RCA REVIEW

a technical journal

Published quarterly by
RCA LABORATORIES
in cooperation with all subsidiaries and divisions of
RADIO CORPORATION OF AMERICA

VOLUME XXV | DECEMBER 1964 | NUMBER 4

CONTENTS

RCA REVIEW is regularly abstracted and indexed by *Abstracts of Photographic Science and Engineering Literature, Applied Science and Technology Index, Bulletin Signalétique des Télécommunications, Chemical Abstracts, Electronic and Radio Engineer, Mathematical Reviews, and Science Abstracts* (I.E.E.-Brit.).

RCA REVIEW

SWITCHING RESPONSE OF COMPLEMENTARY-SYMMETRY MOS TRANSISTOR LOGIC CIRCUITS

By

J. R. Burns

RCA Laboratories
Princeton, New Jersey

Summary—*The complementary-symmetry MOS transistor logic gate is attractive in many applications because of its low standby power, radiation resistance, and adaptability to integrated circuit technology. A computer analysis is used to relate the pair delay time and power dissipation during switching to the static characteristics of the transistors. Simple design formulas for the pair delay and power have been obtained and are applied to an illustrative design of a NOR logic gate with a fan-in and fan-out of five. Excellent agreement between computed and experimental performance has been obtained. All of the results are also generally applicable to TFT (thin-film transistor) complementary logic gates.*

Introduction

THE RECENT development of the MOS (metal-oxide-semiconductor) insulated-gate transistor[1] has made possible integrated logic circuits with active-device packing densities of 2000 per square inch of planar surface. The advantages of such miniaturization, namely reduced cost and size, are obviously desirable in many applications. However, the overall performance of an integrated circuit must be such that these advantages are not offset by practical considerations such as speed, reliability, and power dissipation.

The most promising MOS logic-circuit arrangement in these areas consists of a series connection of an N- and a P-type* transistor as in Figure 1. The circuit performs a logical inversion and offers good noise immunity, and, most important, the standby-power dissipation in either binary state is extremely low, being on the order of microwatts.

---

[1] S. R. Hofstein and F. P. Heiman, "The Silicon Insulated-Gate Field-Effect Transistor," *Proc. IEEE*, Vol. 51, p. 1190, Sept. 1963.

* The terminology N and P type refers to the carriers involved in conduction, i.e., electrons in the N type and holes in the P type.

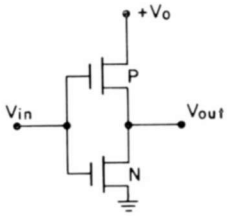


Fig. 1—Basic inverter circuit.

D-C Considerations

Device Operation

The MOS transistor[1] is a majority-carrier device in which the current in a conducting channel between two electrodes (denoted as the *source* and *drain*) is modulated by a voltage applied to a third terminal, the *gate*, which is insulated from the *source* and *drain*. In the N type transistor, the majority carriers are electrons, so that a positive voltage at the gate will *enhance* the number of electrons and increase the conductivity of the channel as indicated by the characteristics shown in Figure 2a. For all gate voltages less than some

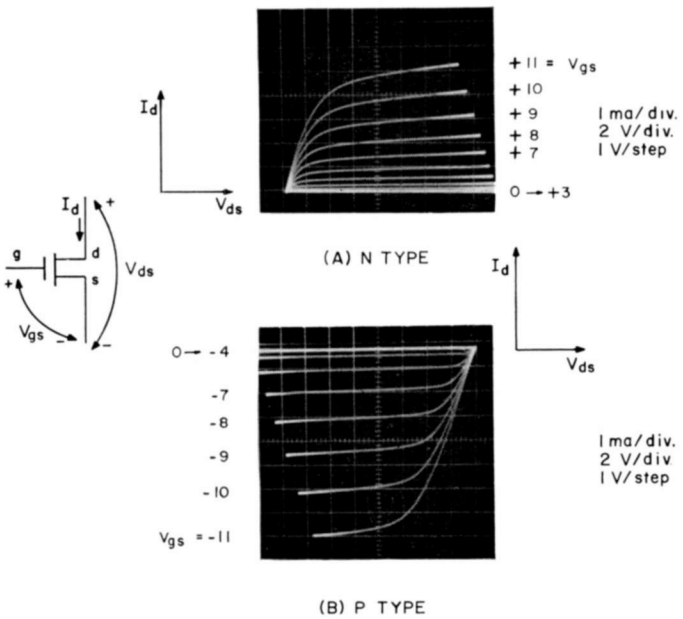


Fig. 2—*V-I* characteristics of MOS transistors.

threshold value V_{T_N}, the entire channel is cut off, resulting in a very low drain current. The threshold voltage of the N-type transistor in Figure 2a is seen to be +3 volts. The operation of the P-type transistor is analogous to the N-type except that the carriers are holes and the voltages are negative. The threshold of the P-type transistor, V_{T_P} in Figure 2b, is —4 volts. An attractive feature of the threshold voltage is that its value may be controlled by impurity doping of the channel and is, therefore, a design parameter of the transistor. Theoretically, for $V_{ds} > V_{gs} - V_{T_N}$ (N type), the current should remain approximately constant for any further increase in V_{ds}. At this drain voltage, the channel is depleted of carriers near the drain end of the channel, and an increase in V_{ds} causes a corresponding increase in channel resistance.

It has been shown that the drain-current characteristic of an N-type MOS transistor (as well as a thin-film transistor) is a well-defined function of the drain and gate voltages;[1,2]

$$I_d = K_N [2 V_{ds}(V_{gs} - V_{T_N}) - V_{ds}^2] \qquad \text{for } V_{ds} < V_{gs} - V_{T_N}, \quad (1a)$$

$$I_d = K_N (V_{gs} - V_{T_N})^2 \qquad \text{for } V_{ds} \geq V_{gs} - V_{T_N}, \quad (1b)$$

$$I_d = 0 \qquad \text{for } V_{gs} \leq V_{T_N}. \quad (1c)$$

The constant $K_N = \mu_N \epsilon_{ox} W/(2LT_{ox})$, where μ_N is the mobility of the electrons, ϵ_{ox} the permittivity of the insulating oxide layer, W the channel width, L the channel length, and T_{ox} the thickness of the insulating oxide. Using $C_{\text{in}} = \epsilon_{ox} LW/T_{ox}$ as the effective capacitance* at the gate of the transistor,

$$K_N = \frac{\mu_N C_{\text{in}}}{2L^2}.$$

The equations for the P-type transistor are similar, with appropriate changes in sign and with V_{T_P} substituted for V_{T_N}.

Operation of N-P Inverter

Basically, the operation of the N-P inverter of Figure 1 is as follows: When the input voltage V_{in} is 0 volts, the P transistor is biased

[2] H. Borkan and P. K. Weimer, "An Analysis of the Characteristics of Insulated-Gate Thin-Film Transistors," *RCA Review*, Vol. 24, p. 153, June 1963.

* This does not include stray capacitance due to gate overlap of the source and drain electrodes.

on by $-V_0$ volts (V_0 is the positive supply voltage) and the N transistor is cut off. The output voltage is now at $+V_0$, as shown in Figure 3(a), since the N transistor has a very high resistance relative to the P transistor. As V_{in} is increased, the current in the N-type rises and the current in the P-type falls; the input voltage at which the two

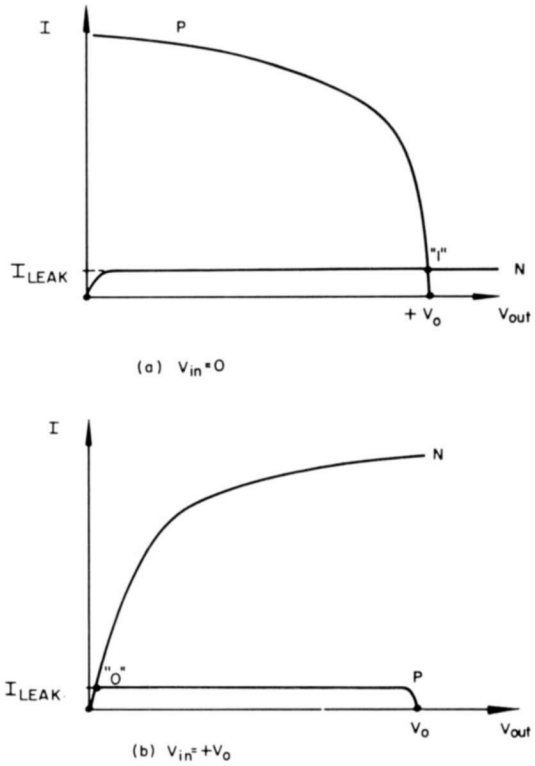


Fig. 3—Operating points of N-P inverter.

transistors are simultaneously in the constant-current regions marks the high-gain or active region of the circuit. As V_{in} is increased beyond this voltage, the output voltage falls rapidly toward zero; ultimately, for $V_{in} = +V_0$, the P unit will be cut off and the N unit will be biased on by $+V_0$ volts, and V_{out} will be 0 volts. The condition for $V_{in} = +V_0$ is illustrated in Figure 3b. Note that for either state, the total current flow is that of a cut off transistor, which is typically only a few microamperes for present units. It is seen that the threshold voltages of the N and P transistors must be positive and negative,

respectively, i.e., both transistors must be enhancement type in order to satisfy this condition.

The static operation of the circuit is conveniently described by a transfer characteristic that relates the output voltage to the input voltage. Such a characteristic, determined experimentally for two typical transistors for $V_0 = +15$ volts, is shown in Figure 4. Current

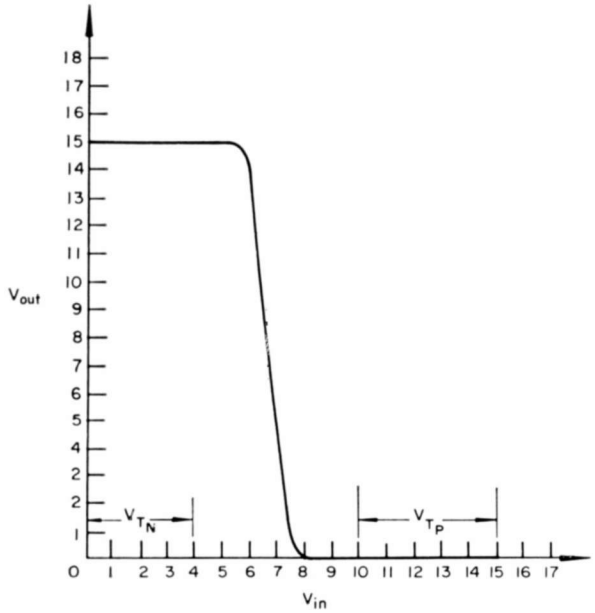


Fig. 4—Circuit transfer characteristic.

flows in the circuit for $4 \leq V_{in} \leq 10$ volts since $V_{T_N} = +4$ volts and $V_{T_P} = -5$ volts. The characteristic shown is typical of digital logic circuits—two regions of low gain separated by a high-gain transition or active region. To minimize the effects of device variations and external noise, this transition region should be midway between the "0" and "1" levels so that variations in either direction may be tolerated.

In the N-P inverter circuit of Figure 1, the current in the P transistor is that for a gate voltage $V_{in} - V_0$ and a drain voltage $V_{out} - V_0$. Thus,

$$I_P = K_P [2(V_0 - V_{out})(V_0 - V_{in} - |V_{T_P}|) - (V_0 - V_{out})^2] \qquad (2a)$$

$$\text{for } V_{out} > V_{in} + |V_{T_P}|,$$

$$I_P = K_P [V_0 - V_{in} - |V_{T_P}|]^2 \quad \text{for } V_{out} \leq V_{in} + |V_{T_P}|, \tag{2b}$$

and

$$I_P = 0 \quad \text{for } V_{in} + |V_{T_P}| \geq V_0.$$

In the high-gain region, both transistors are in the constant-current regions defined by Equations (1b) and (2b), where $V_{in} - V_{T_N} \leq V_{out} \leq V_{in} + |V_{T_P}|$. The condition on the input transition voltage is, therefore,

$$K_P [V_0 - V_{in}^* - |V_{T_P}|]^2 = K_N [V_{in}^* - V_{T_N}]^2. \tag{3a}$$

Defining $\beta \equiv \dfrac{K_P}{K_N}$, $\alpha_N \equiv \dfrac{V_{T_N}}{V_0}$, $\alpha_P \equiv \dfrac{|V_{T_P}|}{V_0}$, $v_{in}^* \equiv \dfrac{V_{in}^*}{V_0}$,

gives

$$\beta [1 - v_{in}^* - \alpha_P]^2 = [v_{in}^* - \alpha_N]^2. \tag{3b}$$

Solving for v_{in}^*, the normalized transition voltage,†

$$v_{in}^* = \frac{\sqrt{\beta}(1 - \alpha_P) + \alpha_N}{1 + \sqrt{\beta}}. \tag{4}$$

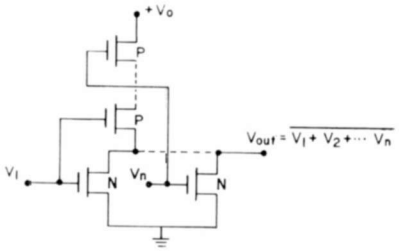
The normalized transition voltage is defined exactly and can therefore be designed for the optimum value of 0.5 by a suitable choice of transistor parameters. In particular, for $\alpha_N = \alpha_P = 0.5$, $v_{in}^* = 0.5$ and is independent of the value of β. This may appear to be an interesting condition, since the transfer characteristic will always be optimum, regardless of fluctuations in the constants K_N and K_P. It will be shown later that this choice of parameters results in a relatively poor time response, but in some applications this consideration may be of secondary importance.

NOR and NAND Logic Circuits

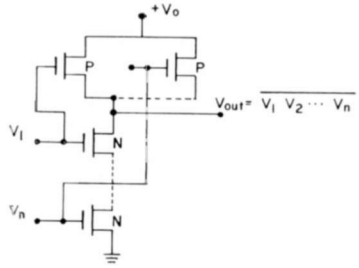
The extension of the single input N-P inverter to multiple input NOR and NAND gates is shown in Figure 5. In the NOR gate when any one of the inputs is high, one N unit will conduct while the associated P transistor in the series string will be cut off thereby opening the whole string. Similarly, for the NAND gate when all the inputs

† This condition does not hold for $\alpha_N + \alpha_P > 1$ since this results in a multivalued transfer characteristic, which in general is not particularly useful.

are high, all the series N transistors are conducting and all the shunt P transistors are cut off. Thus in either connection, for any combination of inputs, the current path to ground is always a high resistance and the standby power is negligibly small. The circuit output impedance is low, however, because of the conducting transistor.



(A) N INPUT NOR GATE



(B) N INPUT NAND GATE

Fig. 5—NOR and NAND circuits using the N-P inverter.

The effect of series and shunt transistors on the overall d-c characteristics of the circuit is to change the constants K_N and K_P. For n identical transistors in parallel, with the same voltage at the inputs, the current is increased by a factor of n, i.e., $K_{eff} = nK$. For n transistors in series, the situation is not as clear-cut since the gate-to-source voltage varies depending on the position of the transistor in the string. However, to a good approximation, the total current is scaled down by a factor of n, particularly at high input voltages. The validity of this approximation for one, two, and three P transistors in series can be seen in Figure 6. The effective β of the NOR and NAND circuits is a function of the number of inputs that are high and of the sequence in which they are applied. For a NOR gate with

n inputs, the worst case for noise immunity is when each input is turned on simultaneously since the current from the parallel N transistors is additive. In this case,

$$\beta_{eff} = \frac{\frac{K_P}{n}}{nK_N} = \frac{\beta_0}{n^2}. \tag{5a}$$

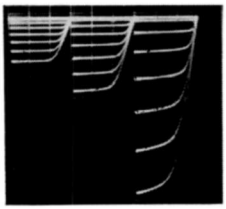


Fig. 6—V-I characteristic of one, two, and three P transistors in series.

However, for switching, the worst case occurs when only one input is applied (which will be shown later) so that

$$\beta_{eff} = \frac{\frac{K_P}{n}}{K_N} = \frac{\beta_0}{n}. \tag{5b}$$

Similarly, for an n-input NAND gate when all inputs are applied simultaneously,

$$\beta_{\text{eff}} = \frac{nK_p}{\frac{K_n}{n}} = n^2\beta_0. \tag{6a}$$

When all inputs but one have been applied, and then the last switches from 0 to $+V_0$ volts,

$$\beta_{\text{eff}} = \frac{K_p}{\frac{K_n}{n}} = n\beta_0. \tag{6b}$$

These relationships may be used in Equation (4) to find v^*_{IN} by replacing β with β_{eff}.

Transient Response

Step Input

The switching response of the N-P logic circuit is determined solely by capacitance loading at the input and output and the amount of current available from the transistors to charge this capacitance. For a useful first approximation to the switching speed, a voltage step input will be assumed. The basic differential equation for the output voltage that must be solved is of the form

$$I_P = I_N + I_C, \tag{7}$$

or

$$C\frac{dV}{dt} = I_P - I_N. \tag{8}$$

Solving, we obtain

$$t = C\int_{V_0}^{V}\frac{dV}{[I_P - I_N]} = C\int_{V_0}^{V}\frac{dV}{I_C}.$$

Evaluation of the integral yields the time elapsed for the output to change from the initial voltage V_0 to any voltage V. Physically, it is evident that the minimum switching time is obtained by maximizing the amount of current available to charge C for any voltage V. This is shown in Figure 7 for the turn-on and turn-off conditions of the inverter. The response times may then be calculated by plotting $1/I_C$ versus V and measuring the area under the curve.

However, when the input is a function of time, this method becomes quite difficult and in most practical cases may even be impossible. Therefore, it is desirable to look for a general analytic solution that can easily be extended to specific cases.

Fall-Time Response

Using a step input of $+V_0$ volts for V_{in} the differential Equation (8) is given by

$$C_{\text{out}}\frac{dV_{\text{out}}}{dt} = -I_N;\ V_{\text{out}}(0) = +V_0.$$

In the interval $V_0 \geq V_{\text{out}} \geq V_0 - V_{T_N}$ and $V_{gs} = +V_0$, I_N is given by

Equation (1b) so that

$$C_{\text{out}} \frac{dV_{\text{out}}}{dt} = -K_N \ (V_0 - V_{T_N})^2 = -K_N V_0^2 \ (1 - \alpha_N)^2,$$

where α_N is defined by Equation (3b), or

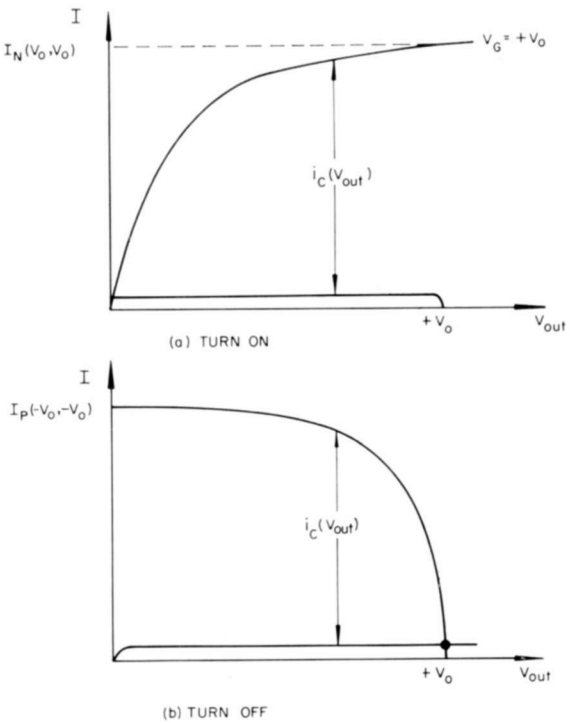


Fig. 7—Switching currents.

$$V_{\text{out}}(t) = V_0 \left[1 - \frac{K_N V_0 (1 - \alpha_N)^2}{C_{\text{out}}} t \right]. \tag{10a}$$

For $V_{\text{out}} \leq V_0 - V_{T_N}$, according to Equation (1a),

$$t(v) = t_0 + C_{\text{out}} \int_{V_0 - V_{T_N}}^{v} \frac{dv'}{K_N V_0^2 \ [2v' \ (1 - \alpha_N) - v'^2]},$$

where $t_0 = C_{out}\,\alpha_N/[K_N V_0\,(1-\alpha_N)^2]$ is the time for V_{out} to fall from V_0 to $V_0 - V_{T_N}$.

The solution of the above is

$$t(v) = t_0 + \frac{C_{out}}{K_N V_0\,(1-\alpha_N)} \left[\tanh^{-1} \frac{V_0 - V_{T_N} - v'}{V_0 - V_{T_N}} \right]_{V_0 - V_{T_N}}^{v}$$

which gives

$$V_{out}(t) = (V_0 - V_{T_N}) \left[1 - \tanh \left(\frac{K_N V_0\,(1-\alpha_N)}{C_{out}} [t - t_0] \right) \right]; \; t > t_0. \tag{11a}$$

Defining a time $\tau_N \equiv C_{out}/(K_N V_0)$ and expressing the results in normalized form, Equations (10a) and (11a) become

$$\frac{V_{out}}{V_0} = 1 - (1-\alpha_N)^2 \frac{t}{\tau_N}; \; 0 \leq t \leq t_0; \; V_{out} > V_0 - V_{T_N} \tag{10b}$$

$$\frac{V_{out}}{V_0} = (1-\alpha_N) \left[1 - \tanh\,(1-\alpha_N) \frac{t - t_0}{\tau_N} \right]; \; t > t_0; \; V_{out} < V_0 - V_{T_N}. \tag{11b}$$

In the special case where $V_{T_N} = 0$,

$$\frac{V_{out}(t)}{V_0} = 1 - \tanh \frac{t}{\tau_N}.$$

This equation is plotted in Figure 8, together with the ramp and exponential functions. The response is slightly better than for a resistor of value $1/(K_N V_0)$ and poorer than a constant-current source of value $K_N V_0^2$ as is to be expected from the V-I characteristics of Figure 9 and Equation (9).

Combining Equations (10b) and (11b), the 10% to 90% fall time for $0.1 \leq \alpha_N \leq 0.9$ is

$$T_F = \tau_N \left[\frac{\alpha_N - 0.1}{(1-\alpha_N)^2} + \frac{\tanh^{-1} \left(1 - \dfrac{0.1}{1-\alpha_N} \right)}{1-\alpha_N} \right]. \tag{12}$$

The normalized fall time is plotted as a function of α_N in Figure

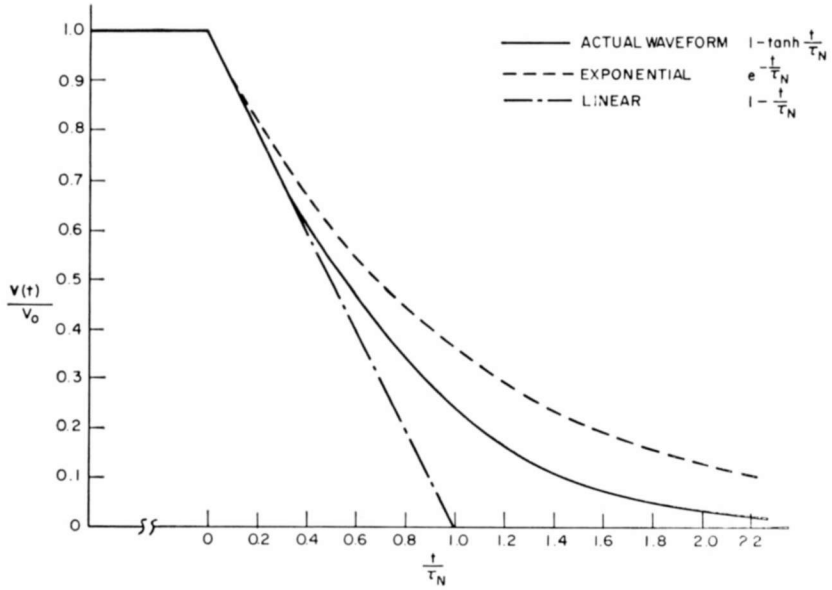


Fig. 8—Normalized turn-on waveform, $V_{TN} = 0$.

19 (dotted curve) and is seen to be a monotonically increasing function of α_N, the normalized threshold voltage.

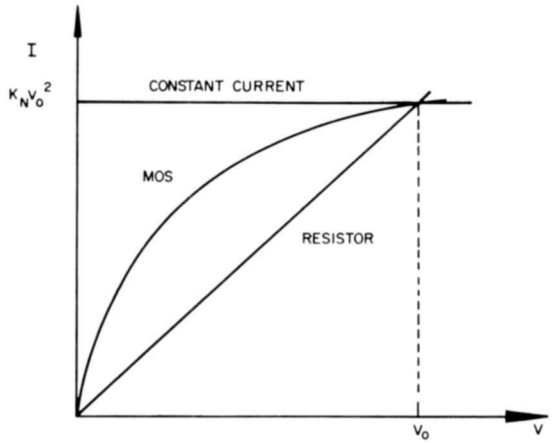


Fig. 9—Comparative V-I characteristics of resistor, MOS, and current source.

Rise-Time Response

When V_{in} switches to 0 volts, the differential Equation (8) becomes

$$C_{out}\frac{dV_{out}}{dt} = I_P.$$

By analysis similar to that for the fall-time response,

$$\frac{V_{out}(t)}{V_0} = (1-\alpha_P)^2\frac{t}{\tau_P}; \quad V_{out} < |V_{T_P}| \tag{13}$$

$$= 1-(1-\alpha_P)\left[1-\tanh\frac{(1-\alpha_P)(t-t_0)}{\tau_P}\right];$$

$$V_{out} \geq |V_{T_P}|, \tag{14}$$

where $\tau_P = C_{out}/(K_P V_0)$ and $t_0 = \tau_P\alpha_P/(1-\alpha_P)^2$.

The 10% to 90% rise time for $0.1 \leq \alpha_P \leq 0.9$ is

$$T_R = \tau_P\left[\frac{\alpha_P - 0.1}{(1-\alpha_P)^2} + \frac{\tanh^{-1}\left(1-\dfrac{0.1}{1-\alpha_P}\right)}{1-\alpha_P}\right]. \tag{15}$$

The rise time is of exactly the same functional form as the fall time, the only essential difference being in the normalizing time, τ_P.

Experimental Verification of Circuit Response to Step Input

The step response is, as has been shown, an exact function of the device constants, K_N and K_P, the threshold voltages, V_{T_N} and V_{T_P}, and the supply voltage, V_0. These constants can be determined from the two-terminal characteristics for N and P transistors with the gate connected to the drain as shown in Figure 10. From Equations (1b) and (2b),

$$K_N = \frac{I_N(V_0, V_0)}{(V_0 - V_{T_N})^2},$$

$$K_P = -\frac{I_P(-V_0, -V_0)}{(V_0 - V_{T_P})^2}. \tag{16}$$

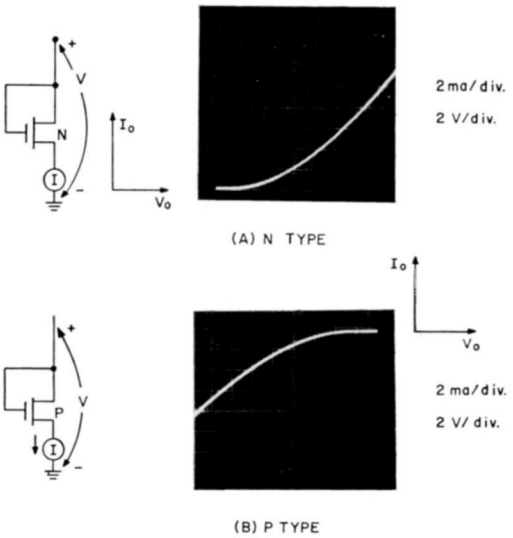


Fig. 10—I_o–V_o characteristics.

Taking $V_0 = +18$ volts in Figure 10,

$$K_N = \frac{12 \times 10^{-3}}{(18-3)^2} = 5.33 \times 10^{-5} \text{ amp/volt}^2,$$

$$K_P = \frac{8 \times 10^{-3}}{(18-4)^2} = 4.1 \times 10^{-5} \text{ amp/volt}^2.$$

The setup shown in Figure 11 was used to measure the response experimentally. The value of capacitance added to the circuit was measured accurately to be 0.014 μf, which is large enough to nullify

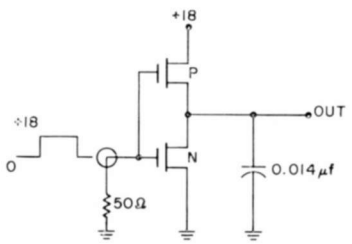


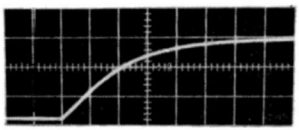
Fig. 11—Test set up for transient response.

the device capacitance as well as any stray wiring capacitance. For $\alpha_P = 4/18 = 0.222$, the rise time is, from Equation (15),

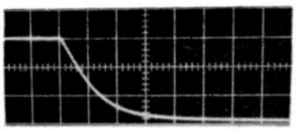
$$T_R = 1.9\,\tau_P = \frac{1.9 \times 1.4 \times 10^{-8}}{73.8 \times 10^{-5}} = 36\ \mu\text{sec},$$

while for $\alpha_N = 3/18 = 0.167$,

$$T_F = 1.7\,\tau_N = \frac{1.7 \times 1.4 \times 10^{-8}}{96.4 \times 10^{-5}} = 25\ \mu\text{sec}.$$



(A) TURN-OFF RESPONSE



(B) TURN-ON RESPONSE

Fig. 12—Experimental results.

These computed results are in very good agreement with the experimental results shown in Figure 12.

Computer Analysis of Transient Response

One of the most important parameters for any logical inverter is the circuit pair delay, which is usually defined as the time delay experienced by the signal in traversing two inversions as measured at the 50% points on the waveform. Obviously, this cannot be determined from the simplified step-response analysis of the previous section, since it requires cascading a number of identical stages and determining the delay between alternate stages. The outputs of the individual

circuits, and thus the inputs to the succeeding stages, are complicated functions of time that are dependent on the device parameters.

For the pair delay time to have precise meaning, the shape of the output waveform must be constant after a signal has propagated through a finite number of stages, the only difference in the waveforms being a time delay that is also constant. This is known as a *standardized signal*, since the shape of this waveform and the pair delay time are sufficient to completely define the performance of the inverter. The following analysis was done using the RCA 601 computer to determine these time parameters of the N-P inverter.

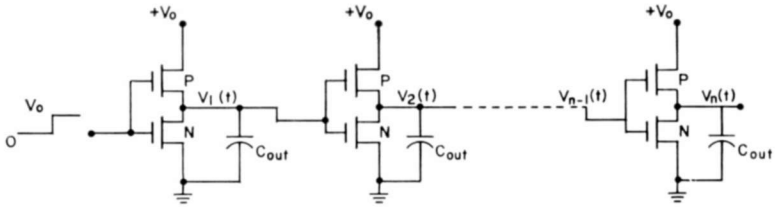


Fig. 13—Cascaded N-P inverters.

The basic problem is to determine the time response of n cascaded N-P inverting stages (see Figure 13) where n is chosen large enough for a given input signal to be standardized. The differential equation for the output V_m of the m^{th} stage in the chain (whose input is V_{m-1}) is

$$C_{\text{out}} \frac{dV_m}{dt} = I_{P_m} - I_{N_m}, \tag{17}$$

where I_{P_m} and I_{N_m} are given in Equations (1) and (2) for $V_G = V_{m-1}$.

By defining the normalized variables as

$$v_m = \frac{V_m}{V_0}, \quad \alpha_N = \frac{V_{T_N}}{V_0}, \quad \alpha_P = \frac{|V_{T_P}|}{V_0}, \quad t' = \frac{t}{\tau}, \quad \tau = \frac{C_{\text{out}}}{K_N V_0} = \tau_N, \tag{18}$$

$$\beta = \frac{K_P}{K_N} = \frac{\tau_N}{\tau_P}, \qquad i_N = \frac{I_N}{K_N V_0^2}, \qquad i_P = \frac{I_P}{K_N V_0^2},$$

Equation (17) reduces to

$$\frac{dv_m}{dt'} = i_P - i_N, \tag{19}$$

with the normalized currents i_P and i_N now given by

$$i_P = 2\beta(1 - v_m)(1 - v_{m-1} - \alpha_P) - \beta(1 - v_m)^2;$$
$$v_m > v_{m-1} + \alpha_P, \tag{20a}$$

$$i_P = \beta(1 - v_{m-1} - \alpha_P)^2; \quad v_m \leq v_{m-1} + \alpha_P, \tag{20b}$$

$$i_P = 0; \quad v_{m-1} + \alpha_P \geq 1, \tag{20c}$$

and

$$i_N = 2v_m(v_{m-1} - \alpha_N) - v_m^2; \quad v_m < v_{m-1} - \alpha_N, \tag{21a}$$

$$i_N = (v_{m-1} - \alpha_N)^2; \quad v_m > v_{m-1} - \alpha_N, \tag{21b}$$

$$i_N = 0; \quad v_{m-1} \leq \alpha_N. \tag{21c}$$

Equations (19), (20), and (21) represent n simultaneous first-order differential equations in the variables $v_1 \rightarrow v_n$ which can range from 0 to 1. These variables are functions of the normalized time as given by Equation (18d) so that the results are independent of the absolute parameters of the device. The results may then be scaled up or down depending on the particular devices.

The computer program was set up for $n = 6$, i.e., 6 cascaded stages, with a unit step applied to the input of the first stage. The parameters α_N and α_P, which represent the normalized threshold voltages of the N and P transistors, were varied from 0 to 0.6 in steps of 0.1, while β, the ratio of the device constants, was varied from 0.2 to 1 in steps of 0.2.* The complete solutions of $v_1(t) \rightarrow v_6(t)$ are shown in Figures 14 and 15 for two different conditions of the parameters α and β. The first case of identical N and P transistors with zero threshold voltages represents the fastest possible circuit in this configuration, and the second (note the different time scale) represents one of the slower, where $\alpha_N = 0$, $\alpha_P = 0.6$, and $\beta = 0.4$. This is to be expected from the step response, since it has been noted previously that the transient

* β was restricted to values < 1 since the cases for $\beta > 1$ can be handled by redefining $\tau = \tau_P$ and interchanging the rising and falling signal waveforms.

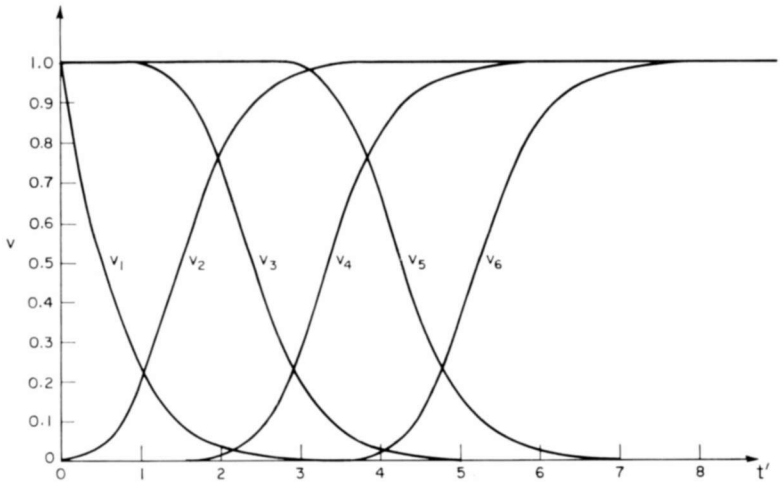


Fig. 14—Solution for $\alpha_N = 0$, $\alpha_P = 0$, $\beta = 1$.

response improves as the amount of current available from the transistors increases. In the second case, Figure 15, the current available from the P transistor is down by a factor of $1/\beta(1 - \alpha_P)^2$ (see Equation (20)), which in this particular case is 15.6.

To get a better feeling for the actual times involved, the value of τ for a typical MOS transistor may be obtained from

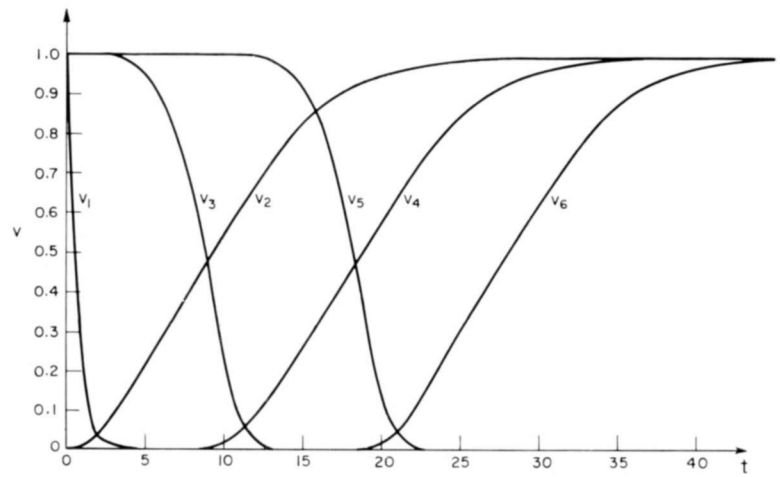


Fig. 15—Solution for $\alpha_N = 0$, $\alpha_P = 0.6$, $\beta = 0.4$.

$$\tau = \frac{C_{out}}{K_N V_0} = \frac{2LT_{ox}C_{out}}{\epsilon_{ox}\mu_N W V_0}. \tag{22}$$

For the MOS transistors under consideration, $L = 10^{-3}$ cm, $W = 12.5 \times 10^{-3}$ cm, $\epsilon_{ox} = 4\epsilon_0 = 0.35 \times 10^{-12}$ farad/cm, $T_{ox} = 10^{-5}$ cm, $\mu_n = 300$ cm^2/volt-sec. With $V_0 = +15$ volts (typical value)

$$\tau = 1000\, C_{out} \text{ sec.}$$

Thus, the MOS transistor has an effective switching resistance of 1000 ohms for $V_0 = +15$ volts. Assuming N and P transistors with input and output capacitances of 2 $\mu\mu$f and 5 $\mu\mu$f each, a circuit fan-out of 1, and stray wiring capacitance of 1 $\mu\mu$f,

$$C_{out} = (2 \times 2) + (2 \times 5) + 1 = 15\ \mu\mu\text{f},$$

or

$$\tau = 15 \text{ nanoseconds.}$$

In Figure 14 the pair delay is 1.8τ, or 27 ns (nanoseconds), with identical rise and fall times of 1.6τ or 24 ns; while in Figure 15 the pair delay is 9.4τ, or 140 ns, with a rise time of 14τ (210 ns) and fall time of 5τ (75 ns). The curves apply equally well for any values of C_{out}, K_N, or V_0, since changes in these values only scale the time up or down proportionally.

With the standardized signal, the pertinent response times are the pair delay* T_D and the rise and fall times, T_R and T_F, between the 10% and 90% points on the standardized waveforms. The pair delay is plotted as a function of the threshold parameters α_N and α_P for $\beta = 1$, 0.6, and 0.2 in Figures 16-18. An approximate relationship for T_D (which holds for $\alpha_N + \alpha_P < 1$, $\beta \geq 0.2$) obtained from the curves is

$$T_D \approx 0.9\tau \left(\frac{1}{(1-\alpha_N)^2} + \frac{1}{\beta(1-\alpha_P)^2} \right) \tag{23}$$

This approximation is within 10% of the actual values obtained.

The quantities $(1-\alpha_N)^2$ and $\beta(1-\alpha_P)^2$ from Equations (20) and (21) are seen to be the maximum values of normalized currents, i_N and i_P, respectively (i.e., for $v_{in} = v_{out} = 1$ and $v_{in} = v_{out} = 0$); so that T_D is given by

* The pair delay is independent of the point on the signal at which it is measured, by definition of a standardized signal.

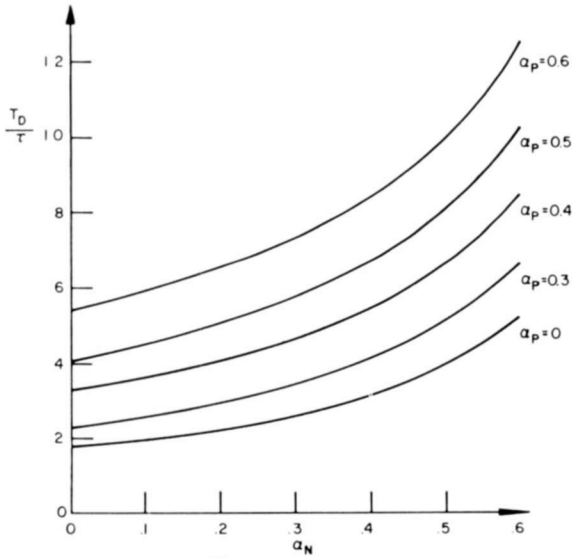


Fig. 16— $\frac{T_D}{\tau}$ versus α_N, α_P; $\beta = 1$.

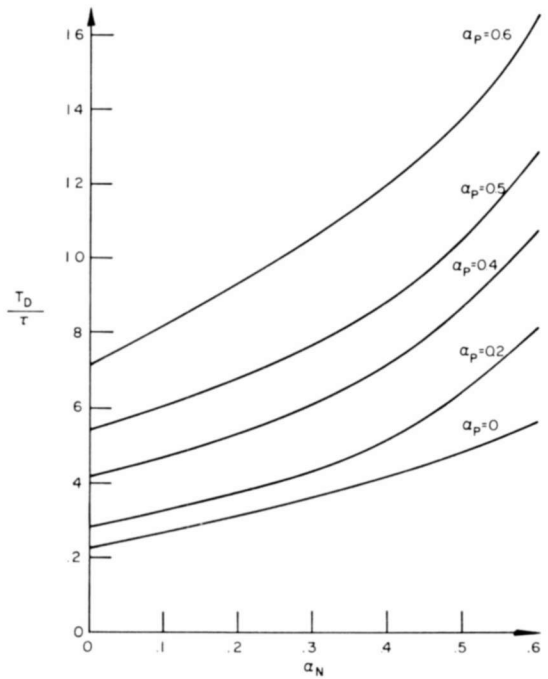


Fig. 17— $\frac{T_D}{\tau}$ versus α_N, α_P; $\beta = 0.6$.

$$T_D \cong 0.9\tau \left(\frac{1}{(i_N)_{\max}} + \frac{1}{(i_P)_{\max}} \right).$$

Intuitively, this is correct since the pair delay must depend on the response of two stages, one of which is primarily determined by the

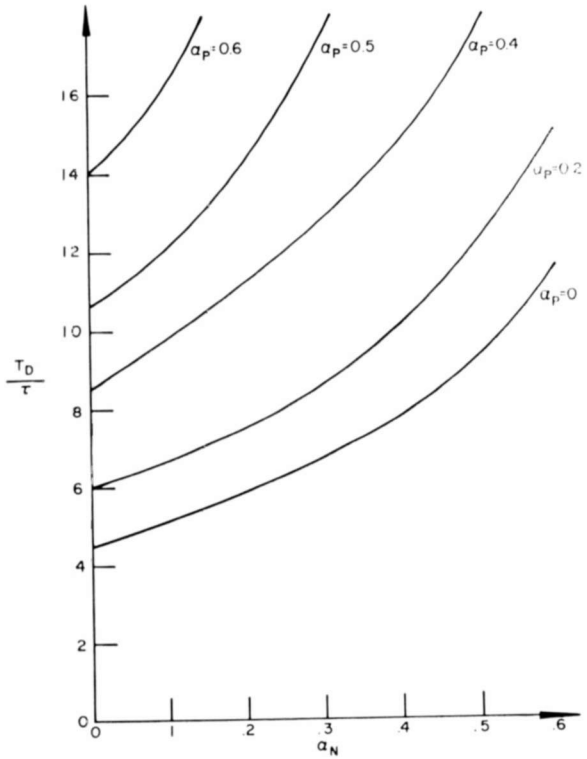


Fig. 18— $\frac{T_D}{\tau}$ versus α_N, α_P; $\beta = 0.2$.

N transistor, the other by the P transistor. If one of the transistors is inherently much faster (higher $i_{\max}$), then the performance of the circuit is determined by the slower of the two transistors.

The rise and fall times are related to the parameters in a somewhat less straightforward manner. With N and P transistors of approximately the same characteristics, the 10% to 90% times are about the same as for a step input (see Figure 14). However, for badly mismatched transistors, as in Figure 15, the response of the faster unit

is deteriorated appreciably. This is seen in Figure 19 where the normalized fall time T_F/τ is shown as a function of α_N and α_P for identical transistors ($\beta = 1$). Note that for small values of α_P, T_F/τ varies little from the step response (dashed curve) of the N transistor. When $\alpha_P = 0.6$, however, the fall time remains substantially independ-

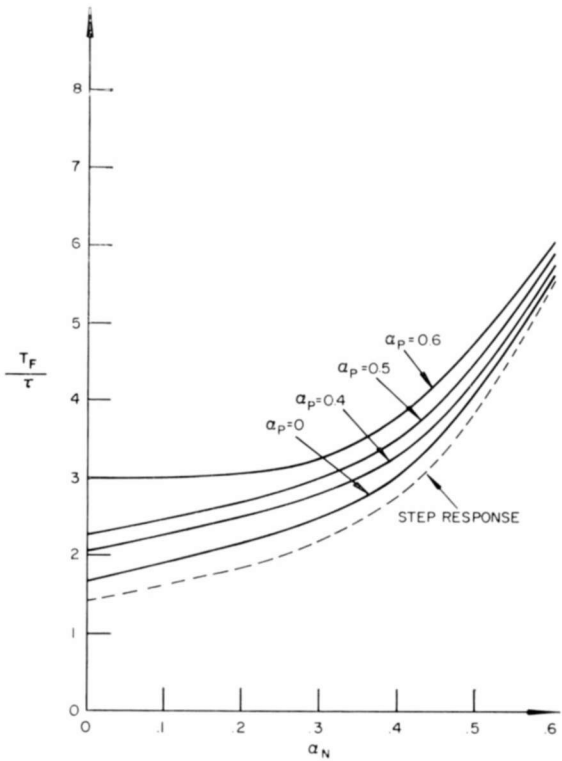


Fig. 19— $\frac{T_F}{\tau}\left(\frac{T_R}{\tau}\right)$ versus α_N, α_P; $\beta = 1$.

ent of α_N until α_N becomes larger than 0.3. Again the circuit response is determined by the slower of the two transistors. The normalized rise time T_R/τ for $\beta = 1$ can also be obtained from Figure 19 by interchanging α_N and α_P on the curves. For $\beta = 0.2$, corresponding to a transistor mismatch by a factor of 5, the time response is completely determined by the P transistor as seen in Figures 22 and 23. The rise time (Figure 23) is essentially equal to the step response for all values of α_N, while the fall time, which bears little relation to the step response, is determined almost entirely by α_P. The rise and fall times for $\beta = 0.6$ are seen in Figures 20 and 21.

