified version of the hybrid used in the monopulse combiner. This hybrid has better than —40 db isolation and better than —20 db reflection coefficient in the 3700 to 6425 megacycle frequency band for all terminals. For the measurement of reflection coefficient of the mode filter, the complete multimode coupler (monopulse combiner–polarizer assembly) was utilized. This was possible because the multimode coupler in itself was very well matched (input reflection coefficient better than —28 db when terminated by a four-port dummy load instead of the mode filter). Under the same loading conditions, the isolation between C and S or between D_{1C} and D_{2C} (see Figure 1) are all better than —35 db, and the axial ratio in the phase shifter is 1 db or better. When the multimode coupler is terminated by the mode-filter–horn assembly, the coupling between C and S measures the reflection coefficient of the mode filter for the sum mode, while the coupling between D_{1C} and D_{2C} indicates the reflection coefficient of the mode filter for the difference modes.

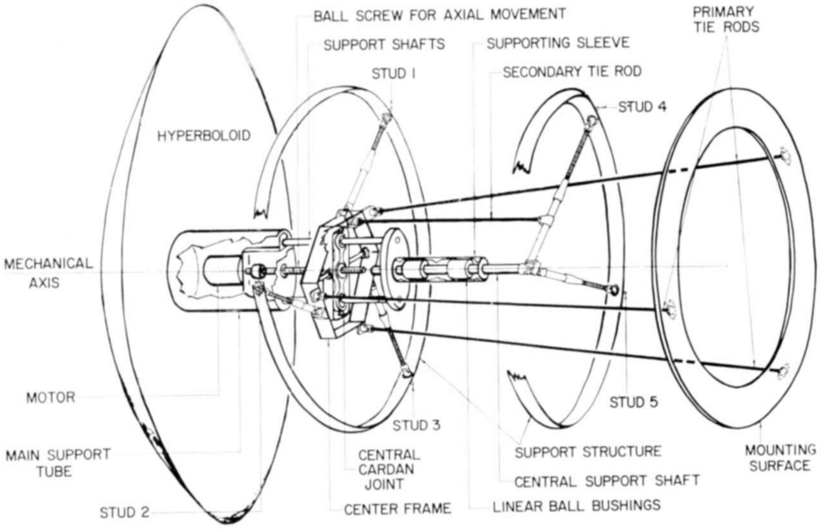


Fig. 8—Hyperboloid adjusting mechanism.

High-Isolation Measurements

In the development of high-rejection filters, it was necessary to monitor the transfer function of these components over a wide frequency band with a wide dynamic range. About 500-megacycle swept-frequency range and 140-db dynamic range was achieved by the use of a circuit in which two sweep generators were locked together with 70-megacycle i-f frequency separation. The signal of one sweep generator was amplified by a 15-watt traveling-wave tube, while the other generator was used as the local oscillator at the receiving end, which has a 30-megacycle-wide i-f strip. Under these conditions the transfer curve of filters can be displayed on an oscilloscope, and tuning over the full frequency and dynamic range of interest is possible.

Low-Loss Measurements

For measurement of low transmission loss values, either dual channel Weinschel[12] equipment or a somewhat modified classical short-circuit technique[13] was used. Both of these techniques resulted in loss-measuring accuracies better than ±0.02 db. To measure such low loss figures, particular attention had to be paid to the waveguide flanges. Generally, lapped joints were used with a flatness of better than 100 microinches and surface roughness of about 8 to 10 microinches. Whenever practical, similar roughness values were maintained on the inside of the waveguide components. All the components were silver plated and the plating was protected by a thin layer of gold. As surface roughness was only about 1/3 of the skin depth, the theoretical conductivity value of a smooth silver surface was nearly realized.

The dual-channel loss-measuring equipment is ideal in the case of an essentially two-terminal component (such as a band-pass filter). However, for multiple-terminal circuits such as hybrids, the short-circuit method is more practical. A special arrangement had to be used for the loss determination of the polarizer, which requires a circularly polarized detector or short-circuit network. This component was measured by connecting two identical polarizers back to back and then terminating the whole assembly by a short circuit. Such an assembly behaves as a linearly polarized component and the measured loss is twice that of one polarizer.

Radiation Characteristics

The feed system was tested in an anechoic chamber 30 feet in length and about 20 feet in cross section. Wall reflections were below —40 db. The antenna was placed on a turntable, which made possible the rotation of the feed about its phase center. In most instances the feed was used as a receiving antenna. Two independent test signals were intro-

---

[12] "Dual Channel Insertion Loss Test Set," Weinschel Engineering Application Note No. 4, 1962.

[13] *Handbook of Microwave Measurements*, edited by M. Sucher and J. Fox, Chapters VI and VII, Vol. 1, Polytechnic Press, New York, 1963.

duced to probe the radiation pattern of the system under test. The main test horn was placed on a mount that permitted remote adjustment of its distance from the feed and of its polarization. This test horn was set at a distance of 18 feet (focal distance of hyperboloid) from the phase center of the feed and was directed along the axis of the chamber. An auxiliary test horn was also placed at the same distance but offset by about 4 feet in the 45° plane relative to the vertical. With this arrangement, the two test horns are approximately in the direction of the D_{1C} and D_{2C} difference-mode maximums when the turntable is rotated toward the auxiliary test horn by about 12° relative to the axis of the chamber. The main test horn is then connected to No. 1 sweep generator and the auxiliary horn to No. 2 sweep generator, both sweeping in the same frequency band. The feed is terminated by four identical detectors at the S, C, D_{1C}, and D_{2C} terminals. The No. 1 sweep generator is synchronized with two dual-beam oscilloscopes. The four detectors are connected to the four available channels of these oscilloscopes, which then show the signals received by the center-sum, side-sum, and two difference-mode terminals.

Under these conditions D_{1C} gives the horizontal difference-mode signal (main polarized) and D_{2C} gives the cross-polarized horizontal difference-mode signal (which ideally should be zero). The No. 2 sweep generator is synchronized to a third oscilloscope that is also connected to the detectors at D_{1C} and D_{2C}. When the turntable is rotated to $\theta = 12°$, this oscilloscope indicates the vertical difference-mode signal at $\theta = 12°$, but D_{1C} now corresponds to the cross-polarized signal and D_{2C} is the main polarized signal.

With this setup, virtually all the pattern characteristics can be monitored simultaneously, on a swept-frequency basis, while adjustments are made on the feed. With $\theta = 0°$ turntable position and continuously rotating main test horn, the oscilloscope pictures reveal the axial ratio and equality of the S and C sum-mode levels and the null depth (symmetry) of the difference modes for all frequencies and incoming polarizations of interest. With $\theta = 12°$ turntable position, where the difference-mode maximums of the center system are located, the axial ratio of the difference modes and the level difference between main and cross-polarized difference signals can be seen. At $\theta = 15°$, which corresponds to the edge of the hyperboloid, both the edge tapering and the axial ratio of the sum modes at that angle can be monitored. This direction of the sum modes corresponds to about −9 db edge tapering. If the axial ratio is within 2 db for the whole 0 to −9 db portion of the sum-mode beam, then two conditions are simultaneously fulfilled: (1) the main beam is nearly axially symmetrical and (2) the side-lobe levels are about 18 db or better for any pattern cut or any incoming polarization. Finally, it was found that the smoothness of the received sum or difference signals (versus frequency) particularly at $\theta = 15°$, is a very good measure of the impedance match.

The only radiation characteristics that cannot be monitored by the above-described display are the phase patterns. For the measurement of these patterns, two separate homodyne[14] bridges were assembled, one for the 4-gigacycle and one for the 6-gigacycle band. Phase-measuring accuracy was about 1°.

For permanent recording of any of the radiation characteristics, discrete frequencies and continuously varying angles were used.

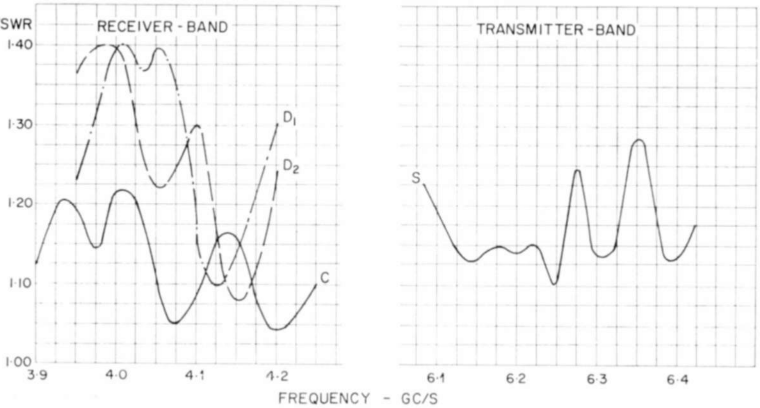


Fig. 9—Measured VSWR of the communication and difference channels (see Figure 1).

EXPERIMENTAL RESULTS

Figure 9 shows the input VSWR measured at the communication and difference channel inputs. It can be seen that the reflection coefficients are generally —20 db or better. It was found that relatively high-quality matching was necessary in order to maintain the proper amplitude and phase relationships between higher order modes in the multimode horn, to keep the internal losses of the feed low, and to maintain good axial ratio for the sum as well as difference modes in the radiation pattern. Figure 10 shows the isolation characteristics of the duplexer. The actual isolation between the Rx and Tx terminals is about 25 to 30 db higher for the circularly polarized operation modes due to the additional isolation available in the orthogonal coupler.

---

[14] F. L. Vernon, Jr., "Application of the Microwave Homodyne," *Trans. IRE PGAP*, Vol. AP-4, p. 110, Nov. 1952.

Table III summarizes the loss values in the individual circuit elements. All values were determined by calculation as well as by measurements; the values given in Table III are measured values, except in the case of the mode filters and horn, where the measurement of loss was not possible. Although there were some minor discrepancies between the calculated and measured loss values for the individual components, these discrepancies averaged out, and the overall measured loss was within 0.02 db of the calculated value. It is significant to note the very low losses, in spite of the complexity of the overall circuit.

Table III—Summary of Loss Measurements and Equivalent Noise Temperature in the Rx Frequency Band

Frequency (megacycles)	3900	3950	4000	4050	4100	4150	4200	Average
Individual Circuit Element Loss (db)								
Power-divider hybrid	.018	.014	.014	.014	.014	.014	.014	
Adj. phase-shifter	.015	.013	.010	.010	.009	.009	.009	
Center-system duplexer #1	.062	.089	.076	.082	.087	.101	.114	
Side-system duplexer #2	.063	.090	.091	.087	.098	.103	.110	
Monopulse combiner	.053	.059	.047	.046	.045	.047	.046	
Side-system 4-way power divider	.090	.082	.077	.073	.085	.094	.083	
Polarizer (based on one channel)	.090	.080	.075	.069	.066	.078	.082	
Mode filter and horn (calculated)	.014	.013	.013	.012	.012	.011	.011	
System Loss, 185-db Isolation (db)								
Center-system	.252	.268	.235	.233	.233	.260	.276	
Side-system	.290	.292	.280	.265	.286	.311	.311	
Average	.271	.280	.253	.249	.260	.285	.293	
Circular polarization (center-system)	.219	.241	.211	.209	.208	.237	.253	.225
Linear polarization	.271	.280	.253	.249	.260	.285	.293	.270
System Loss, 85-db Isolation (db)								
Center-system	.232	.238	.205	.203	.203	.230	.246	
Side-system	.260	.262	.250	.235	.256	.281	.281	
Average	.241	.250	.223	.219	.230	.255	.263	
Circular polarization (center-system)	.189	.211	.181	.179	.178	.207	.223	.195
Linear polarization	.241	.250	.223	.219	.230	.255	.263	.240
Noise Temperature (°K), 185-db Isolation								
Circular polarization	13.9	15.4	13.4	13.2	13.2	15.1	16.1	14.3
Linear polarization	17.3	17.9	16.1	15.9	16.6	18.2	18.7	17.2
Noise Temperature (°K), 85-db Isolation								
Circular polarization	11.9	13.4	11.4	11.3	11.3	13.1	14.2	12.3
Linear polarization	15.4	15.9	14.2	13.9	14.6	16.3	16.8	15.3

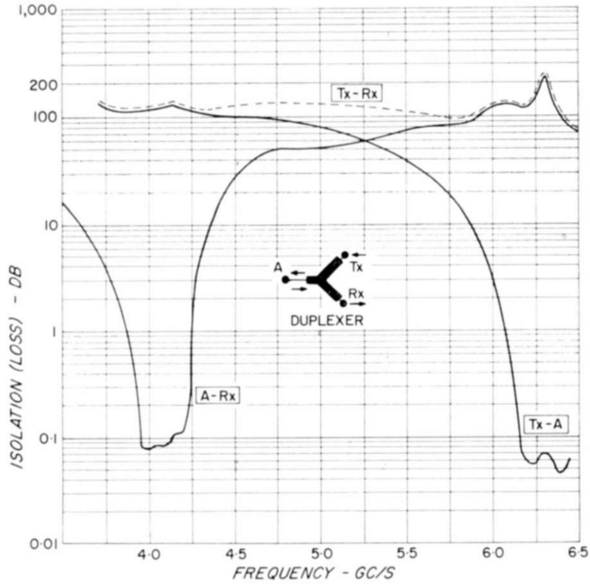


Fig. 10—Isolation (loss) between different terminals of high-power low-loss duplexer.

This indicates the importance of very high surface finish and high-quality impedance matching for low-noise applications.

Figure 11 shows the axial ratio versus frequency for the different radiation modes. The axial ratio for the sum mode is given along the

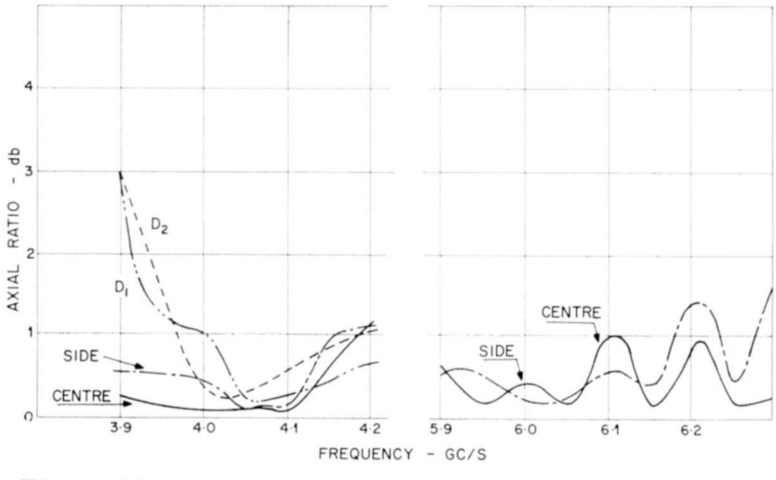


Fig. 11—Measured axial ratios (center and side sum at $\theta = 0°$ and differences at $\theta = 6°$ azimuth angle).

axis of the feed system, while that for the *resultant* difference modes is in the direction of the principal maximum (about $\theta = 6°$). It should be noted that these axial ratios are not the same as the axial ratios in the secondary pattern; however, they are closely related to them. The true secondary axial ratios can be determined from the complete polarization distribution of the primary patterns.

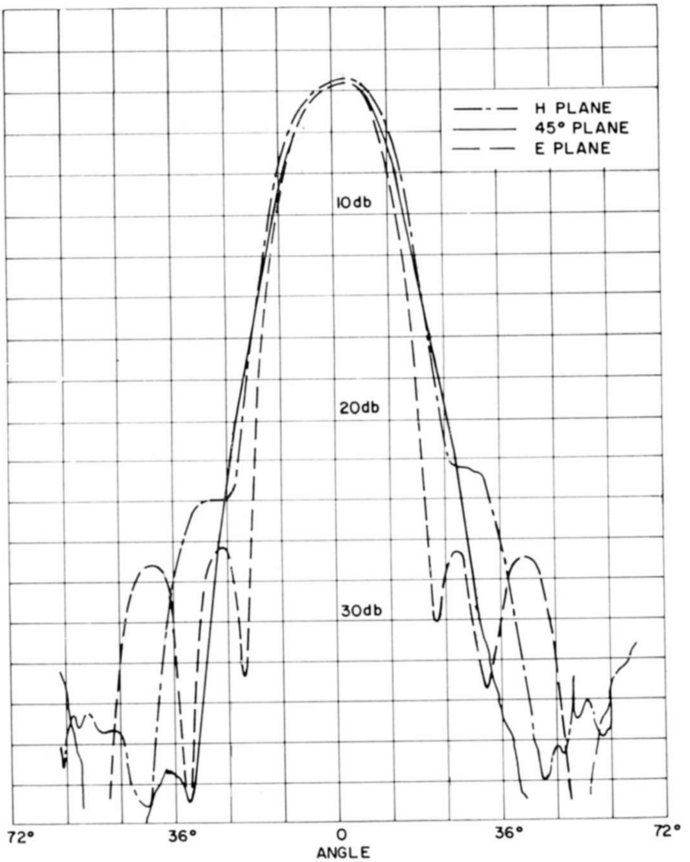


Fig. 12—Measured sum mode patterns in receiver band; $f = 4$ gigacycles, center system.

Figures 12, 13, and 14 show the communication receive (sum) pattern, the tracking receive (difference) pattern, and the communication transmit (sum) pattern. The pattern "cuts" were taken in the 0°, 45° and 90° planes. Note the very high degree of axial symmetry and low side lobes of the receive patterns and the still optimum beam shape of the transmit patterns in spite of the 60% higher frequencies. Another important radiation feature of the feed is the nearly optimum beam width of the resultant difference-mode pattern as compared to the sum mode.

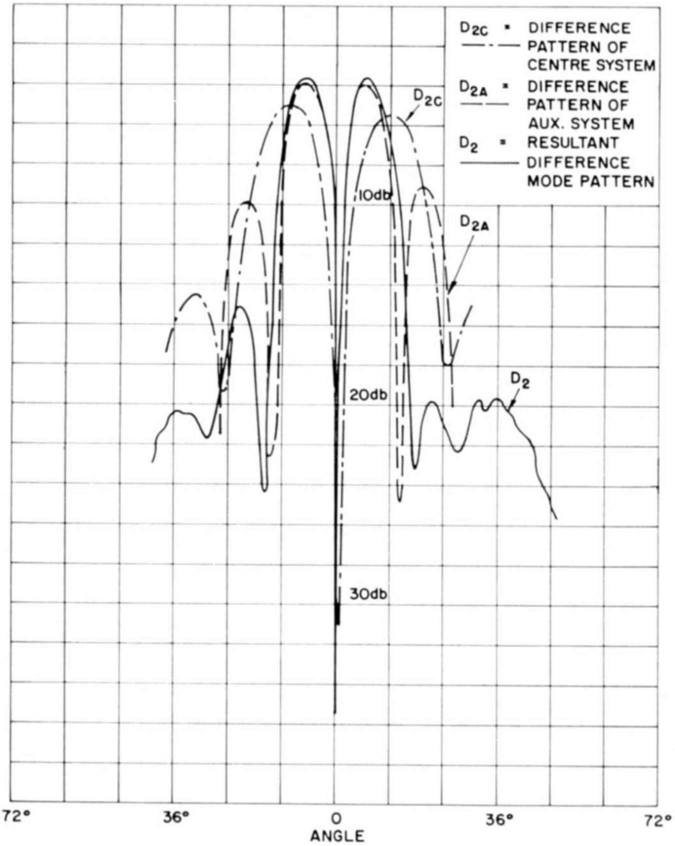


Fig. 13—Measured difference mode patterns; $f = 4$ gigacycles, 45° plane.

Figure 15 shows the effect of phase errors in the sum-mode pattern as a function of frequency. The data actually displayed is an equivalent r-m-s surface tolerance caused by the nonperfection of the source pattern. The equivalent r-m-s error was determined from three plane cuts of the phase patterns. It is interesting to note that the phase center itself is closer to the throat of the horn than for a basic mode horn, but has very small variation with the orientation of the horn or with frequency. It was found that the phase center (P in Figure 15) was 6 ± 1 inches behind the 10-inch horn aperture for both the Rx and Tx frequency bands. The constancy of phase center position is related to the axial symmetry and frequency independence of the main-beam shape. The result is very significant, because it means a sharply defined point source and makes possible optimum focusing of the source relative to the hyperboloid in wide frequency band. Figure 16 indicates another important phase characteristic of the feed, namely the relative phase shifts at the various output terminals as a function of the attitude of incoming polarization. Generally, the phase between a signal at C and D_{1C} or D_{2C} varies slightly with changing input polarizations. This

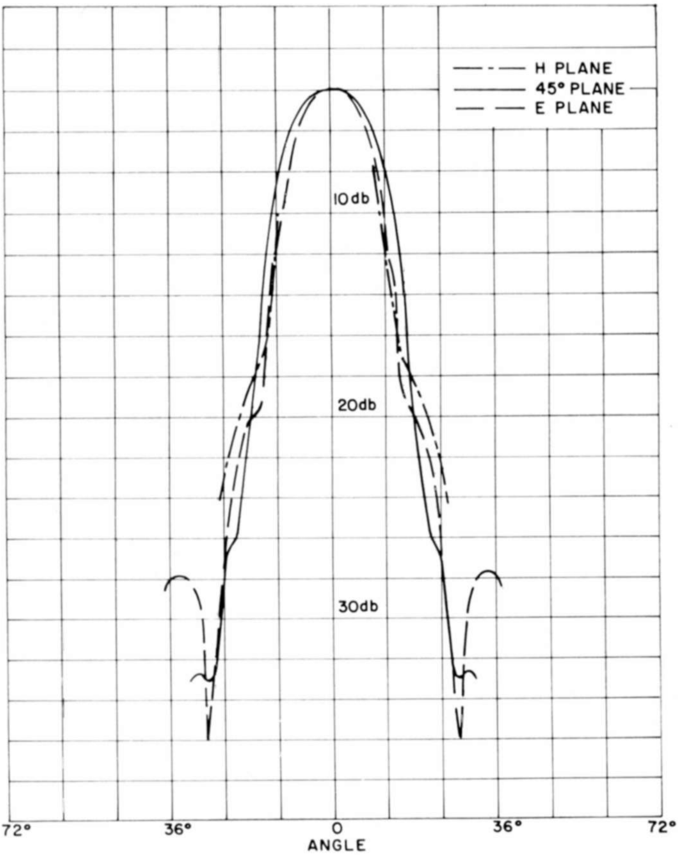


Fig. 14—Measured sum mode pattern in transmit band; $f = 6.2$ gigacycles, side system.

change offsets the optimum alignment of the monopulse receiver, which operates on a phase–amplitude basis. It can be seen that for the combined overall system, in which the phase errors in the center and auxiliary system are averaged, the maximum differential phase error is less than ±5°, which is well within the tolerable practical limit. The graph also shows that the phase of the signals at C and S vary in the opposite direction with a high degree of linearity. The symmetry and linearity of these curves are additional measures of the quality of circular polarization in the system, because for ideal circular polarization the curves become symmetrical straight lines. Note that D_{2C} difference signal should have zero amplitude in an ideal system if the feed is rotated in the plane of the D_{1C} signal; in practice, it has a finite value and opposite sense of polarization to the useful difference signal D_{1C}.

