THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XXXVI NOVEMBER 1957 NUMBER 6

A New Storage Element Suitable for Large-Sized Memory Arrays— The Twistor

By ANDREW H. BOBECK

Three methods have been developed for storing information in a coincident-current manner on magnetic wire. The resulting memory cells have been collectively named the "twistor". Two of these methods utilize the strain sensitivity of magnetic materials and are related to the century old Wertheim or Wiedemann effects; the third utilizes the favorable geometry of a wire.

The effect of an applied torsion on a magnetic wire is to shift the preferred direction of magnetization into a helical path inclined at an angle of 45° with respect to the axis. The coincidence of a circular and a longitudinal magnetic field inserts information into this wire in the form of a polarized helical magnetization. In addition, the magnetic wire itself may be used as a sensing means with a resultant favorable increase in available signal since the lines of flux wrap the magnetic wire many times. Equations concerning the switching performance of a twistor are derived.

An experimental transistor-driven, 320-bit twistor array has been built. The possibility of applying weaving techniques to future arrays makes the twistor approach appear economically attractive.

I. INTRODUCTION

A century ago Wiedemann[1] observed that if a suitable magnetic rod which carries a current is magnetized by an external axial field, a twist of the rod will result. The effect is a consequence of the resultant helical flux field causing a change in length of the rod in a helical sense. Conversely, it was also observed that a rod under torsion will produce a voltage between its ends when the rod is magnetized (see Fig. 1).

Recently, during an investigation of the magnetic properties of nickel wire, it was observed that a voltage was developed across the ends of a nickel wire as its magnetization state was changed. Both the amplitude and the polarity of the observed signal could be varied by movements of the nickel wire. Most surprising, the amplitude of the observed voltage v_2 of Fig. 2, was many times that which would be expected if a conventional pickup loop were used.

After determining experimentally that the observed voltage was generated solely in the nickel wire and was not a result of air flux coupling the sensing loop (nickel wire plus unavoidable copper return wire), it was concluded that the flux in the nickel wire must follow a helical path. This suggested that torsion was the cause of the observed effect, a conclusion verified experimentally. The direction of the applied twist de-

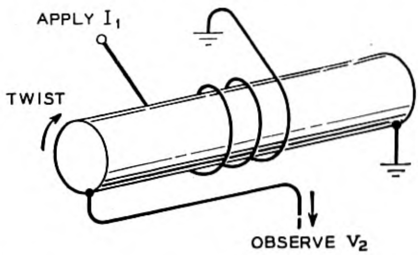


Fig. 1 — Observation of an internally induced voltage v_2 generated by a magnetic wire under torsion.

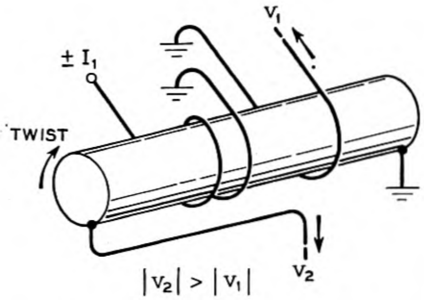


Fig. 2 — Comparison of the internally induced voltage v_2 to the voltage v_1 induced in the pickup loop.

termined the polarity and the amount of the twist determined the magnitude of the observed voltage.

As a consequence of these results, it is possible to build mechanical-to-electrical transducers,[2] transformers with unity turns ratio but possessing a substantial transforming action, and a variety of basic memory cells.

This paper will be concerned with a discussion of the memory cells from both a practical and theoretical viewpoint. It will be shown how these cells can be fabricated into memory arrays. One such configuration consists solely of vertical copper wires and horizontal magnetic wires. Experimental results of the switching behavior of many magnetic materials when operated in the "twisted" manner will be given.

II. A COINCIDENT-CURRENT MEMORY CELL — THE TWISTOR

Consider a wire rigidly held at the far end and subjected to a clockwise torsion applied to the near end. This will result in a stress component of maximum compression[3] at an angle of 45° with respect to the axis of the wire in the right-hand screw sense, and a component of maximum tension following a left-hand screw sense. All magnetic materials are strain sensitive to some degree. This will depend upon both the chemical composition and the mechanical working of the material. For example, if unannealed nickel wire is subjected to a torsion, the preferred direction of magnetization will follow the direction of greatest compression, as would be predicted from the negative magnetostrictive coefficient of nickel. Unannealed nickel wire, then, will have a preferred remanent flux path as shown in Fig. 3.

If the ease of magnetization as measured along the helix is sufficiently lower than that along the axis or circumference, it is possible to insert information into the wire in a manner somewhat analogous to the usual

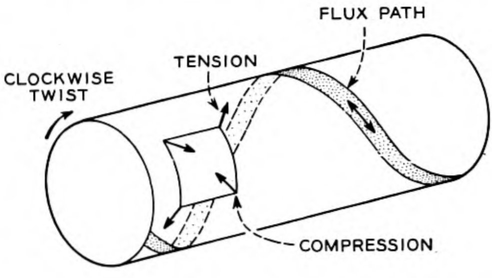


Fig. 3 — Relationship of the mechanical stresses resulting from applied torsion to the preferred magnetic flux path in nickel.

coincident-current method. Consider a current pulse I_1 applied through the nickel wire in such a direction as to enhance the spiraling flux, and a second current pulse I_2 applied by means of an external solenoid (see Fig. 4). Coincidentally, the proper amplitude current pulses will switch the flux state of the wire; either alone will not be sufficient. To sense the state of the stored information it is necessary either to reverse both currents, or to overdrive I_2 in the reverse direction. In an array, the output, in the form of a voltage pulse, would be sensed across the ends of the nickel wire. The solenoid may be replaced by a single copper conductor passing at right angles to the nickel wire. For obvious reasons the memory cell has been named the "twistor". The above method of operation will be referred to as mode A.

Mode B is the use of the magnetic wire as a direct replacement for the conventional coincident-current toroid. Its use here differs only in that the wire itself is used as a sensing winding (refer to Fig. 5). The pulses I_1 and I_2 are equal in value and each alone is chosen to be insufficient to switch the magnetization state of the wire. The coincidence of I_1 and I_2 will, however, result in the writing of a bit of information into the wire. To read, I_1 and I_2 are reversed in polarity and applied coincidently. The output appears as a voltage pulse across the ends of the nickel wire.

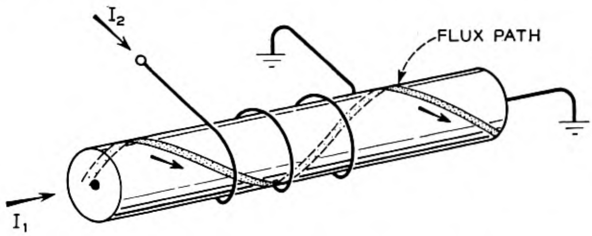


Fig. 4 — Coincident currents for the "write" operation in a twistor operated mode A. Wire under torsion.

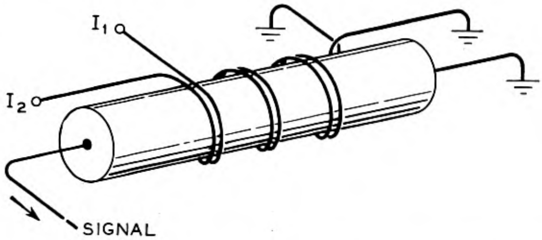


Fig. 5 — For mode B the coincidence of I_1 and I_2 is required to exceed the knee of the φ-NI characteristic. Wire under torsion.

The third method of operating the twistor, mode C, is similar in nature to a method proposed by J. A. Baldwin.[4] In this scheme the wire is not twisted, so that neither screw sense is favorable. By the proper application of external current pulses, information will be stored in the wire in the form of a flux path of a right-hand screw sense for a "1", and a left-hand screw sense for a "0". The operation of the cell is indicated in Figure 6. Note that the writing procedure requires a coincidence of currents; the reading procedure does not. The sign of the output voltage indicates the stored information.

Modes A and C are best suited for moderate sized memory arrays since the reading procedure is not a coincident type selection. Thus to gain access to n^2 storage points, an access switch capable of selecting one of n^2 points is required. For large arrays the use of mode B is indicated. It then becomes possible to select one of n^2 points with a $2n$ position access switch. The crossover point (about 10^5 bits) is determined by access circuitry considerations.

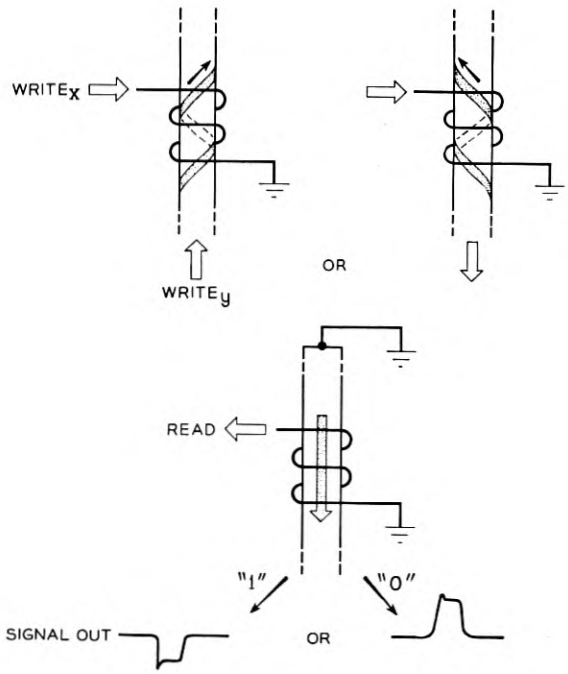


Fig. 6 — Read-write cycle for a twistor operated mode *C*. The wire is not under torsion.

III. ANALYSIS OF THE SWITCHING PROPERTIES OF THE TWISTOR

Section 3.1 will deal with the basic properties of magnetic wire as they pertain to the twistor memory cell. Section 3.2 will be concerned with a composite magnetic wire. The theoretical conclusions will be supported by experimental results wherever possible.

3.1 *Solid Magnetic Wire*

It has been stated above that there is a voltage gain inherent in the operation of the twistor. This voltage gain makes it possible to obtain millivolt signals from wires several mils in diameter. An expression will now be derived relating the axial flux of an *untwisted* wire to the circular flux component of that wire when twisted. Assume that the magnetic wire has been twisted so that the flux spirals at an angle θ (normally $\theta = 45°$) with respect to the axis of the wire. If d and l are the diameter and length of the magnetized region respectively, then, for a complete flux reversal the change in the *circular* flux component is $\varphi_{\text{circ}} = l(d/2)(2B_s \sin \theta)$.

Here, φ_{circ} is the flux change that would be observed on a hypothetical pickup wire which passed down the axis of the magnetic wire. The flux change which would be observed by a single pickup loop around the wire, if the magnetic wire were not twisted, is $\varphi_{\text{longititudinal}} = \pi d^2 B_s/2$. Therefore, $\varphi_{\text{circ}}/\varphi_{\text{long}} = 2\, l \sin \theta/\pi\, d$, and for $\theta = 45°$, this expression reduces to

$$\frac{\varphi_{\text{circ}}}{\varphi_{\text{long}}} = \frac{l/d}{2.22}. \tag{1}$$

Thus, for example, if the storage length on a 3-mil wire is 100 mils, then a 15:1 gain in flux change (or voltage) is obtained.

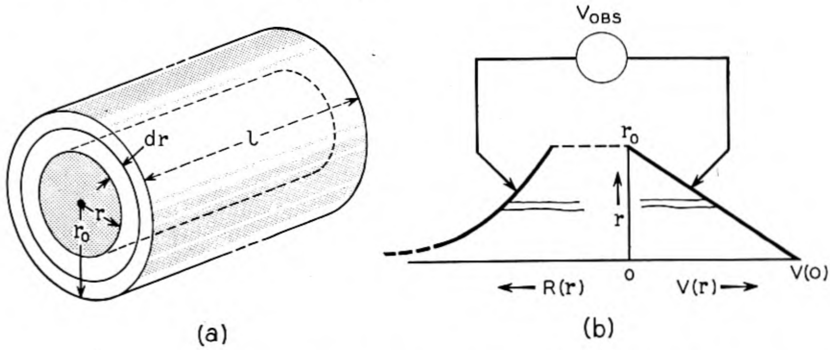


Fig. 7 — (a) Calculation of the observable voltage V_{obs} for a solid magnetic wire. (b) Diagram of induced voltage $V(r)$ and resistance $R(r)$.

The above derivation assumed that the entire circular flux change could be observed externally. Since the magnetic wire must, of necessity, serve as the source of the generated voltage the resultant eddy current flow reduces the *observable* flux change by a factor of three. Consider Fig. 7; assume that the flux reversal takes place in the classical manner and consider the circular component of this flux since it alone contributes to the observable signal. The induced voltage $V(r)$ at any point r is $V(r) = V(0)/(r_0 - r/r_0)$, where r_0 is the radius of the wire and $V(0)$ the voltage at $r = 0$. But $V(r)$, the induced or open-circuit voltage per length of wire l, could only exist if the wire were composed of many concentric tubes of wall thickness dr, each insulated from one another. In a long wire no radial eddy currents can exist. Therefore the wire of length l can be assumed to be faced by a perfect conductor at both of its ends. It remains to calculate the potential between these ends. The resistance of the tube is $R(r) = \rho l/2\pi r\, dr$, where ρ is the resistivity in ohm-cm. The resistance of the wire is given by $R_0 = \rho l/\pi r_0^2$. These resistances form a voltage divider on the *induced* voltage in the tube and the total contribution of all tubes is obtained by integration;

$$\begin{aligned} V_{\text{observed}} &= \int_0^{r_0} V(r) \left(\frac{\rho l/\pi r_0^2}{\rho l/2\pi r dr} \right) \\ &= \int_0^{r_0} V(0) \left(\frac{r_0 - r}{r_0} \right) \frac{2r}{r_0^2}\, dr \qquad (2) \\ &= \frac{V(0)}{3}. \end{aligned}$$

Thus, (1) must be modified by (2) with the voltage step-up per memory cell becoming,

$$\frac{V_{\text{obs}}}{V_{\text{long}}} = \frac{l/d}{6.66}. \qquad (3)$$

3.11 *Bulk Flux Reversal — Classical Case*

The switching performance of a magnetic wire under transient conditions will now be considered. The speed of magnetization reversal of magnetic materials under pulse conditions is best characterized by s_w, the switching coefficient, usually expressed in oersted-microseconds. It is defined as the reciprocal of the slope of the $1/T_s$ versus H curve where T_s is the time required to reverse the magnetization state and H is the applied magnetic field intensity. Only eddy current losses will be considered.

First, consider the case in which the magnetization is entirely circular. Reversal from one flux remanent state to the other is assumed to occur uniformly in time T_s. Axial eddy currents will flow down the center of the wire and return along the surface. The switching coefficient, s_w, will be obtained by equating the input energy to the dissipated energy. The total energy dissipated per unit length is,

$$\mathcal{E} = T_s \int_0^{r_0} 2\pi r P(r)\, dr, \tag{4}$$

and

$$P(r) = \frac{[E(r)]^2}{\rho}, \tag{5}$$

where $P(r)$ the power density is given by $E(r)$ the voltage gradient squared divided by the resistivity. The average energy per unit volume is therefore

$$\begin{aligned}\mathcal{E}_{\mathrm{av/cm^3}} &= \frac{T_s}{\pi r_0^2} \int_0^{r_0} \frac{\left(V(0)\left(\frac{r_0 - r}{r_0}\right) - \frac{V(0)}{3}\right)^2}{\rho l^2}\, 2\pi r\, dr \\ &= \frac{T_s}{18\rho}\left(\frac{V(0)}{l}\right)^2.\end{aligned} \tag{6}$$

Now $V(0) = [(dB/dt)r_0 l]10^{-8}$, so $V(0)/l = (2B_s r_0/T_s)10^{-8}$. Putting this expression into (6) yields

$$\mathcal{E}_{\mathrm{aver/cm^3}} = \frac{2B_s^2 r_0^2}{9\rho T_s}\, 10^{-16}. \tag{7}$$

The input energy per unit volume is

$$\mathcal{E}/\mathrm{cm}^3 = \frac{(2B_s H \cos\theta_1)\, 10^{-7}}{4\pi}, \tag{8}$$

since $\Delta\bar{B}\cdot\bar{H} = 2B_s H \cos\theta_1$ where θ_1 is the angle between the applied field H and the switching flux. The factor $10^{-7}/4\pi$ is a constant relating the energy in joules to the BH product in gauss-oersteds. By equating (7) and (8), and replacing H by $(H - H_0)$, the desired s_w expression is obtained;

$$s_w = T_s(H - H_0) = \frac{(4\pi B_s r_0^2)\, 10^{-3}}{9\rho \cos\theta_1}\ (\text{oe-}\mu\text{sec}). \tag{9}$$

The substitution of $H - H_0$ for H requires some explanation. The switching curve of $1/T_s$ versus H is not a straight line as would be predicted from (9) but generally possesses considerable curvature at low drives. Equation (9) satisfactorily predicts the slope of the switching curve in the high drive region, but H_0 must be determined experimentally. In Section 3.12, flux reversal by wall motion is treated as it is a possible switching mechanism at low drives.

The switching coefficient s_w for the case in which the magnetization is purely axial will now be treated. As above, the flux density will change from $-B_s$ to $+B_s$ uniformly in time T_s. The eddy currents, which are circular, result from an induced voltage $V(r)$ where $V(r) = [V(r_0)(r/r_0)^2]$, and $V(r_0)$ is given by $V(r_0) = [(2B_s/T_s)\pi r_0^2]10^{-8}$. Thus, $E(r) = V(r)/2\pi r$, and $E(r) = (B_s r/T_s)10^{-8}$. Following the procedure used above, the internally dissipated energy density is

$$\begin{aligned} \mathcal{E}_{\mathrm{av/cm}^3} &= \frac{T_s}{\pi r_0^2} \int_0^{r_0} \frac{(B_s r 10^{-8}/T_s)^2}{\rho} 2\pi r \, dr, \\ \mathcal{E}_{\mathrm{av/cm}^3} &= \left(\frac{B_s^2 r_0^2}{2T_s \rho} \right) 10^{-16}. \end{aligned} \tag{10}$$

Equating this expression to (8) yields

$$s_w = (H - H_0) T_s = \frac{\pi B_s r_0^2 10^{-3}}{\rho \cos \theta_2}, \tag{11}$$

where θ_2 is defined as the angle between the applied field and the switching flux now assumed axial.

The helical flux vector in a twistor can be resolved into a circular and an axial component. Fortunately, since the dissipated energy is proportional to the eddy current density squared, and the axial and circular current density vectors are perpendicular to each other, it is possible to write

$$\mathcal{E}_{\mathrm{av/cm}^3}(\text{helical}) = \mathcal{E}_{\mathrm{av/cm}^3}(\text{axial}) + \mathcal{E}_{\mathrm{av/cm}^3}(\text{circular}). \tag{12}$$

It follows, for a 45 degree pitch angle, that

$$s_w(\text{helical}) = \frac{s_w(\text{axial}) + s_w(\text{circular})}{2} \tag{13}$$

where the factor "2" is a consequence of the flux density components being smaller by $1/\sqrt{2}$ than their resultant. Substitution of (9) and (11) into (13) gives the desired switching coefficient

$$s_w(\text{helical}) = (H - H_0) T_s = \frac{13\pi}{18} \frac{B_s r_0^2 10^{-3}}{\rho \cos \theta}. \tag{14}$$

The term $\cos \theta$ requires further explanation. The magnetization vector is constrained by energy considerations to align with the easy direction of magnetization. The angle between the applied field and the easy direction of magnetization is called θ. Equation (14) is valid for any direction of applied field. The angles θ_1 and θ_2 used in deriving (9) and (11) respectively are each 45 degrees for the helical pitch angle assumed above.

Equation (14) indicates that for maximum switching speed a material with low saturation flux density and high resistivity is required. The lower limit on s_w will be determined by internal loss mechanisms not treated here. Experimentally, this lower bound is found to be approximately 0.2 oe-μsec.

3.12 *Reversal by Single Wall*

The switching time of a twistor when operating in a memory array under coincident current conditions will depend upon the low-drive switching coefficient. Experimentally, it is observed that the low drive s_w is several times the high-drive value. In this section, following the method of Williams, Shockley, and Kittel,[5] flux reversal by the movement of a single wall will be treated. Only the circular flux case will be considered.

The technique used to obtain s_w is identical to that used in Section 3.1.1 except it is postulated that a single wall concentric to the wire moves either from the wire surface inward, or from the wire axis outward. The result is independent of the direction in which the wall moves. Assume the wall moves from $r = 0$ to $r = r_0$, as indicated in Fig. 8. In-

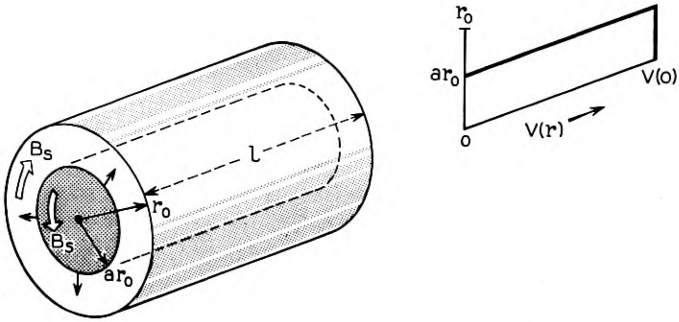


Fig. 8 — Flux reversal by expanding wall instantaneously located at radial position ar_0 moving with velocity v.

stantaneously, the wall is located at ar_0 and is traveling with a velocity v. The induced voltage $V_{oc}(r)$ is,

$$\begin{aligned} V_{oc}(r) &= (2B_s lv)10^{-8} & 0 < r < ar_0 , \\ &= 0 & ar_0 < r < r_0 , \end{aligned} \tag{15}$$

and the observable voltage for a wire of length l is given by

$$\begin{aligned} V_{sc} &= \int_0^{r_0} V_{oc}(r) \frac{\rho l/\pi r_0^2}{\rho l/2\pi r dr} \\ &= \int_0^{ar_0} V_{oc}(r) \frac{2r}{r_0^2} dr \\ &= a^2 V_{oc}(r). \end{aligned}$$

It is clear in the above integration that $V_{oc}(r)$ must be treated as a constant. Using expressions (4) and (5)

$$\frac{\partial \mathcal{E}_{\mathrm{av/cm^3}}}{\partial t} = \frac{1}{\pi r_0^2 \rho} \int_0^{ar_0} \left(\frac{V_{oc}^2}{l} \right) (1 - a^2)^2 2\pi r dr + \frac{1}{\pi r_0^2 \rho} \int_{ar_0}^{r_0} \left(\frac{V_{oc}}{l} \right)^2 (0 - a^2)^2 2\pi r dr, \tag{16}$$

$$\frac{\partial \mathcal{E}_{\mathrm{av/cm^3}}}{\partial t} = \left(\frac{V_{oc}}{l} \right)^2 \frac{(1 - a^2)a^2}{\rho}.$$

The rate of applying energy is

$$\frac{\partial \mathcal{E}}{\partial t} = \left(\frac{\partial B}{\partial t} \frac{(H - H_c) \cos \theta}{4\pi} \right) 10^{-7}. \tag{17}$$

Once again hysteresis losses are not included. Since $\partial B/\partial t = 2B_s v r_0$, and V_{oc} is given by (15), the equation of (16) and (17) yields,

$$(1 - a^2)a^2 v = \frac{\rho(H - H_c) \cos \theta}{8\pi B_s r_0}.$$

Since $v = (dr/dt) = r_0 (da/dt)$,

$$\int_0^a a^2(1 - a^2)\, da = \frac{\rho(H - H_c) \cos \theta}{8\pi B_s r_0^2} \int_0^t dt,$$

$$\frac{a^3}{3} - \frac{a^5}{5} = \frac{\rho(H - H_c) \cos \theta}{(8\pi B_s r_0^2)10^{-9}} (t + K).$$

When a equals 0, t equals 0, so the constant of integration is zero. When a equals 1, t equals T_s, so

$$s_w = T_s(H - H_c) = \left(\frac{16\pi}{15} \frac{B_s r_0^2}{\rho \cos \theta}\right) 10^{-3}(\text{oe-}\mu\text{sec}) \tag{18}$$

Comparison to the corresponding bulk flux reversal case indicates that the wall motion mechanism is more lossy by a factor of 2.4.

3.2 *The Composite Magnetic Wire*

It is apparent from the switching data of Fig. 9 that for reasonably sized solid wires ($r_0 > 1$ mil) the switching coefficient s_w is unreasonably high. The typical ferrite memory toroid, for example, when used as a memory element has an s_w of 0.6 oersted-microseconds. The only possibility for high speed coincident-current operation for solid magnetic wires is that the material have a high coercive force H_c, a conclusion not consistent with the trend toward transistor driven memory systems.

By the use of a composite wire it should be possible to reduce the eddy current losses and still preserve a reasonable wire diameter. A composite wire, by definition, will consist of a non-magnetic inner wire clad with a magnetic skin. It may be fabricated by a plating or an extruding process.

The solid wire analysis of Section 3.1 is a special case of composite

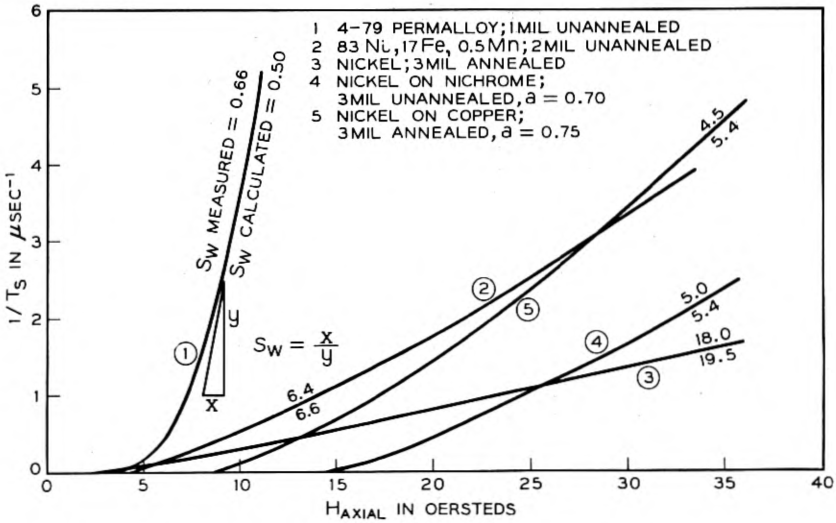


Fig. 9 — Reciprocal of flux reversal time T_S as a function of applied axial drive, H, for solid and composite magnetic wires. Sufficient torsion applied to reach saturation.

wire analysis which is given in Appendix I. Only the results of the composite wire case will be given here. As indicated in Fig. 10, ρ_1 and ρ_2

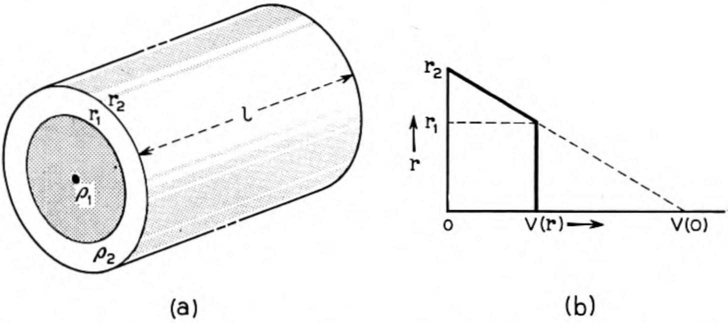


Fig. 10 — (a) Composite wire is composed of non-magnetic core covered by a magnetic skin. (b) A voltage $V(r)$ is induced in the wire during flux reversal.

are the resistivities of the inner (non-magnetic) and outer (magnetic) materials. The inner material is contained within a radius r_1. The overall wire radius is r_2. Defining $a = r_1/r_2$, if V_{obs} is the voltage observed across the ends of the composite wire twistor memory cell, and $V(0)$ is the induced voltage at $r = 0$ for a *solid* magnetic wire of radius r_2, $V_{obs} = bV(0)$, and

$$b = \frac{\frac{\rho_1}{\rho_2}\left(\frac{1}{3} - a^2 + \frac{2}{3}a^3\right) + a^2 - a^3}{\frac{\rho_1}{\rho_2}(1 - a^2) + a^2}. \tag{19}$$

The parameter "b" reduces to $\frac{1}{3}$ for $a = 0$ in agreement with (2) which was derived for the solid wire case. Table I gives "b" for various material and geometry combinations.

TABLE I — THE PARAMETER "b"

ρ_1/ρ_2	a			
	0.9	0.8	0.7	0.0 (Solid Wire)
0	0.100	0.200	0.300	1.000
$\frac{1}{10}$	0.0988	0.194	0.285	.333
$\frac{1}{3}$	0.0963	0.184	0.259	.333
1	0.0903	0.163	0.219	.333
3	0.0790	0.135	0.180	.333
10	0.0643	0.112	0.155	.333
∞	0.0491	0.0963	0.141	.333

TABLE II — THE PARAMETER "C_{circ}"

ρ_1/ρ_2	a			
	0.9	0.8	0.7	0.0 (Solid Wire)
0	0.0858	0.353	0.820	1.396
$\frac{1}{10}$	0.0847	0.339	0.763	1.396
$\frac{1}{3}$	0.0842	0.311	0.657	1.396
1	0.0737	0.256	0.498	1.396
3	0.0595	0.159	0.341	1.396
10	0.0286	0.124	0.243	1.396
∞	0.0209	0.0833	0.186	1.396

The switching coefficient, s_w, for circular flux reversal, is derived as

$$s_w = \frac{(8\pi B_s r_2^2)10^{-3}}{\rho_2(1 - a^2)\cos\theta}\left\{a^2\left[(1 - a - b)^2\frac{\rho_2}{\rho_1} - (1 - b)^2 + \frac{4a(1 - b)}{3} - \frac{a^2}{2}\right] + (1 - b)^2 - \frac{4(1 - b)}{3} + \frac{1}{2}\right\}, \tag{20}$$

or

$$s_w = C_{\text{circ}}\frac{(B_s r_2^2)10^{-3}}{\rho_2 \cos\theta}\ (\text{oe-}\mu sec). \tag{21}$$

Table II gives C_{circ} as a function of a and ρ_1/ρ_2.

The switching coefficient s_w for axial flux reversal is derived as

$$s_w = \left(\frac{\pi B_s r_2^2 10^{-3}}{\rho_2 \cos\theta}\right)\left(\frac{1 - 4a^2 + a^4(3 - 4\ln a)}{1 - a^2}\right) \tag{22}$$

$$= C_{\text{axial}}\left(\frac{B_s r_2^2 10^{-3}}{\rho_2 \cos\theta}\right). \tag{23}$$

Table III gives C_{axial} as a function of "a".

Since, as explained for the solid wire case, the eddy current density vectors for circular and axial flux reversal are in quadrature,

$$s_w(\text{helical}) = \frac{s_w(\text{axial}) + s_w(\text{circular})}{2}. \tag{24}$$

Substitution of (21) and (23) into (24) gives the required expression;

$$s_w(\text{helical}) = \left(\frac{C_{\text{circ}} + C_{\text{axial}}}{2}\right)\frac{(B_S r_2^2)10^{-3}}{\rho_2 \cos\theta}. \tag{25}$$

A number of composite wire samples have been prepared and evaluated. These include nickel on nichrome and nickel on copper. The switch-

TABLE III — THE PARAMETER "C_{axial}"

a	0.9	0.8	0.7	0.0 (Solid Wire)
C_{axial}	0.0795	0.300	0.633	3.142

ing curves for these samples as well as for a number of solid magnetic wires are shown in Fig. 9. The agreement between the measured values of s_w and the calculated values [(14) and (25)] is quite good. Improved composite wire samples are under development.

IV. EXPERIMENTAL MEMORY CELLS AND ARRAYS

The initial experiments were performed using commercially available nickel wire of 3-mil diameter. The φ-NI characteristic of this wire in the helical direction is extremely square. This is a feature of *all* the magnetic materials tested whether annealed or unannealed. As a typical example,

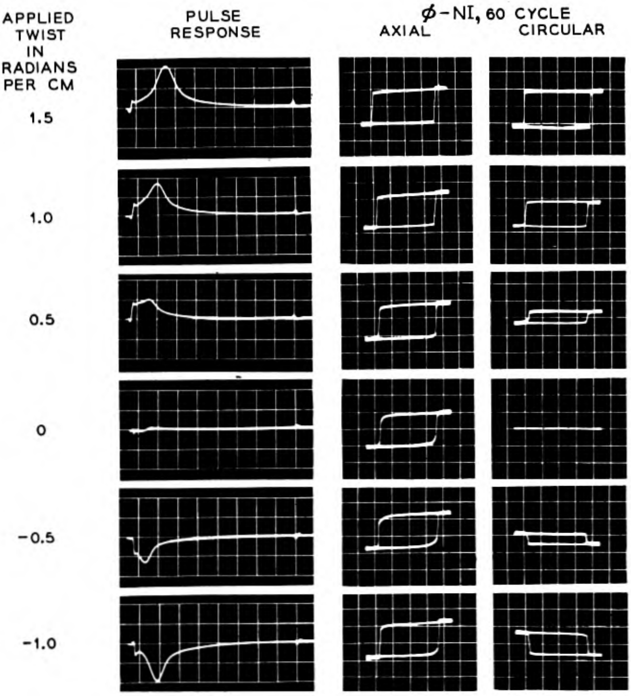


Fig. 11 — Sixty cycle and switching waveforms for 83 Ni, 17 Fe wire (see Fig. 9) as a function of applied torsion.

the 60-cycle characteristics of the axial and circular flux versus axial drive and the switching voltage waveforms under pulse conditions are given in Fig. 11 for 2-mil wire of composition 83 Ni, 17 Fe.

Note the negative prespike on the switching waveforms. By simultaneous observation of both the axial and circular switching voltage waveforms on many different magnetic wires it has been concluded that the negative prespike is due to an initial coherent rotation of the magnetization vector which results in an initial *increase* in the circular flux component. It is during this coherent rotation that the normal positive prespike on the axial switching voltage waveform is observed. Because of the mechanically introduced strain anisotropy, however, the magnetization vector is constrained to remain nearly parallel to the easy direction of magnetization. Thus, the coherent rotation soon ceases and the re-

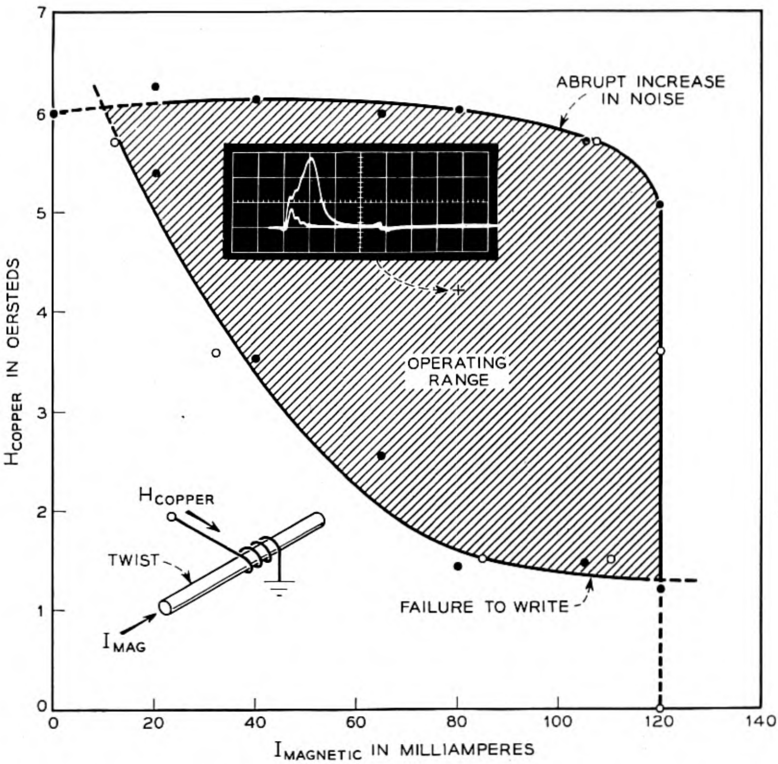


Fig. 12 — Range of writing currents for 83 Ni, 17 Fe wire operated mode. A Read drive held constant at 9 oersteds. Typical signal-to-noise ratio for a read is indicated.

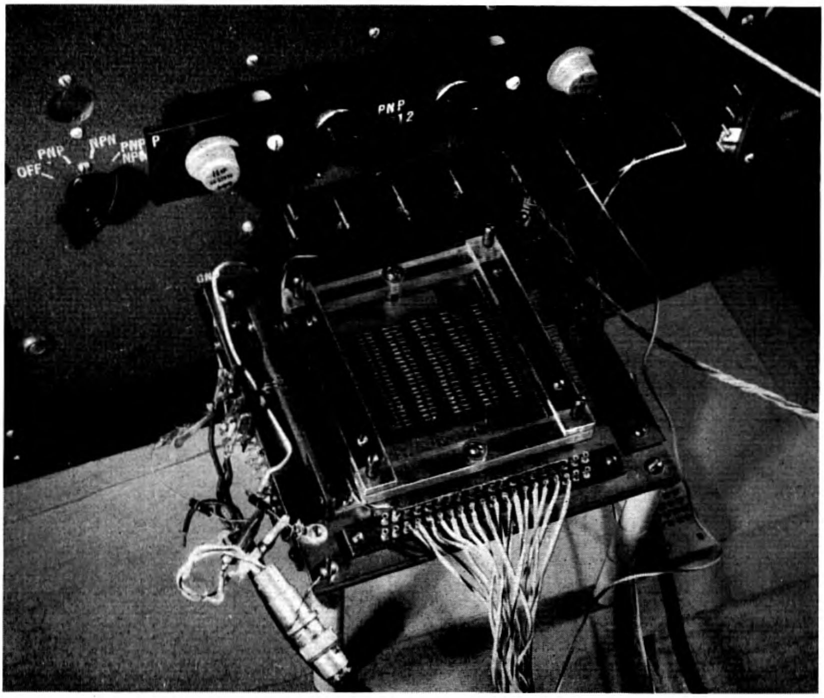


Fig. 13 — A 320-bit experimental twistor memory array. The array is transistor driven.

mainder of the flux reversal process is by an incoherent rotational process. During this latter time the circular and axial voltage waveforms are virtually identical.

Fig. 12 gives the range of operation of 2-mil 83 Ni, 17 Fe wire as a twistor operated by mode A. As a result of the extreme squareness of the φ-NI characteristic in helical direction the range of operation encloses an area nearly the theoretical maximum. The switching times of other memory cells tested ranged from 0.2 μsec for a 1 mil 4-79 molypermalloy wire to 20 μsec for a 5 mil perminvar wire. Thus it is seen that the switching speeds of the twistor compare quite favorably with those of conventional ferrite toroids and sheets.

It is, of course, possible to store many bits of information along a single magnetic wire. The allowable number of bits per inch is related to the coercive force, the saturation flux density, and the diameter of the wire. For the nickel wire, about 10 bits per inch are possible. Predictions as to the storage density for a given material can be made by referring to suitable demagnetization data. There are, however, interference effects between cells which are not completely understood at the present time.

A memory array (16 × 20) has been constructed as a test vehicle. An illustration of this array is shown in Fig. 13. The drive wires have been woven over glass tubes which house the removable magnetic wires. Provision is made for varying the torsion and the tension of the individual magnetic wires.

As an indication of the performance of the twistor, Fig. 14 is a composite photograph showing the minimum and maximum signal over the 16 bits of a given column for 3-mil nickel wire. Also included are the noise pulses for these cells, the so-called disturbed zero signals. The write currents were 2.3 ampere-turns on the solenoid and 130 ma through the magnetic wire. The read current was 6.0 ampere-turns. The array

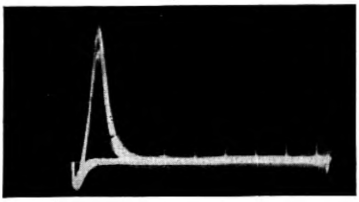


Fig. 14 — Composite photograph of the 16 output signals from a column of the array of Fig. 13. Average output signal about 3.5 millivolts; sweep speed equals 2 μsec/cm.

was transistor driven. A read-write cycle time of 10 microseconds appeared to be possible.

V. DISCUSSION

The twistor is presented as a logical companion to the coincident-current ferrite core and sheet.[6, 7] In many applications it should compete directly with its ferrite equivalents. Perhaps its greatest use will be found in very large ($>10^6$) memory arrays.

From a cost per bit viewpoint the future of the twistor appears quite promising. Fabricating and testing the wire should present no special problems as it is especially suited for rapid, automatic handling. The possibility of applying weaving techniques to the construction of a twistor matrix looks promising.

It is possible that, for both mode A and C operation of the twistor, an array can be built which consists simply of horizontal copper wires and vertical magnetic wires — much like a window screen. Preliminary experiments have shown that single cross wires do operate successfully. The operation of this array would be analogous to a core memory array. Physically it could look just like a core array — but without the cores.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the help of R. S. Title in obtaining the memory array data, R. A. Jensen in constructing the test jigs, and D. H. Wenny, Jr. in supplying the wire samples. The discussions with my associates, in particular D. H. Looney, R. S. Title, and J. A. Baldwin, have been very helpful. The writer would like to acknowledge the interest and encouragement shown by R. C. Fletcher.

APPENDIX I

From Fig. 10, for bulk circular flux reversal in a composite wire, the induced voltage $V(r)$ for a wire length l is

$$\begin{aligned} V(r) &= \left[(r_2 - r_1)l\left(\frac{2B_s}{T_s}\right)\right]10^{-8}, && 0 < r < r_1, \\ &= \left[(r_2 - r)l\left(\frac{2B_s}{T_s}\right)\right]10^{-8}, && r_1 < r < r_2. \end{aligned} \tag{26}$$

For a *solid* magnetic wire of radius r_2

$$V(0) = \left(\frac{2B_s r_2 l}{T_s}\right)10^{-8}. \tag{27}$$

Therefore,

$$\begin{aligned} V(r) &= \left(\frac{r_2 - r_1}{r_2}\right)V(0), && 0 < r < r_1, \\ V(r) &= \left(\frac{r_2 - r}{r_2}\right)V(0), && r_1 < r < r_2. \end{aligned} \tag{28}$$

In general, the resistance of a tube of wall thickness dr is $R(r) = \rho l/$

$2\pi r dr$. The resistance of the wire is $R_T = \rho_1\rho_2 l/\pi[\rho_1(r_2^2 - r_1^2) + \rho_2 r_1^2]$. The observable voltage for a length of wire, l, is

$$V_{\text{obs}} = \int [V(r)] \qquad \text{(Volt-Divider)}$$

$$= \int_0^{r_1} \left(\frac{r_2 - r_1}{r_2}\right) V(0) \left(\frac{\dfrac{\rho_1\rho_2 l}{\pi[\rho_1(r_2^2 - r_1^2) + \rho_2 r_1^2]}}{\dfrac{\rho_1 l}{2\pi r\, dr}}\right)$$

$$+ \int_{r_1}^{r_2} \left(\frac{r_2 - r}{r_2}\right) V(0) \left(\frac{\dfrac{\rho_1\rho_2 l}{\pi[\rho_1(r_2^2 - r_1^2) + \rho_2 r_1^2]}}{\dfrac{\rho_2 l}{2\pi r\, dr}}\right).$$

This reduces to

$$V_{\text{obs}} = V(0) \frac{(\frac{1}{3} - a^2 + \frac{2}{3}a^3)\rho_1/\rho_2 + a^2 - a^3}{(1 - a^2)\rho_1/\rho_2 + a^2} \tag{29}$$

$$= bV(0), \tag{30}$$

where $a = r_1/r_2$ and b is given by reference to (29). The ratio $b/\frac{1}{3} = 3b$ is the relative efficiency of the composite as compared to the solid magnetic wire from an available signal viewpoint. An expression for s_w will now be derived.

The total energy dissipated per unit length l is

$$\xi_T/l = T_s \int_0^{r_2} \rho i_d^2(r) 2\pi r\, dr, \tag{31}$$

where $i_d(r)$ is the current density. Now, $i_d(r) = V(r) - V_{\text{obs}}/\rho l$, therefore

$$\begin{aligned} i_d(r) &= (1 - a - b)\frac{V(0)}{\rho_1 l}, && 0 < r < r_1, \\ &= \left(1 - \frac{r}{r_2} - b\right)\frac{V(0)}{\rho_2 l} && r_1 < r < r_2. \end{aligned} \tag{32}$$

The substitution of (32) into (31) yields, after manipulations,

$$\begin{aligned} \xi_{\text{av}/l} = \frac{\pi T_s r_2^2}{\rho_2}\left(\frac{V(0)}{l}\right)^2 \bigg\{ a^2 \bigg[(1 - a - b)^2 \frac{\rho_2}{\rho_1} - (1 - b)^2 \\ + \frac{4a(1 - b)}{3} - \frac{a^2}{2} \bigg] + (1 - b)^2 - \frac{4(1 - b)}{3} + \frac{1}{2} \bigg\}. \end{aligned} \tag{33}$$

From (8), the applied energy per wire length l is

$$\xi/l = \left(\frac{[B_s(H - H_0) \cos \theta]10^{-7}}{2\pi}\right) \pi(r_2^2 - r_1^2), \tag{34}$$

where only that part of the applied energy associated with the high drive dynamic losses is included. Equating (33) and (34) and replacing $V(0)/l$ by (27) results in

$$\begin{aligned} s_w = (H - H_0)T_S = \frac{8\pi B_S r_2^2 10^{-3}}{\rho_2(1 - a^2) \cos \theta} \Bigg\{ a^2 \Bigg[(1 - a - b)^2 \frac{\rho_2}{\rho_1} \\ - (1 - b)^2 + \frac{4a(1 - b)}{3} - \frac{a^2}{2} \Bigg] + (1 - b)^2 - \frac{4(1 - b)}{3} + \frac{1}{2} \Bigg\}. \end{aligned} \tag{35}$$

This can be expressed as

$$s_w = (H - H_0)T_s = C_{\text{circ}} \frac{(B_s r_2^2)10^{-3}}{\rho_2 \cos \theta} \text{ (oe-}\mu\text{sec).} \tag{36}$$

Bulk axial flux reversal in a composite magnetic wire can be treated in a manner analogous to that used in Section 3.1.1 for the solid wire. The uniform reversal of the axial flux induces a voltage $V(r)$ in the wire where

$$\begin{aligned} V(r) &= V(r_2) \left[\left(\frac{r}{r_2}\right)^2 - \left(\frac{r_1}{r_2}\right)^2 \right], && r_1 < r < r_2, \\ &= 0, && 0 < r < r_1, \end{aligned}$$

and $V(r_2) = [(2B_s/T_s)\pi r_2^2]10^{-8}$. Since $E(r) = V(r)/2\pi r$,

$$E(r) = \frac{B_S}{T_S}\left(r - \frac{r_1^2}{r}\right)10^{-8}. \tag{37}$$

Following the procedure of Section 3.1.1.,

$$\begin{aligned} \xi_{\text{av/cm}^3} &= \frac{T_S 10^{-16}}{\pi(r_2^2 - r_1^2)} \int_{r_1}^{r_2} \frac{B_S^2}{\rho_2 T_S^2} \left(r - \frac{r_1^2}{r}\right)^2 2\pi r\, dr \\ &= \left\{ \frac{2B_S^2 r_2^2 10^{-16}}{\rho_2 T_S} \left[\frac{\frac{1}{4} - a^2 + a^4(\frac{3}{4} - \ln a)}{1 - a^2} \right] \right\}, \end{aligned} \tag{38}$$

where $a = r_1/r_2$ as before. Equating this expression to (8) yields

$$s_w = (H - H_0)T_S = \frac{\pi B_S r_2^2 10^{-3}}{\rho_2 \cos \theta} \left[\frac{1 - 4a^2 + a^4(3 - 4 \ln a)}{1 - a^2} \right] \tag{39}$$

$$= C_{\text{axial}} \left(\frac{B_S r_2^2 10^{-3}}{\rho_2 \cos \theta} \right) \text{ (oe-}\mu\text{sec).} \tag{40}$$

REFERENCES

1. R. M. Bozorth, *Ferromagnetism*, D. Van Nostrand Company, Inc., New York, N. Y., 1951, p. 628.
2. U. F. Gianola, Use of Wiedemann Effect for Magnetostrictive Coupling of Crossed Coils, T. Appl. Phys., **26,** Sept. 1955, pp. 1152–1157.
3. H. F. Girvin, *Strength of Materials*, International Textbook Co., Scranton, Pa., 1944, p. 233.
4. J. A. Baldwin, unpublished report.
5. H. J. Williams, W. Shockley, and C. Kittel, Velocity of a Ferromagnetic Domain Boundary, Phys. Rev., **80,** Dec. 1950, pp. 1090–1094.
6. J. A. Rajchman, Ferrite Apertured Plate for Random Access Memory, Proc. I.R.E., **45,** (March, 1957), pp. 325–334.
7. R. H. Meinken, A Memory Array in a Sheet of Ferrite, presented at Conference on Magnetism and Magnetic Materials, Boston, Mass., Oct. 16–18, 1956.