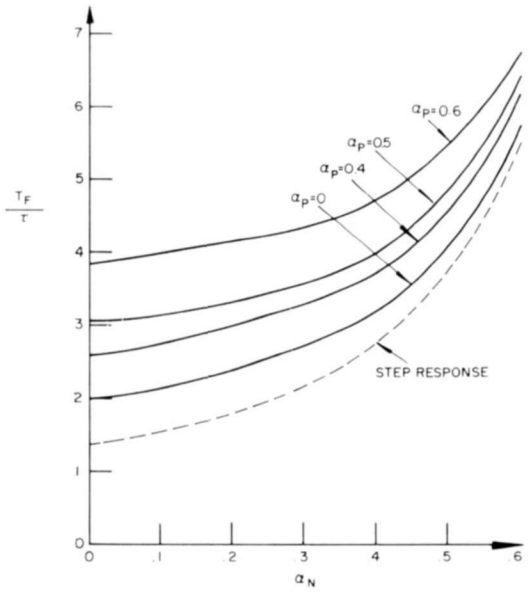


Fig. 20— $\frac{T_F}{\tau}$ versus α_N, α_P; $\beta = 0.6$.

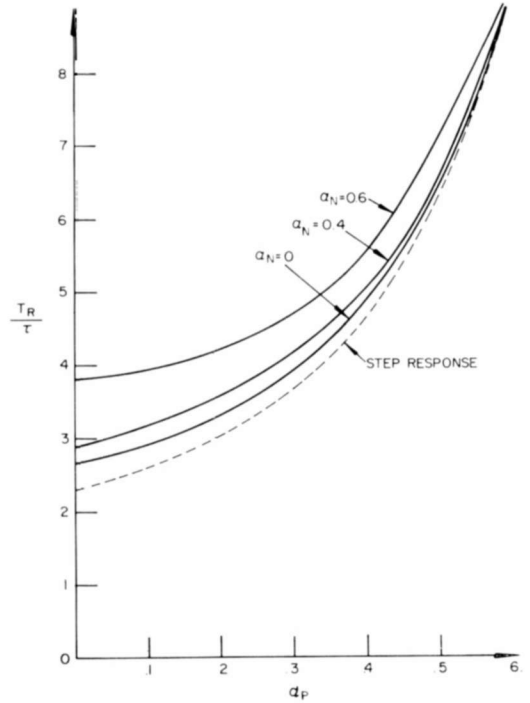


Fig. 21— $\frac{T_R}{\tau}$ versus α_P, α_N; $\beta = 0.6$.

The condition for equal rise and fall times, as determined from Figures 19 through 23, is approximately

$$\beta(1-\alpha_P)^2 \cong (1-\alpha_N)^2. \tag{24}$$

Multiplying both sides by $K_N V_0^2$, the condition becomes $(I_P)_{\max} \cong (I_N)_{\max}$, indicating that the rise and fall times depend on the maximum

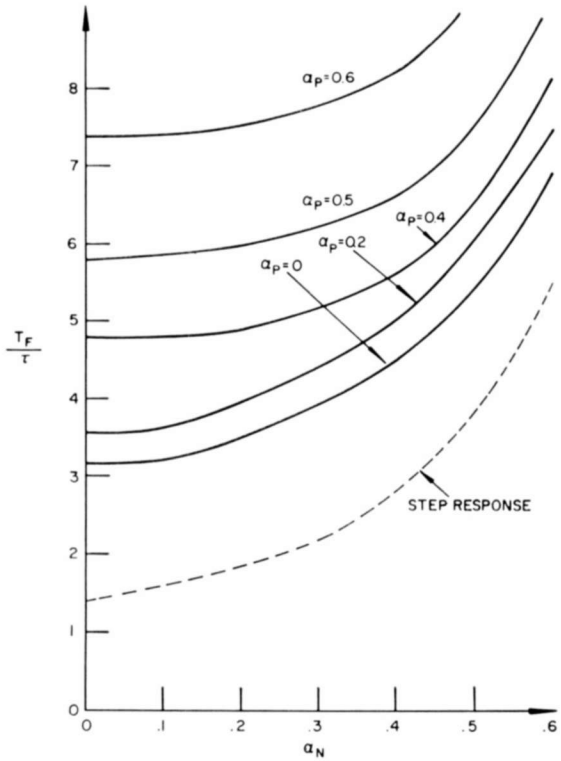


Fig. 22— $\dfrac{T_F}{\tau}$ versus α_N, α_P; $\beta = 0.2$.

available current. For $(I_P)_{\max} > (I_N)_{\max}$, $T_R < T_F$; while for $(I_P)_{\max} < (I_N)_{\max}$, $T_R > T_F$. Thus, a given circuit can be designed (and fabricated) with equal rise and fall times by varying β and/or α_N and α_P.*

The time parameters, and in particular the pair delay times, were measured in a chain of three identical inverting stages whose N and P

* In an integrated circuit, the geometries of the N and P transistors may be the same so that $\beta = \mu_P/\mu_N$ is a constant. The condition can then be met by varying the threshold voltages with appropriate impurity doping of the material.

transistors have the following characteristics:

$$K_N = 5.2 \times 10^{-5} \text{ amp/volt}^2,$$

$$K_P = 26 \times 10^{-5} \text{ amp/volt}^2,$$

$$V_{T_N} = +2 \text{ volts},$$

$$V_{T_P} = -5 \text{ volts}.$$

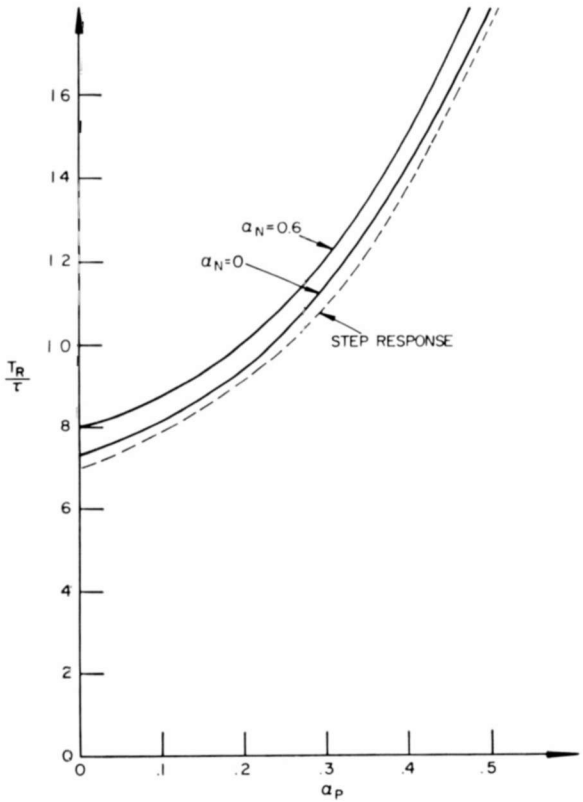


Fig. 23—$\frac{T_R}{\tau}$ versus α_P, α_N; $\beta = 0.2$.

The output of each stage was loaded with a capacitance of 700 $\mu\mu$f to eliminate stray effects. The input to the first stage was a variable-rise-time pulse of height $+ V_0$ volts. In the experiment, the rise time of this pulse was varied until the shape of the pulse and that of the second-stage output were nearly the same, thereby simulating a standardized signal. The pair delay time was then measured between the

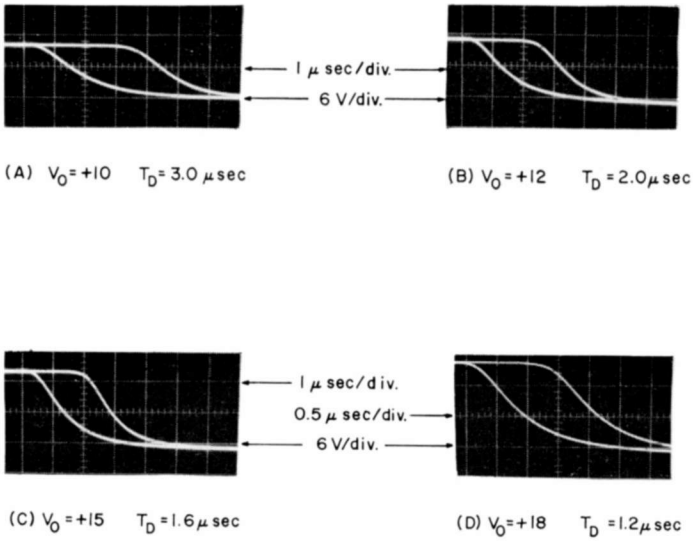


Fig. 24—Pair delay measurements.

outputs of the first and third stages. The measurement was repeated for $V_0 = +10$, $+12$, $+15$ and $+18$ volts; the output waveforms of the first and third stages for these four conditions are shown in Figure 24. Calculated values for T_D, obtained using $\tau = C_{out}/(K_N V_0) = 13.4/V_0$ μsec and $\beta = K_P/K_N = 5$ in Equation (23) are given as a function of V_0 in Table I. The computed results are in excellent agreement with the experimental results of Figure 24.

Power Dissipation

Power dissipation is an extremely important consideration in integrated circuits, and when large numbers of components are involved, it has been shown to be the ultimate limitation in device packing

Table I

V_0 (volts)	α_N	α_P	T_D (calculated) (μsec)	T_D (experimental) (μsec)
10	0.2	0.5	2.9	3.0
12	0.167	0.417	2.1	2.0
15	0.133	0.333	1.5	1.6
18	0.111	0.278	1.1	1.2

density.[3] The N-P inverter is, therefore, extremely advantageous in some applications, since the standby power dissipation is only of the order of microwatts. However, when the circuit switches, a considerable amount of current flows during the transition time. In fact, for

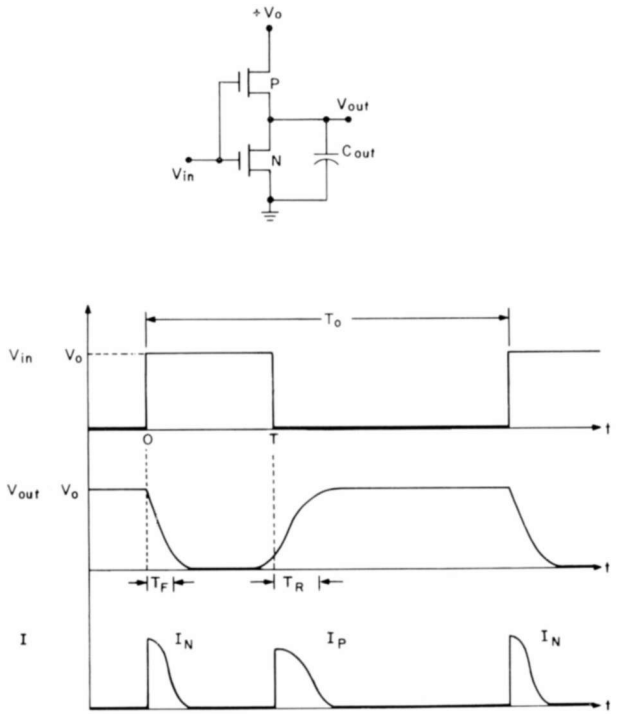


Fig. 25—Model for evaluating power dissipation.

the transition time to be a minimum, it is desirable for the current to be as large as possible. Thus, at very high repetition rates, there is still the fundamental conflict between low power dissipation and high speed.

A measure of the power dissipation of the circuit as a function of frequency can be obtained from the model shown in Figure 25, which assumes step inputs. The average power for the square-wave input voltage shown (repetition rate $f_0 = 1/T_0$) is

[3] J. T. Wallmark, "Design Considerations of Integrated Electronic Devices," *Proc. IRE*, Vol. 48, p. 293, March 1960.

$$P = \frac{1}{T_0}\int_0^{T_0/2} I_N(t)\, V_{\text{out}}\, dt + \frac{1}{T_0}\int_{T_0/2}^{T_0} I_P(t)\,(V_0 - V_{\text{out}})\, dt. \qquad (25)$$

For P, with $I_N(t) = I_P(t) = C_{\text{out}}\,(dV_{\text{out}}/dt)$ (for step inputs only),

$$P = \frac{C_{\text{out}}}{T_0}\int_0^{V_0} V_{\text{out}}\, dV_{\text{out}} + \frac{C_{\text{out}}}{T_0}\int_{V_0}^{0} (V_0 - V_{\text{out}})\, d(V_0 - V_{\text{out}}), \qquad (26)$$

$$P = \frac{C_{\text{out}} V_0^2}{T_0} = C_{\text{out}} V_0^2 f_0. \qquad (27)$$

Thus, for a step input, the average power dissipated is directly related to the energy required to charge and discharge the circuit capacitance to the supply voltage V_0. It should be noted that this power is independent of the device parameters. However, the repetition rate must be less than some maximum value, $f_{\max}$, the frequency for which the voltage swing of the circuit is less than V_0; $f_{\max}$ and, therefore, $P_{\max}$ are determined by the rise and fall times, which are functions of the device parameters. When the input has a nonzero rise time, the situation is not as clear cut, since not all the current is used for charging the capacitance. In particular, consider an N-P inverter with normalized parameters α_N, α_P, and β driven by an input signal of rise time T_R and fall time T_F. For simplicity, assume that the MOS transistor characteristics are those of ideal pentodes where

$$i_N = (v_{IN} - \alpha_N)^2, \qquad (28a)$$

$$i_P = \beta\,(1 - v_{IN} - \alpha_P)^2, \qquad (28b)$$

so that the normalized voltage and current transfer characteristic are as shown in Figure 26. For a rising input signal, it is clear that for $\alpha_N \leq v_{IN} \leq v_{IN}^*$, current is drawn that is not used for discharging the output capacitance, so that the N transistor has current flowing through it with full supply voltage across the terminals. Thus, there is dissipation in the N transistor in addition to that due to capacitive charging; the amount of this excess power increases with the time spent in the region and the current through the N transistor. Similarly, there is additional dissipation for a falling input signal due to current flow with the full supply voltage across the drain–source

terminals of the P transistor. For slowly varying inputs to the circuit, most of the dissipation will stem from the slow transition through these regions.

The total power dissipation of the circuit was evaluated in the computer program for the chain of N-P inverters defined previously.

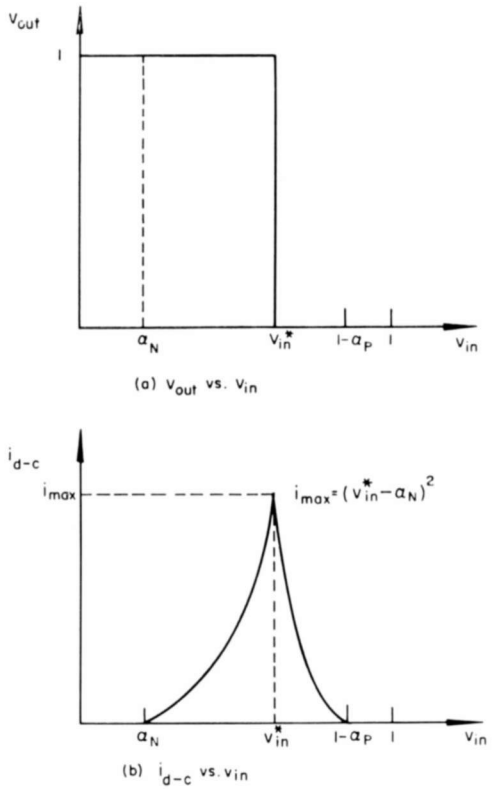


Fig. 26—Normalized transfer characteristics.

This was done by evaluating the energy delivered from the supply voltage for the fifth and sixth stages in the chain whose inputs (and outputs) are the standardized signals of the circuit. The total energy per cycle (i.e., for rising and falling input signals) is then given by

$$E_T = E_{T5} + E_{T6} = V_0 \left[\int_0^\infty I_{P5}(t)\,dt + \int_0^\infty I_{P6}(t)\,dt \right], \qquad (29a)$$

since the current in the P transistor is the same as that from the

supply. Using the normalized parameters of Equation (18), the total energy

$$E_T = C_{\text{out}} V_0^2 \left[\int_0^\infty i_{P5}(t')\,dt' + \int_0^\infty i_{P6}(t')\,dt' \right], \qquad \text{(29b)}$$

which, when multiplied by the frequency of the applied signal, gives the total power dissipation.

The normalized integrals in Equation (29b) were computed for the same values of α_N, α_P, and β as in the time-response analysis (i.e., α_N and α_P ranging from 0 to 0.6 and β ranging in 0.2 to 1). The maximum deviation from the idealized case of a step input, $E_T = C_{\text{out}} V_0^2$, was found to occur for $\alpha_N = \alpha_P = 0$, $\beta = 1$; however, the energy was only 11 percent higher than the ideal case; $(E_T)_{\max} = 1.11\ C_{\text{out}} V_0^2$. Since for all device parameters, $C_{\text{out}} V_0^2 \leq E_T \leq 1.11\ C_{\text{out}} V_0^2$, the power dissipated for any N-P inverting circuit when driven by a standardized signal of frequency f_0 is given, to a good approximation, by

$$P \cong C_{\text{out}} V_0^2 f_0. \qquad (30)$$

The average power dissipation was experimentally determined in a three-stage ring oscillator circuit* (using the same transistors as for the pair-delay measurements) by measuring the total d-c current drawn from the power supply. Each stage was loaded with a capacitance of 700 $\mu\mu$f. The supply voltage was varied from 8 to 18 volts and, in all cases, the measured power was

$$P = V_0 I_{\text{d-c}} \approx 3 C_{\text{out}} V_0^2 f_0,$$

where f_0 is the frequency of oscillation. The agreement between the computed and measured power dissipation is shown in Table II.

Application of Results to Circuit Design

In order to summarize the results presented in the previous sections, it would probably be most instructive to apply them to a hypothetical design problem using the N-P inverter. Suppose that the integrated logic system under consideration required two-level NOR

* The ring oscillator waveforms are the standardized signals defined previously since the waveshape does not vary once the oscillations have built up.

logic and that the basic NOR gate must have the following performance features:

(1) fan-in = fan-out = 5,

(2) circuit pair-delay time as low as possible,

(3) d-c tolerances such as to make $v_{IN}^* \approx 0.5$ for maximum reliability (power dissipation of secondary importance to the speed).

Table II

V_0 volts	f_0 (kilocycles)	$P_{measured} = V_0 I_{d-c}$ (milliwatts)	$P_{comp} = 3C_{out}V_0^2 f_0$ (milliwatts)
8	57	7.6	7.7
10	100	24	21
12	125	41	38
14	167	70	69
16	200	102	107
20	250	200	210

In addition to the above requirements, assume that for maximum yield of the transistors, the construction must be such that $\beta = 0.2$. However, the threshold voltages may be varied (by doping the channels) from 2 to 6 volts. Furthermore, because of voltage breakdown at the gates of the transistors, the maximum supply voltage that may be used is 15 volts.

The two alternative approaches, the series P and series N connections, are shown in Figure 27. Clearly, whichever arrangement is chosen, V_0 must be 15 volts for maximum speed, i.e., $\tau = C_{out}/K_N V_0$ is a minimum. In the series P connection the current available at turn off is lower than that of a single transistor due to the added series resistance of the other P transistors. As shown previously, a good approximation to the current available is obtained by dividing the current available from a single transistor by the number in the series string, which in this case is 5, the circuit fan-in. The effective β of the circuit is then the individual β divided by 5 so that $\beta_{eff} = \beta/5 = 0.04$.

The pair delay time will be dominated by the P transistors and will be of the order of $\tau/\beta_{eff} = 25\tau$ (see Equation (13)) in the worst case with one input present.

On the other hand, the effective β in the series N connection is unity for one input since the amount of turn-off current from the N transistors is scaled down by 5 to equal the turn-on current from the P transistors. In this case, the pair delay time is of the order of 5τ, an improvement by a factor of 5 over the series P connection.

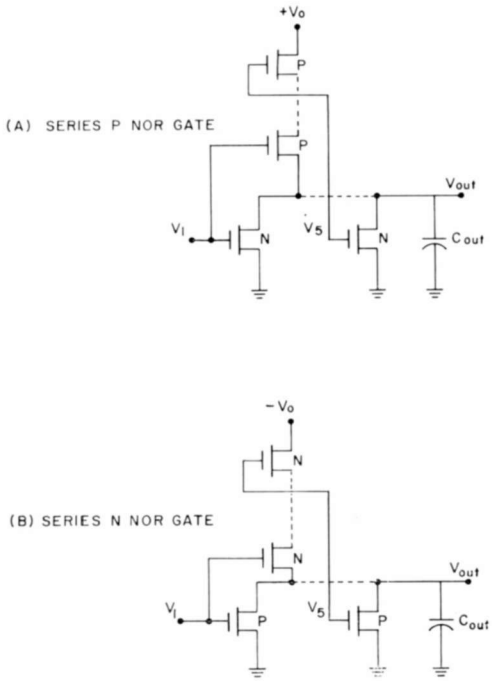


Fig. 27—NOR logic circuits.

Once the circuit configuration is chosen, the remaining condition to be satisfied is that of v_{IN}^*, the fractional threshold voltage of the circuit. The worst case, from a d-c standpoint, occurs when all five inputs are present, since the currents from the five shunt P transistors are additive. Therefore, the effective β for d-c in the worst case is 5, and the normalized equation for v_{IN}^* is

$$\beta_{\text{eff}} (|v_{IN}^*| - \alpha_P)^2 = (1 - |v_{IN}^*| - \alpha_N)^2. \tag{31}$$

Substituting $\beta_{\text{eff}} = 5$, $|v_{IN}^*| = 0.5$,

$$\alpha_N = \sqrt{5}\,\alpha_P - 0.5\,(\sqrt{5} - 1). \tag{32}$$

Since the pair delay time is an increasing function of both α_N and α_P, we choose $V_{T_N} = 2$ volts and $\alpha_N = 2/15 = 0.133$ to optimize the time response. Solving Equation (32), $\alpha_P = 0.333$ or $|V_{T_p}| = 5$ volts. From Figure 16 or Equation (23), the worst-case pair delay time (corresponding to one input present or $\beta = 1$) is 3.4τ where τ is five times larger than the τ of a single transistor, since there are five N transistors in series. For $K_N = 6.67 \times 10^{-5}$ amp/volt^2 (single transistor), $V_0 = -15$ volts.

$$T_D = \frac{5 \times 3.4\, C_{\text{out}}}{6.67 \times 10^{-5} \times 15} = 1.7 \times 10^4\, C_{\text{out}} \text{ sec.}$$

Again taking the output capacitance of each transistor as 5 $\mu\mu$f and the gate capacitance as 2 $\mu\mu$f and neglecting the output capacitances of the series N transistors,

$$C_{\text{out}} = (5 \times 5) + (5 \times 4) = 45\ \mu\mu\text{f for a fan-out of five.}$$

Thus

$$T_D = 1.7 \times 10^4 \times 4.5 \times 10^{-11} = 765 \text{ nanoseconds.}$$

The maximum repetition rate is $T_D/2 = 650$ kc for the two-level logic system described. However, for reliable operation, $f_{\max} = 500$ kc is a more realistic figure. The power dissipation per gate at the frequency $f_{\max}$ is given by

$$P = C_{\text{out}}\, V_0^2\, f_{\max} = 5 \text{ milliwatts.}$$

A summary of the performance of this gate is as follows:

fan-in =	5
fan-out =	5
pair delay time =	765 nanoseconds
maximum repetition rate =	500 kilocycles
power dissipation =	5 milliwatts
supply voltage =	−15 volts
transition voltage =	−7.5 volts

Summary of Circuit Performance

Power Dissipation

The standby power in either binary state is extremely low, being of the order of microwatts. Thus, appreciable power is dissipated only during switching (see Equation (27)). Since power dissipation during switching is directly proportional to the average repetition frequency, f_0, and since f_0 is normally low, this power is still quite small. This is particularly attractive in registers and other low-duty-cycle circuits.

Speed

An approximation of pair delay time of the N-P inverter in a logic chain in terms of the transistor characteristics has been obtained from the analysis (Equation (23)),

$$T_D = 0.9\tau \left[\frac{1}{(1-\alpha_N)^2} + \frac{1}{\beta(1-\alpha_P)^2} \right],$$

where τ is the characteristic time constant of the device (see Equation (22)), α_N and α_P are the normalized threshold voltages of the N and P transistor, respectively, and $\beta = K_P/K_N$ is the relative gain of the transistors. The pair delay time is modified for NAND and NOR gates by an increase or decrease in the device constants K_P and K_N plus an increase in the total output capacitance C_{out}. For an n input NOR gate with series P transistors, $(K_P)_{eff} = K_P/n$ and $(K_N)_{eff} = K_N$ for one applied input, so that $\beta_{eff} = \beta_0/n$. In a five input–five output NOR gate, which is given as an example, the pair delay is 765 nanoseconds. This will probably be improved by as much as 50% with future transistors having lower source-to-drain capacitances.

Noise Immunity

The d-c transfer characteristic, i.e., V_{out} versus V_{in}, is extremely well suited for digital applications, since it has two regions of low gain separated by a high-gain transition region. The high-gain region occurs at a fractional input voltage, v_{in}^*, which can be adjusted for optimum noise immunity by variation of the normalized threshold voltages of the N and P transistors and/or the relative gain β.

General

Certain trade offs can be made between speed, power, and noise immunity, depending on the application. For example, operation at high repetition rates must be accompanied by increased dissipation, since high speed necessitates the use of a large supply voltage. This causes the power to increase as the cube of the operating voltage for

a linear increase in speed. However, in certain circuits where it is desirable to obtain extremely fast switching but at low frequency, such as a full adder, it is possible to keep the power down even at high voltages. It has also been established that the highest speed for a given operating voltage is obtained with zero threshold voltages on the transistors and $\beta = 1$, which gives $v_{in}^* = 0.5$, the condition for optimum noise immunity. Thus, it is possible to achieve fast switching with no sacrifice in noise immunity and at relatively low power levels. These performance features as well as the reduced costs inherent in integrated-circuit technology make complementary-symmetry logic extremely attractive in many digital systems.

Acknowledgments

The author would like to thank J. J. Gibson and G. B. Herzog for many helpful comments during the course of this investigation.

NONLINEAR-ADMITTANCE MIXERS*

By

L. Becker and R. L. Ernst

RCA Communications Systems Division
New York, N. Y.

Summary—*This paper obtains a uniform theory for a nonlinear-admittance frequency converter consisting of a parallel combination of nonlinear conductance and nonlinear capacitance pumped in time phase, and predicts the maximum amount of capacitance variation that will result in performance no worse than that available from nonlinear conductance alone. A general expression for gain and noise figure is derived, and two special cases are treated in detail: (1) conjugate match at both ports, and (2) resonance at both ports. Of the two, conjugate match is the preferred method of loading. The down-converter cannot have any gain, only loss, and adding nonlinear capacitance to a specified nonlinear conductance decreases the loss due to the nonlinear conductance alone. For a large down-conversion ratio the loss becomes excessive if the nonlinear-conductance component is very small and conversion is primarily the result of nonlinear capacitance. For a small down-conversion ratio the amount of nonlinear capacitance that may be added to a specific nonlinear conductance is limited to a narrow range for an inverting converter, while unlimited nonlinear capacitance can be added to almost any nonlinear conductance without appreciably degrading performance of a noninverting converter.*

It is concluded that a junction diode which has nonlinear conductance can have a small amount of nonlinear capacitance shunting it without deteriorating the down-converter gain and noise figure. Because of its larger power-handling capabilities, it will yield lower intermodulation distortion and larger dynamic range than the point-contact resistive diode.

Introduction

LOCAL OSCILLATOR power applied to a mixer diode determines the dynamic range and the amount of distortion caused by cross-modulation and intermodulation.[1] It has been found that a 1-db increase in local oscillator (or pump) power results in a 1-db increase in linear dynamic range and a 2-db decrease in the intermodulation ratio. Therefore, the use of large pump power will result in a mixer that is relatively immune to distortion caused by interfering signals.

Present resistive mixer diodes use a point contact type of construction that limits maximum permissible pump power.[2] The small break-

* This paper is the result of work conducted on U. S. Army Electronics Laboratories Contract DA-36-039-AMC-02345(E), "Interference Reduction Techniques For Receivers".

[1] "Interference Reduction Techniques for Receivers," USAEL Contract DA-36-039-AMC-02345(E), July 1963-June 1965.

[2] H. C. Torrey and C. A. Whitmer, *Crystal Rectifiers*, Vol. 15, Rad. Lab. Series, M.I.T., McGraw-Hill Book Co., New York, 1948.

down voltage of this type diode (typically 4 to 10 volts) restricts the maximum permissible pump voltage that can be applied without causing excessive noise figure due to breakdown. If large pump power is applied to a point-contact diode, the noise figure becomes intolerable, and the probability of diode burnout becomes prohibitively large.

One technique that will permit the use of high power is the use of circuits that incorporate several diodes. Since each diode receives a smaller signal than in a single-diode circuit, the intermodulation will be lowered. But the large number of diodes required to obtain a significant reduction of intermodulation distortion would result in complicated circuitry at microwave frequencies.

An alternative technique requires the design of diodes that can withstand larger pump powers. Junction diodes, such as varactors, are superior to point-contact diodes in this respect. Varactor diodes can handle in excess of 1 watt of pump power because of their large breakdown voltages. Dynamic ranges approaching 160 db and very low spurious responses are typical of varactor circuits.

However, when used as a nonlinear capacitance, the junction diode does not provide a good down-converter.[3,4] For large frequency-conversion ratios, the simple noninverting down-converter is very lossy and noisy, while the inverting down-converter can provide low loss or gain by sacrificing stability. If a capacitive up-converter is used to obtain gain, low noise figure, and large dynamic range,[5] it must be followed by a resistive mixer, which normally has low dynamic range.

It would appear, therefore, that an optimum way of obtaining a good quality down-converter with a large dynamic range and low distortion is to operate a junction diode in a nonlinear-conductance mode. However, there will almost always be a significant amount of nonlinear capacitance effectively shunting the nonlinear conductance.

Edwards analyzed this parallel combination as pumped in time phase, and obtained a gain expression for conjugate-match conditions.[6] Engelbrecht extended the theory by pumping the nonlinear conductance and nonlinear capacitance in time quadrature.[7] Ross and Skalnik

---

[3] L. A. Blackwell and K. L. Kotzebue, *Semiconductor-Diode Parametric Amplifiers*, Prentice-Hall, Englewood Cliffs, N. J., 1961.

[4] P. Penfield, Jr. and R. P. Rafuse, *Varactor Applications*, The M.I.T. Press, Cambridge, Mass., 1962.

[5] B. S. Perlman and B. B. Bossard, "Efficient High Level Parametric Frequency Converters," *I.E.E.E. Int. Convention Record*, Part 3, p. 98, 1963.

[6] C. F. Edwards, "Frequency Conversion by Means of a Nonlinear Admittance," *Bell Syst. Tech. Jour.*, Vol. 35, p. 1403, Nov. 1956.

[7] R. S. Engelbrecht, "Parametric Energy Conversion by Nonlinear Admittances," *Proc. I.R.E.*, Vol. 50, p. 312, March 1962.

analyzed the general case of arbitrary pumping angle.[8] None of these theories included noise figures of the converters. There was some limited analysis of noise due to the nonlinear conductance and nonlinear capacitance of a welded-contact germanium crystal developed during World War II for down-converters.[9] There have also been a number of theories for gain and noise figure of a tunnel-diode mixer, but all assume a constant junction capacitance.[10-12]

It is the purpose of this investigation to obtain a uniform theory for both gain and noise figure of a nonlinear admittance. The study is limited to the consideration of a parallel combination pumped in time phase because this model describes all types of semiconductor diodes, including crystal rectifiers, varactors, and tunnel diodes. The gain and noise-figure expressions are analyzed in order to determine the various parameters for an optimum nonlinear-admittance down-converter.

The procedure used in the analysis is as follows:

(a) Utilizing time-varying functions for the nonlinear conductance and nonlinear capacitance, a two-frequency set of simultaneous equations is developed for the one-port nonlinear admittance.

(b) The general two-port one-frequency network is analyzed for expressions of input and output admittances and transducer power gain.

(c) The nonlinear admittance set of equations obtained in step (a) is applied to the results of the two-port in step (b) to yield converter gain.

(d) Utilizing the conversion properties of the network given in the preceding steps, noise contributions to the output from the equivalent amplitude-modulated noise sources through the frequency spectrum are calculated. The noise figure is then derived from the basic definition.

---

[8] P. W. Ross and J. G. Skalnik, "Parametric Conversion Using Arbitrary Pumping Angles," *Proc. I.E.E.E.*, Vol. 51, p. 239, Jan. 1963.

[9] H. C. Torrey and C. A. Whitmer, "Welded-Contact Germanium Crystals," *Crystal Rectifiers*, Vol. 15, Rad. Lab. Series, M.I.T., Chap. 13, p. 398, McGraw-Hill Book Co., New York, 1948.

[10] K. K. N. Chang, G. H. Heilmeier, and H. J. Prager, "Low-Noise Tunnel-Diode Down Converter Having Conversion Gain," *Proc. I.R.E.*, Vol. 48, p. 854, May 1960.

[11] C. S. Kim, "Tunnel-Diode Converter Analysis," *I.R.E. Trans.*, Vol. ED-8, p. 394, Sept. 1961.

[12] F. Sterzer and A. Presser, "Stable Low-Noise Tunnel-Diode Frequency Converters," *RCA Review*, Vol. 23, p. 3, March 1962.

Nonlinear-Admittance Down-Converter Equations

The voltages and currents associated with the nonlinear admittance are defined in Figure 1. The total current flowing into the nonlinear admittance, i_{na}, may be expressed as the sum of the current flowing into the nonlinear conductance, i_g, and the current flowing into the nonlinear capacitance, i_c;

$$i_{na} = i_g + i_c. \tag{1}$$

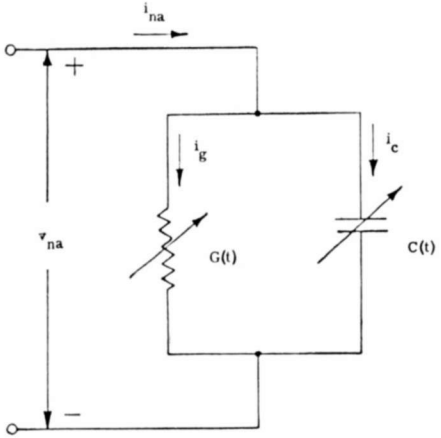


Fig. 1—Equivalent circuit for the nonlinear admittance.

Each current is determined by the voltage across the nonlinear element as given by the relations

$$i_g = G(t)\, v_{na}, \tag{2}$$

and

$$i_c = \frac{d}{dt}\,[C(t)\, v_{na}]. \tag{3}$$

The total current, therefore, can be expressed as

$$i_{na} = G(t)\, v_{na} + \frac{d}{dt}\,[C(t)\, v_{na}]. \tag{4}$$

This nonlinear admittance can be operated as a mixer or down-converter by incorporating it into a circuit such as that shown in

Figure 2. F_s, F_i, and F_p are selective filters that have zero admittance at the signal frequency, f_s, the intermediate frequency, f_i, and the pump frequency, f_p, respectively, and infinite admittance at all other frequencies. In down-conversion, the signal frequency is greater than the intermediate frequency. If it is defined by the relation

$$f_i = f_s - f_p, \tag{5}$$

the down-converter is said to be a noninverting mixer; if

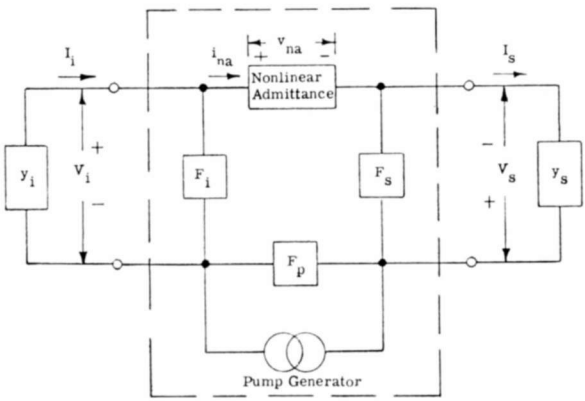


Fig. 2—Nonlinear-admittance down-converter.

$$f_i = f_p - f_s, \tag{6}$$

then it is an inverting mixer. This terminology results from the location of the intermediate frequency spectrum relative to the signal spectrum.

A small-signal analysis is performed that assumes that the pump voltage is much larger than the signal voltage. The time variation of the nonlinear conductance and the nonlinear capacitance is, therefore, essentially determined by the pump voltage alone. Since the conductance and the capacitance are real functions of time, they may be represented by Fourier series as follows:

$$G(t) = \cdots + G^*_2 e^{-j2\omega_p t} + G^*_1 e^{-j\omega_p t} + G_0 + G_1 e^{j\omega_p t} + G_2 e^{j2\omega_p t} + \cdots, \tag{7}$$

$$C(t) = \cdots + C^*_2 e^{-j2\omega_p t} + C^*_1 e^{-j\omega_p t} + C_0 + C_1 e^{j\omega_p t} + C_2 e^{j2\omega_p t} + \cdots, \tag{8}$$

where $\omega_p = 2\pi f_p$. For simplicity, both the conductance and the capacitance will be assumed to be even functions of time. Therefore, $G^*_n = G_n$ and $C^*_n = C_n$ may be substituted into Equations (7) and (8).

The filters in the down-converter circuit restrict the small-signal voltages across the nonlinear admittance to be the sum of the signal and i-f voltages;

$$v_{na} = V_s e^{j\omega_s t} + V^*_s e^{-j\omega_s t} + V_i e^{j\omega_i t} + V^*_i e^{-j\omega_i t}, \tag{9}$$

where V_s and V_i are the signal and i-f voltages, respectively. The components of the current at these frequencies sum to

$$i_{na} = I_s e^{j\omega_s t} + I^*_s e^{-j\omega_s t} + I_i e^{j\omega_i t} + I^*_i e^{-j\omega_i t}, \tag{10}$$

where I_s and I_i are the signal and i-f currents, respectively.

Multiplying Equations (7) and (9) and retaining only terms containing f_s and f_i gives, for the noninverting case,

$$\begin{aligned} i_g = {} & [G_0 V_i + G_1 V_s] e^{j\omega_i t} + [G_1 V_i + G_0 V_s] e^{j\omega_s t} \\ & + [G_0 V^*_i + G_1 V^*_s] e^{-j\omega_i t} + [G_1 V^*_i + G_0 V^*_s] e^{-j\omega_s t}. \end{aligned} \tag{11}$$

Similarly, multiplying Equations (8) and (9) and differentiating the product gives

$$\begin{aligned} i_c = {} & j\omega_i [C_0 V_i + C_1 V_s] e^{j\omega_i t} + j\omega_s [C_1 V_i + C_0 V_s] e^{j\omega_s t} \\ & - j\omega_i [C_0 V^*_i + C_1 V^*_s] e^{-j\omega_i t} - j\omega_s [C_1 V^*_i + C_0 V^*_s] e^{-j\omega_s t}. \end{aligned} \tag{12}$$

Adding the parallel currents obtained in Equations (11) and (12) as indicated by Equation (4) and equating the components to Equation (10) gives

$$I_i = [G_0 + j\omega_i C_0] V_i + [G_1 + j\omega_i C_1] V_s, \tag{13a}$$

$$I_s = [G_1 + j\omega_s C_1] V_i + [G_0 + j\omega_s C_0] V_s. \tag{13b}$$

Repeating the last three steps for the inverting case gives

$$I_i = [G_0 + j\omega_i C_0] V_i + [G_1 + j\omega_i C_1] V^*_s, \tag{14a}$$

$$I^*_s = [G_1 - j\omega_s C_1] V_i + [G_0 - j\omega_s C_0] V^*_s. \tag{14b}$$

These are the basic conversion equations descriptive of the nonlinear admittance mixer.

Two-Port Network Properties

Equations (13) and (14) are in the same form as the admittance equations of the general two-port network shown in Figure 3;

$$I_1 = Y_{11}V_1 + Y_{12}V_2, \tag{15a}$$

$$I_2 = Y_{21}V_1 + Y_{22}V_2, \tag{15b}$$

where I_1, V_1 and I_2, V_2 are the currents and voltages in ports 1 and 2 respectively; and Y_{11}, Y_{12}, Y_{21}, and Y_{22} are the admittances of a two-port.

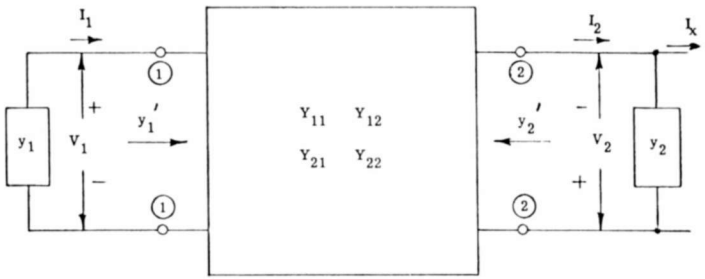


Fig. 3—Two-port network.

Comparing Equations (13) and (14) with (15), it is seen that Equations (13) and (14) use a one-port, two-frequency representation instead of the two-port, one-frequency representation of Equation (15).

Therefore, the network theorems derived for the two-port described by Equation (15), can be utilized for Equations (13) and (14) whenever necessary.

The input admittance, y'_1, at port 1 with y_2 terminating port 2 is

$$y'_1 = Y_{11} - \frac{Y_{12}Y_{21}}{Y_{22} + y_2}. \tag{16}$$

Similarly the admittance, y'_2, looking into port 2 with y_1 terminating port 1 is

$$y'_2 = Y_{22} - \frac{Y_{21}Y_{12}}{Y_{11} + y_1}. \tag{17}$$

The transducer power gain from port 2 to port 1, $P_{G_{21}}$, is the ratio

of power expended in the load y_1 to power available from a source at port 2. Since $y_1 = g_1 + jb_1$ and $y_2 = g_2 + jb_2$, these powers may be easily evaluated. The power dissipated in the output termination is $|V_{1(\text{rms})}|^2 g_1$, where $V_{1(\text{rms})}$ is the root-mean-squared value of the voltage across the conductance g_1. Since V_1 is a Fourier half-amplitude component, it must be multiplied by $\sqrt{2}$ for power calculations. Therefore, the power dissipated in the load y_1 is given by $(\sqrt{2}\,|V_1|)^2 g_1 = 2\,|V_1|^2 g_1$. In a similar fashion, the power available from a current generator of value I_x shunting the admittance y_2 is

$$\text{available power} = \frac{(\sqrt{2}\,|I_x|)^2}{4g_2} = \frac{|I_x|^2}{2g_2}. \tag{18}$$

The transducer power gain is, therefore,

$$P_{G21} = \frac{2|V_1|^2 g_1}{|I_x|^2/2g_2} = 4g_1 g_2 \left|\frac{V_1}{I_x}\right|^2. \tag{19}$$

The ratio V_1/I_x may be determined by realizing that when the current generator is placed at port 2 across an admittance y_2, and port 1 is terminated in an admittance y_1;

$$I_2 = I_x - y_2 V_2, \tag{20a}$$

$$I_1 = -y_1 V_1, \tag{20b}$$

where the minus signs result from the chosen convention for voltage polarities and current directions. By substituting Equation (20) into Equation (15), it is found that

$$\frac{I_x}{V_1} = Y_{21} - \frac{(Y_{11} + y_1)(Y_{22} + y_2)}{Y_{12}}. \tag{21}$$

Substituting the inverse of this ratio into Equation (10) gives the general expression for the two-port transducer power gain;

$$P_{G21} = 4g_1 g_2 \left|\frac{Y_{12}}{Y_{12}Y_{21} - (Y_{11} + y_1)(Y_{22} + y_2)}\right|^2. \tag{22}$$

General Nonlinear-Admittance Properties

The nonlinear-admittance down-converter equations previously derived will be applied to the results of the two-port network to obtain the converter input and output admittances and the transducer power gain.

An equivalence between the two-port admittances and the converter admittances will be established by choosing

$$y_1 = y_i,$$

$$y_2 = y_s. \tag{23}$$

The i-f and signal admittances are defined by

$$y_i = -\frac{I_i}{V_i}, \tag{24a}$$

$$y_s = -\frac{I_s}{V_s} \quad \text{for noninverting case,} \tag{24b}$$

$$= -\frac{I^*_s}{V^*_s} \quad \text{for inverting case,}$$

where these admittances are of the form

$$y_i = g_i + jb_i \tag{25a}$$

$$y_s = g_s \pm jb_s. \tag{25b}$$

The upper sign of multiple symbols ($\pm$ or $\mp$) denotes the noninverting case and the lower sign the inverting case.