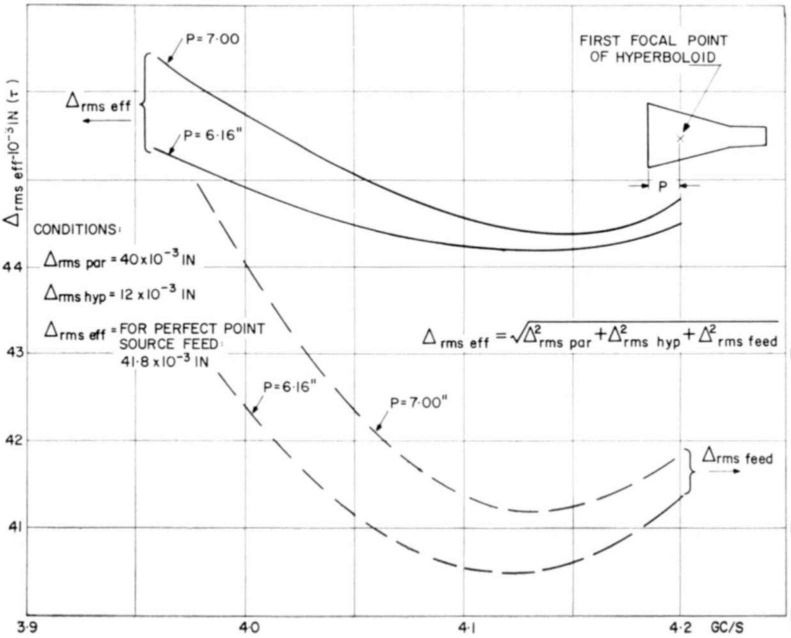


Fig. 15—Effective "surface tolerance" versus frequency.

Calculated Secondary Characteristics

It is generally accepted that the secondary characteristics (gain and noise temperature) of a microwave optics system can be determined to a fair degree of accuracy by calculations based on a sufficient amount of primary (feed) pattern characteristics. The secondary radiation characteristics of large, complex antenna systems are quite difficult to measure directly, and the accuracy of such measurements in many instances does not exceed the accuracy available from prediction based on primary characteristics.

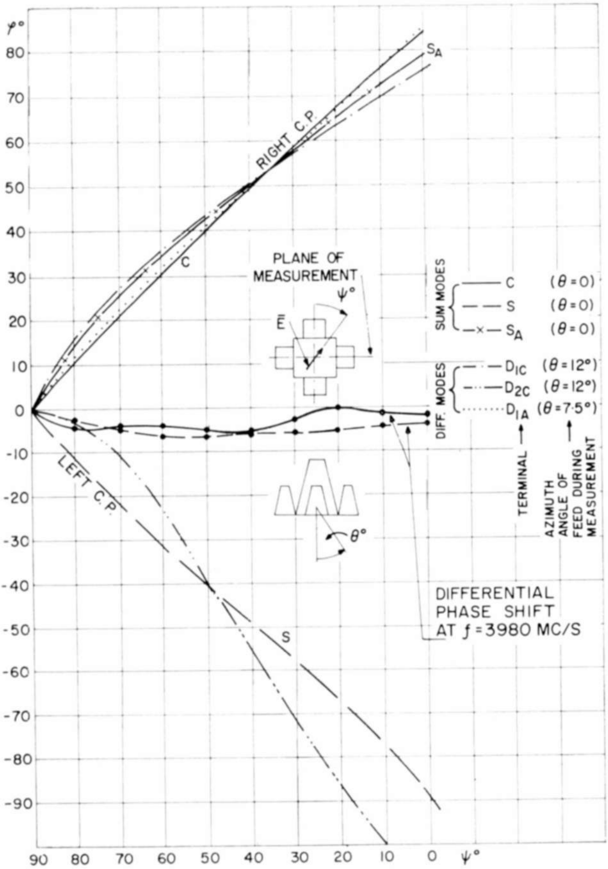


Fig. 16—Relative phase shift versus polarization angle for the different output terminals.

In this section only two calculated characteristics of the secondary pattern are included, namely, the antenna efficiency (gain) as a function of frequency (see Table IV) and the noise temperature as a function of elevation angle (Figure 17).

The antenna efficiency is the most general and important characteristic of the antenna, and is defined as follows: *the antenna efficiency η for the B terminal is the magnitude of the transfer function between the basic waveguide mode in aperture B and the unipolarized plane wave with constant amplitude and zero phase in aperture C.* Here, aperture C is defined as a circular aperture with a nominal diameter of 85 feet. This aperture is perpendicular to the axis of the ideal boresight direction and the antenna and is centrally positioned in the focal plane of the paraboloid. The above definition of C includes the effect of the various blockages, as well as losses in the waveguide circuit. It is assumed that aperture C is in free space, i.e., no effects of the radome, ground, building, etc., are included.

Table IV—Summary of Efficiency Factors and Gain of the Antenna (db)

Frequency (gigacycles)	$-\eta_R$	$-\eta_L$	$-\eta_{SH}$	$-\eta_{SP}$	$-\eta_M$	$-\eta_A$	$-\eta_{PF}$	$-\eta_{PR}$	$-\eta_B$	$-\eta$	G
3.8	.100	.400	.993	.116	.172	.302	.052	.048	.153	2.336	57.92
3.9	.040	.243	.880	.100	.052	.287	.063	.058	.153	1.876	58.61
4.0	.040	.242	.768	.084	.031	.266	.074	.071	.153	1.729	58.98
4.1	.021	.240	.757	.068	.058	.306	.095	.090	.153	1.788	59.13
4.2	.040	.256	.639	.052	.101	.370	.119	.112	.153	1.842	59.28
									Av =	1.869	58.82
5.9	.100	.320	1.656	.100	.237	.405	.125	.140	.158	3.241	60.85
6.0	.020	.320	1.333	.040	.424	.317	.136	.145	.158	2.893	61.24
6.1	.020	.314	.797	.025	.611	.887	.215	.150	.158	3.178	61.19
6.2	.020	.320	.622	.020	.828	1.005	.275	.155	.158	3.403	61.15
6.3	.060	.326	.557	.015	.611	.983	.330	.160	.158	3.201	61.46
6.4	.042	.360	.745	.020	.612	1.075	.380	.165	.158	3.556	61.23
									Av =	3.286	61.15

Note: The above figures are calculated for the linearly polarized operational mode, with an average orientation of polarization. For circular polarization the resultant gain is practically identical. The accuracy of gain figure for the receiver frequency band is better than .25 db; for the transmitter frequency band, it is better than .35 db. $\eta_{SH} + \eta_{SP} = \eta_S$, where η_{SH} and η_{SP} are the efficiencies corresponding to spillover radiation around the hyperboloid and paraboloid reflectors, respectively.

The terminal B can be either a sum or difference mode input terminal.

Only the efficiency of the communication channel is considered here. The restriction of unipolarization in aperture C is not substantial, because any elliptically polarized incoming wave can be expressed in terms of linearly polarized orthogonal components.

The accurate expression for η is fairly complicated, but, for most practical purposes, the following approximation is adequate:

$$\eta = \eta_R \eta_L \eta_S \eta_M \eta_A \eta_P \eta_B,$$

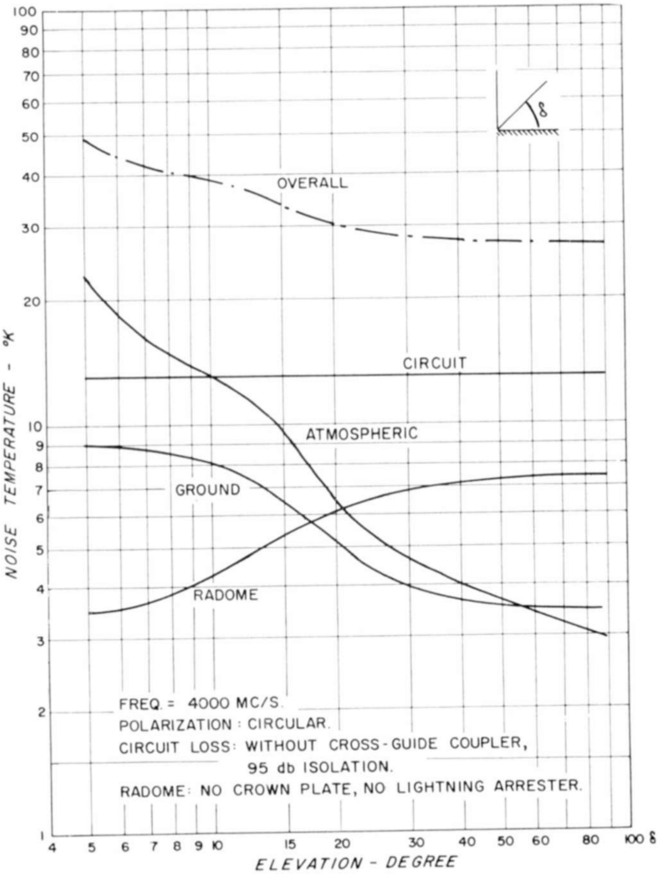


Fig. 17—Noise temperature of the antenna versus elevation angles (components and overall system).

where η_R is the reflection loss at the terminal B,

η_L is the insertion loss between terminal B and C,

η_S is the spillover efficiency, defined as the ratio of the main polarized power intercepted by the paraboloid to the main polarized power in the aperture of the source, assuming a perfectly transparent hyperboloid and supporting quadrupod,

η_M is the main polarized power efficiency, defined as the ratio of main polarized power radiated by aperture C relative to the total power radiated by the same aperture,

η_A is the amplitude (or illumination taper) efficiency, which takes into account the nonuniform-field amplitude distribution of the antenna in the plane C,

η_P is the phase-distribution efficiency, which takes into account the nonuniform-field phase distribution of the antenna in plane C,

η_B is the blockage factor, which takes into account the incomplete utilization of the aperture C.

These factors are determined by a method that utilizes the measured source patterns, loss, and isolation data. The details of this method were worked out by two of the authors (Foldes and Komlos). Table IV gives the individual contributory efficiencies and the overall efficiency in —db, i.e., the given db figure must be deducted from the theoretical gain figure of an ideal antenna with identical aperture size. Because of the multimode properties of the feed system, the η_S and η_A efficiencies vary irregularly, as reflected in the gain figures indicated.

Figure 17 shows the noise temperature of the antenna system versus elevation agle. This figure assumes a Hypalon-coated Dacron radome with a diameter of 120 feet and a wall thickness of 0.055 inch, the terrain existing at the site, spillover radiation characteristics of the feed system, the atmospheric losses, and the internal losses given in Table I. In calculation of noise temperature, the attenuating effect of the feed losses against external noise-temperature sources was included.

CONCLUSIONS

On the basis of measured primary characteristics the following main conclusions can be obtained:

(a) It is possible to control the aperture distribution of a primary source, through the use of higher order modes, in such a way that an approximately —2 db antenna efficiency can be maintained in the large portion of the 500-megacycle receiver frequency band of communication satellite systems.

(b) The relative complexity necessary for mode control and operational flexibility does not exclude the possibility of low internal loss figures.

(c) The resultant system has a high gain/noise-temperature ratio with a relatively simple, conventional paraboloid reflector, but accurate fabrication techniques and sophisticated measuring methods are necessary to produce the feed itself. The penalty of the more complicated feed is more than offset by increased simplicity of the overall antenna system.

ACKNOWLEDGMENTS

The work reported here was performed for the Canadian Department of Transport, whose interest and contractual support made possible the development of the new feed system. As in any large scale team work, it is virtually impossible to acknowledge everyone who contributed either by criticism or by direct work to the success of this work. Nevertheless, the authors wish to express their thanks to all in the Department of Transport and in the Antenna Group of RCA Victor Company, Ltd., whose contribution meant the difference between success and failure. Particular credit should be given to E. Morris, R. Desrochers, H. Schwarz, and K. Flood, who performed many of the individual design tasks.

HIGH-EFFICIENCY, HIGH-ORDER, IDLER-LESS FREQUENCY MULTIPLIERS USING HYPERABRUPT VARACTORS

By

E. Markard and S. Yuan

RCA Communications Systems Division
New York, N. Y.

Summary—*Expressions for the efficiency and power-handling capability of high-order, idler-less frequency multipliers as functions of the nonlinearity coefficient are developed. It is shown that the efficiency of such multipliers can be increased substantially if the varactors possess a high nonlinearity coefficient, γ, i.e., hyperabrupt junction. Since it is a reactive device with a large reverse breakdown voltage, the hyperabrupt varactor can convert a large amount of power at microwave frequencies. Input power in excess of one watt can easily be handled, since hyperabrupt varactors having a reverse breakdown voltage of 75 volts are presently available.*

An experimental times-eight idler-less multiplier using a hyperabrupt varactor (with $\gamma = 0.875$) multiplying from 500 to 4000 megacycles gave an efficiency of 25% including passive circuit losses. The tuning is extremely simple, since the circuit involves only input frequency and output frequency loops. It has also been demonstrated that this multiplier is able to accept wide ranges of drive power without retuning.

INTRODUCTION

THE efficiency and power-handling capability of varactor frequency multipliers are highly dependent upon the nonlinearity coefficient of the capacitance (or elastance) variation. For a graded or an abrupt-junction diode, the efficiency drops off very rapidly as the order of multiplication is increased if no idlers are provided. It is desirable to avoid the use of idlers, since they add to the complexity of the multiplier circuit and can make tuning quite difficult. In order to accomplish high-order efficient multiplication without the use of idlers, varactors with high nonlinearities are required. Such varactors are referred to as being hyperabrupt. Until very recently hyperabrupt varactors were available only in very small quantities as engineering samples. Now, however, they are becoming available as standard products.

At the present time, it appears that when varactor junctions are made hyperabrupt, it is done at the expense of cutoff frequency.[1]

---

[1] J. J. Chang, J. H. Foster and R. M. Ryder, "Semiconductor Junction Varactors with High Voltage-Sensitivity," I.E.E.E. Trans. on Electron Devices, Vol. ED-10, p. 281, July 1963.

However, as γ increases, there is a net increase in the conversion resistance in spite of the effect of the degraded cutoff frequency. Hence, although for low-order multipliers (doublers and triplers) hyperabrupt varactors have no particular advantage over abrupt varactors, they do offer definite advantages when the case of high-order multiplication without idlers is considered.

In this analysis, an expression for the efficiency of a high-order multiplier without idlers as a function of the nonlinearity coefficient is developed. The analysis is carried out in general terms, except where a specific example is introduced to simplify the mathematics involved. In particular, the efficiency of a times-eight multiplier employing a varactor that has a nonlinearity coefficient $\gamma = 0.875$ is optimized.

FUNDAMENTAL VOLTAGE–CHARGE RELATIONSHIP

The general expression for the elastance–voltage relationship for a varactor diode is

$$S = S_m \left(\frac{V_C + V}{V_C + V_B} \right)^{\gamma} \tag{1}$$

where

S is the elastance (reciprocal capacitance) of the reverse biased diode;

V is the applied reverse voltage, which is opposite in sign to the built-in contact potential, V_C, across the junction over most of the operating range;

V_C is the contact potential, and is in the order of a fraction of a volt;

V_B is the breakdown voltage of the varactor diode;

γ is the nonlinearity coefficient of the varactor diode;

$S_m (= S_{\max})$ is the elastance of the reverse-biased diode at breakdown voltage.

Figure 1 shows a plot of normalized elastance as a function of normalized voltage.

The elastance is also given by the slope of the voltage–charge curve,

$$S = \frac{1}{C} = \frac{dV}{dQ}. \tag{2}$$

Charge as a function of voltage can be derived by combining Equations (1) and (2);

$$dQ = \frac{(V_C + V_B)^{\gamma}}{S_m} \frac{dV}{(V_C + V)^{\gamma}}. \tag{3}$$

Integrating Equation (3),

$$Q = \frac{(V_C + V_B)^{\gamma}}{(1 - \gamma) S_m} (V_C + V)^{1-\gamma} + K, \tag{4}$$

where K is the constant of integration; γ, of course, must not be unity.

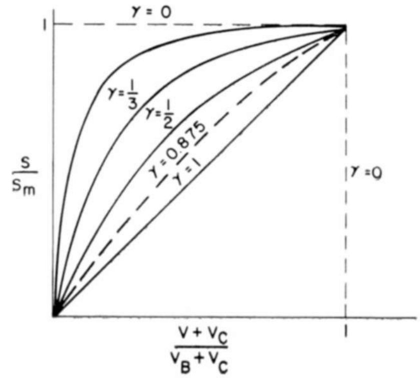


Fig. 1—Normalized elastance as a function of normalized voltage.

When the applied voltage, V, just cancels the contact potential, so that $V_C + V = 0$, K equals the charge on the junction, $-Q_c$. Thus,

$$Q_c + Q = \frac{(V_C + V_B)^{\gamma}}{(1 - \gamma) S_m} (V_C + V)^{1-\gamma}, \tag{5}$$

where

$$Q_c = \frac{V_C^{1-\gamma}}{(1 - \gamma) S_m} (V_C + V_B)^{\gamma}. \tag{6}$$

The maximum charge the junction will support, Q_m, occurs at the breakdown voltage, V_B, and, from Equation (5), is

$$Q_c + Q_m = \frac{(V_C + V_B)}{(1-\gamma)S_m}. \tag{7}$$

Substituting Equation (7) into Equation (5),

$$Q_c + Q = (Q_c + Q_m)\left(\frac{V_C + V}{V_C + V_B}\right)^{1-\gamma}. \tag{8}$$

Equation (8) can be rewritten

$$\frac{V_C + V}{V_C + V_B} = \left(\frac{Q_c + Q}{Q_c + Q_m}\right)^{1/(1-\gamma)}. \tag{9}$$

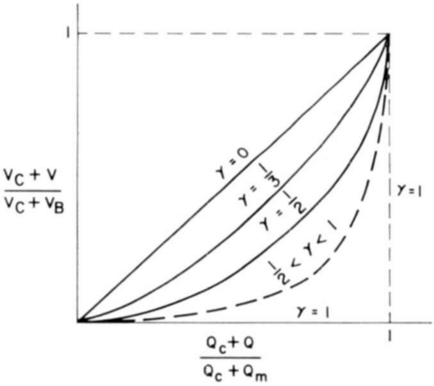


Fig. 2—Normalized voltage as a function of normalized charge.

Figure 2 is a plot of normalized voltage as a function of normalized charge. As γ approaches zero, the voltage–charge curve approaches a straight line, and the diode capacitance becomes constant. As γ increases from zero toward unity, the capacitance deviates more severely from linearity. Since $Q_c + Q_m$ is a function of γ, as expressed by Equation (7), the maximum charge the junction can support increases as γ increases. Therefore, it should be noted that the normalizing factor in Figure 2 is a function of γ. Referring to Equation (7), for an abrupt-junction diode ($\gamma = 1/2$),

$$Q_c + Q_m = \frac{2(V_C + V_B)}{S_m}; \tag{10}$$

and, for a graded-junction diode ($\gamma = 1/3$),

$$Q_c + Q_m = \frac{3}{2} \frac{(V_C + V_B)}{S_m}. \tag{11}$$

The normalized diode voltage, $(V_C + V)/(V_C + V_B)$, can be plotted as a function of charge normalized to $Q_c + Q_m$ for the abrupt junction, as shown in Figure 3.