Non-Binary Error Correction Codes*

By WERNER ULRICH

(Manuscript received April 19, 1957)

If a noisy channel is used to transmit more than two distinct signals, information may have to be specially coded to permit occasional errors to be corrected. If pulse amplitude modulation is used, the most probable error is a small one, e.g., 6 is changed to 7 or 5. Codes for correcting single small errors, and for correcting single small errors and detecting double small errors, in a message of arbitrary length, for an arbitrary number of different signals in the channel, are derived in this paper.

For more specialized situations, the error is not necessarily restricted to a small value. Codes are derived for correcting any single unrestricted error in a message of arbitrary length for an arbitrary number of different signals.

Finally, a set of codes based partially upon the Reed-Muller codes is described for correcting a number of errors in a more restricted class of message lengths for an arbitrary number of different signals.

The described codes are readily implemented. Many techniques are used which have an analog in a binary system. Other techniques are broadly analogous to binary coding techniques or are special adaptations of a binary code.

I. INTRODUCTION

1.1 *Use of Error Correction Codes*

One function of an error correction code is to aid in the correct transmission of digital information over a noisy channel. This process is illustrated in Fig. 1. An information source gives information to an encoder; the encoder converts the information into a message containing sufficient redundancy to permit the message to be slightly mutilated by the noisy channel and still be correctly interpreted at the destination. The message is then sent via the noisy channel to a decoder which will

* This paper was submitted to Columbia University in partial fulfillment of the requirements for the degree of Doctor of Engineering Science in the Faculty of Engineering.

reconstruct the original information if the mutilation has not been excessive. Finally, the information is sent to an information receptor.

One scheme for correcting errors in a binary system is to send each binary digit of information three times and to accept at the receiver that value which is represented by two or three of the received digits. Then, the encoder is simply an instrument for causing each digit to be sent three times, and the decoder consists of a majority organ. However, many methods are available which are considerably more elegant, and which will permit more information to be passed through a noisy channel in a given unit of time. This paper will deal with such methods for channels capable of sending b different symbols instead of the usual 1 and 0 of a binary channel.

The most convenient explanation of an error correction code has been made with respect to the transmission of correct digital information over a noisy channel. This does not imply the restriction of such codes

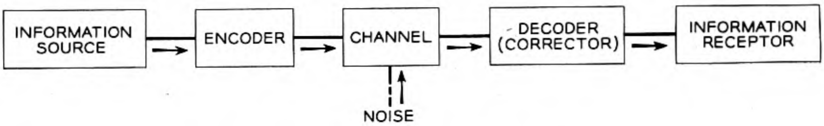


Fig. 1 — Transmission over a noisy channel.

to the noisy channel problem exclusively. Actually, the first application considered for such a code was with respect to computers.[1] Many large high speed computers stop whenever an error is detected in some calculation and must be restarted; with the use of an error correction code this could be avoided by permitting the computer to correct its own random errors directly. To the best knowledge of the author, error correction codes have not yet been used in any major computer. But the storage system of a computer may, in the future, lend itself to the use of error correction codes.

Frequently, very elaborate precautions must be taken in present storage systems to insure that they are free from errors. Magnetic tapes must be specially made and handled to guarantee the absence of defects, magnetic cores must be carefully tested to make sure that no defective cores get into an array, cathode ray tubes used in Williams Tube or Barrier Grid Tube storage systems must be perfect. Probably, there are other storage methods whose development is hampered because of a common requirement for error-free performance in all storage locations. With the use of error correction codes, such storage systems could be used, if they are sufficiently close to perfection, even though not perfect.

It is not unlikely that the near future will see the development of storage systems which will be able to store more than two states at every basic storage location.[2] If such systems are developed, it seems likely that they will be more erratic or noisy than binary storage systems, since each location must store one of b signals instead of one of two. If a cathode ray tube storage system were used, for example, different quantities of charge would have to be distinguished; in a binary storage system, only the presence or absence of charge must be detected. This suggests that error correction codes may become essential with certain types of non-binary storage systems. One object of this paper is to develop codes for this purpose and to discover which number systems are most easily correctable.

Some investigations have been made on the use of computer systems using multi-state elements.[3] A switching algebra has been developed similar to Boolean algebra for handling switching problems in terms of multi-state elements. Single device ring counters (the cold cathode gas stepping tube for example) already exist and might be useful in such systems. But currently, only limited steps in this direction have been made. Another object of this paper is to show the advantages and problems of error correction codes in multi-state systems; it is not unreasonable to predict that error correction codes may be more necessary in multi-state systems than in binary systems.

1.2 *Geometric Concept of Error Correction Codes*

A geometric model of a code was suggested by R. W. Hamming[1] which can be altered slightly to fit the non-binary case. For an n digit message, a particular message is a point in n dimensional space. A single error, however defined, will change the message, and will correspond to another point in n dimensional space. The distance between the original point and the new point is considered to be unity. Thus, the distance d between the points corresponding to any two messages is defined as the minimum number of errors which can convert the first message into the second.

With an error detection and/or correction code, the set of transmitted messages is limited so that those which are correctly received are recognizable; those messages which are received with fewer than a given number of errors are either corrected or the fact that they are wrong is recognized and some other appropriate action (such as stopping a computer) is taken.

In the case of binary codes, an error changes a 1 to 0 or a 0 to 1. In the non-binary case, two definitions of an error are possible and will be used in this paper. A *small* error changes a digit to an adjacent value. In a decimal system, a change from 1 to 2 or 1 to 0 is a small error. An *unrestricted* error changes a digit to any other value. In a decimal system, a change from 1 to 5 is an unrestricted error.

1.3 *Material To Be Presented*

The various types of codes described in this paper and the sections in which they are to be found are summarized in Table I. The techniques which are described are summarized below.

The geometric model suggests the simplest approach to error correction codes. A transmitter has a "codebook" containing all members of the set of transmitted messages. If the message source gives to the encoder the signal that the information to be sent is k (that is to say, the kth

TABLE I — TYPES OF CODES

Type of Code	Distance	Type of Error	Described in Section
Single Error Detection	2	Small and Unrestricted	II
Single Error Correction	3	Small	III and 6.1
Single Error Correction	3		
Prime Number Base		Unrestricted	4.1
Composite Number Base		Unrestricted	4.2
Single Error Correction and Double Error Detection	4	Small	V and 6.1
Multiple Error Correction	—	Small	6.2

output of all the outputs associated with the message source), the encoder chooses the kth member of the set. The decoder will then look up the message it receives in its own codebook which contains all possible received messages, and corresponding to the entry of the received message will find the symbols corresponding to k. Or the receiver may compare the received message with every member of the set of transmitted messages, calculate the distance between the two, and correct the received message to whichever of the transmitted messages is separated from the received message by the smallest distance. (It has been shown by Slepian[4] that this is the message most likely to be correct in a symmetrical binary channel having the property that changes from 1 to 0 and from 0 to 1 as a result of noise in the channel are equally likely.)

The practical difficulty with such a code is the large size of the required codebooks. Most coding schemes try to eliminate such codebooks and substitute a set of rules for encoding, decoding and correcting messages.

One approach toward creating a simple association between the information and the message is to use some of the digits of the message for conveying information directly. The Hamming Code[1] uses this technique.

An *information digit* is a digit of a message that is produced directly by the information source; in a base b code, an information digit may have b different values, the choice between these values representing the information that is to be sent.

A *check digit* is a digit of a message that is calculated as a function of the information digits by the encoder. It is sometimes convenient to represent or calculate a check digit in terms of a recursive formula using previously calculated check digits as well as information digits. In a base b code, a check digit may have b *check states*. When more than one check digit is used, each different combination of check digits corresponds to a different check state for the message; a message with m check digits will have b^m *message check states*.

A *systematic*[5] *code* encoder generates messages containing only information digits and check digits. The information source generates only base b information digits. The Hamming Code is a systematic code.

Section II offers a general method for obtaining single error detection codes for both small and unrestricted errors. The idea of mixed digits (digits which are, in a sense, neither information nor check digits, but a combination of both) is introduced, and it is shown how mixed digits may lead to more efficient coding systems. This idea is believed to be novel. Code systems which use mixed digits are called semi-systematic codes. Semi-systematic codes are used extensively throughout this paper.

Section III offers a general method for obtaining single small error correction codes, including both systematic and semi-systematic codes.

Section IV offers a general method for obtaining the more complicated single unrestricted error correction codes. The problem is divided into two parts. Section 4.1 describes codes for correcting single unrestricted errors in case b, the base of the channel, is a prime number.* Section 4.2 describes a special technique for obtaining the more complex codes for correcting single unrestricted errors in the event b is a composite number.

Section V offers a general method for obtaining semi-systematic codes for correcting single small errors and detecting double small errors. No general solution has been found for obtaining single error correction or double error detection codes for the case of unrestricted errors. No gen-

* This class of codes was previously described in a brief summary by Golay.[6]

eral solution has been found for multiple error correction codes for the unrestricted error case.

In Section VI, a number of techniques are presented for using binary error correction coding schemes for non-binary error correction codes. Section 6.1 shows how such techniques may be used to obtain non-binary single error correction codes, and single error correction double error detection codes, for the small error case. Section 6.2 presents a special technique, involving the use of an adaptation of the Reed-Muller binary code, to obtain a class of non-binary multiple error correction codes, for the small error case.

Section VII shows that an iterative technique of binary coding can be directly applied to non-binary codes. It also shows how an adapted Reed-Muller code can be profitably used in such a system.

Section VIII summarizes the results obtained in Sections II–VII and shows the advantages and shortcomings of many of these codes.

Section IX presents general conclusions which may be drawn from this paper.

II. SINGLE ERROR DETECTION CODES

Single error detection codes require message points separated in n dimensional space by a distance of two.

For the binary case, the only two possible types of errors are the change from a 1 to a 0 and from a 0 to a 1.

A simple technique that is used frequently for binary error detection codes is to encode all messages in such a manner that every message contains an even number of 1's. This is accomplished by adding a *parity check digit* to the information digits of a message; this digit is a 1 if an odd number of 1's exist in the information digits of a message and is a 0 if an even number of 1's exist in the information digits. At least two errors must occur before a message containing an even number of 1's can be converted into another message containing an even number of 1's, since the first error will always cause an odd number of 1's to appear. A message with an odd number of 1's is known to be incorrect.*

An analogous technique may be used for the unrestricted error case in non-binary codes. We can obtain a satisfactory code by adding a complementing digit to a series of information digits to form a message.

A *complementing* digit, base b, is defined as a digit which when added to some other digit will yield a multiple of b.

* Parity check digits may be selected to make the number of 1's in a message always odd, but the principle is the same; in this case, an error is recognized if a received message contains an even number of 1's.

For a single unrestricted error detection code, the complementing digit complements the sum of the information digits. A complementing digit is a check digit. In the binary case, it is a parity check digit.

As an example, consider a decimal code of this type. A message 823 would require a complementing digit 7, making the total message 8237 ($8 + 2 + 3 + 7 = 20$, a multiple of 10). An error in any one digit will mean that the sum of the message digits will not be a multiple of 10.

For the small error case, it is sufficient to make certain that the sum of all digits is even since any error of ± 1 would destroy this property. For the binary case, all errors are small since the only possible error on any digit is a change by ± 1; a simple parity check is adequate. For a non-binary code, it would be wasteful to add a digit just to make sure that the sum of all digits is even. In a decimal code for example, if the sum of the message digits is even, the values 0, 2, 4, 6, 8 for the check digit will satisfy a check, or if the sum of the message digits is odd, the values 1, 3, 5, 7, 9 will satisfy the check. More information could be sent if a choice among these values could be associated with information generated by the information source.

This introduces the concept of a mixed digit; i.e., a digit which conveys both check information and message information.

A *mixed digit* is defined as follows: a mixed digit x, base b, is composed of two components (y, z) where y represents an information component and z represents a check component. The number of information states of a mixed digit is β, with y taking the values $0, 1, \cdots, \beta - 1$; the number of check states of a mixed digit is α, the number base of z. In a message containing m check digits and h mixed digits, the number of check states for the *message* is $b^m \cdot \alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_h$, where α_i is the number of check states of the i'th mixed digit.

If mixed digits are used as part of a code, information must be available in at least two number bases; b, the number base of the channel, and β, the number base of the mixed digit. A situation where this arises naturally is in the case of the algebraic sign of a number; this is a digit of information, base 2, which may be associated with other digits of any base. Similarly, any identification which must be associated with numerical information can be conveniently coded in a number base different from the number base of the numerical information. Thus, a mixed digit can sometimes be used conveniently in an information transmission system without complicating the information source and receptor.

An error detection code for single small errors suggests the use of a mixed digit. In the decimal code for example, the quibinary[7] representa-

TABLE II — QUIBINARY CODE

Quinary Component	Binary Component	Decimal Digit
0	0	0
0	1	1
1	0	2
1	1	3
2	0	4
2	1	5
3	0	6
3	1	7
4	0	8
4	1	9

tion of the mixed digit might be used, letting the quinary component of the mixed digit convey information and the binary component a check. (Table II.)

The information source generates blocks of decimal digits followed by one quinary digit. The messages are then generated in the following way: record all decimal information digits as information message digits and take their sum; if the sum is even, the binary component, z, of the mixed digit is 0, otherwise it is 1. The quinary component, y, of the mixed digit is taken directly from the information source and combined with the calculated binary part by the rules of the quibinary code to form the mixed decimal digit. Thus, x, the value of the mixed digit, is given by the formula:

$$x = 2y + z. \tag{1}$$

For example, if the decimal digits of a message are 289 and the quinary digit of the message is 3, the mixed digit is 7, and the message is 2897. The sum of the decimal information digits is 19, which is odd, so that the binary component of the mixed digit is 1; this is combined with the quinary component, 3, by the rules of the quibinary code table, to form decimal digit 7. The requirement that the sum of all digits be even is satisfied by the binary component of the mixed digit, and the information associated with the mixed digit is contained in the quinary component.

This method is easily extensible to any other number base and is also extensible to the case of slightly larger but still restricted errors (such as ± 1 or ± 2), provided that the maximum single error is less than $(b - 1)/2$.

From the preceding example, it is apparent that mixed digits can be usefully employed in error detection codes. The use of mixed, check and information digits simplified the encoder and decoder. To differentiate among the classes of codes which will be described in this paper, the following terms will be used, in addition to those previously defined.

A *semi-systematic code* encoder produces messages containing only information, mixed and check digits. The information source generates information digits in base b for information digits, and in base β for mixed digits. (The example given above is a semi-systematic code.)

Of two coding schemes in the same channel base b, each working with messages of the same length, and each satisfying a given error detection or correction criterion, the more *efficient* scheme is defined as the one which produces the larger number of different possible messages.

III. SINGLE ERROR CORRECTION CODES, SMALL ERRORS (± 1)

The problems of error correction codes in nonbinary systems are extensive and must be treated in several distinct sections. The basic difference between the error correction problem in binary and non-binary codes is the fact that the sign of the error is important. In a binary code, if the message 11 is received and it is known that the second digit is incorrect, only one correction can be made, to 10. But in a decimal code with errors limited to ± 1, if the message 12 is received and it is known that the second digit is wrong, it can be changed to either 11 or 13.

Consider the following simple code for correcting single small errors. A decimal channel is used, and a message is composed of three information digits and one check digit. Let x_1 represent the check digit and x_2, x_3, x_4 the information digits. Here, x_1 is chosen to satisfy*

$$x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \bmod 10. \tag{2}$$

The encoder calculates x_1, and transmits the message $x_1x_2x_3x_4$. This is received as $x_1'x_2'x_3'x_4'$. The decoder then calculates c given by

$$c = (x_1' + 2x_2' + 3x_3' + 4x_4') \bmod 10. \tag{3}$$

If the assumption is made that at most a single small error exists, then this error can be corrected by using the following rules, which may be verified by inspection.

If $c = 0$, no correction is necessary;

$5 > c > 0$, decrease the cth digit by one;

---

* By definition $a = c \bmod b$ is equivalent to $a = c + nb$, where a, b, c and n are integers. The equality notation is used in preference to the congruence notation throughout this paper, since an addition performed without carry occurs naturally in many circuits; in terms of such a circuit, the mod b signifies only the base of the addition, and a true equality exists between the state of two circuits, with the same output even though one has been cycled more often.

$5 < c$, increase the $(10 - c)$th digit by one;

$c = 5$ implies a multiple error or a larger error.

Since the value of c is used for correcting a received message, it is called the corrector.* For the general case, a corrector is defined as follows.

In a message encoded to satisfy m separate checks, the result of calculating the checks for the received message at the decoder is an m digit word called the *corrector*. There are as many possible values of the corrector as there are check states of the message, although all of the values of the corrector need not correspond to a correctable error.

It is important that, for a given transmitted message, every different error will lead to a different value of the corrector; otherwise there will be no way of knowing which correction corresponds to a particular value of the corrector. The number of correctable errors may be far less than the number of possible values of the corrector, so that not all of these values may be useful for a code to correct a particular class of errors. However, the number of corrector states sets an upper limit to the number of possible corrections.

For many codes, it is convenient to associate a particular value of a corrector for the condition that a particular digit has been received too high by a single increment, for example, a 7 received as an 8.

The *characteristic* of a digit for a particular code is defined as the value of the corrector if that digit is incorrectly received, the error having increased the value of the digit by $+1$, and all other digits are correctly received. Obviously, this definition only applies to those codes having the property that the value of the corrector is independent of the value of the incorrect digit and of the other digits.

A *simple characteristic code* encoder produces messages in which each digit has a distinct characteristic as defined above.

The Hamming code is an example of a simple characteristic code as is the code previously described. In that example, the characteristic of x_i is i.

The advantage of a simple characteristic code for single small error correction is obvious: the association between the calculated checks and the correction to be performed is simple and does not depend on the values of the digits of the message.

The following example of a simple characteristic code will illustrate this principle more fully.

Consider a single small error correction code, working with a quinary

* The terms corrector and characteristic were first used in a more restricted sense in an article on binary coding by Golay.[8]

(base 5) channel. Each message will consist of ten information digits and two check digits.

Let x_1 and x_2 represent the check digits, and $x_3, x_4, \cdots, x_{12}$ represent the information digits.

The equations for calculating x_1 and x_2 are:

$$\begin{aligned} 1x_1 + 0x_2 + 0x_3 + 1x_4 + 1x_5 + 1x_6 + 1x_7 \\ + 2x_8 + 2x_9 + 2x_{10} + 2x_{11} + 2x_{12} = 0 \bmod 5, \end{aligned} \tag{4}$$

$$\begin{aligned} 0x_1 + 1x_2 + 2x_3 + 1x_4 + 2x_5 + 3x_6 + 4x_7 \\ + 0x_8 + 1x_9 + 2x_{10} + 3x_{11} + 4x_{12} = 0 \bmod 5. \end{aligned} \tag{5}$$

At the decoder, the corrector terms, c_1 and c_2, are calculated using x_i', the received value of x_i, in the following formulas:

$$\begin{aligned} 1x_1' + 0x_2' + 0x_3' + 1x_4' + 1x_5' + 1x_6' + 1x_7' \\ + 2x_8' + 2x_9' + 2x_{10}' + 2x_{11}' + 2x_{12}' = c_1 \bmod 5, \end{aligned} \tag{6}$$

$$\begin{aligned} 0x_1' + 1x_2' + 2x_3' + 1x_4' + 2x_5' + 3x_6' + 4x_7' \\ + 0x_8' + 1x_9' + 2x_{10}' + 3x_{11}' + 4x_{12}' = c_2 \bmod 5. \end{aligned} \tag{7}$$

The values of c_1c_2 corresponding to the condition that one and only one digit is too high by 1, $x_i' = x_i + 1$, can be read by reading the coefficients of the ith digit in the corrector formulas. This quantity is therefore the characteristic of the ith digit. If $x_i' = x_i - 1$, then the fives complements of these coefficients will be the value of the corrector. Table III lists the characteristics and characteristic complements associated with each digit.

TABLE III — CHARACTERISTICS AND CHARACTERISTIC COMPLEMENTS SYSTEMATIC QUINARY CODE

Digit	Characteristic	Complement of Characteristic
x_1	10	40
x_2	01	04
x_3	02	03
x_4	11	44
x_5	12	43
x_6	13	42
x_7	14	41
x_8	20	30
x_9	21	34
x_{10}	22	33
x_{11}	23	32
x_{12}	24	31

In this code all the possible values of c_1c_2 correspond to the characteristic of a digit or the complement of this characteristic, except 00 which corresponds to the correct message. (An inspection of equations (4) through (7) reveals that if $x_i' = x_i$ for all values of i, the values of c_1 and c_2 are 0). Thus, we can assign a unique correction to each value of c_1c_2.

The above techniques are extensible to other number bases and different length words provided b, the number base of the channel, is greater than 2. (The equivalent binary channel problem has been treated by Hamming.[1]) The following set of rules and conventions may be used for deriving a satisfactory set of characteristics for a simple characteristic systematic code used to correct single small errors for any length message, and any base, $b \geqq 3$. The rules must be followed, and the conventions (which represent one pair of conventions out of the set of pairs of conventions, which together with Rules 1 and 2 can be used for deriving a code of this class) if followed, will lead to a reasonably simple method for encoding and decoding messages.* Since the rules, not the conventions, limit the efficiency of the code, no set of conventions can be found which will lead to a more efficient code of this class.

Rule 1. For an n digit message (including check digits), m check digits are required and m must satisfy the following inequalities:

$$\text{if } b \text{ is odd,} \qquad \frac{b^m - 1}{2} \geqq n, \tag{8a}$$

$$\text{if } b \text{ is even,} \qquad \frac{b^m - 2^m}{2} \geqq n. \tag{8b}$$

Rule 2. No characteristic may be repeated; i.e., each digit must have a characteristic different from that associated with any other digit.

Convention 1. The various digits of a characteristic are arranged in a set order; i.e., C_{1i}, C_{2i}, $\cdots$, C_{mi}. The first digit which is neither zero, nor (in case b is even) $b/2$, must be less than $b/2$. There must be at least one such digit.

Convention 2. The characteristic of the jth check digit has a 1 in the jth position and 0's elsewhere.

Rule 1 is required since, for a code of this type, we must be prepared to correct any digit in one of two ways (± 1). This implies a minimum of $2n + 1$ values of the corrector, one for each possible correction, and one for the case of no corrections. This means that b^m, the number of possible

* The above distinction between rules and conventions will be observed throughout this paper.

values of the corrector, must be at least $2n + 1$, equation (8a). For even bases, we must reject all values of the corrector containing only the digits 0 and $b/2$ for representing error conditions for the following reasons: a positive error leads to a corrector that is the characteristic of the incorrectly received digit, and a negative error leads to the b-complement of such a characteristic. In order to have unique error correction, we must be able to distinguish between these two conditions. If a characteristic were to contain only the digits 0 and $b/2$, it would be equal to its own b-complement; such combinations of digits are therefore not useable as characteristics or characteristic complements.

Rule 2 is required to permit a unique identification of an incorrect digit in case of a single error.

Convention 1 allows us to distinguish between positive and negative errors. By observing this convention, a characteristic (corresponding to a positive error) can be distinguished from its complement (corresponding to a negative error) by inspecting the first digit of a corrector which is neither 0 nor $b/2$. A characteristic will have this digit less than $b/2$, a characteristic complement will have this digit greater than b/2. If the corrector is a characteristic, the correction is minus one; if it is a characteristic complement, it is plus one.

Once the characteristics have been chosen, the corresponding encoding procedure may be performed in the following manner: Let a_{ij} represent the jth digit of the characteristic of information digit x_i. Let z_j represent the check digit which has a characteristic containing a 1 in the jth position. If convention 2 has been observed, (9) can be used to calculate z_j:

$$\sum_{i=1}^{n-m} a_{ij}x_i = -z_j \bmod b. \tag{9}$$

An encoder calculates each z_j and inserts it into the message in those digit positions which have the characteristic of the jth check digit assigned to them.

In more general terms, we use implicit relations that are equivalent to the explicit equations given by (9). Letting x_i represent an information or a check digit, and letting C_{ij} represent the jth digit of the characteristic of the ith information or check digit, these formulas may be rewritten as

$$\sum_{i=1}^{n} C_{ij}x_i = 0 \bmod b. \tag{10}$$

At the receiver, the decoder calculates m different check sums. Let c_j

represent the check sum corresponding to the jth corrector term, and x_i' represent the received value of x_i : Then,

$$\sum_{i=1}^{n} C_{ij}x_i' = c_j \bmod b. \tag{11}$$

The difference between equations (10) and (11) is the result of any mutilations caused by the channel. If no error has occurred, all the c_j's are 0; if an error of ± 1 has occurred, the m c_j's will form the characteristic or the characteristic complement, respectively, of the incorrectly received digit.

One disadvantage of a systematic code is the discontinuity in the number of check states as a function of m, the number of check digits. For example, in decimal code one check digit is required for a message of up to four digits, and two check digits for up to forty-eight digits. Obviously, for a message of intermediate length, for example, twelve digits, many of the corrector states cannot be used for single error correction since they will not correspond to any single error. A more efficient code would be obtained if the check states were limited to a smaller number.

One method of reducing the number of check states is to perform the check in a different modulus than the modulus of the channel. In the single error detection code using a mixed digit, binary check information and quinary message information was conveyed by this digit. This code was more efficient than a systematic code because each message contained the minimum number of check states which is 2.

If a mixed digit, x, is composed of the two components (y, z) where y is the information state of the digit and z the check state, it is convenient to combine these two components to form x by means of the formula

$$x = \alpha y + z. \tag{12}$$

We calculate z by using a linear congruence equation modulo α.

The use of this formula permits a decoder to act on x', the received value of x, directly, without first resolving x' into y' and z', because (12) insures that $x' = y' \bmod \alpha$. This permits x' to be corrected directly and then resolved into its components.

As an example, consider a semi-systematic code for correcting a single small error in a decimal system, using a twelve digit message; ten of the digits are information digits and two are mixed digits, each conveying binary message information and quinary check information. (One of these binary digits might represent the sign of the number.)

With two quinary checks, twenty-five different check states are possible; for correcting single small errors in a twelve digit message, twenty-

TABLE IV — CHARACTERISTICS AND CHARACTERISTIC COMPLEMENTS, SEMI-SYSTEMATIC DECIMAL CODE

Digit	Characteristics	Characteristic Complements
x_1 (mixed digit)	1 0	4 0
x_2 (mixed digit	0 1	0 4
x_3	0 2	0 3
x_4	1 1	4 4
x_5	1 2	4 3
x_6	1 3	4 2
x_7	1 4	4 1
x_8	2 0	3 0
x_9	2 1	3 4
x_{10}	2 2	3 3
x_{11}	2 3	3 2
x_{12}	2 4	3 1

five corrector states are required, one for each of the two possible corrections (± 1) for each digit, and one for the case of a correctly received message. Characteristics may be chosen for the various digits in accordance with the rules and conventions outlined above in this case, since the check modulus is the same for both check digits. Consequently, it is no accident that these characteristics, shown in Table IV, are the same as those shown in Table III.

Let C_{i1} and C_{i2} represent the characteristic of the ith digit, and let y_1 and y_2 represent the two binary information digits. Then:

$$\sum_{i=3}^{12} C_{i1}x_i = -z_1 \bmod 5, \qquad x_1 = z_1 + 5y, \tag{13}$$

$$\sum_{i=3}^{12} C_{i2}x_i = -z_2 \bmod 5, \qquad x_2 = z_2 + 5y_2 . \tag{14}$$

Because $x_1 = z_1 \bmod 5$ and $x_2 = z_2 \bmod 5$, these relations can be rewritten implicity to resemble equation (10):

$$\sum_{i=1}^{12} C_{i1}x_i = 0 \bmod 5, \tag{15}$$

$$\sum_{i=1}^{12} C_{i2}x_i = 0 \bmod 5. \tag{16}$$

At the decoder, the corrector c_1c_2 is calculated by:

$$\sum_{i=1}^{12} C_{i1}x_i' = c_1 \bmod 5, \tag{17}$$

$$\sum_{i=1}^{12} C_{i2}x_i' = c_2 \bmod 5. \tag{18}$$

If the corrector is 00, the message has been correctly received; otherwise, the corrector is either the characteristic or characteristic complement of the incorrect digit, from which plus one or minus one respectively must be subtracted as a correction.

Consider the general case. Let $x_1, x_2, \cdots, x_k$ represent the k information digits; $y_1, y_2, \cdots, y_m$ represent the information state of the m mixed digits, and $z_1, z_2, \cdots, z_m$ represent the check state of the m mixed digits. In addition, let $\alpha_1, \alpha_2, \cdots, \alpha_m$ represent the number base of $z_1, z_2, \cdots, z_m$ respectively; $\beta_1, \beta_2, \cdots, \beta_m$ represent the number of possible states of $y_1, y_2, \cdots, y_m$ respectively, and $x_{k+1}, x_{k+2}, \cdots, x_{k+m}$ represent the values of the mixed digits after the message has been encoded. (Note that for simplicity, a check digit is considered as a special case of a mixed digit; its information state is permanently 0.) The following encoding procedure may be used in which $x_1, x_2, \cdots, x_k$ are used directly as part of the transmitted message. This is a semi-systematic code, which means that information digits are not changed in coding. To derive the mixed digits, the following formulas are used:

$$a_{11}x_1 + \cdots + a_{1k}x_k = -z_1 \bmod \alpha_1 \tag{19-1}$$

$$x_{(k+1)} = y_1\alpha_1 + z_1 \tag{20-1}$$

$$a_{12}x_1 + \cdots + a_{2k}x_k + a_{2(k+1)}x_{(k+1)} = -z_2 \bmod \alpha_2 \tag{19-2}$$

$$x_{(k+2)} = y_2\alpha_2 + z_2 \tag{20-2}$$

$$\vdots \qquad\qquad \vdots$$

$$a_{j1}x_1 + \cdots + a_{jk}x_k + \cdots + a_{j(k+j-1)}x_{(k+j-1)}$$

$$= -z_j \bmod \alpha_j \tag{19-j}$$

$$z_{(k+j)} = y_j\alpha_j + z_j \tag{20-j}$$

$$\vdots \qquad\qquad \vdots$$

$$a_{m1}x_1 + \cdots + a_{mk}x_k + \cdots + a_{m(k+m-1)}x_{(k+m-1)} = -z_m \bmod \alpha_m \tag{19-m}$$

$$x_{(k+m)} = y_m\alpha_m + z_m . \tag{20-m}$$

In each case, the value of the check component z_j, of a mixed digit $x_{(k+j)}$ is determined by a formula involving the information digits and previously calculated mixed digits. Immediately after z_j has been determined, $x_{(k+j)}$ is calculated for possible use in calculating $z_{(j+1)}$. After the message has been completely encoded the following equations, analogous to (10), will be satisfied.

Let C_{ij} represent a_{ji} in equation (19-j). Then,

$$\sum_{i=1}^{k+m} C_{ij}x_i = 0 \bmod \alpha_j . \tag{21}$$

(Since $x_{(k+j)} = z_j \bmod \alpha_j$, substitution of $x_{(k+j)}$ for z_j in equation (19-j) will continue to satisfy the equation.)

At the decoder, equation (21) is changed to

$$\sum_{i=1}^{k+m} C_{ij}x_i' = c_j \bmod \alpha_j \tag{22}$$

In (22), x_i' represents the received value of x_i, and c_j represents the jth digit of the corrector. If all the digits have been correctly received, i.e., $x_i' = x_i$ for all values of i, then $c_1 = c_2 = \cdots = c_m = 0$; [see equation (21)]. If x_h had been received incorrectly so that $x_h' = x_h + 1$, but all other digits had been correctly received, then the value of c_j (the jth digit of the corrector) would be calculated in the following manner:

$$c_j \bmod \alpha_j = \sum_{i=1}^{k+m} C_{ij}x_i'$$

$$c_j \bmod \alpha_j = \sum_{i=1}^{k+m} C_{ij}x_i + C_{hj} = C_{hj} \tag{23}$$

Equation (23) proves that C_{hj} is actually the jth digit of the characteristic of x_h, because by definition, the characteristic of x_h is the value of the corrector when $x_h' = x_h + 1$, and all other digits have been correctly received. This means that the general term, C_{ij} of (21), is actually the jth digit of the characteristic of the ith digit and that this is a simple characteristic code.

For the case that $x_h' = x_h - 1$, the value of the corrector is such that if it were incremented, digit by digit, by the characteristic of x_h, the corrector would be composed only of zeros. Incrementing the corrector by the characteristic of x_h is equivalent to recalculating the corrector with x_h' increased by one, which in this case would amount to calculating the corrector for the case of a correctly received message. The latter is composed of all zeros [see (21)]. Thus, for the case of a single error of -1, the corrector is the characteristic complement of the digit which is incorrectly received. For a semi-systematic or systematic code, the characteristic complement is an m digit word whose jth digit is the complement modulo a_j of the jth digit of the characteristic.

Equation (20-j) shows that generally $\alpha_j\beta_j$ cannot exceed b. (An exception is given below.) The maximum value of y_j is $\beta_j - 1$ since y is a digit in the number base β_j. The maximum value of z_j is usually $\alpha_j - 1$, since z_j is a digit in the number base α_j. Thus,

$$x_{k+j} = y_j\alpha_j + z_j \leqq b - 1, \tag{24}$$

$$(\beta_j - 1)\alpha_j + \alpha_j - 1 \leqq b - 1, \tag{25}$$

$$\alpha_j\beta_j \leqq b. \tag{26}$$

Equation (24) restates (19-j), and also states that the maximum value of any digit x, is $b - 1$, where b is the number base of the channel. In (25), the maximum values of y_j and z_j are substituted to yield the result shown in (26).

It was stated above that the maximum value of z_j is usually $\alpha_j - 1$. An exception occurs only in case z_j checks only itself and other mixed digits, the latter being restricted to fewer than $b - 1$ states. Under such circumstances, the value of z is sometimes restricted, so that even though z is calculated to satisfy a check, modulo α_j [see equation (19-j)], it cannot assume $\alpha_j - 1$ values. For example, a code for transmitting a single digit message over a decimal channel and permitting the correction of small errors, might use as the set of transmitted messages the digits, 0, 3, 6, 9. In this case, $\alpha = 3$ (any correct message satisfies the check $x = 0 \bmod 3$) and $\beta = 4$ since four different messages may be transmitted. In this case, z is restricted to the value 0 because the mixed digit checks only itself.

In order to correct single errors of ± 1, using a simple characteristic code, it is necessary and sufficient that every characteristic be different from every other characteristic, and that it also be different from the complement of every other characteristic.

The following rules and conventions may be used to derive a set of characteristics which meet the requirements for a simple characteristic semi-systematic or systematic code for correcting small errors for any base $b \geqq 3$ and an arbitrary length message. No set of conventions can be found which will lead to a more efficient code of this class, since the rules, not the conventions limit the efficiency of the code.

Rule 1. For an n digit message, including mixed digits, containing m mixed or check digits of which m_1 are associated with an even modulus, α, the inequality

$$(\alpha_1 \cdot \alpha_2 \cdot \ \ldots \ \cdot \alpha_m - 2^{m_1})/2 \geqq n \tag{27}$$

must be satisfied.

Rule 2. No characteristic may be repeated, i.e., each digit must have a characteristic different from that associated with any other digit.

Rule 3. Since the mth check is the last one to be calculated, and the characteristic of the mth mixed digit must therefore contain only a single digit which is not 0, α_m must be greater than 2.

Convention 1. The various digits of a characteristic are arranged in a set order, i.e., $C_{i1}, C_{i2}, \cdots, C_{im}$. The first digit which is neither 0 nor $\alpha_j/2$ must be less than $\alpha_j/2$. There must be at least one such digit.

Convention 2. The characteristic of the jth mixed digit has a 1 in the jth position and 0's elsewhere, provided that $\alpha_j \neq 2$. If $\alpha_j = 2$, the characteristic of this mixed digit has a 1 in the jth and mth positions, and 0's elsewhere.

Rule 1 is required because the number of possible corrector states is $\alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_m$, of which only those containing at least one digit which is neither 0 nor $\alpha/2$ can be associated with the $2n$ possible errors. The same reasons used for Rule 1 for the systematic code case are equally applicable here; a characteristic containing only the digits 0 or $\alpha_j/2$ in the jth position is not distinguishable from its complement.

Rule 2 is required to permit a unique identification of an incorrect digit.

Rule 3 is necessary to derive the sign of an error on the mth mixed digit.

The reasons for using Conventions 1 and 2 in the case of the systematic code are equally applicable in this case. For the case $\alpha = 2$, however, a special convention must be used to avoid a conflict with Convention 1.

The procedure for converting a set of characteristics into an error correcting code system is the same for a semi-systematic code as for a systematic code except that the following additional functions must be performed: the encoder must combine check states with information states to derive mixed digits, and the decoder must resolve mixed digits into information and check digits *after* it has performed its corrections.

By using these rules and conventions, the most efficient simple characteristic code can be determined. For messages of length n (including mixed or check digits), the following relations must be satisfied:

Let

$$P = \alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_m ,$$

$$Q = \beta_1 \cdot \beta_2 \cdot \ldots \cdot \beta_m ,$$

$$m_1 = \text{number of even } \alpha\text{'s}.$$

Then:

$$(P - 2^{m_1})/2 \geqq n, \tag{28}$$

$$\alpha_i \beta_i \leqq b. \tag{29}^*$$

* For exceptions, see above.

TABLE V — DECIMAL ERROR CORRECTION CODES

n	P	$\frac{10^m}{Q}$	$\alpha_1, \alpha_2, \ldots$	$\beta_1, \beta_2, \ldots$	$2n + 1$
1	3	2.5	3	4	3*
2	5	5	5	2	5
3	10	10	10	1	7
4	10	10	10	1	9
5	15	16.7	5, 3	2, 3	11
6	15	16.7	5, 3	2, 3	13
7	15	16.7	5, 3	2, 3	15
8	20	20	10, 2	1, 5	17
9	25	25	5, 5	2, 2	19
10	25	25	5, 5	2, 2	21
11	25	25	5, 5	2, 2	23
12	25	25	5, 5	2, 2	25
13	30	33.3	10, 3	1, 3	27
14	30	33.3	10, 3	1, 3	29
15	40	40	10, 2, 2	1, 5, 5	31
16	40	40	10, 2, 2	1, 5, 5	33
17	50	50	10, 5	1, 2	35
18	50	50	10, 5	1, 2	37
19	50	50	10, 5	1, 2	39
20	50	50	10, 5	1, 2	41

* The single digit message containing the points 0, 3, 6, 9 is an exception to the inequality $\alpha\beta \leqq b$, because the mixed digit checks only itself.

For the most efficient code b^m/Q should be minimized. This term represents the ratio of the number of possible messages for an n digit message with and without error correction. This is normally at least as great as $2n + 1$, the number of possible corrections on such a message.

Table V shows the most efficient decimal codes of this type for an n digit message, for values of n from 1 to 20. Where two or more different codes are equally efficient, the code with the fewest mixed digits is shown. It is easy to convert from a code using two mixed digits with $\alpha_1 = 5$, $\alpha_2 = 2$, to one using a check digit with $\alpha = 10$, or to make the inverse conversion, and to show that both codes are equally efficient.

IV. SINGLE ERROR CORRECTION CODES, UNRESTRICTED ERROR

The problem of correcting an unrestricted error on one digit of a message must be divided into two categories, depending on whether b is a prime number or a composite number. As will be seen, the error correction problem for prime bases is considerably simpler than that for composite bases. The method for correcting errors in prime number systems was discovered by Golay,[6] although this did not come to the author's attention until after he had worked out the same method. The adaptation to non-prime channel bases is believed to be novel. Since the adaptation makes use of the code for prime bases, both will be described.

4.1 *Prime Number Base, Single Unrestricted Error Correction Code*

This code depends upon a fundamental property of prime numbers, well known in number theory.[9] Let p represent a prime number and d, c, and w represent non-negative integers less than p, related by the expression:

$$dw = c \bmod p. \tag{30}$$

If $d \neq 0$, then d and c uniquely determine w.

In order to have a simple characteristic systematic code for correcting unrestricted errors, it is necessary and sufficient that the set of characteristics shall have the property that all multiples of all characteristics are distinct. Equation (30) implies a unique correspondence between multiples of a characteristic and the characteristic itself, if we consider c to be the multiple, d the multiplying factor and w a digit of the characteristic. An error, d, is simply identifiable if a known digit of a characteristic is always 1. If each characteristic is distinct from every other and if a sufficient number of check digits are available, a simple characteristic code can be obtained. In the following set of rules and conventions which may be used for deriving a set of characteristics for a simple characteristic systematic code for correcting single unrestricted errors, p represents the prime number base of the channel. The number base of the channel must be prime, and the length of the message is arbitrary. Since the rules and not the conventions limit the efficiency of the code, no other set of conventions may be found which will lead to a more efficient code of this class.

Rule 1. For an n digit message, m check digits are required and m must satisfy the inequality

$$n \leqq \frac{p^m - 1}{p - 1}. \tag{31}$$

Rule 2. Each digit must have a different characteristic.

Convention 1. The digits of a characteristic are arranged in a set order, i.e., $C_{i1}C_{i2} \cdots C_{im}$. The first digit which is not 0 must be 1.

Convention 2. The characteristic of the jth check digit has a 1 in the jth position and 0's elsewhere.

Rule 1 is required for a code for correcting single unrestricted errors since any digit must be correctable in one of $p - 1$ ways. This implies a minimum of $n(p - 1) + 1$ states for the corrector, one for each correction and one for the correct message. When m check digits are used, p^m corrector states are obtained.

Rule 2 and Convention 2 are the same for the single small error correction systematic codes. The same reasons apply for both cases.

Convention 1 is changed from the equivalent convention for the small error correction code, because the magnitude of the error, not only its sign, must be derivable for a code for correcting single unrestricted errors.

An encoder first encodes the message according to (32), where C_{ij} represents the jth digit of the characteristic of x_i,

$$\sum C_{ij}x_i = 0 \bmod b. \tag{32}$$

The decoder calculates the corrector using the following formula where x_i' represents the received value of x_1;

$$\sum C_{ij}x_i' = c_j \bmod b. \tag{33}$$

The decoder then examines the digits of the corrector in order. The first digit which is not 0 shows the magnitude, d, of the error. All digits are then divided by d (provided $d \neq 0$). (That division is unique, as shown by (30).) The result of this division is the characteristic of the incorrect digit, which is then corrected by subtracting d.

Consider a code for correcting a single unrestricted error in a six digit message for a base 5 channel:

$$6 \leqq \frac{5^m - 1}{4}. \tag{34}$$

A value of 2 for m will satisfy equation (34). The characteristics are 14, 13, 12, 11, 10 and 01, the last two being check digit characteristics, for x_1, x_2, x_3, x_4, x_5, and x_6 respectively. Here, x_1, x_2, x_3, and x_4 are information digits. The encoding formulas are:

$$x_1 + x_2 + x_3 + x_4 = -x_5 \bmod 5, \tag{35}$$

$$4x_1 + 3x_2 + 2x_3 + x_4 = -x_6 \bmod 5. \tag{36}$$

The decoding and correcting formulas are: (x_i' is the received value of x_i)

$$x_1' + x_2' + x_3' + x_4' + x_5' = c_1 \bmod 5, \tag{37}$$

$$4x_1' + 3x_2' + 2x_3' + x_4' + x_6' = c_2 \bmod 5. \tag{38}$$

The corrector is c_1c_2.

Suppose that a message 221321 is received as 224321. Then:

$$c_1 = 13 = 3 \bmod 5, \tag{39}$$

$$c_2 = 26 = 1 \bmod 5. \tag{40}$$

To find the characteristic of the digit, x_h, that was incorrectly received from the value of the corrector, (41) and (42) must be solved:

$$d\, C_{h1} = c_1 = 3 \bmod 5, \tag{41}$$

$$d\, C_{h2} = c_2 = 1 \bmod 5. \tag{42}$$

Because the first non-zero digit of any characteristic is 1, (41) can be solved for d since $C_{h1} = 1$. This yields the result, $d = 3$. Using this result, (42) is solved for C_{h2}; by inspection, $C_{h2} = 2$, since $3 \cdot 2 = 6 = 1 \bmod 5$. Thus the characteristic of the incorrect digit, $C_{h1}\, C_{h2}$, is 12, and the error d, is 3; x_3' must therefore be reduced by 3 to get the correct value. Since the message was received with x_3' too high by an amount 3, this result confirms our expected correction.

Any correction that is applied must be applied on a modulo b basis. For example, if a correction of -2 is indicated on a digit whose received value is 1, $1 - 2 = 4 \bmod 5$, which means that the digit is corrected to 4.

Codes of this type are restricted in their construction. No mixed digits may be used, and the number base must be prime. For the case of $n = [(p^g - 1)/(p - 1)] + 1$, $g + 1$ check digits are required [see (31)]. This means that the number of information digits for a message of this length is the same as for a message one digit shorter, which requires only g check digits. A comparable binary case is the Hamming Code example of an eight binary digit message (four information digits) compared with a seven digit message (also four information digits). In the binary case, the extra digit is useful for double error detection, but unfortunately, this is not the case for non-binary codes.