Also, the following equalities can be made

$$Y_{11} = G_0 + j\omega_i C_0, \tag{26a}$$

$$Y_{12} = G_1 + j\omega_i C_1, \tag{26b}$$

$$Y_{21} = G_1 \pm j\omega_s C_1, \tag{26c}$$

$$Y_{22} = G_0 \pm j\omega_s C_0. \tag{26d}$$

By substituting Equation (26) into Equation (16), the input admittance, y'_s, of the nonlinear admittance is found to be

$$y'_s = g'_s + jb'_s, \tag{27}$$

where $$g'_s = G_0 - \frac{g_{i0}(G_1^2 \mp \omega_i\omega_s C_1^2) + b_{i0}C_1G_1(\omega_i \pm \omega_s)}{g_{i0}^2 + b_{i0}^2} \tag{28}$$

and $$b'_s = \pm\omega_s C_0 - \frac{g_{i0}C_1G_1(\omega_i \pm \omega_s) - b_{i0}(G_1^2 \mp \omega_i\omega_s C_1^2)}{g_{i0}^2 + b_{i0}^2}, \tag{29}$$

where, for convenience, g_{i0} and b_{i0} are defined by

$$g_{i0} = g_i + G_0, \tag{30}$$

and $$b_{i0} = b_i + \omega_i C_0. \tag{31}$$

Similarly, the output admittance, y'_i, of the nonlinear admittance is

$$y'_i = g'_i + jb'_i, \tag{32}$$

where $$g'_i = G_0 - \frac{g_{s0}(G_1^2 \mp \omega_i\omega_s C_1^2) \pm b_{s0}C_1G_1(\omega_i \pm \omega_s)}{g_{s0}^2 + b_{s0}^2}, \tag{33}$$

and $$b'_i = \omega_i C_0 - \frac{g_{s0}C_1G_1(\omega_i \pm \omega_s) \mp b_{s0}(G_1^2 \mp \omega_i\omega_s C_1^2)}{g_{s0}^2 + b_{s0}^2}, \tag{34}$$

where g_{s0} and b_{s0} are similarly defined by

$$g_{s0} = g_s + G_0, \tag{35}$$

and $$b_{s0} = b_s + \omega_s C_0. \tag{36}$$

By substituting Equations (23), (25), and (26) into Equation (22),

$$P_{G_c} = \frac{4g_i g_s(G_1^2 + \omega_i^2 C_1^2)}{[G_1^2 \mp \omega_i\omega_s C_1^2 - g_{i0}g_{s0} \pm b_{i0}b_{s0}]^2 + [C_1G_1(\omega_i \pm \omega_s) - b_{i0}g_{s0} \mp b_{s0}g_{i0}]^2} \tag{37}$$

This equation is dependent upon the values chosen for the admittances y_i and y_s.

Noise Figure

The noise figure, F_c, of the converter is defined as

$$F_c = \frac{S_s/N_s}{S_i/N_i} = 1 + \frac{N_e}{P_{G_c} N_s}, \tag{38}$$

where S_s = actual signal power delivered to the input,
N_s = actual noise power delivered to the input,
S_i = actual signal power delivered at the output,
N_i = actual noise power delivered at the output,
N_e = excess noise power (noise power delivered at the output, due to internal sources in the network).

The sources of the excess noise are

(a) shot noise or thermal noise in the conductance G_0 at the output frequency, ω_i, and represented by a current generator, I_{ni},

(b) shot noise or thermal noise in the conductance G_0 at the input frequency, ω_s, and represented by a current generator, I_{ns}.

Shot noise is considered in this analysis because the nonlinear conductance will usually be realized by use of a diode. In this case shot noise is proportional to an equivalent direct current, I_{eq}, which is the sum of the forward current, I_f, and the magnitude of the reverse current, I_r. This equivalent current is given by

$$I_{eq} = I_f + |I_r|. \tag{39}$$

The total shot-noise current has a mean-squared value

$$\overline{I_n^2} = 2eI_{eq}\Delta f. \tag{40}$$

In mixer applications, I_{eq} will be determined by the local oscillator voltage and will be of the form

$$I_{eq} = I_{e0} + 2I_{e1}\cos\omega_p t + 2I_{e2}\cos 2\omega_p t + \cdots \tag{41}$$

It can be seen by comparing Equations (40) and (41) that the shot noise becomes amplitude modulated by the pump and will, therefore, be of the form[13]

[13] M. J. O. Strutt, "Noise Figure Reduction in Mixer Stages," *Proc. I.R.E.*, Vol. 34, p. 942, Dec. 1946.

$$I_n(t) = \sum_{n=1}^{\infty} B(t) \cos(\omega_n t + \theta_n), \tag{42}$$

where

$$B(t) = B_0 + 2B_1 \cos \omega_p t + 2B_2 \cos 2\omega_p t + \cdots. \tag{43}$$

It can be shown that

$$\overline{I_n^2} = \frac{B^2(t)}{2} = \frac{A(t)}{2}, \tag{44}$$

where $$A(t) = A_0 + 2A_1 \cos \omega_p t + 2A_2 \cos 2\omega_p t + \cdots. \tag{45}$$

Comparing Equations (40) and (44) shows that

$$\frac{A_0}{2} = 2eI_{e0}\Delta f, \tag{46}$$

$$\frac{A_1}{2} = 2eI_{e1}\Delta f, \text{ etc.} \tag{47}$$

From the definition of $A(t)$ it can be seen that

$$A_0 = B_0^2 + 2B_1^2 + 2B_2^2 + 2B_3^2 + \cdots, \tag{48}$$

$$A_1 = 2(B_0B_1 + B_1B_2 + B_2B_3 + \cdots). \tag{49}$$

The shot noise can be expressed in the following form:

$$I_n(t) = \sum_{n=-\infty}^{\infty} (B_0 + 2B_1 \cos \omega_p t + 2B_2 \cos 2\omega_p t + \cdots) \cos(\omega_n t + \theta_n) \tag{50}$$

$$= \sum_{n=-\infty}^{\infty} \left[B_0 \cos(\omega_n t + \theta_n) + \sum_{r=1}^{\infty} B_r \left\{ \cos[(r\omega_p - \omega_n)t - \theta_n] + \cos[(r\omega_p + \omega_n)t + \theta_n] \right\} \right], \tag{51}$$

and the frequencies of interest will be

$$\omega_n = n\omega_p + \omega_i, \qquad \text{where } \omega_n > n\omega_p, \tag{52a}$$

$$\omega_{-n} = n\omega_p - \omega_i, \qquad \text{where } \omega_{-n} < n\omega_p. \tag{52b}$$

Only the noise contributions at frequencies ω_i and ω_s are of interest. Consider first noise contributions at the frequency band about ω_i. According to Equation (51) there are the following contributions:

(a) Noise generated at frequency ω_i, which is equal to

$$B_0 \cos(\omega_i t + \theta_i);$$

(b) Noise at frequency $(\omega_p - \omega_i)$ that is amplitude modulated by frequency ω_p such that the lower sideband falls within the band ω_i. This component is equal to

$$B_1 \cos\{[\omega_p - (\omega_p - \omega_i)]t - \theta_{-1}\} = B_1 \cos(\omega_i t - \theta_{-1}).$$

This general notation will be retained for the noninverting case. For the inverting case, $\omega_p - \omega_i = \omega_s$ and $\theta_{-1} = \theta_s$.

(c) Noise at frequency $(\omega_p + \omega_i)$ that is amplitude modulated by frequency ω_p such that ω_i receives a contribution of $B_1 \cos(\omega_i t + \theta_1)$. This general notation will be retained for the inverting case. For the noninverting case,

$$\omega_p + \omega_1 = \omega_s \text{ and } \theta_1 = \theta_s.$$

(d) Noise at frequencies $(2\omega_p - \omega_i)$ and $(2\omega_p + \omega_i)$ add components of $B_2 \cos(\omega_i t - \theta_{-2})$ and $B_2 \cos(\omega_i t + \theta_2)$, respectively.

Continuing in this manner the noise at ω_i becomes

$$\begin{aligned} I_{ni} = {} & B_0 \cos(\omega_i t + \theta_i) \\ & + B_1 [\cos(\omega_i t - \theta_{-1}) + \cos(\omega_i t + \theta_1)] \\ & + B_2 [\cos(\omega_i t - \theta_{-2}) + \cos(\omega_i t + \theta_2)] \\ & + B_3 [\cos(\omega_i t - \theta_{-3}) + \cos(\omega_i t + \theta_3)] \\ & + \cdots. \end{aligned} \tag{53}$$

A similar procedure gives the following expression for the noise current at ω_s:

$$
\begin{aligned}
I_{ns} = {} & B_0 \cos(\omega_s t + \theta_s) \\
& + B_1 [\cos(\omega_s t \pm \theta_i) + \cos(\omega_s t + \theta_{\pm 2})] \\
& + B_2 [\cos(\omega_s t - \theta_{\mp 1}) + \cos(\omega_s t + \theta_{\pm 3})] \\
& + B_3 [\cos(\omega_s t - \theta_{\mp 2}) + \cos(\omega_s t + \theta_{\pm 4})] \\
& + \cdots
\end{aligned} \tag{54}
$$

where the upper signs of the $\pm$ and $\mp$ are used for the noninverting converter and the lower signs for the inverting converter.

Consider the complete output circuit using a Norton equivalent-circuit representation as shown in Figure 4. Here I_{ngs} represents noise

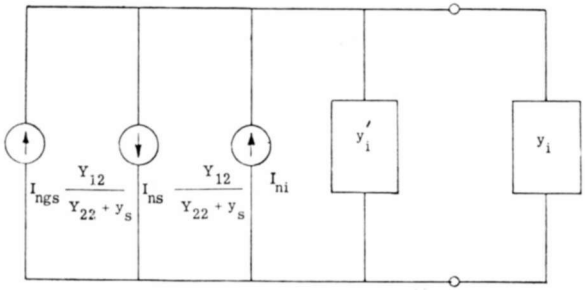


Fig. 4—Output-noise equivalent circuit.

generated in the source conductance and $Y_{12}/(Y_{22} + y_s)$ is the short-circuit current gain. Since all the noise-current generators operate into the same load, the excess noise figure, $F_c - 1$, is equal to the ratio of the squares of the short circuited currents, or

$$
F_c - 1 = \frac{\overline{\left| I_{ni} - I_{ns} \dfrac{Y_{12}}{Y_{22} + y_s} \right|^2}}{\overline{|I_{ngs}|^2} \left| \dfrac{Y_{12}}{Y_{22} + y_s} \right|^2}. \tag{55}
$$

It is therefore necessary to evaluate the expression

$$
\overline{\left| I_{ni} - I_{ns} \frac{Y_{12}}{Y_{22} + y_s} \right|^2}
$$

From Equations (53) and (54) it is seen that the indicated sum is given by

$$I_{ni} - I_{ns}\frac{Y_{12}}{Y_{22}+y_s} = \left(B_0 - B_1\frac{Y_{12}}{Y_{22}+y_s}\right)\cos(\omega_i t + \theta_i)$$
$$+\left(B_1 - B_0\frac{Y_{12}}{Y_{22}+y_s}\right)\cos(\omega_i t \pm \theta_s)$$
$$+\left(B_1 - B_2\frac{Y_{12}}{Y_{22}+y_s}\right)\cos(\omega_i t \mp \theta_{\mp 1})$$
$$+\left(B_2 - B_1\frac{Y_{12}}{Y_{22}+y_s}\right)\cos(\omega_i t \pm \theta_{\pm 2})$$
$$+\left(B_2 - B_3\frac{Y_{12}}{Y_{22}+y_s}\right)\cos(\omega_i t \mp \theta_{\mp 2})$$
$$+\cdots. \tag{56}$$

Taking the mean of the square of the magnitude of this term gives

$$\overline{\left|I_{ni} - I_{ns}\frac{Y_{12}}{Y_{22}+y_s}\right|^2}$$
$$= \frac{1}{2}(B_0^2 + 2B_1^2 + 2B_2^2 + \cdots)\left(1 + \left|\frac{Y_{12}}{Y_{22}+y_s}\right|^2\right)$$
$$-(2B_0B_1 + 2B_1B_2 + 2B_2B_3 + \cdots)\left|\frac{Y_{12}}{Y_{22}+y_s}\right|$$
$$= \frac{A_0}{2}\left(1 + \left|\frac{Y_{12}}{Y_{22}+y_s}\right|^2\right) - 2\left(\frac{A_1}{2}\right)\left|\frac{Y_{12}}{Y_{22}+y_s}\right| \tag{57}$$

For convenience we shall define equivalent noise conductances, G_{n0} and G_{n1}, such that their thermal noise output at the temperature of the nonlinear conductance is equal to the shot-noise output;

$$4kT_G G_{n0}\Delta f = 2eI_{e0}\Delta f = \frac{A_0}{2}, \tag{58}$$

$$4kT_G G_{n1}\Delta f = 2eI_{e1}\Delta f = \frac{A_1}{2}, \tag{59}$$

where k = Boltzmann's constant,

T_G = noise temperature in °K of the nonlinear admittance.

Therefore,

$$\overline{\left| I_{ni} - I_{ns} \frac{Y_{12}}{Y_{22} + y_s} \right|^2} = 4kT_G G_{n0} \Delta f \left[1 + \left| \frac{Y_{12}}{Y_{22} + y_s} \right|^2 \right] - 2(4kT_G G_{n1} \Delta f) \left| \frac{Y_{12}}{Y_{22} + y_s} \right| \tag{60}$$

The signal noise generator has a mean-squared value of

$$\overline{|I_{ngs}|^2} = 4kT_s g_s \Delta f,$$

where T_s is the noise temperature of the source conductance. Noise figure is defined with the source conductance at a standard temperature of

$$T_s = T_0 = 290^\circ \text{K}.$$

The overall noise figure is[14,15]

$$F_c = 1 + \frac{T_i G_{n0}}{T_0 g_s} + \frac{T_G G_{n0}}{T_0 g_s} \left| \frac{Y_{22} + y_s}{Y_{12}} \right|^2 - 2 \frac{T_G G_{n1}}{T_0 g_s} \left| \frac{Y_{22} + y_s}{Y_{12}} \right| \tag{61}$$

When shot-noise effects are not present, only thermal-noise contributions need be considered. The above equation can be applied by substituting

$$G_{n0} = G_0, \tag{62a}$$

$$G_{n1} = 0. \tag{62b}$$

The analysis has yielded general expressions for gain and noise figure, but they are very difficult to apply in their present state. However, if the input and output circuits are specified, the general equations can be simplified to some extent. The two cases that will now be treated are (1) conjugate match at input and output, and (2) resonance at input and output circuits.

For the conjugate match, the source and load admittances will be made equal to the conjugates of their respective input admittances, i.e.,

---

[14] D. G. Peterson, "Tunnel Diode Down Converters," *Proc. I.R.E.*, Vol. 49, p. 1225, July 1961.

[15] S. Lo, "Noise in Tunnel-Diode Mixers," *Proc. I.R.E.*, Vol. 49, p. 1688, Nov. 1961.

$$y_s = y_s'^*$$

$$y_i = y_i'^*.$$

For the resonance case, the susceptances of the source and load admittances will be made equal to the negative susceptances of the respective input admittances, i.e.,

$$b_s = -b_s',$$

$$b_i = -b_i'.$$

Conjugate Match

As previously stated, the input admittances, y_1' and y_2', of a two-port are

$$y_1' = Y_{11} - \frac{Y_{12}Y_{21}}{Y_{22} + y_2}, \tag{16}$$

and

$$y_2' = Y_{22} - \frac{Y_{21}Y_{12}}{Y_{11} + y_1}. \tag{17}$$

These equations may be arranged to form

$$(Y_{11} - y_1')(Y_{22} + y_2) = (Y_{22} - y_2')(Y_{11} + y_1) = Y_{12}Y_{21}. \tag{63}$$

If it is specified that a conjugate match exist at both the signal and the i-f ports, then the source and load admittances are given by

$$y_s' = y_s^* = g_s \mp jb_s, \tag{64}$$

and

$$y_i' = y_i^* = g_i - jb_i. \tag{65}$$

Substituting Equations (26), (64), and (65) into Equation (63) and equating the real parts of the first equality results in the signal and i-f conductances being equal, i.e., $g_i = g_s$. Equating the imaginary parts of the first equality results in the total i-f reactance being equal to the total signal reactance for the noninverting case and the negative of the total signal reactance for the inverting case, i.e.,

$$b_i + \omega_i C_0 = \pm(b_s + \omega_s C_0).$$

Therefore, parameters g and b can be defined as

$$g = g_i = g_s, \tag{66a}$$

$$b = b_i + \omega_i C_0 = \pm (b_s + \omega_s C_0). \tag{66b}$$

Equating the real parts of the second equality and using Equation (66) yields

$$g^2 = G_0^2 - G_1^2 \pm \omega_i \omega_s C_1^2 - b^2. \tag{67}$$

Similarly, for the imaginary parts,

$$b = \frac{C_1 G_1}{2G_0} (\omega_i \pm \omega_s). \tag{68}$$

Substituting Equations (66), (67), and (68) into Equation (22), the transducer power gain for a conjugate match reduces to

$$P_{GCM} = \frac{G_1^2 + \omega_i^2 C_1^2}{(G_0 + g)^2 + b^2}. \tag{69}$$

Similarly, substituting Equations (66), (67), and (68) into Equation (61) and using the same equivalents as before, the noise figure for a conjugate match becomes

$$F_{CM} = 1 + \frac{T_G G_{n0}}{T_0 g_s} \left\{ 1 + \frac{1}{P_{GCM}} - \frac{2}{\sqrt{P_{GCM}}} \frac{G_{n1}}{G_{n0}} \right\}. \tag{70}$$

Loss Contours for Matched Down-Converter

For convenience, the equations will be normalized with respect to G_0. Let

$$\begin{aligned} \frac{\omega_i}{\omega_s} &= p, \quad \frac{b}{G_0} = B, \quad \frac{g}{G_0} = G \\ \frac{\omega_s C_1}{G_0} &= x, \quad \frac{\omega_i C_1}{G_0} = px, \quad \frac{G_1}{G_0} = y. \end{aligned} \tag{71}$$

Then Equations (67), (68), and (69) become

$$B = (p \pm 1) \left(\frac{xy}{2} \right) \tag{72}$$

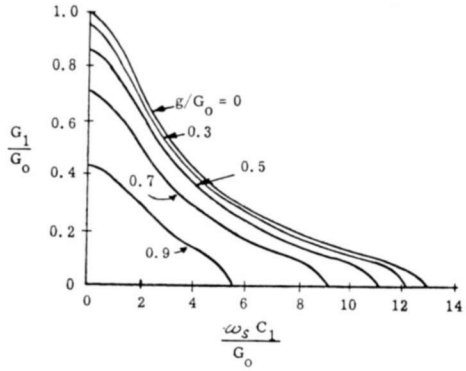


Fig. 5—Conductance contours for inverting down-converter ($\omega_s/\omega_i = 160$).

$$G^2 = 1 - y^2 \pm px^2 - \left[(p \pm 1) \frac{xy}{2} \right]^2, \tag{73}$$

$$P_{GCM} = \frac{y^2 + p^2x^2}{(1+G)^2 + B^2}. \tag{74}$$

The gain expression will be analyzed for $p = 1/160$ and $p = 1/4$. In order to obtain values for gain, the conductance Equation (73) must be specified. Both gain and conductance are functions of the nonlinear conductance, y, and nonlinear capacitance, x.

Figures 5 and 6 show the conductance contours when $p = 1/160$ for the inverting and noninverting down-converters, respectively, while

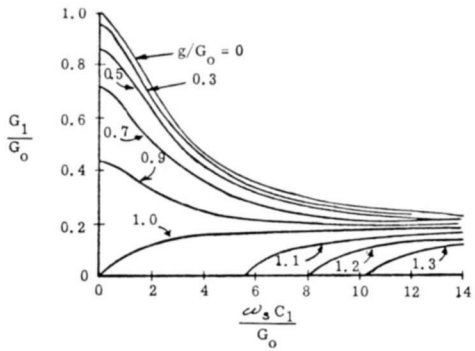


Fig. 6—Conductance contours for noninverting down-converter ($\omega_s/\omega_i = 160$).

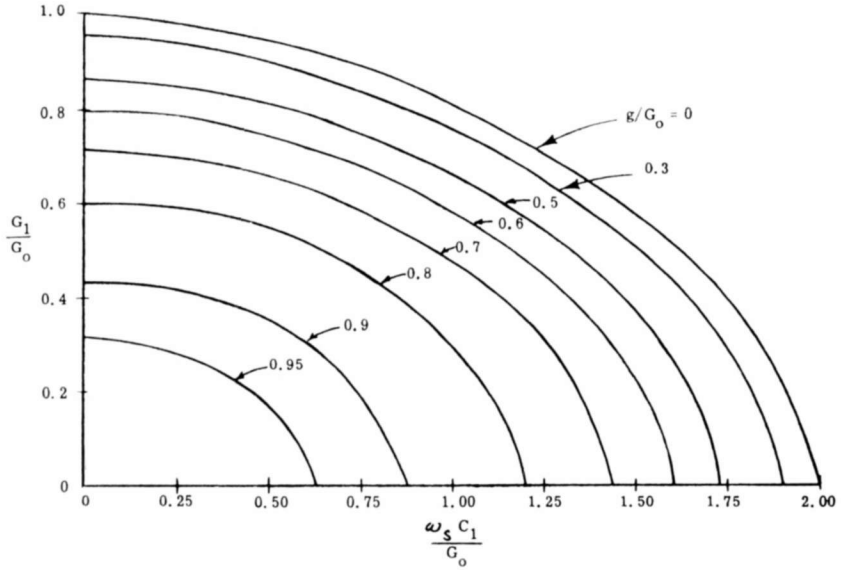


Fig. 7—Conductance contours for inverting down-converter ($\omega_s/\omega_i = 4$).

Figures 7 and 8 correspond to $p = 1/4$. For the inverting down-converter, increasing C_1 requires decreasing source and load conductances. The normalized conductance, G, is always less than unity and becomes

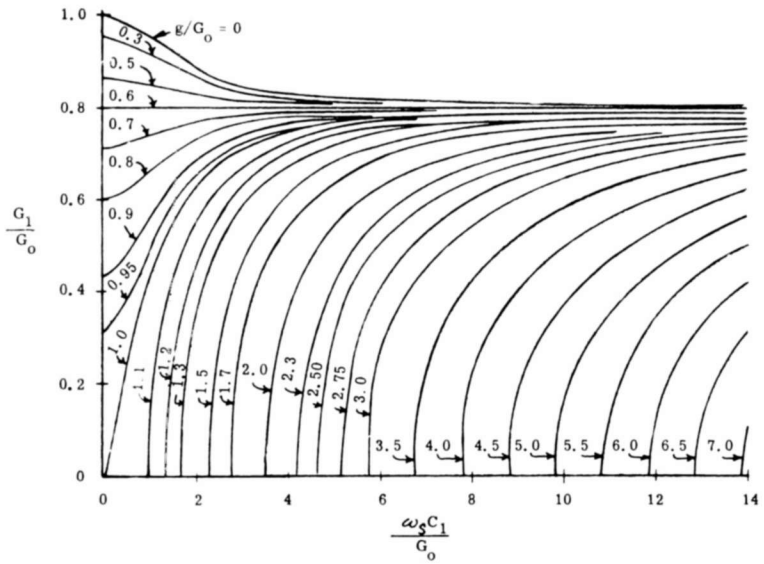


Fig. 8—Conductance contours for noninverting down-converter ($\omega_s/\omega_i = 4$).

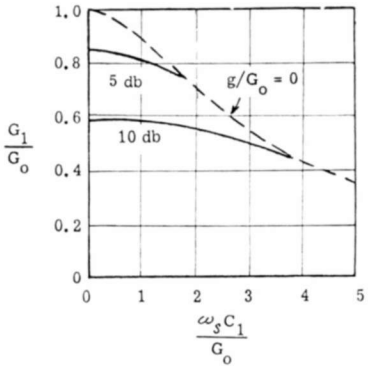


Fig. 9—Loss contours for inverting down-converter ($\omega_s/\omega_i = 160$).

zero when $x = \sqrt{1/p}$. Values of $x > \sqrt{1/p}$ and combinations of x and y above the $G = 0$ line require a negative conductance. A conjugate match is not physically possible when G is negative and operation in that region is not recommended because of stability problems.

For the noninverting down-converter, increasing C_1 also requires a decreasing source and load conductance when $y > 2\sqrt{p}/(p+1)$. For values of $y < 0.157$ for $p = 1/160$, and $y < 0.800$ for $p = 1/4$, increasing C_1 requires increasing source and load conductances. Values of x and y above the $G = 0$ line as before mean a negative conductance, and as with the inverting down-converter, a conjugate match is not possible and there are stability problems.

Figures 9 and 10 are the loss contours when $p = 1/160$ for the inverting and noninverting down-converter, respectively, and Figures 11, 12 and 13 correspond to $p = 1/4$. It is interesting to note that (1)

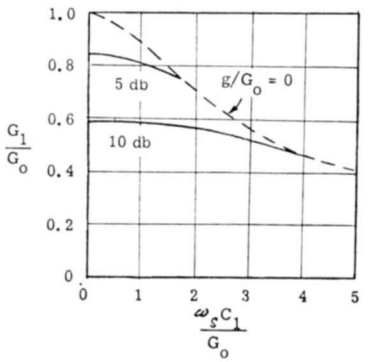


Fig. 10—Loss contours for noninverting down-converter ($\omega_s/\omega_i = 160$).

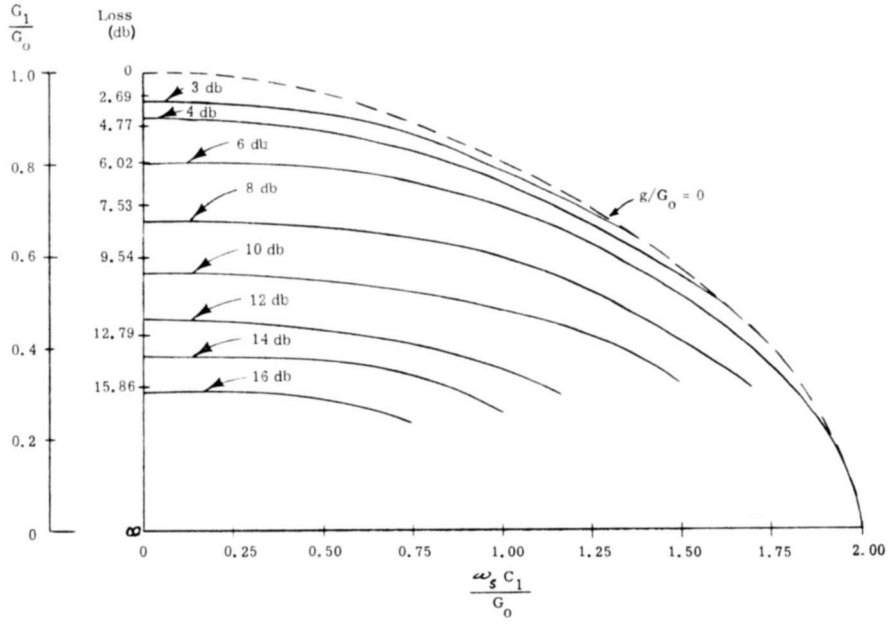


Fig. 11—Loss contours for inverting down-converter ($\omega_s/\omega_i = 4$).

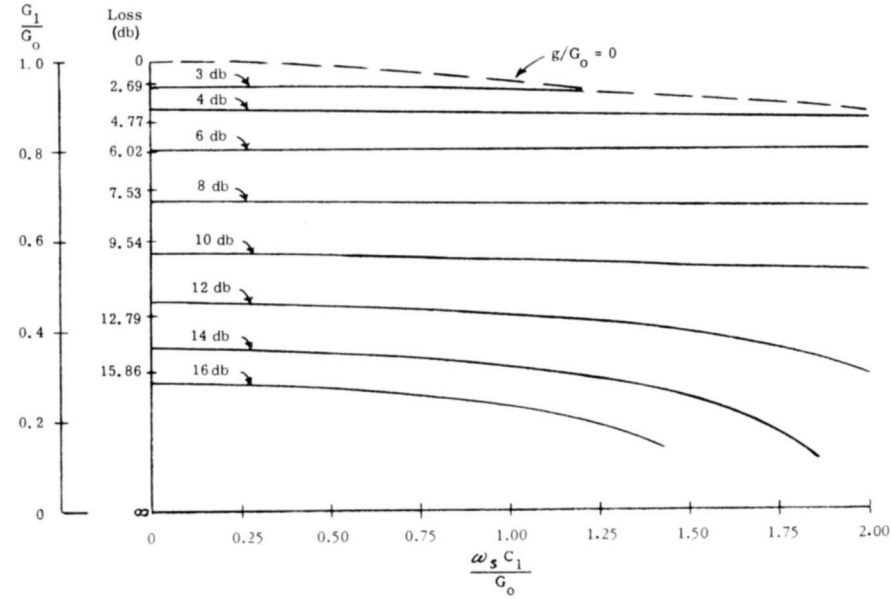


Fig. 12—Loss contours for noninverting down-converter ($\omega_s/\omega_i = 4$).

there is never any gain, only loss; (2) the minimum loss is zero db, which occurs when $x = 0$ and $y = 1$; and (3) both the inverting and noninverting down-converters are identical and independent of p when $x = 0$ (no variable capacitance).

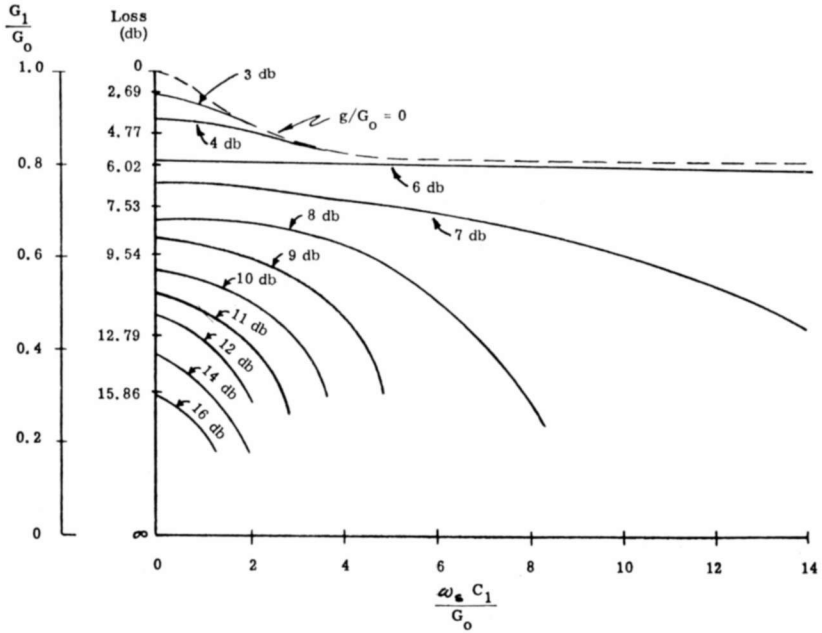


Fig. 13—Loss contours for noninverting down-converter ($\omega_s/\omega_i = 4$) (Figure 12 with abscissa extended).

For $p = 1/160$, the loss contours of the inverting and noninverting down-converters are essentially the same for $x < 3$, with minor differences to $x = 5$. Adding large values of C_1 will require excessive conversion loss in order to maintain a conjugate match. However, a typical possible application exists for $x \approx 1$. For example, there will be a 5-db conversion loss for $y = 0.85$, $x = 0$ and for $y = 0.80$, $x = 1.15$. Therefore, a small amount of nonlinear capacitance will not affect conversion loss.

The change in noise figure is also small because, when considering thermal noise effects at room temperature, Equation (70) indicates that the noise figure changes from 9.7 db at $y = 0.85$, $x = 0$, to 10.6 db at $y = 0.80$, $x = 1.15$.

For $p = 1/4$, the loss contours of the inverting and noninverting

down-converters diverge when variable capacitance is added. The inverting down-converter limits operation up to $x = 2$ (which also requires $y = 0$). The loss at this point is 6 db, which is equal to the down-conversion ratio, and is in agreement with the Manley and Rowe relationship.[16] The noninverting down-converters show very little change in loss in the region $0 < x < 2$ as variable capacitance is increased (Figure 12), but when $x > 2$, the conversion loss approaches 6 db for all $y < 0.8$.

Resonance Conditions

The input susceptances, b_s' and b_i', of the down-converter are

$$b_s' = \pm\omega_s C_0 - \frac{g_{i0}C_1G_1(\omega_i \pm \omega_s) - b_{i0}(G_1^2 \mp \omega_i\omega_s C_1^2)}{g_{i0}^2 + b_{i0}^2}, \tag{29}$$

$$b_i' = \omega_i C_0 - \frac{g_{s0}C_1G_1(\omega_i \pm \omega_s) \mp b_{s0}(G_1^2 \mp \omega_i\omega_s C_1^2)}{g_{s0}^2 + b_{s0}^2}. \tag{34}$$

The specification that resonance exists at both the signal and i-f ports (zero susceptance) is equivalent to

$$b_s' = -(\pm b_s), \tag{75}$$

and

$$b_i' = -b_i. \tag{76}$$

Substituting Equations (75) and (76) into Equations (29) and (34) yields

$$\pm b_{s0} = \frac{g_{i0}C_1G_1(\omega_i \pm \omega_s) - b_{i0}(G_1^2 \mp \omega_i\omega_s C_1^2)}{g_{i0}^2 + b_{i0}^2}, \tag{77}$$

$$b_{i0} = \frac{g_{s0}C_1G_1(\omega_i \pm \omega_s) \mp b_{s0}(G_1^2 \mp \omega_i\omega_s C_1^2)}{g_{s0}^2 + b_{s0}^2}. \tag{78}$$

Let

$$Q_{s0} = \frac{b_{s0}}{g_{s0}} = \frac{b_s + \omega_s C_0}{g_s + G_0}, \tag{79}$$

[16] J. M. Manley and H. E. Rowe, "Some General Properties of Nonlinear Elements, Part I, General Energy Relations," *Proc. I.R.E.*, Vol. 44, p. 904, July 1956.

$$Q_{i0} = \frac{b_{i0}}{g_{i0}} = \frac{b_i + \omega_i C_0}{g_i + G_0}, \tag{80}$$

and

$$U = C_1 G_1 (\omega_i \pm \omega_s), \tag{81}$$

$$V = G_1^2 \mp \omega_i \omega_s C_1^2, \tag{82}$$

$$W = g_{i0} g_{s0} = (g_i + g_0)(g_s + g_0), \tag{83}$$

and substitute these definitions into Equations (77) and (78). This results in

$$\pm Q_{s0} = \frac{U - V Q_{i0}}{W(1 + Q_{i0}^2)} \tag{84}$$

$$Q_{i0} = \frac{U - V(\pm Q_{s0})}{W(1 + Q_{s0}^2)}. \tag{85}$$

The solution of these equations is given by

$$Q_{i0} = \pm Q_{s0}. \tag{86}$$

Using this condition gives a cubic equation of the form

$$Q^3 + KQ + L = 0, \tag{87}$$

where

$$K = 1 + \frac{V}{W}, \tag{88}$$

$$L = -\frac{U}{W}, \tag{89}$$

$$Q = Q_{i0} = \pm Q_{s0}. \tag{90}$$

The solutions of this equation are[17]

$$Q_1 = D + E, \tag{91a}$$

[17] I. S. Sokolnikoff and E. S. Sokolnikoff, *Higher Mathematics for Engineers and Physicists*, 1st ed., p. 20, McGraw-Hill Book Co., New York, 1934.

$$Q_2 = \omega D + \omega^2 E, \tag{91b}$$

$$Q_3 = \omega^2 D + \omega E, \tag{91c}$$

where

$$\omega = -\frac{1}{2} + j\frac{\sqrt{3}}{2}, \tag{92a}$$

$$\omega^2 = -\frac{1}{2} - j\frac{\sqrt{3}}{2}, \tag{92b}$$

$$j = \sqrt{-1}, \tag{92c}$$

$$D = \left[-\frac{L}{2} + \left(\frac{K^3}{27} + \frac{L^2}{4} \right)^{1/2} \right]^{1/3} \tag{92d}$$

$$E = \left[-\frac{L}{2} - \left(\frac{K^3}{27} + \frac{L^2}{4} \right)^{1/2} \right]^{1/3}. \tag{92e}$$

From these solutions it is possible to define a discriminant, Δ, as

$$\Delta \equiv -4K^3 - 27L^2. \tag{93}$$

The value of Δ reveals the following properties about Q:

(1) $\Delta < 0$
Q_1 is real,
Q_2, Q_3 are complex.

(2) $\Delta = 0$
Q_1 is real,
Q_2, Q_3 are real and equal.

(3) $\Delta > 0$,
Q_1, Q_2, Q_3 are real and unequal.

If the resonance condition of Equation (90) is substituted into the general gain expression, Equation (37), then the gain for resonance is

$$P_{GCR} = \frac{4g_i g_s (G_1^2 + \omega_i^2 C_1^2)}{[G_1^2 \mp \omega_i \omega_s C_1^2 - g_{s0} g_{i0}(1 + Q^2)]^2 + [C_1 G_1 (\omega_i \pm \omega_s) - 2Q g_{s0} g_{i0}]^2}. \tag{96}$$

Loss Contours for Resonant Down-Converter

Once again the equations are normalized with respect to G_0 and Equations (71) are substituted into Equations (89), (90) and (94), giving

$$K = 1 + \frac{y^2 \mp px^2}{\left(1 + \frac{g_i}{G_0}\right)\left(1 + \frac{g_s}{G_0}\right)}, \tag{95a}$$

$$L = -\frac{xy(p \pm 1)}{\left(1 + \frac{g_i}{G_0}\right)\left(1 + \frac{g_s}{G_0}\right)}, \tag{95b}$$

$$P_{GCR} = \frac{4\frac{g_i g_s}{G_0 G_0}(y^2 + p^2x^2)}{\left[y^2 \mp px^2 - \left(1 + \frac{g_i}{G_0}\right)\left(1 + \frac{g_s}{G_0}\right)(1 + Q^2)\right]^2 + \left[x(p \pm 1) - 2Q\left(1 + \frac{g_i}{G_0}\right)\left(1 + \frac{g_s}{G_0}\right)\right]^2}. \tag{94}$$

As an illustrative example of the converter properties, the case of $p = 1/160$ will be treated. Also, the normalized signal and i-f conductances will be made equal to 0.5, i.e.,

$$\frac{g_i}{G_0} = \frac{g_s}{G_0} = \frac{1}{2}.$$

Figures 14 and 15 are the resulting Q contours, and Figures 16 and 17 are the loss contours for the inverting and noninverting down-converters, respectively. In both the inverting and noninverting down-converters, the loss curves slope upward, and any addition of variable capacitance will increase the conversion loss. This upward slope of the loss characteristics occurs because the degree of mismatch varies with both x and y. Even with this effect taken into account, the loss at any point is always greater than the value obtained for the conjugate-match case. Therefore, no benefit is obtained from resonance as is possible with conjugate-match conditions, and it will not be considered further.

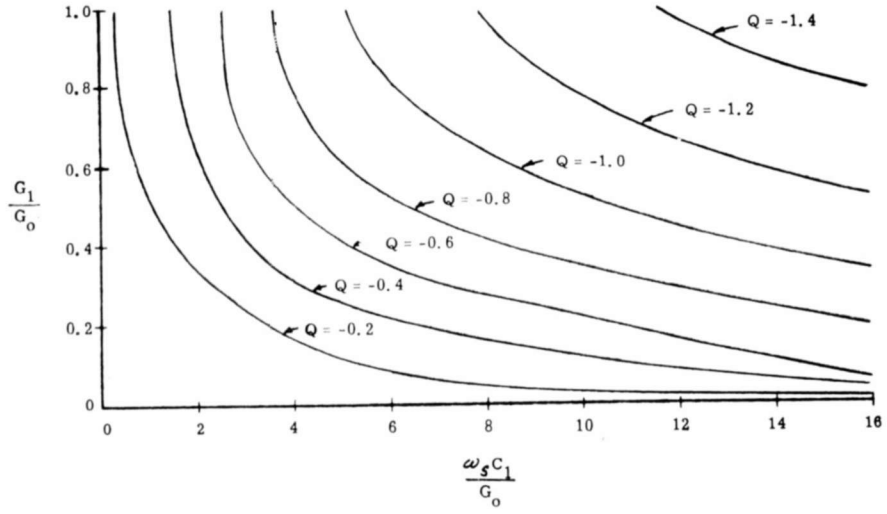


Fig. 14—Q contours for inverting down-converter ($\omega_s/\omega_i = 160$).

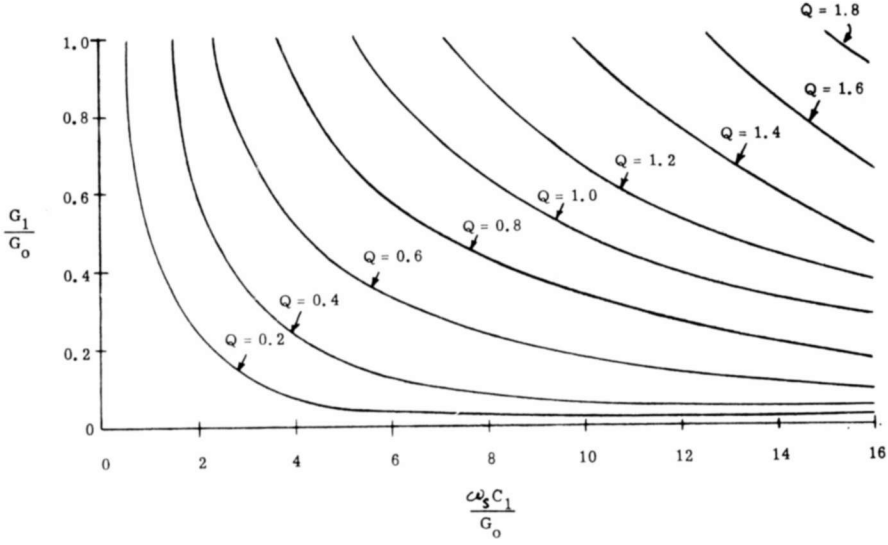


Fig. 15—Q contours for noninverting down-converter ($\omega_s/\omega_i = 160$).

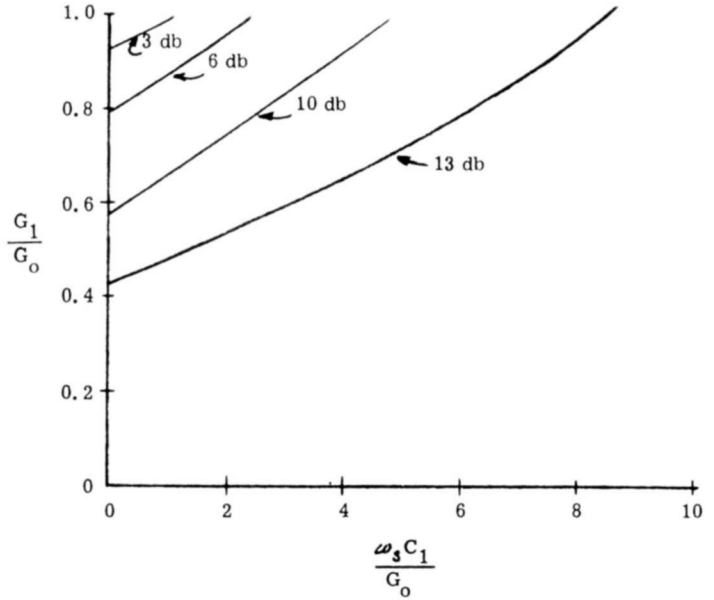


Fig. 16—Resonance loss contours for inverting down-converter ($\omega_s/\omega_i = 160$).

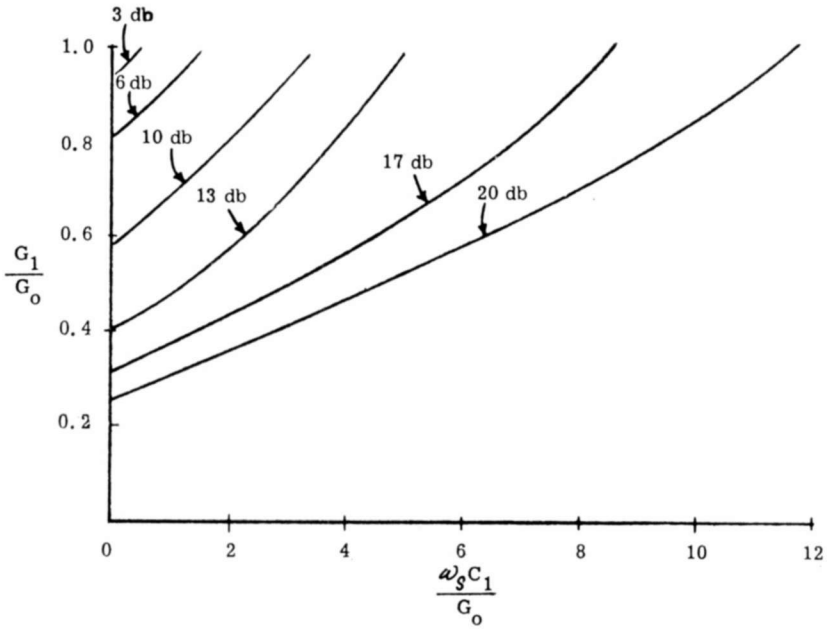


Fig. 17—Resonance loss contours for noninverting down-converter ($\omega_s/\omega_i = 160$).

Conclusions

A nonlinear admittance consisting of a parallel combination of nonlinear conductance and nonlinear capacitance pumped in time phase can be used as a mixer. The addition of nonlinear capacitance to a specified nonlinear conductance does not, in general, degrade down-converter performance. However, if the nonlinear conductance is very small, the addition of a large nonlinear capacitance will result in excessive losses when the frequency ratio of the converter is large. For a small down-conversion ratio, the maximum allowable nonlinear capacitance is limited to a narrow range in the inverting converter, while the noninverting converter can almost always be operated with unlimited nonlinear capacitance without appreciably degrading the loss characteristic.