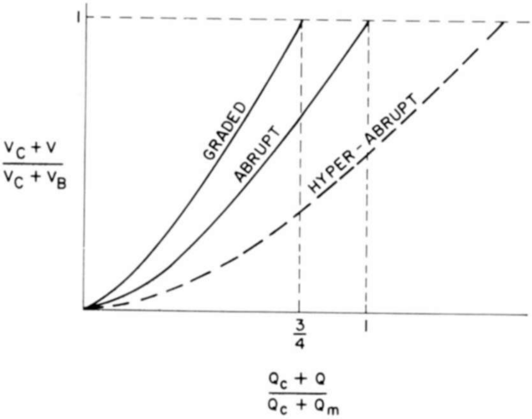


Fig. 3—Normalized voltage as a function of charge normalized to the maximum charge of an abrupt junction.

Since the instantaneous power is the product of voltage and current, and current is the time derivative of the charge, it is obvious that the peak currents and, therefore, the peak power, must be higher for larger γ.

TAYLOR-SERIES REPRESENTATION

The voltage–charge relationship can be represented by a Taylor series expansion about a quiescent point, q_0, with a time-varying component, q, such that

$$Q = q_0 + q. \tag{12}$$

The pertinent Taylor expansion is

$$f(q_0 + q) = f(q_0) + f^{(1)}(q_0)\, q + \frac{f^{(2)}(q_0)}{2!} q^2 + \cdots + \frac{1}{n!} f^{(n)}(q_0)\, q^n, \tag{13}$$

where $f^{(n)}(q_0)$ is the n^{th} derivative evaluated at the bias point, q_0. Rearranging Equation (8),

$$f(q+q_0) = f(Q) = V + V_c = S_m(1-\gamma)\,\frac{(Q_c+Q)^{1/(1-\gamma)}}{(Q_c+Q_m)^{\gamma/(1-\gamma)}}. \tag{14}$$

Evaluating $f^{(n)}(q_0)$ from Equation (14), and setting $q_0 = \tfrac{1}{2}(Q_m - Q_c)$, the Taylor series expansion for the generalized nonlinearity becomes

$$\begin{aligned} V + V_c = {} & \frac{Q_c+Q_m}{2(\Gamma+1)} S_0 + S_0\,q + \frac{\Gamma}{(Q_c+Q_m)} S_0\,q^2 + \frac{4\Gamma(\Gamma-1)}{6(Q_c+Q_m)^2} S_0\,q^3 + \\ & \cdots\cdots + \frac{2^{n-1}}{n!}\,\frac{\Gamma!}{(\Gamma-n+1)!\,(Q_c+Q_m)^{n-1}} S_0\,q^n, \end{aligned} \tag{15}$$

where

$$\Gamma = \frac{\gamma}{1-\gamma}, \tag{16}$$

and

$$S_0 = f^{(1)}(q_0) = \frac{S_m}{2^{\Gamma}}. \tag{17}$$

The corresponding series for the elastance, S, is obtained by differentiating Equation (15):

$$\begin{aligned} S = {} & S_0 + 2\Gamma S_0 \frac{q}{(Q_c+Q_m)} + 2\Gamma(\Gamma-1)\,S_0 \left(\frac{q}{Q_c+Q_m}\right)^2 + \cdots \\ & + \frac{2^{(n-1)}}{(n-1)!}\left[\frac{\Gamma!}{(\Gamma-n+1)!}\right] S_0 \left(\frac{q}{Q_c+Q_m}\right)^{n-1}. \end{aligned} \tag{18}$$

Current-Pumped,* Two-Frequency Direct N-Times Multiplier

Current pumping means that the diode elastance is made to vary in time by controlling the current through and, hence, the charge on, the junction. Current, or the integral of current, charge, then is the independent variable, while the voltage produced across the diode terminals by the current is the dependent variable.

In a current-pumped two-frequency direct N-times multiplier, only

---

* The current pumping method of operation was brought to our attention by M. E. Malchow of RCA.

currents at the input frequency, ω, and the output frequency, $N\omega$, are allowed to flow. Currents at all other frequencies are assumed to be open circuited by filters. Therefore, the current flowing through the diode is

$$i(t) = i_1(t) + i_N(t) = \frac{1}{2}\left[i_1 e^{j\omega t} + i_1^* e^{-j\omega t} + i_N e^{jN\omega t}\right.$$

$$\left. + i_N^* e^{-jN\omega t}\right] = I_1 \cos(\omega t + \theta) + I_N \cos(N\omega t + \phi). \quad (19)$$

For convenience we define

$$m_N = \frac{|i_N|}{N\omega(Q_c + Q_m)}, \quad (N = \text{any integer}) \quad (20)$$

$$M_N = m_N e^{j\phi}, \quad (21)$$

$$M_N^* = m_N e^{-j\phi}. \quad (22)$$

In general, $V_c \ll V$, and by including the effects of the spreading resistance, R_s, the diode terminal voltage becomes, from Equation (15),

$$V_T(t) = \frac{Q_c + Q_m}{2(1+\Gamma)} S_0 + R_s i(t) + S_0\left[\int i(t)\,dt\right]$$

$$+ \frac{\Gamma}{(Q_c + Q_m)} S_0\left[\int i(t)\,dt\right]^2 +$$

$$\cdots + \frac{2^{n-1}}{n!} \frac{\Gamma!}{(\Gamma - n + 1)!(Q_c + Q_m)^{n-1}} S_0\left[\int i(t)\,dt\right]^n. \quad (23)$$

All equations derived up to this point are exact and general. In order to make the mathematical development tenable, a specific example will be given to illustrate the procedure for calculating the multiplier efficiency.

CALCULATION OF N-TIMES MULTIPLIER EFFICIENCY

The general procedure is as follows:

(1) Substitute Equation (19) into Equation (23).

(2) Extract the coefficients of all terms involving the frequencies ω and $N\omega$ for the voltages across the varactor terminals at the respective frequencies.

(3) Add the external circuit and derive the matrix equation that describes the operation of the circuit.

(4) Derive the equation for efficiency as a function of the load resistance and drive level.

(5) Optimize the efficiency as a function of load resistance and drive level.

The above procedure will be demonstrated for an eight times multiplier. The current flowing through the diode is

$$i(t) = I_1 \cos (\omega t + \theta) + I_8 \cos (8\omega t + \phi), \tag{24}$$

and the charge is

$$q = \int i(t)\,dt = \frac{I_1}{\omega} \sin (\omega t + \theta) + \frac{I_8}{8\omega} \sin (8\omega t + \phi). \tag{25}$$

Substituting Equations (24) and (25) into (23), and for convenience letting $U = I_1 \sin (\omega t + \theta) + (I_8/8) \sin (8\omega t + \phi)$

$$\begin{aligned} V_T(t) &= \frac{Q_c + Q_m}{2(1+\Gamma)} S_0 + R_s\left[I_1 \cos (\omega t + \theta) + \frac{I_8}{8} \cos (8\omega t + \phi) \right] \\ &+ \frac{S_0}{\omega} U + \frac{\Gamma}{\omega(Q_c + Q_m)} \frac{S_0}{\omega} U^2 + \frac{4\Gamma(\Gamma - 1)}{6[\omega(Q_c + Q_m)]^2} \frac{S_0}{\omega} U^3 + \cdots\cdots \\ &+ \frac{2^{n-1}}{n!} \frac{\Gamma!}{(\Gamma - n + 1)!\,[\omega(Q_c + Q_m)]^{n-1}} \frac{S_0}{\omega} U^n. \end{aligned} \tag{26}$$

To simplify the analysis further, assume $\gamma = 0.875$. From Equation (15), $\Gamma = 7$. Note that by choosing $\gamma = 0.875$ ($\Gamma = 7$), Equation (26) terminates at $n = 8$, since for $n \geq 9$ the coefficients become zero. Equation (26) becomes

$$V_T(t) = \frac{Q_c + Q_m}{16} + R_s[I_1 \cos (\omega t + \theta) + \frac{I_8}{8} \cos (8\omega t + \phi)] + \frac{S_0}{\omega} U +$$

$$+\frac{7}{\omega(Q_c+Q_m)}\frac{S_0}{\omega}U^2+\frac{28}{[\omega(Q_c+Q_m)]^2}\frac{S_0}{\omega}U^3$$
$$+\frac{70}{[\omega(Q_c+Q_m)]^3}\frac{S_0}{\omega}U^4+\frac{112}{[\omega(Q_c+Q_m)]^4}\frac{S_0}{\omega}U^5$$
$$+\frac{112}{[\omega(Q_c+Q_m)]^5}\frac{S_0}{\omega}U^6+\frac{64}{[\omega(Q_c+Q_m)]^6}\frac{S_0}{\omega}U^7$$
$$+\frac{16}{[\omega(Q_c+Q_m)]^7}\frac{S_0}{\omega}U^8. \tag{27}$$

Equation (27) is the relationship that specifies the terminal voltage across the varactor when the current, $i(t)=I_1\cos(\omega t+\theta)+I_8\cos(8\omega t+\phi)$, flows through it. Figure 4 shows the equivalent circuit representation of the varactor.

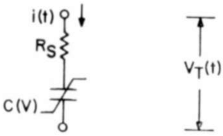


Fig. 4—Equivalent circuit of the varactor diode.

EXTERNAL CIRCUIT CONSTRAINTS

The terminal voltage as expressed by Equation (27) contains many frequency components. However, when the external circuits are introduced, the only voltage terms of importance in Equation (27) are at d-c, ω, and 8ω. The external circuit must have the required filters to present an open circuit to the undesired frequency components. Once the external circuit constraints are satisfied, operation of the multiplier is determined completely. The actual circuit that meets such requirements for an eight-times multiplier is shown in Figure 5.

Having specified the circuit, one can determine circuit performance by first determining the varactor terminal voltages at the fundamental frequency, ω, and the eighth harmonic, 8ω. These voltages are determined by collecting the coefficients of each frequency from Equation (27). Thus, for the fundamental frequency, ω,

$$v_1(t)=V_1\cos(\omega t+\alpha)=\frac{1}{2}(v_1e^{j\omega t}+v_1{}^*e^{-j\omega t})$$

$$= R_s I_1 \cos(\omega t + \theta) + \frac{S_0}{\omega} I_1 \sin(\omega t + \theta)$$

$$+ I_1 \sin(\omega t + \theta) \left\{ \left[\frac{3}{4} I_1^2 + \frac{3}{2} \left(\frac{I_8}{8} \right)^2 \right] \frac{28}{[\omega(Q_c + Q_m)]^2} \frac{S_0}{\omega} \right.$$

$$+ \left[\frac{5}{8} I_1^4 + \frac{15}{4} I_1^2 \left(\frac{I_8}{8} \right)^2 + \frac{15}{8} \left(\frac{I_8}{8} \right)^4 \right] \frac{112}{[\omega(Q_c + Q_m)]^4} \frac{S_0}{\omega}$$

$$+ \left[\frac{35}{64} I_1^6 + \frac{105}{16} I_1^4 \left(\frac{I_8}{8} \right)^2 + \frac{315}{32} I_1^2 \left(\frac{I_8}{8} \right)^4 \right.$$

$$\left. \left. + \frac{35}{16} \left(\frac{I_8}{8} \right)^6 \right] \frac{64}{[\omega(Q_c + Q_m)]^6} \frac{S_0}{\omega} \right\}$$

$$+ \frac{I_8}{8} \sin(8\omega t + \phi) \left[\frac{I_1^7}{8} \frac{e^{j(7\omega t + 7\theta)} - e^{-j(7\omega t + 7\theta)}}{2j} \right] \frac{16}{[\omega(Q_c + Q_m)]^7} \frac{S_0}{\omega}, \tag{28}$$

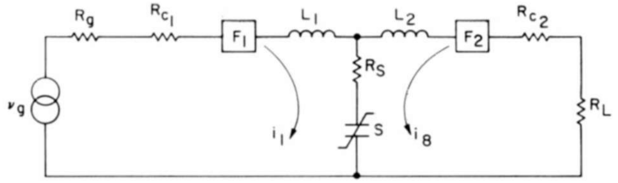


Fig. 5—Circuit configuration of an eight-times multiplier (F_1 and F_2 are bandpass filters that are short circuits at ω and at 8ω, respectively, and that are open circuits at all other frequencies).

where
$$v_1 = |v_1| e^{j\alpha} = V_1 e^{j\alpha}. \tag{29}$$

Similarly, the voltage for the eighth harmonic, 8ω, is

$$v_8(t) = V_8 \cos(8\omega t + \beta) = \frac{1}{2} (v_8 e^{j8\omega t} + v_8^* e^{-j8\omega t})$$

$$= R_s I_8 \cos(8\omega t + \phi) + \frac{S_0}{\omega} \frac{I_8}{8} \sin(8\omega t + \phi)$$

$$+ \frac{I_8}{8} \sin(8\omega t + \phi) \left\{ \left[\frac{3}{2} I_1^2 + \frac{3}{4} \left(\frac{I_8}{8} \right)^2 \right] \frac{28}{[\omega(Q_c + Q_m)]^2} \frac{S_0}{\omega} + \right.$$

$$
+\left[\frac{15}{8}I_1^{4}+\frac{15}{4}I_1^{2}\left(\frac{I_8}{8}\right)^{2}+\frac{5}{8}\left(\frac{I_8}{8}\right)^{4}\right]\frac{112}{[\omega(Q_c+Q_m)]^{4}}\left(\frac{S_0}{\omega}\right)
$$

$$
+\left[\frac{35}{16}I_1^{6}+\frac{315}{32}I_1^{4}\left(\frac{I_8}{8}\right)^{2}+\frac{105}{16}I_1^{2}\left(\frac{I_8}{8}\right)^{4}\right.
$$

$$
\left.+\frac{35}{64}\left(\frac{I_8}{8}\right)^{6}\right]\frac{64}{[\omega(Q_c+Q_m)]^{6}}\frac{S_0}{\omega}\Bigg\}
$$

$$
+I_1\sin(\omega t+\theta)\left[\frac{I_1^{7}}{64}\frac{e^{j7\omega t}-e^{-j7\omega t}}{2}\right]\frac{16}{[\omega(Q_c+Q_m)]^{7}}\frac{S_0}{\omega}, \tag{30}
$$

where

$$
v_8=|v_8|e^{j\beta}=V_8e^{j\beta}. \tag{31}
$$

For a resonance condition, operation of the multiplier is characterized by the following matrix equation:

$$
\begin{bmatrix} v_g \\ 0 \end{bmatrix}=\begin{bmatrix} R_g+R_{c1}+R_s & \dfrac{M_1^{*7}}{8}\dfrac{S_0}{\omega} \\ -\dfrac{M_1^{7}}{8}\dfrac{S_0}{\omega} & R_s+R_{c2}+R_L \end{bmatrix}\begin{bmatrix} i_1 \\ i_8 \end{bmatrix}
$$

$$
=\begin{bmatrix} R_{t1} & B \\ -C & R_{t2} \end{bmatrix}\begin{bmatrix} i_1 \\ i_8 \end{bmatrix} \tag{32}
$$

where

$$
R_{t1}=R_g+R_{c1}+R_s, \tag{33}
$$

$$
R_{t2}=R_s+R_{c2}+R_L, \tag{34}
$$

$$
B=\frac{M_1^{*7}}{8}\frac{S_0}{\omega}, \tag{35}
$$

$$
C=\frac{M_1^{7}}{8}\frac{S_0}{\omega}, \tag{36}
$$

$$
i_1=|i_1|e^{j\theta}=I_1e^{j\theta}, \tag{37}
$$

$$
i_8=|i_8|e^{j\theta}=I_8e^{j\theta}, \tag{38}
$$

$$M_1^* = \frac{I_1}{\omega(Q_c + Q_m)} e^{-j\theta}, \tag{39}$$

$$M_1 = \frac{I_1}{\omega(Q_c + Q_m)} e^{j\theta}. \tag{40}$$

Solving the matrix equation for i_1 and i_8,

$$i_1 = v_g \frac{R_{t2}}{R_{t1}R_{t2} + BC}, \tag{41}$$

$$i_8 = v_g \frac{C}{R_{t1}R_{t2} + BC}. \tag{42}$$

Dividing Equation (42) by (41) gives the phase-angle relationship;

$$i_8 = i_1 \frac{C}{R_{t2}}, \tag{43}$$

$$|i_8| e^{j\phi} = |i_1| e^{j\theta} \frac{m_1{}^7 e^{j7\theta}}{8} \frac{S_0}{\omega}. \tag{44}$$

Thus, $\phi = 8\theta$, and

$$|i_8| = |i_1| \frac{m_1{}^7}{8} \frac{S_0}{\omega}. \tag{45}$$

The equivalent circuit at the fundamental frequency is shown in Figure 6. The input impedance is

$$\begin{aligned} Z_{in} &= \frac{v_g}{i_1} - R_g = R_{t1} + \frac{BC}{R_{t2}} - R_g \\ &= R_{t1} + R_0 - R_g = \mathrm{R}_0 + R_{c1} + R_s \\ &= R_0 + R_1, \end{aligned} \tag{46}$$

where

$$R_0 = \frac{BC}{R_{t2}} = \frac{m_1{}^{14}}{64} \left(\frac{S_0}{\omega} \right)^2 \frac{1}{R_{t2}}, \tag{47}$$

and

$$R_1 = R_{c1} + R_s. \tag{48}$$

R_0 is the equivalent conversion resistance, since it is the resistance reflected into the input loop by the load resistance.

EFFICIENCY RELATIONSHIP

The power delivered to the load is

$$P_L = \frac{|i_8|^2 R_L}{2}. \tag{49}$$

The power available from the generator is

$$P_{AV} = \frac{|v_g|^2}{8R_g}. \tag{50}$$

Equation (50) assumes an input matched condition, so that

$$R_g = R_{in} = R_0 + R_1. \tag{51}$$

The efficiency, η, is defined as the ratio of the power in the load to the available power from the generator;

$$\begin{aligned}
\eta = \frac{P_L}{P_{AV}} &= \frac{|i_8|^2}{|v_g|^2} 4R_g R_L \\
&= \frac{4R_g R_0}{(R_{t1} + R_0)^2} \frac{R_L}{R_{t2}} \\
&= \frac{4R_g R_0}{(R_g + R_1 + R_0)^2} \frac{R_L}{R_{t2}} \\
&= \frac{4R_g (R_1 + R_0)}{(R_g + R_1 + R_0)^2} \frac{R_0}{R_0 + R_1} \frac{R_L}{R_{t2}}.
\end{aligned} \tag{52}$$

For input matched condition,

$$\eta = \frac{R_0}{R_0 + R_1} \frac{R_L}{R_L + R_2}$$

$$= \frac{1}{1 + \dfrac{R_1}{R_0}} \, \frac{1}{1 + \dfrac{R_2}{R_L}}, \tag{53}$$

where

$$R_2 = R_s + R_{c2}. \tag{54}$$

Efficiency Optimization

The condition for excitation may be determined by combining Equations (7) and (9);

$$V_c + V = (1 - \gamma) S_m \frac{(Q_c + Q)^{1/(1-\gamma)}}{(Q_c + Q_m)^{\gamma/(1-\gamma)}}. \tag{55}$$

When the varactor is fully pumped, S/S_m reaches a maximum value of unity. Therefore, from Equation (2), with $\gamma = 0.875$,

$$\left[\frac{Q_c + Q}{Q_c + Q_m} \right]^7 = 1, \quad \text{or} \quad \frac{Q_c + Q}{Q_c + Q_m} = 1. \tag{56}$$

Since $Q = q_0 + q$, and $q_0 = (Q_m - Q_c)/2$,

$$1 = \frac{1}{2} + m_1 \sin(\omega t + \theta) + m_8 \sin(8\omega t + \phi). \tag{57}$$

The optimum value of R_L is found by setting $d\eta/dR_L = 0$; then

$$R_L = \sqrt{\frac{R_2}{R_1} (K + R_1 R_2)}, \tag{58}$$

where

$$K = \frac{{m_1}^{14}}{64} \left(\frac{S_0}{\omega} \right)^2. \tag{59}$$

Here, m_1 was considered to be independent of R_L. While, strictly speaking, this may not be true, it can be shown to be a very good approximation for the idler-less, high-order multiplier.

Average Fundamental Input Power to an Idler-less, High-Order Frequency Multiplier

Considering the equivalent circuit at the fundamental frequency,

Figure 6, the generator voltage, $v_g(t)$, is related to the input current by the expression

$$v_g(t) = i_1(t)\ [R_g + R_{c1} + R_s + R_0]. \tag{60}$$

For a matched condition,

$$v_g(t) = 2i_1(t)\ [R_1 + R_0]. \tag{61}$$

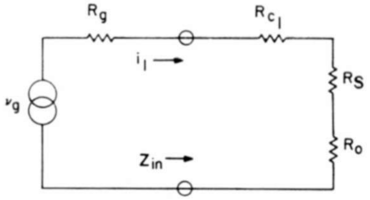


Fig. 6—Equivalent circuit at the fundamental frequency: R_0 = equivalent conversion resistance.