4.2 *Composite Number Base, Single Unrestricted Error Correcting Code*

The problem of correcting an unrestricted error on a single digit, working with a number base b, that is not a prime is much more difficult. Many relatively inefficient techniques exist. For example, characteristics containing only binary numbers (0 and 1) might be used; (this would amount to using the Hamming Code directly). This is obviously inefficient since the corrector associated with any single digit error of amount

d, would contain only the digits 0 and d, thus wasting most of the possible corrector values.*

It is possible to encode and decode using the prime factors of the number base, performing separate and independent corrections on each factor. This is also inefficient, since for many cases, information as to which digit is in error is found independently in two or more ways, while for certain values of the error, it can be found in only one way. Working with mixed digits and check bases, α lower than b, is not satisfactory since certain values of the error (α in particular) will never show up in a particular check. The technique used for primes will not work since multiples of two different characteristics may be identical; for example, base 10, characteristics 11 and 13, error 5, will both yield correctors of 55.

Another technique that is relatively efficient is, however, available. It involves performing all check, encoding and decoding operations in a number base p, where p is some prime number (usually, the lowest) that is equal to or greater than b. (In case b is a prime, we use the procedure outlined above, which is a special case of the procedure to be described below.)

The obvious difficulty in such a procedure is that while the information channel can only handle b levels, the check digits may assume p levels, corresponding to the required p check states. This dilemma can be resolved by adding an *adjustment digit*. The object of this digit is to permit check information to be transmitted in a base greater than b, the channel base. The idea of an adjustment digit can best be illustrated by an example. Suppose for a decimal channel, checks are performed in a unodecimal (base 11) code. Let γ represent the value corresponding to ten. (The consecutive integers in a unodecimal system are then 0, 1, 2, 3, $\cdots$, 9, γ, 10, 11, $\cdots$, 19, 1γ, 20, etc.) Suppose in a particular message, four check digits, z_1 , z_2 , z_3 , z_4 , calculated modulo 11 from decimal information digits are used, whose values are 1, 0, γ, 8. A fifth digit, z_0 is added such that the sums modulo 11 of $z_1 + z_0$, $z_2 + z_0$, $z_3 + z_0$, $z_4 + z_0$ are kept constant at 1, 0, γ, 8 respectively. There are eleven different words satisfying the condition: $[1, 0, \gamma, 8] = [(z_1 + z_0), (z_2 + z_0), (z_3 + z_0), (z_4 + z_0)]$. These are shown in Table VI. Of these words, six do not contain the digit γ, and so may be transmitted over a decimal channel. Thus, an adjustment digit permits check digits which are calculated in a number system of a higher base than b, to be transmitted over a base b channel. When an adjustment digit is used in base p for adjusting m digits so that transmission over a channel in base b is possible, a mini-

* A waste of corrector values is equivalent to an excessive number of check states for a message, which in turn implies an excessive number of check digits.

mum of $b - m(p - b)$ states are allowed for the adjustment digit. (For certain values of the check digits, more states could be allowed, but a code for utilizing these extra states becomes unwieldy.) For the case $b = 10$, $p = 11$, this turns out to be $10 - m$. At least one state must be available for each adjustment digit, to have a workable code.

The characteristic of an adjustment digit is determined in the following way: if an adjustment digit adjusts the jth check digit, then the jth digit of the characteristic of the adjustment digit is 1; otherwise, it is 0. The characteristic of all other digits may be derived using the rules described above for the prime number base channel, except that p, the prime number base of the code must be used instead of b, the number

TABLE VI — ILLUSTRATION OF ADJUSTMENT DIGIT

z_0	z_1	z_2	z_3	z_4
0	1	0	γ	8
1	0	γ	9	7
2	γ	9	8	6
3	9	8	7	5
4	8	7	6	4
5	7	6	5	3
6	6	5	4	2
7	5	4	3	1
8	4	3	2	0
9	3	2	1	γ
γ	2	1	0	9

base of the channel, for generating characteristics. A message is initially encoded using a value of 0 for an adjustment digit. Subsequently, if the adjustment digit always has at least q allowable states, it may be used to transmit one additional information digit, base q, of information. If the value of this information digit is y, the $(y + 1)$st lowest possible value of the adjustment digit (making the lowest value equivalent to $y = 0$) meeting the requirement that all adjusted check digits are no greater than $b - 1$ is transmitted. The adjustment digit in conjunction with its associated check digits conveys a digit, base q, of information.

In the example given above, $q = 6$ and if y is 4, the fifth lowest value of z_0, 7, is transmitted. The lowest value must be associated with $y = 0$. The values of $z_0z_1z_2z_3z_4$ that are sent over the decimal channel are 75431.

An example of such a code is one using a decimal channel working in a unodecimal base for the purposes of encoding and error correction. The word length, n, is twelve, nine decimal information digits, one octal (base

8) information digit associated with the adjustment digit, and two check digits. The characteristics are the following:

x_1 1γ	x_5 16	x_9 12
x_2 19	x_6 15	x_{10} 11 (adjustment digit)
x_3 18	x_7 14	x_{11} 10 (check digit)
x_4 17	x_8 13	x_{12} 01 (check digit)

Let z_{11} and z_{12} represent the values of the check digits x_{11} and x_{12}, originally derived from $x_1, x_2, \cdots, x_8, x_9$:

$$x_1 + x_2 + x_3 + \cdots + x_9 = -z_{11} \bmod 11, \tag{43}$$

$$\gamma x_1 + 9x_2 + 8x_3 + \cdots + 2x_9 = -z_{12} \bmod 11. \tag{44}$$

From z_{11} and z_{12}, the ten different words $(0, z_{11}, z_{12})$, $(1, z_{11} - 1, z_{12} - 1)$, $(2, z_{11} - 2, z_{12} - 2)$, $\cdots$, $(9, z_{11} - 9, z_{12} - 9)$ are formed. If y is the value of the octal information digit, the $(y + 1)$st such word, that does not contain the digit γ, is selected and transmitted as the last three digits of the message. For example, if $z_{11} = 2$, $z_{12} = 1$ and $y = 6$, the ten words are (0, 2, 1), (1, 1, 0), (2, 0, γ), (3, γ, 9), (4, 9, 8), (5, 8, 7), (6, 7, 6), (7, 6, 5), (8, 5, 4), (9, 4, 3); the word (8, 5, 4) is selected since it is the seventh in the sequence that does not contain any γ's. Table VII shows the choice of the three last digits as a function of y, given $z_{11} = 2$, $z_{12} = 1$.

Formula (45) is used for calculating the corrector. Let C_{ij} represent the jth digit of the characteristic of x_i, c_j the jth digit of the corrector, and x_i' the received value of x_i. Then,

$$c_j = \sum_{i=1}^{12} C_{ij} x_i' \bmod 11. \tag{45}$$

The translation from corrector to correction is the same as if the original

TABLE VII — RELATION BETWEEN ADJUSTED DIGIT AND ASSOCIATED INFORMATION

y	x_{10}	x_{11}	x_{12}
0	0	2	1
1	1	1	0
2	4	9	8
3	5	8	7
4	6	7	6
5	7	6	5
6	8	5	4
7	9	4	3

message had been in a unodecimal code. (This has been illustrated in Section 4.1.)

The first step of the encoding procedure is to calculate the unadjusted check digits. Next, the adjusted check digits and adjustment digit are selected according to the value of y, the information digit associated with the adjustment. The message is then ready for transmission.

At the decoder, the message is first corrected as if it had been received as a unodecimal message. The information digits are then in their corrected states. Next, the adjustment digit and the check digits are examined and the inverse of the encoding process used to select a particular set of check and adjustment digits is used to reconstruct the value of y which originally controlled the selection. In the example given above, the values of x_{10}, x_{11}, x_{12} are 8, 5, 4 respectively; the decoder recognizes that this is the seventh lowest value of x_{10}, which means that the value of y, used in selecting x_{10} and the adjusted values of x_{11} and x_{12}, was 6.

The code described above is fairly efficient; about 90 per cent of the corrector values can be associated with corrections; the product of the information states and the check states is about 97 per cent of the total number of states of a twelve decimal digit word. Each of the above factors reduces the efficiency of the code below a possibly unattainable maximum. It will be noted, however, that this reduction is relatively small in both cases, and is very much lower than would be the case for any of the rejected schemes. The scheme is not difficult to instrument; relatively little additional equipment is required in addition to the basic equipment for instrumenting a simple prime number base channel, unrestricted single error correcting code system.

The method of adjustment digits is general and can be used for deriving a single error correction code for correcting unrestricted errors for any channel base. Any convenient prime check base, p, at least as great as b may be used, although the lowest will generally be the most efficient. The only requirements which must be fulfilled are that the number of states of the adjustment digit must be at least 1, and that at least two check digits must be associated with each adjustment digit. An adjustment digit associated with m check digits, working with a channel base b and a check base p, may have $b - m(p - b)$ different states.

V. SINGLE ERROR CORRECTION, DOUBLE ERROR DETECTION CODES FOR CORRECTING SMALL ERRORS

Single error correction, double error detection codes are very useful in situations where a message may occasionally be repeated. In order for a correction code to be reasonably useful in a system with random noise or errors, the errors must be relatively infrequent, which makes double errors still more infrequent. If means are available for an occasional but very infrequent repetition of a message, a single error correction, double error detection code will increase the reliability of a digital system, since a message may be repeated if a double error is recognized.

This section will show how the ideas of the single error correction, double error detection Hamming Code may be combined with the ideas of semi-systematic single small error correction codes (described in Section III) to derive simple and efficient codes for correcting single small errors and detecting double small errors.

In order to derive a simple characteristic code for correcting single small errors, and detecting double small errors, a set of characteristics must be found having the property that the sum or difference of two characteristics or their complements or double the value of one characteristic or its complement be distinguishable from the value of any single characteristic or its complement. The sum of two characteristics represents the value of the corrector for a message with two errors of $+1$, $+1$, the difference represents two errors of $+1$, -1, the sum of their complements represents two errors of -1, -1; double a characteristic represents an error of $+2$, and double a complement represents an error of -2. To have a true single error correction, double error detection code for small errors, all these cases must be distinguished from the case of a single error or no error by making certain that the value of the corrector for any of these cases is different than the value of the corrector corresponding to any single error and no error.

Table VIII gives the characteristics used in the single error correction Hamming Code and the single error correction, double error detection Hamming Code for conveying four digits of information in a message containing seven or eight binary digits respectively.

An inspection ot Table VIII shows that the sum (performed without carries from column to column) of any two characteristics in the right part of the table is distinguished by having at least one 1 in the first three places and a 0 in the last place. This distinguishes it from any single characteristic since all characteristics have a 1 in their last place.

Some difficulties arise in trying to adopt such a scheme directly in a non-binary system. For the code to be efficient, an over-all check would have to be performed using a mixed digit; only two check states are required for an over-all parity check, and if $b > 3$, (b representing the number base of the channel) at least two information states are possible. But the over-all check digit, which performs a binary check, is not checked by any other digit. This means that although errors might be detected in an over-all check digit, difficulties would be encountered in determining the direction of the correction, so that the information conveyed by the mixed digit could be used. Actually, means are available, for accomplishing an adaptation of binary techniques. These methods are described in Section VII but they are less straightforward than the ones described below.

For channels with base b, greater than 3, at least one check may be made using a check base, α_m, that is 4 or greater. If characteristics are used whose last digit (the digit associated with the α_m check) is always 1, and whose only other limitation is that each characteristic is different from every other characteristic, a satisfactory code is obtained. Single errors are corrected in the normal way. If the last digit of the corrector is 1 or $\alpha_m - 1$, the error is ± 1 respectively on the digit whose characteristic or whose characteristic complement is indicated by the corrector. If the last digit of the corrector is 2 or $\alpha_m - 2$, or the last digit is 0 and other digits are not all 0, a double error is indicated. If the entire corrector is made up of 0's, the message is correct as received.

TABLE VIII — CHARACTERISTICS FOR HAMMING CODES

Single Error Correction			Single Error Correction Double Error Detection
001	Check Digit	x_1	0011
010	Check Digit	x_2	0101
011	Information Digit	x_3	0111
100	Check Digit	x_4	1001
101	Information Digit	x_5	1011
110	Information Digit	x_6	1101
111	Information Digit	x_7	1111
	Over-all Check Digit	x_8	0001

An example is a code for a ten digit message, decimal base channel; eight decimal information digits, one mixed digit conveying binary message information (such as the sign of the decimal number) and quaternary (base 4) check information, and one check digit are transmitted in each message. Let x_1 and x_2 represent the mixed and check digit respectively, x_3 through x_{10} the information digits, y_1 the binary information conveyed by x_1, and z_1 the quaternary check information conveyed by x_1. The encoding formulas are:

$$2x_3 + 3x_4 + 4x_5 + 5x_6 + 6x_7 + 7x_8 + 8x_9 + 9x_{10} = -x_2 \bmod 10, \quad (46)$$

$$x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} = -z_1 \bmod 4, \quad (47)$$

$$x_1 = z_1 + 4y_1 . \quad (48)$$

Note that (46), (47) and (48) must be applied consecutively, in that order, since (47) cannot be applied without knowing x_2 obtained from (46), and (48) requires z_1, obtained from (47).

The characteristics are 01, 11, 21, 31, 41, 51, 61, 71, 81, 91 respectively; the complements of the characteristics are 03, 93, 83, 73, 63, 53, 43, 33, 23, 13 respectively. The corrector, c_1c_2, is calculated at the decoder by the following formulas (x_i' is the received value of x_i):

$$c_1 = \sum_{i=2}^{10} (x_i')(i - 1) \bmod 10 \tag{49}$$

$$c_2 = \sum_{i=1}^{10} x_i' \bmod 4 \tag{50}$$

Consider the example of a message with decimal information digits 3 7 5 2 0 6 5 2 and binary information digit 1. Then $x_2 = 3$, $z_1 = 3$, and $y_1 = 1$, yielding a value of 7 for x_1. The message is sent as 7 3 3 7 5 2 0 6 5 2. Suppose that the sixth digit is changed to 1 in transmission. Then the corrector has a value 53; this is the complement of the characteristic of the sixth digit and indicates that the sixth digit should be incremented by 1 according to the rules previously stated. If the sixth digit had been received as 1 and the seventh digit also received as 1 (an error of +1), then the corrector value would be 10, indicating a double error (see rules stated above).

If a multiple of 4 is used as α_m, the last digit of a characteristic may assume all odd values below $\alpha_m/2$. The rule then is that an even value of the last digit of the corrector, or a 0 for the last digit and other digits of the corrector not all 0, indicates a double error.

The following set of rules and conventions may be used with any base $b \geqq 4$, and any length of message, for deriving a set of characteristics for a semi-systematic code for correcting single small errors and detecting double small errors. Since the conventions restrict the efficiency of the code, it is conceivable that a different set of conventions will yield a more efficient code in some cases; (51) may be modified through the use of an alternate set of conventions.

Rule 1. No two digits may have identical characteristics.

Convention 1. Choose for α_m a multiple of 4. Let $\alpha_m/4 = g$.

Convention 2. The characteristic of the mixed digit associated with α_m contains a single 1 in the last position; the rest of its digits are 0.

Convention 3. The characteristics of the jth mixed or check digit contains a 1 in the last position, a 1 in the jth position and 0's elsewhere.

Convention 4. The characteristic of an information digit has an odd number less than $\alpha_m/2$ in its last position. The rest of its digits are arbitrary.

Convention 5. The above conventions restrict the choice of characteristics. In order to have n distinct characteristics, m mixed or check digits, using check bases α_1, α_2, $\cdots$, α_m, are required, and inequality (51) must be satisfied:

$$n \leqq \alpha_1\alpha_2 \cdot \ldots \cdot \alpha_{m-1} \cdot g. \tag{51}$$

Codes may be derived using the above conventions only if $b \geqq 4$. For the ternary case, a relatively efficient code may be obtained by using one ternary digit as an over-all parity check digit. The rest of the message is in a single small error correction code, derived using the rules and conventions of Section III. Any single small error will lead to a failure of the parity check, and a double small error will lead to a failure of other checks but not the parity check.

No general solution has been found for deriving an efficient single error correction double error detection code for the unrestricted error case. Also, no general solution has been found for deriving an efficient multiple error correction code for the unrestricted error case. A reasonably efficient method has been found for correcting multiple errors in the more important small error case; this is discussed in Section 6.2.

VI. THE USE OF BINARY ERROR CORRECTION TECHNIQUES IN NON-BINARY SYSTEMS

In this section, methods for using binary codes for the correction of errors in a non-binary system are described. Although the single small error correction codes obtained in this manner are generally less flexible than the codes obtained in Section III, the class of multiple error correction codes described in Section 6.2 is the only reasonably satisfactory class of such codes that has been found. The codes described in this section are semi-systematic but are not simple characteristic codes.

6.1 *Single Small Error Correction Codes*

Binary codes are most conveniently used for correcting small errors (± 1). Suppose any digit, base b, has an associated pair of binary digits, arranged in such a way that a change of ± 1 in the base b digit will change only one of the two binary digits. For $b = 10$, an association such as the one shown in Table IX might be used. For example, if a 6 is received as a 7, the associated binary message would indicate that the second of the binary digits is incorrect; a 7 can be corrected

TABLE IX — ASSOCIATED BINARY DIGITS FOR CORRECTION OF SMALL ERRORS

Decimal Digit	Associated Binary Digits
0	00
1	01
2	11
3	10
4	00
5	01
6	11
7	10
8	00
9	01

TABLE X — REFLECTED QUIBINARY CODE

Decimal Digit	Quinary Component	Binary Component	Associated Binary Digits
0	0	0	00
1	0	1	01
2	1	1	11
3	1	0	10
4	2	0	00
5	2	1	01
6	3	1	11
7	3	0	10
8	4	0	00
9	4	1	01

to an 8 or a 6, but only the correction to 6 would correspond to a change in the second binary digit of the associated binary message.

If the first of the associated binary digits is the odd or even indication of a quinary component of a decimal digit, a decimal digit can convey ten states rather than the four states of the associated binary digits. The combination of binary and quinary digits shown in Table X may be called a reflected quibinary code because of its analogy with the reflected binary code.*

If a method were available for transmitting without error (e.g., by using an error correcting code) a message composed of the associated binary digits in a base b code, small errors could be corrected in the base b digits.

An examination of Table X for resolving a decimal digit into binary and quinary components, reveals that a change of ± 1 on any decimal

* The reflected binary code has the property that each increment changes only one binary digit; for example, the eight successive words of a three binary digit reflected binary code are 000, 001, 011, 010, 110, 111, 101, 100.

digit will change only one of these two components. Further, an error corresponding to a change in the quinary component can be uniquely corrected if the error in the decimal digit is assumed to be ± 1. For example, if a received 6 is discovered to have an incorrect quinary component, only a decrease in the quinary component making the decimal digit 5 is a possible correction, since an increase in the quinary component would correspond to the decimal digit 9, a change of more than ± 1 from 6.

A system is shown in Fig. 2 for taking advantage of these properties.

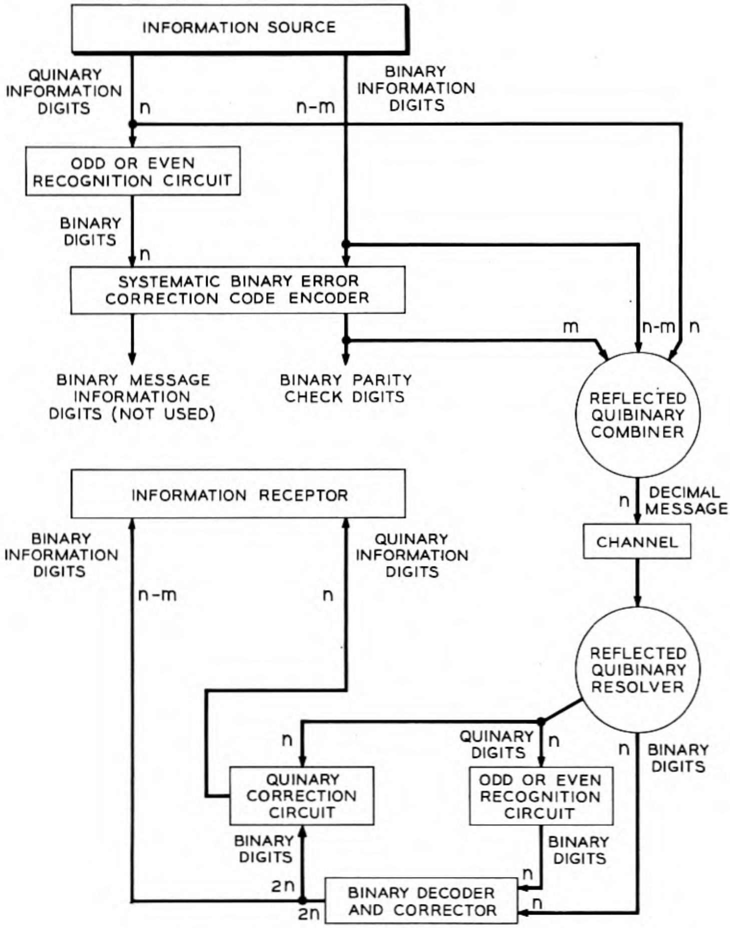


Fig. 2 — Use of binary codes with a decimal channel.

In this example, an information source generates n quinary and $n - m$ binary information digits for each message. All quinary digits go through an odd or even recognition circuit to be converted into binary digits for the purpose of generating a binary error correction code message. These binary digits and the binary digits generated by the information source are fed into a systematic binary error correction code encoder whose output is a binary message containing $2n$ digits, of which m are parity check digits. This output is divided into two parts, $2n - m$ original inputs to the encoder unchanged by the encoding process (this is a systematic encoder which does not change information digits in encoding), and m parity check digits.

The m parity check digits are then combined with m of the quinary information digits through the use of the reflected quibinary combiner to form m of the decimal digits of the decimal message that is transmitted; the other decimal digits are formed by combining the $n - m$ binary information digits with the rest of the quinary information digits.

The decimal message is transmitted over the noisy channel and arrives with one or more (a number limited by the choice of the binary code) errors of ± 1 on decimal digits. It is fed into a reflected quibinary resolver which resolves decimal digits into binary and quinary components in accordance with the reflected quibinary code (Table X). The quinary digits are then fed into an odd or even recognition circuit to form binary digits; these and the binary outputs of the resolver are fed into a binary decoder and corrector, working with the same code as the binary encoder. The output of this corrector should correspond to the output of the original binary encoder.

In the decoder, the binary digits are corrected. When the binary digit derived from a quinary digit is corrected, however, the quinary digit is not yet correct. The correction of the quinary digit is performed by examining both the corrected binary digit derived from the quinary digit and the corrected binary digit which was derived from the same decimal digit as the quinary digit in question. The rules for correcting the quinary digit are given in Table XI.

As an example, consider the application of a Hamming Code for transmitting ten binary digits in a fourteen binary digit message.

Using a code of this type, single errors of ± 1 may be corrected in a seven digit decimal message, transmitting seven quinary digits of information and three binary digits of information. The characteristics required for a fourteen binary digit Hamming Code message are shown in the first column of Table XII.

TABLE XI — CORRECTING QUINARY DIGITS

Q	B_1	B_2	Correction of Quinary Digit
Even	0	0	None
Even	0	1	None
Even	1	0	−1
Even	1	1	+1
Odd	0	0	+1
Odd	0	1	−1
Odd	1	0	None
Odd	1	1	None

TABLE XII — BINARY CODE USED FOR CORRECTING DECIMAL MESSAGE

Binary Characteristics		a	b	Position in Decimal Message
0 0 0 1	Parity Check Digt	(0)	(0)	Binary comp. of 1st digit
0 0 1 0	Parity Check Digit	(0)	(0)	Binary comp. of 2nd digit
0 0 1 1		1	1	Binary comp. of 3rd digit
0 1 0 0	Parity Check Digit	(1)	(1)	Binary comp. of 4th digit
0 1 0 1		0	0	Binary comp. of 5th digit
0 1 1 0		0	0	Binary comp. of 6th digit
0 1 1 1		1	3	Quinary comp. of 7th digit
1 0 0 0	Parity Check Digit	(1)	(1)	Binary comp. of 7th digit
1 0 0 1		1	3	Quinary comp. of 6th digit
1 0 1 0		0	2	Quinary comp. of 5th digit
1 0 1 1		0	4	Quinary comp. of 4th digit
1 1 0 0		1	1	Quinary comp. of 3rd digit
1 1 0 1		1	3	Quinary comp. of 2nd digit
1 1 1 0		0	0	Quinary comp. of 1st digit

To illustrate the method completely, a strictly binary example will first be illustrated, then a related decimal example. In column a of Table XII, the digits of a binary message are indicated and in column b, the binary and quinary information digits. The values of the parity check digits, which are shown in parentheses, are calculated by the usual formula. Let C_{ij} represent the jth digit of the characteristic of the ith digit (including parity check digits):

$$\sum_{i=1}^{14} x_i C_{ij} = 0 \bmod 2. \tag{52}$$

This formula applies for all values of j and in this case will yield four implicit equations each with one unknown term, the value of the parity check digit. Using the given values of the binary information digits, the values of the parity check digits are calculated. These are shown in parentheses in Table XII.

The binary message is

$$0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 0.$$

For this example, the quinary components (quinary information digits) of decimal digits are chosen odd if the corresponding digit of the binary example is 1, even if that digit is 0. The binary and quinary components are then combined by the rules of the reflected quibinary code to form the decimal digits 0 7 2 9 4 7 6. For example, the quinary and binary components of the fifth digit are 2 and 0, respectively; the decimal digit which has these components is 4, the fifth decimal digit of the message.

Consider the binary case. Suppose that the message is mutilated in transmission so that the tenth digit is received incorrectly. The message is mutilated from

$$0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 0$$

to

$$0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 0.$$

The decoder and corrector calculates the corrector by

$$c_j = \sum_{i=1}^{14} x_i' C_{ij} \bmod 2. \tag{53}$$

In this formula, c_j is the jth digit of the corrector and x_i' the received value of x_i. In this example the corrector is 1 0 1 0, which means that the tenth digit, which has this characteristic, is wrong and should be changed to 0.

The corresponding error in the decimal example is a change in the fifth digit from 4 to 3. If the message 0 7 2 9 3 7 6 is received, the resolver and quinary to binary converter delivers the message

$$0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 0$$

to the decoder instead of

$$0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 0$$

corresponding to the correct message. The corrected binary message is produced at the output of the decoder and corrector. When the quinary and binary components of the fifth digit are examined by the quinary correction circuit, the following inputs exist:

Received quinary digit	1 (Odd) (quinary component of received decimal 3)
Corrected binary digit derived from quinary	0 (B_1)
Corrected binary digit from same decimal number	0 (B_2).

Table XI shows that the quinary digit must be increased by 1 to 2, which combined with the binary 0 conveyed by the same decimal digit yields a decimal value of 4, the original transmitted value.

The best semi-systematic simple characteristic code for correcting single small errors in a seven digit message allows 6×10^5 possible messages in a seven digit message (see Table V), whereas this code allows 6.25×10^5. This code is therefore slightly more efficient. In addition, this code has the special advantage that any error of ± 2 on one digit is recognizable since the corrector will have a value of 1111 for the associated binary message. (An inspection of the choice of characteristics and assignment of characteristics to the two components of any decimal digit will confirm this.)

This general technique can be applied to any base b channel, provided

TABLE XIII — COMPONENTS OF QUINARY DIGITS

Mixed Digit			Information Digit		
Quinary Digit	Info. Comp.	Check Comp.	Quinary Digit	Binary Comp.	Ternary Comp.
0	0	0	0	0	0
1	0	1	1	1	0
2	1	1	2	1	1
3	1	0	3	0	1
4	not used	not used	4	0	2*

* If quinary information is initially generated, the combination (1, 2) will not occur.

that b is greater than 3. For odd bases, the digits which convey a parity check component and an information component cannot be utilized efficiently since one state of the base b digit is not available. For example, using a base 5, (see Table XIII), only two information and two parity check states may be conveyed by one digit, since the use of a third information state would require at least six states for the mixed digit. In the case of information digits, however, all states can be used. In the quinary example, the resolution of a digit into two components and the subsequent recombination is subject to the restraint that one of the combinations (1, 2) will not occur, which can be assured if the information source generates quinary digits.

For the case of high redundancy codes having the property that the associated binary code contains more than 50 per cent parity check digits (corresponding to a negative value of $n - m$ in Fig. 2), at least some of the base b digits must convey two or more parity check digits.

This can be easily accomplished: a decimal digit can convey three

TABLE XIV — DECIMAL DIGIT CONVEYING THREE BINARY DIGITS

Decimal Digit	Binary Components
0	0 0 0
1	0 0 1
2	0 1 1
3	0 1 0
4	1 1 0
5	1 1 1
6	1 0 1
7	1 0 0
8	not used
9	not used

parity check digits if a simple reflected binary code correspondence between binary and decimal digits is maintained as shown in Table XIV.

An extension of this idea is the encoding of the original information (i.e., the information that is shown coming out of the information source in Fig. 2) in some error detection or correction code. For example, the decimal to reflected quibinary code resolver will cause both components to be incorrect if an error of ± 2 in a decimal digit occurs. In this case, the system shown in Fig. 2 will automatically make a correction on the decimal digit of either $+2$ or -2 depending upon the value of the received decimal digit, and provided a double error correction binary code is used. Such a correction will be incorrect about half the time. If the received binary digit is compared to the corrected binary digit and the received quinary digit is compared to its corrected odd or even digit, an error of ± 2 can be detected without changing the code. If one extra binary check digit, treated as an information digit by the encoder and decoder, is transmitted in the message, this binary digit can convey the information necessary for determining the sign for a correction of ± 2, provided that only one such correction is required for any one message. A rule for determining the value of this digit is:

$$\begin{aligned} B_c &= 0 \qquad \text{if} \qquad \sum_{i=1}^{n} q_i = (0 \text{ or } 1) \bmod 4, \\ B_c &= 1 \qquad \text{if} \qquad \sum_{i=1}^{n} q_i = (2 \text{ or } 3) \bmod 4, \end{aligned} \tag{54}$$

where q_i represents the ith quinary information digit, and B_c represents the special check digit. If the received message contains one error of ± 2 on a digit, two possible corrections may be made on the quinary component of this digit; ± 1. Obviously, only one of these corrections will satisfy the equation for determining B_c since the two possible corrected values of q are two units apart.

Note that the associated binary codes for performing such a correction must have the property that two binary digits may be corrected since an error of ± 2 corresponds to incorrect values for two associated binary digits. If the noise is such that errors of ± 2 are not very unlikely, it may be desirable to place the binary and the quinary components of any one decimal digit in a different binary error correction code word so as to make the errors independent. In a seven decimal digit message, as an example, the quinary components of the first four decimal digits can be used to generate parity check digits which are conveyed by the binary components of the last three decimal digits. The binary component of the fourth decimal digit (this might be B_c) and the quinary components of the last three decimal digits generate parity check digits conveyed by the binary components of the first three decimal digits. Two separate binary error correction code messages are then conveyed by a single seven digit decimal code message. Each message is in a four information digit, three parity check digit Hamming Code. Through the use of this code, one error in the binary component of any decimal digit, and one error in the quinary component of any decimal digit may be corrected.

In certain cases, the quinary digits themselves might be encoded in an error correction code for single unrestricted errors before the binary process is carried out. This is helpful chiefly for occasional large errors, leading to initial miscorrections.

The variations based upon the principles described, which can be applied to any channel, provided $b \geqq 4$, including the pyramiding of one code scheme upon another, are almost endless. Generally, the last encoding and first decoding step should be able to correct many more errors than the first encoding step. For example, if quinary components are encoded in single unrestricted error correction quinary code, the binary code should probably be a triple or quadruple error correction code; otherwise a correction may not correspond to the most probable error condition, and the correction scheme loses its effectiveness.

These techniques cannot be conveniently applied to the ternary channel, since a ternary digit cannot be resolved into two components efficiently.

6.2 *Multiple Small Error Correction Codes*

One limitation of the above techniques is the requirement for a systematic binary code; i.e., a code in which some of the binary information digits are transmitted directly, and others are determined by parity checks on information and previously calculated check digits. These

TABLE XV — REED-MULLER CODES — 256 DIGIT MESSAGE

Number of Digits of Information per Message	Number of Errors Correctable per Message
255	0
247	1
219	3
163	7
93	15
37	31
9	63
1	127

systematic codes are conveniently applicable only to the correction of single errors and a few special cases of multiple errors.

The Reed-Muller[10] codes are not systematic codes, ("systematic" being used in the narrow sense indicated above, not in the sense of Hamming[11]), but offer the advantage that multiple error correction is relatively straightforward. For this reason, it is desirable to find some way of adapting the binary Reed-Muller codes for correcting a number of small errors in non-binary codes.

To explain the nature of the Reed-Muller codes completely is beyond the scope of this paper; a list of their important features is sufficient. This is:

1. The length of a message is 2^k binary digits for the simpler versions of the code.

2. If C_c^d represents the number of combinations of d items taken c at a time, and $C_c^d = d!/[c!(d - c)!]$, then $2^k - \sum_{i=0}^{m} C_{k-i}^{k}$ information digits may be transmitted correctly in a message containing 2^k digits, if no more than $2^m - 1$ errors occur in the messages; 2^m errors are detected but they are not always correctable. The Reed-Muller codes for correcting a large number of errors will frequently correct more than $2^m - 1$ errors, and will always correct $2^m - 1$ or fewer errors.

These values are given for a 256 digit message in Table XV.

3. Each digit of the transmitted message is a parity check of a group of digits from the information source; the message cannot be broken down into information digits and check digits.

4. The decoding is accomplished by a number of majority decisions among different groups of message digits.

A technique will be described for using a Reed-Muller code efficiently to correct a number of small (± 1) errors for any code base b that is a multiple of 2, and also, at a small sacrifice of efficiency, a number of larger errors.

A theorem, stating that any code which is generated by a set of parity checks will contain the same set of allowable messages as some systematic code, was proved by Hamming.[11] In particular, such a theorem indicates that a Reed-Muller code will contain the same set of allowable messages as some systematic code. This was also proved by Slepian,[12] who has given a simple method of deriving a systematic code generating the same set of messages as a Reed-Muller code. For convenience, such a code will be called an SERM code (Systematic Equivalent Reed-Muller code).

A Reed-Muller decoder serves to derive the information digits from a message in Reed-Muller code which may have been mutilated by noise. If a Reed-Muller decoder is followed by a Reed-Muller encoder, the combination serves as a noise eliminator (provided the noise is within the correction bounds of the code), since the output of the encoder is the noiseless Reed-Muller code message that is equivalent to the noisy message that entered the decoder. This property is useful since it means that any message, drawn from the set of Reed-Muller code messages, which has not been mutilated outside the bounds set up by a particular Reed-Muller code, will be restored to its original form, by a Reed-Muller decoder followed by a Reed-Muller encoder. Since an SERM code will produce only messages included in the set of messages of the corresponding Reed-Muller code, the SERM code can be used in conjunction with a Reed-Muller decoder and encoder to permit transmission over a noisy channel in a systematic code.

The two systems shown in Fig. 3 are therefore equivalent in their error correction properties. In both cases, messages from the set of Reed-Muller code messages are sent, and since the same decoder is used initially, both systems will correct errors in the received message in the same manner. The Reed-Muller encoder in the second system is required because a Reed-Muller decoder does not correct a message but derives information digits from the received message directly. The derived information digits, however, necessarily correspond to some corrected form of the received message and, in effect, the decoder performs the same correction as it would perform by deriving the corrected form of the message first.

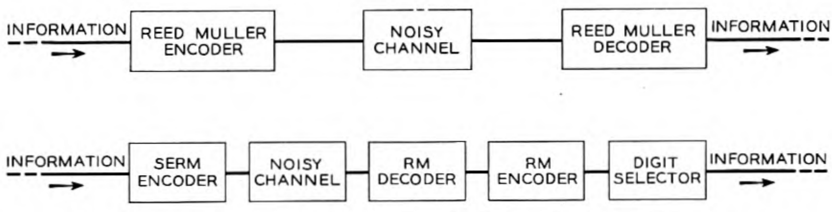


Fig. 3 — Equivalent systems using SERM and Reed-Muller codes.

TABLE XVI — MULTIPLE SMALL ERROR CORRECTION CODE USING SERM CODES WITH DECIMAL CHANNEL

Message Length	Information Digits (Equivalent Decimal Digits)	Check Digits (Equivalent Decimal Digits)	No. of Small Errors Correctable per mes.
128	127.7	.3	0
128	125.3	2.7	1
128	116.9	11.1	3
128	100.0	28.0	7

This means that a Reed-Muller code can be adapted to the system shown in Fig. 2. The Systematic Binary Error Correction Code Encoder is simply an SERM encoder; this is permissible since the SERM codes are systematic. The Binary Decoder and Corrector is simply a Reed-Muller decoder followed by a Reed-Muller encoder. Everything else remains unchanged.

This scheme offers flexibility for the correction of large numbers of small errors. Proper initial error correction encoding of the original information digits will permit correction of a small number of large errors.

Table XVI shows some typical cases of the correction of many small errors in a decimal message as a function of the number of information and check digits in a message of constant length. For convenience, everything is shown in equivalent decimal digits, even though in the actual code, binary and quinary information digits are used. Only the first few entries are considered, since the message composed exclusively of the digits 0, 3, 6, 9 in which any number of small errors in a decimal channel may be corrected (this code is described by the first entry of Table V) is more efficient than the codes corresponding to subsequent entries on Table XVI. This code, which is very easy to instrument, will transmit the equivalent of 77 decimal digits in a 128 decimal digit message.

One problem not efficiently solved by these techniques is the multiple-error correction ternary channel problem. A technique which can be used is a code identical to the regular binary Reed-Muller Code, except that all equations will be modulo 3 instead of modulo 2. In decoding, this will sometimes require subtraction instead of addition; in modulo 2 equations there is no difference between these operations, but in modulo 3 equations, the two operations are distinct. The same procedure can be used for correcting multiple unrestricted errors in any base.

VII. ITERATIVE CODES

All the codes described above have one disadvantage; occasional excessive noise will yield a non-correctable message. In order to approach error free transmission, some iterative coding procedure may be used. This problem has been solved by Elias.[13] His methods are directly applicable to non-binary codes, since nothing restricts the digits to binary values.

In order to minimize the complexity of an iterative coding procedure, systematic codes are desirable. The advantages of the Reed-Muller code are significant however, especially for the case of a relatively noisy channel. A sound procedure for a binary channel would therefore be to use SERM codes, (see Fig. 3); such codes are more efficient than iterated Hamming Codes in a relatively noisy channel.

VIII. SUMMARY AND ANALYSIS

Many codes have been presented in this paper, all constructed by some combination of procedures involving linear congruence or modulo equations.

In most cases, more efficient codes exist. Exhaustive procedures exist for deriving maximum efficiency codes, although the codes derived in this manner usually require an extensive codebook, both at the encoder and at the decoder. Even for simple single error correction binary codes, the most efficient code is not always a systematic code. For example, the best systematic single error correction binary code working with an eight digit message has only 16 different allowable messages; it is known[14] that a non-systematic code with at least 19 allowable messages exists.

In the case of non-binary codes, the situation is somewhat worse. Very few of the codes given in this paper take advantage of the fact that, for most situations, a digit that is incorrectly received as 0 or $b - 1$ is usually corrected only in one direction and no need exists to specify whether the correction is ± 1. Most of the codes are arranged so that any received digit may be corrected either positively or negatively. No codes have been found which take full advantage of such a property, other than codebook codes, except for isolated instances of short message codes having symmetrical properties. For example, the single digit, single small error correction decimal code having 0, 3, 6, 9 as the allowable messages takes full advantage of this property, and is, at the same time, a true semi-systematic code.

It is extremely difficult to find the ultimate limits of efficiency of codebook codes. The exhaustive procedures are totally impractical except for very short messages. If an analysis is restricted to codes which do not take advantage of the property that certain values of digits may be corrected in only one direction, and it is assumed that each possible message is mutilated to the same number of incorrect messages, one limit to the efficiency of codes may be found. This limit can be derived from the fact that an error correction code decoder and correction circuit must be able to convert any message which contains errors within the bounds of the correction performed by the code, into the value of the message as originally transmitted, or must be able to derive the original information which was fed into the encoder. Thus, if each message may be mutilated in w ways, and still be corrected, then at least w messages must be associated with each allowed message. This is indicated diagrammatically in Fig. 4. The messages produced by the encoder are shown at the left; each one fans out to $w - 1$ mutilated messages plus the original message. The decoder converts any of these w messages into the original message.

The value of w can be determined by taking all possible combinations of errors that can be corrected by a coding system. For example, for a code system which can correct up to $(d - 1)/2$ small errors in different digits in an n digit message, w is given by

$$w = \sum_{i=0}^{(d-1)/2} C_i^n 2^i, \tag{55}$$

where d is the minimum distance between messages, and

$$C_i^n = \frac{n!}{(n - i)!i!}.$$

This equation merely signifies that w is the sum of all combinations of positive and negative (accounting for the 2 term) errors in up to $(d - 1)/2$ different digits out of n digits. For single errors, $w = 2n + 1$.

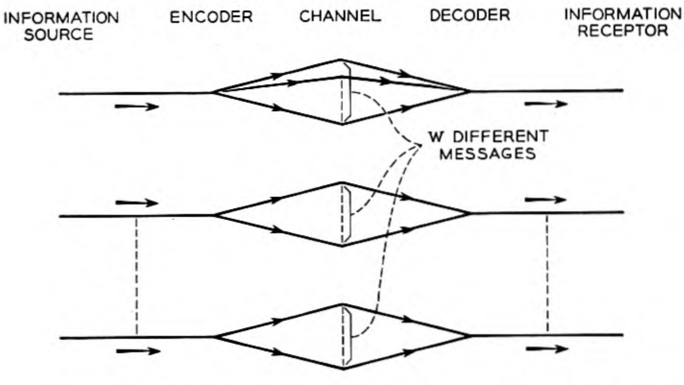


Fig. 4 — Graphical representation of an error correction code.

The number of different messages that can be produced by the encoder must be no greater than b^n/w, subject to the above restriction, b^n representing the maximum number of messages that the decoder may receive as an input. If only systematic and semi-systematic codes are considered, the number of messages is limited to multiples of powers of b and of the information component base β of mixed digits. The number of check states must be at least as large as w, so that w different correctors may be calculated and associated with w different corrections.

Subject to the above restrictions, the following statements may be made.

1. The systematic single small error correction codes derived using the rules of Section III are the most efficient systematic single small error correction codes possible. For those codes in which the two sides of inequality (8a) are equal, no code, not even a non-systematic code, is more efficient.

2. The systematic single unrestricted error correction codes derived using the rules of Section 4.1 are the most efficient systematic single unrestricted error correction codes. For those codes in which the two sides of inequality (31) are equal, no code is more efficient.

3. No codes are more efficient than those semi-systematic codes, derived using the rules of Section III, for which the two sides of inequalities (28) and (29) are equal and $m_1 = 0$. It is difficult to make more general statements about semi-systematic codes, because special techniques (such as those of Section VI), not all of which are known, may be used with these codes.

For multiple error correction codes, other techniques are both simpler and more efficient than the straight systematic and semi-systematic techniques described in Sections III, IV and V. One such scheme has been described in detail in Section VI. No codes have been found which approach the limit set by w, but the codes described in Section 6.2 are moderately efficient.

Throughout this paper, all techniques which involve vast complications at the expense of slight additional efficiency have been avoided. Codebook methods are always possible. If a technique is almost as complicated as a codebook technique with only slightly greater efficiency than a simple technique, the simple technique would always be used in practice, and the codebook satisfies the mathematical and theoretical requirements. In a sense, a really complicated technique is only useful for deriving a better lower limit for the maximum efficiency of a codebook code. In the non-binary case, however, a codebook system is considerably more efficient than any code system which does not take advantage of the fact that all transmitted messages are not mutilatable to an equal number of correctable received messages.

From the point of view of deriving lower limits to the maximum efficiency of a codebook technique, such a consideration is vital. Except for a few relatively trivial cases, no codes have been found which take significant advantage of the above consideration, for deriving such a limit.*

IX. CONCLUSION

In this paper, techniques have been presented for deriving error correction codes for non-binary systems. None of the methods presented are overly complicated, nor do they require excessive storage capacity for either the encoding or decoding and correction system.

The codes are sufficiently simple so that their use with a non-binary storage system may be considered, and the development of such a storage system should not be stopped because a system without flaws or not subject to noise cannot be realized.

An important disadvantage of using error correction codes with such a system is the time requirement. Correction usually requires a significant amount of time. This is probably one reason why the Hamming Code is not used more extensively. The more advanced and complicated codes, such as the Reed-Muller Codes, suffer particularly from the amount of time required for a correction. The codes described in this paper are therefore probably best suited to medium or low speed storages, which are not read too frequently.

A study of this type may be of some interest to those who have been considering the use of multi-state devices for building switching systems and computers, since this paper represents a study of a typical problem. Certain lessons may be derived from this study:

1. Restriction to a single number base for all operations is a severe handicap. The more advanced codes presented in this paper, require extensive use of different number base operations. The ability, inside the computer, to change number bases for different operations, may well be useful.

2. Different problems are best solved using different number bases. For example, the use of an even number base is desirable for multiple small error correction codes, while the use of a prime number base is desirable for correcting single large errors. It is the author's opinion that

* Note that this restriction has less significance in the case of binary codes. In a symmetrical channel with only two available signals, each value of a digit may be changed in as many ways, namely, one, as every other.

number bases which are the product of several small factors are best. Suggested values are six, ten and twelve. Number bases with two different prime factors, may offer an advantage, since they permit simple translation and change of number base among at least three different numbers.

In the comparison between binary and non-binary error correction codes, the following observations may be made:

1. Keeping the amount of information per message fixed, a binary single error correction code is less efficient than a non-binary single small error correction code, provided b, the channel base, is greater than three, but is more efficient than a non-binary single unrestricted error correction code.

2. Non-binary codes are slightly more complicated to implement than binary codes; this applies to multiple error correction codes as well as to single error correction codes. The amount of added complication is in no case really extensive.

It was initially hoped that this study might also produce some additional binary error correction techniques. One such technique was discovered: the use of a systematic equivalent Reed-Muller code to approach error free coding (see Section VII).

Finally, the author wishes to express the hope that further work on non-binary systems will be encouraged by this study.

ACKNOWLEDGEMENTS

This work was performed under the part-time Graduate Study Plan of Bell Telephone Laboratories at the Columbia University School of Engineering under the guidance of Prof. L. A. Zadeh. The author wishes to acknowledge the help of Prof. Zadeh, both in the selection of a dissertation topic and in the subsequent guidance of the study. In addition, a number of helpful discussions with C. Y. Lee and A. C. Rose helped to guide the research into a study of the most significant problems in the field.

REFERENCES

1. R. W. Hamming, Error Detecting and Error Correcting Codes, B.S.T.J., **29,** p. 147–160, April, 1950.
2. Of Current Interest, Elec. Engg., p. 871, Sept., 1956.
3. C. Y. Lee and W. H. Chen, Several-Valued Combinational Switching Circuits, Trans. A.I.E.E., **75,** Part I, p. 278–283, July, 1956.
4. D. Slepian, A Class of Binary Signaling Alphabets, B.S.T.J., **35,** p. 203–234, Jan., 1956.