A junction-diode construction similar to the varactor type can result in a nonlinear admittance with the required parallel combination of nonlinear conductance and nonlinear capacitance. However, care must be taken to restrict the amount of variable capacitance to that dictated by the down-conversion ratio. The down-converter performance compared to a point-contact diode will be the same with regard to gain and noise figure, but lower intermodulation distortion will result. This will be due primarily to the increased power-handling capacity of the junction diode over the point-contact diode.

THERMOPLASTIC ORGANIC PHOTOCONDUCTIVE RECORDING MEDIA — ELECTROPHOTOGRAPHIC CHARACTERISTICS AND PROCESSING TECHNIQUES*

BY

E. C. GIAIMO

RCA Laboratories
Princeton, N. J.

Summary—*Single-layer thermoplastic organic photoconductive material can be processed automatically in air and provide high-resolution images that can be viewed using schlieren-projection readout. An experimental processor providing independent control over processing parameters has been developed. The processor has been used for duplicating photographic transparencies containing half-tone screening. Since the dark decay of charge increases during heat development, the latent charge image produced during exposure is degraded if heating is prolonged during development. Various factors including past light history, past heating history, dark and light charge-decay rates, spectral response in terms of charge decay, and temperature effects have been investigated in order to determine optimum processing conditions. Typical characteristic curves are presented.*

INTRODUCTION

PREVIOUSLY published papers[1,2] describe photoconductive thermoplastic materials that are capable of producing images through surface deformation. The entire processing procedure with these materials is carried out in air.

There is close similarity between the method used for processing the thermoplastic photoconductive layers (TPCL) and that employed in electrophotography; the layer is charged in the dark, exposed to a light image, and then developed. The difference between the two systems is in the development. In electrophotography, the latent charge image is toned with an electroscopic powder to produce a visible image. In the case of TPCL, heat is applied so that the electrostatic forces

* The research reported herein was sponsored by the Aeronautical Systems Division, Air Force Systems Command, Wright-Patterson Air Force Base, Dayton, Ohio under Contract No. AF33(657)-7920 and by RCA Laboratories, Princeton, New Jersey.

[1] H. G. Greig, "An Organic Photoconductive System," *RCA Review*, Vol. XXIII, p. 413, Sept. 1962.

[2] N. E. Wolff, "A Photoconductive Thermoplastic Recording System," *RCA Review*, Vol. XXV, p. 200, June 1964.

existing in the charged areas can deform the surface. The deformed surface can then be viewed with obliquely incident light or projected through a schlieren optical system. If desired, the latent charge image in the TPCL system can be made visible by toning before heating; this is sometimes done as a diagnostic tool to derive information about latent image formation. In the present paper, however, only the heat-development process is considered.

Processing a TPCL consists of charging the layer with either a positive or negative corona discharge in air, exposing it to a light image, which discharges the layer in the illuminated areas, and finally heating it to soften the layer and allow the electrostatic forces acting in the areas containing charge to deform the surface.

Factors Governing Operation of TPCL

In the mode of operation under consideration the generation of a deformation image is governed by an energy balance. The energy that produces the latent image is initially stored in the layer as electrostatic energy. This energy is gradually reduced by charge leakage in the dark; this leakage increases as the temperature of the layer is raised. It is the difference between the energy stored and the energy lost by charge leakage that is available during deformation. Part of this energy is dissipated by internal friction losses in the layer during deformation, and the rest appears as potential energy of the deformed surface.

For the stored electrostatic energy to appear as potential energy of the deformed surface, the rate of change of surface tension and viscosity during the heating and cooling interval must be such that the surface will move rapidly under the action of a decreasing electrostatic force and yet remain in the deformed state upon cooling and removal of the electrostatic force. Ideally, the surface tension and viscosity should have a sharp transition point as the temperature is raised. At this transition temperature the charge loss should be low. At present, these properties can only be determined experimentally, because the published rheological data on properties of the materials used in combination in these layers is not sufficient to predict their behavior.

The physical properties of the layer are determined entirely by the composition. Therefore, the three most easily controlled process parameters are layer thickness, total heat input to the layer during development, and heating rate.

In most cases it was found that layers between 10 and 20 microns thick gave most consistent deformation images with the coating compositions tried. The surface potentials of a typical layer, determined

by electrometer[3] measurements, are shown in Figure 1, where surface potential is plotted versus layer thickness. Layers in excess of 20 microns were not investigated, since the quality of the surface of such layers was poor with present coating techniques.

It has been mentioned that the dark decay of charge is one of the critical factors governing image formation. By measuring the charge decay in the dark as a function of temperature, some idea of the tem-

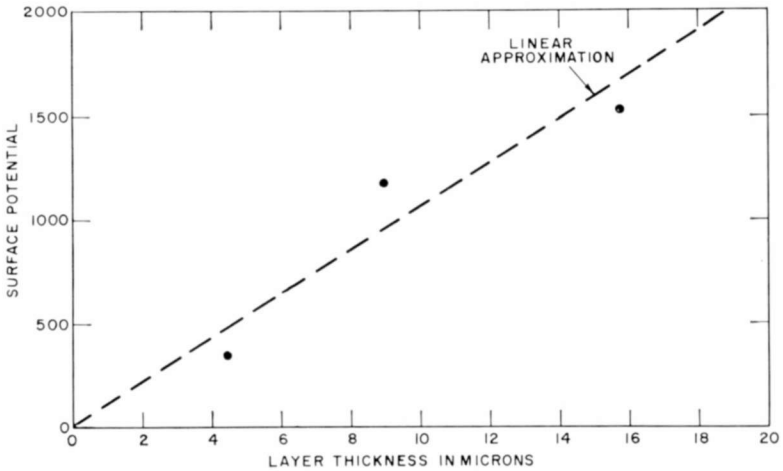


Fig. 1—Surface potential versus thickness of coating.

perature below which the image must form can be obtained. As yet we have not developed satisfactory methods for investigating the interrelationships between image formation, viscosity, and surface tension versus time and temperature on thin layers of the materials being used.

Figures 2 and 3 show dark-decay curves of two TPCL layers having different compositions by weight. The curves were obtained using an electrometer apparatus especially designed to allow discharge curves to be taken on heated samples. The apparatus is described in Appendix I. The composition of the layer 11769-42G shown in Figure 2 is 33.7% leucobase of malachite green, 63.1% polystyrene, and 3.2% chlorinated paraffin. The composition of the layer 16596-5B shown in Figure 3 is 45.5% leucobase of malachite green, 45.4% Piccoflex —115,* and 9.1% Piccotex —120.* The latter composition not only shows lower

[3] E. C. Giaimo, "A Dynamic-Capacitor Electrometer Suitable for Measuring Electrophotographic Recording Media," *RCA Review*, Vol. XXII, p. 780, Dec. 1961.

* Pennsylvania Industrial Chemical Corporation, Clairton, Pennsylvania.

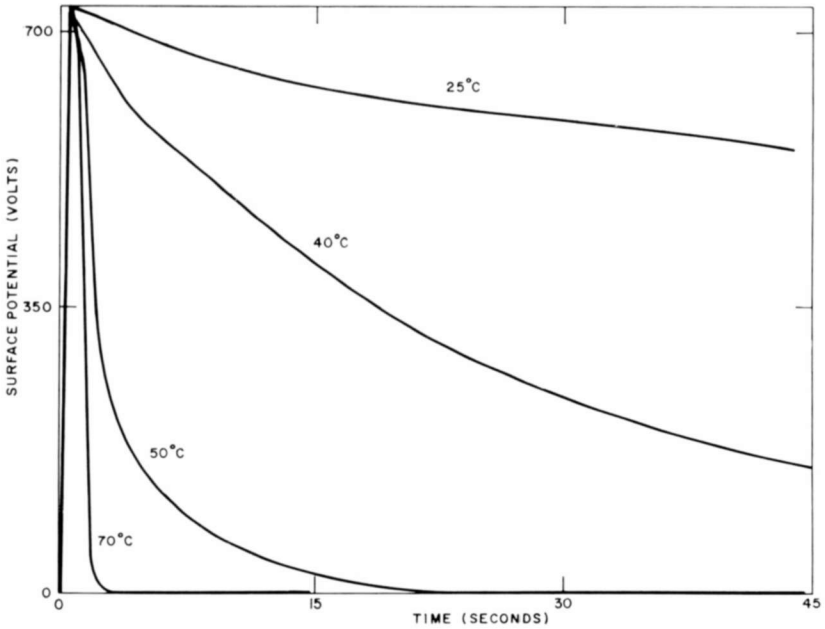


Fig. 2—Thermal dark decay of charge for coating 11769-42G.

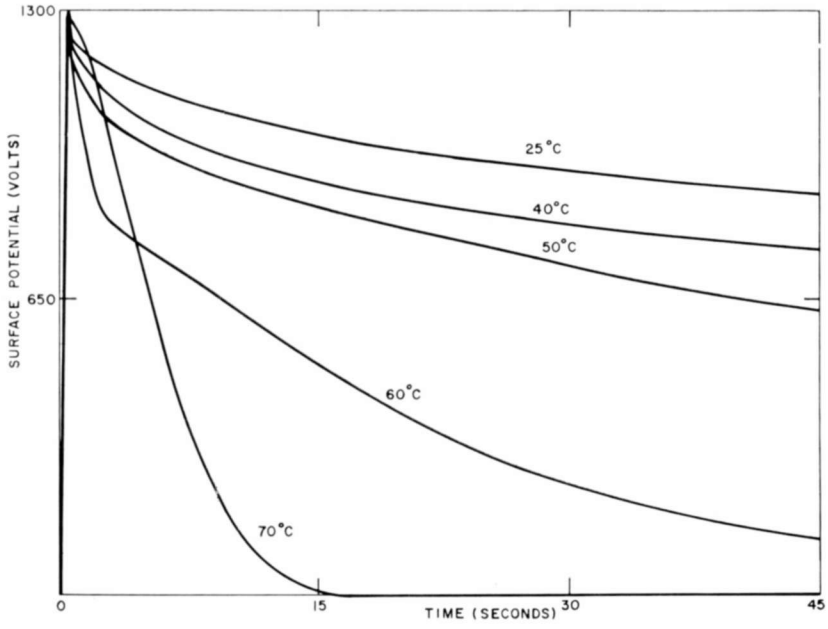


Fig. 3—Thermal dark decay of charge for coating 16596-5B.

dark decay at elevated temperatures but is a physically harder material at room temperature.

Further study of the dark decay and light decay of charge on one of the layers investigated is shown in Figure 4. It was observed that the layer could not be made to discharge to zero but appeared to approach the dark discharge rate after about 30 seconds illumination. Even after several minutes of exposure the surface potential was of

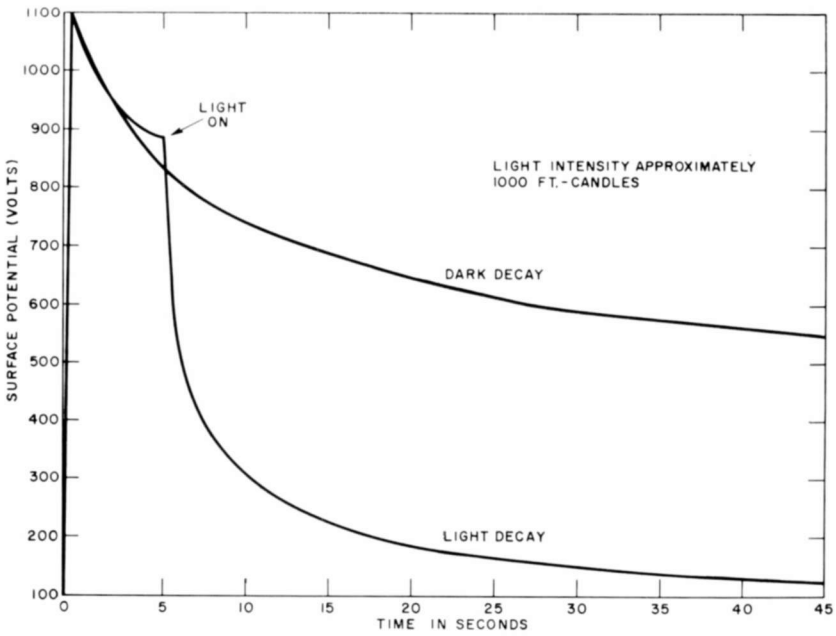


Fig. 4—Light decay of charge for coating 16596-18F.

the order of 100 volts. This "rest charge" could be rapidly dissipated by heating. Simultaneous heating and exposure resulted in no increase in the light-decay rate, indicating that the light and temperature charge decay phenomena are independent processes.

An equivalent circuit describing the discharge characteristics of the layer is shown in Figure 5, where C_1 and C_2 represent the charge storage, R_L is related to the light decay, and R_t is related to the dark decay and is temperature dependent. C_2 accounts for the "rest charge," which is insensitive to light exposure but can be discharged by R_t.

Past Light and Heating History Effects

Knowledge of light and heating history effects are required to deter-

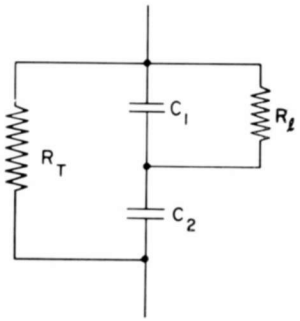


Fig. 5—Equivalent circuit of thermoplastic photoconductive layer.

mine what time cycle must be employed if the TPCL is to be re-used repeatedly. Some preliminary data was taken to determine the effects that can be anticipated with one of the coating formulations used.

Past Heating History Effects

The data shown in Figure 6 was taken to determine whether the dark decay of charge at room temperature is modified by previous heating of the layer. The sample under test was charged with a negative corona, and the decay rate shown in curve No. 1 was recorded at room temperature (25°C). When the sample was recharged and heated to 50°C at a heating rate of 25°C/sec, curve No. 2 resulted. Repetition

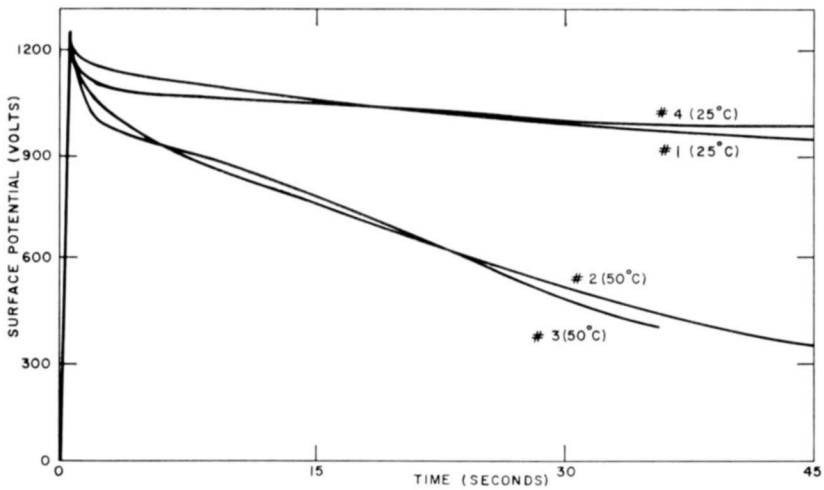


Fig. 6—Past heating history effects on coating 16596-3B.

of the process yielded curve No. 3. Curve No. 4 was obtained after the sample was cooled to 25°C in approximately 10 seconds. It appears that the charge-decay returns immediately to the value it had prior to heating. No permanent or transient change can be observed within experimental error. This indicates that after heat development and cooling to room temperature, one can determine the charge decay rate and the surface potential from a family of thermal decay curves of the type shown in Figures 2 and 3. The data further indicates that the layer can be re-used immediately insofar as the heating cycle is concerned.

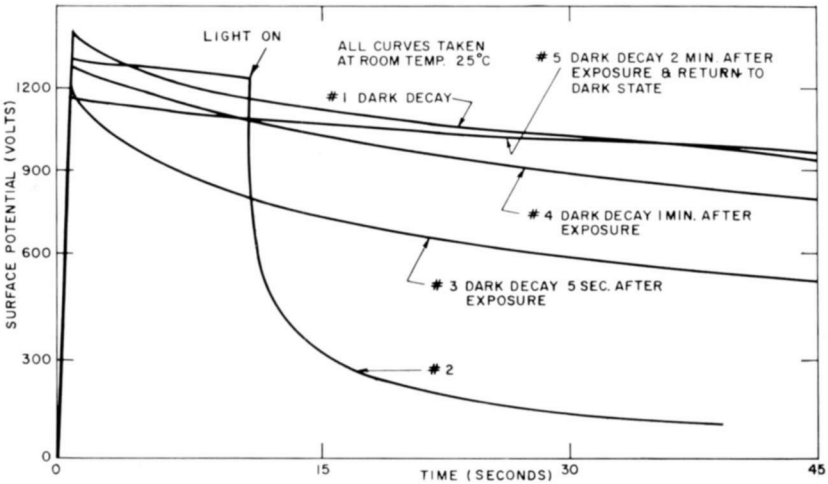


Fig. 7—Past light history effects on coating 16596-3B.

Past Light History Effects

The data shown in Figure 7 was taken to determine the influence of prior light exposures. The data was taken at room temperature. Curve No. 1 is the dark decay before any exposure. Curve No. 2 is the light decay under approximately 1000 ft-candles. Five seconds after exposure ceased, the sample was recharged and the Curve No. 3 was taken. One minute later Curve No. 4 was taken, and Curve No. 5 followed two minutes later. The data shows that this sample, when exposed to approximately 1000 ft-candles for 45 seconds, takes two minutes to return to the same dark decay rate it had prior to exposure. This does not appear to be a heating effect.

The experimental data indicates that the re-use cycle for this sample could be no shorter than two minutes if the same initial conditions are to be maintained for image reproducibility. In practice, the re-use time

probably could be shortened, since Curve No. 4 appears to be quite similar to Curves No. 1 and No. 5.

Spectral Response of Corona-Charged Layers

The processing of the thermoplastic organic photoconductive layers involves corona charging prior to exposure. The spectral response curve shown in Figure 8 was obtained to determine the amount of charge

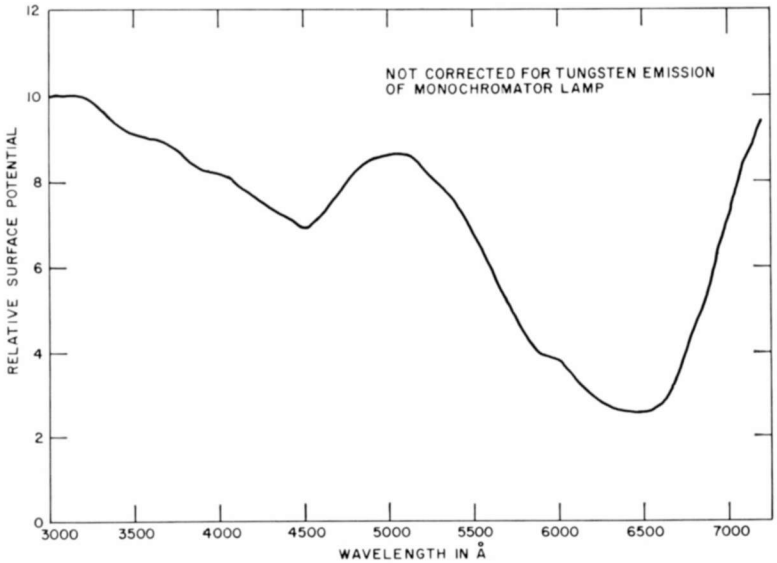


Fig. 8—Spectral curve of transparent PC 11769-59(P69).

lost on corona-charged layers irradiated with light of various wavelengths. Irradiation with light from a tungsten source shows maximum charge loss at approximately 6400 Å with a secondary peak at 4500 Å. Since past experience with these photoconductive layers indicated that past-light-history effects could modify the shape of the spectral curve, care was taken in the measuring technique to eliminate these effects.

The apparatus used for measuring the spectral curve using corona charging is described in Appendix II. This measuring technique is included because it closely simulates the conditions for processing of the TPCL. It is applicable to other electrophotographic photoconductive media that use corona charging during processing.

Processing

Processing of the plate to form a ripple image consists of three

steps. First, the plate is given an electrostatic charge in the dark by exposure to a corona discharge in air. Either a positive or negative charge can be imparted to the surface. Second, the plate is exposed to an optical image to discharge it preferentially in the illuminated areas. Third, it is briefly heated to soften the plastic and allow the electrostatic forces produced by the remaining charge to deform the layer.

Fig. 9—Experimental slide processor.

This form of recording, wherein the photoconductive layer is used directly, imposes critical requirements on the layer in terms of its charge-storage capabilities, thermoplasticity, and photoconductivity. These several parameters must be carefully balanced and the processing conditions carefully controlled if an operative system is to be obtained. The system under consideration is both time and temperature dependent.

To process the TPCL experimentally with some degree of reproducibility, it was necessary to design an experimental processor in which the critical processing parameters could be controlled. An experimental laboratory processor is shown in Figure 9. It was designed to process glass slides $2 \times 2 \times 0.040$ inch coated on one side with a transparent tin oxide conductive layer upon which the thermoplastic photoconductive layer was deposited.

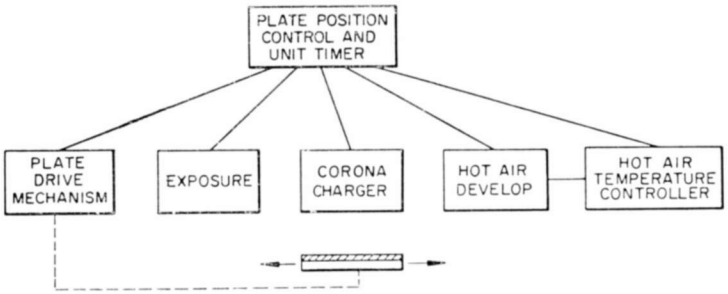


Fig. 10—Block diagram of thermoplastic photoconductive plate processor.

The functions performed automatically by the processor are shown in Figure 10. The photosensitive plate is charged by moving it through a corona discharge. It is then placed in an exposure position where it is held stationary while being exposed for a preset time. The exposure is accomplished with a photographic enlarger fitted with a solenoid-actuated shutter energized through a timer. The enlarger is fitted with suitable optics to allow for either image reduction or 1:1 projection. After exposure, the slide is transported into the development zone where thermostatically controlled filtered forced hot air is directed against the thermoplastic surface. The heating interval is controlled by an air valve located in a hot-air duct that directs the continuously heated air against the plate. The air valve is solenoid actuated and energized by a timer. The sensor for the air temperature control consists of a 0.001-inch-diameter platinum wire resistance element suspended in the hot-air duct. The timing intervals used are shown in Figure 11.

Several heating methods have been tried, including resistance heating of the conductive tin oxide layer and conduction heating by pressing the slide against a heated platen. The filtered hot air method was

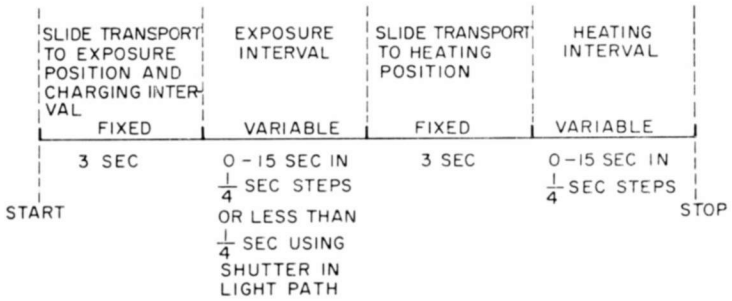


Fig. 11—Experimental-slide-processor timing intervals.

chosen because it provides a means of removing dust that is attracted to and settles on the charged surface. Dust removal prior to softening of the surface reduces the amount of dust imbedded in the surface of the finished print.

Some Photographic Properties of TPCL

Read-Out

Schlieren optics are presently being used to display the ripple images

Fig. 12—Original and projected thermoplastic image. On the left is an enlargement from the photographic screened negative master and on the right is a rephotograph of the projected thermoplastic image (2 × 3 feet) made from that master (see Reference (2), p. 207).

formed on the TPCL. A typical image that can be obtained is shown in Figure 12. It was projected through a schlieren projector at approximately 35 times magnification from a 1-inch square ripple image and then photographed. The original 16 lines/mm screened, negative, master transparency from which the thermoplastic image was made is also shown in Figure 12. The thermoplastic slide was obtained from the processor shown in Figure 9.

The schlieren optical system[2] used for read-out is capable of resolving between 24 and 118 lines/mm and has a maximum contrast ratio of 100:1.

Gray Scale

Variations in charge density in the latent image must correspond to variations in light refraction or scattering centers in the deformed surface of the TPCL to obtain gray-scale rendition of the tonal quality of the input. Any surface of constant refractive index that is perpendicular to the collimated light passing the first schlieren mask in the schlieren projector appears as black in the output. Areas of constant or uniform charge density must be converted by means of screening techniques into a set of surfaces inclined at angles between 0° and 90° to the entrant light rays in order to give scattering that corresponds to differences in the image tonal quality.

The screening techniques employed have included charging through masks to obtain a line screen or dot charge pattern, exposing the TPCL to a screened image, or charging and then exposing to a screen pattern to generate a charged dot pattern prior to exposure to the image to be reproduced. The best results have been obtained with directly screened transparencies.

Sensitivity

The spectral sensitivity of the layers investigated is shown in Figure 8. Variation in the spectral sensitivity can be had using sensitizing dyes.[2]

Since the layers exhibit no gray scale in the form of density variation without screening and schlieren projection techniques, a meaningful ASA rating cannot be assigned to the material alone. Any ASA rating must refer to the material, the particular screening used, and the schlieren projection system. Using the schlieren projector referred to above and a 16 lines/mm screened transparency, exposures were made on both the TPCL and a silver-halide film having an ASA rating of 1.6. The silver halide film was developed under recommended standard conditions to realize ASA 1.6. When it was felt that comparable projected images were obtained on the TPCL and the silver halide film, the ASA rating of the TPCL was computed on the basis of the exposure required for each. This gave an ASA rating of 1.4×10^{-5}.

This seemingly low sensitivity figure must be interpreted in terms of the rather special test that was used to derive it. The material has adequate sensitivity for use with readily available tungsten sources. For example, exposures between 2 and 10 seconds are sufficient if an enlarger fitted with a No. 2 photoflood lamp is used with a lens opening of f:4.5, and a 16 lines/mm screened 35 mm transparency is projected onto 1 square inch of the surface of the TPCL. The light level measured at the exposure plane under the above conditions was of the order of

47 foot-candles. This figure was obtained using a Weston Model 756 illumination meter equipped with a Viscor filter.

Resolution

Resolution tests were made using standard ASA and NBS patterns under various conditions. Here again, the ultimate resolution capability is limited by the overall system, particularly the schlieren projectors. When the plates are used for thermoplastic images, a limit is approached when the lines are so close together there no longer is any appreciable indentation between them, and the line structure disappears when viewed with a schlieren projector. The projector system itself is limited, since the better the resolution, the less light is available on the screen. In our projector the resolution capability of the projector and the plates coincide quite well, both being capable of 56 lines/mm with either type of pattern (NBS or ASA). A typical slide exhibiting more than 56 lines/mm is shown in Figure 13.

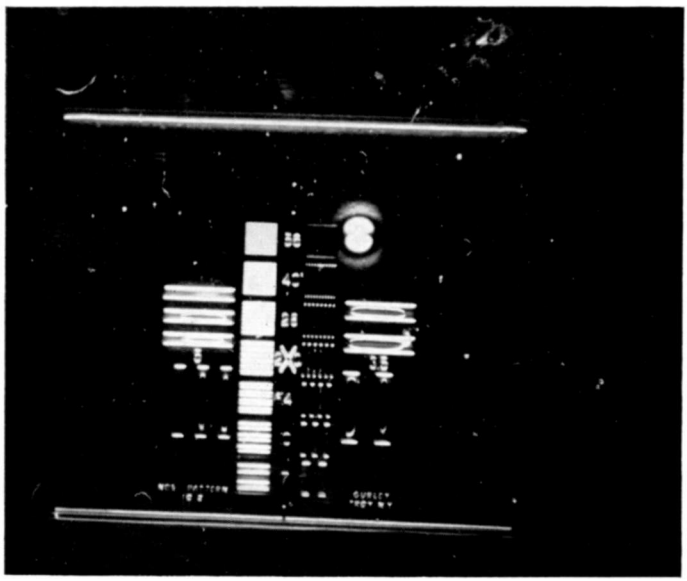


Fig. 13—Thermoplastic reproduction of NBS resolution chart showing that 56 lines/mm are reproduced.

Conclusions

Thermoplastic photoconductive organic layers have been successfully

processed automatically in air to produce deformation type images that can be viewed using schlieren projection techniques. Experimental investigation of dark decay of charge using a specially designed electrometer (Appendix I) reveals that charge leakage through the layer is an important factor during development. The rate at which the layer softens and the charge leakage rate must be such that sufficient electrostatic force is maintained to cause deformation of the surface.

For the materials investigated, it was found that layers thicker than 10 microns are required to obtain the proper energy balance between that initially stored in the electrostatic field and that needed to deform the surface during the development interval. Filtered, thermostatically controlled, forced hot air appears to be a practical means of heating the layer during development. It has yielded uniform development over 2 × 2-inch plates and helps to reduce the amount of dust embedded in the surface, thus improving the quality of the deformation image.

Photographic measurements indicate that resolution of at least 56 lines/mm can be realized. The materials tested exhibit no gray scale, so that half-tone screening is needed to obtain tonal rendition. Exposure of the thermoplastic layer and silver halide film to produce comparable projected images, the former through a schlieren projector and the latter through an ordinary slide projector, yields an ASA rating of 1.4×10^{-5} for the thermoplastic material. The sensitivity is adequate for duplicating photographic transparencies using conventional tungsten projection lamps and exposures between 2 and 10 seconds.

Further, it was found that heat history is not significant in preparation of repeat images; light history also is not significant in preparation of repeat images if intervals of 2 minutes or more are allowed between each use period.

Acknowledgments

The author wishes to thank H. G. Greig for providing the materials used in this work and for his stimulating discussions of the research problems. Thanks are due to F. H. Nicoll for the schlieren projector design and the advice he gave on all aspects of the project based on his early experimental results with this type of recording system, to S. W. Johnson for his photographic measurements, and also to W. Mehl and N. E. Wolff for contributions based on their fundamental investigation of the photoconductive organic system. The author is especially grateful to C. J. Young for providing the environment, encouragement, and support during the course of the program.

Appendix I—Electrometer used for Measuring Thermal Decay Characteristics

The thermal charge decay rates of the various layers under investigation were measured on the apparatus shown in Figure 14. A schematic diagram is shown in Figure 15. The sample under test is charged by a corona discharge and then moved rapidly under the vibrating pickup probe. The probe vibrates at 120 cps, converting the slowly

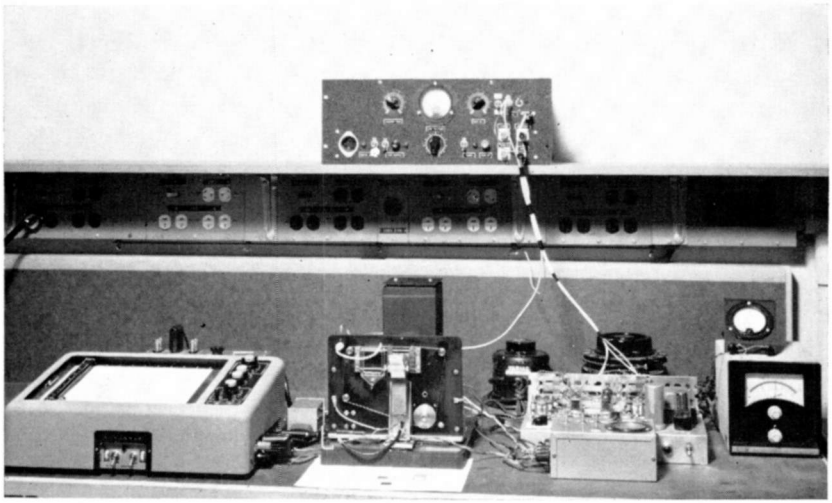


Fig. 14—Apparatus for measuring thermal charge decay.

decaying potentials across the thermoplastic layer to an a-c signal. This signal is amplified with a high-input-impedance amplifier, rectified, and applied to a chart recorder for display.

The sample configuration chosen for testing consists of a $0.75 \times 0.75 \times 0.0035$-inch aluminum plate, coated on one side with Aquadag and on the other with the thermoplastic layer. The thin aluminum substrate is used because it has as small a thermal mass as possible yet is rigid enough to coat and mount easily. The Aquadag increases its absorption of radiant heat.

The test device heats the sample by radiant heating. The image of a 1000-watt (110-volt) projector lamp is focused on the Aquadag-coated side of the test sample. Interposed between the sample and the lamp is a vane-type shutter actuated by a solenoid. The solenoid is energized at the proper time intervals to start and stop the input heating to the

sample. A given temperature excursion of the sample is obtained by timing the opening and closing of the shutter and the voltage applied to the projector lamp. Timing is provided for these two functions by the master program timer. The relative timing of the various functions is shown in Figure 16.

By varying the initial and steady-state voltage applied to the lamp,

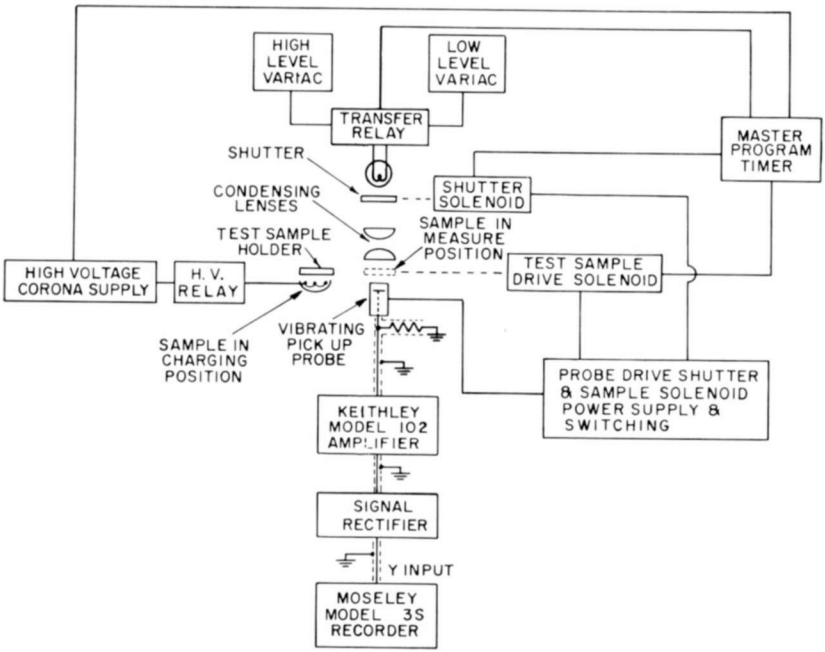


Fig. 15—Diagram of electrometer apparatus for thermal-charge-decay measurements.

the heating ratio and the steady-state temperature can be varied. It was assumed that the thermoplastic coating follows, for all practical purposes, the thermal excursion of the aluminum substrate. To reduce the amount of power needed to maintain a given sample temperature, the sample is sandwiched between the two low-thermal-conductivity mica masks. Electrical connections of the sample to ground during charging is provided by a foil strip.

The heating lamp voltage is programmed so that it is initially at a high value and is lowered during irradiation of the Aquadag-coated side of the aluminum substrate to provide a step-heating function. The lamp voltages and switching times were determined experimentally. The table of values as well as the temperature rise versus time curves are shown in Figure 17. These curves were obtained using an aluminum

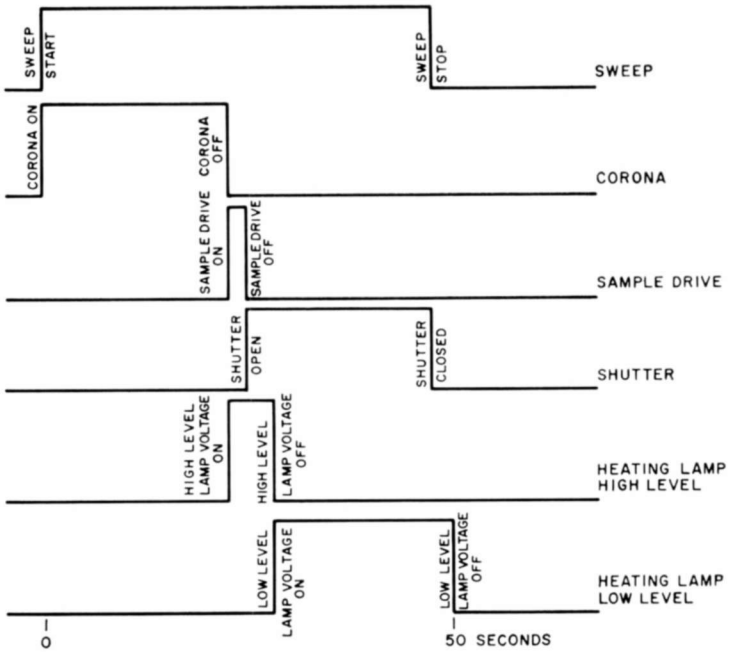


Fig. 16—Typical electrometer timing functions for thermal-charge-decay measurements.

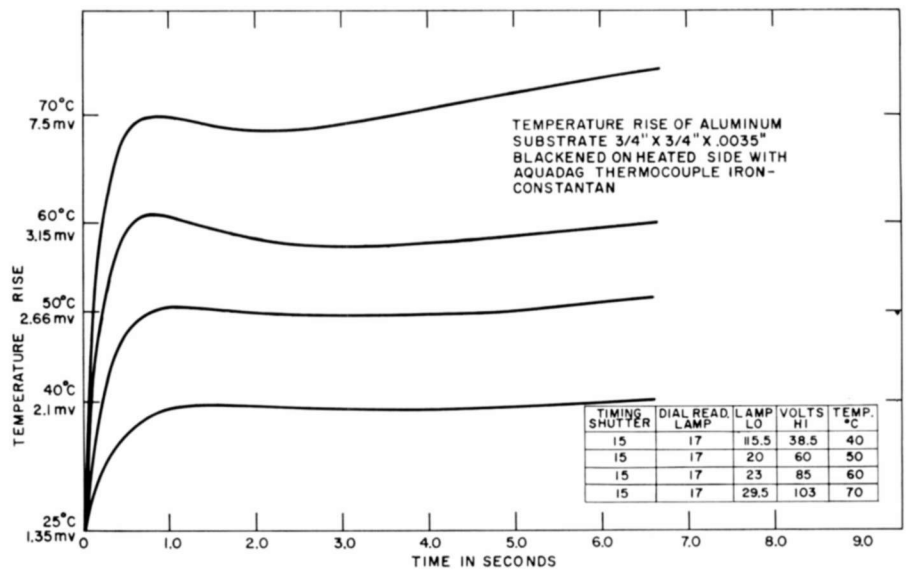


Fig. 17—Temperature rise of sample substrate versus time.

sample having dimensions identical with the test samples but with a thermocouple welded to the side normally carrying the thermoplastic layer.

Voltage calibration of the unit for various probe drive voltages, amplifier and recorder gains, and optimum probe spacing were also made. This provides a means of measuring surface potential or charge density. A metallic surface positioned in the sample location was used to determine the voltage calibration.

Appendix II—Spectral Curve Plotter using Corona Charging

A schematic diagram of the experimental apparatus for measuring spectral curves on charged layers is shown in Figure 18. Past light

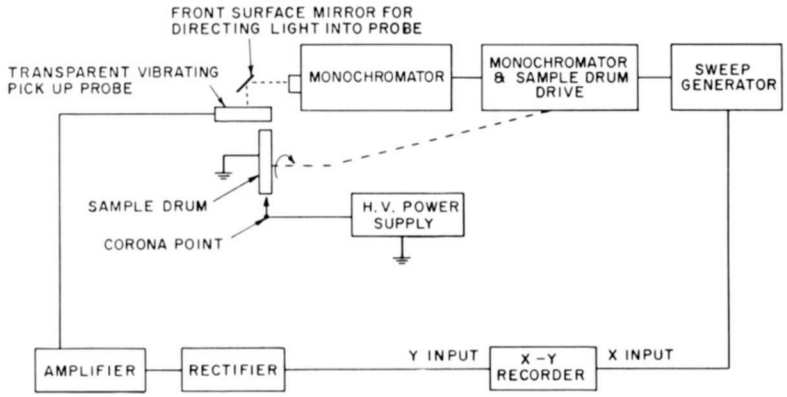


Fig. 18—Diagram of spectral curve plotter using corona-charged layers.

history, which causes temporary and in some cases permanent change in the dark decay of charge, could materially modify the shape of the spectral curves. To eliminate past-light-history effects, the sample strips that were measured were kept in the dark for at least 24 hours. In addition, a particular area of the strip carrying the photoconductor was exposed to a given wavelength only once. The drum carrying the sample moves in synchronism with the monochromator drive.

The photosensitive layer was immersion coated on a 0.0035-inch-thick strip of aluminum 1 inch wide and 20 inches long. The curve is a plot of the charge lost from the sample as it is slowly moved from beneath a point corona discharge and then under a vibrating transparent pickup probe. The light, whose wavelength is varied, passes through the transparent pickup probe. The sample strip is moved in

synchronism with the monochromator drive that changes the wavelength of the output light. The output of the pickup probe is fed to the Y axis of the X-Y recorder. The monochromator drive unit is used to derive an electrical output proportional to wavelength. This signal is applied to the X axis of the recorder. The resulting curve is a plot of surface potential versus wavelength. At the sensitivity peaks where the maximum charge-decay rates occur, minima in the surface potential appear.

Care was taken to obtain uniform coating thicknesses in order to ensure uniform initial surface potentials along the 20-inch-long sample. The uniformity of the layer and the resulting surface potential can be checked by passing the charged sample beneath the probe without admitting the light. This establishes the base line on the plot, from which small corrections can be made in the spectral curve to compensate for minor coating irregularities.

GENERATING NETWORK FUNCTIONS WITH AN INFINITE POTENTIAL ANALOG PLANE

By

F. M. Brock and R. Binks

RCA Broadcast and Communications Products Division
Camden, N. J.

Summary—*The general problem of generating rational-fraction network functions that suitably approximate linear, passive, two-port network transmission characteristics is reviewed. The use of potential theory in creating an analog of the network function is discussed.*

The development, construction, calibration, and use of a double-sided, near-infinite potential-analog plane using dry conducting sheets is described. Details are shown for simulating function roots with point current sources, and for means of detecting potential magnitude and gradient. The plane potential magnitude (representing the function logarithmic magnitude) and gradient (representing the function phase derivative), as well as plane distance, are continuously recorded on a 17-inch X-Y graphical recorder.

The "infinite" potential plane can be used to represent any or all regions of the complex s plane, or any region of a conformally transformed plane; thus, network-transfer or characteristic functions may be generated by appropriately locating point current sources on the plane.

Examples are shown for transfer functions up to the sixteenth order, with a range of over 70 db and pass-band flatness to ±0.75 db. Problems due to conjugate roots and root-shifting can be reduced by suitable mapping transformations.

For double-terminated filter-network synthesis, the characteristic function is often more suitable, and a simple transformation is presented that allows all of the characteristic-function roots to be located on an axis of the potential plane. This simplifies the approximation problem and improves the resolution of the plane in expressing the function magnitude. An example is shown of a twelfth-order band-pass characteristic function of the equal-ripple general-parameter type providing 20-db pass-band return loss and 100-db upper stop-band discrimination.

Introduction

A REQUIRED STARTING point for many network synthesis techniques is the availability of a specified rational-fraction network function for frequency-domain quantities. However, network requirements are seldom initially available in terms of rational functions; they are usually given as graphical plots of desired magnitude, phase, or delay (or of real and imaginary components) of network functions with respect to frequency. Thus, methods and techniques for determining suitable rational-fraction approximations to graphical specified network requirements form an important part of network synthesis.

This paper describes a practical application of the potential-analog theory of network-function approximation to the generation of both network transfer and characteristic functions. This potential-analog system was developed as a design tool during an extensive program of engineering specific filter designs for frequency-division multiplex telephone carrier equipments. Some of the network functions for these filters presented imposing requirements; for example, pass-band attenuation ripples of the order of 0.1 db and stop-band attenuations greater than 70 db.

Network Transfer Functions

For 4-terminal, linear, passive, lumped-element networks, the network transfer function can be expressed as a rational fraction of the form

$$T(s) = \left.\frac{\text{output}}{\text{input}}\right|_{(s)} = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_n s^n}{b_0 + b_1 s + b_2 s^2 + \cdots + b_m s^m} = \frac{\boldsymbol{A}(s)}{\mathcal{B}(s)} \qquad (1)$$

where s = complex frequency $= \pm\sigma \pm j\omega$; $n \leqq m$. $T(s)$ can also be expressed in factor form as

$$T(s) = H\left[\frac{(S - S_{01})(S - S_{02}) \cdots (S - S_{0n})}{(S - S_{p1})(S - S_{p2}) \cdots (S - S_{pm})}\right]. \qquad (2)$$

Thus, the transfer function $T(s)$ can be completely determined from a constant factor (H), the roots of the numerator $\boldsymbol{A}(s)$ (zeros of transfer), and the roots of the denominator $\mathcal{B}(s)$ (poles of transfer).