The instantaneous power input to the multiplier is given by

$$p_1(t) = \frac{1}{2} i_1(t) v_g(t), \tag{62}$$

or

$$p_1(t) = i_1{}^2(t)\ [R_1 + R_0]. \tag{63}$$

Since

$$i_1(t) = I_1 \cos\ (\omega t + \theta), \tag{64}$$

and using Equations (20) and (47), Equation (63) becomes

$$p_1(t) = \left\{ m_1{}^2\omega^2(Q_c + Q_m)^2R_1 + \frac{m_1{}^{16}}{64} S_0{}^2 \frac{(Q_c + Q_m)^2}{R_2 + R_L} \right\} \cos^2\ (\omega t + \theta). \tag{65}$$

The average power is then

$$<P_1> = \frac{\omega}{2\pi} \int_{t=0}^{t=2\pi/\omega} p_1(t)\, dt, \tag{66}$$

or

$$<P_1> = \left\{ \frac{m_1{}^2}{2\pi} \omega^3 (Q_c + Q_m)^2 R_1 + \frac{m_1{}^{16}\omega}{128\pi} S_0{}^2 \frac{(Q_c + Q_m)^2}{R_2 + R_L} \right\} \int_0^{2\pi/\omega} \cos^2 (\omega t + \theta)\, dt$$

$$= \frac{1}{2} \left\{ m_1{}^2 \omega^2 (Q_c + Q_m)^2 R_1 + \frac{m_1{}^{16}}{64} S_0{}^2 \frac{(Q_c + Q_m)^2}{R_2 + R_L} \right\} \qquad (67)$$

Using Equations (7), (16), (58), and (59)

$$<P_1> = \frac{1}{2} m_1{}^2 \omega^2 \frac{(V_C + V_B)^2}{S_m{}^2} [\Gamma + 1]^2 \left[R_1 + \frac{K}{R_2 + \sqrt{\dfrac{R_2}{R_1} (K + R_1 R_2)}} \right] \qquad (68)$$

“REGENERALIZATION” OF THE MULTIPLIER EQUATIONS

The effect of the nonlinearity coefficient, γ, on the efficiency and power-handling capability of the multiplier can be shown explicitly by reintroducing γ into the equations just derived. This will be done empirically, and the results will apply only to the eight times, idler-less multiplier using a varactor of arbitrary γ.

The variable R_L is a function of K (Equation (58)) and

$$R_0 = \frac{K}{R_2 + R_L}. \qquad (69)$$

However, K as defined by Equation (59) is for $\gamma = 0.875$. Consideration of Equations (26), (27), (28), (30), and (32) reveals that the factor 1/64 results from the coefficient of the eighth-order term in the general Taylor expansion, Equation (23). Hence,

$$\tilde{K} = \left[\frac{\Gamma(\Gamma - 1)(\Gamma - 2) \cdots (\Gamma - 6)}{8!} \right]^2 m_1{}^{14} \left(\frac{S_0}{\omega} \right)^2, \qquad (70)$$

where the tilde denotes a “regeneralized” term.

It should be noted here that if $\Gamma > 7$ ($\gamma > 0.875$), the Taylor series will not terminate in the eighth-order term, and there will be some contributions to the eighth harmonic from the additional terms. If these contributions should be too large to be neglected, then the equations for efficiency and power must be derived by taking these terms into account. However, for γ close to 0.875, these contributions are very small and, therefore, can be neglected.

Using the regeneralized $\tilde{K}$, one can rewrite the multiplier equations:

$$\tilde{R}_0 = \frac{\tilde{K}}{R_2 + \tilde{R}_L} = \left[\frac{\Gamma(\Gamma-1)(\Gamma-2)\cdots(\Gamma-6)}{8!} \right]^2 m_1^{14} \left(\frac{S_0}{\omega} \right)^2 \frac{1}{R_2 + \tilde{R}_L}, \tag{71}$$

and

$$\tilde{R}_L = \sqrt{\frac{R_2}{R_1} (\tilde{K} + R_1 R_2)} \tag{72}$$

or

$$\tilde{R}_0 = \frac{\tilde{K}}{R_2 + \sqrt{\frac{R_2}{R_1} (\tilde{K} + R_1 R_2)}}. \tag{73}$$

The efficiency then becomes

$$\tilde{\eta} = \frac{1}{\left(1 + \frac{R_2}{\tilde{R}_L}\right)} \frac{1}{\left(1 + \frac{R_1}{\tilde{R}_0}\right)}, \tag{74}$$

and the expression for the average fundamental input power becomes

$$<\tilde{P}_1> = \frac{1}{2} m_1^2 \omega \frac{(V_C + V_B)^2}{S_m^2} [\Gamma + 1]^2 \left\{ R_1 + \frac{\tilde{K}}{R_2 + \sqrt{\frac{R_2}{R_1} (\tilde{K} + R_1 R_2)}} \right\} \tag{75}$$

or

$$<\tilde{P}_1> = \frac{1}{2} m_1^2 \omega \frac{(V_C + V_B)^2}{S_m^2} [\Gamma + 1]^2 (R_1 + \tilde{R}_0). \tag{76}$$

Since R_1 and R_2 are, in practice, on the order of several ohms, R_1 and $R_2 > 1$, and $\tilde{R}_0$ and $\tilde{R}_L$ increase as $\tilde{K}$ increases. Therefore, the efficiency, $\tilde{\eta}$, and the maximum power input $<\tilde{P}_1>$ increase as $\tilde{K}$ increases. This, in turn, means that these quantities increase as γ increases toward unity. $\tilde{K}$ increases rapidly as $\gamma \rightarrow 1$; for example, $\Gamma = 49$ and $\tilde{K} = 6.25 \times 10^{10} m_1^{14} (S_0/\omega)^2$ when $\gamma = 0.98$.

SOME DESIGN CONSIDERATIONS

Practically speaking, in choosing a varactor for a given high-order multiplier application, several of the varactor parameters will be determined by what is available from diode manufacturers. The predetermined parameters, in general, will be:

(1) The nonlinearity coefficient, γ. The foregoing analysis indicates that the use of a varactor with γ as close to unity as possible gives maximum efficiency and maximum power-handling capability.

(2) The spreading resistance, R_s. R_s is determined by the materials used and the manufacturing process employed. One should choose a varactor with the lowest R_s for a given γ, since R_s is responsible for power loss. But, in general, a choice here will probably not be available.

(3) The cutoff frequency at an arbitrary low reverse voltage, $-V_s$ (generally -6 volts), $f_c\ (-V_s)$. This quantity is related to R_s by the relationship

$$f_c(-V_s) = \frac{1}{2\pi R_s C(-V_s)}, \tag{77}$$

which indicates that f_c should be chosen to be as large as is obtainable for minimum R_s.

(4) The contact potential, V_c. This parameter is determined by the materials used.

Once these parameters are known, the remaining two varactor parameters, namely the junction capacitance and the breakdown voltage, can be calculated:

(1) The junction capacitance at an arbitrary low reverse voltage, $-V_s$ (usually -6 volts), $C(-V_s)$. This parameter can be calculated from Equation (77).

(2) The required breakdown voltage, V_B, is calculated from Equation (76). If the multiplier circuit losses are neglected, the first term of Equation (76) represents the average power lost in the varactor spreading resistance. This represents the power dissipated by the varactor, and $<\tilde{P}_1>$ will be a maximum when the power lost in R_s equals the manufacturer's rated maximum dissipation. Hence,

$$<P_{\text{diss}}> = \frac{1}{2}\, m_1{}^2\omega\, \frac{(V_C + V_B)^2}{S_m{}^2}\, [\Gamma + 1]^2\, R_s. \tag{78}$$

All the terms in Equation (78) are known except S_m and V_B, which are related by Equation (1),

$$S_m = \frac{1}{C(-V_s)} \left(\frac{V_C + V_B}{V_C + V_S} \right)^{(\Gamma+1)/2} . \tag{79}$$

Substituting into Equation (78), and rearranging terms,

$$V_B = \left[\frac{2 \langle P_{\text{diss}} \rangle}{m_1{}^2 \omega [C(-V_s)]^2 (V_c + V_s)^{2\Gamma/(\Gamma+1)} (\Gamma + 1)^2 R_s} \right]^{(\Gamma+1)/2} - V_c, \tag{80}$$

where, with reference to Equation (57) for the fully pumped case, $m_1 \doteq 1/2$.

One additional parameter, a circuit parameter, remains to be found —the optimum applied bias voltage. This is found by setting $Q = q_0 = (Q_m - Q_c)/2$ in Equation (5)

$$\frac{Q_c + Q_m}{2} = \frac{(V_c + V_B)^\gamma}{(1 - \gamma) S_m} (V_C + V_0)^{1-\gamma}, \tag{81}$$

or

$$V_0 = \left[\frac{Q_c + Q_m}{2} \frac{(1 - \gamma) S_m}{(V_c + V_B)^\gamma} \right]^{1/(1-\gamma)} - V_c, \tag{82}$$

where V_0 is the applied d-c bias voltage.

NUMERICAL AND EXPERIMENTAL RESULTS

Figures 7 and 8 show the calculated efficiency of an eight-times multiplier as a function of γ for various values of the circuit parameters. The calculations were based on the following values:

$$S_0 = \frac{1}{1.0 \times 10^{-12}} f_d{}^{-1}$$

$f_c(-6 \text{ volts}) = 100$ gigacycles

$$R_s = \frac{10}{2\pi} = 1.59 \text{ ohms}$$

$$m_1 \leq \frac{1}{2}$$

$R_1 = R_2$

Input frequency = 500 megacycles.

An experimental eight-times idler-less multiplier employing a hyperabrupt varactor, having an input frequency of 500 megacycles, was constructed and tested. The hyperabrupt varactor used has the characteristics shown in Figure 9. An efficiency of 25% was obtained at an input power level of 250 milliwatts. Power output at 4 gigacycles was 62 milliwatts. It can be seen from Figures 7 and 8 that with the proper nonlinearity coefficient and drive level, an efficiency of 55% is possible for an eight-times multiplier without idlers.

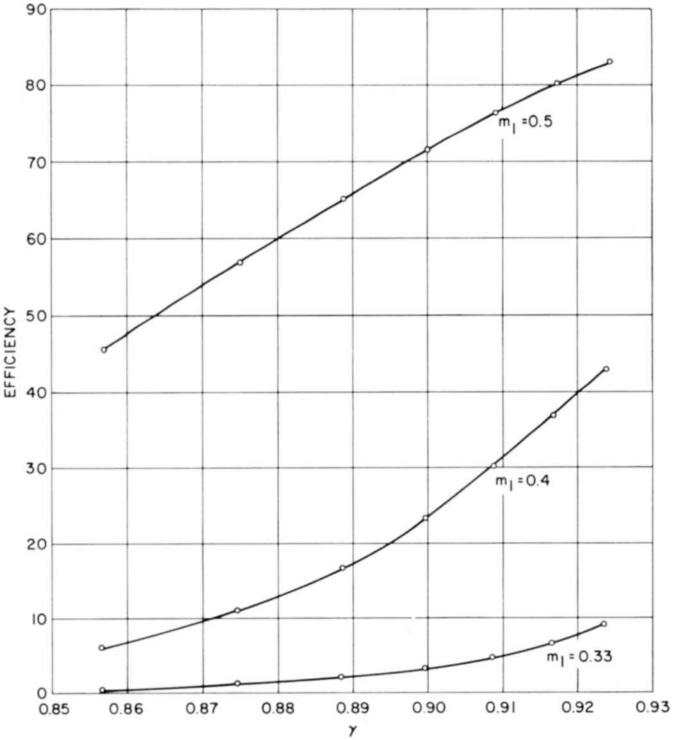


Fig. 7—Efficiency as a function of $\gamma (R_1 = R_2 = 5$ ohms).

Power tests as well as efficiency tests have been performed on the hyperabrupt octupler. The power test was to determine the input power level at which diode burnout occurs. A relatively poor quality diode was chosen for this purpose. Hence, the efficiency obtainable was less than that for the better quality diodes.

For the efficiency test, the relatively good diode used has the following specifications:

$$V_b = 25 \text{ volts}$$

$$f_c(-6 \text{ volts}) = 105 \text{ gigacycles}$$

$$C_j(-6 \text{ volts}) = 0.48 \text{ pf.}$$

For the power test, the above varactor was replaced by one having the following specifications:

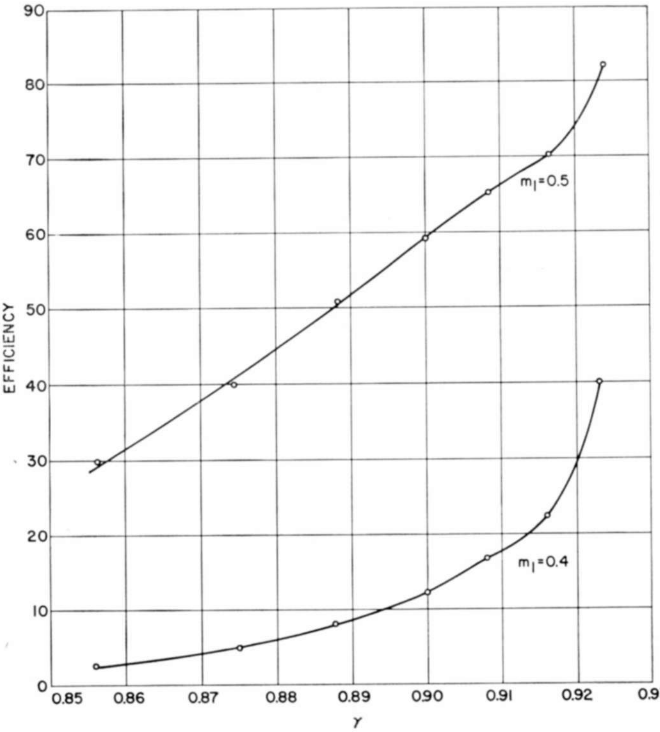


Fig. 8—Efficiency as a function of $\gamma (R_1 = R_2 = 8$ ohms).

$$V_b = 52 \text{ volts}$$

$$f_c(-6 \text{ volts}) = 27 \text{ gigacycles}$$

$$C_j(-6 \text{ volts}) = 1.3 \text{ pf.}$$

It was found that after replacing diodes, the multiplier efficiency was maximized without any retuning of the circuit other than a slight change in d-c bias, V_0. As the input power was increased from 250

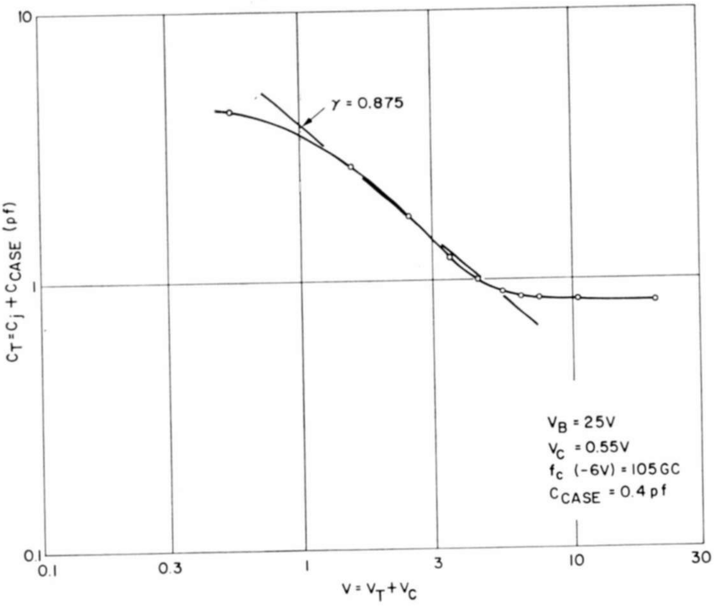


Fig. 9—Characteristics of hyperabrupt varactor used in experimental eight-times idler-less multiplier.

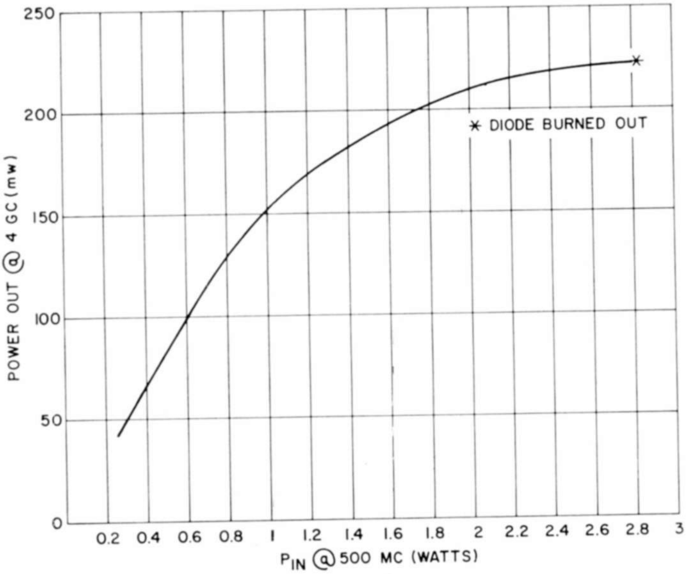


Fig. 10—Output power versus input power curve for experimental eight-times multiplier.

milliwatts, it was also found that to obtain maximum power output at each input power level, only the d-c bias had to be changed. The output power as a function of input power is plotted in Figure 10.

CONCLUSIONS

The generation of broadband or tunable microwave power can be achieved by high-order frequency multipliers with hyperabrupt varactors. High efficiency can be obtained without the use of idler circuits, thus simplifying the circuits considerably. Efficiencies as high as 55% can be achieved for a times-eight multiplier with an input frequency of 500 megacycles and an output frequency of 4 gigacycles if the diode nonlinearity, γ, is 0.875 over the entire range of operation and the diode is fully pumped. Experimental results have yielded a 25% efficiency with the diode nonlinearity, γ, of 0.875 over portions of the operating range. The input power-handling capability is in excess of 2.5 watts and the efficiency remains essentially constant over the range of input power of 0.2 to 1.0 watt. The tuning of the device is extremely simple since only the input circuit and the output circuit need be adjusted. Furthermore, the multiplier can be made broadband since there are no idlers to degrade its bandwidth.

The experimental results described in this paper were obtained using reasonably high-Q varactors. One of these units has a cutoff frequency of 105 gigacycles at —6 volts. While this value is below that of commonly available conventional varactors, it does not pose a serious limitation until the output frequency of the hyperabrupt multiplier is extended well into X-band. It is also not unlikely that further advances in hyperabrupt junction fabrication will yield still higher cutoff frequencies (or Q's).

To summarize, the use of a hyperabrupt varactor in frequency multipliers yields the following advantages:

(1) high efficiency for high order multiplication without the use of idlers,

(2) ease of tuning due to the simplicity of the circuit (only an input frequency loop and an output frequency loop are required),

(3) circuit simplicity makes broadbanding easier to achieve,

(4) higher power handling capability as compared to abrupt varactors with the same breakdown voltages,

(5) greater capability for handling wide variations in input power without requiring retuning.

BIBLIOGRAPHY

P. Penfield, Jr., and R. P. Rafuse, *Varactor Applications,* The MIT Press, 1962.

T. C. Leonard, "Prediction of Power & Efficiency of Frequency Doublers Using Varactors Exhibiting a General Nonlinearity," *Proc. I.E.E.E.,* Vol. 51, p. 1135, Aug. 1963.

P. Canick and S. Yuan, "T-Dependence of Varactor Frequency Doublers," *Proc. IRE,* Correspondence, Vol. 50, p. 1533, June 1962.

J. J. Chang, J. H. Foster and R. M. Ryder, "Semiconductor Junction Varactors with High Voltage-Sensitivity," *I.E.E.E. Trans. on Electron Devices,* Vol. ED-10, No. 4, p. 281, July 1963.

PROJECT RANGER* TELEVISION SYSTEM

By

B. P. Miller

RCA Astro-Electronics Division
Princeton, N. J.

Summary—*The function of the ranger television system was to transmit close-up pictures of the lunar surface to earth. The philosophy, design, and performance of the television system are described. The television system consisted basically of six slow-scan vidicon cameras whose outputs were used to frequency-modulate two sixty-watt transmitters. The two signals, approximately one megacycle apart, were radiated by a high-gain parabolic antenna. The ground receiving station, located at Goldstone, California, used an 85-foot parabolic antenna. Several high-definition photographs of the lunar surface are included to provide a complete description of the performance of the system.*

Introduction

THE Ranger program, which was sponsored by the National Aeronautics and Space Administration (NASA), was the first of a series of projects, the eventual goal of which is to land a man on the moon. The specific purpose of Ranger was to obtain close-up pictures of the lunar surface in order to gain information regarding the surface conditions on the moon. The pictures obtained were to be at least an order of magnitude better in resolution than the best available by earth-based photography. Such information is an essential prerequisite for future projects involving "soft" landings.

The Ranger vehicles were designed to transmit television pictures of the lunar surface in the final few minutes before crashing into the moon. Three such vehicles were successfully flown; they were designated Ranger VII, VIII, and IX. More than 17,000 high-quality photographs of the lunar surface have been obtained from the three vehicles. The picture-taking phase of Ranger has now been completed; processing and interpretation of the information received will continue for some time.

The Spacecraft

The Ranger spacecraft, shown in Figure 2, consists of two major subassemblies—the television system and the spacecraft bus. The bus provides a space-stabilized vehicle for the television system. The total

---

* The Ranger Program was conducted for the National Aeronautics and Space Administration (NASA) by the Jet Propulsion Laboratory (JPL), California Institute of Technology. The Astro-Electronics Division of RCA was a major subcontractor to JPL, responsible for the design, fabrication and test of the TV System.