5. Hamming, *op. cit.*, Section 1.
6. M. J. E. Golay, Notes on Digital Coding, Proc. I.R.E., **37,** p. 657, June, 1949.
7. W. Keister, A. E. Ritchie, and S. H. Washburn, *The Design of Switching Circuits*, D. Van Nostrand Co., Inc., New York, 1951, p. 316.
8. M. J. E. Golay, Binary Coding, 1954 Symposium on Information Theory, Trans. I.R.E., **PGIT-4,** p. 23–28, Sept., 1954.
9. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, Oxford, 1954, p. 51, Theorem 57.
10. I. S. Reed, A Class of Multiple-Error-Correcting Codes and the Decoding Scheme, 1954 Symposium on Information Theory, Trans. I.R.E., **PGIT-4,** p. 38–49, Sept., 1954.
11. Hamming, *op. cit.*, Section 7.
12. D. Slepian, A Note on Two Binary Signaling Alphabets, Trans. I.R.E., **IT-2,** p. 84–86, June, 1956.
13. P. Elias, Error Free Coding, 1954 Symposium on Information Theory, Trans. I.R.E., **PGIT-4,** p. 29–38, Sept., 1954.
14. V. I. Siforov, On Noise Stability of a System with Error-Correcting Codes, Trans. I.R.E., **IT-2,** p. 109–115, Dec., 1956. See Table II, Column 8.

Shortest Connection Networks And Some Generalizations

By R. C. PRIM

(Manuscript received May 8, 1957)

The basic problem considered is that of interconnecting a given set of terminals with a shortest possible network of direct links. Simple and practical procedures are given for solving this problem both graphically and computationally. It develops that these procedures also provide solutions for a much broader class of problems, containing other examples of practical interest.

I. INTRODUCTION

A problem of inherent interest in the planning of large-scale communication, distribution and transportation networks also arises in connection with the current rate structure for Bell System leased-line services. It is the following:

Basic Problem — *Given a set of (point) terminals, connect them by a network of direct terminal-to-terminal links having the smallest possible total length (sum of the link lengths).* (A *set* of terminals is "connected," of course, if and only if there is an unbroken chain of links between every two terminals in the set.) An example of such a *Shortest Connection Network* is shown in Fig. 1. The prescribed terminal set here consists of Washington and the forty-eight state capitals. The distances on the particular map used are accepted as true.

Two simple construction principles will be established below which provide simple, straight-forward and flexible procedures for solving the basic problem. Among the several alternative algorithms whose validity follows from the basic construction principles, one is particularly well adapted for automatic computation. The nature of the construction principles and of the demonstration of their validity leads quite naturally to the consideration, and solution, of a broad class of minimization problems comprising a non-trivial abstraction and generalization of the basic problem. This extended class of problems contains examples of practical

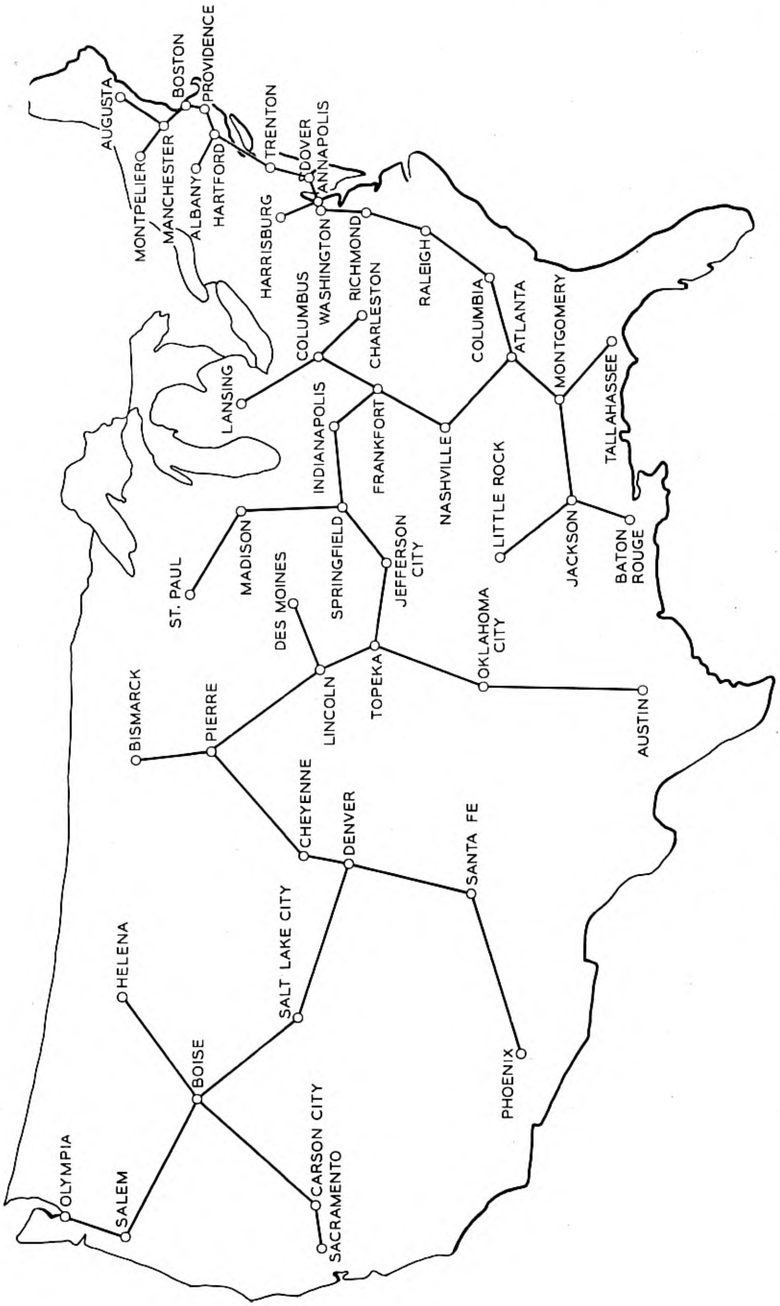


Fig. 1 — Example of a shortest connection network.

interest in quite different contexts from those in which the basic problem had its genesis.

II. CONSTRUCTION PRINCIPLES FOR SHORTEST CONNECTION NETWORKS

In order to state the rules for solution of the basic problem concisely, it is necessary to introduce a few, almost self-explanatory, terms. An *isolated terminal* is a terminal to which, at a given stage of the construction, no connections have yet been made. (In Fig. 2, terminals 2, 4, and 9 are the only isolated ones.) A *fragment* is a terminal subset connected by direct links, between members of the subset. (In Fig. 2, 8-3, 1-6-7-5, 5-6-7, and 1-6 are some of the fragments; 2-4, 4-8-3, 1-5-7, and 1-7 are

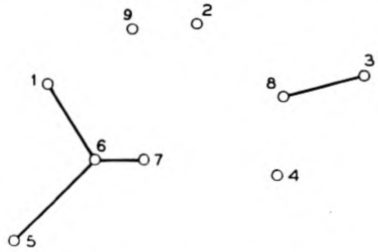


Fig. 2 — Partial connection network.

not fragments.) The *distance of a terminal from a fragment* of which it is not an element is the minimum of its distances from the individual terminals comprising the fragment. An *isolated fragment* is a fragment to which, at a given stage of the construction, no external connections have been made. (In Fig. 2, 8-3 and 1-6-7-5 are the only isolated fragments.) A *nearest neighbor of a terminal* is a terminal whose distance from the specified terminal is at least as small as that of any other. A *nearest neighbor of a fragment*, analogously, is a terminal whose distance from the specified fragment is at least as small as that of any other.

The two fundamental construction principles (P1 and P2) for shortest connection networks can now be stated as follows:

Principle 1 — *Any isolated terminal can be connected to a nearest neighbor.*

Principle 2 — *Any isolated fragment can be connected to a nearest neighbor by a shortest available link.*

For example, the next steps in the incomplete construction of Fig. 2 could be any one of the following:

(1) add link 9-2 (P1 applied to Term. 9)

(2) add link 2-9 (P1 applied to Term. 2)
(3) add link 4-8 (P1 applied to Term. 4)
(4) add link 8-4 (P2 applied to frag. 3-8)
(5) add link 1-9 (P2 applied to frag. 1-6-7-5).

One possible sequence for completing this construction is: 4-8 (P1), 8-2 (P2), 9-2 (P1), and 1-9 (P2). Another is: 1-9 (P2), 9-2 (P2), 2-8 (P2), and 8-4 (P2).

As a second example, the construction of the network of Fig. 1 could have proceeded as follows: Olympia-Salem (P1), Salem-Boise (P2), Boise-Salt Lake City (P2), Helena-Boise (P1), Sacramento-Carson City (P1), Carson City-Boise (P2), Salt Lake City-Denver (P2), Phoenix-Santa Fe (P1), Santa Fe-Denver (P2), and so on.

The kind of intermixture of applications of P1 and P2 demonstrated here is very efficient when the shortest connection network is actually being laid out on a map on which the given terminal set is plotted to scale. With only a few minutes of practice, an example as complex as that of Fig. 1 can be solved in less than 10 minutes. Another mode of procedure, making less use of the flexibility permitted by the construction principles, involves using P1 only once to produce a single fragment, which is then extended by successive applications of P2 until the network is completed. This highly systematic variant, as will emerge later, has advantages for computer mechanization of the solution process. As applied to the example of Fig. 1, this algorithm would proceed as follows if Sacramento were the indicated initial terminal: Sacramento-Carson City, Carson City-Boise, Boise-Salt Lake City, Boise-Helena, Boise-Salem, Salem-Olympia, Salt Lake City-Denver, Denver-Cheyenne, Denver-Santa Fe, and so on.

Since each application of either P1 or P2 reduces the total number of isolated terminals and fragments by one, it is evident that an N-terminal network is connected by N-1 applications.

III. VALIDATION OF CONSTRUCTION PRINCIPLES

The validity of P1 and P2 depends essentially on the establishment of two *necessary* conditions (NC1 and NC2) for a *shortest* connection network (SCN):

Necessary Condition 1 — *Every terminal in a SCN is directly connected to at least one nearest neighbor.*

Necessary Condition 2 — *Every fragment in a SCN is connected to at least one nearest neighbor by a shortest available path.*

To simplify the argument, it will at first be assumed that all distances between terminals are different, so that each terminal or fragment has a single, uniquely defined, nearest neighbor. This restriction will be removed later.

Consider first NC1. Suppose there is a SCN for which it is untrue. Then [Fig. 3(a)] some terminal, t, in this network is not directly joined to its nearest neighbor, n. Since the network is connected, t is necessarily joined directly to some one or more terminals other than n — say f_1, $\cdots$, f_r. For the same reason, n is necessarily joined through some chain, C, of one or more links to one of f_1, $\cdots$, f_r — say to f_k. Now if the link $t - f_k$ is removed from the network and the link $t - n$ is added [Fig. 3(b)], the connectedness of the network is clearly not destroyed — f_k being joined to t through n and C, rather than directly. However, the total length of the network has now been *decreased*, because, by hypothesis, $t - n$ is shorter than $t - f_k$. Hence, contrary to the initial supposition, the network contradicting NC1 could not have been the *shortest*, and the truth of NC1 follows. From NC1 follows P1, which merely *permits* the addition of links which NC1 shows *have* to appear in the final SCN.

Turning now to NC2, the above argument carries over directly if t is thought of as a fragment of the supposed contradictory SCN, rather than as an individual terminal — provided, of course that the $t - n$ link substituted for $t - f_k$ is the shortest link from n to any of the terminals belonging to t. From the validity of NC2 follows P2 — again the links whose addition is permitted by P2 are all necessary, by NC2, in the final SCN.

The temporary restrictive assumption that no two inter-terminal distances are identical must now be removed. A reappraisal of the proofs of NC1 and NC2 shows that they are still valid if n is not the only terminal at distance $t - n$ from t, for in the supposedly contradictory network the distance $t - f_k$ must be *greater* than $t - n$. What remains to be established is that the length of the final connection network resulting

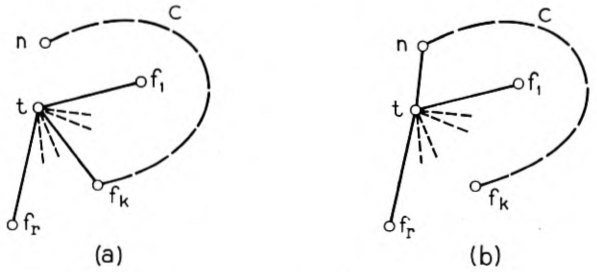


Fig. 3 — Schematic demonstration of NCl.

from successive applications ($N - 1$ for N terminals) of P1 and P2 is independent of *which* nearest neighbor is chosen for connection at a stage when more than one nearest neighbor to an isolated terminal or t is available. This is a consequence of the following considerations: For a prescribed terminal set there are only a *finite* number of connection networks (certainly fewer than $C_{N-1}^{N(N-1)/2}$ — the number of distinct ways of choosing $N - 1$ links from the total of $N(N - 1)/2$ possible links). The length of each one of this finite set of connection networks is a *continuous* function of the individual interterminal distances, d_{ij} (as a matter of f ct, it is a linear function with coefficients 0 and 1). With the d_{ij} specified, the length, L, of a shortest connection network is simply the smallest length in this finite set of connection network lengths. Therefore L is uniquely determined. (It may, of course, be associated with more than one of the connection networks.) Now, if at each stage of construction employing P1 and P2 at which a choice is to be made among two or more nearest neighbors $n_1, \cdots, n_r$ of an isolated terminal (or fragment) t, a small positive quantity, ϵ, is subtracted from any specific one of the distances $d_{tn_1}, \cdots, d_{tn_r}$ — say from d_{tn_k} — the construction will be uniquely determined. The total length, L', of the resulting SCN for the modified problem will now depend on ϵ, as well as on the d_{ij} of the original terminal set. The dependence on ϵ will be *continuous*, however, because the minimum of a finite set of continuous functions of ϵ (the set of lengths of all connection networks of the modified problem) is itself a continuous function of ϵ. Hence, as ϵ is made vanishingly small, L' approaches L, regardless of which "nearest neighbor" links were chosen for shortening to decide the construction.

IV. ABSTRACTION AND GENERALIZATION

In the examples of Figs. 1 and 2, the terminal set to be connected was represented by points on a distance-true map. The decisions involved in applying P1 and P2 could then be based on visual judgements of relative distances, perhaps augmented by application of a pair of dividers in a few close instances. These direct geometric comparisons can of course, be replaced by numerical ones if the inter-terminal distances are entered on the terminal plot, as in Fig. 4(a). The application of P1 and P2 goes through as before, with the relevant "nearest neighbors" determined by a comparison of numerical labels, rather than by a geometric scanning process. For example, P1 applied to Terminal 5 of Fig. 4(a) yields the Link 5-6 of the SCN of Fig. 4(b), because 4.6 is less than 5.6, 8.0, 8.5, and 5.1, and so on.

When the construction of shortest connection networks is thus reduced to processes involving only the numerical distance labels on the various possible links, the *actual location* of the points representing the various terminals in a graphical representation of the problem is, of course. inconsequential. The problem of Fig. 4(a) can just as well be represented by Fig. 5(a), for example, and P1 and P2 applied to obtain the SCN of Fig. 5(b). The original metric problem concerning a set of points in the plane has now been abstracted into a problem concerning *labelled graphs*. The correspondence between the terminology employed thus far and more conventional language of Graph Theory is as follows:

terminal ↔ vertex
possible link ↔ edge
length of link ↔ "length" (or "weight") of edge
connection network ↔ spanning subgraph
(without closed loops) ↔ (spanning subtree)

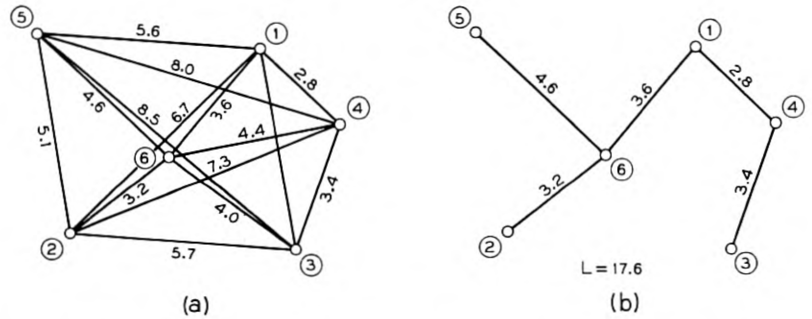


Fig. 4 — Example of a shortest spanning subtree of a complete labelled graph.

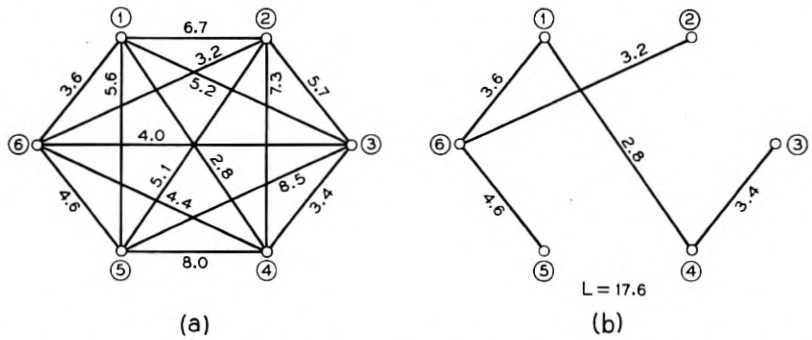


Fig. 5 — Example of a shortest spanning subtree of a complete labelled graph.

shortest connection network ↔ shortest spanning subtree
SCN ↔ SSS

It will be useful and worthwhile to carry over the concepts of "fragment" and "nearest neighbor" into the graph theoretic framework. P1 and P2 can now be regarded as construction principles for finding a shortest spanning subtree of a labelled graph.

In the originating context of connection networks, the graphs from which a shortest spanning subtree is to be extracted are *complete graphs*; that is, graphs having an edge between every pair of vertices. It is natural, now, to generalize the original problem by seeking shortest spanning subtrees for arbitrary connected labelled graphs. Consider, for example, the labelled graph of Fig. 6(a) which is derived from that of Fig. 5(a) by deleting some of the edges. (In the connection network context, this is equivalent to barring direct connections between certain terminal pairs.) It is easily verified that NC1 and NC2, and hence P1 and P2, are valid also in these more general cases. For the example of Fig. 6(a), they yield readily the SSS of Fig. 6(b).

As a further generalization, it is not at all necessary for the validity of P1 and P2 that the edge "lengths" in the given labelled graph be derived, as were those of Figs. 4–6, from the inter-point distances of some point set in the plane. *P1 and P2 will provide a SSS for any connected labelled graph with any set of real edge "lengths."* The "lengths" need not even be positive, or of the same sign. See, for example, the labelled graph of Fig. 7(a) and its SSS, Fig. 7(b). It might be noted in passing that this degree of generality is sufficient to include, among other things, shortest connection networks in an arbitrary number of dimensions.

A further extension of the range of problems solved by P1 and P2 follows trivially from the observation that the *maximum* of a set of

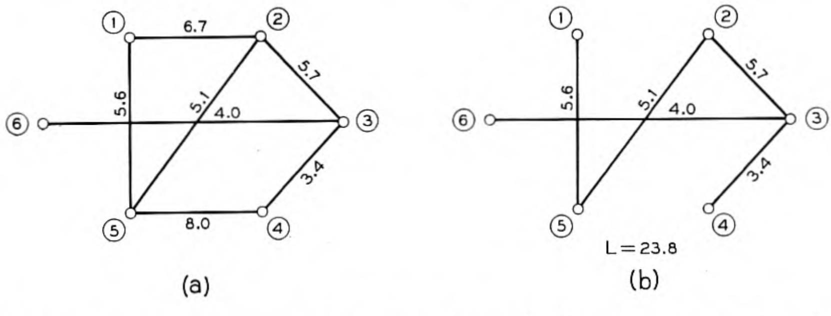


Fig. 6 — Example of a shortest spanning subtree of an incomplete labelled graph.

real numbers is the same as the negative of the *minimum* of the negatives of the set. Therefore, P1 and P2 can be used to construct a *longest spanning subtree* by changing the signs of the "lengths" on the given labelled graph. Fig. 8 gives, as an example, the longest spanning subtree for the labelled graph of Figs. 4(a) and 5(a).

It is easy to extend the arguments in support of NC1 and NC2 from the simple case of minimizing the sum to the more general problems of minimizing an arbitrary increasing symmetric function, or of maximizing an arbitrary decreasing symmetric function, of the edge "lengths" of a spanning subtree. (A symmetric function of n variables is one whose value is unchanged by any interchanges of the variable values; e.g., $x_1 + x_2 + \cdots + x_n$, $x_1 x_2 \cdots x_n$, $\sin 2x_1 + \sin 2x_2 + \cdots + \sin 2x_n$, $(x_1^3 + x_2^3 + \cdots + x_n^3)^{1/2}$, etc.) From this follow the non-trivial generalizations.

> The *shortest spanning subtree* of a connected labelled graph also minimizes all increasing symmetric functions, and maximizes all decreasing symmetric functions, of the edge "lengths."

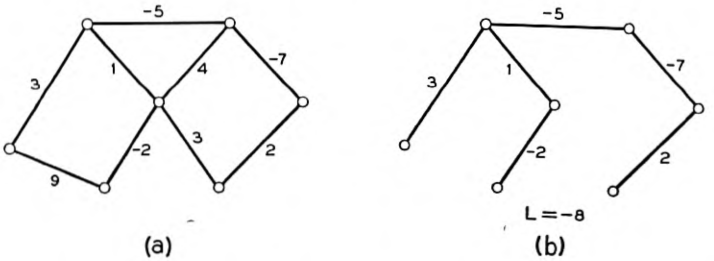


Fig. 7 — Example of a shortest spanning subtree of a labelled graph with some "lengths" negative.

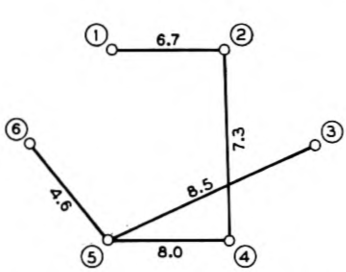


Fig. 8 — The longest spanning subtree of the labeled graph of Figs. 4(a) and 5(a).

The *longest spanning subtree* of a connected labelled graph also maximizes all increasing symmetric functions, and minimizes all decreasing symmetric functions, of the edge "lengths."

For example, with positive "lengths" the same spanning subtree that minimizes the *sum* of the edge "lengths" also minimizes the *product* and the *square root of the sum of the squares.* At the same time, it maximizes the sum of the reciprocals and the product of the arc cotangents.

It seems likely that these extensions of the original class of problems soluble by P1 and P2 contain many examples of practical interest in quite different contexts from the original connection networks. A not entirely facetious example is the following: A message is to be passed to all members of a certain underground organization. Each member knows some of the other members and has procedures for arranging a rendezvous with anyone he knows. Associated with each such possible rendezvous — say between member "i" and member "j" — is a certain probability, p_{ij}, that the message will fall into hostile hands. How is the message to be distributed so as to minimize the over-all chances of its being compromised? If members are represented as vertices, possible rendezvous as edges, and compromise probabilities as "length" labels in a labelled graph, the problem is to find a spanning subtree for which $1 - \Pi(1 - p_{ij})$ is minimized. Since this is an increasing symmetric function of the p_{ij}'s, this is the same as the spanning subtree minimizing $\Sigma\, p_{ij}$, and this is easily found by P1 and P2.

Another application, closer to the original one, is that of minimizing the lengths of wire used in cabling panels of electrical equipment. Restrictions on the permitted wiring patterns lead to shortest connection network problems in which the effective distances between terminals are not the direct terminal-to-terminal distances. Thus the more general viewpoint of the present section is applicable.

V. COMPUTATIONAL TECHNIQUE

Return now to the exemplary shortest connection network problem of Figs. 4(a) and 5(a) which served as the center for discussion of the arithmetizing of the metric factors in the Basic Problem. It is evident that the actual drawing and labelling of all the edges of a complete graph will get very cumbersome as the number of vertices increases — an N-vertex graph has $(1/2)(N^2 - N)$ edges. For large N, it is convenient to organize the numerical metric information in the form of a *distance table,* such as Fig. 9, which is equivalent in content to Fig. 4(a) or Fig.

5(a). (A distance table can also be prepared to represent an incomplete labelled graph by entering the length of non-existent edges as ∞.)

When it is desired to determine a shortest connection network directly from the distance table representation — either manually, or by machine computation — one of the numerous particular algorithms obtainable

	1	2	3	4	5	6
1	—	6.7	5.2	2.8	5.6	3.6
2	6.7	—	5.7	7.3	5.1	3.2
3	5.2	5.7	—	3.4	8.5	4.0
4	2.8	7.3	3.4	—	8.0	4.4
5	5.6	5.1	8.5	8.0	—	4.6
6	3.6	3.2	4.0	4.4	4.6	—

Fig. 9 — Distance table equivalent of Figs. 4(a) and 5(a).

	2	3	4	5	6
F	6.7	5.2	2.8	5.6	3.6
	(1)	(1)	(1)	(1)	(1)

	2	3	5	6
4	7.3	3.4	8.0	4.4

	2	3	5	6
F	6.7	3.4	5.6	3.6
	(1)	(4)	(1)	(1)

	2	5	6
3	5.7	8.5	4.0

	2	5	6
F	5.7	5.6	3.6
	(3)	(1)	(1)

	2	5
6	3.2	4.6

	2	5
F	3.2	4.6
	(6)	(6)

	5
2	5.1

	5
F	4.6
	(6)

Fig. 10 — Illustrative computational determination of a shortest connection network.

by restricting the freedom of choice allowed by P1 and P2 is distinctly superior to other alternatives. This variant is the one in which P1 is used but once to produce a single isolated fragment, which is then extended by repeated applications of P2.

The successive steps of an efficient computational procedure, as applied to the example of Fig. 9, are shown in Fig. 10. The entries in the top rows of the successive F tables are the distances from *the* connected fragment to the unconnected terminals at each stage of fragment growth. The entries in parentheses in the second rows of these tables indicate the nearest neighbor *in the fragment* of the external terminal in question. The computation is started by entering the first row of the distance table into the F table (to start the growing fragment from Terminal 1). A smallest entry in the F table is then selected (in this case, 2.8, associated with External Terminal 4 and Internal Terminal 1). The link 1-4 is deleted from the F table and entered in the Solution Summary (Fig. 11). The remaining entries in the first stage F table are then compared with the corresponding entries in the "4" row of the distance table (reproduced beside the first F table). If any entry of this "added terminal" distance table is smaller than the corresponding F table entry, it is substituted for it, with a corresponding change in the parenthesized index. (Since 3.4 is less than 5.2, the 3 column of the F table becomes 3.4/(4).) This process is repeated to yield the list of successive nearest neighbors to the growing fragment, as entered in Fig. 11. The F and "added terminal" distance tables grow shorter as the number of unconnected terminals is decreased.

This computational procedure is easily programmed for an automatic computer so as to handle quite large-scale problems. One of its advantages is its avoidance of checks for closed cycles and connectedness. Another is that it never requires access to more than two rows of distance data at a time — no matter how large the problem.

SOLUTION SUMMARY	
LINK	LENGTH
1 − 4	2.8
4 − 3	3.4
1 − 6	3.6
6 − 2	3.2
6 − 5	4.6

Fig. 11 — Solution summary for computation of Fig. 10.

VI. RELATED LITERATURE AND PROBLEMS

J. B. Kruskal, Jr.[1] discusses the problem of constructing shortest spanning subtrees for labelled graphs. He considers only distinct and positive sets of edge lengths, and is primarily interested in establishing uniqueness under these conditions. (This follows immediately from NC1 and NC2.) He also, however, gives three different constructions, or algorithms, for finding SSS's. Two of these are contained as special cases in P1 − P2. The third is — "Perform the following step as many times as possible: Among the edges not yet chosen, choose the longest edge whose removal will not disconnect them" While this is not directly a special case of P1 − P2, it is an obvious corollary of the special case in which the *shortest* of the edges permitted by P1 − P2 is selected at each stage. Kruskal refers to an obscure Czech paper[2] as giving a construction and uniqueness proof inferior to his.

The simplicity and power of the solution afforded by P1 and P2 for the Basic Problem of the present paper comes as something of a surprise, because there are well-known problems which *seem* quite similar in nature for which no efficient solution procedure is known.

One of these is *Steiner's Problem*: Find a shortest connection network for a given terminal set, with freedom to add additional terminals wherever desired. A number of necessary properties of these networks are known[3] but do not lead to an effective solution procedure.

Another is the *Traveling Salesman Problem*: Find a closed path of minimum length connecting a prescribed terminal set. Nothing even approaching an effective solution procedure for this problem is now known (early 1957).

REFERENCES

1. J. B. Kruskal, Jr., On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem, Proc. Amer. Math. Soc. **7,** pp. 48–50, 1956.
2. Otakar Borůvka, On a Minimal Problem, Práce Moravské Pridovedecké Spdecnosti, **3,** 1926.
3. R. Courant and H. Robbins, *What is Mathematics*, 4th edition, Oxford Univ. Press, N. Y., 1941, pp. 374 *et* seq.

A Network Containing a Periodically Operated Switch Solved by Successive Approximations

By C. A. DESOER

(Manuscript received June 15, 1956)

This paper concerns itself with the analysis of a type of periodically switched network that might be used in time multiplex systems. The economics of the situation require that the ratio of the switch closure time τ to the switching period T be small. Using this assumption, the analysis is performed by successive approximations. More precisely the zeroth approximation to the transmission is obtained from a block diagram analogous to those used in sampled servomechanisms. From the convergence proof of the successive approximation scheme, it follows that when τ/T is small, the zeroth approximation is very close to the exact transmission. A discussion of some examples is included.

CONTENTS

I. INTRODUCTION

One main contributor to the cost of transmission circuits is the transmission medium itself. Thus it is important to share the transmission medium among as many messages as possible. One possible method is the frequency multiplex where each message utilizes a different frequency band of the whole band available in the medium. An alternate method is the time multiplex where each message is assigned a time slot of duration τ and has access to that time slot once every T seconds. It is obvious that the economics of the situation requires that τ be as small as possible and T as large as possible so that the largest possible number of messages are transmitted over the medium. For this very reason the analysis of periodically switched networks is of special interest in the case where τ/T is small.

W. R. Bennett[4] has published an exact analysis of this problem without any restrictions either on the network or on the ratio τ/T. It is believed, however, that the analysis presented in this paper will, in most practical cases, give the desired answer with a considerable reduction in the amount of calculations. The simplification of the analysis is mainly a result of the assumption that τ/T is small.

First the successive approximation method of solution will be discussed in general terms. Next it will be shown that the zeroth approximation to the transmission through the network can be obtained from the gain of a block diagram analogous to those used in the analysis of sampled servomechanisms. The nature of the zeroth approximation is further clarified by some general discussion and some examples. Next it is shown that the successive approximations converge. The convergence proof then suggests some slight modifications of the block diagram to obtain a more accurate solution.

II. DESCRIPTION OF THE SYSTEM

The system under consideration is shown on Fig. 1. It consists of two reactive networks N_1 and N_2 connected through a switch S which is itself in series with an inductance ℓ. N_1 is driven at its terminal pair (1)

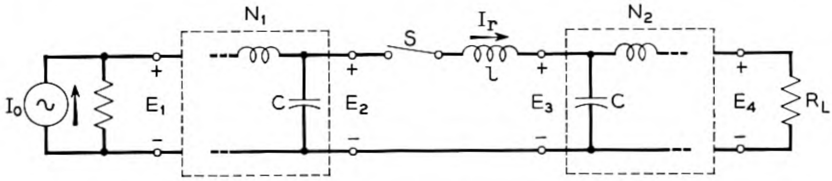
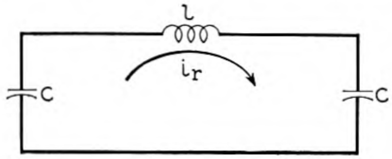


Fig. 1 — System under consideration.



$$f_r = \frac{1}{2\pi\sqrt{\ell \frac{C}{2}}} = \frac{\omega_0}{2\pi} \qquad \tau = \frac{1}{2f_r} = \frac{\pi}{\omega_0} = \pi\sqrt{\ell \frac{C}{2}}$$

Fig. 2 — Resonant circuit.

by a current source I_0 which is shunted by a one ohm resistor. N_2 is also terminated at its terminal pair (1) by a one ohm resistor R_L which is the load resistor of the system. The switch S is periodically closed for a duration τ. The switching period is T. Thus if the switch is closed during the interval $(0, \tau)$ it will be closed during the intervals $(nT, nT + \tau)$ for $n = 1, 2, 3, \cdots$. The inductance ℓ is selected so that the series circuit shown on Fig. 2 has a resonant frequency $f_r = 1/2\tau$; i.e., the time τ during which the switch is closed is exactly one-half period of the circuit of Fig. 2.

The switch S acts as a sampler and, as a result of the well-known modulating properties of sampled systems, the sampling period T must be chosen such that the frequency $1/2T$ is larger than any of the frequencies present in the signals generated by I_0. Furthermore, in order to eliminate all the sidebands generated by the switching, N_2 must have a high insertion loss for all frequencies above $1/2T$ cps.

In the analysis that follows networks N_1 and N_2 will be assumed to be identical: it should, however, be stressed that this assumption is not necessary for the proposed method of analysis.* This assumption is made because in the practical situation which motivated this analysis N_1 and N_2 were identical since transmission in both directions was required.

In order for the system under consideration to achieve the maximum degree of multiplexing, the closure time τ of the switch will be taken as small as practically possible and the switching period T as large as possible (consistent with the bandwidth of the signals to be transmitted). As a result the ratio τ/T is very small, of the order of 10^{-2} or less in practical cases. Consequently the resonant frequency f_r of the series resonant circuit shown on Fig. 2, is many times larger than any of the natural frequencies of N_1 and N_2.

* The modifications required for the case where N_1 is not identical to N_2 are given in Appendix IV.

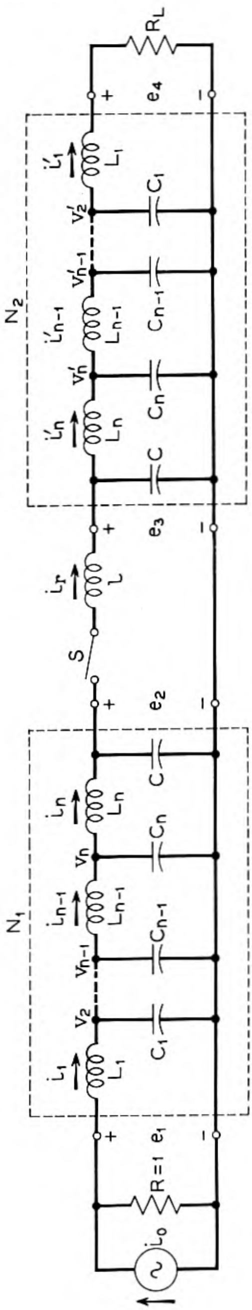


Fig. 3 — System under consideration when N_1 and N_2 are lossless ladders.

The problem is to determine the relation between E_4, the voltage across R_L, and I_0.

III. METHOD OF SOLUTION

Let us first write the equations of the system. Obviously the equations will depend on the exact configuration of the networks N_1 and N_2. For simplicity we shall write them for the case where N_1 and N_2 are dissipationless low-pass ladder networks. As will become apparent later this assumption is not essential to the argument. What is essential, however, is the fact that both N_1 and N_2 should start (looking in from the switch) with a shunt capacitor C and a series inductance L_n, the element value of L_n being much larger than ℓ. Using a method of analysis advocated by T. R. Bashkow,[5] we obtain, for the network of Fig. 3, the equations:

$$
\left.\begin{aligned}
L_1 \frac{di_1}{dt} &= -Ri_1 - v_2 + Ri_0 \\
C_1 \frac{dv_2}{dt} &= i_1 - i_2 \\
&\vdots \\
C_n \frac{dv_n}{dt} &= i_{n-1} - i_n \\
L_n \frac{di_n}{dt} &= v_n - e_2 \qquad (1.\text{a})
\end{aligned}\right\} I_a
$$

$$
\left.\begin{aligned}
C \frac{de_2}{dt} &= i_n - i_r\Delta(t) \qquad (1.\text{b}) \\
\ell \frac{di_r}{dt} &= [e_2 - e_3]\Delta(t) \qquad (1.\text{c}) \\
C \frac{de_3}{dt} &= i_r\Delta(t) - i_n' \qquad (1.\text{d})
\end{aligned}\right\} R \qquad (1)
$$

$$
\left.\begin{aligned}
L_n \frac{di_n'}{dt} &= e_3 - v_n' \\
C_n \frac{dv_n'}{dt} &= i_n' - i_n'_{-1} \\
&\vdots \\
C_1 \frac{dv_2'}{dt} &= i_2' - i_1' \\
L_1 \frac{di_1'}{dt} &= v_2' - R_L i_1'
\end{aligned}\right\} I_b
$$

where

$$\Delta(t) = \sum_{k=-\infty}^{+\infty} [u(t - kT) - u(t - kT - \tau)], \tag{2}$$

with $u(t) = 1$ for $t > 0$, and $u(t) = 0$ for $t < 0$.

This system of linear time varying equations may be broken up into three sub-systems I_a, R and I_b. It is this subdivision that suggests a successive approximation scheme that will be shown to converge to the exact solution.

The zeroth approximation is obtained as follows: when the switch is closed, i.e., $\Delta(t) = 1$, the resonant current i_r is much larger than the currents i_n and i_n'. Thus, during the switch closure time, i_n and i_n' are neglected with respect to i_r in (1.b) and (1.d). Hence when $\Delta(t) = 1$ the system R may be solved for $i_r(t)$, $e_2(t)$ and $e_3(t)$ in terms of the initial conditions. The resulting function $e_2(t)$ and given function $i_0(t)$ are then the forcing functions of the system I_a. The other function $e_3(t)$ is the forcing function of the system I_b. Under these assumptions, the periodic steady-state solution corresponding to an applied current $i_0(t) = I_0 e^{i\omega t}$ is easily obtained.

The zeroth approximation will be distinguished by a subscript "0". Thus $i_{r0}(t)$ is the (steady state) zeroth approximation to the exact solution $i_r(t)$.

The first approximation will be the solution of the system (1), provided that during the switch closure time the functions $i_n(t)$ and $i_n'(t)$ in (1.b) and (1.d) are respectively replaced by the known functions $i_{n0}(t)$ and $i_{n0}'(t)$. And, more generally, the $(k + 1)$th approximation will be the solution of (1) provided that during the switch closure time, the functions $i_n(t)$ and $i_n'(t)$, in (1.b) and (1.d), are respectively replaced by the known solutions for $i_n(t)$, and $i_n'(t)$ given by the kth approximation It will be shown later that this successive approximation scheme converges. Let us first describe a simple method for obtaining the zeroth approximation.

IV. THE ZEROTH APPROXIMATION

4.1 *Introduction*

The problem is to obtain the steady-state solution of (1) under the excitation $i_0(t) = I_0 e^{i\omega t}$. Using the approximations indicated above, during the switch closure time (that is when $\Delta(t) = 1$) the system R becomes

$$C\,\frac{de_2}{dt} = -i_r(t)\Delta(t), \tag{3}$$

$$\ell\,\frac{di_r}{dt} = [e_2 - e_3]\Delta(t), \tag{4}$$

$$C\,\frac{de_3}{dt} = i_r(t)\Delta(t). \tag{5}$$

Differentiating the middle equation and eliminating de_2/dt and de_3/dt we get for $0 \leqq t < \tau$:

$$\frac{d^2 i_r}{dt^2} = -\frac{2}{\ell C}\,i_r\Delta(t) + \frac{1}{\ell}\,[e_2(t) - e_3(t)]\delta(t) \tag{6}$$

in which we used the notation $\delta(t)$ for the Dirac function and the knowledge that

$$\frac{d\Delta(t)}{dt} = \delta(t) - \delta(t - \tau). \tag{7}$$

Equation (6) represents the behavior of the resonant circuit of Fig. 2 for the following initial conditions:

$$i_r(0+) = 0, \tag{8}$$

$$\frac{di_r(0+)}{dt} = \frac{e_2(0) - e_3(0)}{\ell}. \tag{9}$$

In Appendix I it is shown that the resulting current $i_r(t)$ is, for the interval $0 \leqq t < \tau$,

$$i_r(t) = C[e_2(0) - e_3(0)]s_1(t), \tag{10}$$

where

$$s_1(\ ) = \begin{cases} \dfrac{\pi}{2\tau}\sin\dfrac{\pi t}{\tau} = \dfrac{1}{2}\,\omega_0 \sin \omega_0 t & \text{for} \quad 0 \leqq t < \tau \\ 0 \text{ elsewhere} \end{cases} \tag{11}$$

with

$$\omega_0 = \frac{\pi}{\tau} = \sqrt{\frac{2}{\ell C}}. \tag{12}$$

Thus the zeroth approximation to the exact $i_r(t)$ is given for the interval $0 \leqq t \leqq T$ by

$$i_{r0}(t) = C[e_2(0) - e_3(0)]s_1(t). \tag{13}$$

We shall now show that the zeroth approximation may be conveniently obtained from the block diagram of Fig. 4.

4.2 *Description of the Block Diagram*

All the blocks of the block diagram are unilateral and their corresponding transfer functions are defined in the following. Capital symbols represent $\mathfrak{L}$-transform of the corresponding time functions, thus $I_0(p)$ is the $\mathfrak{L}$-transform of $i_0(t)$.

Referring to Fig. 1,

$$Z_{12}(p) = \left.\frac{E_2(p)}{I_0(p)}\right|_{I_r=0}.$$

Thus $Z_{12}(p)$ represents the transfer impedance of N_1 when its output is open-circuited (i.e., $I_r = 0$). Since N_1 and N_2 are identical we also have, from $R_L = 1$ and reciprocity, $Z_{12}(p) = E_4/I_r$, where I_r is the current entering N_2.

The impulse modulator is periodically operated every T seconds, and has the property that if its input is a continuous function $f(t)$ its output is a sequence of impulses:

$$\sum_{-\infty}^{+\infty} f(t)\,\delta(t - kT).$$

The transfer function $S_1(p)$ is defined by

$$S_1(p) = \mathfrak{L}[s_1(t)] = \frac{\omega_0^2}{p^2 + \omega_0^2}\cosh\frac{p\tau}{2}\,e^{-p\tau/2}. \tag{14}$$

Let $Z(p)$ be the driving point impedance at the terminal pair (2) of N_1. It is also that of N_2 since N_1 and N_2 are assumed to be identical.

Let $V(p)$ be the output of the first block, then, by definition, $V(p) = Z_{12}(p)I_0$. Let $v(t)$ be the corresponding time function. The voltage $v(t)$

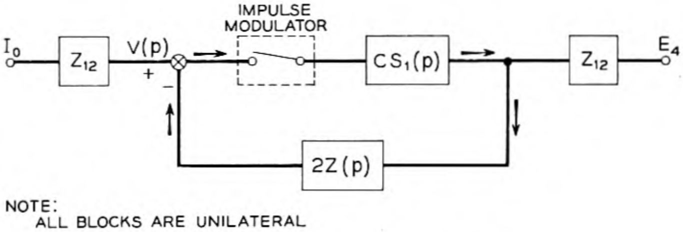


Fig. 4 — Zeroth-approximation block diagram.

is the output voltage of N_1, when N_1 is excited by the current source I_0 and the switch S remains open at all times.

4.3 *Analysis of the Block Diagram*

For simplicity, suppose that the system starts from a relaxed condition (i.e., no energy stored) at $t = 0$. Let $z(t) = \mathfrak{L}^{-1}[Z(p)]$. Considering the network N_1 as driven by i_0 and i_{r0}, it follows that the voltage $e_2(t)$ shown on Fig. 3 is given by

$$e_{20}(t) = v(t) - \int_0^t i_{r0}(t')z(t - t')\,dt'. \tag{15}$$

Similarly

$$e_{30}(t) = \int_0^t i_{r0}(t')z(t - t')\,dt'. \tag{16}$$

Thus

$$e_{20}(t) - e_{30}(t) = v(t) - 2\int_0^t i_{r0}(t')z(t - t')\,dt'. \tag{17}$$

These equations have been derived by considering Fig. 1. They could have been also derived from the block diagram of Fig. 4 as follows: let $I_{r0}(p)$ be the output of $CS_1(p)$. As a result, the output of the block $2Z(p)$ is $2Z(p)I_{r0}(p)$. When this latter quantity is subtracted from $V(p)$ one gets $V(p) - 2Z(p)I_{r0}(p)$, which is the $\mathfrak{L}$-transform of the right-hand side of (17). Referring to the block diagram it is also seen that this quantity is the input to the impulse modulator.

Thus we see that if $I_{r0}(p)$ is the output of $CS_1(p)$, then the input of the impulse modulator is $e_{20}(t) - e_{30}(t)$ by virtue of (17). If this is the case the output of $CS_1(p)$ is given by $C[e_{20}(0) - e_{30}(0)]s_1(t)$, for $0 \leqq t < T$, which, according to (9), is $i_{r0}(t)$.

Thus the block diagram of Fig. 4 is a convenient way of obtaining the zeroth approximation to the periodic steady-state solution.

In order to use the techniques developed for sampled data systems,[1, 2] we introduce the following notation.[2] If $f(t) = \mathfrak{L}^{-1}[F(p)]$, then we define $F^*(p)$ by the relation

$$F^*(p) = \frac{1}{T}\sum_{n=-\infty}^{+\infty} F(p + jn\omega_s), \tag{18}$$

where

$$\omega_s = \frac{2\pi}{T}. \tag{19}$$

If $f(0+)$ is defined by $\lim_{\epsilon \to 0} f(\epsilon^2)$, then, provided $f(0+) = 0$,*

$$F^*(p) = \mathfrak{L}\left[\sum_{n=0}^{\infty} f(t)\,\delta(t - nT)\right]. \tag{20}$$

Going back to the system of Fig. 4 we get[2]

$$E_{40}(p) = \frac{[Z_{12}(p)I_0(p)]^* C S_1(p) Z_{12}(p)}{1 + 2C[S_1(p)Z(p)]^*}, \tag{21}$$

and

$$I_{ro}(p) = \frac{[Z_{12}(p)I_0(p)]^* C S_1(p)}{1 + 2C[S_1(p)Z(p)]^*}, \tag{22}$$

where according to the notation defined by (18)

$$[Z_{12}(p)I_0(p)]^* = \frac{1}{T}\sum_{n=-\infty}^{\infty} Z_{12}(p + jn\omega_s)I_0(p + jn\omega_s),$$

$$[S_1(p)Z(p)]^* = \frac{1}{T}\sum_{n=-\infty}^{+\infty} S_1(p + jn\omega_s)Z(p + jn\omega_s).$$

It should be stressed that (21) and (22) are not valid when τ is made identical to zero. When $\tau = 0$, $S_1(p) = 1$ for all p's and since $Z(p) \sim 1/Cp$ as $p \to \infty$ the time function whose transform is $Z(p)S_1(p)$ is different from zero at $t = 0$. In such a case (20) does not hold. From a physical point of view, the feedback loop of Fig. 4 is unstable when τ is identically zero since an impulse generated by the impulse modulator produces instantaneously a step at the input of the impulse modulator. This step causes an instantaneous jump in the measure of the impulse at the output of the impulse modulator and so on. In short the feedback loop is unstable.