For evaluation along the $j\omega$ axis of the complex s-plane, the transfer function can be written

$$\begin{aligned} T(j\omega) &= H\left[\frac{(j\omega - S_{01})(j\omega - S_{02}) \cdots (j\omega - S_{0n})}{(j\omega - S_{p1})(j\omega - S_{p2}) \cdots (j\omega - S_{pm})}\right] \\ &= R(\omega) + jX(\omega) = |T(j\omega)| \exp\{j\Phi(\omega)\} \\ &= \exp\{A(\omega) + j\Phi(\omega)\}. \end{aligned} \qquad (3)$$

In logarithmic form,

$$A(\omega) = \ln |T(j\omega)| = \ln H + \sum_{i=1}^{n} \ln (j\omega - S_{0i}) - \sum_{i=1}^{m} \ln (j\omega - S_{pi}), \qquad (4)$$

$$A(\omega) \text{ db} = 20 \log |T(j\omega)| = 20 \log H + 20 \sum_{i=1}^{n} \log (j\omega - S_{0i}) - 20 \sum_{i=1}^{m} \log (j\omega - S_{pi}), \quad (5)$$

$$\Phi(\omega) = \arg T(j\omega) = \pm (m - n)\, \pi + \sum_{i=1}^{n} \arg (j\omega - S_{0i}) - \sum_{i=1}^{m} \arg (j\omega - S_{pi}), \quad (6)$$

where $A(\omega)$ is the attenuation function and $\Phi(\omega)$ is the phase function.

The network transmission delay will be defined as

$$D(\omega) = \frac{d\,\Phi(\omega)}{d\omega}. \quad (7)$$

Since the given graphical requirements will be in terms of $A(\omega)$ (gain), and/or $\Phi(\omega)$ (phase), and/or $D(\omega)$ (delay), the approximation problem resolves to the selection of pole–zero factors for $T(s)$ that will provide an "acceptable" match to the desired $A(\omega)$, $\Phi(\omega)$, and $D(\omega)$. The factors selected must also meet any special network or structure restrictions placed on the pole–zero locations so that the final transfer function can be realized by physical elements, is stable, and meets other practical requirements for the type of network structure desired (ladder, lattice, bridged-T, parallel-T, etc.).

In general, for the transfer-function approximations considered in this paper, these restrictions may be summarized as follows:

(1) Poles of $T(s)$ cannot have positive real parts (i.e., they must be located in the left half of the s-plane).

(2) Poles of $T(s)$ on the $j\omega$ axis are not allowed.

(3) Zeros of $T(s)$ will also be located in the left half s-plane (i.e., minimum phase functions).

(4) Zeros of $T(s)$ on the $j\omega$ axis may be multiple.

(5) All poles or zeros of $T(s)$ not occurring on the negative real $(-\sigma)$ axis, will occur in complex conjugate pairs.

Network Function Approximation

The approximation problem has been studied extensively by many workers in the field of circuit theory. A brief listing of some of the important techniques used to develop or generate approximating functions would include

(a) approximation by point coincident polynomials[1,2]

(b) the Fourier series polynomial method[2]

(c) approximation by Butterworth, Chebychev, Legendre, Bessel polynomials[3-7]

(d) approximation by elliptic functions[8,9]

(e) semigraphical methods[10-12]

(f) approximation by potential analog theory.[1,2,13]

Of the techniques listed, the potential analog theory has attracted considerable interest because of the flexibility and insight possible with this method.

It can be demonstrated that the gain and phase functions of linear networks are the real and imaginary components of analytic functions of a complex (frequency) variable, and satisfy Laplace's equation

$$\nabla^2\psi(x, y) = 0. \tag{8}$$

Thus, an analog exists between network functions and two-dimensional potential functions, and the many applications of potential function theory may be used in network function approximation. The electric

---

[1] J. E. Storer, *Passive Network Synthesis*, McGraw-Hill Book Co., New York, 1957.

[2] S. Winkler, "The Approximation Problem in Network Synthesis," *I.R.E. Trans. on Circuit Theory*, Vol. CT 1, p. 5, Sept. 1954.

[3] S. Butterworth, "On the Theory of Filter Amplifiers," *Experimental Wireless*, Vol. 7, p. 536, Oct. 1930.

[4] S. Darlington, "Network Synthesis Using Tchebycheff Polynomial Series," *Bell Sys. Tech. Jour.*, Vol. 31, p. 613, July 1952.

[5] Y. H. Ku and M. Drubin, "Network Synthesis Using Legendre and Hermite Polynomials," *Jour. Frank. Inst.*, Vol. 273, p. 138, Feb. 1962.

[6] W. E. Thomson, "Networks with Maximally-Flat Delay," *Wireless Engineer*, Vol. 29, p. 256, Oct. 1952.

[7] A. Papoulis, "A New Class of Filters," *Proc. I.R.E.*, Vol. 46, p. 649, March 1958.

[8] W. Cauer, *Theorie Der Linearen Wechselstromschaltung*, Akademie-Verlag, Berlin, Germany, 2nd Ed., 1954.

[9] A. Grossman, "Synthesis of Tchebycheff Parameter Symmetrical Filters," *Proc. I.R.E.*, Vol. 45, p. 454, April 1957.

[10] H. W. Bode, *Network Analysis and Feedback Amplifier Design*, Van Nostrand Co., Inc., New York, 1945.

[11] R. F. Baum, "A Contribution to the Approximation Problem," *Proc. I.R.E.*, Vol. 36, p. 863, July 1948.

[12] J. M. Linke, "A Graphical Approach to the Synthesis of General Insertion Attenuation Functions," *Proc. I.E.E.*, Vol. 97, Pt. III, p. 179, May 1950.

[13] S. Darlington, "The Potential Analogue Method of Network Synthesis," *Bell Sys. Tech. Jour.*, Vol. 30, p. 315, April 1951.

potential distribution set up in a thin conducting sheet by an array of line charges piercing the sheet, as measured along a specified axis on the sheet, is analogous to the network-function attenuation characteristic measured along the $j\omega$ axis. The locations of the line charges correspond to the locations of the network-function poles and zeros, and thus provide a flexible physical model of the network function to expedite solving the approximation problem.

Two-dimensional potential theory ideally assumes point charge sources and a uniformly conductive sheet of infinite extent and infinitesmal thickness. Under these assumptions, the electric current flow from (or towards) a single point source is perfectly radial, and the potential distribution around this point is a logarithmic function in two dimensions.

If the "potential plane" is given dimensions

$$\lambda = x + jy, \tag{9}$$

the potential in the plane is proportional to the electric charge at the point source and to the logarithm of the distance from the point source;

$$V = Q \log |\lambda - \lambda_s| + K. \tag{10}$$

Consider the complex potential function

$$W = Q \log (\lambda - \lambda_s) + K = V + j\psi = |V| e^{j\psi}, \tag{11}$$

where $V(\lambda)$ is the potential function and $\psi(\lambda)$ is the stream function. W is defined as an analytic function except at the point charge sources. Therefore, the theory of analytic functions of a complex variable may be used to establish the properties of W and to present an analogy with network functions.

The complex potential function has a derivative defined by

$$\frac{dW}{d\lambda} = \frac{\partial V}{\partial x} + j\frac{\partial \psi}{\partial x} = \frac{\partial \psi}{\partial y} - j\frac{\partial V}{\partial y}. \tag{12}$$

The potential and stream functions satisfy the Cauchy–Riemann relations for analytic functions;

$$\frac{\partial \psi}{\partial x} = \frac{-\partial V}{\partial y}, \qquad \frac{\partial \psi}{\partial y} = \frac{\partial V}{\partial x}. \tag{13}$$

Other properties of potential functions may be established, but the analogy to network functions may now be made. For an ideal potential system of point sources (and/or sinks), the complex potential can be expressed in summation form as

$$W = 20 \log C + 20 K_1 \sum_{i=1}^{n} \log (\lambda - \lambda_{0i}) - 20 K_1 \sum_{i=1}^{m} \log (\lambda - \lambda_{pi}), \quad (14)$$

where C and K_1 are constants dependent on the voltage–current relations of the conducting sheet, and the point sources and sinks are located at λ_{0i} and λ_{pi}.

The potential function V along the Y potential axis is given by

$$V(y) = 20 \log C + 20 K_1 \sum_{i=1}^{n} \log (jy - \lambda_{0i}) - 20 K_1 \sum_{i=1}^{m} \log (jy - \lambda_{pi}) \quad (15)$$

and the stream function along this axis is

$$\psi(y) = \pm K_2 + \sum_{i=1}^{n} \arg (jy - \lambda_{0i}) - \sum_{i=1}^{m} \arg (jy - \lambda_{pi}). \quad (16)$$

Comparison with Equations (5) and (6), shows that the network gain function $A(\omega)$ is analogous to the potential function $V(y)$, and the network phase function $\Phi(\omega)$ to the stream function $\psi(y)$.

Using the Cauchy–Riemann relations the network delay function corresponds (in the limit) to the orthogonal potential function partial derivative;

$$D(\omega) = \frac{d\,\Phi(\omega)}{d\omega} \simeq \frac{\partial V}{\partial x}. \quad (17)$$

Thus network functions can be created by suitable point-charge distributions on an ideal potential plane.

All of the above has been known for many years, and considerable mathematical effort has been expended[13-16] to produce theories and techniques for discrete charge distributions whose potential functions approximate the required $A(\omega)$ (and/or $\Phi(\omega)$, $D(\omega)$) to within speci-

---

[14] T. R. Bashkow, "A Contribution to Network Synthesis by Potential Analogy," Tech. Report No. 25, Electr. Res. Labs., Stanford Univ., June 30, 1950.

[15] G. L. Matthaei, "A General Method for Synthesis of Filter Transfer Functions as Applied to L-C and R-C Filter Examples," Tech. Report No. 39, Electr. Res. Labs., Stanford Univ., Aug. 31, 1951.

[16] A. R. Boothroyd, "Design of Electric Wave Filters with the Aid of the Electrolytic Tank," *Proc. I.E.E.*, Vol. 98, Pt. IV, p. 65, Oct. 1951.

fied error tolerances. However, the physical imperfections in the structure of various types of potential-plane systems devised previously has limited their use to research laboratories, and the range of derivable functions to those with maximum magnitude variations of 30 to 40 db, and "flatness" requirements of ± (0.5 to 1.0) db. The network functions required during the previously mentioned wave-filter design program had specified magnitude ranges of 70 to 100 db, with "flatness" requirements of ± (0.1 to 0.5) db. Therefore, efforts were extended to improve and refine potential-plane construction and operational techniques to meet these requirements.

Before elaborating on the development of the potential plane and the results obtained, a very brief background on previous potential-plane systems is presented.

Liquid and Dry Conducting-Sheet Potential-Plane Systems

Potential-plane systems have been built using both liquid and dry conducting sheets. Liquid electrolytic tanks, or "planes," have a lengthy historical background. Detailed treatments of particular tank constructions and techniques have been presented.[2,17,18]

The electrolytic tank as a potential analog for network functions has been the basis for much of the original potential analog work. However there are many problems in construction and use that place limits on the accuracy of the analog, and cause inconveniences in the use and storage of the equipment.

The dry conducting-sheet plane is a more recent development. Various types of conducting sheets, using conductive plastic or colloidal graphite coatings, have been investigated, but for uniformity of resistivity, long-term stability, availability, and low cost, the carbon-impregnated paper developed for facsimile reproduction[19] ("Teledeltos" paper) has proven superior to any other known dry conducting sheet. Potential planes using Teledeltos sheets have been described in the literature.[20,21]

---

[17] W. W. Hansen and O. C. Lundstrom, "Experimental Determination of Impedance Functions by the Use of an Electrolytic Tank," *Proc. I.R.E.*, Vol. 33, p. 528, Aug. 1945.

[18] A. R. Boothroyd, E. C. Cherry, and R. Makar, "An Electrolytic Tank for the Measurement of Steady-State Response, Transient Response and Allied Properties of Networks," *Jour. I.E.E.*, Vol. 96, Pt. I, p. 163, May 1949.

[19] T. F. Cofer, "Analog Field Mapping on Teledeltos Recording Paper," Information Bull., Marketing Dept., Western Union Telegraph Co., April 1, 1960.

[20] R. E. Scott, "An Analog Device for Solving the Approximation Problem of Network Synthesis," Tech. Report No. 137, Res. Lab. of Electr., M.I.T., June 30, 1951.

[21] P. Liebman, "Simultaneous Gain-Phase Approximation with a Potential Analog Computer," AD-146246, Poly. Tech. Inst. of Brooklyn, 1954.

The conducting plane is a sheet of suitably selected Teledeltos paper, stretched over a smooth, flat surface. A conducting rim of copper or conducting silver paint provides the infinity point, and needle electrodes are used to supply the pole and zero currents. The potential-measuring electrode is usually a smooth-pointed small-diameter rod capable of exploring the potential field on the Teledeltos paper, with constant pressure supplied, for example, by spring loading.

The Teledeltos sheet has none of the liquid-electrolyte problems of corrosion, polarization, leveling, chemical contamination, or "ripples" due to electrode movement or table vibration. Also, d-c can easily be used, thus simplifying both the electrode current supplies and the potential-measuring system. However, some problem areas still remain, such as finite-plane errors, uniformity of the Teledeltos sheet, and susceptibility of the sheet to humidity and surface damage.

Finite-plane errors: As with the liquid electrolytic tank, these errors can be minimized by determining and applying correction factors to the potential measurement.[17] Also, conformal transformation into a complex plane where the physical sheet rim is mathematically located at infinity can result in negligible boundary errors.[22,23] However, this latter approach is not always convenient for the type of function to be approximated (e.g., a narrow-band-pass function). Determinations of the electrode (function root) locations are complicated by the fact that the coordinates of the transformed plane will be curvilinear rather than orthogonal, and require a de-transformation operation to obtain the s-plane co-ordinates.

Uniformity of the Teledeltos Sheet: Variations of resistivity and surface faults can be expected in the material. To minimize these errors, a reasonably large area of Teledeltos paper may be obtained and the surface examined for a suitable area. This section of the paper can then be used for the plane. Small, localized areas of high-resistivity variations provide the greatest errors when this area occurs near or between two current electrodes or along the potential-measuring axis. Fortunately, Teledeltos paper is relatively free from this defect, and gradual resistivity variations can be accommodated by a system of electrically locating the root (current) electrodes, as described later.

Susceptibility to humidity: Since the porous Teledeltos paper absorbs moisture easily, the surface may expand and contract in a non-

[22] W. H. Huggins, "A Note on Frequency Transformations for Use with the Electrolytic Tank," *Proc. I.R.E.*, Vol. 36, p. 421, March 1948.

[23] G. L. Matthaei, "Conformal Mappings for Filter Transfer Function Synthesis," *Proc. I.R.E.*, Vol. 41, p. 1658, Nov. 1953.

uniform manner. This will cause potential-measurement errors due to variations of apparent electrical surface-resistance "length." Also, the loose, expanded surface is more susceptible to tearing and puncture damage from the thin wire electrodes. This problem can be minimized by locating the potential plane (and the supply of Teledeltos paper) in an air-conditioned area.

Susceptibility to surface damage: The surface of the Teledeltos paper is relatively soft, and it is easily abraded or torn by pointed metal electrodes. Since the current electrodes must have a small point diameter to approximate a point charge, care must be exercised in moving these electrodes so as to minimize surface damage to the paper. The potential-measuring (exploring) electrode should have a suitably rounded, smooth point of relatively soft metal; contact pressure should be kept at a minimum.

Because of the greater ease with which it can be used, the Teledeltosheet system has definite advantages over the electrolytic tank system, provided a suitable means can be devised to eliminate boundary error. The next section describes the construction of such a plane, made effectively infinite through the use of a double-sided plane of Teledeltos sheets.

Development of an Infinite-Potential Analog Plane

It has been shown[18] that two finite conducting electrolytic sheets could be combined to simulate an infinite conducting sheet.

Let $v(r, \theta)$ be the potential distribution over the surface of an ideal infinite conducting sheet with m current sources and n current sinks, and $v_1(r, \theta)$ be the potential distribution over the surface of an ideal infinite conducting sheet with one current sink of order $(m-n)$ at the center. Now if these two sheets are made contiguous along a circular boundary at a radius r_0 from the center of each, the potentials at this boundary on each sheet are equal, and the potential distributions on each sheet may be expressed

$$V(r, \theta) = v(r, \theta) + \delta(r, \theta), \text{ for the first sheet,} \tag{18}$$

$$V_1(r, \theta) = v_1(r, \theta) + \delta_1(r, \theta), \text{ for the second sheet.} \tag{19}$$

Potential functions V, v, δ, V_1, v_1, δ_1 all must satisfy Laplace's equation everywhere. Applying the boundary conditions,

$$\left.\frac{\partial V(r,\theta)}{\partial r}\right|_{r=r_0} = -\left.\frac{\partial V_1(r,\theta)}{\partial r}\right|_{r=r_0}, \tag{20}$$

$$V(r_0,\theta) = V_1(r_0,\theta), \tag{21}$$

and the fact that $v_1(r,\theta)$ is a polar symmetric function, it can be shown that all conditions are satisfied on the boundary if

$$\delta(r_0,\theta) = 0. \tag{22}$$

If $\delta(r,\theta) = 0$ at r_0, and $V(r_0,\theta) = v(r_0,\theta) = V_1(r_0,\theta) = v_1(r_0,\theta) + \delta_1(r_0,\theta)$, then the potential function at the boundary radius r_0 for the first sheet is the same before and after the sheets are joined. Since the m current sources and n current sinks still exist in the same radial locations on the first sheet and

$$v(r_0,\theta) = K\sum_{i=1}^{n}\log|p_{0i}-r_0| - K\sum_{i=1}^{m}\log|p_{pi}-r_0|, \tag{23}$$

then

$$v(r,\theta) = V(r,\theta) \tag{24}$$

and $\delta(r,\theta) = 0$ everywhere on the first sheet.

Thus the potential distribution on the first sheet behaves as if the dimensions of the first sheet were truly infinite. This will be true only if the two sheets can be electrically joined without creating a surface-resistivity discontinuity. It has been empirically determined that a very close approximation to such a "join" of Teledeltos conducting sheets may be accomplished by suitable over-lapping of the two sheets.

Teledeltos type L paper consists of a carbon-filled paper base coated on one side with a lacquer-bound aluminum pigment. For facsimile recording, the paper is coated on the other side with an electro-sensitive pigment. For use as a potential mapping sheet, however, the electro-sensitive coating is not present, so the paper presents a black (carbon) side and a silver (aluminum coated) side. The sheet thickness is approximately 0.0035 inch. The sheet resistance per square is approximately 1500 ohms on both sides. The black side is used for potential-mapping purposes, and resistivity variations over reasonably large areas may be as much as 15 percent. The paper can be obtained in rolls 3 feet wide and 10 or 20 feet long. Three-foot-square sheets are cut out and tested for regional resistivity variations by placing conductors along opposite sides of the sheet, and accurately measuring

the potential on the paper surface at intervals of distance between the two conductor edges. Sheets are selected whose resistivity variations expressed as potential variations from a straight-line function are no greater than ± 1 percent.

If such a sheet is cut across the middle and the two cut edges are joined by overlapping approximately 3/8 inch, with the black surface mated to the silver surface under strip pressure clamps, the potential distribution across the "joined" section is approximately a straight-line function with potential (and, therefore, resistivity) variations no greater than ± 1 percent, to within 0.25 inch of the "join." Thus, two discs of Teledeltos paper can be similarly joined at their rims, with assurance that to within 0.25 inch of the join, the resistivity variations of the mapping surface are no greater than those that would be presented by a similar area of selected Teledeltos sheet.

Mechanical and Electrical Details of the Potential Analog System

Apart from the potential plane and associated mechanics, standard items of laboratory equipment are used in the system. A general view of the equipment is shown in Figure 1, and circuit details are given in Figure 2.

An 11×17 inch X–Y recorder is used to display both the graphical specification requirements and the magnitude or delay function corresponding to the configuration of poles and zeros on the plane. The X–Y recorder offers many attractions as a display device. For instance, it has the major advantages over an oscilloscope of high resolution on a large display area and simplicity in obtaining permanent records of the graphical requirements with a superimposed approximating function. Also, information can be plotted, on the same sheet, from which the plane locations of the poles and zeros of the approximating function are determined.

The Conductive-Plane Assembly

The potential-analog plane is formed by two discs of Teledeltos paper 31⅝ inches in diameter separated from each other, by a thin insulating disc assembly, except for an annular region at the rim over which they are maintained in pressure contact.

This annular region is formed by clamping the paper discs between plexiglass annuli of 31.5 inches outside diameter and ⅜ inch radial width. Pressure is applied over the region by a ring of 77 metal spring clips, each 1.25 inches wide. The assembly separating the paper sheets consists of two thin discs of plexiglass 30⅝ inches in diameter, be-

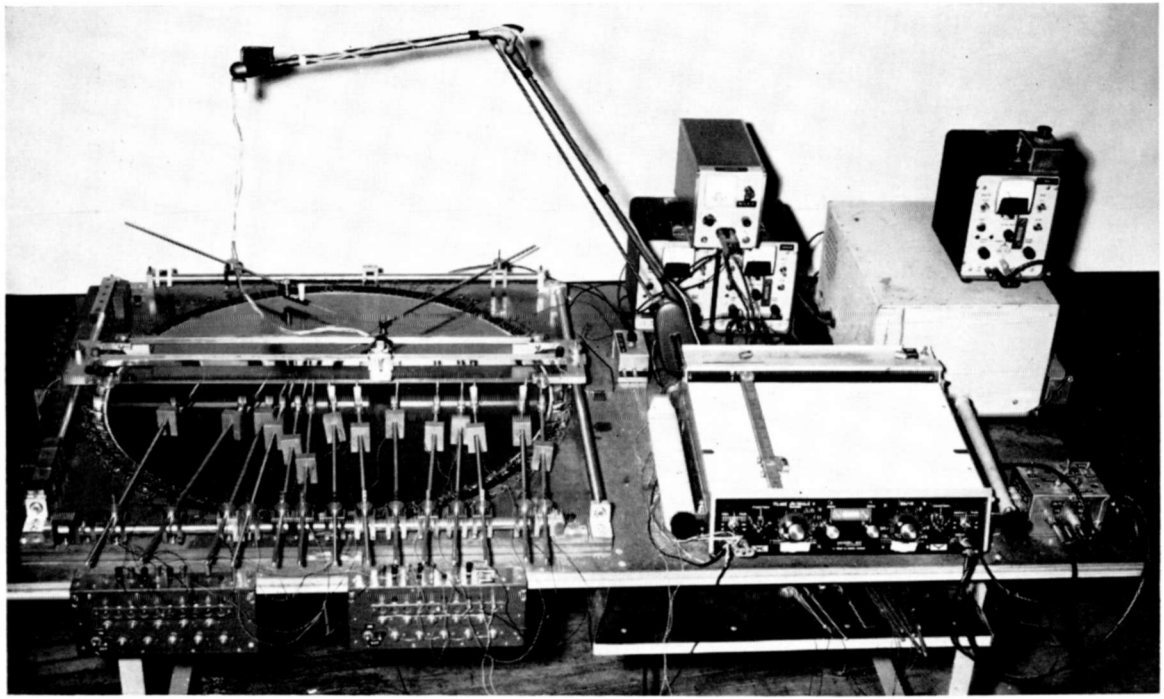


Fig. 1—Complete infinite potential plane system.

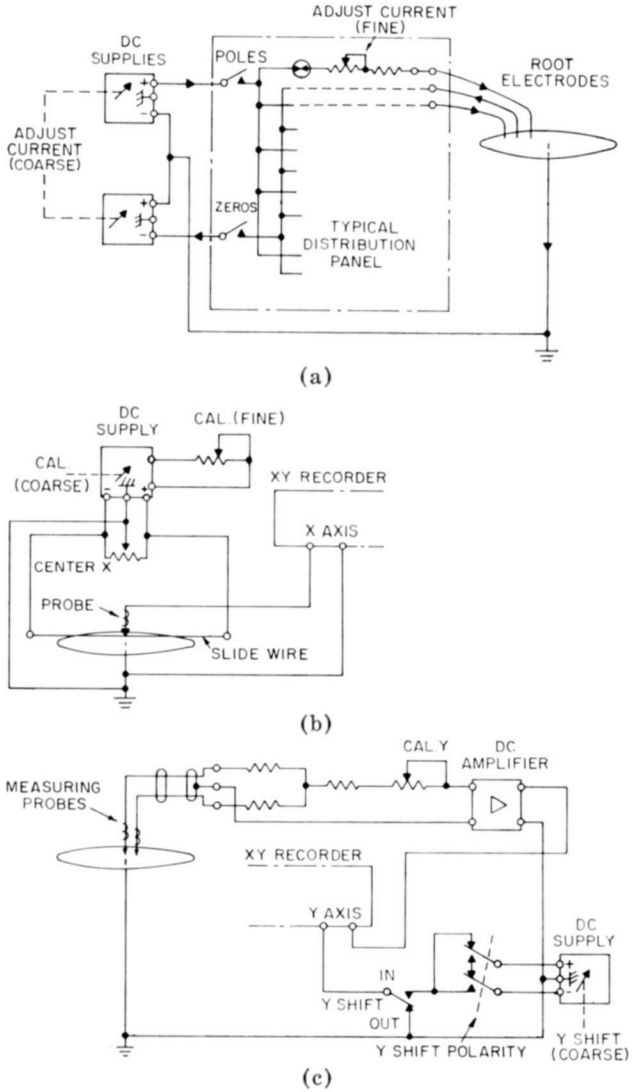


Fig. 2—Basic circuits of potential plane system: (a) Root-electrode current supply, (b) X axis (frequency) and (c) Y axis (log magnitude and delay).

tween which is cemented a thin copper disc; the copper disc is grounded and acts as an electric shield between the paper discs. Both of the paper discs have the black side upper-most. The lower annulus is cemented to a plexiglass disc 31¼ inches in diameter that is cemented to a plywood base ⅝-inch thick and 31 inches in diameter. Figure 3 shows a cross section of the plane assembly.

Application of the potential analog requires, for exactness, that the conducting plane be of infinitesmal thickness and have no finite boundary. The thinness of the paper gives a good approximation to

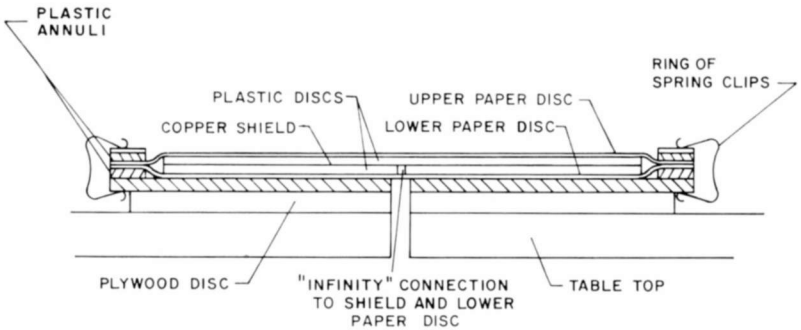


Fig. 3—Sectional view of plane assembly.

the first requirement and the joining of the two paper discs in the manner described gives a good approximation to the second requirement. The upper disc is effectively any circular region in a plane of infinite extent, and the lower disc is, in effect, the rest of the infinite plane with the point representing infinity located at its center.

The other requirement for exactness of the analog is uniform plane resistivity. This is reasonably met by selecting the upper Teledeltos disc by measurements of uniformity of resistivity. It is not necessary to specially select paper for the lower disc.

The Method of Representing Function Roots

Poles and zeros of the approximating function are represented by current-carrying electrodes consisting of silver-plated sewing-machine needles ground flat at one end to form a 0.030-inch-diameter contact area.

Each needle is mounted on the end of an insulating hinged arm carried by a long steel rod. This rod passes through a rotatable turret mounted on an inverted U-shaped block that may be moved along one

of four rods forming a square about the plane assembly. The hinged arm, turret and U-block are fitted with locking screws. With this arrangement electrodes may be manipulated conveniently about the plane during the approximating process and locked into position when required. U-shaped lead weights straddling the electrode rods serve to maintain suitable contact pressure between the electrode needles and the plane.

The electrodes may be supplied by either alternating or direct current. Normally, direct current is used, and poles and zeros are represented by feeding the electrodes with positive and negative current, respectively.

Six current distribution panels to which the electrodes may be connected are mounted around the plane. Each panel, except the calibration panel, is arranged to provide current for five poles and five zeros. Each bank of five pole and five zero outlets has an associated on/off switch mounted on the distribution panel. Thus the current to the electrodes may be readily switched off so that calibration checks can be made at any time.

It is essential, in order to maintain the analog, that all electrodes pass the same amount of current. The magnitude of current adopted is about 0.9 milliampere and all electrode currents may be adjusted to the same magnitude by means of variable resistors provided on each current distribution panel. To minimize the change of current as the electrodes are maneuvered about the plane, the source for each electrode is made to approximate a constant current source by the inclusion of fixed resistors of 50,000 ohms.

Multiple or fractional poles or zeros may be represented by passing through the representing electrode a current that is a corresponding multiple or fraction of the standard current. This is achieved by appropriate "strapping" on the current-distribution panels.

The electrode currents are monitored via miniature jacks. The monitoring plug lead is terminated with a 100-ohm resistor and the current is determined by measuring the voltage drop across this resistor.

Display of the Approximating Functions

The approximating attenuation or delay function is displayed by manually sliding a carriage, incorporating voltage-measuring probes, along a bar that spans the plane. There are three probes on the carriage. One is in contact with a resistive slide wire stretched taut along the length of the bar. The wire and probe are connected into the bridge circuit shown in Figure 2(b), so that movement of the carriage will

result in a proportional movement of the X–Y recorder pen carriage along the X axis.

The other two probes, which are connected to the Y axis of the recorder, are mounted in a subassembly that may be slid vertically in the carriage to either place the two probes in contact with, or lift them clear of, the plane. This arrangement allows the measuring probes to be retracted to clear root electrodes on or very close to the measuring axis. These two probes are mounted at right angles to the line of travel and 0.1 inch apart. For delay measurements, the line passing centrally between the probes is considered to be the $j\omega$ axis, and the voltage difference between the probes is considered to be proportional to the delay magnitude. The voltage at the junction of two equal resistors connected in series between the probes is considered to be proportional to the attenuation. When the delay function need not be measured, only one of the two probes need be used to measure the attenuation function. This has the advantage of allowing continuous plots of the attenuation function with root electrodes nearer the $j\omega$ axis. With a single measuring probe, the nearest any root needle can come to the $j\omega$ axis, without collision between the measuring probe and the needle is about 0.12 inch. For roots on the $j\omega$ axis or roots closer to the $j\omega$ axis than 0.12 inch, a discontinuous plot may be obtained by using the retractable feature of the measuring-probe arrangement to "hop over" these roots.

A plane scale factor F (in cycles per second per inch) must be chosen to include on the plane the interval of frequency over which the function must be plotted. As low a scale factor as possible is generally desirable to obtain maximum separation of the roots, particularly those at the edge of the passband, for high-order transfer functions. The minimum separation for the root needles, without touching, is about 0.1 inch. For a given scale factor, F, the maximum frequency range that can be displayed, using a diameter of the plane to represent the $j\omega$ axis, is $30F$ cps.

By using network criteria other than transfer functions (e.g., characteristic functions) and conformal transformations of the plane, the restrictions imposed by the size of the plane and root needles may be minimized.

The relationship of plane-potential difference to log-magnitude, based on a "standard" electrode current, I, of 0.9 milliampere and resistance per square, R_s, for the plane of 1500 ohms, is about 25 millivolts/db. The "output resistance" of the plane is about 3200 ohms. Thus, the display equipment should have an input impedance about a hundred times this, i.e., about 300,000 ohms. The X–Y recorder avail-

able had an input impedance less than 300,000 ohms below 200 millivolts/inch. Since a sensitivity of $25y$ millivolts/inch is required for y db/inch, this X–Y recorder can only be connected directly to the plane when scales of 10 db/inch or greater are desired. For greater sensitivity, an impedance converter is required. For this purpose, a d-c amplifier with an input impedance of 100,000 ohms was used with the input circuit shown in Figure 2(c). The two potentiometers included in this circuit provide both a coarse and fine control of Y-axis gain. To compensate for the loss introduced by the input circuit, the amplifier gain is set to $\times$ 10. With this arrangement the Y-axis sensitivity can be adjusted in convenient steps after initial calibration, which is generally done at 10 db/inch with the X–Y recorder at 200 millivolts/inch. Of the available scales, the most often used are 0.5, 1.0, 2.0, 5, 10 and 20 db/inch.

Although there is virtually no limit to the lowest sensitivity that may be used, a sensitivity of 20 db/inch gives a total available range of 200 db. This was the maximum required for the network functions created.

At the higher sensitivities it is found that a limiting factor in the increased resolution obtained from greater sensitivity is the amplifier noise and irregular voltage pickup from the plane due to the granularity of the surface. Surface granularity causes small random deflections of the recorded line with an average peak-to-peak magnitude of about 0.05 db.

Experiments with different kinds of measuring-probe materials, finishes, and shapes to minimize the granularity effect led to the conclusion that use of detachable unfinished rounded copper tips on the ends of copper rods is the most satisfactory arrangement.

When making measurements of passband or stop-band regions at high sensitivity, the absolute magnitude of the potentials measured is generally high relative to the full-scale deflection of the recorder, and it is thus necessary to have means for bucking this potential to obtain an on-scale recording. The shift control provided on the recorder is generally insufficient and is used mostly as a fine control. The Y-shift power unit, shown in Figure 2(c), may be switched in series with the Y-axis input if required. The polarity may also be switched. This unit therefore provides means for large Y-axis shifts and is used as a coarse control.

The relationship of plane potential to delay is about $135F$ microvolt/millisecond so that high sensitivity is generally required for delay measurement. Ideally, a differential amplifier would be used for

delay measurement, but the delay can also be obtained by taking the difference of plots with first one and then the other of the two voltage-measuring probes connected into the measuring system. This latter method is used for the delay example.

Calibration

Frequency Axis (cps/inch)

Calibration of the frequency axis is accomplished with the aid of a template bar for positioning the traveling carriage. This bar may be positioned above the bar on which the traveling carriage is mounted and located in position by studs protruding from the carriage-bar mounting assemblies. There are holes in the template bar, accurately located at its center and 14 inches on either side of center, that accommodate a stud protruding from the carriage.

With no input to the X axis, the pen is adjusted to the center of the X axis by means of the recorder X-shift control. By means of the template bar, the measuring probe carriage is positioned at the center of the $j\omega$ axis. The center X control is then adjusted so that, with the bridge output connected to the X axis, the pen remains at its central position. The carriage is moved to either of the locations 14 inches from center, and the voltage to the bridge circuit is then adjusted so that the pen deflects exactly 7 inches from center.

Logarithmic Magnitude (db/inch)

The method adopted for calibrating the Y axis in terms of db/inch is as follows. Two root electrodes, situated on a plane diameter normal to the $j\omega$ axis at distances of 3.000 inches and $3\sqrt{10}$ inches from the plane center, are supplied in turn with "standard" current. The measuring probes are positioned at the plane center by means of the template bar. When both measuring probes are used, the difference between the deflections of the pen, as current is switched between the root electrodes, is adjusted by means of the potentiometers in the amplifier input circuit to correspond to 10 db. Thus, for example, on the 10 db/inch scale, the difference in pen deflections would be adjusted to 1 inch.

This procedure is based on the fact that the db difference, N, in the potential magnitude between two points at distances x_1 and x_2 from a single root is given by

$$N = 20 \log \frac{x_1}{x_2}. \tag{25}$$

Thus, for $x_1 = x_2\sqrt{10}$, $N = 10$.

To avoid moving the measuring probes during calibration, two roots are used instead of one. This considerably simplifies the procedure.

The change in plane potential, V_d, at the mid-point between the measuring probes as current is switched between the roots is given by

$$V_d = \frac{IR_s}{1.7372\pi}, \tag{26}$$

and the corresponding pen-deflection difference, D, is given by

$$D = \frac{GIR_s}{1.7372\pi}, \tag{27}$$

where G is the gain relating the recorder pen movement to the $j\omega$ axis potential. G has units of inches/volt. Therefore, differences in R_s, which may be experienced between different sheets of Teledeltos paper, and different values of I, are automatically accommodated by the adjustment of G. However, the current I used for calibration must be used for all roots during the approximation procedure.

A semicircular plexiglass template, with appropriately positioned holes, is used to simplify the positioning of the two roots. This template mounts on the assemblies holding the carriage bar in position. By means of fine adjustments provided in the mounting arrangements the template may be accurately located relative to the $j\omega$ axis. Also, it may be mounted on either side of this axis. It is usually positioned in the right half plane, so that the calibration electrodes remain in position and calibration checks may be made during the approximating process.

Delay (milliseconds/inch)

Unlike the frequency and log-magnitude calibrations, the delay calibration involves the plane frequency-scale factor, F. The calibration procedure is similar to that for log-magnitude in that a reference delay magnitude is obtained by positioning the measuring probes in line with a root electrode a known distance away, and the system gain is adjusted to provide some convenient milliseconds/inch scale. The method is based on the relationships

$$D = \frac{d\phi}{d\omega} = \frac{1}{\sigma} \text{ seconds},$$

where D is the absolute delay and σ is distance of root from $j\omega$ axis in radians/second. Therefore, $D = 1/(2\pi F\sigma)$ seconds, where F is in cps/inch and σ is in inches.

For example, using the root electrode spaced 3 inches from the $j\omega$ axis and a plane frequency-scale factor of 500 cps/inch, $D = 106.2$ microseconds. This corresponds to approximately 7 millivolts potential difference between the measuring probes.

Determination of Root Locations

When an arrangement of root electrodes has been obtained that indicates a satisfactory approximation to the specified log-magnitude and/or delay requirements, the s plane locations of the roots must then be determined.

Designating the location of a root as $\sigma + j\omega$, σ and $j\omega$ are determined for each root by passing current through each root electrode in turn and moving the measuring probes past the current-carrying electrode to record the peak of the log-magnitude response. To determine the σ locations of roots not on the $j\omega$ axis, the standard current must be used. The location σ_x of root x is obtained from the relation

$$\sigma_x = \text{antilog}\left[\log \sigma_r - \frac{(D_x - D_r)}{2}\right] \text{inches}, \tag{28}$$

where σ_r is the distance between a reference electrode and the $j\omega$ axis in inches. D_x and D_r are the peak pen deflections due to the root-x electrode and the reference electrode, respectively, in inches on the 10 db/inch range.

This identity may be derived as follows. The plane-potential difference between two points at distances σ_1 and σ_2 from an electrode is

$$V_1 - V_2 = \frac{IR_s}{1.7372\pi} \log\left(\frac{\sigma_2}{\sigma_1}\right)^2 \text{volts}, \tag{29}$$

and the corresponding pen-deflection difference is

$$D_1 - D_2 = \frac{GIR_s}{1.7372\pi} \log\left(\frac{\sigma_2}{\sigma_1}\right)^2 \text{inches}. \tag{30}$$

Since, in the calibration procedure for log-magnitude, $GIR_s/(1.7372\pi)$ is made equal to one inch at a sensitivity of 10 db/inch, after calibration

$$D_1 - D_2 = \log \left(\frac{\sigma_2}{\sigma_1} \right)^2 \text{ inches.} \tag{31}$$

By replacing subscripts 1 and 2 in Equation (31) by r and x, respectively, Equation (28) may be obtained.

The electrode that is positioned 3 inches from the $j\omega$ axis and used in the log-magnitude calibration procedure is normally used as the reference electrode.

The $j\omega$ root location is obtained by noting the frequency axis location of the response peak on the X-Y recorder graph. For roots on the $j\omega$ axis, the carriage bar is shifted parallel to the $j\omega$ axis so that a response peak may be recorded.

The response peak, and hence the $j\omega$ location, is more easily resolved the closer the measuring probes are to the root electrode. For roots with large σ values, it is advantageous to move the carriage bar toward them before recording the response.

Transfer Function Approximation

General

The number and arrangement of pole–zero electrodes selected to form a suitable approximating function depends on whether the approximation is

(a) primarily concerned with magnitude or delay,

(b) intended to fit a prescribed curve by
  (1) least-mean-square error criteria,
  (2) maximum peak deviation error criteria,
  (3) equal ripple approximation,

(c) intended to fit within a set of arbitrary limits (such as is usually the case for filter functions) by
  (1) maximum peak deviation error criteria,
  (2) equal ripple approximation.

For the purposes of this paper, it is assumed that the functions required are of the filter-function type, and that either (c1) or (c2) is intended. For further simplification, it is assumed that magnitude characteristics are of primary interest, and that the delay characteristic is to be examined after a satisfactory magnitude characteristic has been created.

It should be noted that equalizer-type transfer-function magnitudes have been successfully approximated (in the maximum peak deviation

sense) for relatively low-order functions (e.g., fifth order). However, the complexity of the pole–zero selection problem increases rapidly for higher-order transfer functions. An equalizer-type transfer-function magnitude problem is defined as approximating, to within certain criteria, a given magnitude curve expressed as log magnitude versus frequency over a specified frequency range. A filter-function magnitude problem is defined as creating a rational function whose magnitude characteristic falls within specified log-magnitude *limits* over a specified frequency range.

In either case, the network-function poles (and perhaps also the zeros) will have both real and imaginary parts, and so require manipulation in two dimensions. Since all function roots contribute to the function magnitude at any particular frequency, manipulation of the pole-zero electrodes effectively involves solving simultaneous equations by trial adjustments and so, to a large extent, is dependent upon the experience and intuitive ability of the operator. However, several aids have been devised to assist such pole–zero manipulations. These aids are based on two general observations.

First, the potential magnitude contribution of a pole (or zero) q, located a distance r from a measuring point x on the $j\omega$ axis, is unchanged if the pole (or zero) is moved so as to keep a constant distance r from the point x. However, the phase (and delay) at the measuring point x will be altered.

Second, using partial derivatives of the individual log functions of the network function in factored form, sensitivity functions may be plotted over a significant range of (normalized) frequency. This can provide a means for pre-estimating the overall effects of specific shifts in both the real and imaginary parts of pole (zero) locations. Similar partial derivative curves were originally presented by Linvill[24] and have been prepared in normalized form as pole–zero deviation functions for real and quadratic roots by one of the authors.[25] Figure 4 shows a typical deviation function "family" plot for $j\omega$ axis magnitude of a pair of complex conjugate roots with various damping factors, having root shifts in the σ direction as a parameter.

Conjugate Roots

The infinite potential plane may be regarded as a near-perfect representation of any or all regions of the s-plane or of any region

[24] J. G. Linvill, "The Selection of Network Functions to Approximate Prescribed Frequency Characteristics," Tech. Report No. 145, Res. Lab. for Electr., M.I.T., March 1950.

[25] F. M. Brock, *The Transfer Function Slide Rule as an Aid to the Approximation Problem in Network Synthesis*, M.Sc. Thesis, Moore School, Univ. of Pennsylvania, April 1962.

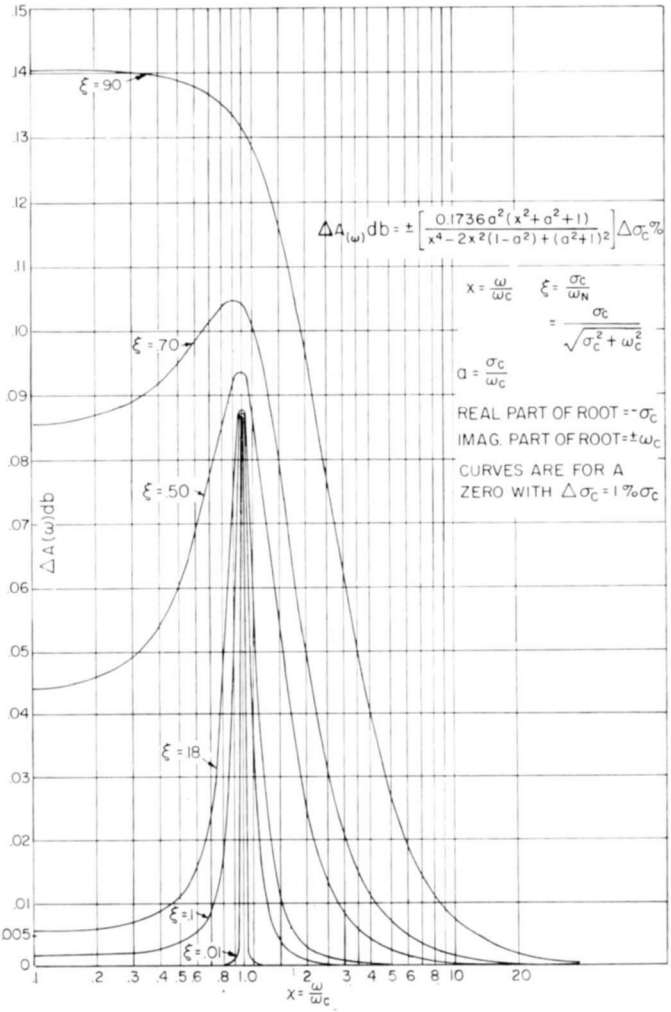


Fig. 4—Graph of deviation function for a root real part shift.

of any plane into which the *s*-plane has been conformally transformed. Thus, for the *s*-plane, the region presented on the potential-plane top surface can be considered either to include the origin or to be far removed from the origin, and to include or not include the region of the conjugate roots. If the frequency scale factor desired for reasons of accuracy causes the conjugate root locations and the roots at the origin to fall "off the plane," these must be allowed for or an error will result in the final approximation. For many narrow-band-pass

transfer functions, the conjugate and origin roots are effectively located at an extreme distance relative to the region presented by the potential plane, and their contributions in the measuring region have an insignificant effect on the approximated magnitude or delay. For band-pass functions of significant relative bandwidth this may not be true, and some consideration as to conjugate-root effects may be required.

In many cases a first trial location of poles–zeros, ignoring the conjugates, may be made. Then the response limits of approximation can be deliberately pre-distorted on the basis of the effects of the origin roots, plus a single conjugate root of order $(m-n)$ located a distance $2\omega_0$ from the center of the band-pass region. Here m and n are the numbers of poles and zeros, respectively, in the band-pass region of the trial approximation and ω_0 is the center of the band-pass function. A second approximation is then obtained by minor shifts of the first trial approximation pole–zeros to meet the pre-distorted limits.

When the band-pass function required has a very large relative bandwidth, the above methods are not satisfactory. Where magnitude approximation only is required, a simple conformal transformation of the s-plane may be used for restricted classes of transfer functions. The magnitude approximation would be done in the W plane, where

$$W = s^2. \tag{32}$$

Squaring the s-plane transforms conjugate roots in the s-plane into double roots in the W-plane. This, and other transformations intended to eliminate the conjugate roots are still under investigation by the authors.