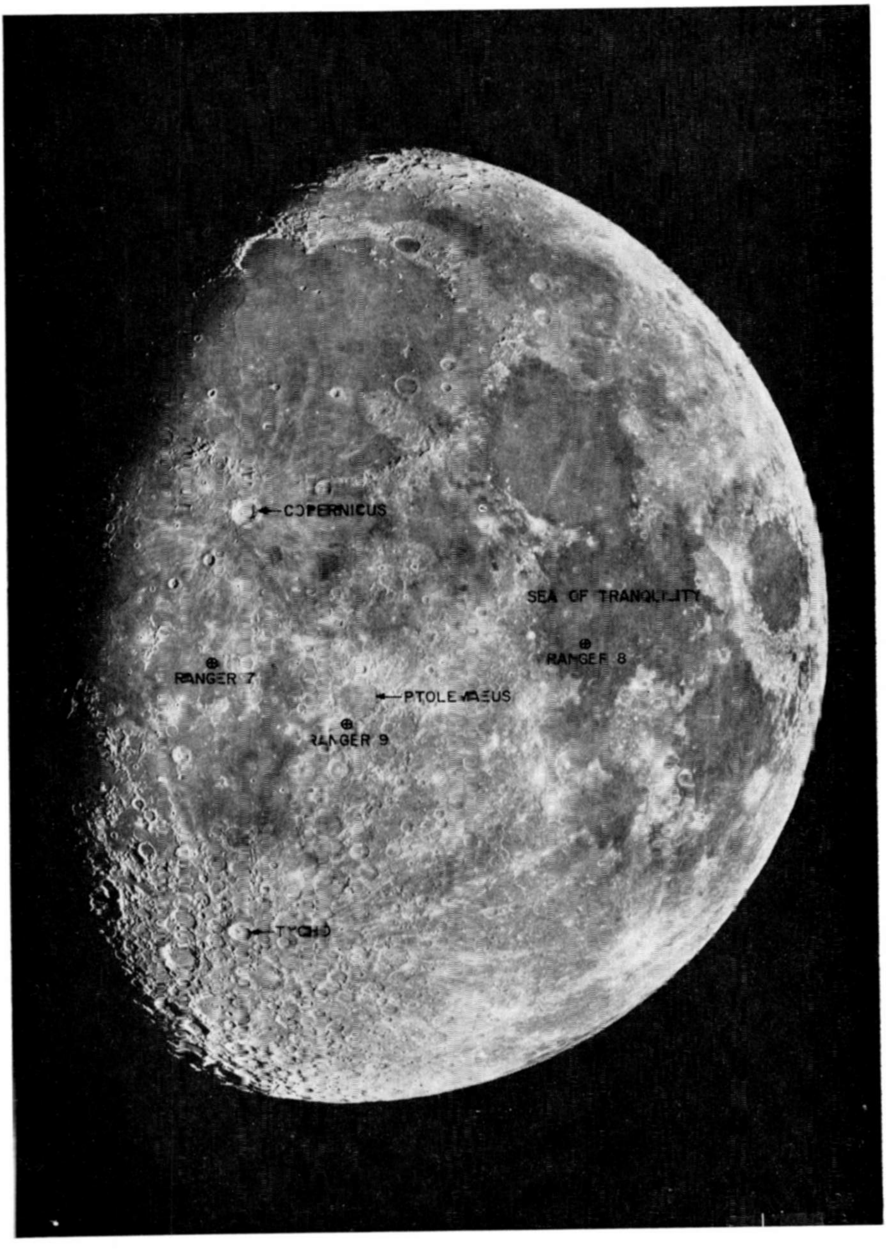


Fig. 1—Telescopic photograph of the moon taken from earth showing the impact points of Rangers VII, VIII, and IX.

spacecraft (bus and television system) weighs about 810 lb.; the weight of the television system is about 380 lb. The bus consists of several systems; these include the radio, data encoder, and command systems; the central computer and sequencer (CC&S) system; the attitude-control system; the power system; and the midcourse propulsion system. The power system, which derives its power from two solar panels when the spacecraft is oriented toward the sun and from batteries during other periods, supplies all the power required by the bus portion of the spacecraft. The attitude-control system provides for attitude stabilization during the major portion of the flight, so that the solar panels will be oriented toward the sun at the same time that the high-gain antenna of the radio system is pointed toward the earth. It also provides the capability for maneuvers to initially acquire the sun and the earth, to correct any trajectory errors during flight in

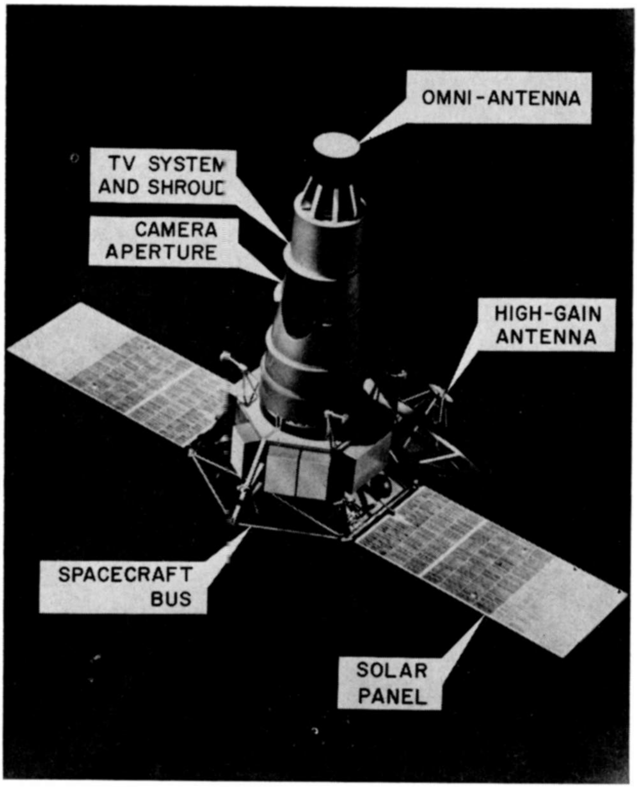


Fig. 2—Ranger spacecraft.

conjunction with the midcourse propulsion system, and to align the cameras of the television system in relation to the lunar surface prior to impacting the moon. The radio system is part of a two-way link between the earth and the spacecraft for transmission of telemetry from the data encoder system to earth, and to receive command information from earth to the spacecraft system. The high-gain antenna of the radio system is also used to transmit the processed signal outputs of the television system transmitters. The CC&S provides all of the on-board timing for the various spacecraft functions.

THE TELEVISION SYSTEM

The television system is the means by which the primary data-gathering function of the Ranger mission is performed. This self-contained unit provides the power, control, and communications equipment for collecting, processing, and transmitting the photographic information of the lunar surface. In a normal mission, the picture-taking sequence is initiated by an earth-generated command, with the first picture taken at an altitude of approximately 2000 kilometers from the lunar surface, and continues uninterrupted until lunar impact. The initial pictures cover a wide area of the moon at resolutions comparable to that obtained by earth-based telescopes. Area coverage is exchanged for increasing resolution as the spacecraft approaches impact at a terminal velocity of 2700 meters per second; resolutions of 0.5 meter, or better, are achieved in the final picture sequence.

The television system is separated into two essentially independent channels of operation to ensure a maximum probability of mission success. The channel arrangement is illustrated in Figure 3. Each channel has all of the components required to collect and process the video information and to supply a 60-watt frequency-modulated signal to the common high-gain antenna for transmission back to earth. Channel separation is on the basis of the two modes of camera operation employed—full scan and partial scan. There are two full-scan and four partial-scan cameras. Both types employ ruggedized vidicon pickup tubes with a highly sensitive photoconductive target. The basic difference between the two modes of camera operation is in the size of the area scanned on the vidicon target, and thus the time required to read out the stored image. Because of the high sensitivity of these vidicons and the image retention characteristic needed for slow-scan readout, special provision is made to erase the residual picture after frame readout and prior to a subsequent exposure so as to prevent a double image. To provide for maximum utilization of the communications equipment, multiple cameras, operating during each other's erase

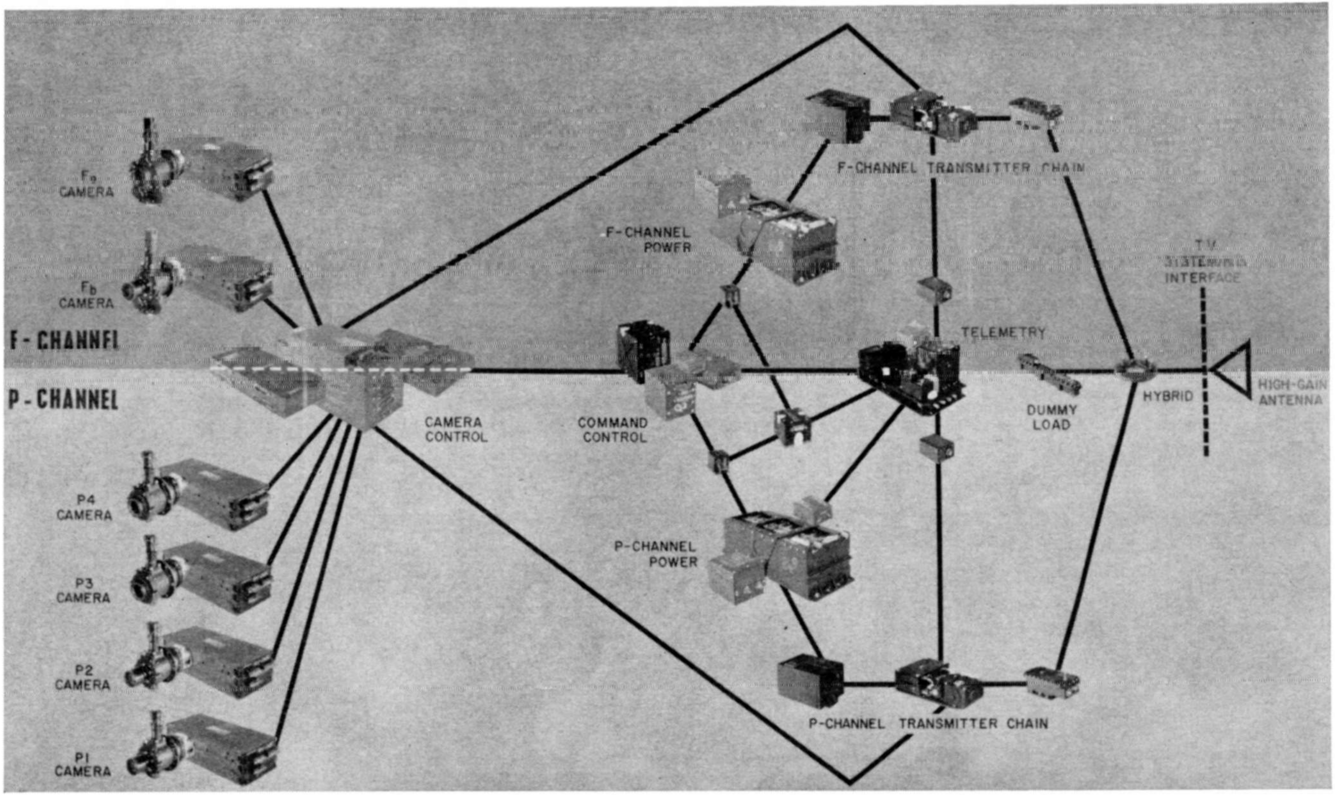


Fig. 3—Channel separation of the Ranger television system.

cycles, are used to fill out what would otherwise be nonproductive portions of the available transmission time.

In order to satisfy the requirements of high resolution, it was desirable that the final sequence of pictures be taken from a minimum altitude above the lunar surface. The minimum altitude for the final pictures is determined by the time required to read out the video data from the camera and by the spacecraft terminal velocity. Minimum altitude, and thus highest resolution with a given optical system and camera design, is therefore obtained by minimizing the camera readout time. However, the minimum readout time is also constrained by the number of scanning lines (or line density) required to provide the desired picture resolution, and the permissible system video bandwidth, which in this case was fixed at a maximum of 200 kc for each channel. In the Ranger television system, high line density and minimum readout time were achieved by the use of a special method of raster formation with a high-resolution, slow-scan vidicon camera tube. Minimum readout time is achieved in the four partial-scan cameras by scanning only the central 282 lines of a nominal 1125 line raster, enabling a 0.2-second frame time with a corresponding altitude of less than 600 meters for the exposure and transmission of the last complete frame. A set of four partial-scan cameras (designated P1, P2, P3, and P4) are exposed and read out sequentially at 0.2-second intervals to ensure continuous coverage until impact with the partial-scan channel, based on a complete picture cycle time (exposure, read, erase, and prepare) of 0.8 second for one camera.

The area of the vidicon target scanned in the partial-scan cameras measures 0.11 by 0.11 inch. The P1 and P2 cameras are equipped with 76-mm f/2.0, narrow-angle lenses to provide fields of view of 2.1 degrees, and the P3 and P4 cameras with 25 mm f/0.95, wide-angle lenses to provide fields of view of 6.3 degrees. The sequence of operation is P1, P3, P2, and P4, so that photographs are alternately taken by a narrow-angle lens and wide-angle lens. The combination of the optics, reduced scan area, and minimum readout time provides the capability to achieve a resolution of 0.5 meter per optical line pair in the final sequence of pictures.

The television system also contains two cameras that utilize the full 1125 line raster with a correspondingly longer readout time to provide coverage of wider areas of the lunar surface at a level of resolution that is sufficient to locate the data observed in the final high-resolution pictures. These cameras, designated F_a and F_b, are exposed and read out sequentially at 2.5-second intervals, and scan a 0.44-inch-square area. The F_a-camera is equipped with a 25-mm f/0.95, wide-angle lens to provide a 25-degree field of view, while the F_b-camera has a 76-mm f/2.0, narrow-angle lens to provide an 8.4-degree field. One reason for having several cameras with different lens apertures is that prior to the flight of Ranger the lighting conditions on the moon could not be precisely determined from earth. The different lenses employed provide greater exposure latitude, allowing picture taking over a wide range of lighting conditions. The range of lunar lighting conditions covered by the lenses and the dynamic range of the vidicons was set for 30 to 2600 foot-lamberts during the flight of Ranger VII, which roughly corresponds to lighting conditions on earth (on an average day) from noon to dusk. On the basis of the Ranger VII flight data, the upper limit of the camera dynamic range corresponding to the peak video signal was reduced to 1600 foot-lamberts for Rangers VIII and IX. The full-scan cameras provide pictures with slightly different pointing vectors to achieve a partially overlapping pattern. The four partial-scan cameras are arranged with slightly divergent pointing angles to provide a series of photographs whose coverage will result in a nesting pattern around a common optical axis as shown in Figure 4. This figure shows the theoretical pointing angles and coverage for each of the six cameras. Prior to the launch, an accurate determination is made of the actual pointing angles and camera overlap so that the relation of each picture received from the spacecraft during the actual mission will be known. A special camera array alignment procedure was used for this purpose. The television system, with the cameras in flight configuration, was positioned with respect to a stationary target to within 1 minute of arc using an autocollimating theodolite. Pictures were taken of the target with each of the cameras, and the video signal was processed and displayed on a kinescope. The resultant image was then photographed with a 35-mm film camera and the film developed for viewing and analysis. The actual pointing angle for each camera was then obtained by measuring the reticle position and picture edge on the developed film with respect to markings on the target.

Each television camera employs a magnetically focused and deflected vidicon tube having a special photoconductive surface with excellent sensitivity, high retentivity, and good erasure characteristics. The sensitivity of the vidicon used in the Ranger cameras is similar to photographic film with an ASA rating of 5 to 20 (DIN of 8 to 14). Each vidicon is capable of supplying a signal current greater than 10 nanoamperes at an exposure of 0.3 foot-candle second. The peak-to-peak residual signal after erasure is less than 5 percent of the peak-to-peak signal before erasure. The light transfer characteristic (gamma) is shown in Figure 5 for typical Ranger cameras. A sample

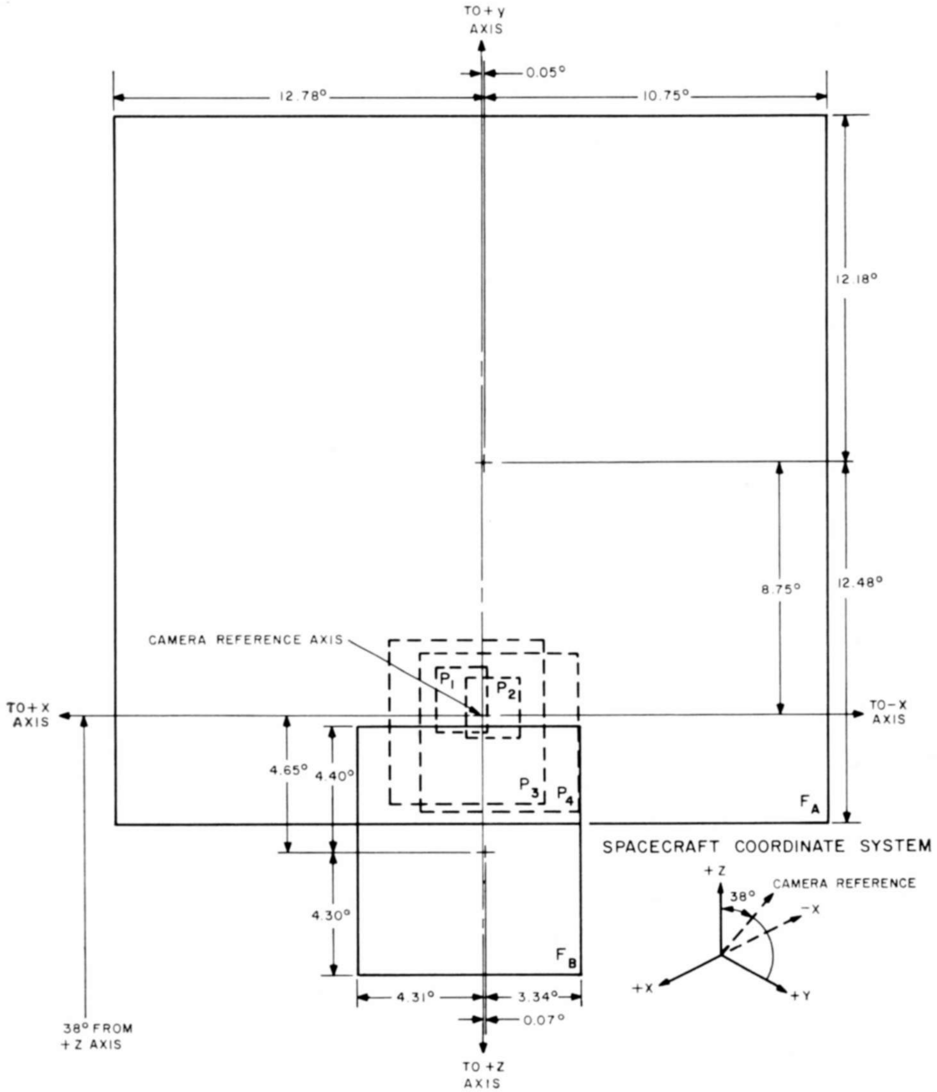


Fig. 4—Picture nesting pattern provided by the Ranger television cameras.

of the horizontal square-wave response curve for both a full-scan and a partial-scan camera is shown in Figure 6. The spectral response of the tube compared to the human eye is shown in Figure 7. In operation, the photoconductive surface is prepared for exposure by uniformly charging the surface with the electron scanning beam. Optical exposure modifies the stored charge in accordance with the photoconductive resistance changes induced by the image. The electron beam then

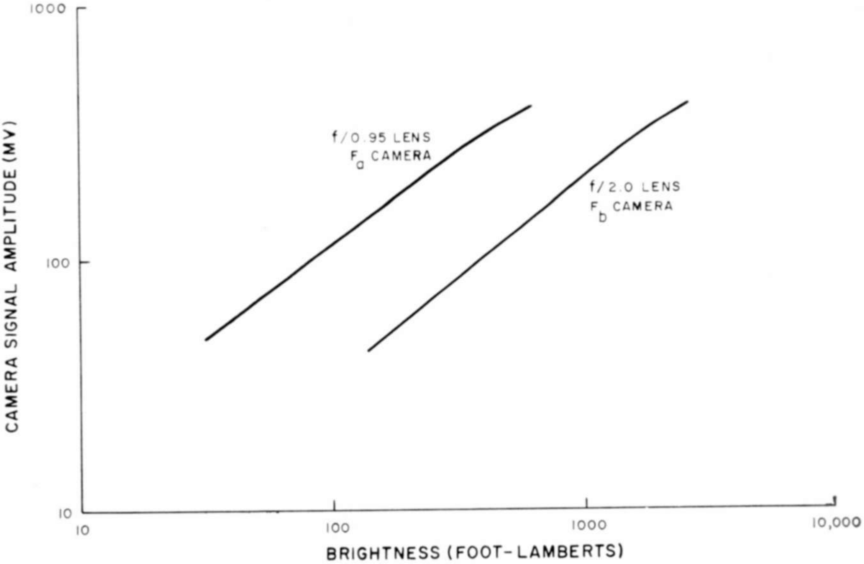


Fig. 5—Sample light transfer characteristics of Ranger cameras.