It should be pointed out that if the power density spectrum of I_0 is zero for frequencies higher than $\omega_s/2$, (21) reduces to

$$\frac{E_{40}}{I_0} = \frac{1}{T}\frac{C S_1(p) Z_{12}^{\,2}(p)}{1 + 2C[S_1(p)Z(p)]^*} \quad \text{for} \quad |\,p\,| < \frac{\omega_s}{2}. \tag{23}$$

For certain applications it is convenient to rewrite (21) in a slightly

---

* When $f(0)$, as defined above, is different from zero, (20) should be replaced by

$$\mathfrak{L}\left[\sum_{n=0}^{\infty} f(t)\,\delta(t - nT)\right] = F^*(p) + \frac{1}{2}f(0+).$$

different form. Advancing the time function $s_1(t)$ by $\tau/2$ seconds, one gets the function $s_0(t)$ which is even in t. As a result its transform $S_0(p)$ is purely real, that is,

$$S_0(p) = \frac{\omega_0^2}{p^2 + \omega_0^2} \cosh \frac{p\tau}{2}.$$

From an analysis carried out in detail in Appendix IV we finally obtain

$$E_{40}(p) = \frac{[Z_{12}(p)I_0(p)]^* S_0(p) Z_{12}(p)}{2[Z(p)S_0(p)]^*}. \tag{24}$$

It should be pointed out that (23) is still valid when $\tau = 0$. Equations (20) and (23) give the zeroth approximation to the gain of the system for any driving current $i_0(t)$.

In many cases it is sufficient to know only the steady-state response $E_{40}(p)$ to an input $i_0(t) = I_0 e^{j\omega_0 t}$. The response $E_{40}(p)$, as given by (23) [or (20)] includes both transient and steady-state terms. Since $I_0(p) = \dfrac{I_0}{p - j\omega_0}$ equation (24) gives

$$E_{40}(p) = \frac{\left(\dfrac{1}{T} \displaystyle\sum_{n=-\infty}^{+\infty} Z_{12}(p + jn\omega_s) \frac{I_0}{p + jn\omega_s - j\omega_0}\right) S_0(p) Z_{12}(p)}{2[Z(p)S_0(p)]^*}. \tag{25}$$

Since neither $S_0(p)$ nor $Z_{12}(p)$ have poles on the imaginary axis, the steady state includes only the terms corresponding to the imaginary axis poles of the summation terms. Thus the steady-state response is of the form

$$\sum_{-\infty}^{+\infty} A_n\, e^{j(\omega_0 - n\omega_s)t},$$

where, from (25),

$$A_n = \frac{I_0 Z_{12}(j\omega_0) S_0(j\omega_0 - jn\omega_s) Z_{12}(j\omega_0 - jn\omega_s)}{2 \displaystyle\sum_{k=-\infty}^{+\infty} Z[j\omega_0 + j(k - n)\omega_s] S_0[j\omega_0 + j(k - n)\omega_s]}. \tag{26}$$

V. TRANSMISSION LOSS

A practically important question is to find out *a priori* whether a switched filter necessarily introduces some transmission loss.

The following considerations apply exclusively to the zeroth order approximation. It will be shown that assuming ideal elements, the transmission at dc may have as small a loss as desired.

By transmission at dc we mean the ratio of the dc component of the steady state output voltage to the intensity of the applied direct current. Thus we refer to (26) and set $\omega_0 = 0$ and $n = 0$. Suppose the lossless networks N_1 and N_2 are designed so that their transfer impedance Z_{12} is of the Butterworth type, that is

$$| Z_{12}(j\omega) |^2 = \frac{1}{1 + \omega^{2M}},$$

where for our purposes M is a large integer.

In the following sum, which is the denominator of (26) when $\omega_0 = n = 0$

$$2 \sum_{k=-\infty}^{+\infty} Z(jk\omega_s)S_0(jk\omega_s),$$

(where $\omega_s > 2$ since the cutoff of the networks N occurs at $\omega = 1$), the terms corresponding to values of $k \neq 0$ will make a contribution that vanishes as $M \to \infty$. This is a consequence of the following facts:

(a) $\mathrm{Re}[Z(jk\omega_s)] = | Z_{12}(jk\omega_s) |^2$, since the networks N_1 and N_2 are dissipationless. Hence for $k \neq 0$ and as $M \to \infty$ $\mathrm{Re}[Z(jk\omega_s)] \to 0$,

(b) $\mathrm{Im}[Z(jk\omega_s)] = -\mathrm{Im}[Z(-jk\omega_s)]$,

(c) $S_0(j\omega)$ is real.

Thus the imaginary part of the products $Z(jk\omega_s)S_0(jk\omega_s)$ cancel out and the real part (for $k \neq 0$) decreases exponentially to zero as $M \to \infty$. Hence for sufficiently large M the denominator of (26) may be made as close to two as desired.

It is easy to check that the numerator of (26) reduces to I_0, the intensity of the applied direct current. Therefore the ratio of A_0, the dc component of the output voltage to I_0 may be made as close to one-half as desired.

VI. A SIMPLE EXAMPLE

Since the approximate formulae derived in Section IV are somewhat unfamiliar it seems proper to consider in a rather detailed manner a simple example.*

Consider the system of Fig. 5. Assume that the current source applies a constant current to the system and assume that the steady state is reached. For simplicity let $I_0R = E$.

The steady-state behavior of the voltages $e_2(t)$ and $e_3(t) = e_4(t)$ is

* In addition, the limiting case of the sampling rate $\to \infty$, i.e., $T \to 0$, is treated in Appendix II.

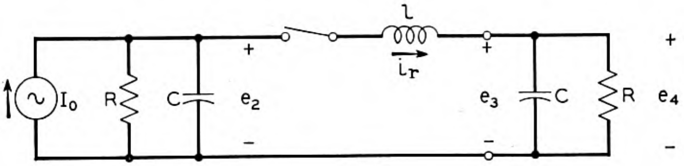


Fig. 5 — A simple example.

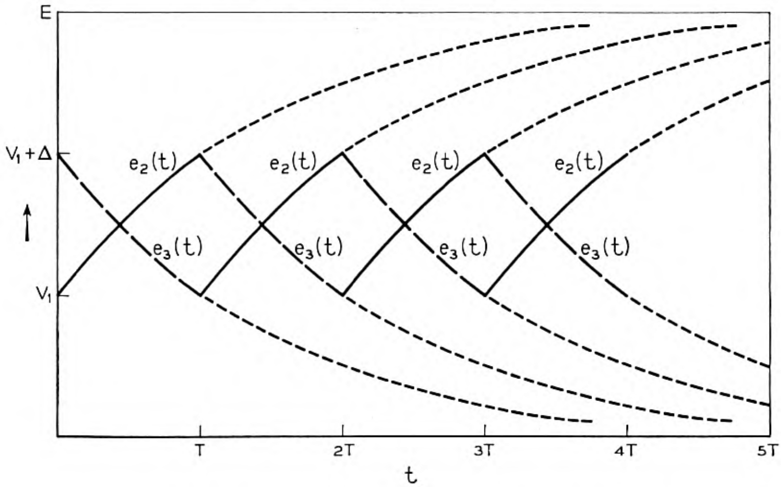


Fig. 6 — Waveforms of the network of Fig. 5.

shown on Fig. 6. It is further assumed that the duration τ during which the switch is closed is negligible compared to T, the interval between two successive closures.

Let $\bar{e}_4$ be the average value of the steady-state voltage $e_4(t)$. Thus $\bar{e}_4$ is equal to A_0 as given by (26) with $\omega_0 = n = 0$, namely,

$$e_4 = I_0 \frac{Z_{12}^2(0)\ S_0(0)}{2 \sum_{-\infty}^{+\infty} Z(jk\omega_s)\ S_0(jk\omega_s)} .$$

In this particular case

$$Z(p) = Z_{12}(p) = \frac{R}{1 + pRC} = \frac{C^{-1}}{p + \frac{1}{RC}} .$$

Since we assume τ to be infinitesimal $S_0(p)$ and $S_1(p)$ may be considered equal to unity over the band of interest. Using the expansion

$$\coth z = \frac{1}{z} + \sum_{1}^{\infty} \frac{2z}{z^2 + n^2\pi^2},$$

we obtain

$$Z^*(p) = \frac{1}{T} \sum_{-\infty}^{+\infty} \frac{C^{-1}}{(p + jn\omega_s) + \dfrac{1}{RC}} = \frac{1}{2C} \coth \left[\left(p + \frac{1}{RC} \right) \frac{T}{2} \right].$$

Hence

$$\frac{2}{T} \sum_{-\infty}^{+\infty} Z(jk\omega_s) = 2Z^*(p) \bigg|_{p=0} = \frac{1}{C} \coth \left(\frac{T}{2RC} \right).$$

Thus finally

$$\bar{e}_4 = \frac{I_0 R^2}{T \dfrac{1}{C} \coth \left(\dfrac{T}{2RC} \right)} = E \frac{RC}{T} \frac{1}{\coth \left(\dfrac{T}{2RC} \right)}. \tag{27}$$

This last result obtained from the theory developed above is now going to be checked directly. Referring to Fig. 6, where the notation is defined, and noting the periodicity of the boundary conditions, we get

$$(V_1 + \Delta)e^{-(T/RC)} = V_1,$$
$$(E - V_1)e^{-(T/RC)} = E - (V_1 + \Delta).$$

Noting that $e_4(t) = (V_1 + \Delta)e^{-t/RC}$, and solving for V_1 and Δ we finally get

$$e_4(t) = \frac{e^{-t/RC}}{1 + e^{-T/RC}}.$$

By definition

$$\overline{e_4} = \frac{1}{T} \int_0^T e_4(t)\, dt = \frac{E\, RC}{T} \frac{1 - \bar{e}^{-T/RC}}{1 + \bar{e}^{-T/RC}},$$

or

$$\overline{e_4} = E \frac{RC}{T} \frac{1}{\coth \left(\dfrac{T}{2RC} \right)}.$$

This last equation checks with (27).

VII. NUMERICAL EXAMPLES

Consider the network of Fig. 7. The cutoff of both N_1 and N_2 occurs at $\omega = 1$. In view of the sampling theorem good transmission requires

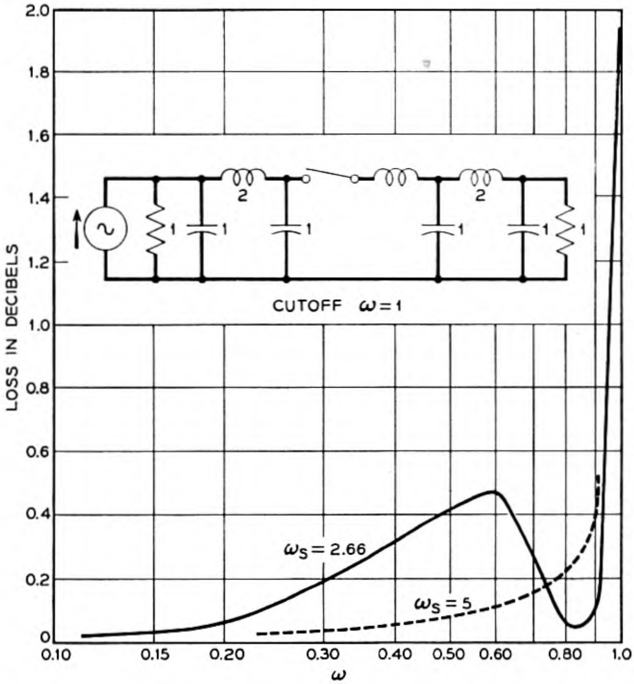


Fig. 7 — Computed transmission loss.

that the signal be sampled at a rate at least twice as large as its highest frequency component. Since the cutoff occurs at $\omega = 1$, the sampling angular frequency should at least be equal to 2. For illustration purposes we have taken $\omega_s = 2.67$ and $\omega_s = 5$ for the angular sampling frequency The value $\omega_s = 2.67$ corresponds to a cutoff at 3 kc and a sampling rate. of 8 kc. The ratio τ/T was taken to be 1/125. The transmission through the switched network as given by the zeroth approximation is shown for both cases on Fig. 7.

As expected the transmittance of the switched filter gets closer to that of an ordinary filter as the switching frequency increases.

VIII. THE SUCCESSIVE APPROXIMATION SCHEME

The ideas involved in the successive approximation scheme are simple and straightforward. One point remains to be settled, namely the convergence of the procedure.

We shall assign a subscript 1 to the correction to be applied to the zeroth approximation in order to obtain the first approximation. Thus adding $i_{r1}(t)$ to $i_{r0}(t)$ we get the first approximation $i_{r0}(t) + i_{r1}(t)$. More generally the kth approximation is $\sum_{n=0}^{k} i_{rn}(t)$. The procedure will converge if, in particular, the infinite series $\sum_{n=0}^{\infty} i_{rn}(t)$ converges.

8.1 *Preliminary Steps*

(a) Let us normalize the frequency (and consequently the time) so that the switching period T is unity. Since the networks N_1 and N_2 must have high insertion loss for $\omega > \frac{1}{2}(2\pi/T) = \pi$, the pass band of N_1 and N_2 must be the order of 1 radian/sec. As a result the element values of the capacitor C and the inductance L_n (see Fig. 3) are also $0(1)$.

(b) For the excitation $i_0 = e^{i\omega t}$, the zeroth approximation derived above may be written in terms of Fourier components:

$$i_{r0}(t) = e^{i\omega t} \sum_{k=-\infty}^{+\infty} I_{r0,k} e^{i2\pi kt},$$

$$i_{n0}(t) = e^{i\omega t} \sum_{k=-\infty}^{+\infty} I_{n0,k} e^{i2\pi kt}.$$

Let $\bar{i}_{r0}$ denote the complex conjugate of i_{r0}, then

$$i_{r0}(t)\bar{i}_{r0}(t) = \sum_{k,\ell=-\infty}^{+\infty} I_{r0,k} \bar{I}_{r0,\ell} e^{2\pi i(k-\ell)t}.$$

Since the functions $e^{2\pi ikt}[k = \cdots -1, 0, 1, \cdots]$ are orthonormal over the interval $(0, 1)$ and form a complete set,[3] we have from Bessel's equality:[3]

$$\int_0^1 | i_{r0}(t) |^2 \, dt = \sum_{k=-\infty}^{+\infty} | I_{r0,k} |^2 = N(I_{r0}),$$

where $N(I_{r0})$ denotes the norm of the vector I_{r0} which is defined by its components $I_{r0,k}(k = \cdots -1, 0, +1 \cdots)$. Similarly,

$$\int_0^1 | i_{n0}(t) |^2 \, dt = \sum_{k=-\infty}^{+\infty} | I_{n0,k} |^2 = N(I_{n0}).$$

(c) Since the switch is periodically closed we shall be interested in the Fourier series expansion of $\Delta(t)$:

$$\Delta(t) = u(t) - u(t - \tau) = \sum_{-\infty}^{+\infty} \Delta_k e^{ik2\pi t},$$

where $\Delta_0 = \tau$ and

$$\sum_{-\infty}^{+\infty} | \Delta_k |^2 = \tau.$$

Since $\tau/T \ll 1$, and since the frequency has been normalized so that $T = 1$, we have $\tau \ll 1$.

Using the convolution in the frequency domain, we have

$$i_{n0}(t)\Delta(t) = \sum_{k=-\infty}^{+\infty} \left(\sum_{\alpha=-\infty}^{+\infty} \Delta_{k-\alpha} I_{n0,\alpha} \right) e^{i(\omega+2\pi k)t}.$$

If we introduce the infinite matrix G defined by

$$G_{ik} = \Delta_{i-k} \; (i, k = -\infty, \cdots, -1, 0, +1, \cdots \infty),$$

the convolution may be represented by the product, GI_{n0}, where I_{n0} is the vector whose components are $I_{n0,k}(k = \cdots -1, 0, 1, \cdots)$.

(d) Considering the network shown on Fig. 3, let $E(p)$ be the ratio of $I_n'(p)$ to $I_r(p)$. Taking into account the assumed identity between N_1 and N_2 it follows that

$$\left. \frac{+I_n(p)}{I_r(p)} \right|_{I_0=0} = \frac{I_n'(p)}{I_r(p)} = E(p).$$

Using the system of (1) and, for example, by Neumann series expansion of the inverse matrix, we get

$$E(p) = \frac{1}{L_n C p^2} - \frac{2}{L_n^2 C^2 p^4} + \cdots$$

(e) Considering now the effect of $i_n(t)$ and $i_n'(t)$ on $i_r(t)$, (42) of Appendix I gives $I_r(p)$ as a function of $I_n(p)$ and $I_n'(p)$. In the present discussion where we are interested in the steady state of $i_r(t)$ it is essential to keep in mind that since the switch opens at $t = \tau$, the memory of the resonant circuit extends only over an interval $0 < t \leqq \tau$. To take this into account we must modify the factor $(\omega_0^2/2)/(p^2 + \omega_0^2)$ of (40), because the impulse response (which represents this memory) must be identically zero for $t > \tau$. The resulting new expression is

$$F(p) = \frac{\frac{\omega_0^2}{2}}{p^2 + \omega_0^2} e^{-p\tau/2} [e^{p\tau/2} + e^{-p\tau/2}],$$

or

$$F(p) = \frac{\omega_0^2}{p^2 + \omega_0^2} e^{-p\tau/2} \cosh \frac{p\tau}{2}.$$

Since the time function whose transform is $F(p)$ is non-negative for all t's and since $F(0) = 1$, it follows that

$$|F(j\omega)| \leqq 1. \tag{28}$$

8.2 *Matrix description of the successive approximations*

From the developments of Section IV, we know $i_{r0}(t)$, $i_{n0}(t)$ and $i_{n0}'(t)$ or what is equivalent, the vectors I_{r0}, I_{n0} and I_{n0}'. The first approximation takes into account the effect of $i_{n0}(t)$ and $i_{n0}'(t)$ on $i_r(t)$. [See equation (1.c) and (1.d)]. The time functions $i_{n0}(t)$ and $i_{n0}'(t)$ affect the system R only during the interval $(0, \tau)$. Therefore we must consider the vector $G(I_{n0} + I_{n0}')$ which corresponds to the excitation of the resonant circuit. Since the opening of the switch after a closure time τ forcibly brings $i_r(t)$ to zero we have

$$I_{r1} = G\,F\,G(I_{n0} + I_{n0}'), \tag{29}$$

where the matrix G has been defined above and the matrix F is a diagonal matrix whose diagonal elements F_k ($k = \cdots -1, 0, +1 \cdots$) are defined by $F_k = F(j\omega_s + j2\pi k)$. Note that (28) implies that $|F_k| \leqq 1$ for all k's. It should be kept in mind that $I_{r0} + I_{r1}$ is the first approximation to the exact $I_r(p)$.

The next iteration is obtained by first taking into account the effect of I_{r1} on the rest of the network:

$$\begin{aligned} I_{n1} &= E\,I_{r1}, \\ I_{n1}' &= E\,I_{r1}, \end{aligned} \tag{30}$$

where E is a diagonal matrix whose elements E_k ($k = \cdots, -1, 0, +1, \cdots$) are defined by $E_k = E(j\omega_s + 2\pi kj)$, and then the effects of I_{n2} and I_{n2}' on I_r, that is,

$$I_{r2} = G\,F\,G\,(I_{n1} + I_{n1}'), \tag{31}$$

combining (30) and (31), $I_{r2} = 2\,G\,F\,G\,E\,I_{r1}$. A repetition of the same procedure would lead to $I_{r3} = 2\,G\,F\,G\,E\,I_{r2}$, and in general $I_{rn+1} = 2\,G\,F\,G\,E\,I_{rn}$.

Since the nth approximation to $I_r(p)$ is given by the sum $\sum_{k=0}^{n} I_{rk}$, the successive approximation scheme will be convergent only if the series

$$\sum_{k=0}^{\infty} I_{rk}$$

converges. This will be the case if and only if the series

$$[1 + 2\,G\,F\,G\,E + \cdots + (2\,G\,F\,G\,E)^n + \cdots]I_{r1} \tag{32}$$

converges.

8.3 *Convergence Proof*

Consider a vector X of bounded norm corresponding to a time function $x(t)$ having the property that $x(t) = 0$ for $\tau \leqq t \leqq T$ and $x(t) \neq 0$ for $0 < t < \tau$. In the above scheme, the vector X would be I_{rn}. Let us define the vectors Y, Z, U and V by the relations

$$Y = EX, \tag{33}$$

$$Z = GY, \tag{34}$$

$$U = FZ, \tag{35}$$

$$V = 2GU, \tag{36}$$

hence

$$V = 2\,G\,F\,G\,E\,X. \tag{37}$$

We wish to show that $N(V) \leqq \alpha N(X)$ with $\alpha < 1$, since these inequalities imply that the infinite series (32) converges.

Since (a) N_1 and N_2 are low-pass filters with cutoff $\leqq \pi$ radians/sec, (b) $E(p) = 1$ for $p = 0$, (c) $E(p) \propto 1/L_nCp^2$ for $p \gg 1$, only a few of the E_k's will be of the order of unity In most cases E_{-1}, E_0, E_1 will be smaller than unity, thus,

$$N(Y) \leqq N(X). \tag{38}$$

In view of the pulsating character of $x(t)$ the power spectrum of $x(t)$ is almost constant up to frequencies of the order of π/τ radians/sec. Because of the low-pass characteristic of $E(p)$, the function $y(t)$ associated with the vector Y is smooth in comparison to $x(t)$, thus from (34),

$$N(Z) = \int_0^1 |\, z(t)\,|^2\, dt = \int_0^\tau |\, y(t)\,|^2\, dt = a\tau N(Y),$$

where $a = 0(1)$.

Since $|\, F_k\,| \leqq 1$ for all k's, from (33), $N(U) \leqq N(Z)$, hence $N(U) = b\tau N(Y)$ with $b = 0(1)$.

$$N(U) = b\tau N(Y) \quad \text{with} \quad b = 0(1).$$

From (36) we have

$$N(V) = 2\int_0^\tau |\, u(t)\,|^2\, dt \leqq 2\int_0^1 |\, u(t)\,|^2\, dt = 2N(U).$$

Thus we finally get

$$N(V) = 2b\tau N(Y) \quad \text{where} \quad b = 0(1), \tag{39}$$

and since $\tau \ll 1$ we get from (38) and (39) $N(V) = \alpha N(X)$ with $\alpha < 1$. Hence the convergence is established.

IX. A MODIFICATION OF THE BLOCK DIAGRAM TO IMPROVE THE ZEROTH APPROXIMATION

In principle it is possible to obtain a block diagram whose transmission characteristic is equal to the first approximation. In many cases it is not necessary to go that far. The first approximation takes into account the effect of the currents $i_{n0}(t)$ and $i_{n0}'(t)$ on the resonant circuit of Fig. 2. Since during the switch closure time the currents i_{n0} and i_{n0}' cannot vary much, let us assume that they remain constant for the duration of the switch closure.

Referring to the analysis of Appendix I and to (42) in particular, we see that the current i_r is increased by

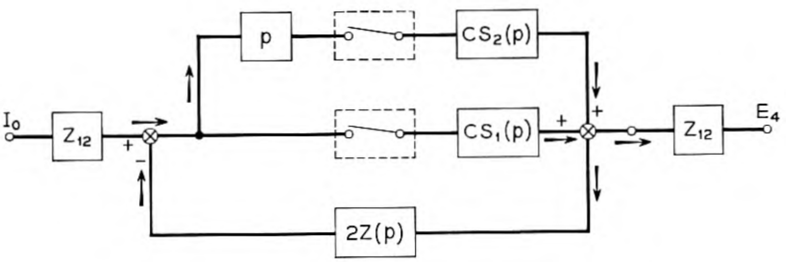
$$\delta i_n(p) = \frac{\omega_0^2}{p^2 + \omega_0^2} \frac{i_n(0-) + i_n'(0-)}{2p},$$

or

$$\delta i(t) = C \frac{\dot{e}_2(0-) - \dot{e}_3(0-)}{2} (1 - \cos \omega_0 t) \qquad 0 \leqq t < \tau.$$

Defining $S_2(p) = \mathfrak{L}^{-1}\{\tfrac{1}{2}(1 - \cos \omega_0 t)[u(t) - u(t - \tau)]\}$, or

$$S_2(p) = \frac{\omega_0^2}{2} \frac{1}{p(p^2 + \omega_0^2)} - \frac{(2p^2 + \omega_0^2)}{2p(p^2 + \omega_0^2)} e^{-p\tau},$$



$$E_4(p) = \frac{C[Z_{12}(p)I_0(p)]^* S_1(p) + C[pZ_{12}(p)I_0(p)]^* S_2(p)}{1 + 2C\left\{[Z(p)S_1(p)]^* + [pZ(p)S_2(p)]^*\right\}}$$

Fig. 8 — Modified block diagram.

and recalling that the input of the impulse modulator of Fig. 4 is $e_2(0) - e_3(0)$, it becomes obvious that the modified block diagram should be that given by Fig. 8. The output of the modified block diagram is given by[2]

$$E_4(p) = \frac{C[Z_{12}(p)I_0(p)]^*S_1(p) + C[pZ_{12}(p)I_0(p)]^*S_2(p)}{1 + 2C\{[Z(p)S_1(p)]^* + [pZ(p)S_2(p)]^*\}}.$$

X. CONCLUSION

Let us compare the method of solution presented above with the more formal approach proposed by Bennett. The latter method leads to the exact steady-state transmission through a network containing periodically operated switches. This method is perfectly general in that it does not require any assumption relative to the properties of the network nor to the ratio of τ/T. As expected this generality implies a lot of detailed computations. In particular it requires, for each reactance of the network, the computation of the voltage across it due to any initial condition. The method presented in this paper is not so general because it assumes first that the ratio τ/T is small; second the value of the inductance ℓ is very much smaller than that of L_n (see Fig. 3). The result of these assumptions is that the system of time varying equations may be solved by successive approximations with the further advantage that the convergence proof guarantees that, for very small τ/T, the zeroth approximation will be a close estimate of the exact solution.

The zeroth approximation may conveniently be obtained by considering a block-diagram analogous to those used in the analysis of sampled servomechanisms. Further the proposed method leads directly to some interesting results, for example, as far as the zeroth approximation is concerned, the dc transmission may be achieved with as small a loss as desired provided the lossless networks N_1 and N_2 are suitably designed. Another advantage of the proposed method is that the simplicity of the analysis permits the designer to investigate at a small cost a large number of possible designs.

Finally it should be pointed out that this approach to the solution of a system of time-varying linear differential equations may find applications in many other physical problems.

APPENDIX I

ANALYSIS OF THE RESONANT CIRCUIT

Consider the resonant circuit of Fig. 2. Suppose that at $t = 0$, the left-hand capacitor has a potential $e_2(0)$ and the right-hand capacitor has a potential $e_3(0)$ and that at $t = 0$ the current i_r through the inductance ℓ is zero.

The network equation is

$$\ell \frac{d}{dt} i_r + \frac{2}{C} \int i_r \, dt = 0.$$

Now $i_r(0) = 0$ and $di_r(0)/dt = [e_2(0) - e_3(0)]/\ell$. Let $2/\ell C = \omega_0^2$, then $di_r(0)/dt = \omega_0^2 C[e_2(0) - e_3(0)]/2$.
Using Laplace transforms,

$$\begin{aligned} (p^2 + \omega_0^2) I_r(p) &= p i_r(0) + \frac{di_r(0)}{dt}, \\ I_r(p) &= \frac{\omega_0^2 C}{2} [e_2(0) - e_3(0)] \frac{1}{p^2 + \omega_0^2}, \end{aligned} \tag{40}$$

hence

$$i_r(t) = \omega_0 C \frac{e_2(0) - e_3(0)}{2} \sin \omega_0 t \tag{41}$$

and

$$q(t) = \int_0^t i_r(t) \, dt = C \frac{e_2(0) - e_3(0)}{2} [1 - \cos \omega_0 t].$$

If $2\pi/\omega_0 = 2\tau$, i.e., $\tau = \pi\sqrt{\ell C/2}$, which means that the duration of the switch closure is a half-period of the resonance of the tuned circuit, then

$$i_r(t) = \frac{\pi}{\tau} \frac{C[e_2(0) - e_3(0)]}{2} \sin \frac{\pi t}{\tau},$$

$$q(t) = \frac{C[e_2(0) - e_3(0)]}{2} \left[1 - \cos \frac{\pi t}{\tau}\right].$$

It is clear then that, during the period τ, the charge transferred onto the

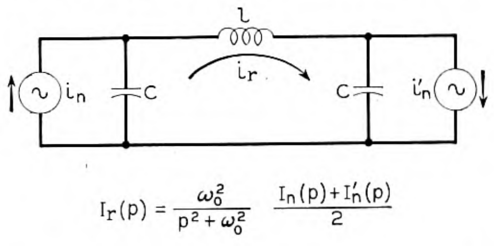


Fig. 9 — Resonant circuit excited by current sources I_n and I_n'.

right-hand capacitor is $q(\tau) = C[e_2(0) - e_3(0)]$ and as a result at time $t = \tau$ the right-hand capacitor has a voltage $e_2(0)$ and the left-hand capacitor has a voltage $e_3(0)$. Considering now the network of Fig. 9, the equation is

$$\ell \frac{d^2 i_r}{dt^2} + \frac{2}{C} i_r = \frac{1}{C} [i_n(t) + i_n'(t)].$$

Assuming all initial conditions* to be zero we get,

$$I_r(p) = \frac{\omega_0^2}{p^2 + \omega_0^2} \frac{I_n(p) + I_n'(p)}{2}. \tag{42}$$

Appendix II

STUDY OF THE LIMITING CASE $T \to 0$

We expect that if the sampling period $T \to 0$, which is equivalent to stating that the sampling frequency $\omega_s \to \infty$, then the inductance $\ell \to 0$ and as a result the voltage $e_3(t)$ will be infinitely close, at all times, to the voltage $e_2(t)$. Thus, in the limit, everything happens as if the terminal pairs (2) of N_1 and N_2 were directly connected. In that case the gain of the system is

$$\frac{Z_{12}(p)}{2Z(p)} Z_{12}(p),$$

as is easily seen by referring to the Thevenin equivalent circuit of N_1.

Let us show that as $T \to 0$, (21) leads to the same result. First note that both $Z_{12}I_0$ and ZS_1 go to zero at least as fast as $1/p^2$ for $p \to \infty$. Hence the summations in (21) reduce to the term corresponding to $n = 0$. Therefore,

$$E_4(p) = \frac{C[Z_{12}I_0]^* S_1}{1 + 2C[ZS_1]^*} Z_{12} \to \frac{CZ_{12}I_0Z_{12}S_1}{T + 2CZS_1} \to \frac{Z_{12}^2}{2Z} I_0 .$$

Appendix III

ZEROTH APPROXIMATION IN THE CASE WHERE N_1 IS NOT IDENTICAL TO N_2

Let, for $k = 1, 2$; C_k be the shunt capacitor at the terminal pair 2 of N_k, $Z_k(p)$ be the driving point impedance of N_k, and $Z_{12}^{(k)}(p)$ be the transfer impedance of N_k. In the present case the capacitors C_1 and C_2 are in series in the resonant circuit of Fig. 2. It can be shown that the

* Their contribution has been found in (40).

charge exchanged during one-half period of the resonance is

$$\frac{2C_1C_2}{C_1 + C_2} [e_2(0) - e_3(0)].$$

For the present case, (16) and (17) become

$$e_2(t) = v(t) - \int_{-\infty}^{t} i_{r0}(\tau) z_1(t - \tau) \, d\tau,$$

$$e_3(t) = \int_{-\infty}^{t} i_{r0}(\tau) z_2(t - \tau) \, d\tau.$$

Following the same procedure as before we are finally led to the block diagram of Fig. 10 whose output is given by

$$E_4(p) = \frac{[Z_{12}^{(1)}(p) I_0(p)]^* \dfrac{2C_1C_2}{C_1 + C_2} S_1(p) Z_{12}^{(2)}(p)}{1 + \dfrac{2C_1C_2}{C_1 + C_2} \{[Z_1(p) + Z_2(p)] S_1(p)\}^*}.$$

Appendix IV

THE DERIVATION OF EQUATION (24)

Considering the method used in Section IV to derive the zeroth approximation, it is clear that during the switch closure the voltages $e_2(t)$ and $e_3(t)$ vary sinusoidally, that is,

$$e_2(t) = e_2(0) - \frac{e_2(0) - e_3(0)}{2} \left[1 - \cos \frac{\pi t}{\tau}\right],$$

$$e_3(t) = e_3(0) + \frac{e_2(0) - e_3(0)}{2} \left[1 - \cos \frac{\pi t}{\tau}\right].$$

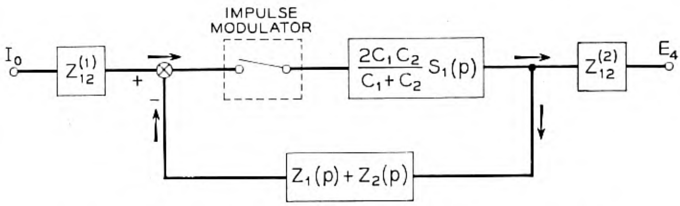


Fig. 10 — Zeroth approximation: modified block diagram for the case where N_1 and N_2 are not identical.

Thus, it always happens that for $t = \tau/2$, i.e., at the middle of switch closure time, $e_2(t) - e_3(t) = 0$.

Therefore if we consider the time function $e_2(t) - e_3(t)$ we have for all k's $(-\infty, \cdots 0, \cdots + \infty)$,

$$e_2\left(kT + \frac{\tau}{2}\right) - e_3\left(kT + \frac{\tau}{2}\right) = 0.$$

If, for simplicity of analysis, we assume that the switch is closed during the intervals $-(\tau/2) + kT \leqq t \leqq +\tau/2 + kT$, then for all k's,

$$e_2(kT) - e_3(kT) = 0.$$

Using (17), this condition implies that $[V(p)]^* - 2[I_{r0}(p)Z(p)]^* = 0$.

Now, remembering that $i_{r0}(t)$ consists of a sequence of half sine waves whose shape is defined by $s_0(t)$ (which is by definition identical to $s_1(t)$ except for an advance in time of $\tau/2$) it follows that $I_{r0}(p) = B(p)S_0(p)$, where $B(p)$ is the $\mathfrak{L}$-transform of the sequence of impulses whose measure is equal to the charge interchanged between N_1 and N_2 at each switch closure. Since $[B(p)S_0(p)Z(p)]^* = B(p)[S_0(p)Z(p)]^*$, then

$$B(p) = \frac{[Z_{12}(p)I_0(p)]^*}{2[S_0(p)Z(p)]^*}.$$

From which it immediately follows that

$$I_{r0}(p) = \frac{[Z_{12}(p)I_0(p)]^*S_0(p)}{2[S_0(p)Z(p)]^*}$$

and

$$E_{40}(p) = \frac{[Z_{12}(p)I_0(p)]^*S_0(p)Z_{12}(p)}{2[S_0(p)Z(p)]^*},$$

where

$$[S_0(p)\ Z(p)]^* = \frac{1}{T}\sum_{n=-\infty}^{+\infty} S_0(p + jn\omega_s)\ Z(p + jn\omega_s)\,.$$

REFERENCES

1. L. A. MacColl, *Fundamental Theory of Servomechanisms*, D. Van Nostrand Co. Inc., New York, 1945.

2. John G. Truxal, *Automatic Feedback Control System Synthesis*, McGraw-Hill Book Co. Inc. New York, 1955.
3. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Interscience Publishers, Inc. New York, 1953.
4. W. R. Bennett, Steady-state Transmission Through Networks Containing Periodically Operated Switches, I.R.E. Trans., **PGCT-2,** No. 1, pp. 17–22, March, 1955.
5. T. R. Bashkow, The A-matrix, New Network Description, to be published in I.R.E. Trans., **PGCT-4,** No. 3, 1957.

Experimental Transversal Equalizer for TD-2 Radio Relay System

By B. C. BELLOWS and R. S. GRAHAM

(Manuscript received February 26, 1957)

To determine the effect of improved equalization on the performance of TD-2 radio relay systems, an experimental adjustable transversal equalizer has been developed. The equalizer is based on the echo principle as used in transversal filters, and operates in the 60- to 80-mc frequency band. Seven pairs of adjustable leading and lagging echo terms provide flexibility for simultaneous gain and delay equalization. Directional couplers are used for tapping and controlling the echo voltages. Field experiments have shown that system equalization can be improved appreciably by the use of such equalizers.

INTRODUCTION

The TD-2 radio relay system[1] employs frequency modulation to transmit multichannel telephony or television. The frequency modulated signal requires a transmission system whose gain and envelope delay are constant over the frequency band, 20 mc wide, used to transmit the signal. Deviations from constant gain or delay result in non-linear distortion of the demodulated signal. For television this results in distorted images, and for multichannel telephone transmission this introduces cross modulation among the voice channels.

Basic equalization is provided in each repeater. In addition, certain fixed equalizers have been used on a mop-up basis. However, there remains some residual distortion of random shape. This paper discusses the design of an adjustable equalizer to compensate for this distortion.

I. BASIC EQUALIZATION

The transmission path through a TD-2 repeater consists of an intermediate frequency portion covering the 60- to 80-mc band and radio frequency channels 20-mc wide in the 3,700- to 4,200-mc band. At each repeater the RF channels are separated by wave guide filters and each is demodulated to the IF frequency for amplification and equalization. The outgoing signal is then modulated back to an RF channel, at a different frequency from the incoming signal to reduce interference. The RF and IF portions of the repeater were designed so that the combination would have flat gain to within the closest practicable limits over the 20-mc band, and a number of adjustments are provided in the IF amplifier to maintain this flatness under field conditions. The unequalized repeater, however, has an envelope delay distortion characteristic shown in Fig. 1 which is approximately parabolic. To minimize this distortion, each repeater contains a 315A equalizer, which has approximately the inverse of the delay distortion of a typical repeater.

II. SUPPLEMENTARY EQUALIZATION

In a TD-2 system consisting of many repeaters in tandem, both gain and delay distortions may accumulate to the point where additional equalization is necessary. If the pass band of the repeaters shifts slightly in frequency, due to changes in temperature or adjustment, the difference between the repeater delay and the delay of its equalizer will result in delay distortion which has approximately a linear slope with frequency. This may be corrected at main repeater stations by combina-

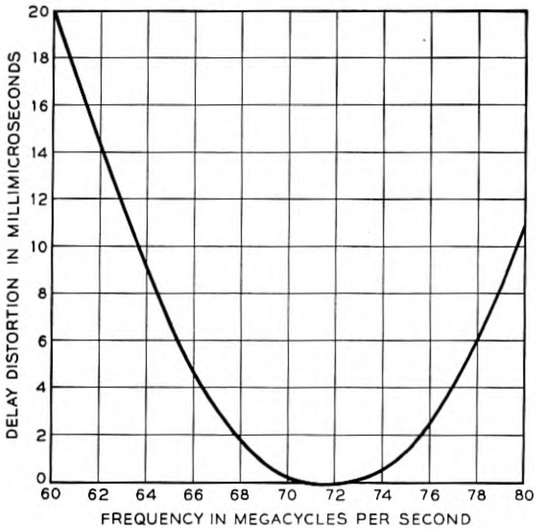


Fig. 1 — Over-all delay distortion of a typical microwave repeater, TD-2 system.

tions of delay slope equalizers. The characteristics of the 319A, B and C equalizers provided for this purpose are shown in Fig. 2. Each equalizer consists of two bridged-T all-pass sections. There is also some variation in bandwidth of the TD-2 repeaters, resulting in part from the fact that the waveguide filters used in the higher frequency radio channels are somewhat broader than those in the lower frequency channels. This variation in bandwidth results in delay distortion which has a parabolic shape with frequency. Use of a larger or smaller number of the basic 315A equalizers corrects this.

Over long circuits, small distortions in the gain shape of the TD-2

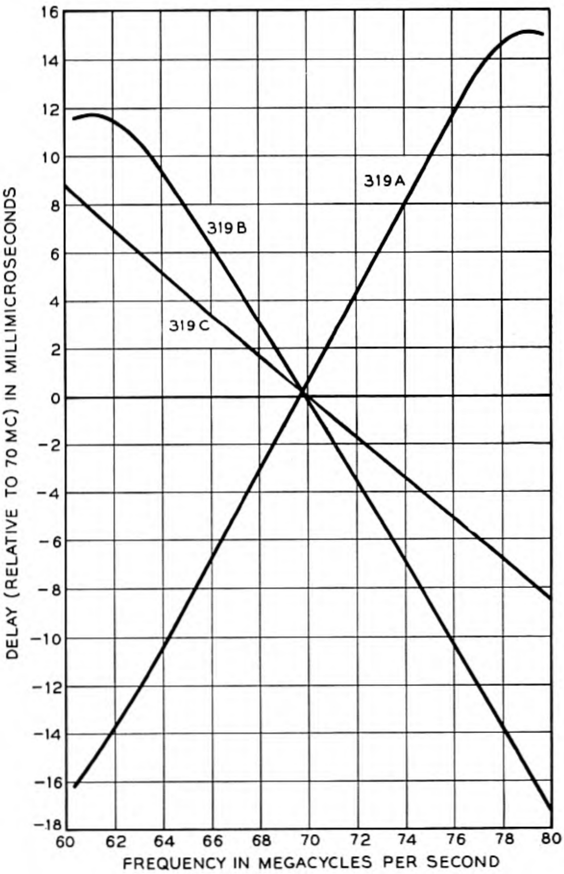


Fig. 2 — Delay characteristics of 319 type equalizers. Combinations of these are used at main stations to equalize delay slope of the system.

equipment produce a cumulative gain-frequency distortion which is noticeable in television circuits. Present practice is to correct for this by standard video equalizers after the FM signal has been demodulated to baseband. In connection with the experimental equalizing program to be described, parabolic gain equalizers operating on the FM signal before demodulation were used.

III. RESIDUAL DISTORTION

After correction of the known shapes discussed above, there remains a certain residual gain and delay distortion which results from a random summation of many minor sources. The shape of this distortion is not predictable, but its statistics are known. Examination of typical delay versus frequency characteristics have shown that these may be reasonably well approximated by six cosine terms: a 40-mc fundamental and the next five harmonics. Similar gain terms are needed. However, the gain and delay distortion, when examined within the 20-mc band of interest, do not have a minimum phase relationship. This is to be expected because of the presence in the system of the delay equalizers, which are non-minimum-phase networks, and of amplifiers with compression.

The magnitude of the residual distortion is small enough so that transcontinental TD-2 circuits provide television and telephone transmission of commercial quality. Some effects, such as cross modulation, are sufficiently marginal so that improvement would be desirable. To determine whether this could be achieved by improved gain and delay equalization, the development of an experimental adjustable equalizer was undertaken. The considerations outlined show that such an equalizer should approximate the desired characteristics with independent gain and delay terms of the harmonically related cosine type. Equalization to reduce cross modulation in telephone channels and differential phase in color television must be performed before demodulation of the FM signal to base band. The equalizer was, therefore, built to operate in the 60- to 80-mc IF band.

IV. TRANSVERSAL EQUALIZER

One method of obtaining independent control of the loss and delay characteristics of a network has been achieved in the transversal filter.[2] Equalizers have been designed on this principle for the equalization of television circuits.[3, 4] This type of equalizer, referred to here as a transversal equalizer, provides a flexible means of synthesizing any loss characteristic and any delay characteristic limited only by the number of harmonics that are provided and the range of each.

Basically, the transversal equalizer consists of a delay line with equally spaced taps, with a means for independently controlling the amount of signal fed through each of the taps to a summing circuit, as shown schematically on Fig. 3. The input signal is fed into one end of the delay line which is terminated at the other end. The center tap is fed to the output and forms the main transmission path.

The operation of the equalizer can best be described using the "time domain" analysis based on the theory of paired echos.[5] Portions of the signal tapped off the "leading" or first half of the delay line will not be delayed as much as the main signal and will introduce leading "echos". Similarly, lagging echos can be obtained from the taps on the lagging or second half of the delay line. Combinations of both types of echos, either positive or negative as required, can be added to cancel out, to a first approximation, distortion present in the input signal.

This analysis can also be carried out in the frequency domain. To obtain a family of cosine loss versus frequency characteristics without any appreciable delay characteristic, equal leading and lagging echos of the same polarity are added to the main signal in the summing circuit. To obtain a corresponding family of cosine delay versus frequency characteristics without loss distortion, leading and lagging echos equal

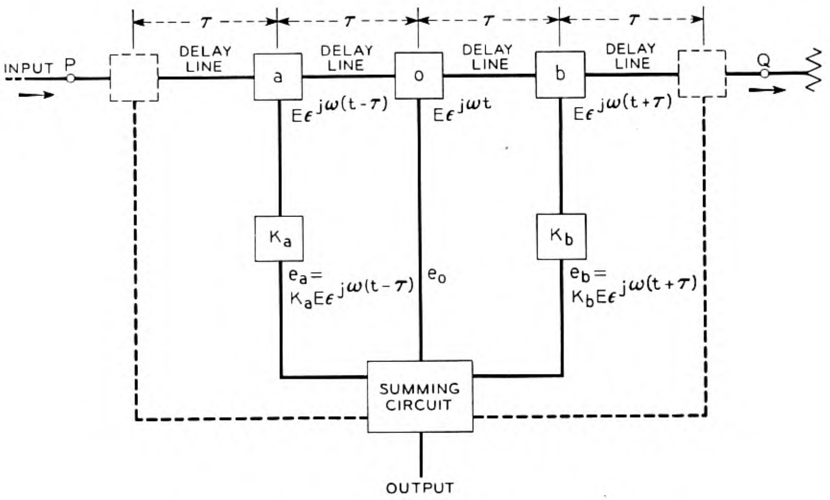


Fig. 3 — Block schematic of transversal equalizer.

in magnitude but of opposite polarity are added in the summing circuit to the main signal.

To achieve a practical equalizer for operation over the 60- to 80-mc band requires the following components: delay line or delay networks, means for tapping off small portions of the signal controlled both in amount and polarity, and a suitable summing circuit.

V. DERIVATION OF THE EQUALIZER CIRCUIT

A brief analysis of the operation of the equalizer will be given at this point as a basis for discussion of the method of tapping the signal and controlling the amplitude and polarity of the tapped portion.

Fig. 3 shows the basic delay line PQ as well as the means used for producing the main signal and a single pair of leading and lagging echos. The tap labeled "o" in the center of the line produces the main signal. The tap "a", being closer to the input, produces a signal which leads the main signal by time τ. The tap "b" produces a signal which lags the main signal by the same amount. The boxes "K_a" and "K_b" control the amplitude and polarity of the leading and lagging signals which are to be combined with the main signal to produce one term of the desired equalization characteristic. It will be shown that these three signals will provide one cosine gain term and one cosine delay term, both having the same period, but being independently controllable as to amplitude and polarity.

We will choose as our reference point for phase the main output signal, $e_0 = E\varepsilon^{j\omega t}$. The output from tap "$a$" is then

$$E\varepsilon^{j\omega(t-\tau)}.$$

After passing through box "K_a", this becomes

$$e_a = K_a E\varepsilon^{j\omega(t-\tau)}.$$

Similarly,

$$e_b = K_b E\varepsilon^{j\omega(t+\tau)}.$$

Here the terms K_a and K_b are of the form

$$K = \pm\varepsilon^{-\alpha}$$

where α is the attentuation in nepers of the box K. Note that $|K|$ is less than unity, assuming the box represents a passive network. Combining these two signals with the main signal, we have

$$e_r = e_0 + e_a + e_b = E\varepsilon^{j\omega t}[1 + K_a\varepsilon^{-j\omega\tau} + K_b\varepsilon^{j\omega\tau}]. \tag{1}$$

Now it will be shown that, by the adjustment of the two parameters K_a and K_b, it is possible to realize independent control of a cosine gain term and a cosine delay term.