Examples of Transfer Function Approximation

With an analog as flexible as the infinite potential plane, a wide range of types of transfer-function approximations could be used as examples, even if only filter transfer functions are considered. For brevity, only three types of filter transfer-function examples are discussed, and only one of these in detail. These examples are fairly representative of filter-transfer functions most commonly encountered, but of course do not describe the complete scope of transfer-function approximations possible.

Example 1

The first example is a fifth-order, maximally flat magnitude, low-pass transfer-function approximation. The potential-plane plot of both

log magnitude and delay is shown in Figure 5. The 3-db pass-band width is set at 0.75 cps, and log-magnitude plots at scale settings of both 10 db/inch and 5 db/inch are shown. To check this function, a digital computer was used to compute the log magnitude at various check frequencies. The computed values are marked on the plot and indicate maximum errors of 0.2 db at the 3-db point and 0.6 db in the stop band at a 31-db attenuation point.

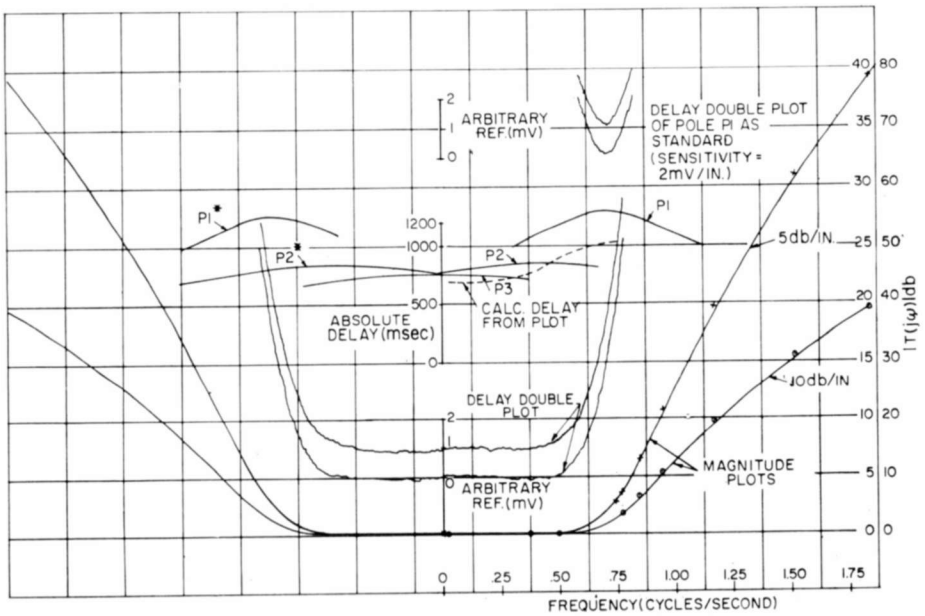


Fig. 5—Low-pass transfer-function approximation plot.

The absolute delay of this function was calculated from the delay "double-plot," and is shown as a dotted curve in Figure 5. This curve has been checked against the computed delay for a five-pole maximally flat low-pass function and found to have less than 6% error in delay from zero frequency to the frequency for 3-db attenuation.

Example 2

Figure 6 shows the X-Y recorder log-magnitude plot for a wide-band transfer function with pre-distorted limits to allow for 13 conjugate roots and 4 zeros at the origin. The magnitude is flat to within —0.5 to +1.5 db from 60 to 108 kilocycles, and provides asymmetrical stop-band requirements of >45 db at 0 to 38 kilocycles; >30 db at 112

kilocycles; ≧70 db from 120 to 168 kilocycles. This function required 8 complex conjugate poles and 5 $j\omega$-axis conjugate zeros. Computed checks on the log-magnitude of the transfer function, including conjugate roots and 4 zeros at the origin, are also marked on Figure 6 and indicate the correcting effects of approximating to pre-distorted attenuation limits.

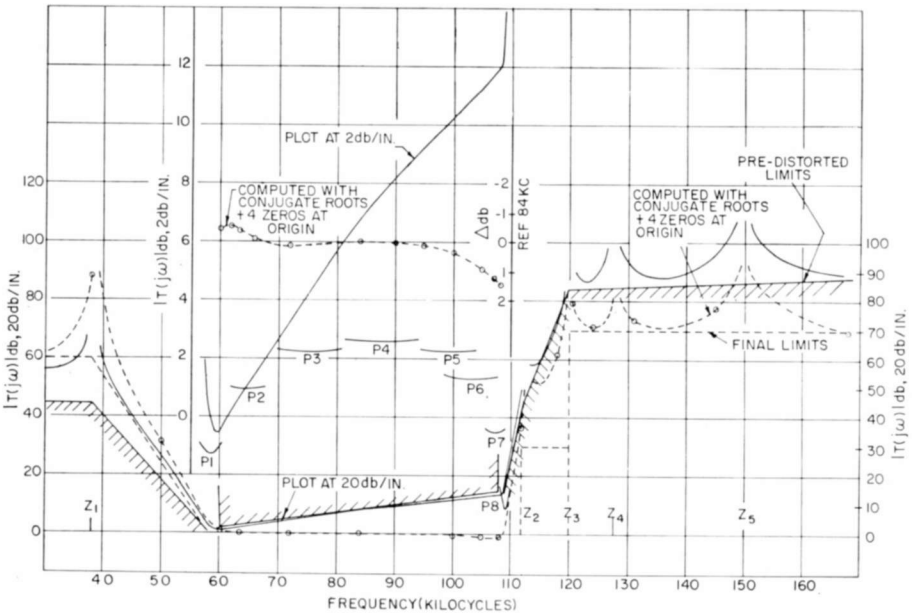


Fig. 6—Wide-band-pass transfer-function approximation plot.

Example 3

An equal-ripple approximation for a frequency-symmetrical narrow band-pass function (or for a low-pass function with conjugate roots) is shown as a magnitude plot in Figure 7 and a delay plot in Figure 8. If this approximation is considered as a narrow band-pass function with a bandwidth of 2.0 radians, the pass-band magnitude is equal ripple with a peak ripple of 1.5 db, and the stop-band minimum attenuation is 70 db for frequency ≧2.0 radians from the center frequency. Five conjugate complex poles and four $j\omega$-axis conjugate zeros are required. The magnitude characteristics plotted by the potential plane X-Y recorder demonstrate the equal-ripple magnitude properties of the well-known elliptic functions. The pass-band delay characteristics

for this transfer function (as a double-potential magnitude plot), together with a calibrating delay plot from a "standard" pole at the center frequency, are displayed in Figure 8. Using the standard pole delay plot for calibration purposes, the delay characteristic of the function pass band was measured and transferred to Figure 9.

As a check on the accuracy of both magnitude and delay plots taken from the potential plane, as well as the accuracy of locating the func-

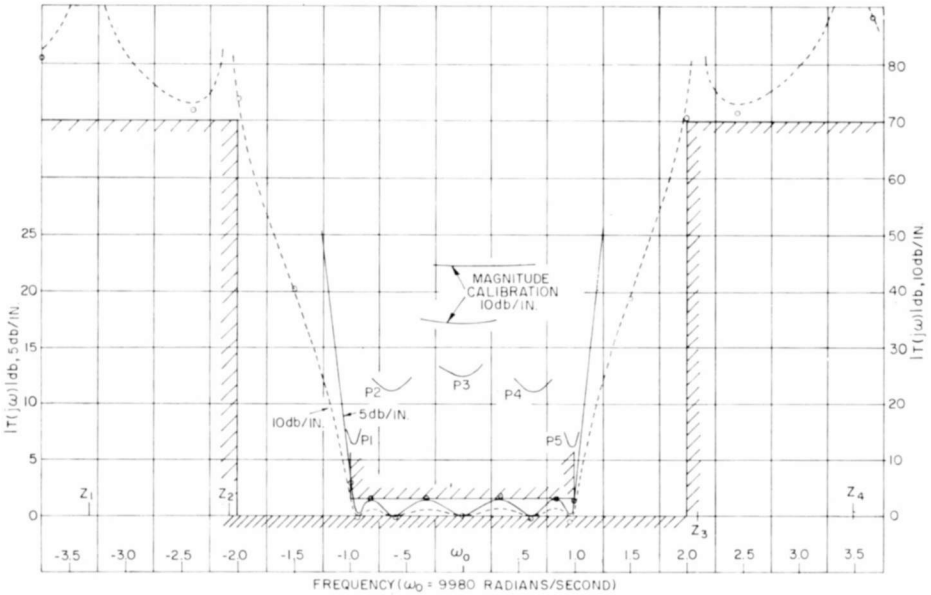


Fig. 7—Equal-ripple transfer-function approximation plot.

tion roots, a digital computer was programmed to compute, at specified frequencies, the magnitude and phase of the rational-fraction transfer function created. For this purpose, it was assumed that the function pass-band center frequency was located at 9980 radians per second (to produce a narrow band-pass function of less than 0.1% relative bandwidth). The computed magnitude points are shown on Figure 7, where the maximum error is approximately 2.0 db at a 70 db attenuation point, and 0.25 db in the pass band.

The computed delay characteristics shown in Figure 9 were determined from the phase computations by the approximation

$$D \simeq \frac{\Delta\beta}{\Delta\omega} \tag{33}$$

where D is in seconds, $\Delta\beta$ is the increment in phase (radians), and $\Delta\omega$ is the increment in frequency (radians/sec). The frequency incre-

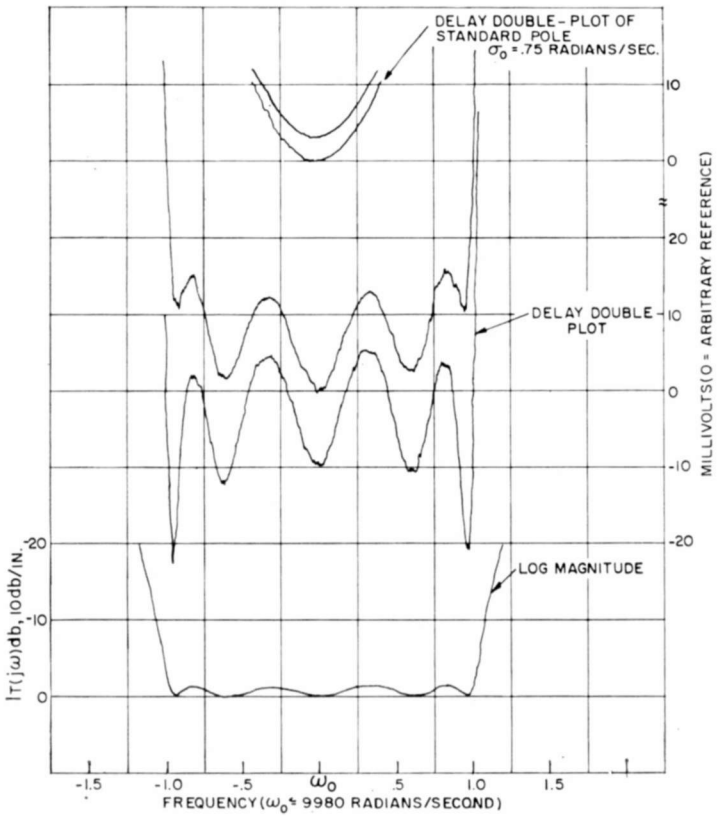


Fig. 8—Pass-band delay plot for equal-ripple transfer function.

ments used for this delay determination were 0.05 radian per second steps. Examination of the phase-versus-frequency computations later revealed that these frequency increments were large enough to introduce error into the approximate calculation of D by Equation (33). Therefore, some of the difference between the computed and potential-plane plots of delay shown on Figure 9, may be attributed to inaccuracy in the delay computation. Even ignoring these inaccuracies, the

maximum apparent error between the computed and potential plane delays is 750 milliseconds at a 6500-millisecond point, or 11.5%.

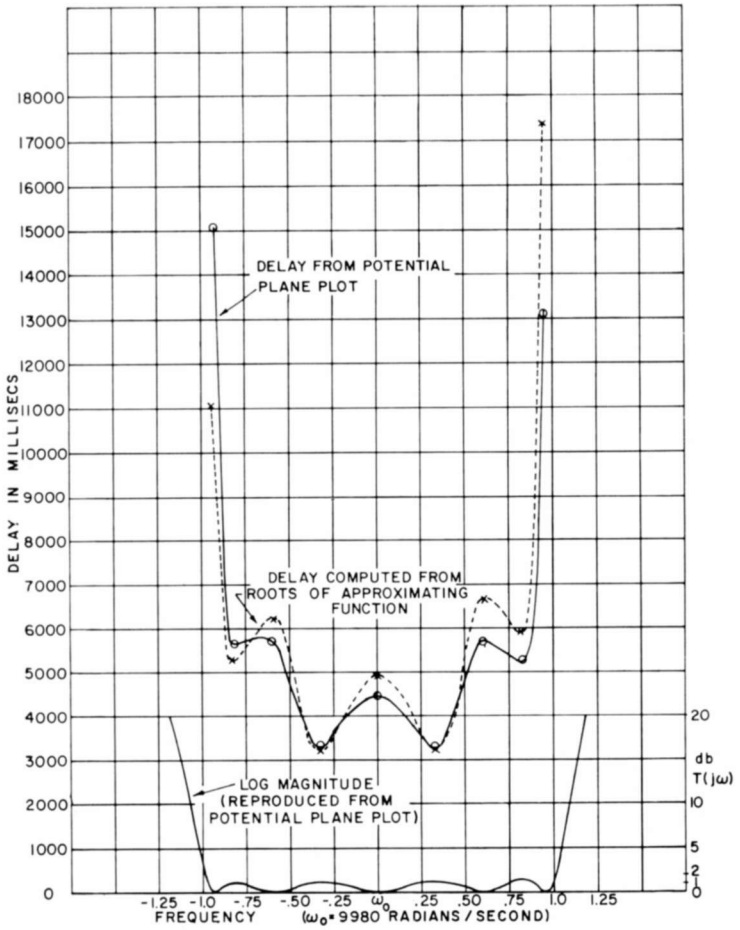


Fig. 9—Plotted and computed delay characteristics.

Characteristic Function Approximation

The characteristic network function can provide a useful method to synthesize doubly terminated lossless reactance networks, particularly filter networks where a low and controlled pass-band impedance mismatch or reflection coefficient is desired together with specific stop-band attenuation limits. From Figure 10, which shows an equally

terminated network, and known basic relations,[26,27] the characteristic function can be related to the transfer function by

$$|H(s)|^2 = 1 + |K(s)|^2, \tag{34}$$

where $H(s)$ = transmission function $= 1/T(s) = E(s)/P(s)$, $K(s)$ = characteristic function $= F(s)/P(s)$. Also,

$$|K(s)|^2 = |\rho(s)|^2 |H(s)|^2, \tag{35}$$

where $\rho(s)$ = reflection coefficient function $= A(s)/E(s)$. For equally

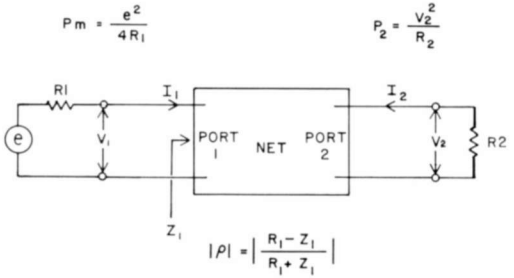


Fig. 10—Doubly terminated two-port reactance networks.

terminated reactance networks,

$$K(s) = \frac{A(s)}{E(s)}\frac{E(s)}{P(s)} = \frac{A(s)}{P(s)} = \frac{F(s)}{P(s)}. \tag{36}$$

$P(s)$ = zeros of the transfer function $T(s)$, all of which occur on the $j\omega$ axis for lossless LC networks. $A(s)$ = zeros of the reflection function $\rho(s)$; these can be arranged to occur on the $j\omega$ axis for lossless reactance networks resistively terminated.

Therefore, for certain network restrictions the characteristic function $K(s)$ is a rational-fraction network function whose numerator, $F(s)$, roots all occur on the $j\omega$ axis in the pass-band region and whose denominator, $P(s)$, roots also all occur on the $j\omega$ axis in the stop-band

[26] R. Saal and E. Ulbrich, "On the Design of Filters by Synthesis," *I.R.E. Trans. on Circuit Theory*, Vol. CT-5, p. 284, Dec. 1958.

[27] J. Zdunek, "The Network Synthesis on the Insertion-Loss Basis," *I.E.E. Monograph* No. 278R, Vol. 105, Pt. C, p. 259, Jan. 1958.

region at frequencies of infinite insertion loss. There are other mathematical restrictions placed on the nature of the $K(s)$ numerator and denominator polynomials, but for the purposes of the potential-plane analog, we may create approximations to required characteristic filter-function magnitude limits by suitably locating pole–zero electrodes all on the same $(j\omega)$ axis on the potential plane. These poles and zeros have only one degree of freedom, thus expediting the approximation procedure.

In a similar manner to that for transfer functions, the log-magnitude characteristic function may be expressed as a summation of the log-magnitude contributions from the numerator factors (zeros) and denominator factors (poles) of $K(s)$:

$$20 \log |K(j\omega)| = 20 \log C + 20 \sum_{i=1}^{n} \log (j\omega - j\omega_{0i}) - 20 \sum_{i=1}^{m} \log (j\omega - j\omega_{pi}). \quad (37)$$

The constant factor C can be ignored for approximation purposes. The log-magnitude plot of $K(s)$ along the $j\omega$ axis as indicated in Figure 11, provides a direct measure of the pass-band return loss and the stop-band attenuation.

Return loss is defined as

$$\text{return loss} = 20 \log \left| \frac{R_1 + Z_1}{R_1 - Z_1} \right| = 20 \log \left| \frac{1}{\rho} \right| \quad (38)$$

The return loss (or reflection coefficient) is a measure of the degree of input-impedance mismatch, and for lossless networks, a measure of the pass-band attenuation of the network. These relations are shown graphically in Figure 12.

Conjugate Roots

The roots of $K(s)$ are complex conjugates, and so for relatively wide-band functions (or low-pass functions) some practical method to compensate for the effects of conjugate roots must be found. Since the roots all lie on the same $(j\omega)$ axis in the s-plane, a simple conformal mapping into the W plane, where

$$W = s^2 = (j\omega)^2 = -\omega^2, \quad (39)$$

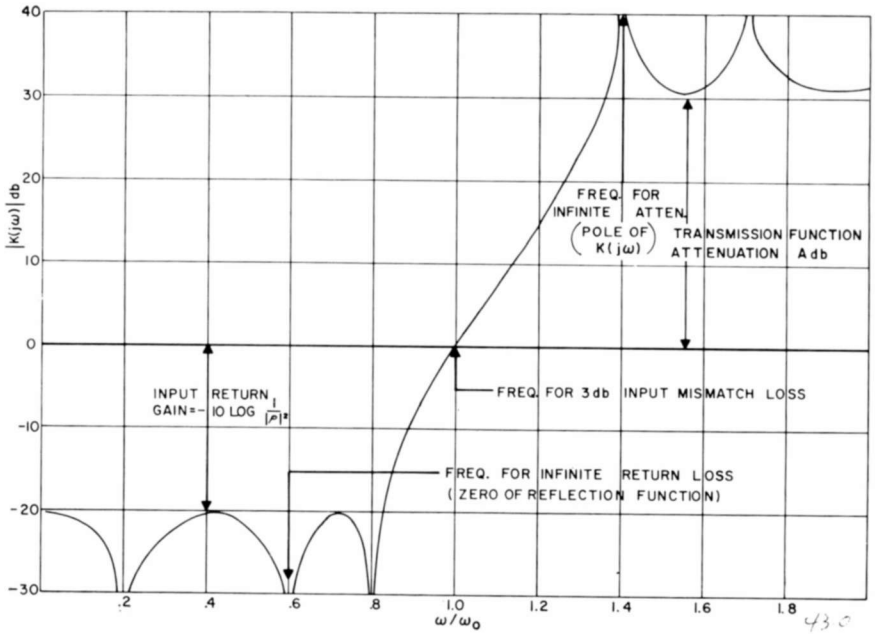


Fig. 11—Typical log magnitude characteristics of $K(j\omega)$.

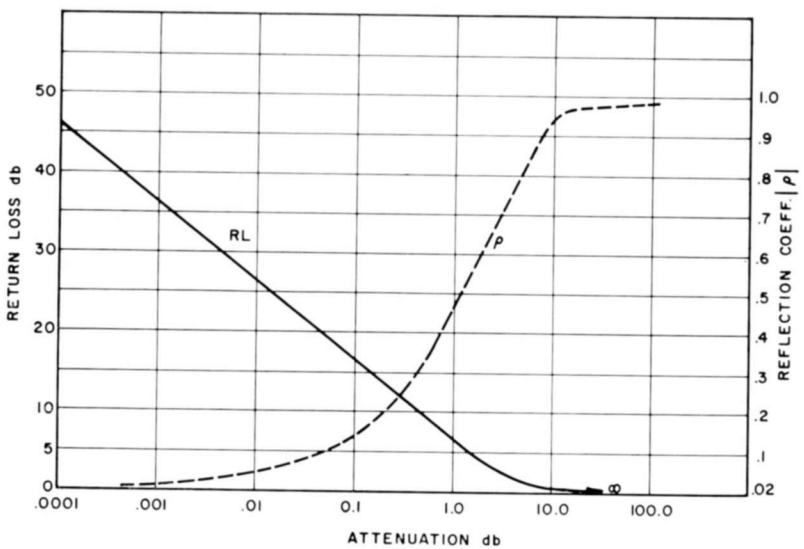


Fig. 12—Reflection coefficient, return loss, and attenuation (return loss (db) $= 20 \log |1/\rho|$, attenuation (db) $= -10 \log (1 - |\rho|^2)$).

results in double roots located all on the negative real u axis in the W plane where

$$W = u + jv. \tag{40}$$

Roots at the origin of the s-plane map onto the origin of the W plane as single roots.

Since the potential plane may represent any region of any analytic plane, an arbitrary axis (i.e., a diameter of the upper sheet) may be selected as the u axis of the W plane. Normally, double roots would be represented by doubling the pole–zero electrode currents. However, since the log-magnitude potential measured on the W-plane representation of the potential plane is a squared function, it will be doubled in log magnitude. Therefore, if the W-plane root electrodes are supplied with the standard current (and single roots at the origin supplied with half standard current), the potential measured along the u axis on the potential plane will represent the log magnitude of $K(j\omega)$ in the s-plane. The frequency scale is of course squared and the approximating limits must be located correctly on this squared frequency axis. For many network characteristic functions, squaring the frequency axis has the salutary effect of "spreading out" the pass-band frequency range and the pass-band roots.

Characteristic Function Approximation Examples

Again a wide choice of types of possible examples exists. A typical band-pass function approximation with a normalized pass-band width of approximately 15% is shown on Figure 13. This function has the following limits:

(a) Lower stop band $>$ 20 db to $\omega = 0.839$,

(b) Pass-band return loss $>$ 20 db (max. attenuation $<$ 0.043 db) for $0.859 \leq \omega \leq 0.991$,

(c) Upper stop band $\geq$ 100 db for $1.080 \leq \omega \leq 1.169$.

A total of six zeros and four poles in the W-plane were required. The $K(s)$ roots were located on the u-axis, recorded as described previously, and their $j\omega$-axis locations calculated by conversion from the squared frequency axis locations determined from the X-Y recorder plot.

As a check on the accuracy of approximation, the log-magnitude of the above $K(s)$ was computed at numerous test frequencies using a digital computer. The computed results are also shown in Figure 13. The maximum error in the stop band is approximately 2.2 db at a 100-db point (i.e., 2.2%), and in the pass band is a maximum of 0.95

db at a 21-db return loss point. For these check computations, the log-magnitudes of the potential-plane plot, and of the rational fraction $K(s)$ created, were taken as identical at the $\omega_0 = 1.00$ point.

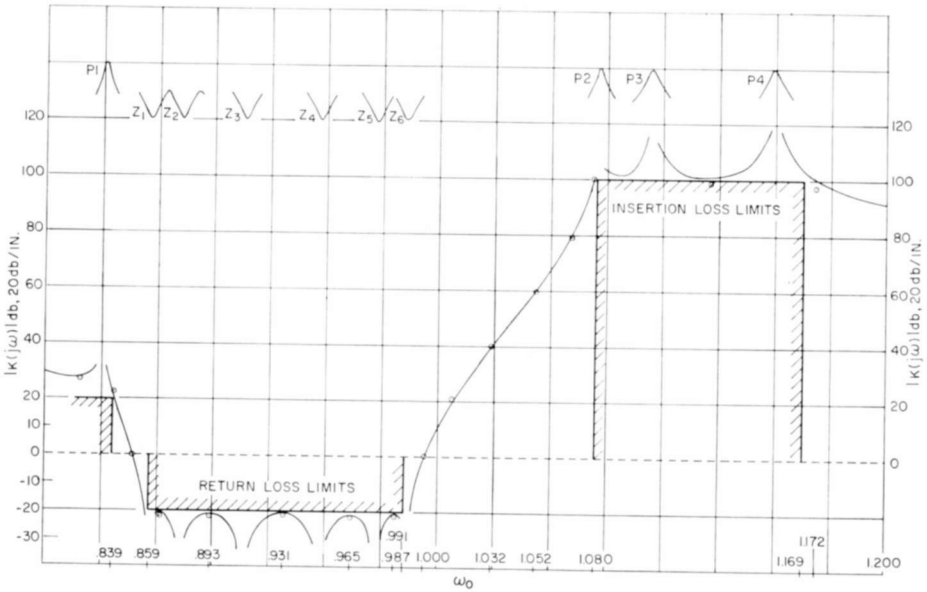


Fig. 13—Band-pass characteristic-function approximation plot.

Assessment of Errors in the Approximation

Imperfections in the system are manifested in two ways: (1) the plotted log-magnitude or delay responses do not coincide exactly with those that should result from the root configurations represented on the plane, and (2) the root locations as determined from measurements on the plane will include small errors.

A consequence of these limitations is that the responses corresponding to the determined root locations will be different from those plotted. Factors that may contribute to errors in the approximation are:

Nonuniformity of the paper resistivity,

Imperfections in the "infiniteness" of the plane,

The finite size of the root electrodes,

Inaccuracy in the setting of root electrode currents,

The finite size, and method of use, of the two measuring probes,

Errors in calibration,

Errors in determining the root locations.

Nonuniformity of the Paper Resistivity

No satisfactory relationship has been evolved to relate nonuniformity of paper resistivity to errors in approximation. The subject is complicated by the fact that the effects of nonuniformity are dependent on the characteristic being approximated. However, selection

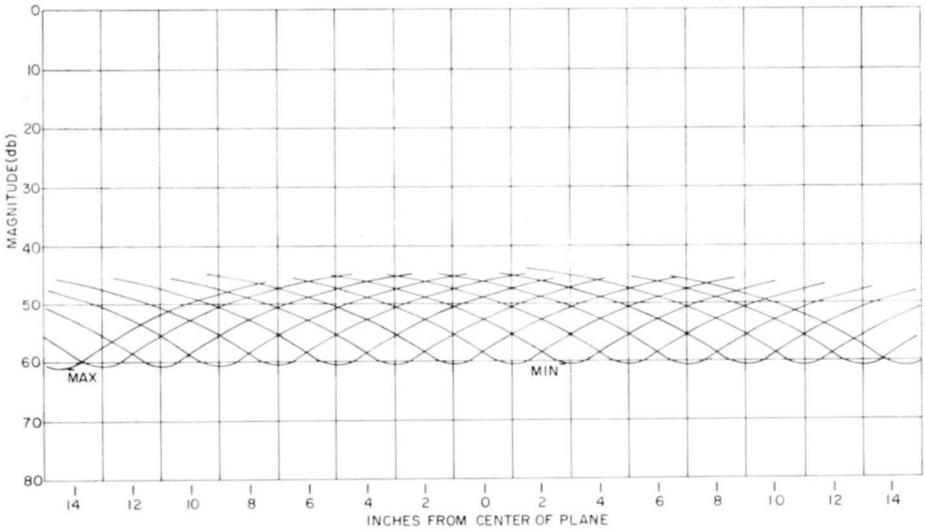


Fig. 14—Measurement of the "infiniteness" of the plane (max–min $\simeq 1\%$ potential magnitude $\simeq$ 61.1-60.5 db $\simeq$ 0.1 db/10 db).

of paper on the basis of nonuniformity measurements, as described earlier, should minimize errors due to this source.

Imperfections in the "Infiniteness" of the Plane

The accuracy of the "infinite" representation is so good that errors from this source are negligible. Figure 14, showing the peak log-magnitude responses from a row of root electrodes accurately positioned parallel to the $j\omega$ axis, and Figure 15, showing a plot of equipotential lines about a single root electrode positioned one inch from the rim of the plane, indicate the achievement of a very close approximation to a truly infinite plane.

The Finite Size of the Root Electrodes

It has been shown[20,28] that the distortion of the field, and hence errors in potential magnitude, due to the finite size of the root electrodes is less than 1% (i.e., 0.1 db) at a distance from the root of ten times the electrode radius. The radius of the root electrodes used in the

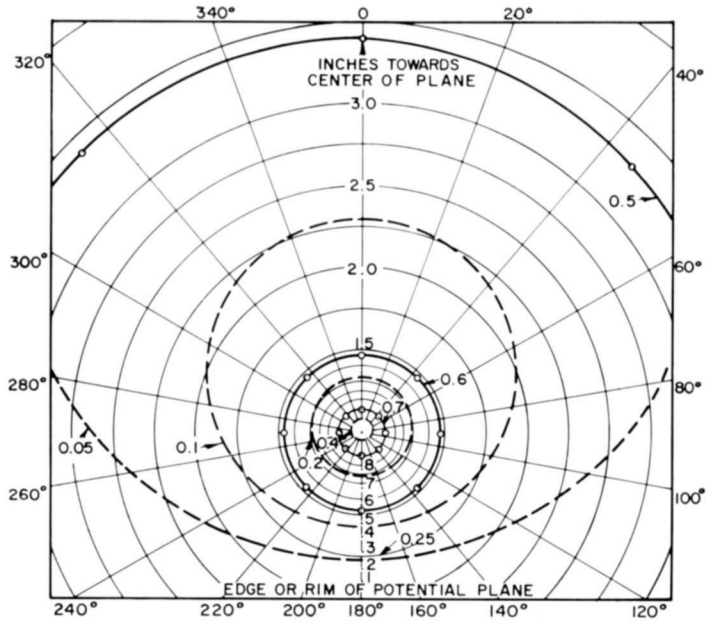


Fig. 15—Equipotential plot on the infinite plane. The solid lines represent measurements on a double-sided plane; the dashed lines were computed for a plane with a conducting rim; the figures indicate relative potential magnitudes.

present system is about 0.015 inch and the nearest a measuring probe can approach a root is about 0.12 inch. Thus, the maximum possible error from this source is about 0.1 db.

Inaccuracy in Setting the Root Electrode Currents

A fractional error in a single root-electrode current may be considered as the superposition of a fractional root of the same magnitude upon a root of unit magnitude. The potential-magnitude error along the $j\omega$ axis will have the same shape (but scaled) as the potential magnitude due to the unit root. The fractional root may represent a

[28] P. Moon and D. E. Spencer, *Field Theory for Engineers*, D. Van Nostrand Co., Inc., Princeton, N. J., 1961.

fractional pole or a fractional zero according to the sense of the current error. The potential-magnitude error is a maximum directly opposite the root and may be evaluated as follows.

The voltage drop V_d between an electrode radius r and a point P at a distance x from the electrode center is given by

$$V_d = \frac{IR_s}{0.8686\pi} \log \frac{x}{r}. \tag{41}$$

The potential, V, at point P then is given by

$$V = I\left[R_T - \frac{R_s}{0.8686\pi} \log \frac{x}{r}\right], \tag{42}$$

where R_T is the resistance of the plane. An error ΔV at P due to a root electrode current error of ΔI is given by

$$\Delta V = \Delta I\left[R_T - \frac{R_s}{0.8686\pi} \log \frac{x}{r}\right]. \tag{43}$$

ΔV will have a maximum value when $x = r$ of

$$\Delta V = \Delta I R_T. \tag{44}$$

For this system, $R_T = 3300$ ohms and $I = 0.9$ milliampere; taking $\Delta = 0.01$ (1%), then $\Delta V = 30$ millivolts, which corresponds to 1.2 db.

The root electrode currents are set by measuring the voltage drop across a resistor, in series with the electrode, with a precision differential voltmeter. The accuracy and resolution of this instrument are such that the currents can be set to $\pm 0.05\%$. Since the measuring probes can only approach within 0.12 inch of the root electrode, the maximum voltage and corresponding db errors are less than 150 microvolts, or 0.06 db.

The Finite Size and Method of Using the Two Measuring Probes

A measuring probe placed on the plane distorts the field in its immediate vicinity, but takes up the potential that existed beneath its center before it was placed on the plane.[20,28] Therefore, there is no error in the measurement of potential at this point.

When two measuring probes are used they are each situated in the distorted field due to the other so that each assumes a potential in error

by about 1%. However, because of the symmetry of the distortions, the potential mid-way between the probes remains unchanged.

For log-magnitude measurements using two probes, the method of measuring assumes a linear potential gradient between the probes. Because of the small distance between the probes this assumption will generally introduce negligible error.

For delay measurements, however, the errors in probe potentials add to give about 2% error in absolute delay measurement. Additional errors in absolute delay due to the finite spacing of the measuring probes will be small if all root electrodes are at least one inch from the $j\omega$ axis.

Errors in Calibration

Frequency Axis: When calibrating the frequency axis, the greatest pen deflection error noted is about 0.02 inch, which corresponds to 0.04 inch on the plane. With a plane frequency-scale factor of F cycles per second per inch this error represents $0.04F$ cps. The maximum inaccuracy due to this error occurs when the center of the plane represents zero frequency. Then the maximum frequency represented on the plans is $15F$ cps and, assuming the error to be linearly distributed, the frequency error is $0.02F/15F$ (i.e., approximately 0.13%). When the zero frequency point is far removed from the center of the plane, the percentage frequency error is much less.

Log-Magnitude: The db/inch calibration is based on the accurate location of the two calibration root electrodes relative to the measuring probes and visual resolution of setting. The dimensional accuracy is better than 0.01 inch and the error due to visual resolution is less than 0.02 inch. The calibration error will be less than 0.2 db/10 db.

Delay: The delay calibration is based on the same considerations as the log-magnitude calibration, but with additional inaccuracy introduced due to the "noisy" nature of the plotted response. However, an accuracy of better than 5% should be achievable.

Errors in Determining the Root Locations

$j\omega$ Location: With the method adopted, the limitation on accuracy is the visual resolution of the response peak. By recording with the measuring probe close to the root electrode, an accuracy of resolution of better than 0.02 inch is possible. As noted in the discussion of frequency-axis calibration errors, this corresponds to a maximum error of about 0.13% in $j\omega$ location.

σ Location: The resolution of the response peak is also a factor in σ location accuracy, and a maximum resolution error of 0.02 inch may

again be assumed. From Equation (28)

$$\log \sigma_x = \log \sigma_r - \left(\frac{D_x - D_r}{2} \right) \tag{45}$$

Assuming resolution errors in D_x and D_r, each of 0.02 inch and of opposite sign resulting in a computed error in σ_x of $\Delta\sigma_x$, then

$$\log \sigma_x(1 + \Delta) = \log \sigma_r - \left(\frac{D_x - D_r}{2} \right) \pm \frac{0.04}{2}. \tag{46}$$

Subtracting the above equations, $\log (1 + \Delta) = \pm 0.02$; from which $\Delta = \pm 0.047$, i.e., a maximum of 4.7% error in σ location. This can be halved by using the 5 db/inch scale instead of the 10 db/inch scale for plotting the peaks.

Summary of Error Contributions

Errors due to imperfections in the "infiniteness" of the plane are negligible, and for log-magnitude measurements, errors due to the finite size and method of use of the two measuring probes are also negligible. Errors due to the finite size of the root electrodes and inaccuracy in the setting of root-electrode currents are also negligible except in regions where root electrodes are extremely close to the $j\omega$ axis. Calibration errors that can be specified in percentage of absolute value may be minimized by taking a reference on the recorded plot in the region where greatest accuracy is required. This has the effect of translating the maximum errors to the regions where accuracy requirements are less. For example, by taking the relative attenuation reference for a filter in the pass band, the accuracy in the pass band will be high and the accuracy in the stop band will be lower.

Errors in determining root locations are shown to be small for $j\omega$ locations. It is seen, however, that great care must be exercised in locating the response peak for σ location determination in order to keep σ location errors low.

Errors due to nonuniformity of the paper resistivity are difficult to define. However, the method of determining root σ locations tends to reduce errors in the approximation due to nonuniformity. For instance, consider a root electrode in a region of relatively high resistivity. Its contribution to the $j\omega$-axis potential will be lower than it should be, so that it may be closely represented as a root further away from the $j\omega$ axis in a region of normal resistivity. Since the σ

location is determined from the $j\omega$-axis potential, computation based on this potential will indicate that the root is further from the $j\omega$ axis than it is physically. This is in the right direction to reduce the error. In effect, the electrical rather than the physical σ location is determined.

Conclusions

While the basic principles and applications of potential analog theory to electric network function approximation have been known for some time, the general use of the potential analog as an engineering design tool has been restricted. The infinite-potential-plane analog, using the double-sided dry conducting-sheet concept has been used to generate more than 50 filter transfer and characteristic functions during the past two years. Most of these functions presented unique and stringent requirements in db attenuation range and pass-band flatness as well as arbitrary limitations as to frequency symmetry, etc., and could only have been otherwise produced by a digital computer using extensive programs.

Work on digital-computer programs for general type filter-function approximation by nonlinear programming techniques has recently been presented.[29-31] To provide sufficient generality, flexibility, and range, such computer programs are extremely complex and require considerable time and effort to prepare. When programmed for a medium or high-speed computing system, they provide the obvious advantages of rapid problem solution, accuracy, and automatic operation. For very-high-order functions (e.g., greater than 20th order) the use of such digital computer programs appears mandatory to achieve the required accuracy in any reasonable time. For network functions of order lower than about the 20th, however, the infinite potential plane method has many attractive features.

As indicated in the assessment of errors, accuracies in representing approximating functions, of the order 2 to 3% in log-magnitude and 5 to 10% in delay, are achievable. Also, as noted, the magnitude of the errors can be minimized in specific frequency regions at the expense of accuracy in other regions.

---

[29] C. L. Semmelman, "Experience with a Steepest Descent Computer Program for Designing Delay Networks," *I.R.E. Convention Record*, Pt. 2, p. 206, 1962.

[30] T. Fujisawa, "Optimization of Low-Pass Attenuation Characteristics by a Digital Computer," *Proc. Sixth Midwest Symposium on Circuit Theory*, Wisconsin, May 1963.

[31] B. R. Smith and G. C. Temes, "An Iterative Approximation Procedure for Automatic Filter Synthesis," *I.E.E.E. Convention Record*, Pt. I, p. 270, March 1964.

The display and auxiliary equipments required are all standard laboratory items of reasonable cost. The fabrication of the plane requires only standard machine shop work, and the Teledeltos paper is relatively inexpensive.

While the time required to generate a network function depends primarily on the order of the function required to meet the approximation specifications, the experience and ingenuity of the operator is a contributing factor. For example, the characteristic band-pass function shown on Figure 13 required approximately 3 hours to complete, including calibration of the plane and determination of the final root locations.

Use of the plane provides a powerful insight into the nature and behavior of network functions (e.g., pole–zero sensitivity; conformal transformations; relations of attenuation, phase, and delay).

The infinite plane may also be used in other areas of two dimensional field mapping. Some interesting subjects are polynomial root solving,[32] root-locus plotting (with a traveling measuring probe), time-domain approximation using pole residues, heat and fluid flow problems, and electric or magnetic field problems. In these and other applications of potential theory, the infinite potential plane can provide an excellent engineering design tool of reasonable accuracy and cost.

ACKNOWLEDGMENTS

The authors acknowledge with appreciation the many contributions to the mechanical details of this potential plane provided by R. A. Pasquarella, J. H. Peirce, and J. Brittain, as well as the many useful discussions with T. G. Marshall, and the encouragement of H. F. L. Cameron and N. E. Edwards.

[32] W. J. Karplus and W. A. Soroka, *Analog Methods, Computation and Simulation*, McGraw-Hill Book Co., 2nd Ed., 1959.

DIELECTRIC BEHAVIOR OF NONRIGID MOLECULES II. INTRAMOLECULAR INTERACTIONS AND DIELECTRIC RELAXATION

By

Francis K. Fong*

RCA Laboratories
Princeton, N. J.

Summary—*The concept of simultaneous relaxation mechanisms has been extended to a broader class of nonrigid aromatic molecules. Various types of energy barrier to intramolecular group rotations in these molecules are described. Distribution of relaxation times associated with intramolecular group rotations are discussed in terms of the intramolecular interactions that give rise to potential barrier to such rotations. Expressions involving such a distribution are derived for the real and imaginary parts of the dielectric constant. Conventional methods for the analyses of dielectric data in terms of two relaxation times, one associated with the over-all molecular rotation and the other with group rotation, are critically evaluated. Several current problems, including the dielectric behavior of p-dimethoxybenzene are discussed in some detail.*

Introduction

If the assumption is made that the transition from one orientation at time t to another orientation at time $t + dt$ occurs by a succession of many small jumps, the theory of polar molecules in a viscous medium can be developed by the methods used in describing Brownian motions.[1-3] The relaxation time will then be related to a frictional constant that appears in the basic equations of Brownian motion theory. In this approach, the dielectric relaxation time of a rigid, polar aromatic molecule in the presence of an applied field is related to the dimensions of the molecule, and can be written[4]

$$\tau_m = 4\pi\eta_m a_m{}^3/(kT), \tag{1}$$

* Present address, North American Aviation Science Center, Thousand Oaks, California 91360.

[1] S. Chandrasekhar, "Stochastic Problems in Physics and Astronomy," *Rev. Mod. Phys.*, Vol. 15, p. 1, Jan. 1943.

[2] M. C. Wang and G. E. Uhlenbeck, "On the Theory of the Brownian Motion II," *Rev. Mod. Phys.*, Vol. 17, p. 323, April-July 1945.

[3] W. Kauzmann, "Dielectric Relaxation as a Chemical Rate Process," *Rev. Mod. Phys.*, Vol. 14, p. 12, Jan. 1942.

[4] P. Debye, *Polar Molecules*, The Chemical Catalogue Company, 1929.

where the subscript m denotes the molecular rotational process, η_m is the coefficient of internal friction, and $4\pi a_m^3/3$ is the volume swept out by rotation of the spherical molecule having a radius a_m; k and T have their usual significance. The relaxation time of a rigid molecule can thus be estimated according to the size and shape of the molecule. Nonrigidity in a polar molecule often results in the lowering of the relaxation time from the value estimated for the rigid model similar in size and shape to the real molecule. The study of dielectric relaxation times of polar molecules thus provides a means of investigating the molecular structure and intramolecular interactions of such molecules.

The anomalously low relaxation time of diphenyl ether first reported by Fischer[5] has been discussed in Part I of this series,[6] where the concept of simultaneous relaxation processes has been shown to be of importance in the study of nonrigid polar molecules. Thus, for diphenyl ether-like molecules, the apparent relaxation time, τ_0, can be written according to the equation:

$$\frac{1}{\tau_0} = \frac{1}{\tau_m} + \sum_i \frac{1}{\tau_i}, \tag{2}$$

where τ_m is the relaxation time associated with the overall molecular rotation, and relaxation times τ_i are those associated with i simultaneous intramolecular relaxation processes. Equation (2) is suitable for nonrigid molecules in which the direction of the molecular moment bisects the molecule. For all other nonrigid molecules, overall molecular rotation will occur independently of the intramolecular mechanisms, since in these molecules, dissimilar orientations of the molecular dipole are effected by molecular rotation and rotations involving intramolecular motions.

Intramolecular group rotations in aromatic molecules are subject to potential-energy barriers which may be caused by resonance of the π electrons in these molecules, by steric repulsion of the molecular constituents neighboring the rotatable group, and by hydrogen bonding to neighboring groups. Examples of these are discussed in the following section.

---

[5] E. Fischer, "Untersuchung der Innermolekularen Beweglichkeit Von Dipolmolekülen in Flüssigkeiten Mittels der Dielektrischen Relaxation," *Z. Naturforschung*, Vol. 4a, p. 707, 1949.

[6] F. K. Fong, "Dielectric Behavior of Nonrigid Molecules. I. The Simultaneous Relaxation Mechanisms of Diphenyl Ethers and Analogous Compounds," *Jour. Chem. Phys.*, Vol. 40, p. 132, 1 Jan. 1964.

Energy Barriers to Intramolecular Group Rotations

In a phenomenological description of the dielectric relaxation time, the dielectric phenomenon has been likened to a unimolecular chemical reaction.[3] Following the arguments of the absolute rate theory,[7] the dielectric relaxation time may be written according to the equation

$$\tau = \frac{h}{kT} \exp\left\{-\frac{\Delta S^{\ddagger}}{R}\right\} \exp\left\{\frac{\Delta E^{\ddagger}}{RT}\right\}, \quad \Delta F^{\ddagger} = \Delta E^{\ddagger} - T\Delta S^{\ddagger}, \tag{3}$$

where $\Delta F^{\ddagger}$, $\Delta S^{\ddagger}$ and $\Delta E^{\ddagger}$ are the molar free energy, entropy, and energy increases, respectively, required to convert one mole of the "normal state" dipoles into one mole of the "activated state," and h, k, and T have their usual significance. Although such a concept has been extended to the interpretation of experimental data in obtaining various thermodynamic quantities, it is sometimes difficult to visualize the involvement of a "chemical" activated state in the physical process of rotation of individual molecules and groups. The relaxation time τ is perhaps more suitably represented by the more general and simpler expression,

$$\tau = A \exp\left\{\frac{E}{RT}\right\}, \tag{4}$$

where A is a frequency factor that is assumed to be constant in most cases. The quantity E is the energy barrier to dipolar reorientation. In the case of overall molecular rotation, E is related to η_m in Equation (1).