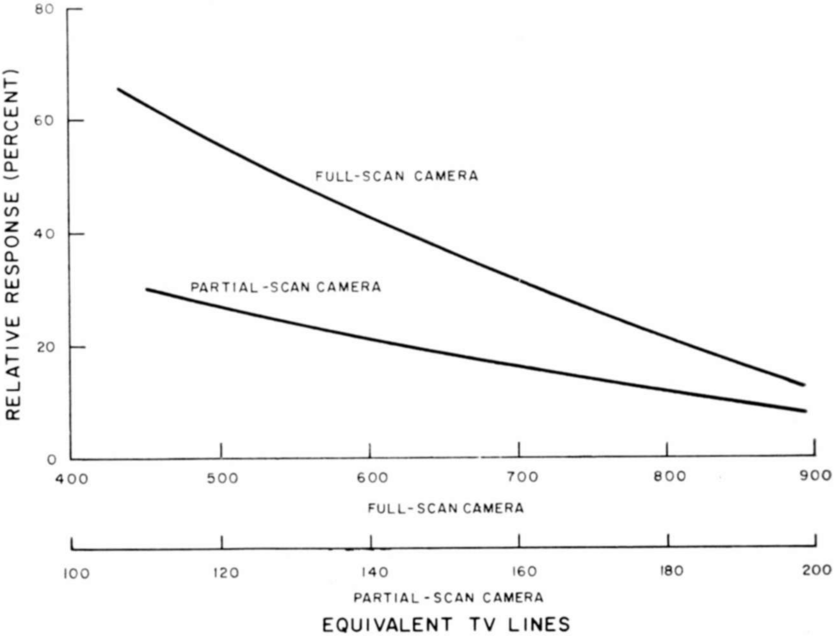


Fig. 6—Sample horizontal square wave response of Ranger cameras.

scans the surface. The video signal is dependent upon the discharge that occurs during exposure. After the picture has been read out, the target surface is erased to eliminate the residual image.

Each camera is provided with a metal focal-plane shutter. This shutter is a solenoid-operated, sliding-aperture type that moves from one side of the lens to the other each time a picture is taken. The moving blade is located as close to the focal plane as possible. The shutter is designed to provide the full-scan cameras with a 4-millisecond exposure and the P-cameras with an exposure of 2 milliseconds to minimize image motion. The faster sequencing of the P-cameras enables photographs from a much lower altitude than is possible with the F-cameras, so that image motion becomes a critical factor in terminal resolution. The shutter is a unique high-reliability design with a demonstrated capability of 1 million operations—many more than could possibly occur during the testing and flight of a Ranger spacecraft.

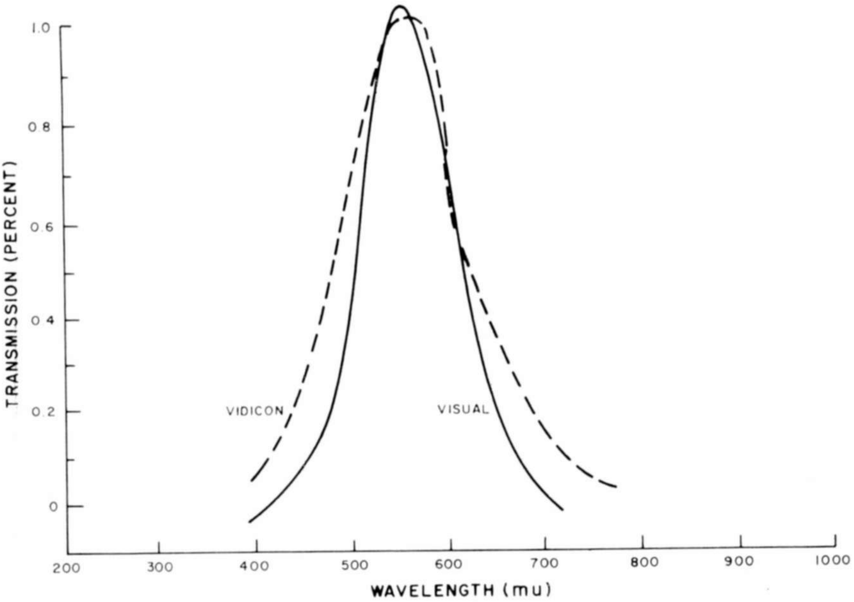


Fig. 7—Spectral response of Ranger vidicon compared to human eye.

The F_a and F_b cameras along with their associated electronic circuits, sequencing circuits, video combiner, a control circuit, a power-distribution circuit and transmitter comprise the F-channel of the Ranger television system. The P-channel contains independent circuits, identical to the F-channel, to control, process, and transmit the video information obtained by the four partial-scan television cameras. In addition, the P-channel contains the secondary synchronizing and electronic circuits necessary to provide the P1 camera with the capability of a "free-running" mode of operation. This capability would permit the P1 camera to operate in the event of a failure of the P-channel sequencer and thus ensure that the high-resolution pictures would be obtained.

Each camera of the television system has an associated unit that amplifies the video signal from the vidicon and then feeds the signal to a video combiner for processing. The operating voltages, sweep signals, and focus signals for the vidicon and yoke assembly of the associated camera are all generated in this unit. A camera sequencer circuit in each channel provides the synchronizing and sequencing signals to the cameras, and also provides gating signals to the video combiner circuit. In the video combiner the video signals from each camera are amplified, processed, and applied as a composite video signal to the transmitter.

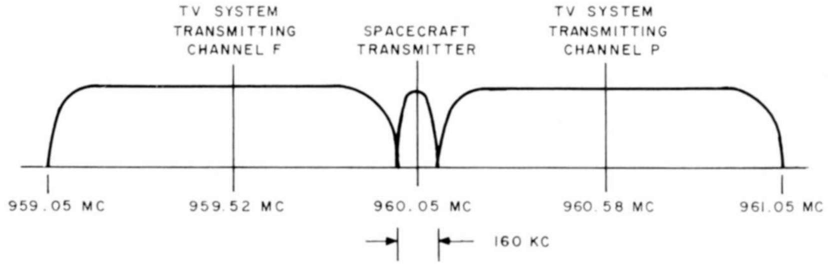


Fig. 8—Ranger spacecraft frequency allocations.

The two transmitters employed were designed to operate in the frequency band 960.05 ± 1 megacycle. The nominal center frequency of the F-channel transmitter is 959.52 mc, and that of the P-channel transmitter is 960.58 mc. The transmission spectrum is shown in Figure 8. The two transmitters are identical except for center frequency. The composite video signal that is applied to the transmitter has a 187-kc baseband. The output of the transmitter is a 60-watt frequency-modulated signal. The two 60-watt power amplifier outputs are combined in a four-port hybrid ring to provide two outputs with 30 decibels of isolation. One output is dissipated in a dummy load; the other signal is applied to the 4-foot-diameter high-gain antenna and transmitted to earth.

GROUND EQUIPMENT

The spacecraft-transmitted r-f signal is received by the 85-foot parabolic antenna at the Ranger ground station located at the Deep Space Instrumentation Facility (DSIF), at Goldstone, California. The antenna receiver system, using a maser pre-amplifier, converts the 960.05 mc f-m signal to a 30-mc, dual-channel signal and feeds it to the operational support equipment (OSE) through a 30-mc preamplifier for processing, recording, and display. A picture of the OSE is shown in Figure 9. A simplified block diagram of the OSE is shown in Figure 10. The signal is again frequency-converted to the 5-mc region in the dual-channel limiter amplifier of the OSE receiver, and separated into two individual video channels for further processing. The channel containing the full-scan video signal is centered at 4.47 mc, and the channel containing the partial-scan video signal at 5.53 mc. These 4.47- and 5.53-mc signals are then applied to detector amplifiers, and "record i-f" amplifiers.

Fig. 9—Operational support equipment at the ground station in Goldstone, California.

In the detector amplifiers, the signals are amplified, limited, detected, and de-emphasized, and then applied through a low-pass filter to the television recording and display equipment for on-line, or real-time, display and 35-mm film recording of the video. The outputs of the detector amplifiers are also applied to a 225-kc discriminator that rejects all video information from the signal and allows real-time display of the television-system telemetry data through a d-c coupled amplifier, on a strip-chart recorder.

The signals applied to each record i-f amplifier from the limiter amplifiers are frequency converted to the 0.5-mc region and made available to the television recording and display equipment for the predetection tape recording of the composite video and 225-kc telemetry data.

Television recording and display equipment provides interim storage of predetected i-f data, and reduction, display, and archival storage

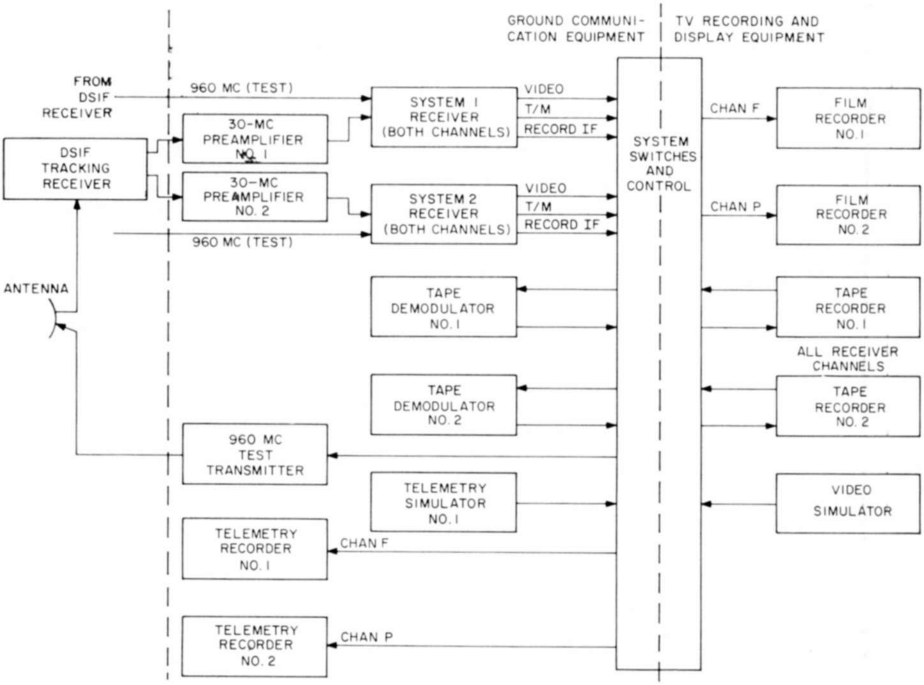


Fig. 10—Goldstone ground receiving equipment (simplified).

of the transmitted video information. Archival storage of the video data is performed automatically by 35-mm film recorders that provide photographic records of the kinescope displays of the F- and P-channel video. In addition, photographs of selected video displays are taken semi-automatically.

In the film-recorder equipment, kinescopes that utilize 5-inch cathode ray tubes (CRT) display a full frame of F_a or F_b video (Channel F), or four frames of camera P1, P2, P3, P4 video (Channel P). The kinescope for P-channel presentation displays the video frames of channel P in sequence and divided equally in area and aspect ratio in the four quadrants of the tube. A 35-mm camera then photographs

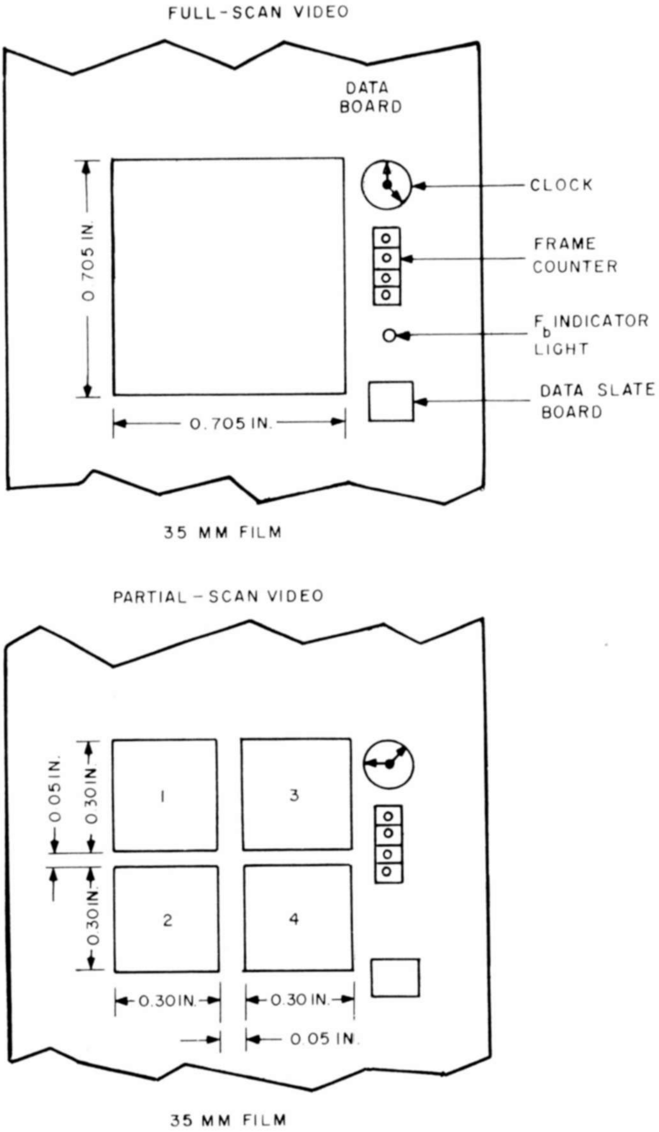


Fig. 11—Film recording format.

the tube face and a data board, which identifies the video display through use of a real-time clock, an indicator light to identify a F_b-camera display, a frame counter, and a slate board for handwritten information. Additionally, a Polaroid camera is used for single-frame, sampling photographs. The film-recording formats are illustrated in Figure 11.

The tape recording-equipment employed for predetection recording of the composite video and 225-kc telemetry signals are seven-channel, wide-band, magnetic-tape devices. The recorders operate at a tape speed of 120 inches per second, and an interconnection technique is used whereby each recorder records both full- and partial-scan video.

Additional redundancy is provided through the use of two completely separate ground stations including antennas, receiving, processing, and display equipment at the JPL Goldstone, California Deep Space Instrumentation Facility. These operational-support chains operate simultaneously during the terminal mode of the mission.

Extensive tests and calibrations of the television recording and display equipment are performed using electrically generated test standards to assure that picture detail and photometric characteristics are faithfully reproduced. Magnetic tape recordings of actual video data from various spacecraft tests are used to confirm the operating parameters of the video display equipment and the film recorders. Resolution capabilities, linearity, cathode-ray-tube display characteristics (optical distortion, sizing, etc.), and the integrity of the display are checked. Further test and calibration of the entire operational support equipment are carried out up to the time of the terminal-mode operation of the spacecraft when the video signals from the spacecraft are transmitted. Additionally, post-flight calibration and check are performed so that proper evaluation of the received data can be made, based on the actual performance of the operational support equipment during the mission.

PERFORMANCE OF THE RANGER TELEVISION SYSTEM

Ranger VII

Ranger VII was launched from Cape Kennedy, Florida at 16.50 GMT on July 28, 1964. At approximately 13.06 on July 31, the F channel was placed in warmup by a command from the television system clock. The F channel went into full power operation 80 seconds later and began taking pictures. The P channel was placed in warmup at approximately 13.09 and was placed in full-power operation 80 seconds later by a command stored in the spacecraft. Picture-taking continued for approximately 17 minutes until impact in the Mare Nubium just south-west of the crater Guericke at 13.25.50 GMT. The target area, a lunar maria, appeared from earth to represent an area typical of the smoothest parts of the lunar surface. The topological features of the lunar maria as viewed from earth presented the type of surface considered to be desirable for the "soft" landing of manned spacecraft. The actual impact area at 10.6° South and 20.6° West was outside of the zone of preferred manned landing areas of plus or minus 5 degrees of the lunar equator, but was in an area that was considered to be representative of the large, relatively smooth mare to be explored. Because the lighting conditions of the moon could not be precisely determined from earth, the gains of the F_b, P1 and P2 cameras were adjusted for a peak scene luminance of 2700 foot Lamberts and the gains of the F_a, P3 and P4 cameras were adjusted for a peak scene luminance of 650 foot Lamberts. Analysis of the video data obtained indicated that the choice of impact area for the Ranger VII spacecraft resulted in maximum use of the optical-to-video conversion capabilities of all six cameras. A total of 4308 pictures of the lunar surface were taken, transmitted, and successfully recovered on earth. The pictures have been published in atlas form,* and distributed on a world-wide basis.

The Ranger VII spacecraft and television system met or surpassed all mission goals and objectives. The level of performance of some portions of the system can be evaluated on the basis of specific information received during the mission. This is particularly true in regard to the cameras and their associated electronic equipment, since the pictures themselves are evidence of the performance of this equipment. It is also true, although to a lesser degree, of the telecommunications and OSE equipment. The performance of the command and control circuitry, however, can only be evaluated in absolute terms; that is, all commands were processed and executed on-time and in good order.

The last partial-scan picture, shown in Figure 16, was a fragment from the P3 camera, which was equipped with a 25-mm, f/0,95 lens. A simple geometrical relationship for the 25-mm optical system and for a spacecraft traveling at 2.64 km/sec shows that the final picture was taken at an altitude of 530 meters above the point of impact on the lunar surface and covered an area of about 42 meters by 28 meters. Craters can be recognized that occupy 1/50 of picture height, a diameter of 0.8 meter. The bright rim and shadowed rim constitute image elements of 0.4 meter; hence, the high-resolution goal was attained.

The requirements of wide-angle coverage and picture nesting were imposed as a basis for correlation of the expected data output from the Ranger camera with the well-defined telescopic photographs presently available. The success with which Ranger VII met these objectives is illustrated by Figure 12. This image was produced by the F_a-Camera,

---

* Ranger VII Photographs of the Moon: Part 1, Camera A Series, NASA SP-61 (Sept. 1964); Part 2, Camera B series, NASA SP-62 (Feb. 1965); and Part 3, Camera P Series (to be released).

which has a 25-degree field of view, and was taken 480 miles before impact. It covers a wide area where several well known features, such as the crater Lubiniezky and the Riphaeus mountains, can be distinguished by their characteristic images and their geographical relationships. Also indicated on this figure are the areas covered by pictures taken later in the picture-taking sequence at a higher level of resolution, including the point of impact as shown in Figures 13, 14 and 15. The realization of the design to satisfy picture nesting is contained in the interplay of camera fields of view, the camera mounting bracket that physically ties the six-camera array into appropriate

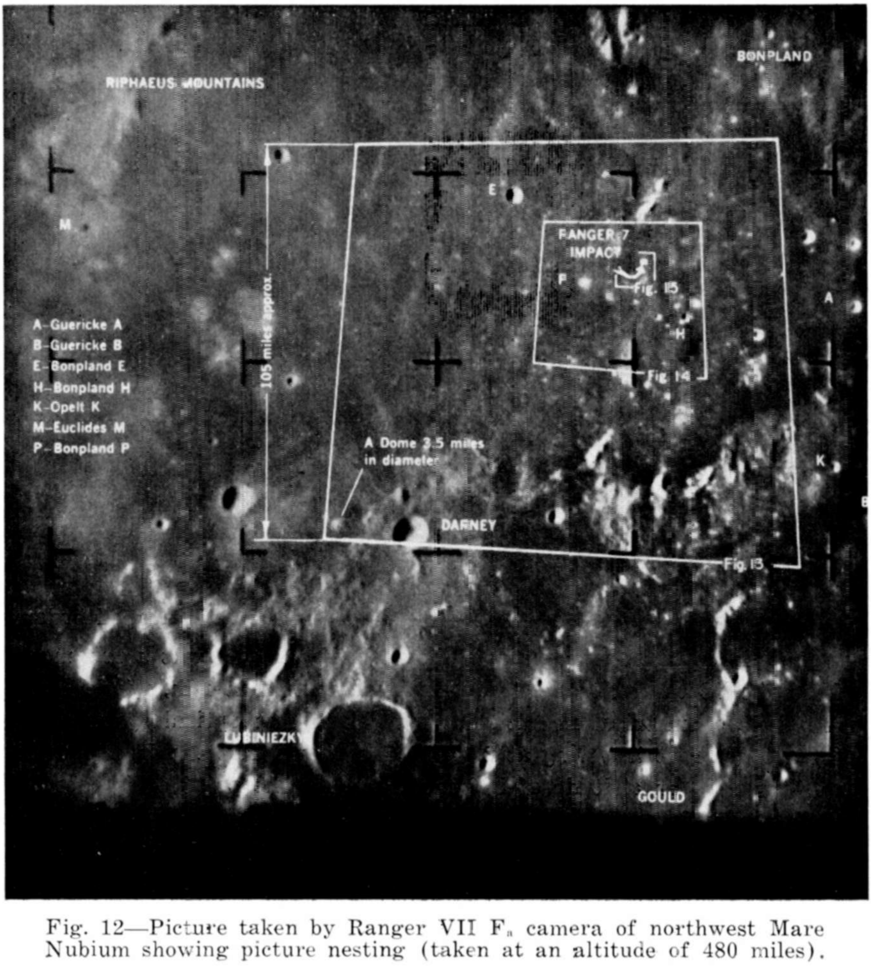


Fig. 12—Picture taken by Ranger VII F_a camera of northwest Mare Nubium showing picture nesting (taken at an altitude of 480 miles).

overlapping fields of view, the camera exposure sequence, and the design of the impacting trajectory.

Review of the received video data indicated that the choice of lunar impact area resulted in maximum use of the optical-to-video conversion capabilities of all six cameras. The video data indicated that on the three 25-mm, f/0.95 cameras the peak illuminations were occasionally near the saturation level of the cameras, but the cameras did not actually saturate nor was the signal clipped. Maximum exposure without saturation was then achieved on the F_a, P3, and P4 cameras. The F_b, P1, and P2 cameras had average exposures with signal-to-noise ratios that were 3 to 4 decibels lower than for the F_a, P3, and P4 cameras.