Since, in general, $K_a \neq K_b$, let us define

$$K_a = K_g + K_p$$

and

$$K_b = K_g - K_p. \tag{2}$$

Then

$$e_r = Ee^{j\omega t}[1 + K_g(e^{-j\omega\tau} + e^{j\omega\tau}) + K_p(e^{-j\omega\tau} - e^{j\omega\tau})]. \tag{3}$$

Substituting the trigonometric form:

$$e_r = Ee^{j\omega t}[1 + 2K_g \cos \omega\tau + j\, 2K_p \sin \omega\varepsilon]. \tag{4}$$

Note that for K_a equal to K_b, the sine phase term is zero and that for K_a equal to $-K_b$ the cosine gain term is zero. Similarly, by proper proportioning of K_a and K_b, K_g and K_p may be assigned any desired values.

If we normalize (3) by setting $e_0 = E\varepsilon^{j\omega t} = 1$, the expression in brackets can yield two vector diagrams which are useful in explaining the functioning of the equalizer. To obtain the diagram shown in Fig. 4(a), we have set $K_p = 0$. We then have a unit vector, representing the main signal, a leading echo $K_g\varepsilon^{-j\omega\tau}$, and a lagging echo $K_g\varepsilon^{j\omega\tau}$. The

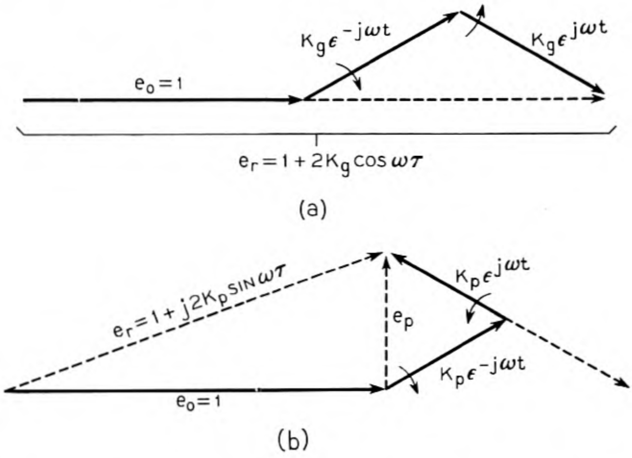


Fig. 4 — Vector diagrams of paired echos. (a) Equal echos of same polarity produce magnitude change without phase change. (b) Equal echos of opposite polarity produce a change in phase shift with a minor change in magnitude.

vector representing the leading echo rotates clockwise with respect to the main signal when the frequency increases, whereas the vector representing the lagging echo rotates counterclockwise by the same amount, and the resultant thus varies in magnitude but not in phase. The magnitude of the resultant is given, for this case, by the first two terms in parentheses in (4).

If, on the other hand, if we set $K_g = 0$, we have the three vectors shown in Fig. 4(b), identical with those in Fig. 4(a) except that the polarity of the lagging echo has been reversed. In this case, the two echos produce a resultant, e_p, which is in quadrature with the main signal. For small echos, e_r is thus shifted in phase from the main signal, with substantially no change in magnitude. The resultant in this case is given by the first and third terms in parentheses in (4). This gives a sinusoidal variation in the phase of the resultant. Since envelope delay is defined as $d\beta/d\omega$, where β is the phase shift through the circuit in question, the sinusoidal phase ripple will be seen to yield, after differentiation, a cosine delay ripple.

The period of the ripple can be seen from the above expressions to depend on τ, the delay between the leading and the main tap, and between the main tap and the lagging tap. Other pairs of echos, each pair symmetrically disposed about the main tap, but with different values for τ, will give transmission ripples of different periods. To provide a series of orthogonal terms, the values of τ must be integral multiples of a common value, normally that required to produce 180° phase shift across the band of interest.

A complete equalizer must, of course, sum up the various echos and the main signal, taking care that the delay between the tap and the summing point is the same for each echo and the main signal, that parasitic losses such as losses in cabling are the same for each path through the equalizer, and that any frequency characteristic in the tapping device or other parts of the equalizer is properly equalized out so that the over-all equalizer introduces a minimum of distortion of its own.

VI. DIRECTIONAL COUPLER

To reduce incidental distortion, it is desirable that the device used to tap the delay line for the main signal and the echos introduce substantially no discontinuity in the main line. The device chosen for this purpose is a directional coupler. It is shown symbolically in Fig. 5. The directional coupler is a four port device having properties similar to a hybrid coil. Power entering one port divides (not necessarily equally) between two other ports, but none of it reaches the fourth, or conjugate port. In Fig. 5, the power entering at 1 divides between 2 and 4, that entering at 2 divides between 1 and 3, that entering at 3 divides between 2 and 4, and that entering at 4 divides between 1 and 3. Directional couplers inherently provide an impedance match at all four ports. Thus, such a coupler sets up no reflections in the main line. Its insertion loss in this line may be kept small by having nearly all the power entering at 1 come out at 2; then only a small fraction is diverted to 4. Coaxial directional couplers have been discussed in the literature[6, 7, 8] and will not be dealt with in detail here.

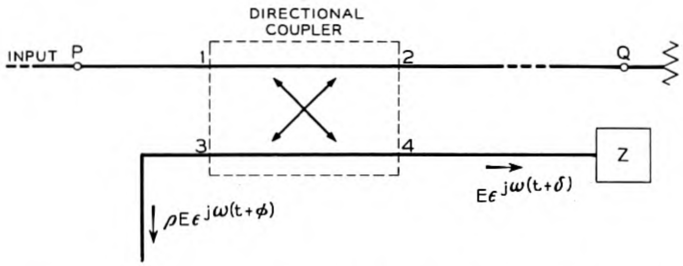


Fig. 5 — Diagram of directional coupler. Input signal divides between Ports 2 and 4 with no output at Port 3. Termination Z at Port 4 reflects some signal to Port 3, proportional to the reflection coefficient, ρ.

The coupler used here (J68333C) is one originally developed to measure reflections on IF transmission lines in the TD-2 system. The directivity of a coupler is defined as the coupling loss between main line and branch line in the undesired direction less the loss in the desired direction (loss from 1 to 3 less the loss from 1 to 4, for example). In the J68333C coupler, the directivity can be adjusted to exceed 45 db over the band of interest. This can be done by adjusting two screws, shown on model in Fig. 6, to obtain the optimum spacing between the coupling elements. The loss between the main line and the branch line in the desired direction is about 23 db at mid-band (70 mc), and decreases 6 db per octave with increasing frequency. The loss along one of the coupled lines (1 to 2 or 3 to 4) is very small.

Use has been made of the directional properties of the coupler in providing a simple means of controlling the amplitude and polarity of the tapped signal. Referring to Fig. 5, and keeping in mind the properties of the coupler, it will be noted that a small portion of the input signal appears at Port 4 of the coupler, but none at Port 3. If the impedance Z

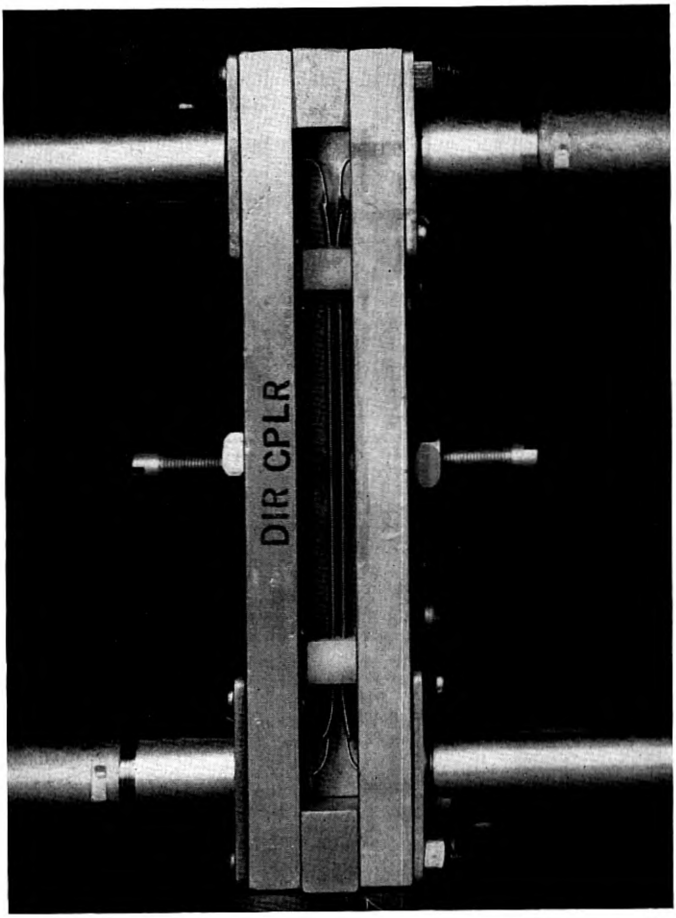


Fig. 6 — J-68333C directional coupler with cover plate removed to show coupling elements. Optimum spacing for maximum directivity can be obtained by adjusting screws.

matches the impedance seen looking into Port 4, all of the small signal will be absorbed in Z. If Z is an adjustable resistance, then a controllable portion of the small signal can be reflected back into Port 4, whence most of it will come out Port 3. A small portion of the reflected signal will emerge at Port 1, headed toward the input. This portion will be attenuated by twice the coupling loss between the main and the branch line plus the return loss of the reflection at Z, and can be made negligible. The interactions caused by the reflected signal on the main delay line entering previous couplers are also reduced by twice the coupler loss.

The coupler in Fig. 5 represents any one of the taps on the line PQ in Fig. 3. The signal emerging from Port 4 can be written as $E\varepsilon^{j\omega(t+\delta)}$, where δ is a time delay dependent on the location of the tap on the line PQ. The signal emerging from Port 3 is then $\rho E\varepsilon^{j\omega(t+\delta)}$, where ρ is the voltage reflection coefficient at Z, and is given by

$$\rho = \frac{R - R_0}{R + R_0}.$$

Here R is the value of resistance used to provide the impedance Z, and R_0 is the impedance seen looking into Port 4. An examination of the signal from Port 3 shows that it is the same as the signals e_a or e_b in Fig. 3, with the reflection coefficient ρ substituted for variable K_a or K_b in Fig. 3. Thus, it is seen that we may use the reflection at Z, variable by controlling the value of R, to perform the function of the box K in Fig. 3. Neglecting parasitic losses we may then write:

$$K = \rho = \frac{R - R_0}{R + R_0}$$

and

$$R = R_0 \frac{1 + K}{1 - K}. \tag{5}$$

This gives us the value of R to use for any desired value of K for any of the taps which derive echos, assuming the summing circuit has equal attenuation in all paths. In the case of the main central tap, the signal from Port 4 of the coupler is seen to be the same as the main signal e_0 in Fig. 3, and is used as such directly.

VII. METHOD OF ADJUSTMENT

The detailed design of a manually adjustable equalizer is materially influenced by the method to be used in the field for determining the setting of its controls. The present equalizer with 14 independent controls would present a complex problem of field adjustment unless special procedures were developed to simplify the adjustment. To adjust the equalizer, the radio circuit being equalized must be taken out of commercial service, so any reasonable measures to simplify the adjustment or reduce the time required are justified.

Two methods appeared to be feasible at the time the development was started. One would be to use existing gain and delay sweep test circuits. These present a visual display of the circuit gain or delay distortion versus frequency. These displays are not available simultaneously with present equipment. To adjust the equalizer controls using this equipment, it must be possible to adjust either gain or delay without affecting the other. Thus, all the gain terms can be adjusted in succession using the gain display. Then the procedure is repeated with the delay display, adjusting the delay terms. Since a combination of leading and lagging echos in equal amounts is required for this procedure, an arrangement of the controls to facilitate this is required. One way to achieve this is to use stepped rheostats with the steps proportioned to introduce equal amplitude changes in the echo voltage. With this arrangement, gain changes can be introduced by rotating the two switches corresponding to a pair of echos in the same direction an equal number of steps. Delay changes can be obtained by similar rotation in opposite directions. A further refinement consisting of mechanically ganging the controls is possible but this was not done on these experimental models.

The second method of adjustment would be to develop a special test set similar to the one developed for the L3 system.[9] This could produce a meter reading proportional to the amount of gain and delay distortion present in the circuit. Successive controls could then be adjusted for minimum meter readings. Experience with the L3 system cosine equalizers has shown the desirability of continuously adjustable controls for such a method.

To test both methods under field-trial conditions, two versions of the equalizer were built — one with stepped rheostats and one with continuously adjustable rheostats.

VIII. COAXIAL RHEOSTAT

Since there were no available continuously adjustable rheostats satisfactory for operation at 70 mc, a special rheostat was developed for

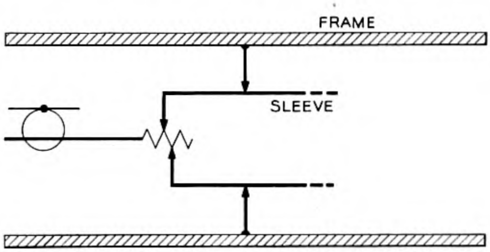


Fig. 7 — Schematic of coaxial rheostat. Moveable sleeve changes position of inner contacts touching ceramic rod, changing resistance. Fixed outer contacts maintain constant path length to frame.

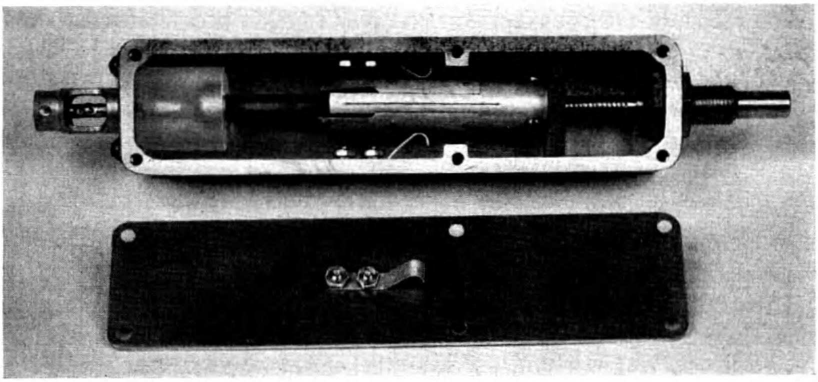


Fig. 8 — Model of coaxial rheostat with cover removed.

this purpose. It employs a ceramic rod, $\frac{1}{4}$ inch in diameter, coated with a pyrolytic carbon film, as a center member of a coaxial structure. A metal sleeve which is moved longitudinally by a lead screw carries sliding contacts along the rod. These parts are supported inside a rectangular housing which forms the outer conductor. A second set of fixed contacts attached to the rectangular housing makes contact with the sleeve. This arrangement maintains a substantially constant length of path from the input end of the rod to the housing, which forms the ground, independent of the position of the sleeve. The schematic of the rheostat is shown in Fig. 7. A model of the rheostat is shown in Fig. 8.

To obtain uniform adjustment of amplitude in decibels, a resistance that varies exponentially with length or with rotation of the lead screw is required. Such a resistance characteristic is realized by varying the thickness of the carbon film along the rod. This produced a total resistance which varied from 20 ohms at the low setting to about 350 ohms at the high resistance setting. After an initial wearing-in period of 1,000 cycles of moving the contacts over their full travel, the resistance was changed less than 1 per cent by another 9,000 cycles. This amount of wear is estimated to be greater than that encountered in twenty years of normal operation.

The housing and rod were dimensioned to form a 75-ohm transmission line. Measurements of the impedance at the input connector, made at frequencies from 60 to 80 mc, showed that this impedance can be approximated by a resistor terminating 6.7 cm of 75-ohm coaxial cable. For the 75-ohm setting, the reflection coefficient of the rheostat is less than 2 per cent across this frequency band.

One model of the equalizer was completely equipped with these rheostats. By allowing for the equivalent length of cable within the rheostat, an essentially pure resistive termination was obtained.

IX. OTHER COMPONENTS

Other components required for the equalizer included stepped rheostats, delay line or delay networks, a suitable summing network, and a loss equalizer.

The stepped-switch rheostats were made from standard switch parts with eleven positions. Deposited carbon resistors, 205D, were used for the steps. The mid-position corresponded to the circuit impedance level, 75 ohms. The other steps were arranged to provide equal increments of echo amplitude measured in decibels. With careful control of lead lengths, special shielding and a coaxial cable connector, satisfactory control of the return loss of this rheostat was obtained.

Resistance pads were added to the switch assemblies associated with each of the echo terms. The loss of each pad was determined so that the corresponding term would have the desired maximum amplitude. In addition, the losses of the pads associated with the leading echo terms were increased to compensate for the midband loss of the delay line between the leading coupler and the corresponding lagging coupler. This insured that the two echos would have equal amplitudes.

The delay required between taps in the delay line is 0.025 microsecond, corresponding to a change in phase shift of 180° from 60 to 80 mc. In order for the equalizer cosine characteristics to have maxima at the band edges, the total phase shift at 60 and 80 mc must be successive integral multiples of 180°. Since the phase shift of coaxial patch cable is closely linear and proportional to length, it could be used for the delay line. Lumped-element delay networks consisting of two or more all-pass sections are a feasible alternative and would reduce the over-all size and weight. In view of the additional development effort involved to produce these and the experimental nature of this equalizer, it was decided to use coaxial patch cord. The type selected, 728A cable, has a polyethylene dielectric and is tested during production for return loss in the 50- to 95-mc band. The length required for each section is about 15.6 feet. This much cable has a loss of about 0.3 db at 70 mc.

It was originally proposed to use a series of directional couplers for summing the echo voltages with the main signal. This would provide additional isolation between terms. However, tests on a preliminary model indicated this isolation was not required in this application. In-

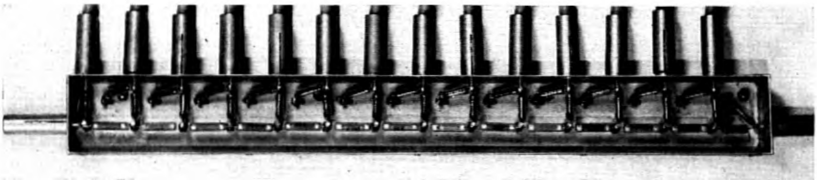


Fig. 9. — Summing network with cover removed. Main signal input is at right, echo signal inputs on top, and output is at left..

stead, a resistance summing network was developed using deposited carbon resistors. An L-pad is used in each echo path and a series resistor is added to the main path to preserve the 75-ohm impedance level, introducing a main path loss of about 0.4 db per tap. A model with cover removed is shown on Fig. 9. The main signal is introduced at one end of the structure, the echo voltages are connected along the side and the sum is taken off the other end. The return loss measured at any of the connectors with the others terminated was of the order of 40 db over the 60- to 80-mc band.

An attentuation equalizer is required to make the transmission through the main path constant. This path consists of about 108 feet of 728 cable, the straight-through loss of six couplers, and the coupling loss of the main coupler. The net distortion over the band is a slope of about 1.5 db and is corrected for by a constant resistance equalizer. Return losses exceeding 34 db were obtained over the frequency band.

X. ASSEMBLY

All the components were mounted on the rear of a standard relay rack panel. The rheostat controls are arranged on the front of the panel as shown on Fig. 10. This is a front view of the completed equalizer. The controls for the leading echo terms are on the left and for the lagging echo terms on the right. They are arranged vertically in numerical order with the first terms (shortest time separation from the main signal) at the top.

The rear of the panel is shown on Fig. 11. The directional couplers are mounted horizontally in two vertical columns. The cables forming the delay line sections are terminated in a plug and a jack and these are inserted in successive couplers, from the second port of one to the first port of the next. The third port of each coupler is connected through a short cable to its corresponding rheostat assembly. The fourth port of each coupler is connected to the summing network. An exception is

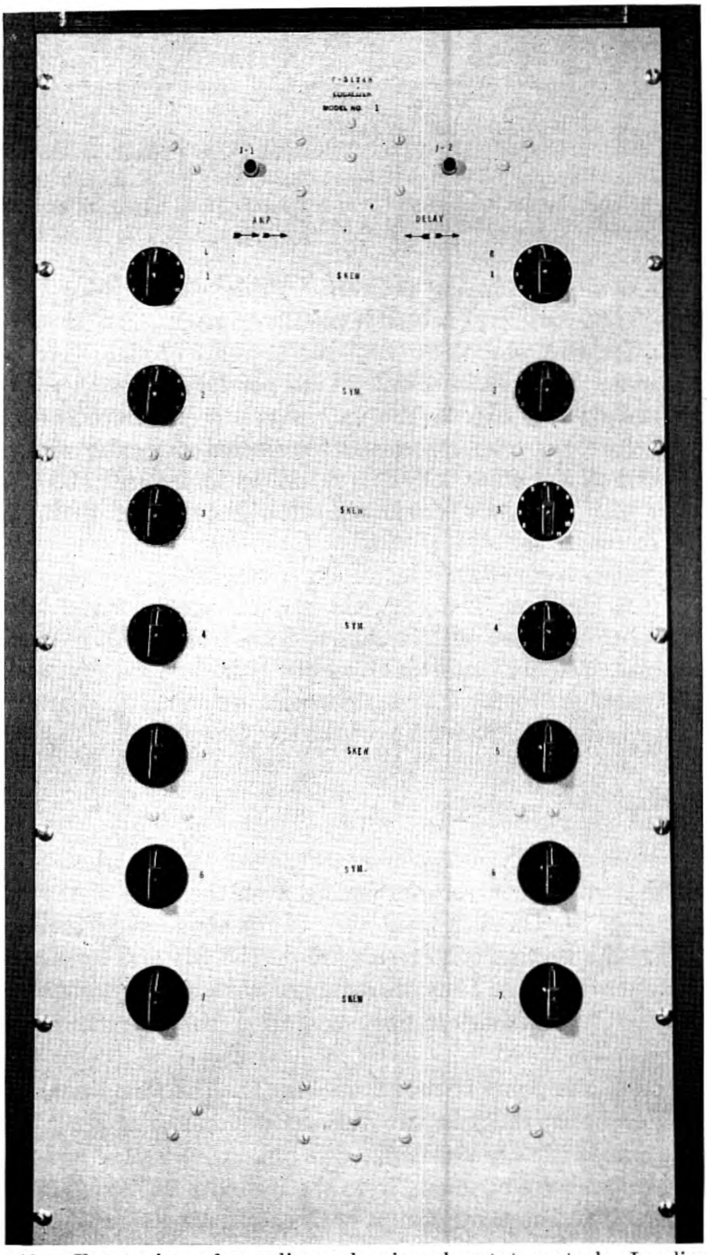


Fig. 10 — Front view of equalizer, showing rheostat controls. Leading echo controls are on left, lagging ones on right. First harmonic controls are at top.

the middle coupler, the fourth port of which is terminated in 75 ohms and the third port connected to the summing network.

The envelope delay in the cables connecting each coupler to the rheostat and to the summing network appears as delay for the particular echo path. Since these delays are not negligible compared to the 25 millimicrosecond delay between echos, the cable lengths were controlled so that the same amount of additional delay was introduced into each path including the main path.

XI. ADJUSTMENT AND PERFORMANCE

After the equalizer was assembled, the length of each of the cables connecting the couplers to the summing network was adjusted so that

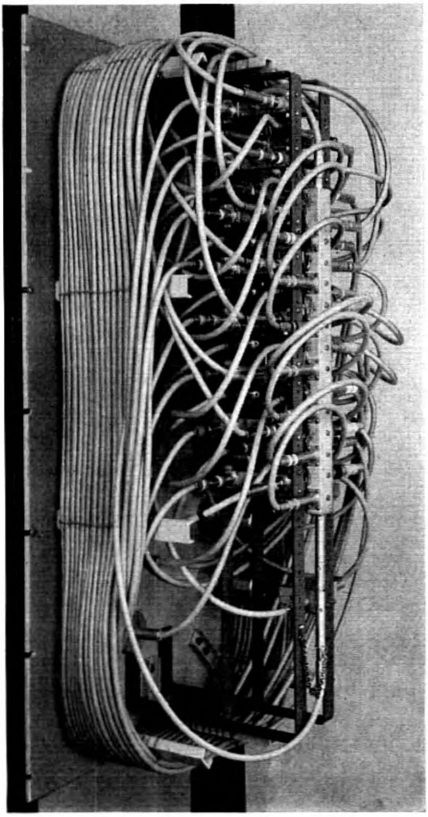


Fig. 11 — Equalizer with cover removed. Cables wound outside frame are the delay line sections. Summing network and directional couplers are in center. Rheostat cases are mounted on panel with coaxial connection in rear.

the zeros of the cosine shape occurred at the proper frequencies, as observed when the associated rheostat was set at maximum and minimum positions, with all other rheostats set at midrange, or no-echo, setting.

Some reflections were present in the main signal path as evidenced by ripples in the gain characteristic when all rheostats were set at midrange, corresponding to the "flat" loss condition. These reflections were reduced to some extent by minor readjustments of the balancing screws on the directional couplers. The over-all flat gain characteristic obtained after these adjustments is shown on Fig. 12.

This figure also shows the seven gain characteristics obtained when each pair of rheostats is set for maximum gain. The markers on the reference trace correspond to the band edges. A sharp gain bump resulting

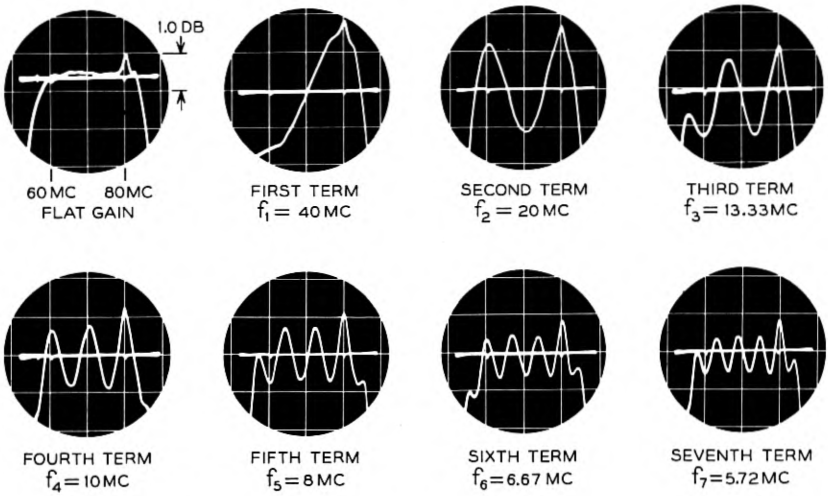


Fig. 12 — Measured gain characteristics of equalizer.

Fig. 13 — Gain change introduced by changing delay terms from zero to maximum. Left, normal case. Middle, third harmonic term at maximum. Right, fifth harmonic term at maximum.

from reflections on the delay line occurs just above 80 mc and distorts each characteristic near this frequency. The delay characteristics obtained closely resemble the corresponding gain characteristics. The delay characteristic with all rheostats on midrange was flat to within about ±1.5 millimicroseconds.

Another measure of the performance is the amount of interaction between gain and delay characteristics. As shown on Fig. 13 the gain changes less than 0.2 db when the third or fifth harmonic delay terms are set at their maximum values of 11 and 12 millimicroseconds, respectively. Similarly, when a pair of rheostats are set for the maximum gain characteristic, the effect on delay is of the order of two millimicroseconds. These results, which are typical, indicate the interaction effect is of the order of 20 per cent using one neper of gain distortion ripple as equivalent to a ripple of one radian of phase shift amplitude. The effect of this interaction on the field use of the equalizer is to require a second round of equalization to correct for interactions after the gross distortions in a circuit have been equalized.

XII. FIELD EXPERIMENTS

Models of this equalizer were installed in two channels of the TD-2 system between Denver, Colorado, and Omaha, Nebraska, early in 1956. This route is about 500 miles long and includes 18 microwave links. One equalizer was installed at the center of the route and a second one at the receiving end. The results of a typical set of characteristics obtained are shown in Figs. 14 and 15. The first shows the delay characteristic of the whole channel measured at the receiving intermediate

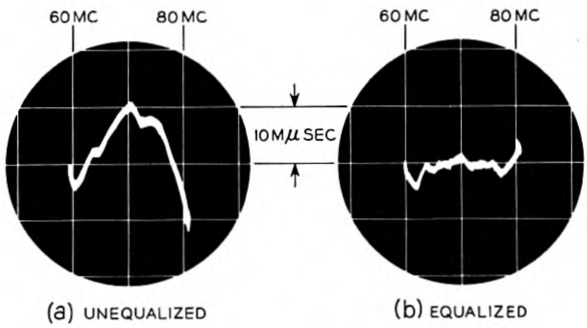


Fig. 14 — Measurements on Denver-Omaha route. Envelope delay distortion measured at intermediate frequency point. Left, unequalized; right with equalizer adjusted.

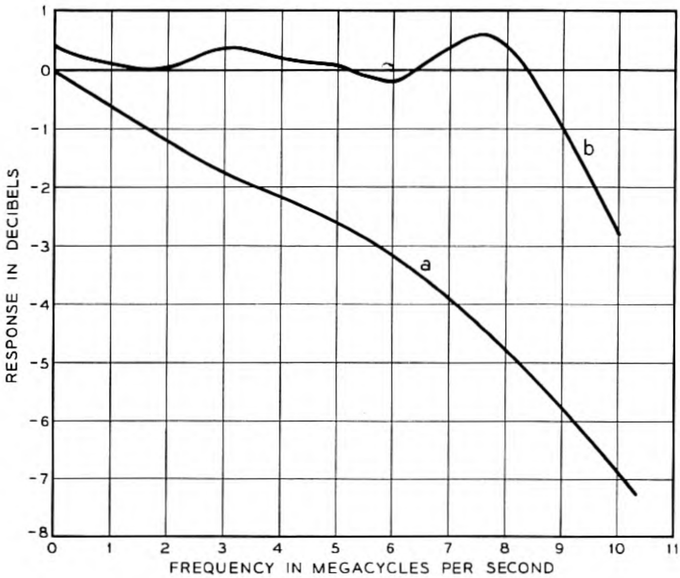


Fig. 15 — Measurement on Denver-Omaha route. Transmission characteristic measured at video (a) unequalized, (b) with equalizers adjusted.

frequency point. The "unequalized" characteristic is the circuit delay distortion immediately after standard line-up procedures without video equalization. The "equalized" characteristic shows the same circuit distortion corrected by the addition of the transversal equalizer. The first two delay terms of this equalizer were supplemented by the use of 319 type (linear slope) and 315A (parabolic) equalizers. The second loss term was supplemented by the use of experimental fixed parabolic loss equalizers. An attempt was made to obtain the best performance over the center 10 mc of the band, corresponding to the first order modulation band.

The gain distortion was adjusted on a demodulated video basis as this was the only type of sweep gain circuit available. As shown in Fig. 15, the unequalized circuit had more than 0.5 db per mc of slope up to 10 mc. The addition of the equalizer produced a flat band to about 8.5 mc.

The effect of this equalization on cross modulation, measured by simulating the message load with a flat band of noise, is shown in Fig. 16, for two sample channels. In these curves, normal drive represents the noise load whose power is 12 db below the power in a sine wave giving 4-mc peak deviation of the TD-2 carrier. Channel A showed 5-db im-

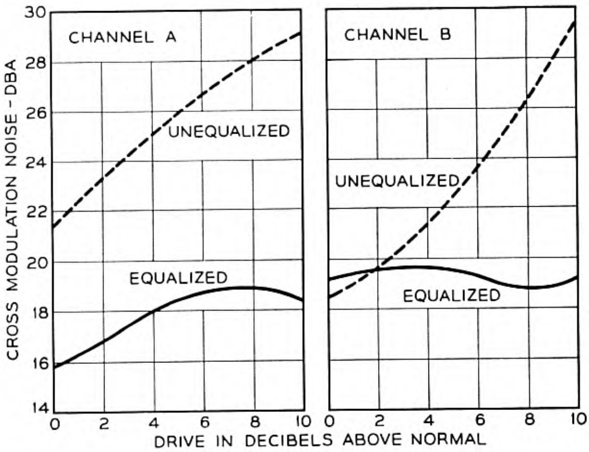


Fig. 16 — Effect of improved equalization on cross modulation noise, Denver-Omaha route. Measurements on two channels, referred to − 9 db transmission level.

provement at normal drive, and a somewhat greater improvement at higher drives. Channel B, however, showed a slight degradation at normal drive, which becomes a 9- or 10-db improvement at high drive. In the case of the latter channel, the improvement at normal drive was limited by the presence of a delay ripple, due to waveguide echoes, which was of too short a period to be equalized by the present equalizer.

The realization of such improvements in a working system is limited by such waveguide echoes, which are not stable enough for ready equalization, as well as by other instabilities in the transmission characteristic which have been attributed to antenna and air path effects.

ACKNOWLEDGMENTS

Several of the authors' colleagues contributed materially to this development. C. H. Dagnall developed the 315 and 319 type equalizers. A. H. Volz and J. K. Werner developed the coaxial rheostat, based on suggestions by S. Bobis. Mrs. Grace L. Ebbe was responsible for much of the equalizer design and the assembly and testing of models. J. G. Chaffee and I. Welber directed the field tests.

REFERENCES

1. A. A. Roetken, K. D. Smith, and R. W. Friis, The TD-2 System, B.S.T.J., **30,** pp. 1041–1077, Oct., 1951.

2. H. E. Kallman, Transversal Filters, Proc. I.R.E., **28,** pp. 302–310, 1940.
3. J. M. Linke, A Variable Time Equalizer for Video Frequency Waveform Correction, Proc. I.E.E., **99,** Part IIIa, pp. 427–435.
4. R. S. Graham and R. V. Sperry, Equalizer, U. S. Patent No. 2,760,164.
5. H. A. Wheeler, The Interpretation of Amplitude and Phase Distortion in Terms in Terms of Paired Echos, Proc. I.R.E., **27,** pp. 359–385, 1939.
6. G. D. Monteath, Coupled Transmission Lines as Symmetrical Directional Couplers, Proc. I.E.E., Part B, No. 3, **102,** pp. 383–392, May, 1955.
7. B. M. Oliver, Directional Electromagnetic Couplers, Proc. I.R.E., **42,** p. 1686, Nov., 1954.
8. B. C. Bellows, Directional Coupler, U. S. Patent No. 2,679,632.
9. R. W. Ketchledge and T. R. Finch, The L-3 System - Equalization and Regulation, B.S.T.J. **32,** pp. 833-878, July 1953.

Transmission Aspects of Data Transmission Service Using Private Line Voice Telephone Channels

By P. MERTZ and D. MITCHELL

(Manuscript received February 2, 1957)

An exploration is reported of the possibilities of a moderately fast data transmission system to use private line message facilities. A comparatively conventional system was desired to permit expeditious application. An auxiliary "word start" signal was necessary for the system considered.

Transmission characteristics of a number of arrangements were examined. These included several exploratory AM vestigial sideband systems (using a spectrum similar to telephotography), double sideband AM systems, various telegraph and other multichannel systems.

It was concluded that the 1600 bits per second usable over the AM vestigial sideband arrangement was about as fast as could be expected with the data system contemplated. This transmission is not as "rugged," with respect to impulse noise and sudden level changes, as other slower arrangements, but it is expected to be satisfactory. It will require delay correction, and simple methods are considered for carrying this out.

I. INTRODUCTION

The Bell System has been approached on a number of occasions in regard to the transmission of computing machine and similar data over its telephone circuits. This has reached the point where specific possibilities for private line data transmission have been given serious consideration.

The telephone network was developed for speech transmission, and its characteristics were designed to fit that objective. Hence, it is recognized that the use of it for a distinctly different purpose, such as data transmission, may impose compromises both in the medium and in the special service contemplated.

A short time ago the authors were assigned the problem of examining possibilities for such an adaptation aimed at high speed, and exploring the nature of the transmission compromises that might be needed. As a result, a variety of data transmission systems in different stages of development have been investigated including some telegraph systems. Certain conclusions have been reached regarding their suitability for use over private line telephone facilities. (By "private line" facilities are meant facilities leased to the subscriber on a more or less permanent private basis, and that are not set up by operators at a telephone central office switchboard.)

These conclusions are summarized below. In the later text there is included the background material for the conclusions which includes a brief characterization of the facilities. Finally there is included a discussion of the line treatment which may be needed for the best transmission.

It should be emphasized that not all possible data systems for use over telephone circuits are covered. The problem considered covers particularly some recently proposed applications, for which the need of a relatively high bit rate is important. Also, only the more promising comparatively conventional systems, which have been relatively well tested and can be readily applied, are considered. More radical designs are conceivable but they would require more extensive investigation before conclusions could be reached concerning them. It is clear also that the designs involved in the choice of a system are determined by the type of service it is to provide.

1.1 *Conclusions*

It is concluded that about the fastest transmission of data which can be accomplished with the present art over message-type telephone facilities is obtainable with an amplitude modulated vestigial sideband system. Such a system will be presented in some greater detail below. Its frequency spectrum is similar to that of a telephotograph signal[2] and a number of the transmission problems involved are the same.

This system will provide about 1600 binary digits of data (or "bits") per second, but it requires some special selection and considerable treatment of many types of circuits. This treatment is necessary to reduce noise, particularly of the impulsive type, and to correct for delay distortion.

Such a system is therefore considered suitable where a high bit rate is essential. It will be developed in the text that beyond the matter of the delay correction and other treatment of circuits required, the vestigial sideband process imposes a certain signal-to-noise penalty. A further signal-to-noise penalty comes from the method of multiplexing an auxiliary channel needed for word start indications.

Where a lower bit rate is acceptable, a 750 bit per second amplitude-modulated, double-sideband system which is now available, and has been tested extensively, is somewhat more rugged. It permits operation over the great majority of telephone facilities and will also be described in more detail.

The most rugged system considered uses frequency-shift transmission. This form of transmission has been used extensively for some time on a multiple-channel, voice-frequency telegraph basis in which each channel is capable of 46 to 74 bits per second. When used in this form to handle high speed data signals, it requires relatively complicated terminal devices because the several channels have to be merged to provide one high speed data system. However, when frequency shift is used as a single channel over an untreated telephone message facility the system promises to be relatively simple and to give a total possibility up to 1,200 bits per second according to the type of facility.

1.2 *Summary Table*

These findings are concisely grouped in a summary table, Table I. These entries are based upon present knowledge and are believed to be reasonably accurate, although estimates regarding impulse noise need more extensive checking, particularly in the case of the broad-band, frequency-shift system.

The table compares relative estimated performances of three broad-band systems among themselves, and with a subdivided channel (or telegraph) system. The performances considered cover the effects of noise and delay distortion, and the bit rates considered achievable in a 1,000-cycle band and in a 300- to 2,800-cycle telephone band. Some crude approximations covering speech are also given as a matter of interest.

The noise performance is given in terms of relative total power capacity required in the line for a given error rate in the presence of a given noise. The double sideband system is taken as reference. Allowance is made in the multiple channel or telegraph system for occasional peaking caused by temporary unfavorable phasing. In this part of the comparison a 12 channel system has been assumed. This is about as many channels as can be used on a telephone facility that has heavy impulsive noise.

The delay distortion figures represent some present ideas on good engineering design in the allowable impairment of signal-to-noise ratio.

TABLE I — SUMMARY OF VARIOUS DATA SYSTEM CHARACTERISTICS

Use in Band	Relative Peak Power Required in Line for Equal Performance* in Presence of Noise, db			Approx. Max. Delay Distortion in Millisec. Allowable in Band (for 3 db noise penalty)		Approx. Bits per Second in 1000 Cycle Band (25-db S/N Random Noise for DSB)	Approx. Max. Bits per Second in Telephone Facility 300–2,800 Cycles
	SF	Random	Impulse	Within One Channel	Over 1000 Cycle Band		
Broadband VSB	+6	+6	+6	±0.4	±0.4	1000	1600**
Broadband DSB	0	0	0	±0.55	±0.55	700	800 to 1400**
Broadband FS	−3	−3	−7†	±0.5	±0.5	650	750 to 1200**
Telegraph Comparison							
43A1 Channel	+10††	0††	−10††	±5	±60	450 (6 Channels)	1100 (15 Channels)
	(——12 Channels——)						
Speech Comparison							
Speech	+10	+10	+10	±10			50

* High grade performance; i.e., less than 1 error per 100,000 bits.
** High figure assumes accurate delay correction and control of nonlinearity.
† Depends on precise line-up of filter and carrier.
†† Allowance for peak factor of 12 frequencies.
VSB = Vestigial sideband; DSB = Double sideband; FS = frequency shift.

This impairment is here assumed to be about 3 db. A rather larger impairment has on occasion been assumed by other authors.[3] In the case of the telegraph system some arrangements for merging the signals in parallel channels into a single high speed channel require a certain exactness in timing correlation. The need for this may be overcome by the use of a small amount of data storage in the receiver of each channel. The delay distortion figures quoted in the table assume no need for this timing exactness. As in the other part of the comparison twelve channels are assumed.

The last two columns indicate the bit rate that can be expected of the various systems, first per 1,000 cycles of band, and second for a telephone facility of somewhat narrow (but frequently encountered) bandwidth. Some of these figures assume a careful control of delay distortion and of nonlinear distortion. In the 1,000-cycle band only six channels of the telegraph assumed may be accommodated. In a 2,500-cycle band the number can be extended to 15. The band that can actually be used for telegraph, over a given facility, depends upon the nature of that facility.

Some comparison data are indicated for telephone speech. The bit rate given assumes particularly message communication and not finer shades of artistic expression. For this a collection of phonemes (or elementary sounds) in the fifties, needing six bits for identification, and a typical speech rate of eight phonemes per second, require 48 bits per second. A few bits are also needed for pitch indication, and the figure is rounded to 50.

The low bit rate obtained for telephone speech communication indicates considerable redundancy in the speech signal sent over the telephone channel. This suggests why substantial transmission impairments can be tolerated without destroying the intelligibility of speech, as compared with telegraph or data signals.

1.3 *Nature of Study*

The procedure followed has consisted of examining systems of binary or similar signal transmission which appeared suitable for the sending of data. The systems studied are listed here, and some description of them is given later in the text. As noted, part of the information has come from outside the Bell System.

1. Exploratory 1,650 bit per second vestigial sideband system studied at Bell Telephone Laboratories.
2. Exploratory 1,600 bit per second vestigial sideband system studied at Lincoln Laboratory of M.I.T.[1]
3. Exploratory 750 bit per second double sideband system, of general type reported by Horton and Vaughan.[3]
4. Voice frequency telegraph channels.[5]
5. "Polytonic" signaling system reported by Lovell, McGuigan and Murphy.[4]

The transmission problem of applying these various systems to the various types of telephone message facilities employed for private line service has been considered. These are also listed, and a brief characterization of them is given in the text.

1. Voice frequency circuits, over cable and open wire.
2. Type-C carrier circuits for open wire.[16]
3. Type-N carrier circuits for cable.[6]
4. Broad-band carrier systems[15] using A channel banks for paired cable, coaxial cable, open wire, and microwave radio.
5. Other broad-band carrier systems.

1.4 *Outline of Paper*

As a result of the study, some recommendations were made. The discussion of this material is presented in Section II. It covers first the sort of service, with comparatively high bit rate, for which there has been user demand, and which can be furnished over private line facilities. Secondly, the recommendations cover the broad features of the signal characteristics which appear promising for use over such facilities. These recommendations are at the present time evolving into an exploratory system. The recommendations themselves, together with some general remarks, are presented in Sections 2.0 and 2.1. The background material on which these were based, and which covers consideration of the five systems which have been listed, is presented in Sections 2.2 to 2.5.

The discussion on the nature of the problems involved in transmission of the signals over the telephone plant to be used (which above has been listed in five categories) is finally covered in Section III.

II. SYSTEMS OF BINARY SIGNAL TRANSMISSION

2.0 *General Remarks*

There are a number of arrangements in the present art, both experimental and commercial, which are capable of sending digital data information under something like the conditions required for the service considered. Study of these has led to a set of recommendations which are outlined below. The specific arrangements are then discussed in more detail. The description covers only the essential transmission characteristics of the arrangments.

The arrangments generally divide into two groups, those which are essentially short-pulse single channel (though some of these may include a slow auxiliary channel), and those which use frequency division multiplex channels and therefore employ longer individual pulses. One important advantage of the multiple channel systems is noted in Section 3.3 as consisting of an increased immunity against impulse noise.

The single channel group shows much similarity among the systems. The principal difference is that of bit rate. The faster systems use vestigial sideband transmission, at the expense of a certain increase in vulnerability to noise as compared with those which use double sideband transmission, and also at the expense of a more general need for delay distortion correction because of the speed.

The systems in this group which use vestigial sideband are expected to be able to operate, with a very low error rate (some 1 error per 100,000 bits), over telephotograph type circuits, which employ delay distortion correction. This is a "normal" condition error rate, and does not include abnormal circuit conditions, such as static, trouble conditions, etc. It also assumes a more or less even error distribution.

One of these systems gives very desirable word start indications by using an additional level of signal. "Words" are groups of signal elements of fixed total number each. A distinctive separation signal greatly simplifies their recognition at the receiver, particularly after short line interruptions.

As the price, however, of both the speed and the third level signal indication, this system is more vulnerable than the others to impulse noise, of the kind encountered in hitherto installed N1 carrier and open-wire circuits, which have not been treated for impulse noise. It is also more vulnerable than the other systems to sudden level changes in the received signal.

The multiple channel group is generally characterized by a lower bit rate. Of all of them only one, the frequency-shift carrier telegraph, shows capability of consistent operation over untreated N1 carrier. On this type of carrier the frequency-shift telegraph permits a bit rate of the order of only half that obtainable with the vestigial-sideband systems. This telegraph system performs much better over other types of telephone facilities but even then its bit rate is only about three fourths that desired.

2.1 *Digital Data Transmission Service*

2.1.1 *Some Desirable Major Requirements*

1. Transmission of a maximum of 1,600 bits per second with an error rate not to exceed one in every 100,000 bits (or once per minute).

2. Applicability to most telephone facilities for private line use. Some selection of circuits and some treatment of those selected will be acceptable. The circuits carrying data are to be one-way terminal circuits only (i.e., not to be linked in tandem). It is expected that for the bulk of the service the circuits would not be likely to run over some 200 to 300 miles in length, but a small number of 3,000-mile circuits is considered possible.

2.1.2 *Characteristics of the System*

The system which has been considered and recommended as promising for the service outlined above has the following essential characteristics:

1. Carrier at 2,000 cycles, with vestigial band extending up to some 2,400 cycles. Lower nominal effective band (half the bit rate in width)

extends down to 1,200 cycles, and roll-off band down to about 1,000 cycles. The frequency space above the vestigial band and below the roll-off band is essentially "dead" space, free of signal energy.

2. The signal comprises three components:

(a) *synchronization*, (or start), to indicate word separations,

(b) *data*, or the actual information, and

(c) *timing*, to mark out the successive bit intervals in the data.