In rotation around the bond joining the group to its parent molecule, the group encounters an energy barrier due to viscous friction similar to that experienced by the overall molecular rotation. Also, there exists an additional barrier due to intramolecular interactions such as conjugation of the π-electrons in the group with those of the benzene nucleus, steric repulsion of molecular constituents neighboring the rotating group, and intramolecular hydrogen bonding. The dielectric behavior of *p*-diacetobenzene[8] and *p*-dimethoxybenzene[9] paral-

---

[7] H. Eyring, "The Activated Complex in Chemical Reactions," *Jour. Chem. Phys.*, Vol. 3, p. 107, Feb. 1935.

[8] F. K. Fong and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. L. The Dielectric Relaxation of Several Substituted Naphthalenes and 4-Acetyl-o-terphenyl," *Jour. Amer. Chem. Soc.*, Vol. 85, p. 548, Jan. 30, 1963.

[9] F. K. Fong and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. LI. The Molecular and Intramolecular Relaxations of Two Substituted Biphenyls, 1-Naphthol, 2,6-Dimethylphenol and p-Diacetobenzene," *Jour. Amer. Chem. Soc.*, Vol. 85, p. 1565, 5 June 1963.

lels that of benzophenone and diphenyl ether,[6,10] respectively. Whereas benzophenone and *p*-diacetobenzene relax predominantly by overall molecular rotation due to restrictions of intramolecular motions by resonance, diphenyl ether and *p*-dimethoxybenzene relax largely by intramolecular mechanisms due to negligibly low energy barriers to internal rotations. Similarly, resonance energy makes an important contribution to the energy barrier associated with the internal acetyl group rotation in 4-acetyl-*o*-terphenyl and 2-acetonaphthone.[11] Steric repulsion constituting an energy barrier to internal group rotation is evidenced in the dielectric behavior of 2,6-dimethylanisole and 3,5-dimethylanisole.[12] Whereas these two molecules are of approximately the same size, the apparent relaxation time of 3,5-dimethylanisole ($\tau_0 = 16 \times 10^{-12}$ sec) is only roughly two thirds of that of 2,6-dimethylanisole ($\tau_0 = 25 \times 10^{-12}$ sec). Apparently, the methoxy group rotation in the 3,5-dimethylanisole molecule is relatively free of any steric hindrance compared to that in the 2,6-dimethylanisole molecule. This interpretation finds support in the Stuart–Briegleb models of these molecules. The strong steric hindrance exerted on the acetyl group in 1-acetonaphthone by the adjacent hydrogen atom on the 8-position has been noted.[11] It has also been established[13] that intramolecular hydrogen bonding in molecules such as 2,6-dibromophenol may lengthen considerably the relaxation time of the hydroxy group rotation. Thus, whereas the OH group rotation is associated with a small relaxation time of 3.4×10^{-12} sec in 2,6-dimethylphenol[8] and 2,6-*t*-butyl-*p*-cresol,[14] it has a relaxation time of 12.7×10^{-12} sec in 2,6-dibromophenol,[13] where the OH group can be intramolecularly hydrogen bonded to one of the neighboring polar groups.

---

[10] E. L. Grubb and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XLII. Molecular and Group Rotation in Aromatic Methoxy and Ethoxy Compounds," *Jour. Amer. Chem. Soc.*, Vol. 83, p. 4873, 20 Dec. 1961.

[11] F. K. Fong, Ph.D. Dissertation, Princeton University, 1962.

[12] F. K. Fong and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. LVII. Interpretation of the Parameters for the Dielectric Relaxations of 2,6-Dimethylanisole and 3,5-Dimethylanisole and Dipole Moments of 17 Aromatic Compounds," *Jour. Chem. Phys.*, Vol. 40, p. 2404, 15 April 1964.

[13] A. A. Antony, F. K. Fong, and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. LVIII. The Dielectric Relaxations, Infrared Spectra, and Intramolecular Hydrogen Bonding of 2,6-Dichloro-p-nitroaniline and Four Substituted Phenols," *Jour. Phys. Chem.*, Vol. 68, p. 2035, Aug. 1964.

[14] F. K. Fong and C. P. Smyth (to be published).

Interpretation of the Parameters of the Dielectric Relaxations of Nonrigid Molecules

The one parameter provided by a single relaxation time often proves insufficient for the representation of the dielectric data of nonrigid molecules. For the majority of nonrigid molecules, the introduction of a second parameter describing a distribution around a most probable relaxation time is frequently necessary to provide adequate representation of the data. For molecules in which rotation of one polar group around a bond in the molecule or of two or more identical groups around identical bonds can occur, a great many data may usually be treated by the Cole–Cole representation[15] with distribution parameters α and the most probable relaxation times τ_0 according to the equation,

$$\epsilon^* = \epsilon' - i\epsilon'' = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 + (\omega\tau_0)^{1-\alpha}}, \tag{5}$$

where ϵ_0 and ϵ_∞ are static and optical dielectric constants, ω is the angular frequency, and $i = \sqrt{-1}$. Since dipole relaxation can occur by molecular and by intramolecular orientation, each mechanism having its own relaxation time, the two parameters given by these two relaxation times may also provide adequate representation of the dielectric dispersion data according to the equations

$$\frac{\epsilon' - \epsilon_\infty}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{1}{1 + (\omega\tau_1)^2} + c_2 \frac{1}{1 + (\omega\tau_2)^2}, \tag{6}$$

$$\frac{\epsilon''}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{\omega\tau_1}{1 + (\omega\tau_1)^2} + c_2 \frac{\omega\tau_2}{1 + (\omega\tau_2)^2}, \tag{7}$$

where $c_1 + c_2 = 1$, τ_1 is the relaxation time corresponding to the rotation of the molecule as a whole, and τ_2 is the relaxation time associated with the group, the significance of which is discussed in the next section. This frequency dependence of the dielectric constant and loss is predicted by the theory of Budo[16] for the dielectric relaxation of spherical molecules containing freely rotating polar groups,

---

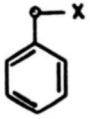
[15] K. S. Cole and R. H. Cole, "Dispersion and Absorption in Dielectrics. I. Alternating Current Characteristics," *Jour. Chem. Phys.*, Vol. 9, p. 341, April 1941.

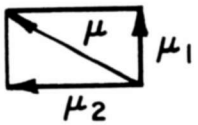
[16] A. Budo, *Physik Z.*, Vol. 39, p. 706, 1938.

and has been adapted[17-19] for the representation of dielectric dispersion data of aromatic compounds with rotatable groups in the liquid state, in benzene or some more viscous solvents.[9-14,17-24] Although effective in the illustration of the role of intramolecular relaxation, Equations (6) and (7) are only applicable for molecules with freely rotating groups, and are unsuitable for molecules in which intramolecular interactions give rise to hindered group rotations. An attempt is made in the following sections to derive more general expressions for the representation of overall and intramolecular relaxations of nonrigid aromatic molecules.

The Parameter τ_2 in Equations (6) and (7)

The special case of a molecule with one freely rotating group will now be discussed. Consider a nonpolar medium such as benzene, containing N_0 nonrigid molecules of the type





[17] K. Bergmann, D. M. Roberti, and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XXXI. Analysis in Terms of Two Relaxation Times for Some Aromatic Ethers," *Jour. Phys. Chem.*, Vol. 64, p. 665, May 1960.

[18] F. K. Fong and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XLIX. A Method of Analyzing Simultaneous Relaxation Processes and Its Application to 2-Naphthol and 2-Naphthalenethiol, *Jour. Phys. Chem.*, Vol. 67, p. 226, Feb. 1963.

[19] W. E. Vaughan, S. B. W. Roeder, and T. Provder, "Dielectric Relaxation of Some Aromatic Molecules Containing Methoxy Groups," *Jour. Chem. Phys.*, Vol. 39, p. 701, 1 Aug. 1963.

[20] W. P. Purcell and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XXXIX. The Relaxation Times of Some Substituted Phenyl Ethers and Analogous Molecules," *Jour. Amer. Chem. Soc.*, Vol. 83, p. 1063, 5 Mar. 1961.

[21] E. N. DiCarlo and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XLVIII. The Dielectric Relaxation of Diphenyl Sulfide, Triphenylamine and Diphenylmethane," *Jour. Amer. Chem. Soc.*, Vol. 84, p. 3638, 5 Oct. 1962.

[22] W. E. Vaughan, W. S. Lovell, and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XLIV. Investigation of Dielectric Relaxation by an Interferometric Method for the Measurement of Dielectric Constant and Loss at 2.2-mm Wavelength," *Jour. Chem. Phys.*, Vol. 36, p. 535, 15 Jan. 1962.

[23] W. E. Vaughan and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XXXV. Absorption by Pure Polar Liquids at 4.3 MM. Wave Length," *Jour. Phys. Chem.*, Vol. 65, p. 98, Jan. 1961.

[24] D. M. Roberti, O. F. Kalman, and C. P. Smyth, "Microwave Absorption and Molecular Structure in Liquids. XXX. The Anomalous Dielectric Relaxation of Diphenyl Ether and Some Similar Molecules," *Jour. Amer. Chem. Soc.*, Vol. 82, p. 3523, 20 July 1960.

where μ is the overall molecular dipole moment; μ_2, associated with the rotating OX group, changes its direction with the intramolecular rotation; and μ_1 is associated with overall molecular rotation around an axis parallel to the direction of μ_2. Let each of the dipoles, μ_1 or μ_2, point in a direction having a polar angle ϕ_1 or ϕ_2 and longitude ζ_1 or ζ_2, respectively, with respect to some fixed set of axes. The distribution functions can be written as

$$N_{\Omega_1} = N(\phi_1,\zeta_1,t)\,d\Omega,$$

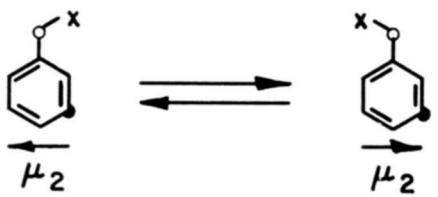
and (8)

$$N_{\Omega_2} = N(\phi_2,\zeta_2,t)\,d\Omega,$$

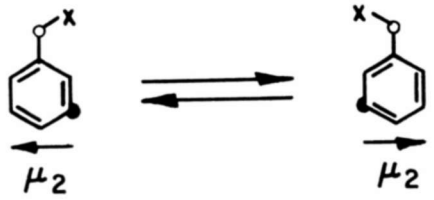
for dipoles μ_1 and μ_2 at the various directions in space at any given time, where N_{Ω_1} and N_{Ω_2} are the numbers of dipoles in 1 cc whose directions are included in an element of solid angle $d\Omega$ around the directions ϕ_1, ζ_1 and ϕ_2, ζ_2, respectively. The polarization along the direction $\phi = 0$ contributed by the two types of dipoles at a given time will be

$$P(t) = P_1(t) + P_2(t) = \mu_1 \int N_{\Omega_1} \cos\phi_1 d\Omega + \mu_2 \int N_{\Omega_2} \cos\phi_2 d\Omega. \quad (9)$$

Let $(\tau'_m)^{-1}(\phi_1,\zeta_1 \rightarrow \phi'_1,\zeta'_1)\,dtd\Omega'$ be the probability that the molecular moment μ_1 will rotate from a position ϕ_1, ζ_1, to a new orientation, ϕ'_1, ζ'_1, within a solid angle $d\Omega'$ in an interval of time dt by overall molecular rotation, and let $(\tau'_m)^{-1}(\phi_2,\zeta_2 \rightarrow \phi'_2,\zeta'_2)\,dtd\Omega'$ and $(\tau'_1)^{-1}(\phi_2,\zeta_2 \rightarrow \phi'_2,\zeta'_2)\,dtd\Omega'$ be the probabilities that the moment μ_2 will rotate from a position ϕ_2, ζ_2, to a new orientation, ϕ'_2, ζ'_2, within a solid angle $d\Omega'$ in an interval of time dt by overall molecular rotation and intramolecular group rotation, respectively. The orientation of the moment μ_2 is effected by an intramolecular rotation as follows:



However, the direction of μ_2 can also be changed by an overall molecular rotation as follows:



the *meta* hydrogen having been marked to indicate the relative positions of the benzene nucleus. Since there is no cause for the molecule to remain rigid while it is rotating as a whole, or for overall molecular rotation to cease while the molecule is rotating intramolecularly, both molecular and intramolecular mechanisms can and must occur simultaneously in the orientation of μ_2. It is therefore necessary to consider both $(\tau'_m)^{-1}(\phi_2,\zeta_2\rightarrow\phi'_2,\zeta'_2)\,dtd\Omega'$ and $(\tau'_1)^{-1}(\phi_2,\zeta_2\rightarrow\phi'_2,\zeta'_2)\,dtd\Omega'$ when considering the rotation of the dipole μ_2. Following the arguments of the previous work,[6] it can be easily shown that

$$\begin{aligned}
\frac{dP(t)}{dt} &= \frac{dP_1(t)}{dt} + \frac{dP_2(t)}{dt} \\
&= \left[\frac{dP_1(t)}{dt}\right]_m + \left[\frac{dP_2(t)}{dt}\right]_m + \left[\frac{dP_2(t)}{dt}\right]_i \\
&= \mu_1 \int \left[\frac{dN_{\Omega_1}}{dt}\right]_m \cos\phi + \mu_2 \left\{ \int \left[\frac{dN_{\Omega_2}}{dt}\right]_m \cos\phi + \int \left[\frac{dN_{\Omega_2}}{dt}\right]_i \cos\phi \right\} \\
&= \mu_1 \int\int [N(\phi'_1,\zeta'_1,t)\,(\tau'_m)^{-1}(\phi'_1,\zeta'_1\rightarrow\phi_1,\zeta_1) \\
&\quad - N(\phi_1,\zeta_1,t)\,(\tau'_m)^{-1}(\phi_1,\zeta_1\rightarrow\phi'_1,\zeta'_1)]\,d\Omega d\Omega' \cos\phi_1 \\
&\quad + \mu_2 \left\{ \int\int [N(\phi'_2,\zeta'_2,t)\,(\tau'_m)^{-1}(\phi'_2,\zeta'_2\rightarrow\phi_2,\zeta_2) \right. \\
&\quad - N(\phi_2,\zeta_2,t)\,(\tau'_m)^{-1}(\phi_2,\zeta_2\rightarrow\phi'_2,\zeta'_2)]\,d\Omega d\Omega' \cos\phi_2 \\
&\quad \left. + \int\int [N(\phi'_2,\zeta'_2,t)\,(\tau'_i)^{-1}(\phi'_2,\zeta'_2\rightarrow\phi_2,\zeta_2) \right. \\
&\quad \left. - N(\phi_2,\zeta_2,t)\,(\tau'_i)^{-1}(\phi_2,\zeta_2\rightarrow\phi'_2,\zeta'_2)]\,d\Omega d\Omega' \cos\phi_2 \right\}. \qquad (10)
\end{aligned}$$

This equation gives the rate of change of the macroscopic polarization in terms of the overall molecular and intramolecular motions of the individual molecules. Its complexity can be drastically reduced if it is assumed that both $(\tau'_m)^{-1}$ and $(\tau'_i)^{-1}$ are constants.[3] Under this condition and noting that the net amount of polarization re-established by jumps of dipoles from other orientations must be zero due to the

more or less symmetrical arrangement of possible directions in the absence of an external field, Equation (10) becomes

$$\frac{dP(t)}{dt} = -4\pi P_1(t)\,(\tau'_m)^{-1} - 4\pi P_2(t)\,[\,(\tau'_m)^{-1} + (\tau'_i)^{-1}]$$
$$= -P_1(t)\,\tau_1^{-1} - P_2(t)\,(\tau_1^{-1} + \tau_i^{-1}), \tag{11}$$

where $\tau_1 = 4\pi\tau'_m$ and $\tau_i = 4\pi\tau'_i$. Solving the appropriate differential equations,

$$P(t) = P_1(t) + P_2(t) = P_{01}\exp\left\{-\frac{t}{\tau_1}\right\} + P_{02}\exp\left\{-\left(\frac{t}{\tau_1} - \frac{t}{\tau_2}\right)\right\}$$
$$= P_{01}\exp\left\{-\frac{t}{\tau_1}\right\} + P_{02}\exp\left\{-\frac{t}{\tau_2}\right\}, \tag{12}$$

where

$$\tau_2^{-1} = \tau_1^{-1} + \tau_i^{-1},$$

or

$$\tau_2 = \frac{\tau_1\tau_i}{\tau_1 + \tau_i}. \tag{13}$$

At $t = 0$, $P_0 = P_{01} + P_{02}$, and since $P_{01}/P_0 = \mu^2{}_1/\mu^2$ and $P_{02}/P_0 = \mu^2{}_2/\mu^2$,

$$P_0 = \frac{(\mu_1{}^2 + \mu_2{}^2)P_0}{\mu^2} = (c_1 + c_2)P_0 \tag{14}$$

where

$$\frac{c_1}{c_2} = \left(\frac{\mu_1}{\mu_2}\right)^2, \quad c_1 + c_2 = 1. \tag{15}$$

Equation (12) may now be rewritten as

$$P(t) = \left(c_1\exp\left\{-\frac{t}{\tau_2}\right\} + c_2\exp\left\{-\frac{t}{\tau_1}\right\}\right)P_0. \tag{16}$$

It is evident from Equation (13) that τ_2, contrary to the conventional interpretation in the literature, is not the group relaxation time. Rather, it is a function of the group relaxation time τ_i and the overall

molecular relaxation time τ_m. A significant conclusion from this result is that whereas the group relaxation time τ_i may be very large due to high energy barriers, τ_2 may in no case exceed the value for the overall molecular relaxation τ_1. This simple concept has been useful in the analysis of the dispersion data of 2,6-dimethylanisole in which the two methyl groups at the 2,6-positions hinder rotation of the methoxy group so much, that intramolecular rotation is limited to a librational motion.[12]

Hindered Rotation and the Distribution of τ_i

For free group rotations, the intramolecular group relaxation time τ_i may be approximated as that associated with rotation of a free molecule, and Equation (1) may be employed in the evaluation of its magnitude. The energy barrier to such a rotation can be related to η_m, the coefficient of internal friction. For a great majority of aromatic molecules, however, group rotations are not free due to intramolecular interactions such as those discussed in the second section; therefore, τ_i written in the form of Equation (4) must include an energy term due to such intramolecular interactions. Thus,

$$\tau_i = A \exp\left\{\frac{E}{RT}\right\}$$

$$= A \exp\left\{\frac{E_\eta + E_i}{RT}\right\} = \tau_\eta \exp\left\{\frac{E_i}{RT}\right\}, \tag{17}$$

where $\tau_\eta = A \exp\{E_\eta/RT\}$ is the group relaxation time in the absence of intramolecular interactions, and E_i is the contribution to the energy barrier by intramolecular interactions.

Assuming that the number of groups N is evenly distributed throughout the height E_0 of the potential barrier,

$$dN = N_0 \frac{dE_i}{E_0} \tag{18}$$

is the fraction of groups with values in a range dE_i, near E_i, if N_0 is the total number of groups per unit volume. From Equation (17), it is evident that the individual relaxation time τ_i covers the range

$$\tau_\eta \leq \tau_i \leq \tau_E, \tag{19}$$

where $\tau_E = \tau_\eta \exp\left\{\dfrac{E_0}{RT}\right\}$.

In the special case when $\tau_1 \gg \tau_E \geqq \tau_i \geqq \tau_\eta$, $\tau_2 = \tau_i$ according to Equation (13), and the distribution of τ_2 is approximately the same as that of τ_i. Assuming such a distribution, $f(\tau_2) = f(\tau_i)$, Equations (6) and (7) are now rewritten as

$$\frac{\epsilon' - \epsilon_\infty}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{1}{1 + (\omega\tau_1)^2} + c_2 \int_{\tau_\eta}^{\tau_E} \frac{f(\tau_2)}{1 + (\omega\tau_2)^2}\, d\tau_2, \tag{20}$$

$$\frac{\epsilon''}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{\omega\tau_1}{1 + (\omega\tau_1)^2} + c_2 \int_{\tau_\eta}^{\tau_E} \frac{\omega\tau_2 f(\tau_2)}{1 + (\omega\tau_2)^2}\, d\tau_2, \tag{21}$$

where it can be readily shown[25] that

$$f(\tau_2) = \frac{RT}{E_0}\frac{1}{\tau_2}, \qquad \text{if} \quad \tau_\eta \leqq \tau_i \leqq \tau_E = \tau_\eta \exp\left\{\frac{E_0}{RT}\right\},$$

$$f(\tau_2) = 0, \qquad \text{if} \quad \tau_E < \tau_i < \tau_\eta, \tag{22}$$

provided that Equations (17) and (18) are assumed and $\tau_1 \gg \tau_E$. Integrations of Equations (20) and (21) lead immediately to the equations

$$\frac{\epsilon' - \epsilon_0}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{1}{1 + (\omega\tau_1)^2} + c_2 \left(1 + \frac{RT}{2E_0} \log \frac{1 + \omega^2\tau_\eta^2 \exp\left\{\dfrac{2E_0}{RT}\right\}}{1 + \omega^2\tau_\eta^2}\right), \tag{23}$$

$$\frac{\epsilon''}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{\omega\tau_1}{1 + (\omega\tau_1)^2} + c_2 \frac{RT}{E_0}\left[\tan^{-1}\left(\omega\tau_\eta \exp\left\{\frac{E_0}{RT}\right\}\right) - \tan^{-1}\omega\tau_\eta\right], \tag{24}$$

[25] H. Frohlich, *Theory of Dielectrics, Clarendon Press*, Oxford, p. 91, 1949.

the terms containing c_2 being identical with the corresponding equations derived by Frohlich.[25]

The condition that $\tau_1 \gg \tau_E \geqq \tau_i \geqq \tau_\eta$, which implies that $\tau_2 \cong \tau_i$, is seldom realistic for most aromatic molecules with rotatable groups. Indeed, for this condition to exist, two essentially separate dispersion regions are to be expected. With only a few exceptions,[26] however, the two dispersion regions of a great many nonrigid aromatic molecules measured have been found to be overlapping, and τ_2 must therefore be expressed according to Equation (13). In this case, the distribution of τ_2 is still limited by two values according to the following:

$$\tau_{2\eta} \leqq \tau_2 \leqq \tau_{2E}, \tag{25}$$

where

$$\tau_{2\eta} = \tau_1\tau_\eta(\tau_1 + \tau_\eta)^{-1}, \tag{26}$$

$$\tau_2 = \tau_1\tau_\eta \exp\left\{\frac{E_i}{RT}\right\}\left(\tau_1 + \tau_\eta \exp\left\{\frac{E_i}{RT}\right\}\right)^{-1}. \tag{27}$$

and

$$\tau_{2E} = \tau_1\tau_\eta \exp\left\{\frac{E_0}{RT}\right\}\left(\tau_1 + \tau_\eta \exp\left\{\frac{E_0}{RT}\right\}\right)^{-1}, \tag{28}$$

Also, the assumption that the number of groups is evenly distributed throughout the height E_0 of the intramolecular barrier [Equation (18)] may not be justified in real systems in which a Maxwell–Boltzmann distribution is probably more appropriate. Equation (18) is thus replaced by

$$dN = c \exp\left\{-\frac{E'_i}{RT}\right\} dE'_i, \tag{29}$$

where $E'_i = E_0 - E_i$, and the constant c may be evaluated by integrating the equation over all values of E'_i from 0 to E_0. Hence

$$N_0 = cRT\left[1 - \exp\left\{-\frac{E_0}{RT}\right\}\right]$$

or

[26] M. Davies and R. J. Meakins, "Multiple Relaxation Times in Nonrigid Solute Molecules," *Jour. Chem. Phys.*, Vol. 26, p. 1584, June 1957.

$$c = \frac{N_0}{RT\left(1 - \exp\left\{-\frac{E_0}{RT}\right\}\right)}. \tag{30}$$

Since $f(\tau_2)\,d\tau_2$ is the probability of the assembly of groups having individual relaxation times in a range $d\tau_2$ near τ_2,

$$\int_{\tau_{2\eta}}^{\tau_{2E}} f(\tau_2)\,d\tau_2 = 1. \tag{31}$$

Expressing τ_2 as a function of E_i and using Equations (27) and (29), we arrive at the equation

$$\begin{aligned}\frac{1}{N_0}\int_0^{E_0} c\exp\left\{\frac{-(E_0 - E_i)}{RT}\right\} dE_i &= \int_0^{E_0} f(\tau_2)\,\frac{d\tau_2}{dE_i}\,dE_i \\ &= \frac{1}{RT\tau_\eta}\int_0^{E_0} f(\tau_2)\,\frac{\tau_2^2}{\exp\left\{\frac{E_i}{RT}\right\}}\,dE_i.\end{aligned} \tag{32}$$

It follows, therefore, that

$$f(\tau_2)\,\frac{\tau_2^2}{RT\tau_\eta \exp\left\{\frac{E_i}{RT}\right\}}\,dE_i = \frac{1}{N_0}\,c\exp\left\{\frac{-(E_0 - E_i)}{RT}\right\} dE_i, \tag{33}$$

or, remembering Equation (30),

$$\begin{aligned} f(\tau_2) &= \left(\tau_\eta \exp\left\{\frac{E_0}{RT}\right\} - \tau_\eta\right)^{-1} (1 - \tau_2/\tau_1)^{-2} = (\tau_E - \tau_\eta)^{-1}(1 - \tau_2/\tau_1)^{-2} \\ &\qquad \text{if} \quad \tau_{2\eta} \leq \tau_2 \leq \tau_{2E}, \\ f(\tau_2) &= 0, \qquad \text{if} \quad \tau_{2E} < \tau_2 < \tau_{2\eta}. \end{aligned} \tag{34}$$

The dielectric constants ϵ' and ϵ'' of the nonrigid molecule can now be expressed by the general equations obtained by the insertion of $f(\tau_2)$

from Equation (34) into Equations (20) and (21);

$$\frac{\epsilon' - \epsilon_\infty}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{1}{1 + (\omega\tau_1)^2} + \frac{c_2}{K} \int_{\tau_{2\eta}}^{\tau_{2E}} \frac{1}{(1 - \tau_2/\tau_1)[1 + (\omega\tau_2)^2]^2} d\tau_2, \quad (35)$$

$$\frac{\epsilon''}{\epsilon_0 - \epsilon_\infty} = c_1 \frac{\omega\tau_1}{1 + (\omega\tau_1)^2} + \frac{c_2}{K} \int_{\tau_{2\eta}}^{\tau_{2E}} \frac{\omega\tau_2}{(1 - \tau_2/\tau_1)^2[1 + (\omega\tau_2)^2]^2} d\tau_2, \quad (36)$$

where $K = \tau_E - \tau_\eta = \tau_\eta(\exp\{E_0/RT\} - 1)$ and $\tau_{2\eta}$ and τ_{2E} are given by Equations (26) and (28), respectively. The validity of Equations (35) and (36) rests upon the exactness of the distribution assumed in Equation (29), and the above derivation represents a general approach to the treatment of dielectric data of nonrigid molecules in which energy barrier to internal rotation has non-zero values. The real and imaginary parts of the dielectric constant, ϵ' and ϵ'', considered as functions of frequency, depend on four independent parameters, namely, $c_1 = 1 - c_2$, the overall molecular relaxation time τ_1, the re-relaxation time $\tau_{2\eta}$ as defined by Equation (26), and the factor E_0/kT, which determines the extent of the range of intramolecular relaxation times τ_2 according to Equation (25). In the event that $E_0/kT = 0$, Equations (6) and (7) are obtained. Equations (6) and (7), therefore, contain three independent parameters only. In the event that $E_0/kT = \infty$, the Debye formulas for rigid spherical molecules containing only one parameter are obtained.

Some Practical Considerations

The four unknown parameters τ_1, τ_2, c_2 and ϵ_∞ in Equations (6) and (7) have been evaluated by the chord method,[17] the double-arc method,[18] and by fitting the data to Equations (6) and (7) using a computer.[19] The interpretation of the dielectric data of nonrigid molecules by these methods without proper consideration of the intramolecular potential barrier, the corresponding distribution of τ_2, and the plausibility of separation of molecular and intramolecular relaxation times has led to analytical difficulties. It has been shown that the dielectric data of 2,6-dimethoxyanisole can be analyzed in terms of two relaxation processes by the double-arc method.[12] The parameter τ_2

has the surprisingly large value of 40×10^{-12} sec at 20° in benzene solution, whereas the parameter τ_1 associated with the overall molecular relaxation process is only 14×10^{-12} sec under the same conditions. Since it is not possible for the value of τ_2 to exceed that of overall molecular rotation according to Equation (13), it becomes evident that the mere adequate representation of the data in terms of two parameters does not establish the existence of intramolecular group rotation with a relaxation time corresponding to one of the parameters.

The very short relaxation times observed for *p*-dimethoxybenzene indicate that the molecule relaxes predominantly by an intramolecular process.[19] This is to be expected[9] since the permanent components of the dipole moment along the C-O bonds cancel each other. The observed relaxation time, τ_0, however, is large compared to the values of τ_2 for the other compounds with the methoxy group.[9,19] This large value has been explained by Grubb and Smyth[9] as being due to a fluctuation of the mesomeric moment as the group rotates, in which case a small contribution to the orientational polarization could be made by the overall end-over-end rotation of the molecule. According to the analysis of Vaughan, Roeder, and Provder, however, a decomposition of the data into a superposition of terms of the Debye type by the employment of Equations (6) and (7) shows that the additional dispersion is at the higher rather than lower frequency.[19] In addition to the questionable "large group relaxation time," 8.5×10^{-12} sec, a suspiciously small relaxation time, 0.7×10^{-12} sec, of unknown origin is obtained.[19] In view of these difficulties, treatment of the dielectric data of *p*-dimethoxybenzene in terms of τ_1 and τ_2 according to Equations (6) and (7) appears to be inappropriate, since only if a plausible physical mechanism can be associated with each of the two relaxation times is it justifiable to represent the data in terms of such relaxation times.

The fluctuation of the mesomeric moment accompanying the rotation of the group as a possible cause for the "large" relaxation time of *p*-dimethoxybenzene is questionable, since the effect of the fluctuating mesomeric moment accompanying the rotation of one group, if it exists, should be equal to and, therefore, cancelled by that of the other group. Since *p*-dimethoxybenzene resembles diphenyl ether[6] in that the direction of the molecular moment bisects the molecule, it is plausible to discuss the mechanism for the internal rotations of this molecule in terms of coupled group rotations not unlike those described for diphenyl etherlike molecules.[6] Rotation of the two methoxy groups with the benzene nucleus stationary is not acceptable, since such a mechanism involves the displacement of the center of gravity of the

molecule. Thus, the simultaneous rotation of the methoxy groups must necessarily be coupled by suitable motions of the benzene nucleus so that the center of gravity of the molecule may remain stationary during the rotations. Since the volume swept out by such a rotation is considerably larger than that by the rotation of a single methoxy group, the apparent relaxation time of *p*-dimethoxybenzene should be somewhat larger than the relaxation time associated with the intramolecular mechanism involving the rotation of a single methoxy group. While this coupled internal rotation predominates when the long axis of the molecule is perpendicular to the direction of the applied field, end-over-end molecular rotation can result if this axis is parallel to the direction of the field. Such a situation may be described according to the Perrin theory[27] for an ellipsoidal molecule, and the dielectric data may conceivably be represented by a distribution of relaxation times with a "most probable relaxation time," and with the relaxation time associated with overall molecular rotation and the relaxation time associated with the intramolecular process described above as the two limiting values. Of course, the presence of any potential barrier to intramolecular rotation may also contribute to the distribution of relaxation times in the manner discussed in the preceding section, although such a barrier may be comparatively unimportant in molecules such as diphenyl ether and *p*-dimethoxybenzene.[6]

Another difficulty of the analyses based on Equations (6) and (7) may be illustrated in the evaluation of c_1 or c_2. Whereas it is well established that $c_2/c_1 = (\mu_2/\mu_1)^2$ according to Equation (15) for nonrigid molecules in which group rotations are unhindered, this relation very often does not hold even qualitatively. c_2 frequently appears to be smaller than expected in molecules where group rotation is hindered by π-electronic resonance,[8,11] or by weak hydrogen bonding to π-electron clouds of the solvent benzene molecules.[8,18] In the 2,6-dimethylphenol molecule,[8] the OH group is partially screened from the benzene molecules by the neighboring methyl groups, and larger contribution to the total loss by OH group rotation is made possible. The value 0.83 for c_2 in the case of 2,6-dimethylphenol is more than three times those for *p*-phenylphenol,[8] 1-naphthol,[8] and 2-naphthol,[18] for example. From the heights of the two loss maxima obtained by Davies and Meakins[26] for 2,4,6-tri-*t*-butylphenol, c_2/c_1 is estimated to be 8.2, which is in fair agreement with the hypothetical value 7.5 calculated from the assumed bond moments of 0.7 and 1.6D for the C-O and O-H bonds, respectively,[28]

---

[27] F. Perrin, *Jour. Phys. Radium*, Vol. 5, p. 497, 1934.

[28] R. J. W. LeFevre, *Dipole Moments*, Methuen and Company, Ltd., London, 1938.

with an angle of 110° between them, 0.4D for the aromatic C-H bond,[29] and 0.4D for the aromatic C_4H_9-C bond.[30] These values of c_2/c_1 are larger than the value 4.9 for 2,6-dimethylphenol[8] probably because the OH group in 2,4,6-tri-*t*-butylphenol is freer from molecular interaction than is that in 2,6-dimethylphenol as the result of better screening by the large *t*-butyl groups. It is evident from these considerations that whereas the representation of the dielectric data by Equations (6) and (7) makes good physical sense for molecules with freely rotating groups, it is physically unjustified for molecules with hindered group rotations. In the latter case, the dielectric data may best be represented by equations such as Equations (35) and (36), in which an appropriate distribution of relaxation times due to intramolecular potential barriers has been accounted for.

Budo showed[31] that in the event of deviations from free rotation, fluctuating bond moments, deviations of the molecular shape from ellipsoidal, and other possible deviations from the idealized models, the complex reduced dielectric constant should have the form

$$\frac{\epsilon^* - \epsilon_\infty}{\epsilon_0 - \epsilon_\infty} = \sum_{i=1}^{\infty} \frac{c_i}{1 + i\omega\tau}, \tag{37}$$

where

$$\sum c_i = 1.$$

The parameters may be calculated from the bond moments, molecular size and shape, shape of the barrier to intramolecular rotation, and other appropriate factors. While the data presently available in the literature do not allow a plausible examination of such a theory, it should be possible to represent the dielectric data of a great many nonrigid molecules by the use of Equations (35) and (36), in which only deviations from free rotation have been taken into account. Since τ_2 is related to τ_1 by the ratio of the relative sizes of the group and the parent molecule according to Equation (1), only four unknown parameters, τ_1, E_0, c_2 and ϵ_∞ need be evaluated for Equations (35) and (36) instead of the four parameters τ_1, τ_2, c_2 and ϵ_∞ in the case of Equations (6) and (7). Application of these equations will be discussed in a future paper.

[29] J. W. Smith, *Electric Dipole Moments*, Butterworths, London, 1955.

[30] C. P. Smyth, *Dielectric Behavior and Structure*, p. 253, McGraw-Hill Book Co., New York, N. Y., 1955.

[31] A. Budo, "Dielectric Relaxation of Molecules Containing Rotating Polar Groups," *Jour. Chem. Phys.*, Vol. 17, p. 686, Aug. 1949.

THEORY AND APPLICATION OF THE B-CHART

By

J. Breckman

RCA Defense Electronic Products
Camden, New Jersey

Summary—*This paper presents a new projection, well suited to a wide class of trajectory and coverage problems in the space era. The Breckman Projection, or* B*-chart has the following basic properties:*

1. *Equal horizontal distances on the chart take an equal time to traverse, whether the trajectory is circular, elliptical, or even non-ballistic. The horizontal coordinate may be used interchangeably or simultaneously as longitude or time, and the chart displays a continuous ephemeris of the object.*
2. *The orbital plane cuts the chart in a vertical line, and this cut moves across the face of the chart at a rate combining earth rotation, nodal precession and whatever higher derivatives of nodal motion are important.*
3. *The ground track on the chart is a replica of the angle-versus-time characteristic of the object in its own plane, for any trajectory—circular, elliptical, or arbitrary. Hence, perigee precession is accommodated by vertical shifts of the track with respect to the chart.*
4. *The fraction of time an orbiting object spends above an arbitrary area on the earth is numerically equal to that area on the chart. (The chart rectangle has unit area.) This property makes the charts particularly useful for analyzing and improving the statistical performance of the ground environment against the space population.*

Introduction

THE precipitous advent of the space age in October 1957 suddenly presented a class of trajectory, coverage, and operational problems to a technical community that did not have the benefit of an evolutionary process of trying, testing, and finally selecting an appropriate set of analytical and graphical aids to meet the immediate challenges of space engineering. There were, it is true, the centuries-old analytics of celestial mechanics, and the traditional unhurried approaches to orbital calculations. There was also a small and dedicated group of space and rocket scientists whose profound and prophetic pioneering work now underlies most of our current space activity, but these men and women were not in a position where an urgent need existed for fast calculations.

The urgencies of near-earth satellite activity has tremendously compressed the time scale between the appearance of some space situ-

ation, and the preparation of a response. A new and increasing army of enthusiastic space workers, recruited from the widest diversity of non-astronautical backgrounds is engaged in mission planning, orbit determination, guidance system design, lunar flight preparation, military space exploitation, field-station operation, satellite communications, and literally hundreds of allied and supporting space programs. A new set of requirements is evident for making the complex relationships between orbital motion and mission fulfillment tractable at all levels of planning, development, design, and operation of the final systems.

These requirements certainly include at times rapidity, precision, automaticity and all the other niceties of response commonly associated with large-scale digital computers, but anyone who has solved satellite trajectory problems on a day-to-day basis at some field installation that has a richly varied space activity knows that even if a large computer is part of the station complement, there are many occasions where desk-side aids are needed. Every mission planner, systems programmer, feasibility analyst, preliminary designer, and optimizations engineer — in short, everyone for whom the difference between a design trajectory and a precision trajectory is clear — knows the value of graphic aids. Thus there exists a need for an intermediate tool with a response time measured in minutes and with an accuracy comparable to that of a slide rule.

One such tool has become known as the Breckman Projection, or B-chart, which has been in use since 1960. A B-chart is a zone of the earth mapped in such a way that the track of a satellite in a circular orbit appears as a straight line. In this respect, it provides a ground track presentation to the space analyst much like the rhumb line provided by a Mercator Projection. The mapped zone is symmetrical about the equator, its terminal latitudes corresponding numerically to the inclination of the orbits that are displayed on the chart.

The chart displays the full geographic and time history of the object in a nodal day. Cardinal lines on the earth may provide background grids against which the course of the object is traced. These include latitude, longitude, local azimuth, ground range (with respect to particular sensing sites), political boundaries, and geographic features.

Summary of B-Chart Structure

Basic Mapping Procedure

The B-chart is a map of a zone of the earth, using the Breckman Projection. The basic mapping process is illustrated in Figures 1, 2, and 3.

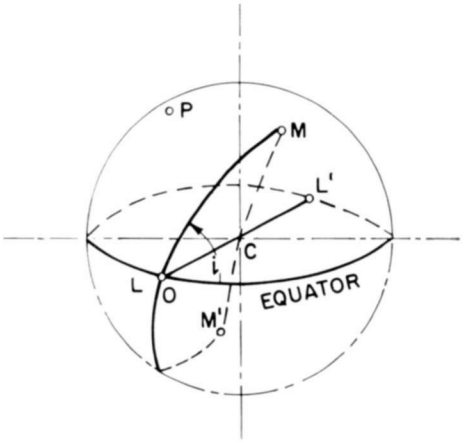


Fig. 1—Initial position of mapping arc.

In Figure 1, point O is the intersection of the Greenwich meridian (not shown) with the equator, C is the center of the earth, and MLM′ is half of a great circle with ML = LM′ = 90°. L is shown coincident with O. The angle i is the mapping angle or inclination angle, and is established with the observer facing O and rotating the arc MLM′ out of the equator counterclockwise for positive i. The arc MLM′, called the mapping arc, is now in mapping position. Points on the earth will be mapped by moving this arc so that i is kept constant, and L remains on the equator. L is called the ascending node; its antinode L′ is the descending node.

In Figure 2, the mapping arc has been moved until it contains P, the point to be mapped. The coordinates of P are l (the longitude of

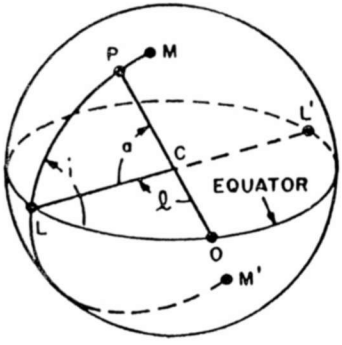


Fig. 2—Final position of mapping arc.

L) and a (the angle LCP); a is positive if P is in the northern hemisphere and negative if P is in the southern hemisphere; l is positive or negative depending on whether L is east or west of O. In Figure 3, the point P is shown mapped onto a plane surface, using l and a as abscissa and ordinate, respectively, each with a linear scale.

The following observations are of interest. First, the process is a generalization of the conventional longitude–latitude system of designating a point on the earth's surface. In the conventional system, i is taken as 90°; but, in the present system, i may be any angle suitable

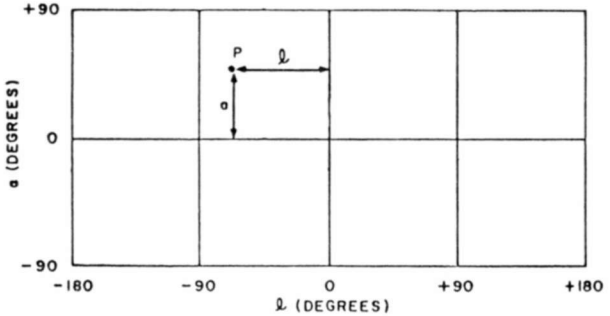


Fig. 3—Rudimentary B-chart for a particular inclination angle.

to the problem at hand. Second, only points between the latitudes corresponding to M and M′ may be mapped for a particular i; the region between these latitudes is called the map zone. Third, there is a one-to-one correspondence between the points in the map zone and the points on the map (Figure 3). Finally, the choice of Cartesian coordinates for plotting the number pair (l, a) is purely arbitrary, and is made with a view to convenience in the application.

Two Points for One

For the pertinent applications of B-charts, it is convenient to have two mapped points on the chart for each point in the map zone. Thus, a mapping great circle, MLM′L′ in Figure 4, is used instead of the half of a great circle in Figure 1. Figure 5 shows the two possible positions of the mapping circle to include the point P. Two number pairs result, (l_1, a_1) and (l_2, a_2). The pair in which a is an acute angle is a designation of the first kind; the other pair is a designation of the second kind. The resultant B-chart is shown in Figure 6.

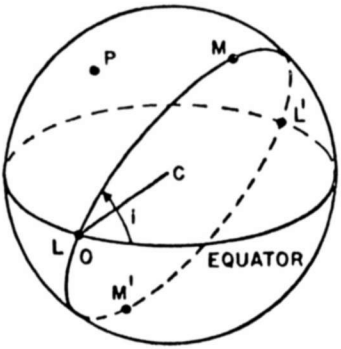


Fig. 4—Initial position of mapping circle.

Fig. 5—Developing the two designations.

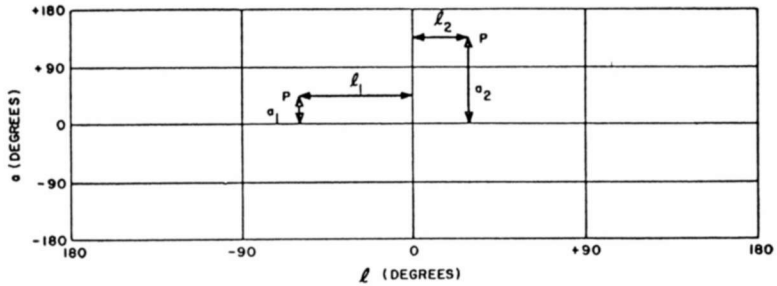


Fig. 6—B-chart showing one-to-two correspondence.

Shifts in the Coordinate Origins

It is convenient in the applications to use positive angles for both a and l. Accordingly, a will be taken as an angle between 0° and 360°, positive in the sense LML′M′; l is taken as an angle between 0° and 360°, positive when measured west of O. In other words, l is the *west longitude of the ascending node*, and a is the *argument from the ascending node*. The new situation is shown in Figure 7.

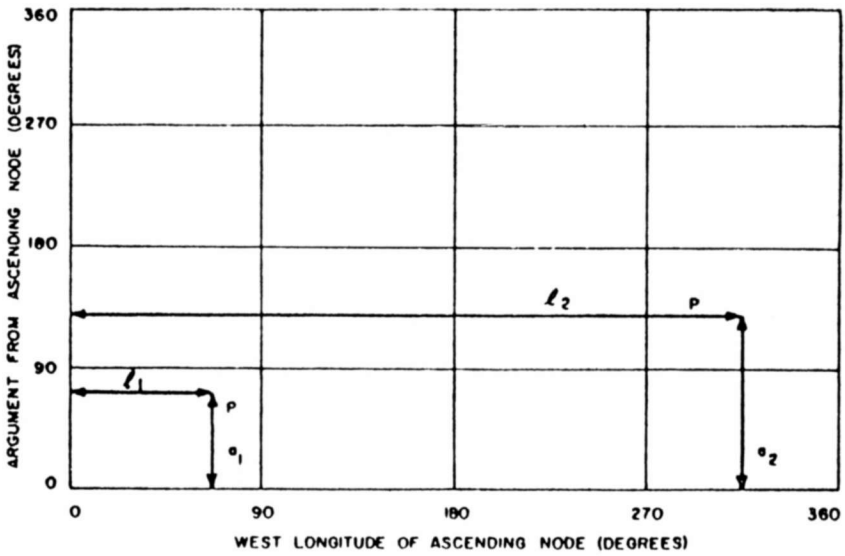


Fig. 7—B-chart with all angles positive.