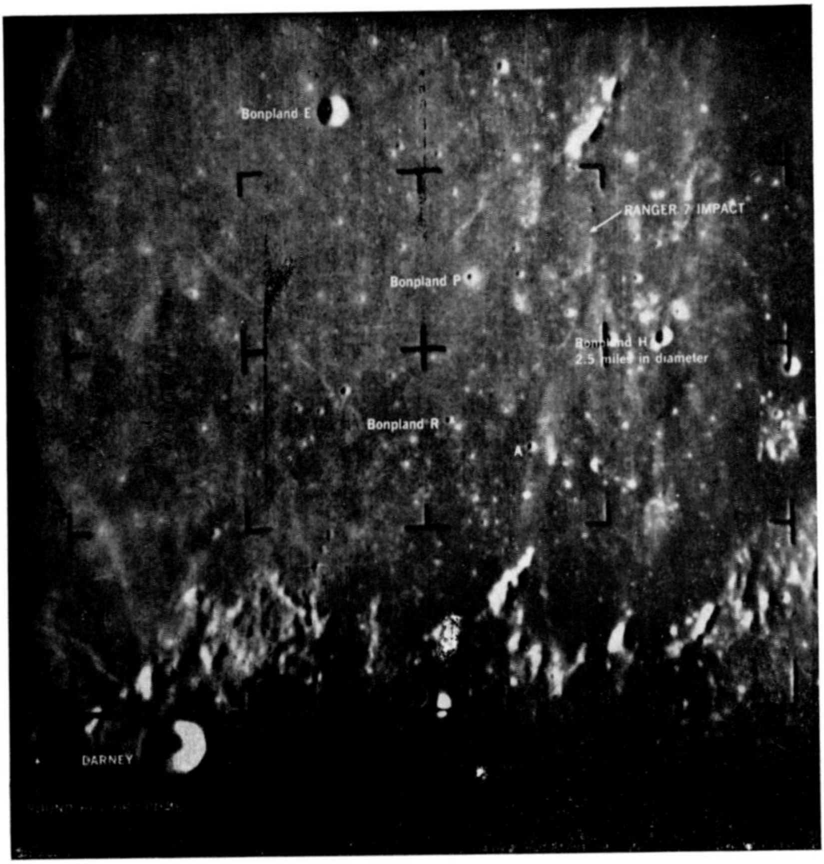


Fig. 13—F_a camera picture taken at 235 miles altitude, 2 minutes 40 seconds before impact.

Video-line selections from the final F_a and P3 camera frames show that the requirement of high signal-to-noise ratio was also satisfied. A received video signal with a signal-to-noise ratio in excess of 30 db (peak-to-peak signal to r-m-s noise) will produce a picture acceptable as having high quality. The average signal-to-noise ratio was 35 decibels for the F_a, P3, and P4 cameras, and 31 decibels for the F_b, P1, and P2 cameras. Each of the six cameras, then, provided a video signal with more-than-adequate signal-to-noise ratios to satisfy that specification and the video criterion for high-quality images.

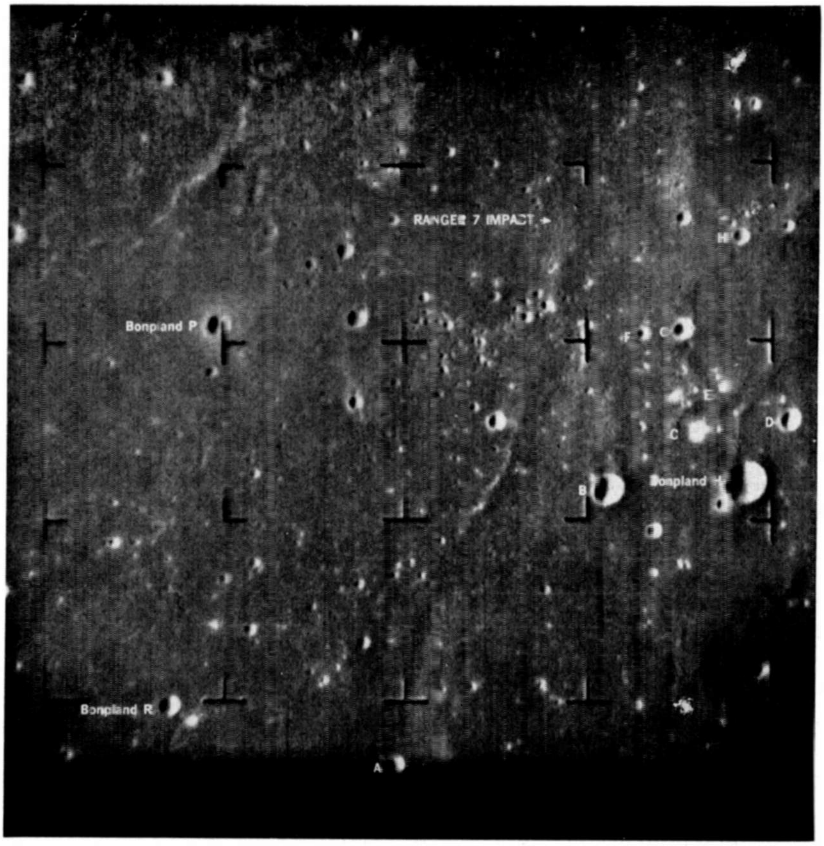


Fig. 14—F_a camera picture taken at 85 miles altitude, last of previously known craters.

Ranger VIII

The selected Ranger VII impact area, a region of lunar maria located between Oceanus Procellarum and Mare Nubium was classified by astronomers as a "red" mare. The success of Ranger VII in classifying the small-scale topography of this typical "red" mare raised questions with respect to its similarity with other maria. The February launch period chosen for Ranger VIII permitted the selection of a "blue" mare, Mare Tranquillitatis, as the target impact area. In addition, this launch period permitted the selection of a target area within the Apollo landing zone, thus providing photographs of a potential landing site.

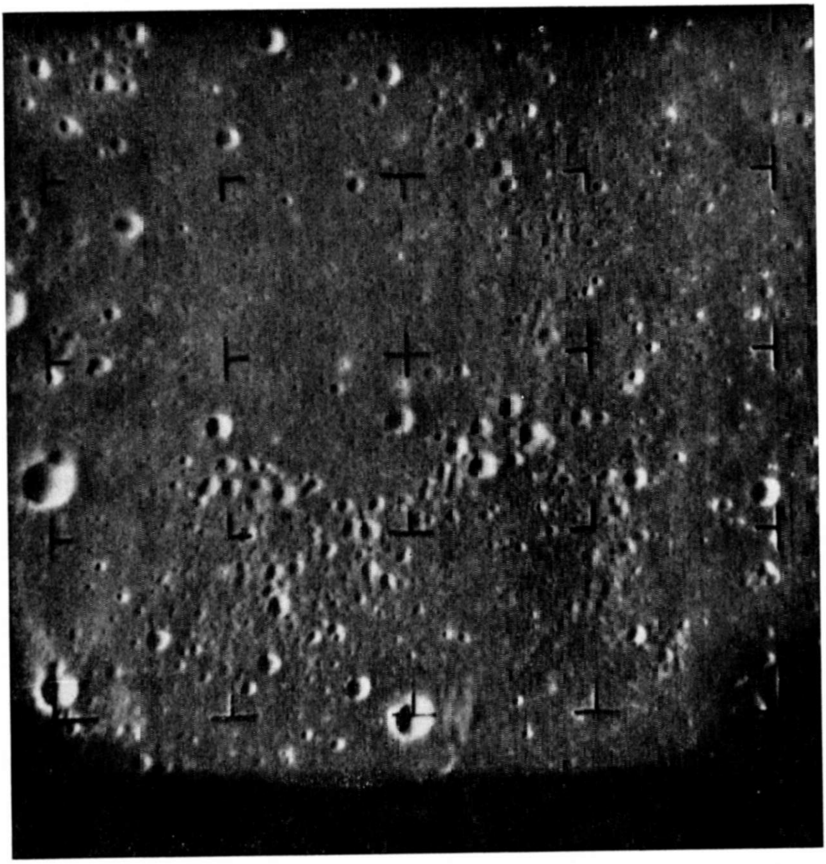


Fig. 15—F_a camera picture taken at an altitude of 34 miles showing "rays" of crater Copernicus.

The Ranger VIII spacecraft impacted on the lunar surface at 09.57.36 GMT on February 20, 1965. The spacecraft obtained 7140 pictures of the moon that covered an area extending from Mare Nubium, across the lunar highlands, to Mare Tranquillitatis. The final pictures were taken in Mare Tranquillitatis before Ranger VIII impacted at lunar co-ordinates 2.6° North and 24.8° East.

Based on an increased ability to define the anticipated levels of lunar luminance as a result of information gained from camera performance of the Ranger VII mission, the gains of the P1, P2 and F_b cameras on Ranger VIII were adjusted for a peak scene luminance of 1600 foot lamberts; the P3, P4 and F_a cameras remained at a saturation level of 650 foot lamberts.

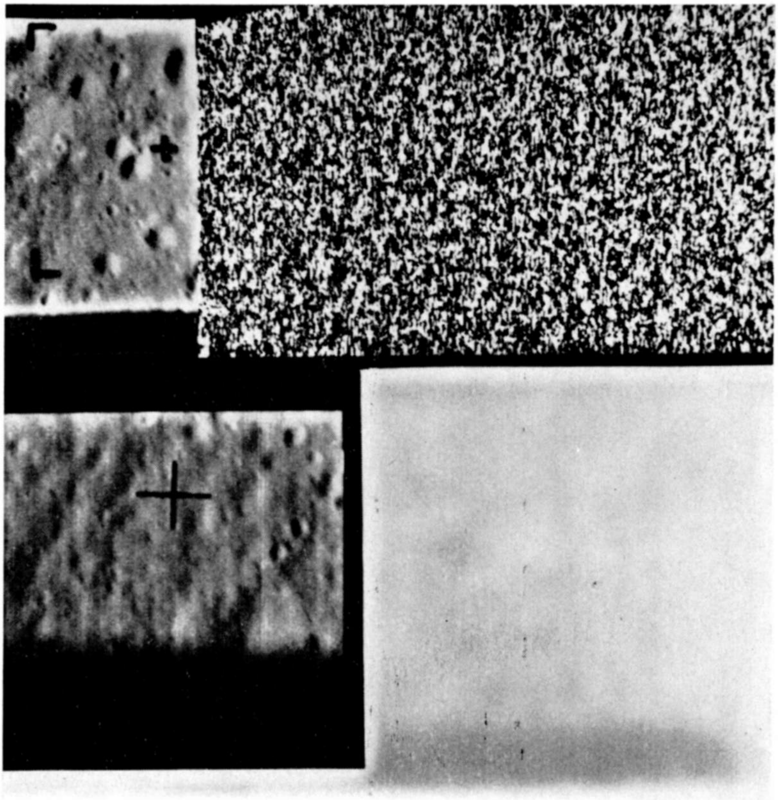


Fig. 16—Final P1 and P3 camera frames taken by Ranger VII.

The picture-taking sequence was initiated when the spacecraft was still 1550 miles from the surface of the moon, and all six television cameras continued to transmit pictures for 23 minutes until the spacecraft impacted at almost 6000 miles per hour. The initial pictures transmitted covered an area about 400 miles long and 200 miles wide.

The Ranger spacecraft was designed to provide attitude orientation through a terminal maneuver to minimize the effects of uncompensated image motion in the final pictures. The close alignment of the camera axis and velocity vector in the Ranger VII flight did not require the performance of a terminal maneuver; however, the Ranger VIII trajectory required a terminal maneuver if the maximum resolution capability of the system was to be realized. Analysis of the flight path of the Ranger VIII spacecraft indicated, however, that without a terminal maneuver, a wide swath of overlapping coverage of the lunar highlands and maria would be obtained, providing stereo coverage of the areas photographed. The stereo photographs were deemed to be of more interest than the final high resolution pictures and a terminal maneuver was not performed.

The pictures provided by Ranger VIII showed a great degree of similarity in topography between the "red" mare photographed by Ranger VII and the "blue" Mare Tranquillitatis. The stereo photographs obtained permit a far more accurate height determination for the areas photographed than was previously possible.

Ranger IX

The success of Rangers VII and VIII in photographing the lunar maria and the similarity of the data obtained in these two missions, enabled the use of Ranger IX to photograph a lunar crater region. The crater Alphonsus was selected as the primary target for this launch because of the numerous dark haloed craters and elaborate system of rills on the floor of Alphonsus. The specific selection of a target on the floor of Alphonsus represented a compromise between the central peak where the Russian astronomer Kosyrev obtained a spectrogram of an apparent gaseous discharge and the rills and dark haloed craters along the eastern wall of Alphonsus.

Ranger IX impacted on the moon at 14.08.20 GMT on March 24, 1965. The impact occurred at lunar coordinates 12.9° south and 2.4° west, near the eastern wall of Alphonsus. Ranger IX returned 5814 pictures with excellent contrast and high shadowing. The spacecraft terminal maneuver discussed in connection with Ranger VIII was performed by Ranger IX and the resolutions obtained in these pictures were the best of the three successful Ranger missions. Figure 17 is a picture transmitted 19 minutes before impact at an altitude of 1485 miles. Figures 18 through 21 correspond approximately to the successively smaller areas outlined in black in Figure 17. Figure 22 shows the last three frames transmitted by the P1 camera. The smallest crater visible is 2½ feet across and features as small as 12 inches across are visible.

During the 18-minute picture-taking period before impact, the two full-scan and four partial-scan cameras of the Ranger IX television system continuously obtained pictures of the lunar surface. The first full-scan picture was taken at an altitude of 1500 miles and the last partial-scan picture was transmitted between 0.2 and 0.05 second before impact. The flight of Ranger IX was a "textbook" mission, in which all systems performed normally and from which sharp-detailed video data of the area about and within the 75 mile-diameter crater Alphonsus were obtained.

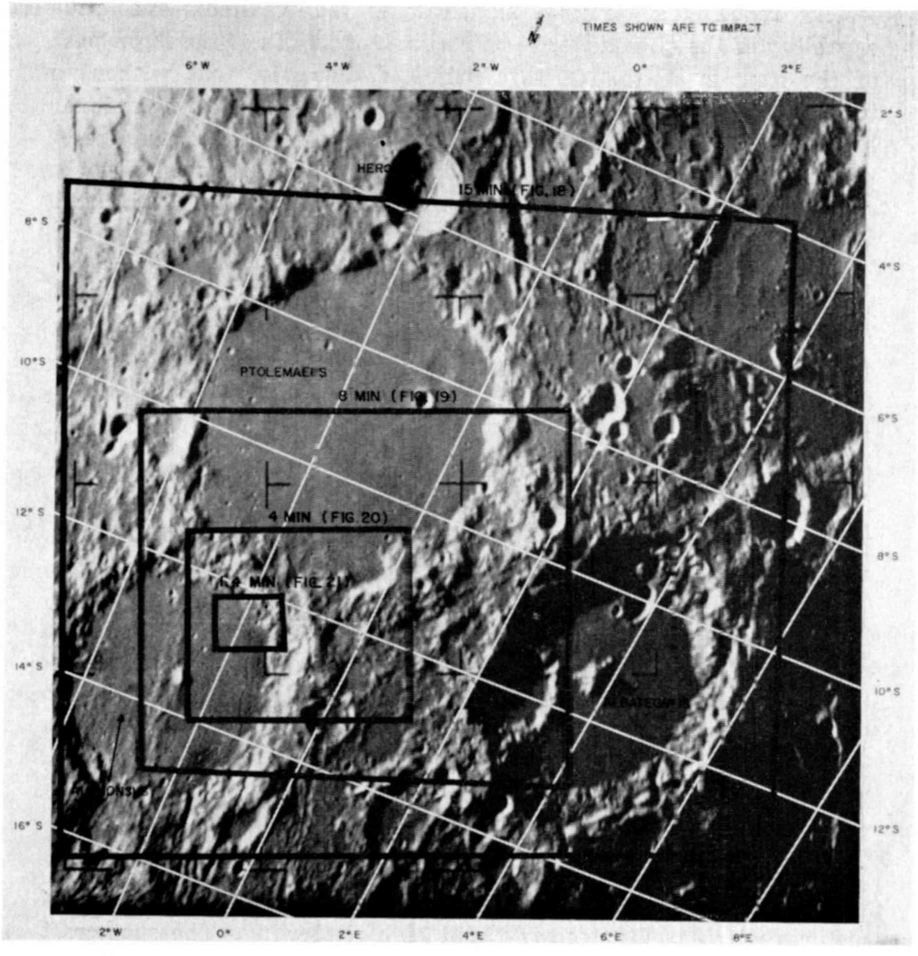


Fig. 17—Photograph taken by Ranger IX F_b camera 19 minutes prior to impact at an altitude of 1485 miles. Nesting pattern corresponds to Figures 18 through 21.

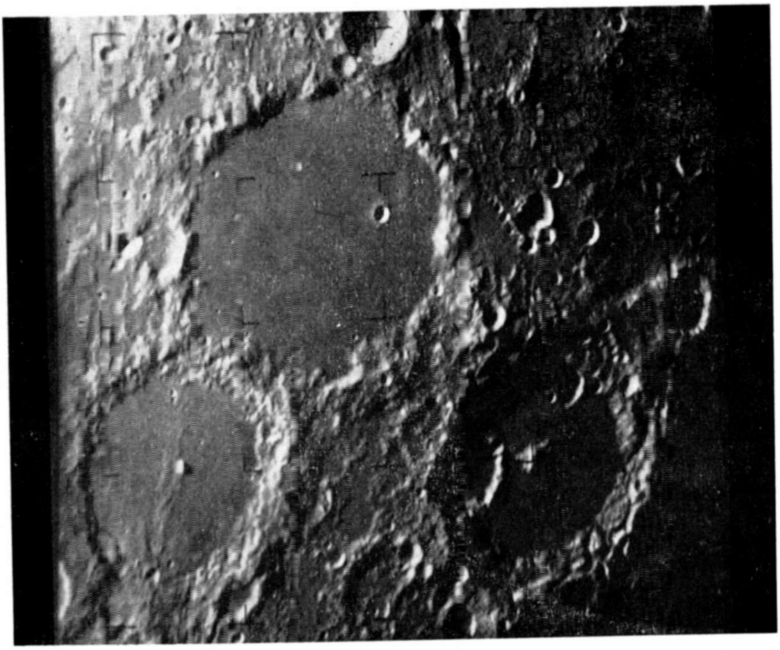


Fig. 18—Photograph taken by Ranger IX F_b camera 15 minutes before impact (altitude 1208 miles).

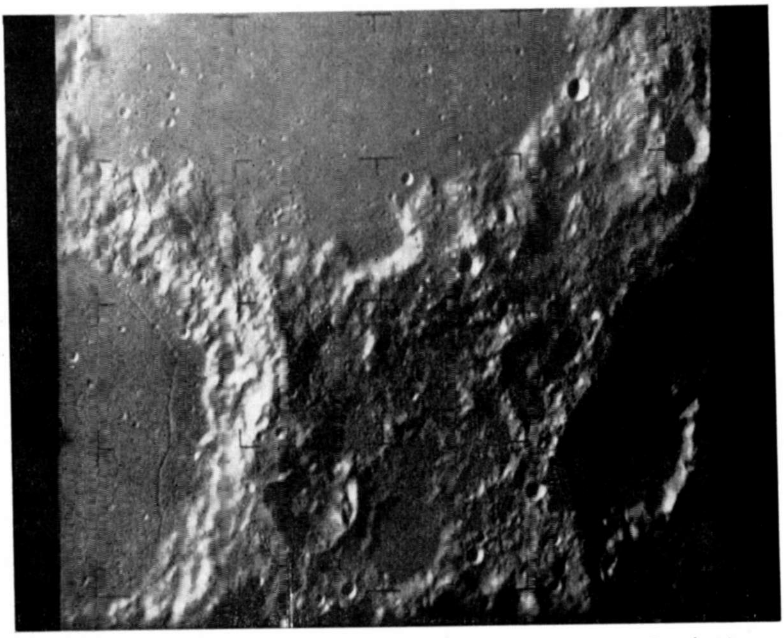


Fig. 19—Photograph taken by Ranger IX F_b camera 8 minutes before impact (altitude 683 miles).

Fig. 20—Photograph taken by Ranger IX F_b camera 4 minutes before impact (altitude 349 miles).

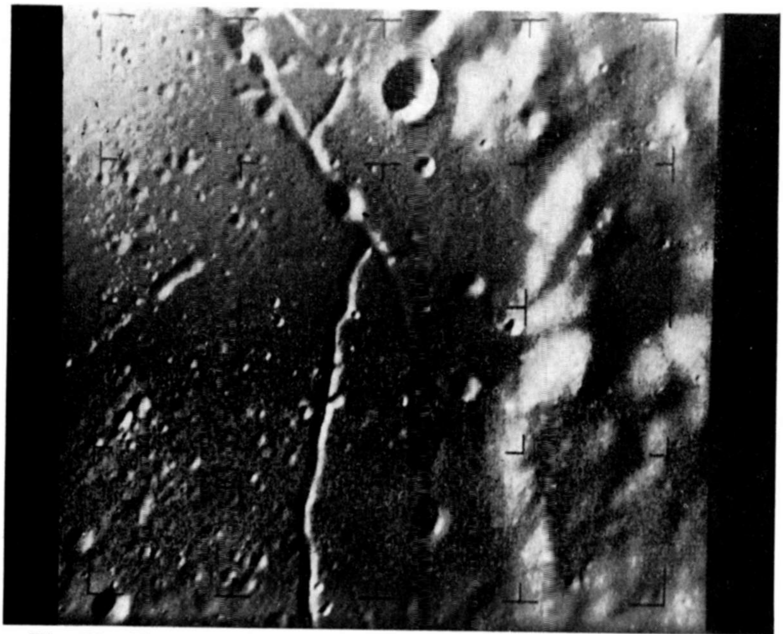


Fig. 21—Photograph taken by Ranger IX F_b camera 1.4 minutes before impact (altitude 130 miles).