The signal is represented in Fig. 1 as the envelope of a carrier. The synchronizing signal has an amplitude of one and a duration of one signal element. Data spacing signals have an amplitude of about 0.63 and a duration each of one signal element. Data marking signals have an amplitude of about 0.25 and also one signal element duration. It will be noted that this is an "upset" signal, in which spaces have more power than marks. The two signal elements immediately before, and again the two signal elements immediately after, the synchronization signal are always to be spacing. A timing signal is not actually sent, but instead the synchronizing signal is used to pull the phase of a highly

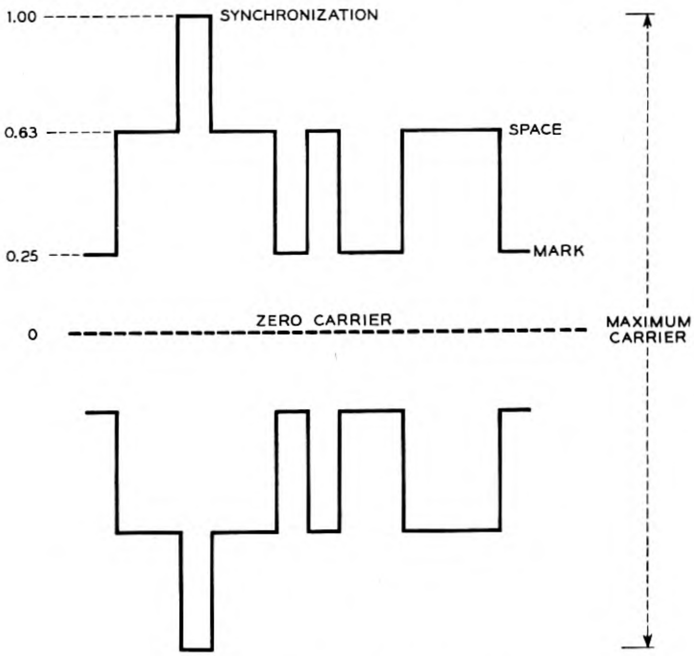


Fig. 1 — Multiple level signal.

accurate oscillator or "clock" at the receiver into the proper phase relation at the beginning of each word.

The phase of the oscillator is readjusted slightly as necessary at the beginning of each word but this adjustment is purposely made sluggish so that an occasional noise burst will not throw timing out badly. Thus the timing information is, in effect, by its repetitive characteristic, highly redundant. The receiving device is capable of getting in step without any manual adjustment but it may require as many as 10 words (or that number of synchronizing pulses) before it gets into exact phase, on an initial connection or after losing synchronization.

2.1.3 *Transmission Considerations*

Such a signal leads to the following transmission considerations:

1. The signal is rather similar in its spectrum to a telephotograph signal, and as a first order approximation requires about the same type of facility for its transmission.

2. In particular, the signal is expected to require something like the same over-all delay equalization as the telephotograph signal. A brief discussion of this point was given in a paper by one of the authors.[7] The conclusions there are that the telephotograph equalization limits of ± 0.4 times a signal element duration in envelope delay are generally reasonable, although these are probably overly severe with respect to very fine structure irregularities in the residual envelope delay characteristic. The formal limits which have been set on the envelope delay distortion for the 1,600 bit per second signal are ± 250 microseconds.

3. These limits constitute a rather less severe problem for the bulk of the expected circuits (200–300 miles) than they present for telephotograph circuits which are equalized for 3,000- to 5,000-mile lengths. For the small expected number of very long circuits the problem would of course be the same as for the telephotograph service.

4. The delay equalization problem requires special consideration in view of the nature of practices which have developed in the telephotograph art. The number of circuits which have been equipped for telephotography has in the past been rather small. The adjustment of the delay equalization has involved arithmetical calculation in the process of fitting various manufactured delay equalizer sections to the measured delay distortion. These methods are not generally suitable for large scale operations. It is believed that a process of equalization by prescription will be useable for the 200–300 mile circuits. This is discussed in more detail in Section 3.3.

5. So far, it has not been found expedient to transmit telephotograph signals generally over unmodified N1 carrier or other compandored

circuits, and the same problem is appearing with the proposed data signal. The problem is discussed in more detail in subsequent sections.

6. The specific signal suggested has the disadvantage that multiple levels must be discriminated by the receiver. This increases vulnerability to noise and level changes, as will now be discussed in some more detail.

The specific signal which has been suggested, as noted before, is outlined in Fig. 1. This signal comprises several levels. In the first place it includes a word start indication, on a separate level from the mark and space indications. In the second place the lowest amplitude level (normally a space, but a mark in the "upset" signal, as has been noted) is not made zero, but 0.25 the amplitude of the word start indication or "synchronizing" pulse. The need for this is explained in the discussion on vestigial sideband, below.

Consider for the moment a signal of only two levels; these can be taken as 1 volt and 0 volts respectively. The discrimination between them without error requires that an instantaneous noise pulse at this time be kept to less than $\frac{1}{2}$ volt. In these terms this represents an S/N ratio of 6 db.

The use of 0.25 volt minimum signal means that the amplitude range between maximum and minimum is reduced from 1 volt to $1 - 0.25 = 0.75$ volts. The maximum allowable noise pulse must be $0.75/2 = 0.375$ volts for this signal. This is an 8.8-db S/N ratio, as compared with the previous 6 db. It represents a handicap of approximately 3 db which must be accepted as part of the price of the increased bit transmission rate permitted by the use of vestigial sideband as compared with double sideband transmission.

An additional penalty comes from the use of three as against two signal levels. The spacing signal level is set at 0.625 volt, midway between 1 and 0.25 volt. Discrimination between synchronization and spacing signals can tolerate noise pulses of $(1 - 0.625)/2 = 0.1875$ volt. This amplitude is 15 db below synchronization level. Discrimination between spacing and marking signals tolerates maximum noise pulses of $(0.625 - 0.25)/2 = 0.1875$ volt. This is again 15 db below synchronization level. Thus the signal tolerates a 15-db S/N ratio between synchronization level and the level of maximum noise pulses.

The difference between the approximate 9-db S/N for the two level vestigial sideband signal and the 15 db represents the 6-db handicap caused by the multiple level discrimination in the signal. This is the price paid for a distinctive word start indication.

The price also applies to sudden level changes. In a two level signal between 1 and 0 volt, a sudden drop of 6 db (without compensating change in the "slicing" or critical level) causes error.

In the two level signal between 1 and 0.25 volt, the permissible drop is reduced to 4 db (ratio of 0.625/1).

In the three level signal, the permissible drop is reduced to 1.8 db (ratio of $(0.625 + 0.1875)/1$).

Automatic gain control and corresponding adjustment of the slicing level ameliorate these conditions to some extent, but the problem is still a serious one; a sudden rise is also serious.

The use of a compandor in the transmission facility exaggerates the situation. Some discussion of the action of a compandor to improve the signal-to-noise ratio for speech is given in Section 3.1. For the moment it can be said that, at the transmitting end, the compandor compresses the range of amplitudes in the speech signal at approximately a syllabic rate. At the receiving end an expansion restores the original amplitude relationships in a complementary manner.

The compression and expansion are matched almost perfectly as far as the human ear is concerned. However data transmission is more vulnerable to short time level irregularities. This allows small imperfections in the amplitude restoration to impose a further penalty in the form of error hazards.

2.1.4 *Other Considerations*

Present ideas call for the acceptance of the signal by the telephone company, and redelivery to the subscriber not in the form indicated by Fig. 1, but in the form of the three separate components listed in 2.1.2. This presents some problems regarding the transmission of such signals over the connecting loops between the subscriber and the telephone office. These are not, however, germane to the general transmission problem and will not be considered further here.

2.2 *Bell Telephone Laboratories and Lincoln Laboratory Vestigial Sideband Systems*

The first of these represents an unpublished exploration, particularly by C. B. H. Feldman and A. C. Norwine, of the possibilities of transmitting moderately high speed data pulses over telephone facilities.

The exploratory system used a carrier at 2,200 cycles; a vestigial band extended up to 2,600 cycles; the nominal effective band extended down to 1,375 cycles; and a roll-off band continued down to 1,100 cycles. In order to reduce the quadrature component resulting from the vestigial sideband the spacing signal was made equal to one-third the marking signal amplitude. The receiver used AVC both for amplitude control and space-to-mark slicing level adjustment.

The timing of the receiver sampling instants to determine mark and space indications was determined by a flywheel circuit which operated from space-to-mark and mark-to-space transitions in the signal. This is similar in principle to the old Baudot quadruplex arrangements in telegraphy. It was a synchronous system and required a dummy signal for a lining-up period previous to the actual transmission of information. The lining-up was automatic but required some 15 to 50 milliseconds (25 to 80 signal elements).

A number of experiments with the system were made on actual lines. Most of these showed successful transmission, though the error rates were not measured quantitatively. The test showed the signal margin against error on a cathode ray oscilloscope. The wave trace indicated displacements both in the signal amplitude and in the timing of the sampling instant. This margin has been found to correlate reasonably well (in an inverse relationship) with the calculated delay distortions of the circuits used. It also corresponds reasonably well to theoretical expectations.[7]

The system reported on by the Lincoln Laboratory of MIT[1] shows much general similarity to the above. Perhaps the most distinctive feature of difference is in the use of a word start indicating or synchronizing signal, in the form of the high level pulse discussed above and somewhat similar to that used in television for scanning-line synchronization.

2.3 *Double Sideband Systems*

A prototype of these systems has been described by Horton and Vaughan.[3] Several models have been derived from the prototype which differ from it, somewhat, in several respects. Large portions of the systems are not germane to the present discussion, and only a brief description will be given of the signals.

An outline of the signal spectrum for the most recent of these derived models is illustrated in Fig. 2. The main signal is handled on a carrier at 1,500 cycles. The bit rate is 750 per second, and the nominal effective band is shown as ± 375 cycles. A schematic roll-off is indicated. The words in this system are of about 100-signal element length, of which 8 are used for synchronization. The synchronization pattern involves a 3-signal element ready pulse on the 600-cycle carrier, simultaneous with the first 3 of the 6-signal element marking pulse on the 1,500-cycle carrier. Following this comes a spacing bit, then another marking bit on the 1,500 cycles only. The next bit is the first information bit.

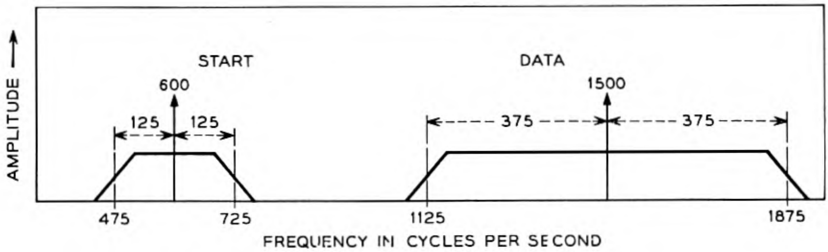


Fig. 2 — Double sideband signal with auxiliary start channel.

Marking consists of maximum carrier amplitude and spacing is zero carrier amplitude. The pulses are shaped both at the transmitting end and at the receiver before sampling.

The earlier prototype system as described by Horton and Vaughan was tested over a variety of telephone facilities (not including N1 carrier) running up to over 12,000 miles in length and found to be quite rugged. For reasons which are to be discussed later, the use of these systems over compandored circuits presents certain extra transmission problems regarding noise and level changes.

2.4 *Voice Frequency (VF) Carrier Telegraph*

The opposite extreme to the use of the telephone facility as a single band, to carry short duration pulses, is to subdivide the band into a large number of subchannels, each using longer duration pulses. As already noted, this has advantages against impulse noise, and also delay distortion.

There is available for this use the VF telegraph system.[5] In an AM form (40C1) this subdivides the space into 18 telegraph channels, which can each carry 100 words per minute (or 74 bits per second). A frequency shift form (43A1) is available, to give 17 channels, each to carry 74 bits per second.

The 18 channels use a band of 200 to 3,200 cycles in the telephone facility, and the lowest channel permits only a lower word speed. The 17 channels occupy the band from 350 to 3,200 cycles.

It has been mentioned that the telephone channels which provide the most serious problem for data transmission are those using compandors. The untreated N1 carrier is a principal example of such. Further, the type of circuits which use compandors are apt to be placed in plant which is relatively exposed to impulsive noise.

The principal interest in the telegraph channels, therefore, is to examine how they fare in application over untreated N1 carrier. There have been some studies of this point. The general conclusion, to be elaborated below, is that although the frequency subdivision (and also the frequency shift) helps against impulsive noise, fewer telegraph channels may be used than over good non-compandored facilities; and at best, the transmission is accompanied by more distortion than expected in a telegraph link of the highest grade. However, a serious possibility for data transmission is indicated.

2.4.1 *Telegraph Tests on N1 Carrier*

Tests on this subject have been carried out by S. I. Cory, J. M. Fraser and others and reported in unpublished memoranda. Before presenting the results of the tests, some background is necessary on the terms in which the results are reported.

The performance is usually evaluated in tests of this kind in terms of a "maximum checkable" telegraph distortion over a short time (about 5 minutes). The "telegraph distortion" represents the displacement from correct timing, of received signal transitions, after the initial mark-to-space transition in the "start" element. These displacements are measured in percentages of the signal element duration. By "maximum checkable" distortion is meant the maximum such displacement that is consistently reproduced in repetitions of the short testing period. This is somewhat larger than the root-mean-square distortion. A larger displacement is, of course, obtainable over a longer testing period. Although this measure of performance has had long use in the telegraph art, other measures are perhaps more readily grasped by and probably of more value to the data transmission engineer. Such a measure, for example, is the error rate.

The error rate may be estimated through the use of the telegraph transmission coefficient.[8] This is a figure which has been designed by telegraph engineers to indicate the performance of a telegraph circuit, particularly when it is made up of several sections. It is more or less proportional to the square of the distortion, and has the property that, when carefully chosen, it can be added for circuits connected in tandem. A small coefficient thus characterizes good transmission, and correspondingly, a large coefficient characterizes poor transmission.

The correlation between peak distortion over 5-minute intervals and error rate, through the telegraph transmission coefficient, is indicated in Reference 8 and Table II. It is to be understood that the correlation is only a rough one. Particularly at the two extreme ends of the scale, the entries in the table can serve only as a general guide to the performance, and the specific numbers are not to be taken too literally.

TABLE II — CHARACTERIZATION OF TELEGRAPH DISTORTION AND AND ERROR RATES BY TELEGRAPH TRANSMISSION COEFFICIENT

Distortion		Transmission Coefficient	Errors	
RMS	5 min. peak		1 in n characters	1 in m bits
			(n)	(m)
13.9	30	30	40	2.9×10^2
12.6	27	25	87	6.2×10^2
11.2	24	20	2.5×10^2	1.8×10^3
9.8	21	15	1.5×10^3	1.1×10^4
8.0	17	10	4.4×10^5	3.2×10^5
5.6	12	5	10^9	7.4×10^9
4.3	9	3	10^{12}	7.4×10^{12}
2.5	5	1	—	—

A very brief summary of some of the experiments in the use of VF telegraph over untreated N carrier is given in Table III. This portion of the results covers N1 circuits with compandors which have slightly more noise than the objective which is set for the telephone use of such circuits. The noise was 28 dba at the zero transmission level point, as against an objective of 26 dba at that point. It is noted in Section 3.3.2 that the measurement of noise in these terms is not altogether reliable in the evaluation of its effects on transmission systems that use pulses. Thus, these experiments must be considered as giving only a general indication of the situation.

TABLE III — SUMMARY OF TELEGRAPH PEFORMANCE OVER NOISY N-I CARRIER LINK

	40C1 (AM)	43A1 (FS)	
1. Number of channels	6	12	
2. Frequency space used	1020	2040 cycles	
3. Words per minute	75	75	100
4. Total bits per second	342	684	888
Average Channel			
5. Peak distortion	16	8	18 per cent
6. Estimated Transmission coeff.	9	2.5	11.5
7. Estimated errors, 1 in	10^6	10^{14}	10^5 bits
Worst Channel			
8. Peak distortion	25	17	22 per cent
9. Estimated transmission coeff.	21	10	17
10. Estimated errors, 1 in	1.5×10^3	4×10^5	5×10^3 bits

The tabulation is first given for an average channel. The performance of the worst channel has, however, also been included to give an indication of its contribution to over-all operation.

Many of the N1 carrier telephone circuits in the plant show lower noise than the objective, and to this extent Table III is somewhat pessimistic. Also some of the tests have shown, particularly for AM, that the performance is somewhat improved by removing the compandors. Thus, allowing for both points, better performance can be expected from the average N1 circuit (less than 200 miles) in the plant. The transmission coefficient of 11.5 listed for Item 6 at 100 words per minute might go down to say 9. At 75 words per minute with FS or AM, it might not be over 4.5. It is clear, of course, that with further modifications of the N1 channels, such as to reduce noise exposures, better performance could be obtained.

The broad conclusions that can be derived from these considerations are:

1. The subdivision of the frequency band into telegraph channels, and the use of FS, permit a workable system to be operated over a compandored facility like the N1 carrier without modification. This occurs even when the latter has noise up to and a little over the telephone objective.

2. This workable system under such noisy conditions transmits up to some 350 bits per second with AM, and some 800 bits per second with FS. It is accomplished with an error rate of the order that has been implied for data transmission, even in the worst channel.

3. There is a relatively wide range of performance of the system over different N1 circuits, and the average performance is sensibly better than that under the limiting conditions which have been considered.

2.4.2 *Distribution of Signal in Allocated Bandwidth*

A more extensive discussion of the use of bandwidth is given in Section 3.1, below. However, a few specific points are appropriate here on the band use in telegraph channels.

The spectrum of the original voice frequency telegraph system, based

TABLE IV — USE OF FREQUENCY SPECTRUM IN TELEGRAPH CHANNEL

	AM	FS
1. Channel spacing	170	170 cycles
2. Nominal effective bands	74	74
3. Roll-off band (both sides)	37	26
4. FM swing	0	70
5. Guard band (both sides)	59	0

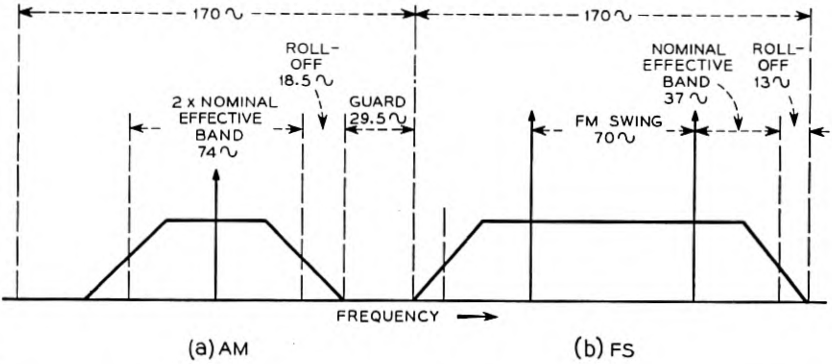


Fig. 3 — Utilization of telegraph channels.

on 170 cycles between carriers, was conservatively developed for the 60 word per minute speed of the time. The 100 word per minute speed has used up some of this conservatism. The use of frequency shift in the same channels has, however, used up the spectrum space even more.

An outline of the band allowances is given in Table IV, and illustrated in Fig. 3. Item 2 of the table is based on the 100 word per minute speed, using double sideband. On this basis, the number is equal to the number of bits per second. This is the minimum double sideband over which that number of bits can be transmitted, according to the Nyquist theory. Each such sideband is sometimes called a "nominal effective band." In practice various allowances are necessary over this minimum.

In the first place a roll-off is necessary because filters are not infinitely sharp, and in addition the nature of the modulation itself forms a roll-off. Roll-off also leads to a signal which is more free of overshoots and generally "cleaner" than when a sharp cutoff is used. Item 3 and Fig. 3(a) and (b) show an allowance for roll-off. For the AM case this amounts to half the nominal effective band. For the FS case there is not quite that much space available.

For the FS signal it is necessary to allow for the frequency swing as Item 4. For the 43A1 system this amounts to 70 cycles. In Fig. 3(b) the spectrum includes the region comprised by the FM swing, and upper and lower sidebands. The upper and lower sidebands as formed by the modulation of a random signal are shown extending respectively above and below the extremities of the swing (instead of only above and below a central carrier, as they would with AM).

A final allowance in Item 5 is a "guard band." This is taken to mean a region in which the signal energy is negligible, but at the same time a region in which no appreciable interference is tolerated from the adjacent band. The allowance for this is generous in the AM case. In the FS case the roll-off band of Item 3 tends to use up all the space not employed by the nominal effective band and the swing, and nothing is indicated as left for guard band. This is illustrated in Fig. 3(b) by the extremities of the roll-off band reaching the extremities of the 170-cycle spacing. It is to be recognized that the illustrations are diagrammatic. However, the extremities mentioned measure the bands occupied by power from random signal transitions between marking and spacing (as distinguished from mere mark-space reversals), analogous to the bands occupied by power from AM signal transitions. Comparison of Figs. 3(a) and 3(b) suggests how modulated signal components of significant intensity are displaced farther from the edges of the channel with FS than with AM.

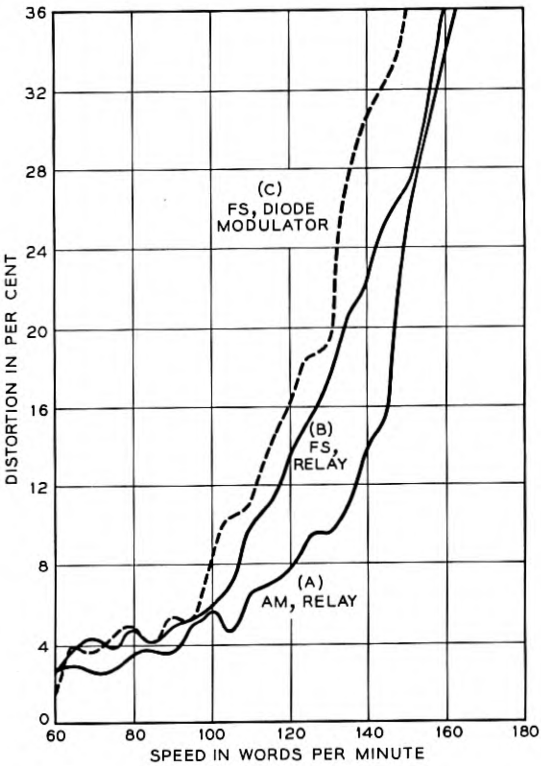


Fig. 4 — Experimental relation between telegraph distortion and speed, with fixed channel filters (after Jones and Pfleger[9]).

The conclusion is reached that there is hardly any excess conservatism in the 43A1 system. Some confirmation of this is indicated in a paper by Jones and Pfleger.[9] Fig. 4 reproduces some curves presented in that paper. The curve at (B) (from Fig. 5 of that paper) shows that the FS telegraph rises rapidly in distortion above the 100 word/min speed. The curve at (A) (taken from the same Fig. 5) for level-compensated AM shows a substantially broader and somewhat lower curve in this region. From curve (C) (taken from Fig. 3 of the paper) for diode modulated signals, it is seen that the sharp rise in distortion for FS is even more accentuated than for the relay modulator. This indicates how much more characteristic distortion FS exhibits than AM because of the sharper roll-off of its signal bands within the confines of the same 170-cycle channel spacings. Of course, as the word speed is raised beyond the practically usable values, either type of modulation leads to so much power in the filter cutoff regions that the characteristic distortions become more or less indistinguishable.

2.5 "*Polytonic*" *System*

This is a frequency discrimination system experimentally proposed for toll and local signaling.[4] It works on 5 channels at speeds of 100 and 300 decimal digits per second (or the equivalent of some 330 to slightly under 1,000 bits per second). It has some similarities to an earlier multi-frequency system,[10] but is faster.

The distinguishing characteristic of the polytonic system lies in the mathematical theory which has been followed to reduce interchannel interference. This analysis makes use of the theory of orthogonal functions, and is similar to that used in the computation of Fourier components. The mathematical analysis leads to an ideal receiver design for minimum interchannel interference. The ideal detector in this receiver closely resembles the conventional homodyne detector. The detector actually used, however, represents a practical simplification of the latter. A complete description cannot be given here, but it may be noted that the theory leads to a need for synchronization of the signal elements in the five channels, and to the setting of an exact timing instant for the sampling of the received wave to obtain minimum interchannel interference. The indication for this instant is obtained from the use of a sharp wave-front pulse in the marking channels.

Tests with the 100 decimal digit per second system (100 decimal digits normally correspond to 332 binary digits) indicated that it gave good operation[4] over toll circuits of comparatively limited length (350 miles for four-wire voice frequency, 1,900 miles for K carrier). The higher speed system was designed only for local plant.

It is clear that this system has too low a bit rate for the application contemplated, even in the faster form. It is also not generally adaptable to the variety of circuit lengths which are expected to be encountered.

III. UTILIZATION OF THE TELEPHONE CHANNEL

The discussion herewith covers a broad examination of some major characteristics of telephone communications facilities, to evaluate their bearing on the choice of a system for the data transmission service outlined before. It is of course clear that different conclusions might be reached for other types of service.

The first item is an outline review of the different types of message telephone facilities in the plant. This is followed by an analysis of the different possibilities in the use of the frequency spectrum, of noise, and of delay distortion, in the application of the data signals.

3.1 *Telephone Facilities*

There is a rather wide variety of facilities to be found in the telephone plant, to be examined with respect to the factors that are germane to the present question.

The first of these factors is the frequency bandwidth capability of the facility. For message-type voice circuits, this is generally characterized as being three kilocycles (with the exception of "emergency banks," which are substantially narrower, and are not to be considered as useable for data transmission).[17] However, some of the telephone circuits in the plant, aside from the emergency banks, are also somewhat narrower than 3-kc, and in any case, not all the band is effectively usable for data transmission. As will be seen, the net available band is, in practice, about half of the 3kc.

Part of the reason that not all of the frequency band is effectively usable is that the circuit shows delay distortion. This tends to become large at both the lower and the upper edges of the band. Some details of the delay correction are discussed further below.

Another impairing factor in telephone facilities is the nonlinear distortion encountered. In voice frequency facilities this comes from amplifiers and loading coils, and increases progressively with circuit length. In carrier facilities the nonlinear distortion arises almost exclusively at carrier terminals, in the part of the circuit where the signal is at voice frequency. In such a case, the distortion increases with the number of times in the telephone facility that the signal is modulated down to voice frequency. Second order nonlinear distortion tends to develop modulation products in the lower portion of the transmission band which are a source of potential interference with the signal.

Another impairment encountered in carrier telephone channels is a slight frequency shift; that is, a 1,000 cycle input may appear at the output, say at 998 cycles. This occurs because modulator and demodulator frequencies are not identical. With independent oscillators on recent systems this shift may amount to some two cycles. With older systems it can run from 5 to 10 times as much. The effects of this shift are discussed below. The frequency shift may be avoided by working double or vestigial sideband and using an envelope detector, or in some carrier systems by locking the oscillators in a constant frequency network. This locking may or may not result in close phase synchronization of the carriers, depending on the method used to lock and the particular carrier system involved.

Still another factor is the use, or not, of "compandors." A compandor compresses the range of speech volume in the impressed line signal and correspondingly expands this range at the receiver. This raises the line signal level during periods of low speech power, and lowers it during periods of high speech power without, in principle, affecting the final received level. The effect is to reduce the final noise in periods of low speech power, and increase it during periods of high speech power. A listener is less perceptive to the noise during high speech power levels than low. By this means, it has been found that the telephone circuit can be engineered to some 23 db more noise (and also crosstalk and similar forms of interference) than it can without the use of a compandor.

In the case of data signals, however, the influence of noise in causing error is not very much different whether the signal is marking or spacing. Thus, there is no "compandor advantage"* (indeed there is a certain disadvantage as pointed out earlier), and facilities that have been engineered to be entirely satisfactory for voice transmission are effectively some 23 db more noisy for data transmission. As a practical matter it appears desirable to remove compandors from circuits used exclusively for data.

A short listing is presented here of the various types of message facili-

* Perhaps a simpler way to think of it is that all possible "compandor advantage" has already been obtained in a data system by using the best combination of amplitudes for mark, space, start, etc.

ties most frequently found in the telephone plant, and some comments are made on each.

3.1.1 *Voice Frequency Circuits*

There is a variety of open-wire facilities of this type. They are mostly short, and two-wire. Thus, repeaters can to advantage be turned one-way for data service. Delay correction is discussed later.

Voice frequency cable facilities over more than a very short distance are loaded. This gives appreciable delay distortion. The loading used is indicated by a letter denoting the spacing, followed by a number denoting the loading coil inductance. Thus "H-44" means 6,000-foot spacing, of 44-millihenry coils, and "B-88", 3,000-foot spacing of 88-millihenry coils. Conductor capacities range from 0.62 microfarads per mile for toll circuits, to 0.82 microfarads per mile or sometimes even higher for local circuits. This affects the delay distortion.

3.1.2 *Type-C Carrier Circuits*[16]

This is an open-wire three-channel system operating at different frequencies in opposite directions, over the same pair. Historically there has been a variety of C systems developed, but only the C-5 system exists in any extensive quantity. The upper frequency cutoff in the voice channel is well under 3 kc. The delay distortion varies widely with the specific channel and direction of transmission. There is a variety of channel frequency allocations, and the distortion varies with this also. The delay distortion over some channels increases rapidly above 2,400 cycles. The frequency shift discussed before may be as much as ± 20 cycles.

3.1.3 *Type-N Carrier Circuits*[6]

This is a short-haul twelve-channel system for use over cables. Because of its economy it has been extensively introduced. Its principal characteristic, in the application of data circuits, is that it uses compandors. It therefore presents a noise problem. The delay distortion, introduced almost exclusively by the terminals is not excessive, and depends very little upon circuit length between the terminals. The N system uses double sideband transmission, and therefore exhibits no frequency shift between input and output signals.

3.1.4 *Broadband Carrier Systems Using A Channel Banks*[15]

There is a variety of carrier systems designed for paired cable, coaxial cable, open wire, and radio, that use a standard grouping of twelve channels with associated filters, known as an "A channel bank." The delay distortion in these associated filters constitutes nearly all of the distortion measurable over the complete system. These are single sideband systems, and unless the local modulator and demodulator carrier supplies are locked in a constant frequency network, frequency shifts of some ±2 cycles may be expected between input and output.

For paired cables, these are known as K1 and K2 systems. For coaxial, they are L1 and L3, for open-wire, J, and for microwave radio, TD-2.

3.1.5 *Other Broadband Carrier Systems*

An O carrier system has been developed for open wire, and combinations of it are used with N for open-wire and carrier. These are compandored systems.

3.2 *Use of Bandwidth*

This section examines the more important factors which affect the choice of how the available bandwidth of a facility is to be used, either in one band or a subdivided band.

3.2.1 *Baseband Transmission*

This is the simplest type of transmission. It is used in telegraph loops and other short distance telegraph transmission. A mark is indicated by placing marking voltage across the wire line, and a space by placing spacing voltage. In the simplest systems the latter is zero. In "polar" systems it is the negative of marking voltage.

The frequency spectrum of the signal runs down to and includes dc, as illustrated by the solid lines in (a) of Fig. 5.

With many transmission facilities it is difficult or impossible to transmit the dc; i.e., the circuit cuts off as is illustrated by the dotted lines. In such cases it is impossible to distinguish between a permanent mark and a permanent space.

Extra pulses can, however, be added to the signal to insure that marks or spaces are not permanent, but are relieved by the opposite signal in some maximum interval of time. In such cases the received signals can be clamped on mark or space signals and the opposite condition can be readily distinguished. This is sometimes called "dc restoration," and strictly speaking the system ceases to use baseband transmission. It may be designated as "modified baseband transmission." Methods other than clamping have been suggested for dc restoration.

Reverse pulses can be systematically inserted after each mark or space pulse, according to various patterns.[11] Two suggested are "dipulse" and "dicode" pulses. Such signals approach carrier signals, which are discussed below.

The principal weakness of baseband transmission appears when it is sent over C carrier or other single sideband telephone facilities, where the recovered signal may vary in frequency from that sent. This causes a distortion of the received pulse which makes it difficult to recognize.

An analysis of this point is given in Appendix I. It is concluded that while there may be long range possibilities in baseband transmission, it requires more study, and it will not be considered further in this paper.

3.2.2 *AM Carrier Transmission*

The simplest of this type is double sideband transmission, as illustrated by the full line of Fig. 5(b).

A comparison of the susceptibility to noise of this arrangement, with that of baseband transmission, is considered in the next section.

A further consideration required is susceptibility to nonlinearity in the facility. Second order modulation leads among other things to a rectification of the signal back to baseband. This is indicated by the dotted lines in Fig. 5(b). After such rectification of the signal by the facility, it is impossible to separate any overlapping portions of the signal between the baseband and lower sideband. Some overlap is shown. This interference was first considered in telephotography[12] and is known as "Kendall effect."

The possibility of Kendall effect may be eliminated insofar as second order modulation is concerned by moving the carrier frequency high

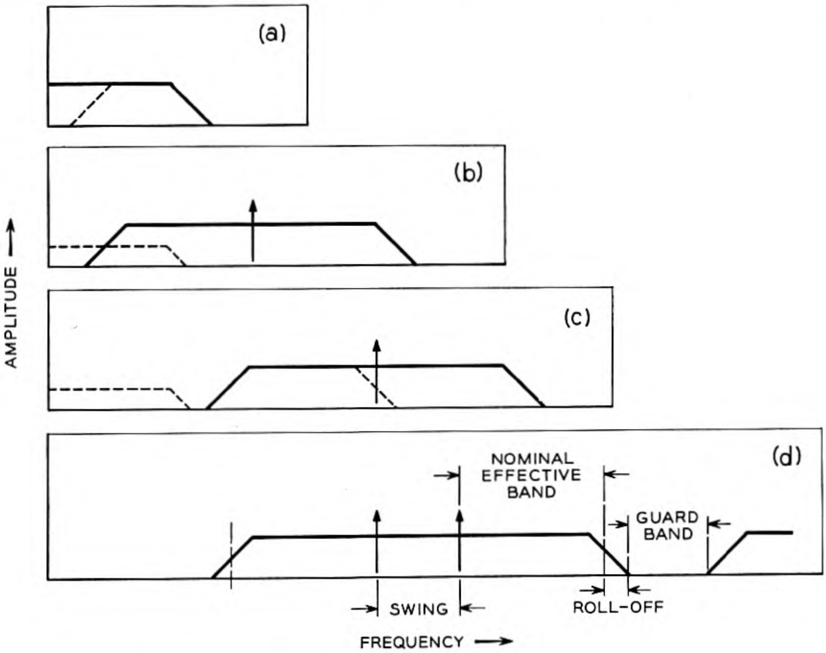


Fig. 5 — Spectra of signals with various modulations.

enough to prevent such an overlap. This is indicated in Fig. 5(c). It does not prevent third order modulation effects.

It has been found necessary to allocate frequency bands thus to avoid the overlap discussed in the transmission of high grade telephotography over telephone type facilities. It has also been noted that allocation with such overlap was undesirable in some data transmission experience. However, the question has not been resolved in complete detail. For the data service under consideration, which is expected to show a very low error rate, it is deemed conservative to allocate bands without the overlap.

This conclusion then leads to wasting a certain part of the lower frequency range. It is still possible, however, to use this range for an auxiliary signal, as in the system illustrated in Fig. 2, if the auxiliary signal occurs only during word starts.

The double sideband frequency range, as indicated in Fig. 5(b) or 5(c), is about twice the baseband range of Fig. 5(a). It is possible to reduce this extension by cutting down one of the sidebands to a "vestige" of itself and sending carrier at a reduced level, as indicated by the diagonal dotted line about the carrier in Fig. 5(c). This was proposed by Nyquist.[13] It is done at the expense of an increased vulnerability to noise, which in total amounts to some 5 or 6 db in certain typical cases.[7, 11] In Section 2.1.3 discussion was given to account for 3 db of this. In the references cited herewith it is noted that vestigial sideband transmission is accompanied by an interfering component (called a "quadrature component") which accounts for the other 2 or 3 db.

3.2.3 *FM Carrier Transmission*

Certain additional immunity to noise is gained by the use of frequency modulation (or "frequency shift") of the carrier. The immunity which can be obtained against impulse noise can be even greater than that against random noise, provided that the receiver is precisely tuned. This was noted in Section 2.4 in connection with voice frequency telegraph.

The noise immunity obtained from FS is in part due to the use of a higher average power and is at the expense of a wider frequency band as illustrated in Fig. 5(d). In addition to all the band that is used for a double sideband system, a space must be allowed for the swing. FS is also much less vulnerable to sudden level changes than DSB and thus may well be preferable to DSB for medium bit-rate service. As shown in Table I, these advantages can be obtained with only a small sacrifice of bit rate compared to DSB for equal bandwidths.

So far the single channel broadband FS system has not been generally used over wire circuits. Hence its exact performance, particularly with impulse noise, is only estimated.

On occasion not all of the band illustrated in Fig. 5(d) is allowed for. This leads to an increase in distortion of the signal, which has some similarity to a very close-in echo. It uses up some of the additional noise immunity provided by the FM, as an engineering compromise.

Another direction along which the FM system may be practically extended is to use four instead of two (marking and spacing) frequencies. This would double the bit capacity at the expense of only a moderate widening of the frequency band and somewhat tighter requirements on noise and delay distortion (but not of level regulation, which would be required for a similar extension of the AM signal).

3.2.4 *Multichannel Systems*

It is possible to divide up the entire frequency band available into a number of separate channels and use any one of the various carrier systems which have been described, in each individual channel. This may be done because the nature of the information transmitted may be better adapted to the narrow channel, as in conventional telegraph. It permits certain elements of flexibility in layout, and offers certain noise advantages (and also disadvantages) as discussed in the next section.

In an idealized way one can proportion the various allowances for nominal effective band, roll-off, guard band, and swing (FS) in the same proportion in which they would occur in a single broad channel over the whole facility. Thus no frequency space would be lost by the subdivision. In practice, however, subdivision usually does lead to some actual loss in the frequency space.

A significant limitation to frequency subdivision lies in nonlinearity of the facility. This leads to modulation products between the various channels, which interfere with other channels.

In the case of voice frequency telegraph and other multiple channel systems the modulation effects are mitigated by allocating the carriers at odd multiples of a basic frequency. That is, any given carrier f is set at $f = nk$, where n is odd and k is a basic figure. Then the three second order modulation products are

$$2f = 2nk,$$

$$f_1 + f_2 = (n_1 + n_2)\, k,$$

$$f_1 - f_2 = (n_1 - n_2)\, k.$$

Table V — Use of Telephone Band by Various Data Systems

System	Modulation	Max. No. of Channels	Bits/Sec. per Channel	Total Bits/Sec.
1. Proposed	VSB	1	1600	1600
2. Exploratory	VSB	1	1650	1650
3. Exploratory	DSB	1	750	750
4. Polytonic toll	DSB	5	100	500*
4. Polytonic local	DSB	5	300	1500*
5. 40Cl Teg.	DSB	18†	74††	1332††
6. 43A1 Teg.	FS	17†	74††	1258††

* Not realizable with 2 out of 5 codes used.
† Not realizable over some facilities.
†† Based on 100 word per minute channel capability.

All these products are necessarily even multiples of k and therefore always fall half way between carriers. This allocation, however does not permit mitigation of third order modulation effects.

3.2.5 *Experience*

Table V reviews systems which have been discussed earlier in this paper, to indicate the extent to which they use a general telephone channel facility, in terms of the bit rate output. It is clear, of course, that the various systems are not engineered to the same conservatism. These differences have already been commented upon.

The general conclusion which one can reach from the table is that the use of the medium in a single channel gives possibilities of a higher bit rate than subdivision. However, it is to be kept in mind that the telegraph facilities are conservatively engineered. Further as noted, they can be used to the full extent indicated only over the broader band telephone channels. For example, the full 18 telegraph channels can not be used over a C carrier telephone channel.

3.3 *Noise*

A general theory regarding the influence of noise on digital systems was presented in 1948 by Oliver, Pierce and Shannon.[14] This was discussed further at a symposium.[7] Several additional points are considered here, relating to the effect of noise on a data transmission service of the type contemplated.

3.3.1 *Effect of Channel Subdivision on Vulnerability to Noise*

This has been suggested earlier at several points. A more detailed discussion is given of the effects in Appendix II.

The conclusions reached there are, broadly:

1. Channel subdivision has comparatively small effect on vulnera-

bility to random noise. The small effect which does occur is a disadvantage which can run up to some 3 or 4 db for ten or twelve channels, for the multichannel as compared to the single channel system.

2. Channel subdivision is advantageous over single channel use, with regard to vulnerability to impulse noise.

3. Channel subdivision is disadvantageous compared to single channel use, with regard to vulnerability to single frequency noise.

3.3.2 *Noise Effects in Vestigial Sideband System*

Some brief discussion of the noise problems in the vestigial sideband system under consideration has already been presented in Sections 2.1.3 and 3.1. That such problems may be important in the use of actual telephone facilities has generally been checked by some unpublished tests carried out by J. Mallett, of Bell Telephone Laboratories.

The noise problem is most serious in the application of data transmission over N1 carrier. It is particularly significant for the N1 installations previous to the most recent. The most recent installations are engineered with distinctly more conservatism in regard to noise performance. As is noted in Section 3.1, the principal characteristic of the N1 system that affects this application is that it uses a compandor and that its design for telephone use assumes a reduction of the noise by this compandor. The reduction then is not realized for data signals. A second characteristic is that N1 channels are exposed to impulse noise. The channels are of course designed to limit such noise to the extent that it will not sensibly impair telephone speech. But short data pulses are more vulnerable to the impulse noise than speech. The 2B noise meter, which is normally used for telephone noise measurements, does not read sharp noise impulses according to their effect on data signals, and other methods of measurement have been explored.

A summary of some of Mallett's results is plotted in Fig. 6. Here the reading on the 2B meter is compared with that on a level distribution recorder which records peaks of 1 millisecond or longer. The "one per cent" point is noted, which means that one per cent of the one second intervals in the period of measurement contained one or more peaks (of 1 millisecond or longer duration) of the amplitude indicated. This is found convenient as a measure of the error frequency tolerated in the system considered (one in 10^5 bits).

The heavy dots indicate the correlation between the two sets of noise readings on idle tested channels of an operating N1 system (other channels in the system being busy due to normal use). All channels had compandors in. Dots that mark the boundaries of a group of tests (usually a single system in the group) are connected by straight lines. The open

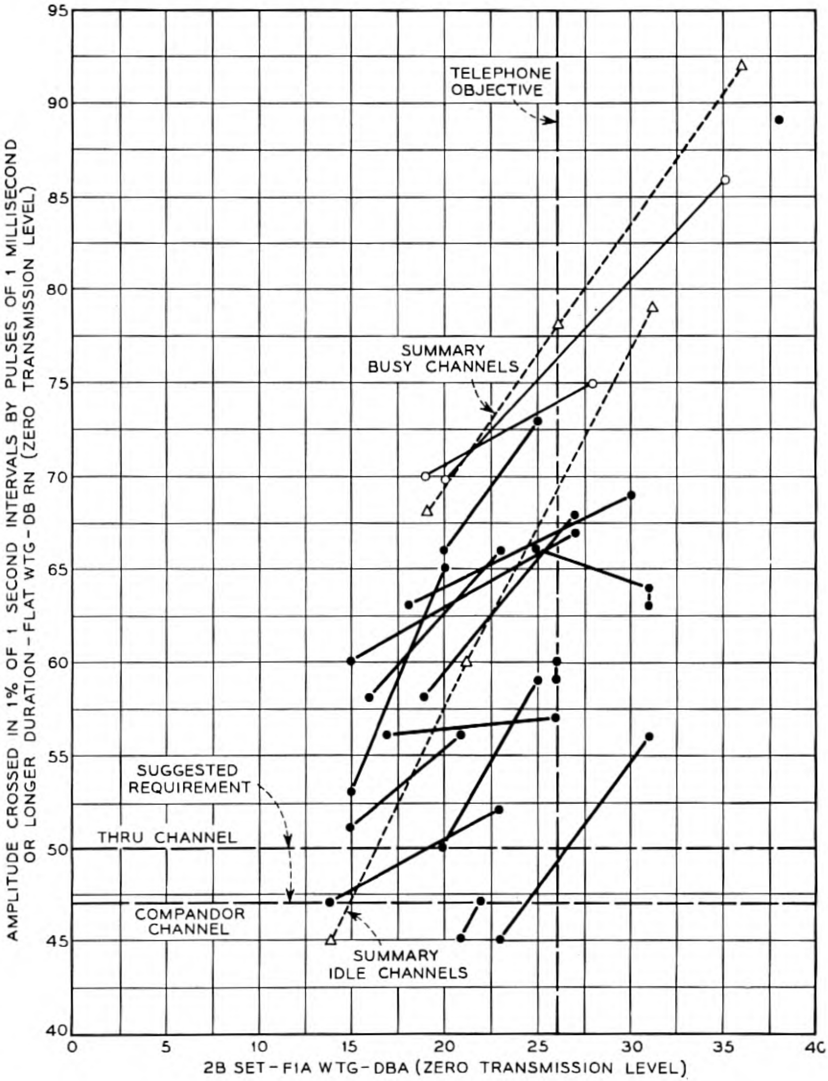


Fig. 6 — Noise measurements in N1 carrier facilities.

dots refer to an estimate (computed from the known properties of the compandor) of what the noise peak levels would have been, had the channel tested been busy with an operating data transmission signal (of the general type illustrated in Fig. 1). Under this condition the compandor setting would of course be different from that of an idle channel.

In such cases the 2B meter reading is also adjusted to the effective message circuit noise which would have existed in the presence of speech on the tested channel.

Open triangles connected by fine dotted lines indicate summary plots for the idle channels, and for the busy channels.

Examination of the plot shows that there is only a general correlation between the 2B set and level distribution recorder readings, as summarized by the dotted lines. Thus the 2B meter is not a reliable instrument to denote how noisy a circuit is for data transmission of the speed used in the proposed project.

The telephone objective for N1 circuits, as read on the 2B meter, is indicated by the vertical dotted line. This shows that most of the circuits (actually 55 out of 62) met the objective.

The suggested requirements for data are shown by the horizontal dotted lines. The "through" or noncompandored channel has a 3 db more lenient requirement (as indicated during some of the tests) than the compandored one. This results from the penalty mentioned earlier which compandors impose on in a data circuit, as compared with the 23 db or so advantage that it introduces for voice transmission. Only a few of the circuits measured (actually 10 out of 62) met the suggested requirements.

The principal conclusion reached from these measurements is that where used for a data service of the type considered, with vestigial sideband transmission, most of the hitherto installed N1 circuits (and probably other compandored circuits) will require modification to reduce noise exposure. Also noise measurement will be more complicated for such a service than for telephony.

3.4 *Envelope Delay Distortion*

Some simple theoretical considerations[7] have shown that the envelope delay distortion limits for a telephotograph circuit generally also hold for a data transmission circuit of the same speed. The principal difference is that less emphasis need be placed for data circuits upon fine structure deviations of the envelope delay as plotted on a frequency scale. In general, distortion of ± 0.4 signal element in the important part of the band has been found to give a signal-to-noise impairment of some 3 db in signal reception. This has been assumed here as a tentative engineering objective.

In accordance with this, the envelope delay requirements for service with the vestigial sideband signal consideration have been set at not to exceed 500 microseconds (± 250 microseconds) between 1,000 and 2,500 cycles. This contains an element of conservatism inasmuch as the strict requirement is really fully implied only on the nominal effective band (1,200–2,000 cycles). The signal power is reduced in the roll-off and vestigial bands, respectively 1,000–1,200 and 2,000–2,500 cycles, and some corresponding liberality may be expected there.