In anticipation of the applications, some changes were made; the west longitude, l, is run from right to left (i.e., increasing to the west) and the argument, a, is started at the quarter line. Thus, the two kinds of designations appear in the lower and upper halves of the sheet as shown in Figure 8.

In the applications, i will be a first- or second-quadrant angle; that is $0° \leq i \leq 180°$. Now it is easily shown for supplementary angles i_1 and i_2, that

$$[l_1\ (i_1) - l_2\ (i_2)]\ \text{mod}\ 360° = 180°,$$

$$[a_1\ (i_1) + a_2\ (i_2)]\ \text{mod}\ 360° = 180°,$$

in which $l_1(i_1)$, $a_1(i_1)$ refer to the designation of the first kind for the i_1 inclination. Hence, the same chart may be made to serve for

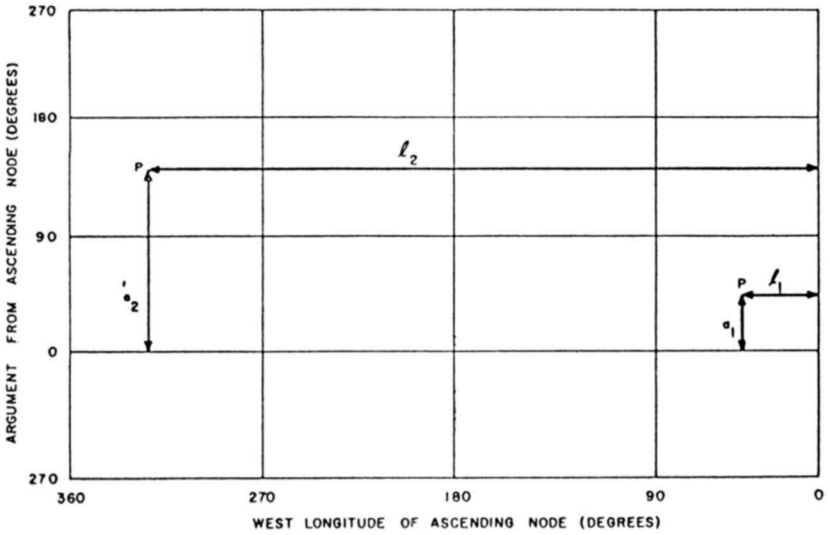


Fig. 8—B-chart with linear scale transformations.

two angles, i and $180°-i$, by entering a second pair of Cartesian coordinates, as in Figure 9. Note that the point P on the earth is plotted at the same two places on the B-chart for both i_1 and i_2. In using the chart of Figure 9, i_1 is arbitrarily taken as the acute angle, and i_2 as the obtuse angle. Hence the lower and left-hand scales are to be used for acute inclinations, and the upper and right-hand scales for obtuse inclinations.

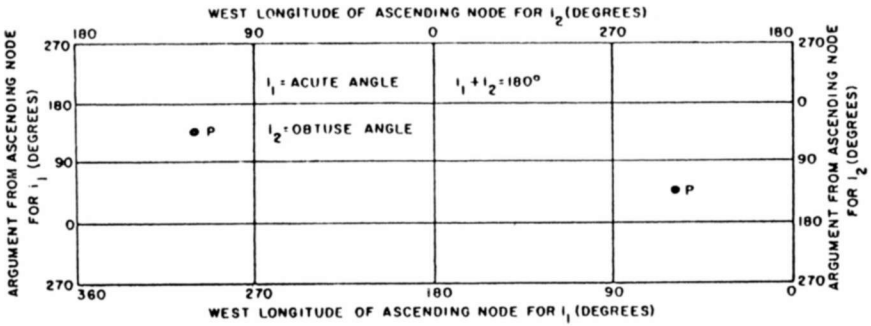


Fig. 9—Basic B-chart showing both sets of Cartesian coordinates.

Addition of Conventional Latitude and Longitude

The problems to which the B-chart will be applied will require frequent conversions between the B-coordinates (l, a) and the conventional designations of longitude and latitude. To make the chart self-converting, latitude tic marks and a longitude index line are superimposed on the chart of Figure 9, resulting in Figure 10.

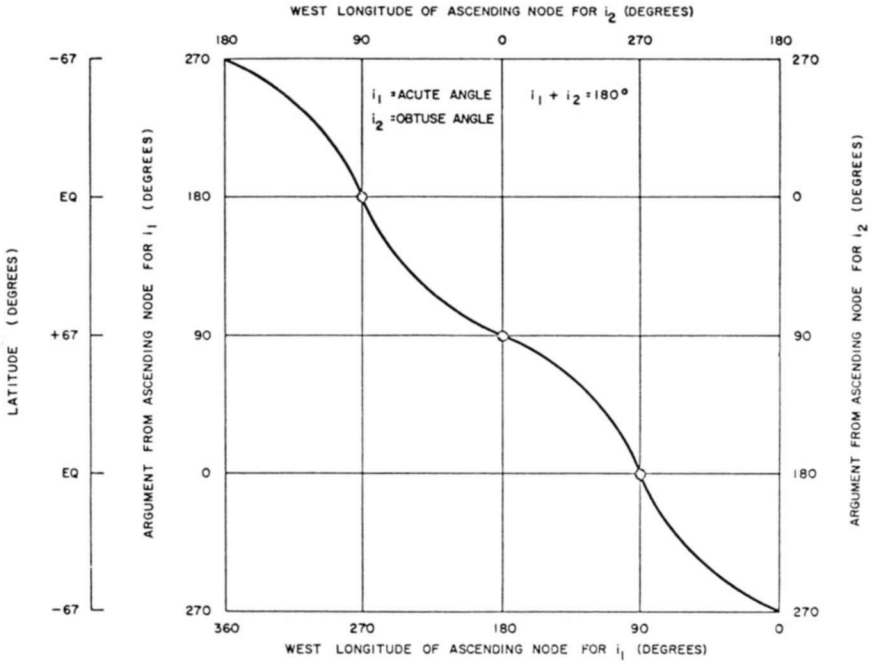


Fig. 10—B-chart with latitude and longitude superimposed. (Inclinations of 67°/113° used as example.)

The longitude index line is used in conjunction with a sliding rule that is an adjunct of the charts, the same rule serving for all inclination angles. The index line used corresponds to the meridian of 90° west longitude. Its shape on any chart, and the spacing of the latitudinal tic marks, depend on the angle i (and will be the same for the supplementary angle $180°—i$). Any line of latitude in the mapped zone will appear as a horizontal line on the B-chart, and uniformly spaced meridians on the earth will map as uniformly spaced sinuous lines on the B-chart. The only mapping angles for which meridians appear as straight lines on the B-chart are $i = 90°$, when meridians appear as

vertical, and the degenerate case where $i = 0°$ (or 180°), when meridians appear as oblique straight lines.

The Local Grid

The discussion thus far has shown two ways of designating a point on the earth. The B-coordinates (l, a), and the conventional longitude–latitude system. There is a third way to be encountered in

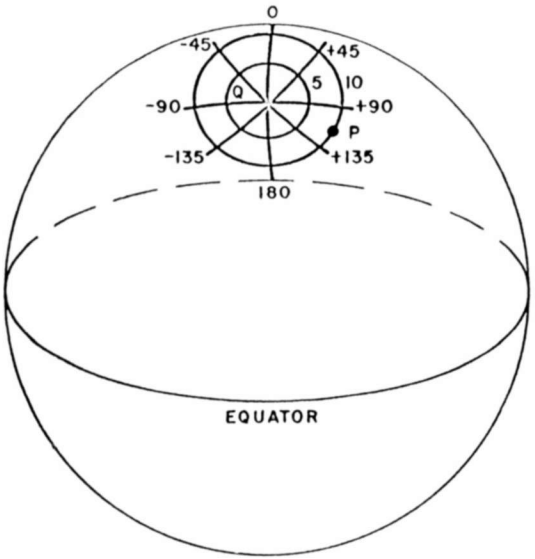


Fig. 11—The local grid.

the applications, which is analogous to the conventional latitude–longitude system, except that the pole of this third system is an arbitrary point called the *site*, instead of the north pole. Figure 11 illustrates the system.

The site is at point Q. A family of great circles pass through Q, generating a system of azimuths, θ, with the north direction as 0° azimuth. A family of small circles, concentric with Q constitute an orthogonal coordinate. This is the ground angle, g, from the site. Note that azimuth is analogous to longitude, and ground angle to co-latitude, with the site as pole. In the figure, the point P is at $\theta = 120°$, $g = 10°$, that is, at (120°, 10°).

The essential sameness of the three coordinate systems discussed so far can be seen. Previously it was indicated that the (l, a) system was really a generalization of the longitude–latitude system with re-

spect to the mapping angle. It can now be seen that the (θ, g) system is also a generalization of the longitude–latitude system with respect to the effective equator (or equivalently, the effective pole). There are many other coordinate systems and generalizations; however, these three systems are met frequently in the applications, and the B-charts provide a self-converting mechanism from any one system to any other

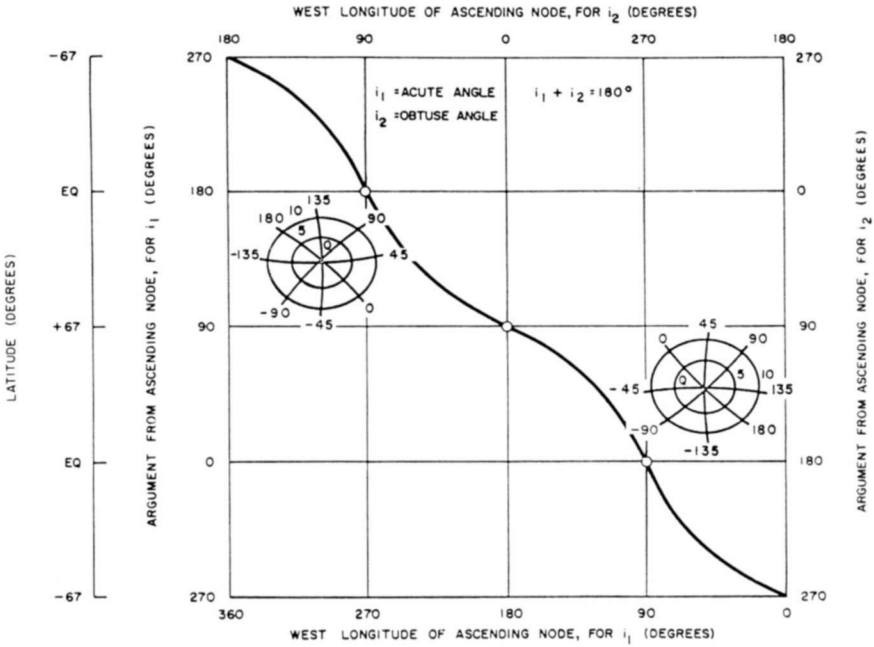


Fig. 12—Compound B-chart, including local grid.

when the local grid is superimposed on the basic chart of Figure 10. This situation is illustrated in Figure 12. The only part of a particular local grid that appears on a B-chart is the part contained in the map zone. Also, in many applications, a practical upper limit on g is about 40°, and because of the physical nature in which g arises in most problems, $g = 90°$ is the theoretical limit. Note that any point or line in the map zone may be plotted on the B-chart, including cities, political boundaries, etc. as the occasion requires.

Satellites and B-Charts

The Orbital Cut

The instantaneous intersection of the orbital plane of a satellite

with the earth's surface, sometimes called the orbital trace, is here called an orbital cut. The orbit need not be either elliptical or ballistic. The orbital cut is a great circle. Also, the angle this great circle makes with the equatorial plane is the inclination angle of the orbit. Hence the orbital inclination angle is the same as the inclination or mapping angle used in generating the B-charts. The orbital cut, being a line on the earth, is mappable onto appropriate B-charts; in particular, if it is mapped onto a chart whose mapping angle corresponds to the orbital inclination, and with the ascending node of the mapping circle corresponding to the ascending node of the orbit, it maps as a vertical straight line. When the B-chart map of satellite behavior is discussed, it is understood to mean the chart whose mapping angle equals the inclination of the orbit, and that the ascending nodes of mapping circle and orbit are aligned. Thus, *the orbital cut is a vertical line on the B-chart.*

The Flight Line

The ground point of a point in, on, or above the earth is where the earth's radius to the point pierces the earth's surface. The path of the ground point of a satellite, sometimes called the subsatellite trace, is here called the flight line.

Suppose an object is moving in a plane in space containing the earth's center, as satellites do. Then if a is the angle the radius vector makes at any instant with some reference line in the plane, and if a periodic angular motion is assumed, an argument-versus-time characteristic can be developed as shown in Figure 13 where time is shown increasing to the left. In the figure, one cycle of activity is shown. Also, at $t = 0$, the object is at the reference radius. The shape of the curve in this figure corresponds roughly to ballistic motion in an ellipse. However, ballistic or elliptical motion is not required in this discussion.

Consider, for a moment, the orbital cut. The earth is in more or less uniform rotational motion with respect to the orbital plane, so that the orbital cut appears to move across the face of the B-chart, as shown in Figure 14. For a stationary plane, or one whose rotation is westward (or, if eastward, not faster than the earth's motion), this drift across the chart is to the left, as in the figure. For ballistic objects, this is always the case.

Note that the horizontal coordinate, being proportional to time, constitutes a uniform time scale, while the vertical coordinate is just the orbital argument. In other words, except for a horizontal scale

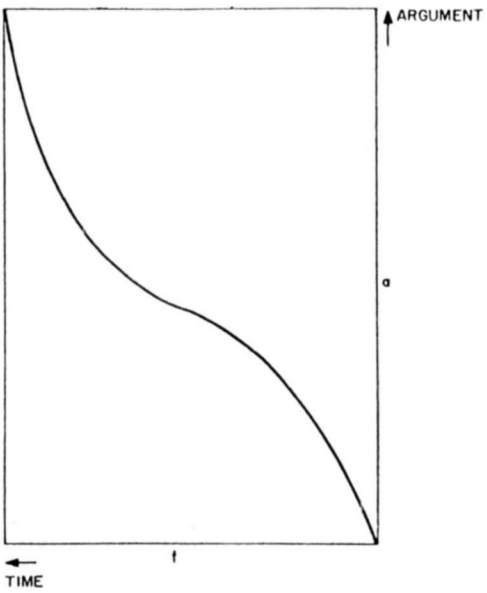


Fig. 13—Argument versus time characteristic.

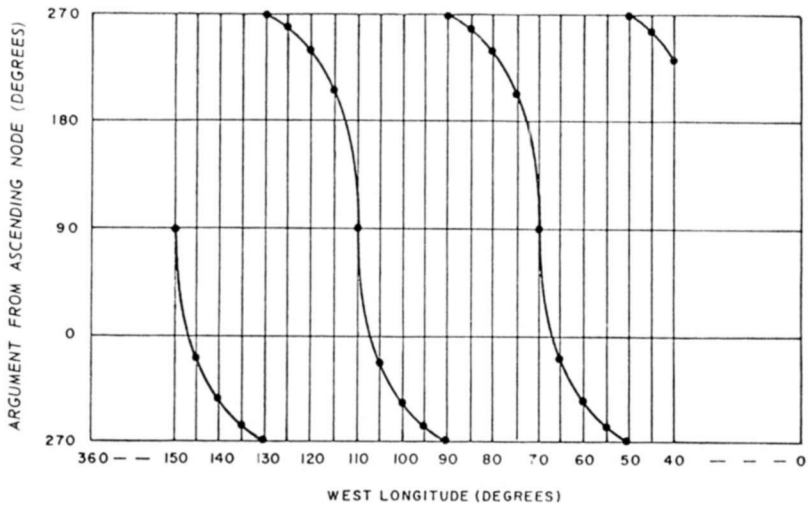


Fig. 14—B-chart showing drift of the orbital cut and development of the flight line. The succession of vertical lines represents the westward movement of the orbital cut over the earth; the succession of heavy dots represents the movement of the object in its orbit. The curve through the heavy dots then constitutes the flight line.

factor *the flight line of the B-chart across the geography of the earth is merely the argument-versus-time characteristic in the orbital plane.*

In a circular orbit, the flight line is a straight line since the argument-versus-time characteristic is a straight line. In the analytical problems for which the B-charts were developed, the problems may almost always be analyzed in terms of circular orbits, so that drawing the flight lines becomes a matter of ruling a family of oblique parallel, properly spaced, straight lines. Notice that for comparatively inertial

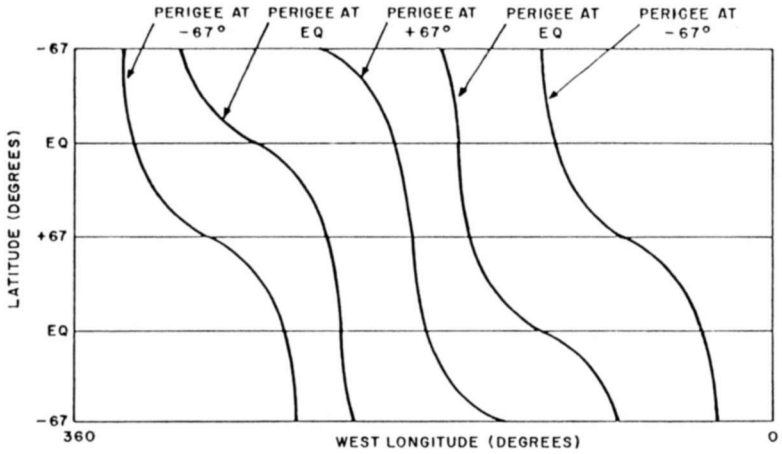


Fig. 15—Flight lines corresponding to several arguments of perigee.

orbits, the flight lines always proceed to the left; for acute inclinations they go upward to the left, and for obtuse inclinations they go downward to the left. Note also that *for acute inclinations, northbound motion appears in the lower half of chart, southbound in the upper half; for obtuse inclinations, northbound motion appears in the upper half of chart, southbound in the lower half.*

Argument of Perigee

The argument-versus-time characteristic is generally drawn in the orbital plane with the radius to perigee as zero argument. However, the B-charts have zero argument corresponding to the equator. In the event that the perigee is not at the equator, the appropriate flight line is drawn by aligning the argument of perigee to the proper latitude on the B-chart, taking care to keep in mind whether perigee is traversed during the northbound or southbound motion of the object. In other words *the flight line is adjusted to motion of perigee by vertical movement of the B-chart.*

Time

One of the more important properties of the B-chart has been implied in the discussion of flight lines. Another important consequence of Figure 14 is that *equal horizontal displacements on the flight line require equal time.* In other words, the primarily geographic, horizontal coordinate is also a uniform time scale.

Statistics

Ordinarily, it might be expected that after sufficient launchings of objects at or near a given inclination, the orbits would be distributed more or less uniformly in right ascension, and at any given moment the objects within the orbits would be distributed more or less uniformly in argument. However, for a given inclination, the west longitudes of the orbits and the right ascensions have similar distributions; hence, *equal areas on the B-chart have equal probability of containing a given object.*

Alternatively, the expected number of objects within an area is proportional to the area. Of course, where the right ascension and/or argument distributions are not uniform, but their shapes are known or assumed, it is a relatively simple matter to generate the distribution on the B-chart.

Scan Lines

It has been pointed out that events in space may be referred to the earth and thence to the B-chart by means of the ground points corresponding to the point above the earth. This is the principle used in generating the so-called scan line. Briefly, the scan line is the locus of ground points corresponding to the set of points in space at which penetrations of a given radar scan with a particular object (or class of object) might occur. To find the scan line, the geometry of the proposed scan and the orbit of the object must be known. The scan line, like any other line on the earth, has two plots on the B-chart; one covers the northbound excursions of the object, the other covers the southbound excursions. It was primarily to show this total relationship of the scan to the flight line, in separate parts, that the map zone was plotted twice.

B-Chart Mathematics

To find the B-chart coordinates (l, a) of a point P, given its latitude ψ and its longitude $\mathcal{L}$, when the inclination is i, we note that

$$\sin a = \sin \psi / \sin i$$

$$\cos a = \pm (1 - \sin^2 a)^{\frac{1}{2}}.$$

Hence there are two solutions, a_1 and a_2. Also

$$\sin \phi = \tan \psi \cot i$$

$$\cos \phi_j = \cos a_j / \cos \psi, \qquad j = 1, 2,$$

where ϕ is the longitudinal displacement of P from the ascending node.

Finally,

$$l_j = \mathcal{L} + \phi_j$$

so that we have proceeded from the single earth point P $(\psi, \mathcal{L})$ to two B-chart points P_1 (l_1, a_1) and P_2 (l_2, a_2).

Uses of the Charts

In general, the uses of the charts fall into two broad categories — design uses (synthesis, analysis and performance evaluation) and field uses. Some of the problems to which the charts have already been applied are listed below.

Design Applications (statistics of objects and ground environment)

(1) Design of radar scans to match a particular site to an assigned mission, such as monitoring space for new objects in a designated domain of height and inclination, updating already catalogued objects, rediscovering "lost" objects, and detecting remnant companions of newly launched objects.

(2) Design of radar scans to match a group of sites to an assigned mission, so that each sensor works as part of a team, specification of the radar characteristics needed at each site for this mission, and evaluation of the performance of the resulting network configuration against given mission criteria. This network approach avoids the overdesign of individual radars that may result from trying to make each site do the total job.

(3) Optimization of site locations to meet given mission requirements. This problem is particularly important in programs involving space surveillance, satellite inspection, translunar shots and deep space probes, orbital rendezvous or other "critical event" condition, recovery problems and landing stations.

(4) Computation of the average and peak data-processing burdens of space surveillance stations.

(5) Specification of the cluster of satellites needed for global communication missions, with respect to designing the orbits, placing a family of objects within an orbit, and arranging a family of such orbiting families; determination of the deterioration of the network performance as the geometric relationships among the various objects change with time.

(6) Evaluation of the performance of existing and postulated ground networks for detecting ballistic missile attacks, considering all physically feasible energies, geometrically feasible launch angles, and politically feasible launch areas.

(7) Conduct of feasibility studies on mid-course intercept of ballistic missiles.

(8) Determination of optimum ground tracks and hence flight profiles (and then thrust timetables) for critical phases of special missions. The ability to do this stems from the fact that the B-chart ground track is a replica of the angle-versus-time characteristic.

Field applications (particulars of objects and ground environment)

(1) Generation of local "look" angles, range, range rate, and time under field conditions, conditions of urgency, or any other situation where a computer solution is not available or practical.

(2) Provision of a continuous ephemeris in topocentric or geocentric coordinates.

(3) Development of on-board navigational procedures using B-charts based on the earth, moon, sun, planets, or other celestial spheres.

(4) Determination of the flight profile of a maneuvering or highly perturbed object, from ground observations.

(5) Development of the range, azimuth, elevation, and range rate programs (all versus time) for a particular sensor, to follow characteristic points on a given orbit such as closest point on the orbit, point of closest approach, point of constant range, etc.

(6) Determination and display of the conditions for optical visibility of a given object from a given site.

(7) Generation of displays of the space situation for control centers and on-line flight analysts.

(8) Provision of quick-fix aids to auxiliary groups such as moonwatch teams, amateur radio operators, etc.

EFFECT OF HIGH MAGNETIC FIELD ON ELECTRON-BEAM NOISE

By

J. M. Hammer and C. P. Wen

RCA Laboratories
Princeton, New Jersey

Summary—*High magnetic fields, applied either to the helix input region or over the entire tube, reduce the electron-beam noise and traveling-wave-tube noise figure to very low values. The high fields act to suppress an additive noise term that is due to transverse velocity and position fluctuations in the electron beam. The fact that the peak field values may be applied to the input region of the helix is of practical significance in allowing a relatively light focusing magnet to be used without suffering from the beam expansion that accompanies the peaked cathode field formerly thought to be necessary.*

It has recently been shown that very high magnetic focusing fields reduce the equivalent noise temperature of microwave electron-beam amplifiers to very low values.[1] The present study indicates that the magnetic field suppresses transverse noise fluctuations that couple to the principal fields of the amplifier.

Noise-figure measurements have been made on an S-band traveling-wave tube focused by solenoids capable of providing axially variable magnetic fields. Figure 1 contrasts the noise behavior of the tube for two widely different conditions of electron flow in the cathode vicinity as the magnetic field strength is varied. In these measurements the magnetic field is peaked at the cathode, but both the peak and main field are varied proportionally; the average field over the input region of the helix is maintained at 0.75 times the value at the peak. The noise effect of the beam-forming electrode and the first-anode voltages, V_{BF} and V_1, has been treated at length.[2,3] As expected, with V_{BF} at -12 volts the excess noise $F-1=T_n/T_0$ is of the order of 7.0, while with V_{BF} at $+8$ volts the excess noise is below 1.0. Despite the wide difference in the cathode flow condition and hence in the excess noise,

[1] J. M. Hammer and E. E. Thomas, "Traveling-Wave-Tube Noise Figures of 1.0 db at S-Band," *Proc. IEEE*, Vol. 52, p. 207, Feb. 1964.

[2] J. M. Hammer, "Power Spectra Measurements on Ultralow-Noise Beams," *Jour. Appl. Phys.*, Vol. 35, p. 1147, April 1964.

[3] M. R. Currie and D. C. Forster, "Conditions for Minimum Noise Generation in Backward-Wave Amplifiers," *IRE Trans. Electron Devices*, ED-5, p. 88, April 1958.

the rate of change of noise with magnetic field is almost the same for the two cases. If the effect of the high magnetic field was either to alter the cathode flow conditions or to cause a direct interaction with the longitudinal noise flow in the electron gun, the reduction would be proportional to the excess noise, which is contrary to the observation. It is clear, then, that the magnetic field is acting on a noise term that

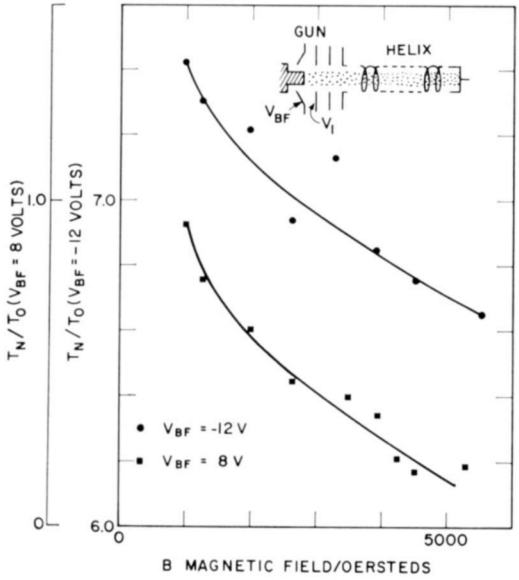


Fig. 1—Excess noise $F - 1 = T_n/T_0$ versus peak value of magnetic field for two different operating conditions of the multivelocity region of the gun. At $V_{BF} = -12$ volts, $V_1 = 26.3$ volts; at $V_{BF} = 8.0$ volts, $V_1 = 3.4$ volts; beam current = 102 microamperes.

is additive to the longitudinal noise which, as is well known, is controlled by V_{BF} and V_1.

The nature of this additive term is probed by a series of noise measurements with various arrangements of magnetic field. Increasing the magnetic field over the output section of the helix, while maintaining the remainder of the tube in a constant field, results in no change in noise figure. Thus, the reduction is not occurring in the output region of the tube. Figure 2 shows the results of measurements made by varying the magnetic field over the input region of the helix while holding the field over the remainder of the tube (B_0) sensibly constant. The results of varying a uniform field over the entire tube are also shown. The data of Figure 2 were taken after the tube had

aged considerably and some degradation in noise performance as compared with the earlier measurements[1] had occurred. In this figure the excess noise is plotted versus the reciprocal of the value of the magnetic field in the input section of the helix. The results for both the uniform field and the various peaked-field combinations are well fitted by a straight line. Thus, the additive term is proportional to $1/B$ and fur-

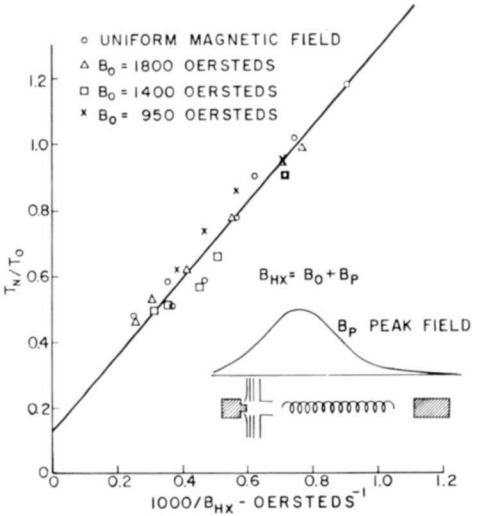


Fig. 2—Excess noise $F - 1 = T_n/T_0$ versus reciprocal of field over input section of helix (average), B_{HX}. The tube is immersed in a uniform field of value B_0. The value of the peak field B_P is then varied for different values of B_0.

ther reduction in noise is indicated at even higher values of magnetic field. It should be mentioned that uniform-field data taken before the tube noise had degraded showed that the straight line in the $1/B$ plot passes through the origin. This is interesting in that it indicates that the longitudinal noise term (S-II) was zero.

In a number of treatments of purely transverse-wave amplifiers,[4-8]

---

[4] G. Wade, K. Amo, and D. A. Watkins, "Noise in Transverse-Field Traveling-Wave Tubes," *Jour. Appl. Phys.*, Vol. 25, p. 1514, Dec. 1954.

[5] A. E. Siegman, "Waves on a Filamentary Electron Beam in a Transverse-Field Slow-Wave Circuit," *Jour. Appl. Phys.*, Vol. 31, p. 17, Jan. 1960.

[6] E. I. Gordon, "Charged-Particle Orbits in Varying Magnetic Fields," *Jour. Appl. Phys.*, Vol. 31, p. 1187, July 1960.

[7] R. Adler and G. Wade, "Beam Refrigeration by Means of Large Magnetic Fields," *Jour. Appl. Phys.*, Vol. 31, p. 1201, July 1960.

[8] T. Wessel-Berg and K. Blötekjaer, "Noise Reduction Schemes in Transverse Modulation Tubes," *Low Noise Electronics*, The Macmillan Co., New York, p. 142, 1962.

it is shown that the noise power follows a $1/B$ law. In the case of an actual longitudinal-wave amplifier such as a traveling-wave tube, there are large radial field components associated with the growing wave. These components can be excited by transverse noise and give rise to just such an additive term as seems to be indicated by the measurements reported here. Thus one might expect the proper form of the noise expression for an electron-beam amplifier to be

$$F-1=T_n/T_0=\frac{S-\mathrm{II}}{S_0}\frac{T_c}{T_0}f(\gamma a)+\frac{T_{HX}}{T_0}g(\gamma a)+\frac{\omega T_c}{\omega_c T_0}h(\gamma a). \quad (1)$$

T_n is the equivalent amplifier noise temperature, T_0 is room temperature, T_{HX} is the temperature of the amplifying structure, T_c is cathode temperature, ω_c the cyclotron frequency (eB/m), γ is the propagation constant, and a the beam diameter. The first term is the familiar longitudinal noise term;[9] the second term is the contribution of the thermal noise on the amplifying structure;[10] and the last term is the contribution due to the transverse noise, assuming that a term of the form $\omega\omega_c$ due to synchronous waves does not couple to the amplifying structure. $f(\gamma a)$, $g(\gamma a)$, and $h(\gamma a)$ are the appropriate loss and coupling terms.

If, for the traveling-wave tube studied here, one takes the transverse coupling term to be simply the square of the ratio of the average transverse to the average longitudinal electric field (i.e., $h(\gamma a)=<I_1^2(\gamma a)>/<I_0^2(\gamma a)>$ where I_0 and I_1 are the modified Bessel functions), the third term of Equation (1) contributes 0.65 at 1000 oersteds and 0.16 at 4000 oersteds. These values are in reasonable agreement with the observed noise figures. Note that the 57°K beam noise of Reference (1) corresponds to $T_n/T_0=0.19$.

It is clear from these measurements that the application of high magnetic fields to either the input region of the amplifying structure or over the entire tube reduces an additive noise term that falls as $1/B$. Because of these properties there can be little doubt that the term originates in transverse velocity and position fluctuations in the current emitted from the thermionic cathode. The transverse fluctuations are statistically independent of the longitudinal fluctuations, and thus enter as an added term. At the same time the existing transverse noise theory, albeit not specifically directed to the problem at hand, leads one to expect the $1/B$ field dependence found. The question of the relative importance of the field at the cathode to that at the helix

[9] H. A. Haus and F. N. H. Robinson, "The Minimum Noise Figure of Microwave Beam Amplifiers," *Proc. IRE*, Vol. 43, p. 981, Aug. 1955.

[10] S. Bloom, "Effect of Distributed-Loss Noise Generators on Traveling-Wave-Tube Noise Factor," *RCA Review*, Vol. 22, p. 347, June 1961.

input is not fully resolved since, in the nonuniform field measurements, the beam diameter is reduced on entering the peaked helix field. The reduced beam diameter results in a reduced coupling of the transverse noise, which can partly account for the observed lowering of T_n. Under certain conditions the reduction in the coupling factor can go as $1/B_{HX}$ for fixed B_0.

Finally, the practical significance of being able to suppress noise by applying a peaked field over the interacting structure should be mentioned. Weight can be saved by not applying the high field over the entire tube. Beam expansion is avoided by not having a peak at the cathode. These are both real advantages in the construction of practical low-noise devices.

Acknowledgment

The authors wish to thank E. E. Thomas, who carefully performed the measurements reported on in this paper. Much gratitude is due to numerous members of the Microwave Research Laboratory, especially S. Bloom and L. S. Nergaard, for many illuminating discussions and helpful suggestions.

RCA Technical Papers†

Third Quarter, 1964

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

"Accuracy in Automatic Test Equipment," M. C. Kidd, *Trans. IEEE PTGAS* (July) 1964
"A Broad Tunable Bandwidth Traveling-Wave Maser," L. C. Morris and D. J. Miller, *Trans. IEEE PTGMTT* (July) 1964
"Cinerama Theatre Acoustics," M. Rettinger, *Jour. S.M.P.T.E.* (July) 1964
"Circular Polarization at Millimeter Waves by Total Internal Reflection," H. Buizert, *Trans. IEEE PTGMTT* (Correspondence) (July) 1964
"Continuous Pumping of Cesium Vapor Devices," J. R. Fendley, Jr., *Rev. Sci. Instr.* (Notes) (July) 1964
"Effect of the Self-Magnetic Field on Galvanomagnetic Effects in Bismuth," S. Tosima and R. Hirota, *IBM Journal of Research and Development* (July) 1964
"The Engineer's Responsibilities in Technical Documents," M. Hollander, *Trans. IEEE PTGAS* (July) 1964
"Microwave Amplification with Superconductors," A. S. Clorfeine, *Proc. IEEE* (Correspondence) (July) 1964
"Using the Computer for Integrated Circuit Analysis," L. C. Drew and A. G. Atwood, *Electronic Industries* (July) 1964
"Gamma-Induced Divalent Dysprosium in Calcium Fluoride," F. K. Fong, *Jour. Chem. Phys.* (1 July) 1964
"Conductivity Tensor of an Anisotropic Quantum Plasma in a Uniform Magnetic Field," J. J. Quinn, *Phys. Rev.* (6 July) ... 1964
"Density Matrix Formulation of Small-Polaron Motion," L. Friedman, *Phys. Rev.* (6 July) 1964
"Two-Photon Ionization of Atomic Hydrogen," W. Zernik, *Phys. Rev.* (6 July) 1964
"Switching Times of the Current-Induced, Superconducting-to-Normal Transition in Filaments of Tin and Indium," J. I. Gittleman and S. Bozowski, *Phys. Rev.* (20 July) 1964
"Hollow Center CdS Crystals," A. Dreeben, *Jour. Appl. Phys.* (Communications) (August) 1964
"The Effect of Crossed Field on the Velocity Distribution of Hollow Electron Beams," J. Pearl, *Trans. IEEE PTGED* (August) . 1964
"Elastic Constants of Barium Fluoride Between 4.2 and 300°K," D. Gerlich, *Phys. Rev.* (31 August) 1964
"High-Band TV Tape Recording," H. H. Klerx, *Broadcast News* (August) 1964
"How the TK-22 Gets the Best from Film," N. P. Kellaway and J. C. Adison, *Broadcast News* (August) 1964
"Ionization Mechanism and Electron Temperature in Cesium Arcs," J. R. Fendley, Jr. and K. G. Hernqvist (Correspondence) *Proc. IEEE* (August) 1964
"Large Power Potentials," K. Fischbeck, *Signal* (August) 1964

† Report all corrections to *RCA Review*, RCA Laboratories, Princeton, New Jersey.

"Magnetic Field Penetration into Niobium-Stannide Discs," K. G. Petzinger and J. J. Hanak, *RCA Review* (September) 1964
"Magnetization of Niobium-Stannide Films in Transverse Fields," J. J. Hanak, *RCA Review* (September) 1964
"Microwave Studies of Niobium Stannide," B. Rosenblum, M. Cardona, and G. Fischer, *RCA Review* (September) 1964
"Parametric Solution of the Equations of the Schwarzschild Two-Reflector Optical System," G. E. Roberts, *Jour. Opt. Soc. Amer.* (September) 1964
"The Point-Contract Flexode," B. E. Tompkins, *Proc. IEEE* (Correspondence) (September) 1964
"Preparation and Properties of Vapor-Deposited Niobium Stannide," J. J. Hanak, K. Strater, and G. W. Cullen, *RCA Review* (September) 1964
"Simplified Rating System for Silicon Rectifiers Used in Capacitive Load Circuits," B. J. Roman and J. M. S. Neilson, *Semiconductor Products and Solid State Technology* (September) 1964
"Surge-Magnetic-Field and Pulse-Current Effects in Niobium Stannide," W. H. Cherry, *RCA Review* (September) 1964
"Superconducting Energy Gap and Net Electron Drift Velocity as Functions of Temperature and Cooper-Pair Drift Velocity," R. H. Parmenter and L. J. Berton, *RCA Review* (September) 1964
"The Superconducting Energy Gap of Niobium Stannide," G. D. Cody, Y. Goldstein, and R. Cohen, *RCA Review* (September) 1964
"The Superconducting Penetration Depth of Niobium Stannide," G. D. Cody, *RCA Review* (September) 1964
"Superconducting Properties of the $(Nb, Ta, V)_3Sn$ System," G. D. Cody, J. J. Hanak, J. P. McConville, and F. D. Rosi, *RCA Review* (September) 1964
"30-GC Multiplier as an Ultrastable Parametric Amplifier Pump," R. J. Kampf, R. S. Forman, O. J. Hanas, and D. H. Knapschaefer, *Proc. IEEE* (Correspondence) (September) 1964
"Transition Temperature of Niobium Stannide," J. L. Cooper, *RCA Review* (September) 1964
"VSWR Measurement of Attenuators for Helix-Type Traveling-Wave Tubes," E. Belohoubek (Correspondence) *Trans. IEEE PTGED* (September) 1964
"A Wide Dynamic Range Tunnel Diode Amplifier Employing Automatic Gain Control Approach for Phase Tracking Systems," V. Stachejko, *Proc. IEEE* (Correspondence) (September) .. 1964
"Crystal Growth and Color Centers of Alkaline-Earth Halides," F. K. Fong and P. N. Yocom, *Jour. Chem. Phys.* (1 September) 1964
"Stable Divalent Rare-Earth-Alkaline-Earth Halide Systems," Z. J. Kiss and P. N. Yocom, *Jour. Chem. Phys.* (Letters to the Editor) (1 September) 1964
"Zeeman Effect of the $2F_{5/2}$, $E_{5/2} \rightarrow 2F_{7/2}$, $E_{5/2}$ Transition in $CaF_2:Tm^{2+}$," H. A. Weakliem and Z. J. Kiss, *Jour. Chem. Phys.* (1 September) 1964
"Designing Noise Immunity Into High-Speed Circuits," D. R. Gipp, *Electronics* (September 7) 1964
"Tunnel-Diode Circuits Invert Direct to Alternating Current," F. M. Carlson, *Electronics* (September 21) 1964

AUTHORS

LAZAR BECKER received the B.E.E. degree (summa cum laude) in 1953 and the M.E.E. degree in 1955 from Polytechnic Institute of Brooklyn. From 1953 to 1956 he was employed by Sylvania Electric Products where he developed applications and recommended modifications for experimental receiving and miniature tubes. With the Sperry Gyroscope Company, from 1956 to 1959, he was responsible for design, development, production refinement, and applications of traveling wave tubes. Mr. Becker joined PRD Electronics in 1959 where he designed microwave components and supervised a microwave storage project which utilized cryogenic techniques. In 1963, he joined the technical staff of RCA Communications Systems Division in New York as a senior member, specializing in the development of advanced solid-state r-f techniques. His current activities include interference reduction techniques for X-band receivers and investigation of crystal rectifier, parametric and tunnel diode downconverters. Mr. Becker is a senior member of the Institute of Electrical and Electronics Engineers.

RONALD BINKS received the H.N.C.E.E. degree in England in 1952. He received his early training with the Automatic Telephone & Electric Co. Ltd., where he worked on test equipment design and development and multichannel V.H.F. Radio equipment engineering. He worked with British Telecommunications Research Ltd. on the design of equipment for 4 and 12 mc frequency-division multiplex communications systems and also on the design of specialized test equipment to aid the design, manufacture and customer maintenance of these systems. In 1961 he joined the Broadcast & Communications group of RCA where he has worked on the design of equipment for the CV240 multiplex development program. He has been concerned with test and measuring requirements of this program and with the filter design and development aspects. At present Mr. Binks is studying for a master's degree at the University of Pennsylvania.

JACK BRECKMAN received the B. S. degree in Electrical Engineering from Cooper Union in 1943 and the M.S. degree from the University of Pennsylvania in 1962. He is presently engaged in the doctorate program at that University. Until 1949 he designed test equipment and taught mathematics. While with the U.S. Signal Corps (1949-1954) he designed analog and digital communication and control systems. Since joining RCA in 1954, Mr. Breckman designed digital control and data-processing equipment for BIZMAC and has been engaged in the application of computer programs in several projects including BMEWS, PAGE, and EAST. He invented the B-Chart (Breckman Projection) and developed its applications to radar scan design, communications, rendezvous, navigation, optics, and on-site operation. He is presently Systems Manager, Space Operations, in the SEER organization.

FRANK M. BROCK received the B.S. degree in Electrical Engineering from the University of Alberta, Canada, in 1950. He received the M.S. degree from the University of Pennsylvania in 1962. Following graduation in 1950, he spent a year on the G.E. Test Course with the Canadian General Electric Co. in Canada, and then joined the Electronics Department of C.G.E., working on radar and microwave communications equipment. In 1953 he joined the Microwave Engineering Department of RCA in Camden, N.J., to participate in the design and development of frequency-division multiplex communications equipment and systems. He was involved in the original development program for the AN/GRC-50 radio relay equipment, and he acted as the Project Engineer for the development of the CT-42 solid state FSK telegraph data equipment, and for the development of precision filters and networks for the CV-240 multiplex equipment. Mr. Brock is a Registered Professional Engineer of the Province of Ontario, and a member of the Institute of Electrical and Electronic Engineers.

JOSEPH R. BURNS received the B.S. and M.S. degrees in Electrical Engineering from Princeton University in 1959 and 1962, respectively. Since 1959, he has been employed at the RCA Laboratories in Princeton, New Jersey where he has engaged in research on magnetic switching and high-speed digital logic circuits. Mr. Burns is a member of the Institute of Electrical and Electronics Engineers.

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.

FRANCIS K. FONG (see *RCA Review*, Vol. XXV, No. 2, June 1964, p. 324.)

E. C. GIAIMO served in the U.S. Navy as an Electronic Technician from 1943 to 1946. In this capacity, he was a laboratory assistant in microwave research at the Naval Research Laboratory, Washington, D. C., from 1944 to 1946. He received a B.S.E. in 1950 and an M.S.E. in 1955 in electrical engineering from Princeton University. He joined RCA Laboratories in 1951, where he made contributions to ruggedization of traveling-wave-tube structures, germanium purification and testing devices. He has been actively associated with electrophotographic research since 1952. From 1957 to 1959 he was assigned to the technical staff of C Stellarator Associates to perform system specification and coordination duties in the design of a thermonuclear fusion reactor. He was project engineer for a research program on the investigation of thermoplastic organic photoconductive recording media. Mr. Giaimo is a member of the American Association for the Advancement of Science, Sigma Xi, and a senior member of the IEEE.

J. M. HAMMER received his B.S. in Engineering Physics from New York University in 1950. In 1951 he received the M.S. in Physics from the University of Illinois. He received the Ph.D. in Physics from New York University in 1956. He taught physics at New York University from 1953 to 1956 and did research in the field of low-energy electron scattering from various atomic systems. From 1956 to 1959, he worked at the Bell Telephone Laboratories on microwave noise problems. In 1959, he joined RCA Laboratories as a member of the technical staff and has been working on low-noise microwave research and atomic interactions. Dr. Hammer is a senior member of the Institute of Radio Engineers and a member of the American Physical Society and Sigma Xi.

CHENG P. WEN received the B.S. and M.S. degrees in electrical engineering from the University of Michigan in 1956 and 1957, respectively. From 1956 to 1963, he was employed in the Electron Physics Laboratory of the University, and he received the Ph.D. degree in electrical engineering in 1963. In March 1963, he joined RCA Laboratories, Princeton, New Jersey, as a member of the Technical Staff. Dr. Wen is a member of Eta Kappa Nu, Tau Beta Pi and Sigma Xi.