CONCLUSIONS

The Ranger program fulfilled its mission objectives of obtaining both wide area coverage and high resolution photographs of the lunar surface. The television system met or surpassed all design objectives in the flight of Rangers VII, VIII, and IX. More than 17,000 high quality photographs of three widely separated regions of the lunar surface were obtained. Although the photographs cannot provide direct information on the surface hardness, they do indicate that the detailed topological features do not preclude a "soft" landing by a manned spacecraft. The photographs indicate that craters are dominant features of the lunar topology down to the limit of the resolution achieved, and it may reasonably be extrapolated that even smaller craters exist down to a size where the grain size of the lunar surface material equals the crater size.

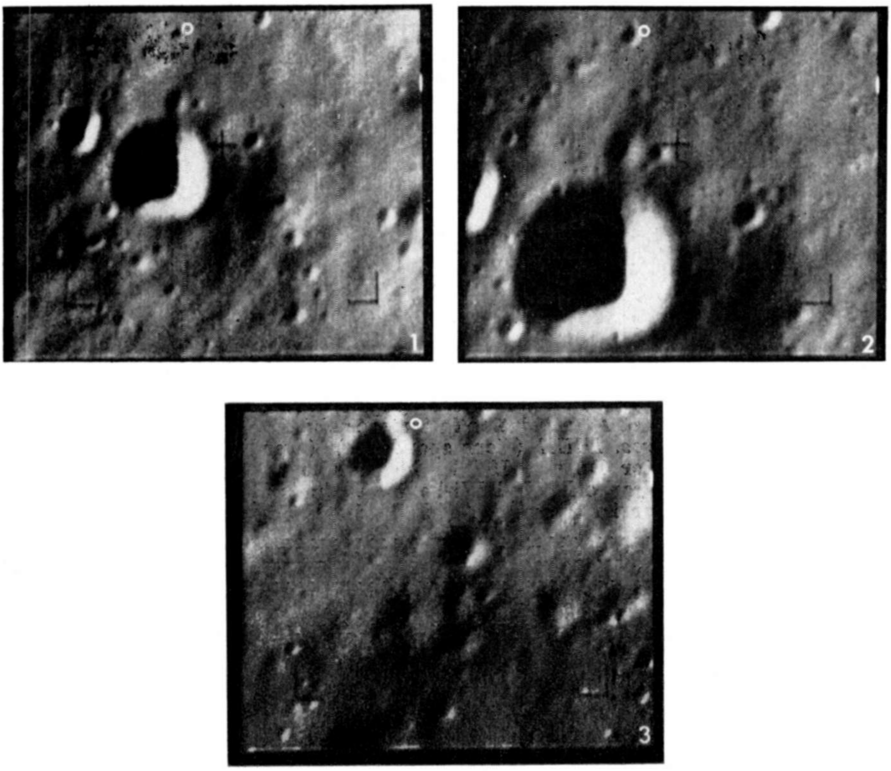


Fig. 22—Last three camera frames from Ranger IX P1 camera. The lower photograph (3) is the last in the sequence. Ranger IX impacted in the circled area near upper margin. The smallest visible crater is 2½ feet across.

RCA Technical Papers[†]

Second Quarter, 1965

Any request for copies of papers listed herein should be addressed to the publication to which credited.

"Calculations of Speed of Ladder Network for Superconductive Associative Memories," R. W. Ahrons, *Trans. IEEE PTGEC* (Short Notes) (April) 1965

"Chebyshev Approximation and Threshold Functions," K. R. Kaplan and R. O. Winder, *Trans. IEEE PTGEC* (Short Notes) (April) 1965

"Communications: A Challenge to Leadership," D. Sarnoff, *Signal* (April) 1965

"The Effects of Oxide Traps on the MOS Capacitance," F. P. Heiman and Coauthor, *Trans. IEEE PTGED* (April) 1965

"Limitations of the MOS Capacitance Method for the Determination of Semiconductor Surface Properties," K. H. Zaininger and Coauthor, *Trans. IEEE PTGED* (April) 1965

"On the Mechanism of Amplification in Superconducting Films," A. S. Clorfeine, *Proc. IEEE* (Correspondence) (April) 1965

"Metal-Semiconductor Barrier Height Measurement by the Differential Capacitance Method Without an Ohmic Reference Contact—One-Carrier System," A. M. Goodman, *Jour. Appl. Phys.* (April) 1965

"Multiple-Diode Theorems," R. L. Ernst, *Proc. IEEE* (Correspondence) (April) 1965

"The Probable Reliability of a Measurement," W. D. Moon, *Electronic Industries* (April) 1965

"Properties of Threshold Functions," R. O. Winder, *Trans. IEEE PTGEC* (Short Notes) (April) 1965

"The Spiral Top-Loaded Antenna," B. E. Keiser, *Proc. IEEE* (Correspondence) (April) 1965

"Suction versus Pressure Forced Air Cooling," G. Rezek, *Trans. IEEE PTGPEP* (April) 1965

"Supersonic Domain Wall Motion in Triglycine Sulfate," B. Binggeli and E. Fatuzzo, *Jour. Appl. Phys.* (April) 1965

"Temperature Evaluation in Microelectronic Structures Using Analog Networks," G. Rezek, *Trans. IEEE PTGPEP* (April) 1965

"Using a Decision-Table Structure as the Input Language Format for Programming Automatic Test Equipment Systems," B. Scheff, *Trans. IEEE PTGEC* (Short Notes) (April) 1965

"Microwave Phonon-Assisted Tunneling in Superconducting Diodes," B. Abeles and Y. Goldstein, *Phys. Rev. Letters* (12 April) .. 1965

"Klystron Life-Span," G. K. Hahn, Jr., *Electronics* (Readers Comment) (April 19) 1965

"Dynamics of Oxidation-Reduction Processes in Rare-Earth-Doped CaF_2," Z. J. Kiss and D. L. Staebler, *Phys. Rev. Letters* (26 April) 1965

[†] Report all corrections to *RCA Review*, RCA Laboratories, Princeton, New Jersey, 08540.

"Analysis of GaAs Tunnel Diode Oscillators," F. Sterzer, *Trans. IEEE PTGED* (May)	1965
"Beam Noise Reduction in High Magnetic Fields," B. Vural, *Proc. I.E.E.E.* (Correspondence) (May)	1965
"Comments on 'Excess Current in Gallium-Arsenide Tunnel Diodes'," R. D. Gold, *Proc. I.E.E.E.* (Correspondence) (May)	1965
"Effects of γ Irradiation on the Divalent Holmium Concentration in CaF_2," E. S. Sabisky, *Jour. Appl. Phys.* (Communications) (May)	1965
"Electronic Heat Controls for Appliances and Domestic Heating," G. D. Hanchett, *Trans. IEEE PTGIGA* (May/June)	1965
"Formation of Dislocations Around Precipitates in Single Crystals of (Zn,Cd)S:Er," M. S. Abrahams and A. Dreeben, *Jour. Appl. Phys.* (May)	1965
"Long-Wire Antennas," E. A. Laport, *Electronic Industries* (May)	1965
"Microwave Impedance of Semiconductor Posts in Waveguides. Part I," R. D. Larrabee, *Jour. Appl. Phys.* (May)	1965
"Optical Beam Deflection by Pulsed Temperature Gradients in Bulk GaAs," S. G. Liu and W. L. Walters, *Proc. I.E.E.E.* (Correspondence) (May)	1965
"Threshold Dependency on Reabsorption Loss in Injection Lasers," M. F. Lamorte, H. Junker, and T. Gonda, *IEEE Jour. Quantum Electronics* (Correspondence) (May)	1965
"A Two-Position Probe Method of Microwave Impedance Measurements for the Determination of the Electrical Properties of Indium Antimonide. Part II," R. D. Larrabee, D. W. Woodard, and W. A. Hicinbothem, Jr., *Jour. Appl. Phys.* (May)	1965
"Retrieving Data from Gemini," R. G. Erdmann, *Electronics* (May 3)	1965
"Blurred Color," C. J. Hirsch, *Electronics* (Readers Comment) (May 17)	1965
"Effect of Traps on Acoustoelectric Current Saturation in CdS," A. R. Moore and R. W. Smith, *Phys. Rev.* (17 May)	1965
"Electron and Phonon Scattering in GaAs at High Temperatures," A. Amith, I. Kudman, and E. F. Steigmeier, *Phys. Rev.* (17 May)	1965
"Television's Toughest Challenge," M. H. Mesner, *Electronics* (May 17)	1965
"Tunneling in Lead Salt p-n Junctions," R. M. Williams and Coauthor, *Phys. Rev. Letters* (17 May)	1965
"Computing Irredundant Normal Forms from Abbreviated Presence Functions," D. M. Y. Chang and Coauthor, *Trans. IEEE PTGEC* (June)	1965
"Design of Solid-State Radar Transmitter Units for the Lunar Excursion Module," E. Bliss and M. Fromer, *Trans. IEEE PTGPMP* (June)	1965
"Enumeration of Seven-Argument Threshold Functions," R. O. Winder, *Trans. IEEE PTGEC* (June)	1965
"Fixed Resistor-Card Memory," M. H. Lewin, H. R. Beelitz, and J. Guarracini, *Trans. IEEE PTGEC* (June)	1965
"Generalized Functions and Laplace-Transform Initial Conditions," R. B. Schilling and P. Schnitzler, *Proc. IEEE* (Correspondence) (June)	1965
"Generation of Microwave Power by Parametric Frequency Multiplication in a Single Transistor," M. Caulton, H. Sobol, and R. L. Ernst, *RCA Review* (June)	1965
"Grid-Controlled Power Tubes in Particle-Accelerator Applications," M. V. Hoover, *Trans. IEEE PTGNS* (June)	1965
"Linear Signal Stretching in a Time-Variant System," H. Weinstein, *Trans. IEEE PTGCT* (June)	1965
"Low-Energy-Electron Scattering from Photoconductors Applied to a Camera Tube," A. D. Cope and W. Bruce, *RCA Review* (June)	1965
"Multiple-Input Large-Signal Mixer Analysis," P. Torrione and S. Yuan, *RCA Review* (June)	1965

"Physical Limitations on Frequency and Power Parameters of Transistors," E. O. Johnson, *RCA Review* (June) 1965

"Plasma Acceleration by Electron Cyclotron Resonance," H. Hendel, T. Faith, and E. C. Hutter, *RCA Review* (June) 1965

"Plasma Anodized Lanthanum Titanate Films," R. E. Whitmore and J. L. Vossen, *Trans. IEEE PTGPMP* (June) 1965

"Retarding-Field Analyzers for the Measurement of Axial-Velocity Distributions in Electron Beams," M. Caulton, *RCA Review* (June) 1965

"Signal and Noise Propagation on a Drifting Multivelocity Beam with both Low- and High-Velocity Tails," B. Vural, *Jour. Appl. Phys.* (Communications) (June) 1965

"A 60-Megacycle Video Chain for High-Definition Television Systems," O. H. Schade, Sr., *RCA Review* (June) 1965

"Stress Analysis of Butt-Type Ceramic-To-Metal Seals," S. W. Kessler, *RCA Review* (June) 1965

"Thermoelectric Properties of Niobium in the Temperature Range 300°-1200°K," V. Raag and H. V. Kowger, *Jour. Appl. Phys.* (June) 1965

"Ultra-Low-Noise Beams from Synthesized Plasma Cathodes," A. L. Eichenbaum, *Proc. IEEE* (Correspondence) (June) 1965

"The Improved Gray Scale and the Coarse-Fine PCM Systems, Two New Digital TV Bandwidth Reduction Techniques," W. T. Bisignani, G. P. Richards, and J. W. Whelan, *IEEE International Convention Record Part 4, Aerospace and Military Electronics* 1965

"Learning Systems for Automatic Control," J. Sklansky, *IEEE International Convention Record, Part 6, Symposium on Automatic Control; Systems Science; Cybernetics; Human Factors* 1965

"Novel Frequency Dividers for TV Sync Generators," A. J. Banks and F. I. Johnson, *IEEE International Convention Record, Part 2, Communications II; Radio Communication, Broadcasting, and Audio* 1965

"Physical Limitations on Frequency and Power Parameters of Transistors," E. O. Johnson, *IEEE International Convention Record, Part 5, Antennas; Microwaves; Electron Devices* 1965

"Progress in Missile and Aerospace Automatic Control," R. Lieber, S. Kolodkin, and H. Perkel, *IEEE International Convention Record, Part 6, Symposium on Automatic Control; Systems Science; Cybernetics; Human Factors* 1965

"Spectral Analysis—An RFI Prediction Tool," R. F. Ficchi, *IEEE International Convention Record, Part 2, Communications II; Radio Communication, Broadcasting, and Audio* 1965

"'Threat Indifference' A Theory of Communications Survivability," L. Siegel and K. M. Curtin, *IEEE International Convention Record, Part 2, Communications II; Radio Communication, Broadcasting, and Audio* 1965

"Vehicular Interference Radiation Measurement Technique," J. R. Neubauer, *IEEE International Convention Record, Part 2, Communications II; Radio Communication, Broadcasting, and Audio* 1965

AUTHORS

N. K. **Chitre** received the B.Sc. and M.Sc. degrees in Physics from the Banaras Hindu University in 1953 and 1955, respectively, and the M. Tech. degree in Electronics from the Indian Institute of Technology, Kharagpur, in 1956. After a year with Van der Heem, N.V. The Hague, he studied at the Imperial College, London, where he received his diploma (D.I.C.) and, in 1962, the Ph.D. degree of the University of London. In March 1963, Dr. Chitre joined the RCA Victor Co. Ltd., Montreal, where he has worked on microwave communications systems, both overland and satellite, and designed various microwave multiplexing networks, filters, and ferrite components. Dr. Chitre is a Member of the I.E.E.

R. **Feryszka** received the masters degree in Electrical Engineering from the Polytechnic of Lwow, Poland in 1936. From 1936 to 1939 he was employed as a design engineer on radio receivers at Elektrit, Wilno, Poland. From 1949 to 1956 Mr. Feryszka was manager of Design Engineering and later assistant to the Chief Engineer at "KASPRZAK" Electronic Plant in Warsaw. In 1956 he transferred to the Tele Radio Institute in Warsaw, where he carried out environmental investigations on electronic components. In 1960 Mr. Feryszka joined RCA Electronic Components and Devices as an applications engineer on the micromodule program. In 1963 he joined the advanced development group at Somerville, where he was responsible for the development of high-power tunnel-diode inverters. He is presently working on the development of UHF MOS transistors for television applications. Mr. Feryszka is a member of the Institute of Electrical and Electronics Engineers.

Peter Foldes received the B.S. degree in electrical engineering from the Technical University of Budapest in Budapest in 1950. From 1950 to 1956 he was a research engineer at the Hungarian Telecommunication Research Institute and from 1953 to 1956, on a part-time basis, was also a lecturer on antennas at the Technical University of Budapest. In 1957 he joined RCA Victor Co., Ltd., Montreal, Canada. He has worked mostly in the field of antenna, propagation, and system engineering studies. Since 1958 he has been responsible for the theoretical aspects of the microwave subsystem in a wide-band microwave communication system. Mr. Foldes is a member of the American Institute of Electrical Engineers.

P. Gardner received the diploma in Technology in Applied Physics from Brunel College of Technology in London, England, in 1960. From 1960 to 1961 he worked at Transitron Electronic Corporation. Mr. Gardner joined the RCA Semiconductor and Materials Division in 1961. There, he has worked on high-speed germanium tunnel diodes and microcurrent and high-current germanium tunnel diodes. In 1963, he joined the Advanced Devices Development Activity where he has been engaged in the development of high-current Ge and GaAs tunnel diodes and MOS transistors.

Steven G. Komlos received the B.S. degree in electronics engineering from the Technical University of Budapest in 1952. From 1952 to 1956 he worked as a lecturer at the Technical University of Budapest, and from 1954 to 1956 he was also consulting engineer for the Hungarian Television Co. In 1957, he joined RCA Victor Company, Ltd., Montreal, where he carried out development work in TV Broadcasting. At present, he is engaged in antenna development work.

EDGAR W. MARKARD received the BEE Degree from City College of New York in June 1961. He joined RCA and after a short training period was assigned to the advanced microwave techniques group of RCA Communications System Division, New York, where he worked on S band, broadband parametric amplifiers. At the completion of that program, he worked on the development of varactor frequency multipliers, and was responsible for the development of a high-power microwave varactor multiplier chain which yielded two watts at C-band. Hhe was then assigned to the cryoelectric program which has as its objective the development of a parametric amplifier based on the superconducting properties of thin films. He is also working on high power tunable and broadband varactor frequency multipliers, and has participated in the development of the theory of high nonlinearity (hyper-abrupt) varactor devices.

Mr. Markard is a member of the Institute of Electrical and Electronics Engineers, Eta Kappa Nu, and the American Institute of Aeronautics and Astronautics.

BERNARD P. MILLER received a B.S. in Aeronautical Engineering from Pennsylvania State University in 1950, and did post-graduate work, also in Aeronautical Engineering, at the U.S. Air Force Institute of Technology in 1953 and 1954 and at Princeton University from 1957 to 1959. From 1950 to 1953, as a captain in the USAF, he was assigned to the Aircraft Laboratory, Wright Air Development Division. In 1954, he taught at the U.S. Naval Academy, Annapolis, as an instructor of Fluid Dynamics and Thermodynamics. In 1956 he was appointed Head of the Thermodynamics group of the Academy's Department of Marine Engineering. He joined the RCA Special Systems and Development Department in Princeton in 1957, and transferred to the Astro-Electronics Division in 1958. There, on Project JANUS, he performed trajectory and orbital analyses and ascertained propulsion-system requirements. From 1959 through 1961, he was Project Engineer on a series of studies examining military applications of space systems. In 1961, Mr. Miller participated in the series of studies that eventually led to the present RCA Ranger Project. With the organization of the RCA Ranger Project, he was appointed Group Leader, Mechanical Systems, and in 1963 was made Ranger Project Manager. Mr. Miller received a NASA Public Service Award in October 1964 for work on the six-camera television system used aboard Ranger 7. Mr. Miller is a member of the Institute of Aerospace Sciences and is a registered Professional Aeronautical Engineer.

John A. Olmstead received the B.S. degree in electrical engineering from the University of Buffalo in 1952 and the M.S. degree in electrical engineering in 1957 from Newark College of Engineering. He has also taken graduate courses in mathematics and physics at the University of Buffalo. In 1952 he joined the Radio Corporation of America as a specialized trainee. From 1952 to 1954 he worked on gaseous discharge devices at RCA Laboratories, Princeton. In 1954 he joined Electronic Components and Devices, Harrison, N. J., where he worked on the design and development of gas-filled and other specialized industrial tubes. Between 1950 and 1960 he alternated between assignments at RCA and the University of Buffalo where he was an Assistant Professor of Electrical Engineering. In 1960 he returned to RCA on a full time basis, joining the Semiconductor and Materials Division. He has worked on GaAs solar cells, p-n junction field-effect devices, high-frequency power bipolar transistors and MOS field-effect transistors. He is responsible for the multi-emitter concept basic to the "overlay" transistor technique. Mr. Olmstead is a member of the Institute of Electrical and Electronics Engineers and Sigma Xi.

R. Schwerdtfeger graduated from McGill University in 1963 with a B.E.E. Since this time he has been a member of the RCA Victor Company, Ltd., Montreal, working on the design and development of the multimode feed antenna to be used in Canada's first satellite communication station.

Joseph H. Scott received the A.B. degree in chemistry from Lincoln University, Pennsylvania. In 1958 he attended graduate school at Howard University, Washington, D.C. and is presently pursuing a degree in Electrical Engineering at Newark College of Engineering. In 1959, he joined RCA, where he has been engaged in research and development of semiconductor devices. He has worked on GaAs solar cells, p-n junction unipolar transistors, high-frequency high-power bipolar transistors and, most recently, MOS transistors. Together with J. A. Olmstead he did the initial work on multi-emitter "overlay" transistors. Mr. Scott is a member of the Electrochemical Society.

T. Szirtes received his Diploma of Mechanical Engineering at the Technical University of Budapest in 1950. In the same year he joined the Hungarian Research Institute for Development of Engines and Vehicles as a research engineer. In 1956 he emigrated to Canada where he became an instructor at McGill University, Montreal, in the Department of Mechanical Engineering. In 1960 he received his Master of Mechanical Engineering degree at McGill. In early 1961 he joined RCA Victor Company, Ltd., Montreal, as a mechanical engineer designing antennas, feed systems and mechanisms. He has also worked on problems of heat transfer, structures, and applied and celestial mechanics. He is a member of ASME.

S. Yuan (See *RCA Review*, Vol. XXVI, June, 1965, p. 319.)