The delay distortion constitutes a more serious problem with a faster system as compared with a slower one, in part because of the wider frequency band occupied by it, and in part because 0.4 signal element represents a more severe tolerance in microseconds for a shorter element than for a longer one. Consequently, the limits given represent about as severe tolerances as may be expected to be needed with the use of a telephone channel.

The distortions of various circuits have been considered to estimate the order of the problem involved in meeting the proposed requirements over links of 100 to 500 miles.

The following conclusions are reached first for the vestigial sideband signal, and after this for the slower systems.

3.4.1 *Facilities Requiring No Treatment*

As already noted, K2, L1, and L3 carrier, and TD-2 microwave, use "A" channel banks to separate the individual channels, and these give the dominant delay distortion. This amounts to a maximum of about 200, and a minimum of 150 microseconds, according to the exact combination of filters used. This figure is for one link of transmitter and receiver. A single section delay equilizer can cut the maximum residual to about 80 microseconds. It is concluded that these facilities present no important delay distortion problems. An N1 carrier link gives a maximum delay distortion of 220 microseconds, which can be reduced to 50 microseconds by one section of equalizer. This, then, also presents no serious problem.

3.4.2 *Facilities Treated by Simple Prescription*

The delay distortion of H-44 voice frequency cable in the 1,000- to 2,400-cycle range runs to slightly under 900 microseconds for 300 miles, if the cable is of standard toll capacitance (0.062 mf per mile), and to slightly under 2,000 microseconds if of higher local plant capacitance (0.084 mf per mile). The use of about one section of equalization per 100 miles reduces the residual to less than 330 microseconds for the low capacitance cable. For the higher capacitance cable, about three sections are needed per 110 miles. The J-2 carrier uses A channel banks, but has, in addition, directional separation filters at each repeater. This gives maximum and minimum distortions, respectively, for 100 miles, of slightly under 300 and slightly under 160 microseconds. The precise distortion in any given channel depends upon its proximity to the cutoff of the directional filter. For 500 miles the figures are slightly over 500 and slightly under 50 microseconds. With the same single section of equalization the maximum figures are reduced to about 100 microseconds for 100 miles, and about 300 microseconds for 500 miles. To carry out this equalization requires only rudimentary information on the general nature and correction of delay distortion. If moderate care is used in the prescription of equalization on a packaged basis no delay measurement of the circuit would in general be needed, though it is recognized that some difficult cases may arise.

3.4.3 *Facilities Requiring More Involved Prescription*

The delay distortion in C-5 carrier[16] is influenced to a dominating extent both by channel and directional separation filters. It varies in a complex fashion from channel to channel, and according to the direction of transmission. Its correction thus requires more involved prescription than is required for the other types of circuit. In some few cases measurement may be necessary. The distortion of H-88 voice-frequency cable runs from some 1,400 to over 3,000 microseconds per 100 miles according to capacitance. For 20 miles of H-174 toll cable the distortion is slightly under 1,400 microseconds, and its use is not contemplated.

3.4.4 *Data Systems Requiring No Corrections*

The delay distortion problem is practically non-existent for the slower systems. For the double sideband systems some delay correction may be needed if long heavily loaded circuits are used or perhaps for some other rare unfavorable situations, but otherwise no correction is necessary. No correction is needed for the telegraph systems.

APPENDIX I — BASEBAND SIGNAL DISTORTION CAUSED BY CARRIER FREQUENCY SHIFT

A simple analysis of the phenomenon may be considered. Let the voltage input, as in Fig. 7(a), be a raised cosine pulse between the angular arguments of $-\pi$ and $+\pi$. That is

$$V_i = 1 + \cos \Omega t, \qquad (1)$$

where Ω/π is the envelope frequency. When this is transmitted on the carrier, $\cos \omega t$, the carrier signal voltage is

$$V_c = (1 + \cos \Omega t) \cos \omega t, \qquad (2)$$

$$= \cos \omega t + \tfrac{1}{2} \cos (\omega - \Omega)t + \tfrac{1}{2} \cos (\omega + \Omega)t. \qquad (3)$$

When the carrier and one sideband are removed (say the lower sideband), V_c becomes (neglecting the factor $\frac{1}{2}$)

$$V_c = \cos(\omega + \Omega)t. \qquad (4)$$

At the receiving end, V_c is modulated with a carrier which may momentarily differ in phase from the signal carrier by angle φ, giving

$$V_o = \cos(\omega + \Omega)t \cos(\omega t + \varphi), \qquad (5)$$

$$= \tfrac{1}{2} \cos[(\omega + \Omega)t - (\omega t + \varphi)] + \tfrac{1}{2} \cos[(\omega + \Omega)t + (\omega t + \varphi)]. \qquad (6)$$

The lower frequency part of V_o constitutes the recovered signal, and it is extracted by a filter that attenuates the higher frequency part. Thus, again neglecting the factor of $\frac{1}{2}$,

$$\begin{aligned} V_r &= \cos(\omega t + \Omega t - \omega t - \varphi), \\ V_r &= \cos(\Omega t - \varphi). \end{aligned} \qquad (7)$$

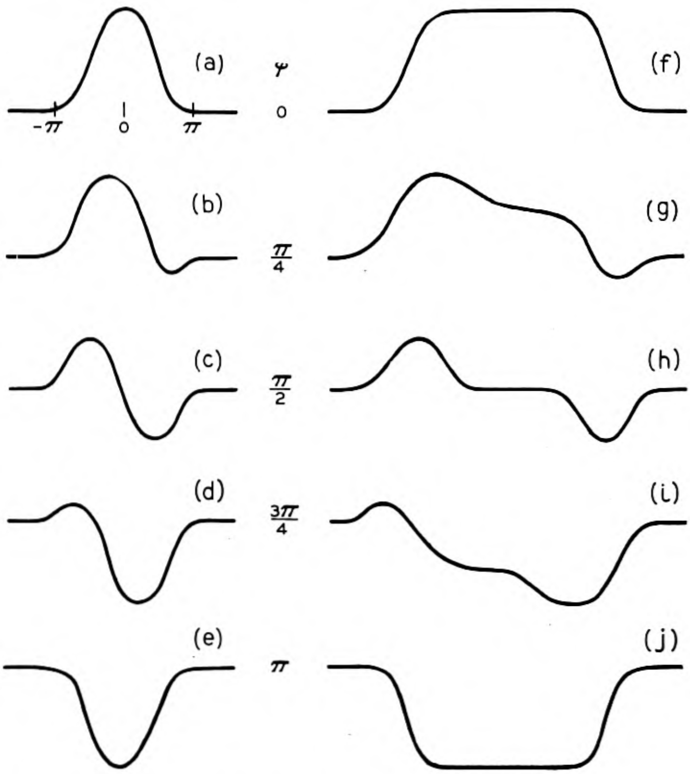


Fig. 7 — Distortion of baseband pulse signal in single sideband carrier facility.

The angle φ progresses through 0 to 2π per cycle of difference between the original and recovered frequencies. That is, if this difference is 2 cycles, then φ progresses from 0 to 2π twice each second.

The effect of the progression from 0 to π is illustrated in Figs. 7(a) to 7(e). When the phase is equal to π, as at 7(e), the signal is "upset"; i.e., marks are changed to spaces, and vice versa. When the phase is equal to $\pi/2$, the signal is effectively differentiated, as at 7(c).

In general, the signal as specified in (1) includes harmonics of Ω. An illustration of such a signal several dots long is shown in Fig. 7(f). As before, when $\varphi = \pi$, the complete signal is upset, as illustrated in Fig. 7(j). When $\varphi = \pi/2$, as at 7(h), the signal is more or less differentiated. The differentiation is not exact because the successive harmonics are not weighted according to order, as in true differentiation.

The distortions caused by the progression of φ are what make it difficult to recognize the signal at the receiving point. Some suggestions have been made for correcting the indication when the signal is upset, as in Figs. 7(e) and 7(j). It is more difficult, however, to take care of the intermediate cases, particularly 7(c) and 7(h).

APPENDIX II — EFFECT OF CHANNEL SUBDIVISION ON VULNERABILITY TO NOISE

There are three broad categories of noise to which the system may be exposed. These are:

1. Random noise which is not localized in time nor frequency.
2. Impulse noise which is highly localized in time but covers a broad frequency spectrum.
3. Single-frequency noise, which is highly localized in frequency but which lasts a significant time (or substantial number of bits).

We can assume two systems, A used as a single-frequency band, and B divided into ten channels. Correspondingly, therefore, signal pulses of unit duration over system A, are of 10 units duration in each channel of system B.

Random noise having uniform spectral distribution, and power W in each channel of system B, cumulates to power $10W$ in system A. Signal power P in each channel of system B cumulates to $10P$ for the total system. If signal power $10P$ is used in system A, the two systems are at a stand-off in signal-to-noise ratio for this type of noise.

In practice, the power capacity to handle the signals for system B must be made a few db higher than indicated by $10P$ to allow for occasional peaking caused during instants of unfavorable phasing among the various channels. Thus, the multichannel system B is really worse off, by that small amount, than system A. This may amount to some 3 or 4 db in ten or twelve channels.

There are occasions where crosstalk into other facilities sets the level permitted for the signal. It may be that concentration of the signal onto a single carrier aggravates this interference, as compared with that from a multi-carrier signal. In such cases system A may be penalized by a few db, as compared with system B.

Impulse noise shows correlation among the phases of its spectral components. Thus a noise pulse of voltage amplitude N in each channel of system B, cumulates to voltage amplitude $10N$ in system A or 20 db greater. On the other hand, a signal of amplitude S in each channel of system B, cumulates to an RMS voltage amplitude $\sqrt{10}\, S$, or 10 db greater, over the total system (plus a peaking correction which may be positive or negative as just mentioned). Thus, the single channel A system is at a disadvantage of 10 db, less the peaking adjustment, with respect to the B system.

A further adjustment may be needed, because a single noise peak that affects all 10 channels of B, or 10 bits, may affect only one bit of A. This adjustment depends upon the word grouping of bits which is used. It may be neglected if all the 10 bits of B are in the same word, and if, in one word, an error of 10 bits is effectively no worse than an error of 1 bit.

Single-frequency noise lies at the opposite extreme of the gamut from impulse noise. The vulnerability of a signal pulse to single-frequency noise varies according to the relationship between the frequencies of the noise and of the signal carrier in the utilized signal band.

The pattern of sensitivity to noise over an individual channel can be expected to be about the same for a narrow band as for a wide-band channel. Thus the pattern of sensitivities in the single channel of system A is repeated in each channel of system B on a 10 times finer frequency scale. The required S/N ratio in any one channel of B remains the same as that for A.

In system B each of the ten channels must put out only one tenth of the power of the single channel of system A (less the correction for peaking which as before may be positive or negative). Thus any one of the ten channels of B is 10 db (plus a peaking correction) more vulnerable to single frequency noise than the single channel of A.

It must be noted that there are occasional special circumstances where the single frequency noise may be persistent and steady. The multichannel system B may in such cases have an advantage in permitting the one channel affected to be dropped, and the others to be worked entirely free from this interference. This of course reduces the total bit rate.

To summarize the discussion in a general philosophical way, it can be said that there is advantage in multiplexing the signal in the manner that makes it as different as possible from the type of noise to which it is expected to be the most exposed. If the predominant noise is in short duration pulses, the most advantageous signal is in long duration pulses with frequency discrimination multiplex. If the noise is in longer duration single frequencies, the most advantageous signal is in very short pulses with time discrimination multiplex.

VI. REFERENCES

1. J. V. Harrington, P. Rosen, D A. Spaeth, Some Results on the Transmission of Pulses Over Lines, *Symposium on Information Networks*, **III,** pp. 115–130, Polythechnic Institute of Brooklyn, April, 1954.
2. F. W. Reynolds, A New Telephotograph System, B.S.T.J., **15,** pp. 549–475, October, 1936.
3. A. W. Horton and H. E. Vaughan, Transmission of Digital Information Over Telephone Circuits, B.S.T.J., **34,** pp. 511–528, May, 1955.
4. C. A. Lovell, J. H. McGuigan and O. J. Murphy, An Experimental Polytonic Signaling System, B.S.T.J., **34,** pp. 783–806, July, 1955.
5. A. L. Matte, Advances in Carrier Telegraph Transmission, B.S.T.J., **19,** pp. 161–208, April, 1940; J. R. Davey and A. L. Matte, Frequency Shift Telegraphy — Radio and Wire Applications, B.S.T.J., **27,** pp. 265–304, April 1948.
6. R. S. Caruthers, The Type N1 Carrier Telephone System; Objectives and Transmission Features, B.S.T.J., **30,** pp. 1–32, January, 1951.
7. P. Mertz, Transmission Line Characteristics and Effects on Pulse Transmission, *Symposium on Information Networks*, **III,** pp. 115–130, Polytechnic Institute of Brooklyn, April, 1954.
8. S. I. Cory, Telegraph Transmission Coefficients, Bell Laboratories Record, **33,** pp. 11–15, January, 1955.
9. T. A. Jones and K. W. Pfleger, Performance Characteristics of Various Carrier Telegraph Methods, B.S.J.T., **25,** pp. 483–531, July, 1946.
10. C. A. Dahlbom, A. W. Horton, and D. L. Moody, Application of Multifrequency Pulsing in Switching, Trans. A.I.E.E., **68,** pp. 392–396, 1949.
11. E. D. Sunde, Theoretical Fundamentals on Pulse Transmission, B.S.T.J., **33,** pp. 721–788 and 987–1010, May and July, 1954.
12. I. E. Lattimer, The Use of Telephone Circuits for Picture and Facismile Service, Long Lines Department, 1948.
13. H. Nyquist, Certain Topics in Telegraph Transmission Theory, Trans. A.I.E.E., **47,** pp. 617–644, April, 1928.
14. B. M. Oliver, J. R. Pierce and C. E. Shannon, The Philosophy of PCM, Proc. I.R.E., **36,** pp. 1324–1331, November, 1948.
15. R. E. Crane, J. T. Dixon, and G. H. Huber, Frequency Division Techniques for a Coaxial Cable Network, Trans. A.I.E.E., **66,** pp. 1451–1459, 1947.
16. J. T. O'Leary, E. C. Blessing and J. W. Beyer, An Improved Three-Channel Carrier Telephone System, B.S.T.J., **17,** pp. 162–183, January, 1938.
17. C. W. Carter, U. S. Patent No. 2,390,869.

Design, Performance and Application of the Vernier Resolver*

By G. KRONACHER

(Manuscript received May 29, 1957)

The Vernier Resolver is a precision angle transducer which, from the stand-point of performance, resembles a geared up synchro resolver, except that the step-up ratio between the mechanical angle and the electrical signal is obtained electrically.

Vernier resolvers with step up ratios of 26, 27, 32 and 33 have been designed and built.

The unit is a reluctance type, variable coupling transformer. By placing all windings on the stator, sliding contacts are eliminated. Both the stator and the rotor are laminated. Because of the averaging effect inherent in a laminated construction, the accuracy of the unit exceeds by many times the machining accuracy.

The performance of present experimental units is characterized by a repeatability of better than ±3 seconds of shaft angle, and a standard deviation error over one full revolution of less than 10 seconds of arc.

I. INTRODUCTION

The precise measurement of an angle is a basic operation in many technical fields. The observation of stars, mapping of land, machining in the factory are all operations which require angle measurements. Of course, an angle can be measured by reading a calibrated dial. However, in automatically controlled operations the angular position of a shaft has to be sensed electrically. The instrument which performs the conversion from a mechanical angle to an electrical output is called an angle transducer. One commonly used angle transducer is the synchro resolver.

Basically this is a variable coupling transformer with one primary winding and two output windings displaced 90 degrees from each other. The variable electrical coupling is accomplished by placing the primary

* The Vernier Resolver was developed under the sponsorship of the Wright Air Development Center.

winding on the rotor and the secondary windings on the stator or vice versa. The primary winding is excited from an alternating voltage source of, say, 400 cycles per second. The amplitudes of the induced secondary voltages of the synchro resolver are ideally proportional to the sine and cosine of the rotor orientation. These two induced, amplitude modulated voltages are the resolver output.

The accuracy of commercially available synchros is, at best, three minutes of arc. Certainly, this accuracy is sufficient for many applications. In the machining of precision parts and in field applications involving the measurement of elevation and azimuth of distant targets, however, accuracies down to 10 seconds of arc are required.

One might be tempted to try to meet this requirement by merely refining the present standard synchro. However, even if this refinement were possible, it still would be a difficult task to transmit this near-perfect synchro output and also to convert it into other analog forms without losing most of the added accuracy because of noise in the system. The transmission and conversion problem can be side-stepped by going to a so-called "two speed" or "vernier" representation of the angle. This representation is obtained by using two synchros; one, the low speed synchro, is positioned directly to the particular angle and the other, the high-speed synchro, is geared up with respect to the former. The angle is now represented by two synchro outputs. Assuming perfect gears the accuracy of this system is improved by the step-up ratio in the gearing.

This approach has been adopted in the past, but unfortunately, it has major disadvantages to it. First, precision gears of better than one minute of arc are expensive, relatively large and of limited life due to wear. Second, considerable torque is required to overcome the gear friction and the inertia effect of the high speed synchro. For these reasons, it is desirable to replace the geared-up synchro by a transducer which performs the step-up between input and output electrically. The vernier resolver is such an angle transducer.

The unit is a reluctance type, variable coupling transformer. By placing all windings on the stator, sliding contacts are eliminated. Both the stator and the rotor are built up of laminations. The step-up ratio is equal to the number of teeth on the rotor lamination. Prototype units have been built with step-up ratios of 26, 27, 32 and 33. The accuracy of these units is characterized by a standard deviation error of less than 10 seconds of arc. This high degree of accuracy is due largely to the averaging effect inherent in a laminated construction. The unit may be regarded, simply, as a device which senses the average orientation of all rotor laminations with respect to the stator. Because of the great number of laminations (one hundred in the present units) the effect of individual imperfections in laminations is greatly reduced.

In preparation for a close study of the vernier resolver we shall describe the performance of an ideal unit, and also introduce some technical terms. The output of the vernier resolver consists of two amplitude modulated voltages one of which is called the sine-voltage and the other the cosine voltage. The amplitudes of these voltages are proportional to the sine and cosine of "n"-times the rotor orientation. The factor "n" which, of course, is a function of the rotor configuration will be called the order of the resolver. We shall call the arctangent of the ratio of the secondary voltages — sine-voltage over cosine-voltage — the "signal-angle". Furthermore, to define a positive sense of rotation and to make the signal-angle definition unambiguous, we shall assume that, with continuous positive shaft rotation, the signal-angle runs through a sequence of cycles, each going from zero to 360°. Thus, one signal-angle cycle corresponds to a shaft rotation of $(1/n)$th; of one revolution. This angular interval is called the "vernier" interval.

II. DESIGN PRINCIPLES

2.1 *A Simplified Description*

Fig. 1 represents a simplified model of a third order vernier resolver. The unit consists of a laminated rotor with three equally spaced teeth and a laminated 4-pole-shoe stator. Each pole-shoe bears one exciting coil (not shown in the figure) and one output coil. Successive exciting coils are wound in opposite directions, connected in series, and energized from an ac source. Thus, successive pole-shoe fields alternate in phase. The two output windings each consist of two diametrical output coils connected in phase opposition.

If the rotor were a circular cylinder, the net voltage in either output winding would be zero. However, because of the three rotor teeth, the induced voltage of either output winding goes through three identical cycles per rotor revolution. Consequently, the amplitude E_c of the induced cosine-voltage e_c can be represented as a Fourier series of three times the shaft angle θ_m,

$$E_c = E_1 \cos (3\theta_m) + E_3 \cos [3(3\theta_m)] + \cdots , \tag{1}$$

where E_1, E_3 are the Fourier components of E_c with respect to $(3\theta_m)$.

The series is free of even harmonic terms because of the symmetry between positive and negative half-cycles. The expression for the amplitude E_s of the sine-voltage e_s is obtained by substituting $[\theta_m - (\pi/6)]$

for θ_m in (1);

$$E_s = E_1 \sin (3\theta_m) - E_3 \sin [3(3\theta_m)] + \cdots . \tag{2}$$

The magnitudes of the Fourier components depend on the design details of the unit. In a properly designed unit all higher order Fourier components are sufficiently small so that the induced signal voltages closely approximate those of an ideal third order resolver as expressed by the following equations:

$$E_c = E_1 \cos (3\theta_m), \tag{3}$$

$$E_s = E_1 \sin (3\theta_m), \tag{4}$$

$$\theta_S \equiv \tan^{-1} \frac{E_s}{E_c} = 3\theta_m , \tag{5}$$

where θ_S is the signal-angle.

2.2 *Analysis of Practical Case*

Fig. 2 shows an assembled unit, and typical stator and rotor laminations are illustrated in Figs. 3 and 4.

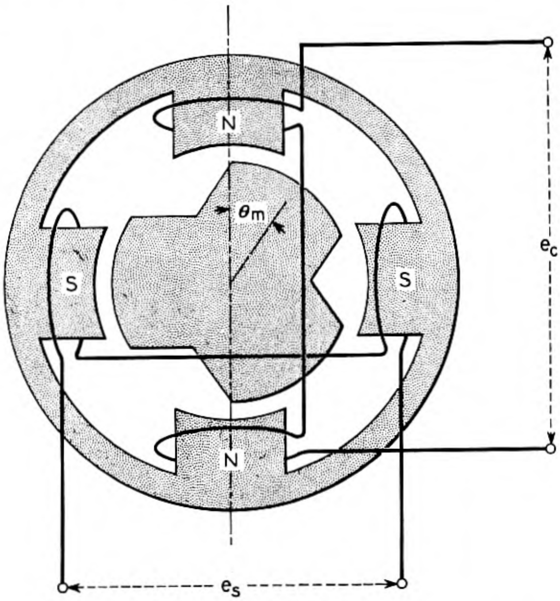


Fig. 1 — Schematic of a 3rd order vernier resolver.

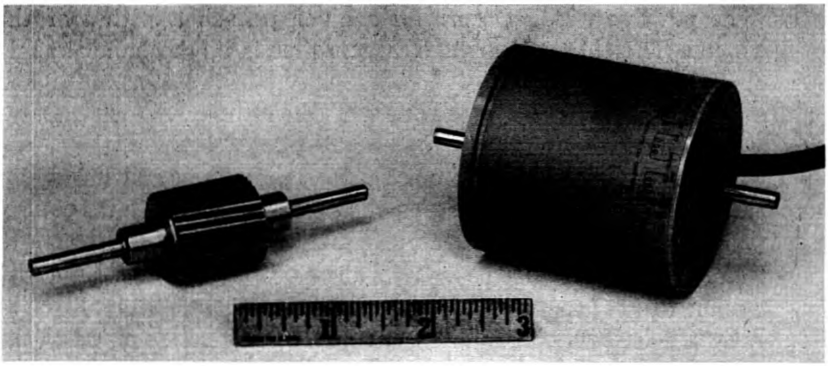


Fig. 2 — View of the assembled vernier resolver and its rotor.

The stator lamination is of ten pole-shoes, each pole-face having three teeth. All 30 teeth of the stator are equally spaced. The number of equally spaced teeth of the rotor lamination is equal to the order "n" of the resultant vernier resolver.

The exciting winding, as in the simplified model, produces ac magnetic fields alternating in phase from one pole-shoe to the next. Each of the two output windings is distributed over all ten pole-shoes.

To describe the turns distribution of these windings it is necessary to define the positive winding sense and the electrical angle of a pole-

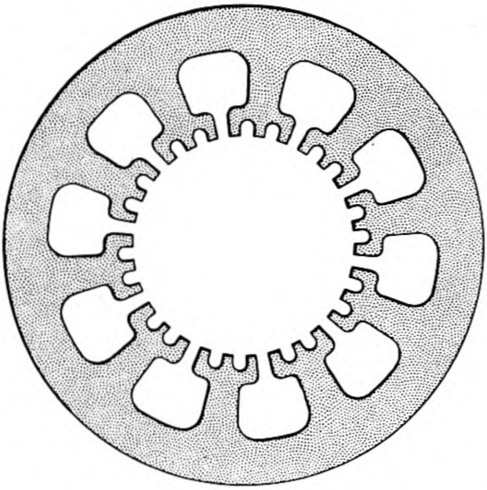


Fig. 3 — Stator lamination of the vernier resolver.

shoe. The positive winding sense for a given pole-shoe is that of the exciting coil. The electrical angle of a pole-shoe is its mechanical angle measured clockwise, with respect to a reference on the stator, multiplied by the number of rotor teeth (the order of this resolver).

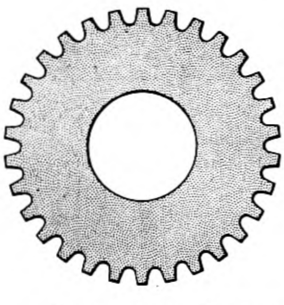
The winding which produces the cosine voltage, e_c, consists of ten coils, one on each pole-shoe, connected in series. Each coil has a different number of turns depending on the pole-shoe to which it belongs. Specifically, this number of turns is equal to a design constant "t" multiplied by the cosine of the electrical angle of the pole-shoe. Similarly, the winding which produces the sine-voltage, e_s, consists of ten coils, one on each pole-shoe. The number of turns of each coil is equal to the same constant "t" multiplied by the sine of the electrical angle of the particular pole-shoe.

With α_e being the electrical angle between adjoining pole-shoes and with the electrical angle of pole-shoe No. 0 equalling zero, the turns of the coils of the cosine-winding on pole-shoes No. 0 through 9 are:

$$\begin{aligned} t_{C0} &= t \cos (0) \\ t_{C1} &= t \cos (\alpha_e) \\ \vdots\ &\quad \vdots\ \vdots \\ t_{C9} &= t \cos (9\alpha_e) \ . \end{aligned} \tag{6}$$

Similarly the turns distribution, t_s, of the sine winding is:

$$\begin{aligned} t_{s0} &= \mathrm{t} \sin (0) \\ t_{s1} &= t \sin (\alpha_e) \\ \vdots\ &\quad \vdots\ \vdots\ \vdots \\ t_{s9} &= t \sin (9\alpha_e) \end{aligned} \tag{7}$$



ORDER OF RESOLVER	NUMBER OF ROTOR TEETH
26	26
27	27
32	32
33	33

Fig. 4 — Rotor lamination of the vernier resolver.

To obtain the voltage induced in the output coils, the flux amplitude for each pole-shoe must be established. Defining the electrical angle of the rotor, θ_e, as its mechanical angle multiplied by its number of teeth and choosing θ_e to be zero when the center of a rotor tooth lines up with the center of pole-shoe No. 0, one can write for the flux amplitudes ϕ_0 through ϕ_9 of pole-shoes No. 0 through No. 9:

$$\begin{aligned} \phi_0 &= A_0 + A_1 \cos \theta_e + A_2 \cos 2\theta_e + \cdots \\ \phi_1 &= A_0 + A_1 \cos (\theta_e - \alpha_e) + \cdots \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \phi_9 &= A_0 + A_1 \cos (\theta_e - 9\alpha_e) + \cdots \end{aligned} \tag{8}$$

where $A_0, A_1, A_2, \ldots$ are the Fourier Components of ϕ.

The amplitude, E_c, of the voltage induced in the cosine winding is the sum of the products of the pole-shoe flux, ϕ_ν, measured in [volt sec] and the coil turns, $t_{c\nu}$, multiplied by the exciting current frequency in radians per second, ω:

$$E_c = \omega \sum_{\nu=0}^{9} \phi_\nu t_{c\nu} \,. \tag{9}$$

Substituting the values of $t_{c\nu}$ and ϕ_ν from (6) and (7) and neglecting all higher order Fourier components of ϕ, one obtains

$$E_c = \frac{10}{2} \omega A_1 \cos \theta_e \,. \tag{10}$$

Similarly one obtains for the amplitude E_s of the sine voltage:

$$E_s = \frac{10}{2} \omega A_1 \sin \theta_e \,. \tag{11}$$

As required, the two induced secondary voltages are proportional to the cosine and sine of the electrical rotor angle.

As shown in the appendix, an analysis which takes into account the higher order Fourier components of the pole-shoe flux shows that a sinusoidally distributed winding is sensitive solely to the so-called slot-harmonics. The order "m" of these harmonics is given by the expression

$$m = kq \pm 1 \tag{12}$$

where "k" is any integral positive number and q is the number of pole-shoes divided by the largest integral factor common to the number of pole-shoes and the number of rotor teeth. For instance, in the case of a

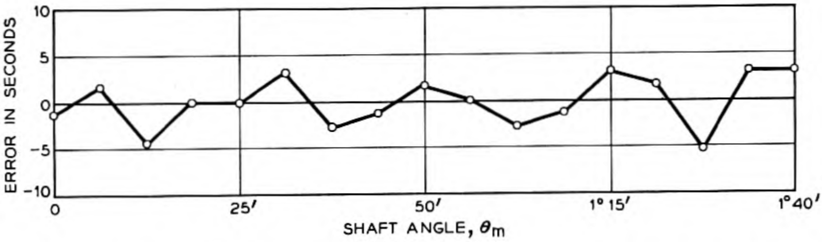


Fig. 5 — Error of a 27th order vernier resolver over $\frac{1}{8}$ vernier interval.

27th order vernier resolver this common factor is 1 and consequently the slot harmonics are of order: 9, 11, 19, 21, etc.

The effect of the slot harmonics can be reduced by the following means:

a) Selecting the dimensions as well as the number of rotor and stator teeth such as to keep the higher order flux components low.

b) Using a "skewed" rotor or stator, in which successive laminations are progressively displaced with respect to their angular orientation.

III. PERFORMANCE

Clifton Precision Products Co. built experimental resolver models of order 26, 27, 32 and 33 using the laminations shown in Figs. 3 and 4. The best results were obtained with 27th order resolvers. Their performance is described in the following sections.

3.1 *Repeatability and Accuracy*

The repeatability is better than ±3 seconds of shaft angle.

Figs. 5, 6 and 7 show the error curves taken on a 27th order vernier resolver after compensating with trimming resistors for the fundamental and second harmonic error with respect to the vernier interval. (In essence, the effect of these trimming resistors is either to add or to subtract a small voltage to one or both of the resolver signals.) Fig. 8 shows an error curve before trimming.*

3.2 *Temperature Sensitivity*

The error introduced by a temperature change of 70°C is less than 25 seconds of shaft rotation.

* The error curves really represent the combined error of the tested resolver itself plus that of the testing apparatus.

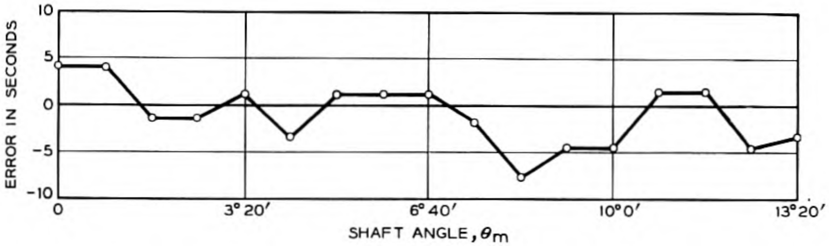


Fig. 6 — Error of a 27th order vernier resolver over one vernier interval.

It may be pointed out that the housing of the tested unit was of aluminum. A unit with a non-magnetic steel housing should be of lower temperature drift, because stator stack and housing would then have the same temperature coefficient of expansion.

3.3 *Transformation Ratio, Input and Output Impedances*

At maximum coupling the induced output voltage is 0.123 times the exciting voltage and is leading in phase by 6°.

The impedance of the input winding with the output windings open is $117 + j\,781$ ohms.

The impedance of the output windings with the primary winding shorted is $235 + j\,920$ ohms.

The effect of the rotor position on this impedance is hardly noticeable.

3.4 *Output Signal Distortions*

The harmonic content of the output signal at maximum coupling is:

fundamental 1.7 volts
2nd harmonic 0.2 mv
3rd harmonic 13.5 mv
5th harmonic 5.4 mv

The harmonic content of the output signal at minimum coupling (null voltage) is:

fundamental 1.6 mv
2nd harmonic 0.05 mv
3rd harmonic 2.0 mv

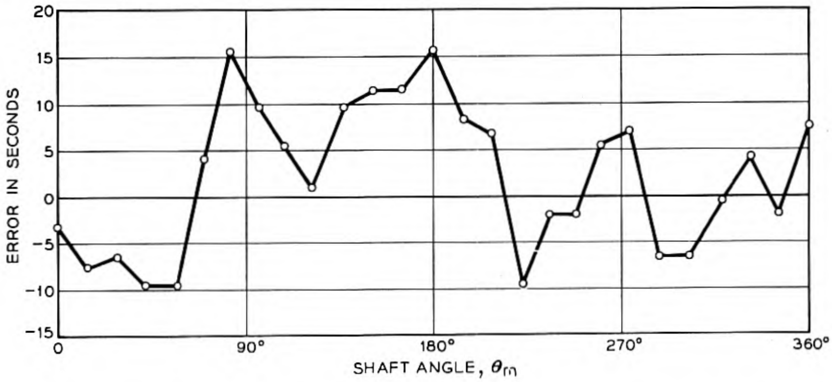


Fig. 7 — Error of a 27th order vernier resolver over one shaft revolution.

Fig. 8 — Error of a 27th order vernier resolver before trimming.

3.5 *Moment of Inertia and Friction Torque*

The moment of inertia of the rotor of a 27th order harmonic resolver is 63 gram cm sq.

The maximum break-away friction torque among five units was 0.027 in. oz. No change in this torque due to excitation of the unit could be detected.

IV. APPLICATION

In its application, the vernier resolver is usually directly coupled to a standard resolver or some other coarse angle transducer. Such a system which represents a variable, in this case the shaft angle, in two scales, coarse and fine, will be called a vernier system.

The following sections describe applications using the vernier resolver in an encoder, a follow-up system and an angle-reading system.

4.1 *Vernier Angle Encoder*

A vernier angle encoder converts a shaft angle into a pair of digital numbers, one being the coarse and the other being the vernier number. This type of encoder can be built by mechanically coupling a standard resolver directly to a vernier resolver. The outputs of the two resolvers, after encoding, represent the coarse and the vernier number.

The output of a resolver may be encoded, for instance, by the following method. The primary winding of the resolver is excited from an a-c source of, say, 400 cycles per second. The two induced secondary voltages are in phase with each other. Their amplitudes are proportional to the cosine and sine of the electrical rotor angle, θ_e .

These two amplitude modulated voltages are combined by means of two phase-shifting networks into two phase-modulated voltages. One network first advances the sine voltage by 90° and then adds it to the cosine voltage. The other network performs the same addition after retarding the sine voltage by 90°. The result is two constant amplitude voltages with relative phase shift of twice the electrical rotor angle. The time interval between the respective zero crossings of these two voltages is con-

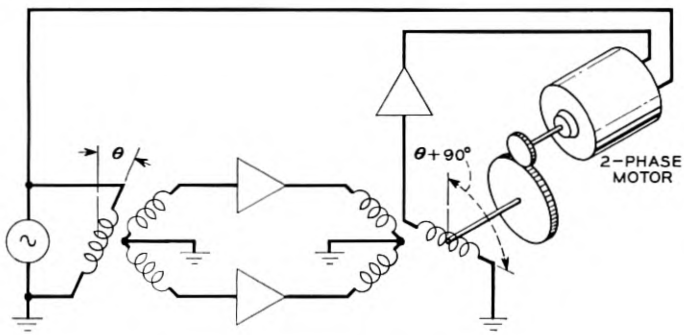


Fig. 9 — Resolver servo system.

verted into digital representation by means of an electronic stop-watch (time encoder).

4.2 *Vernier Follow-up System*

A vernier follow-up system can be built much like the present two speed synchro control-transformer system, except that the geared up synchros are replaced by vernier resolvers. Fig. 9 illustrates the vernier portion of this system. Since the output impedance of the vernier resolver is fairly large, it may be desirable to use amplifiers, as shown in Fig. 9, to energize that vernier resolver which plays the part of the control-transformer.

4.3 *Vernier Angle-Reading System*

A visual vernier angle-reading system as required to read the position of a rotary table can be built by using the output of a vernier resolver to position a standard resolver.

The coarse angle can be read as usual from calibration lines marked directly on the rotary table. The vernier angle reading is obtained by coupling a vernier resolver directly to the rotary table. The output of this resolver is used to position a standard control transformer. This control transformer will go through "n" revolutions for each revolution of the rotary table, where "n" is the order of the vernier resolver. The reading of a dial coupled either directly or through gears to the control transformer provides the vernier reading.

V. SUMMARY

Vernier resolvers of order 26, 27, 32 and 33 have been designed, built and tested.

The construction of the unit is very simple because all windings are located on the stator. The absence of brushes and slip rings makes the unit inexpensive in production and reliable in performance.

The performance of present experimental models is characterized by a repeatability of better than ± 3 seconds of arc and by a standard deviation error over one revolution of less than 10 seconds of arc.

Production units should be of even higher accuracy because better tooling fixtures would be used and minor design improvements would be incorporated.

The principal forseeable application of the resolver lies in vernier systems. Vernier encoder, vernier servo and vernier angle-reading systems are readily obtained by applying existing techniques to the vernier resolver.

ACKNOWLEDGEMENT

The development of the vernier resolver was undertaken under the sponsorship of the Wright Air Development Center. The work was encouraged and furthered by J. C. Lozier of Bell Telephone Laboratories. Valuable design contributions were made by J. Glass of Clifton Precision Products Co. All testing and evaluating of test results was done by T. W. Wakai of Bell Telephone Laboratories.

APPENDIX

Symbols

- E Amplitude of induced voltage
- p Number of pole-shoes
- n Number of rotor teeth
- θ_e Electrical rotor angle, equal to its geometrical angular position multiplied by n
- q p divided by largest integral factor common to n and p
- n' n divided by the same factor
- ν Pole-shoe number running from 0 to $(p - 1)$
- m Order of Fourier component representing the pole-shoe flux as a function of the electrical rotor angle, θ_e
- α_e Electrical angle between adjoining pole-shoes, equal to the geometrical angle multiplied by n
- k A number equal to zero or to any positive integer.

In accordance with equations (7) and (8) the voltage induced in the sine-winding coil on the νth pole-shoe by the mth flux harmonic is:

$$E_{sm\nu} = \omega[A_m \cos m\,(\theta_e - \nu\alpha_e)\,t \sin\,(\nu\alpha_e)]. \tag{13}$$

After trigonometric transformation:

$$E_{sm\nu} = \tfrac{1}{2}A_m t\omega\,[\sin\,(m\theta_e - (m-1)\nu\alpha_e) + \sin\,(-m\theta_e + (m+1)\nu\alpha_e)]\,. \tag{14}$$

The voltage, E_{sm}, induced in the sine winding is obtained by summing the expression of (14) over all values of ν. Since $(p\alpha_e)$ is a multiple of 2π the summing of all sine terms from $\nu = 0$ to $\nu = (p - 1)$ results in zero unless the angle $(m \pm 1)\,\alpha_e$ is an integral multiple, k, of 2π. This condition is spelled out in the following equation:

$$(m \pm 1)\, \alpha_e = k2\pi\,. \tag{15}$$

The electrical angle α_e being the mechanical angle between successive pole-shoes divided by the number of rotor teeth is:

$$\alpha_e = \frac{2\pi n}{p}. \tag{16}$$

Dividing p and n by the largest common integral factor, one can write

$$\alpha_e = \frac{2\pi n'}{q}. \tag{17}$$

Substituting this expression into (15) and solving for m, one obtains

$$m = \frac{kq}{n'} \pm 1 \tag{18}$$

where m and k are integers or zero. Consequently, (18) is satisfied for the following values of m:

$$m = 1; \qquad q \pm 1; \qquad 2q \pm 1; \cdots . \tag{19}$$

The amplitude of the voltage, E_{sm}, induced in the sine winding by flux harmonics of order m, where m is specified by (19), is

$$E_{sm} = \frac{p}{2} A_m t\omega \sin (m\theta_e). \tag{20}$$

Similarly one obtains for the voltages, E_{cm}, induced in the cosine winding

$$E_{cm} = \frac{p}{2} A_m t\omega \cos (m\theta_e). \tag{21}$$

Bell System Technical Papers Not Published in this Journal

ANDERSON, O. L.,[1] CHRISTENSEN, H.,[1] and ANDREATCH, P., JR.[1]

A Technique for Connecting Electrical Leads to Semiconductors, J. Appl. Phys., Letter to the Editor, **28,** p. 923, August, 1957.

ANDREATCH, P., JR., see Anderson, O. L.

BENSON, K. E., see Pfann, W. G.

Benson, K. E., see Wernick, J. H.

BIRDSALL, H. A.[1]

Insulating Films, in "Digest of Literature on Dielectrics", Publication 503, National Academy of Sciences, National Research Council, **19,** pp. 209–220, June, 1957.

BOGERT, B. P.[1]

Response of an Electrical Model of the Cochlear Partition with Different Potentials of Excitation, J. Acous. Soc. Am., **29,** pp. 789–792, July, 1957.

BOHM, D., see Huang, K.

BOYET, H., see Weisbaum, S.

BOZORTH, R. M.[1]

The Physics of Magnetic Materials, in "The Science of Engineering Materials", John Wiley & Sons, New York, pp. 302–335, 1957.

BRADY, G. W.[1]

Structure of Tellurium Oxide Glass, J. Chem. Phys., **27,** pp. 300–303, July, 1957.

[1] Bell Telephone Laboratories, Inc.

BRADY, G. W.,[1] and KRAUSE, J. T.[1]

Structure in Ionic Solutions. I, J. Chem. Phys., **27,** pp. 304–308, July, 1957.

BREIDT, P., JR., see Greiner, E. S.

CHYNOWETH, A. G.,[1] and McKAY, K. G.[1]

Internal Field Emission in Silicon p-n Junctions, Phys. Rev., **106,** pp. 418–426, May 1, 1957.

CHRISTENSEN, H., see Anderson, O. L.

DACEY, G. C.,[1] and THURMOND, C. D.[1]

p-n Junctions in Silicon and Germanium: Principles, Metallurgy, and Applications, Metallurgical Reviews, **2,** pp. 157–193, July, 1957.

ELLIS, W. C., see Greiner, E. S.

FRISCH, H. L.[1]

The Time Lag in Nucleation, J. Chem. Phys., **27,** pp. 90–94, July, 1957.

FULLER, C. S.,[1] and REISS, H.[1]

Solubility of Lithium in Silicon, J. Chem. Phys., Letter to the Editor, **27,** pp. 318–319, July, 1957.

GIBBONS, D. F.[1]

Acoustic Relaxations in Ferrite Single Crystals, J. Appl. Phys., **28,** pp. 810–814, July, 1957.

GIANOLA, U. F.[1]

Damage in Silicon Produced by Low Energy Ion Bombardment, J. Appl. Phys., **28,** pp. 868–873, Aug., 1957.

GITHENS, J. A.[1]

The TRADIC LEPRECHAUN Computer, Proc. Eastern Joint Computer Issue, A.I.E.E. Special Publication, **T-92,** pp. 29–33, Dec. 10–12, 1956.

GLASS, M. S.[1]

Straight Field Permanent Magnets of Minimum Weight for TWT — Design and Graphic Aids in Design, Proc. I.R.E., **45,** pp. 1100–1105, Aug., 1957.

[1] Bell Telephone Laboratories, Inc.

GREEN, E. I.[1]

Evaluating Scientific Personnel, Elec. Engg., **76,** pp. 578–584, July, 1957.

GREINER, E. S.,[1] BREIDT, P., JR.,[1] HOBSTETTER, J. N.,[1] and ELLIS, W. C.[1]

Effects of Compression and Annealing on the Structure and Electrical Properties of Germanium, J. Metals, **9,** pp. 813–818, July, 1957.

HAMMING, R. W., see Hopkins, I. L.

HARKER, K. J.[1]

Nonlaminar Flow of Cylindrical Electron Beams, J. Appl. Phys., **28,** pp. 645–650, June, 1957.

HERMAN, H. C.[1]

Jumbo Case Considerations, J. Patent Office Society, **39,** pp. 515–523, July, 1957.

HOBSTETTER, J. N., see Greiner, E. S.

HOPKINS, I. L.,[1] and HAMMING, R. W.[1]

On Creep and Relaxation, J. Appl. Phys., **28,** pp. 906–909, Aug. 1957.

HUANG, K.,[1] BOHM, D.,[8] and PINES, D.[9]

Role of Subsidiary Conditions in the Collective Description of Electron Interactions, Phys. Rev., **107,** pp. 71–80, July 1, 1957.

INGRAM, S. B.[1]

Graduate Study in Industry — The Communications Development Training Program of the Bell Telephone Laboratories, Eng. J., **40,** pp. 993–996, July, 1957.

JACCARINO, V.,[1] SHULMAN, R. G.,[1] and STOUT, R. W.[10]

Nuclear Magnetic Resonance in Paramagnetic Iron Group Fluorides, Phys. Rev., Letter to the Editor, **106,** pp. 602–603, May 1, 1957.

JONES, W. D., see Turrell, G. C.

[1] Bell Telephone Laboratories, Inc.
[8] Technion, Haifa, Israel.
[9] Princeton University, Princeton, N. J.
[10] University of Chicago, Chicago, Ill.

KIERNAN, W. J.[1]

Appearance Specifications and Control Methods, Elec. Manufacturing, **60,** pp. 126–129, 294, 296, 298, July, 1957.

KRAUSE, J. T., see Brady, G. W.

LANDER, J. J., see Morrison, J.

LEGG, V. E.[1]

Survey of Square Loop Magnetic Materials, Proc. Conference on Magnetic Amplifiers, A.I.E.E. Special Publication, **T-98,** pp. 69–77, Sept. 4, 1957.

LOGAN, R. A.,[1] and PETERS, A. J.[1]

Diffusion of Oxygen in Silicon, J. Appl. Phys., **28,** pp. 819–820, July, 1957, Letter to the Editor.

LUKE, C. L.[1]

Determination of Sulfur in Nickel by the Evolution Method, Anal. Chem., **29,** pp. 1227–1228, Aug. 1957.

MAKI, A., see Turrell, G. C.

MARCATILI, E. A.[1]

Heat Loss in Grooved Metallic Surface, Proc. I.R.E., **45,** pp. 1134–1139, Aug., 1957.

MERTZ, P.[1]

Information Theory Impact on Modern Communications, Elec. Engg., **76,** pp. 659–664, Aug., 1957.

McCALL, D. W.[1]

Dielectric Properties of Polythene, in "Polythene, The Technology and Uses of Ethylene Polymers", edited by A. Renfrew and P. Morgan, published for "British Plastics" by Iliffe and Sons Limited, London, 1957.

McCALL, D. W., see Slichter, W. P.

McKAY, K. G., see Chynoweth, A. G.

[1] Bell Telephone Laboratories, Inc.

MEITZLER, A. H.[1]

A Procedure for Determining the Equivalent Circuit Elements Representing Ceramic Transducers Used in Delay Lines, Proc. 1957 Electronic Components Symposium, pp. 210–219, May, 1957.

MILLER, R. C.,[1] and SMITS, F. M.[1]

Diffusion of Antimony Out of Germanium and Some Properties of the Antimony-Germanium System, Phys. Rev., **107,** pp. 65–70, July 1, 1957.

MONTGOMERY, H. C.[1]

Field Effect in Germanium at High Frequencies, Phys. Rev., **106,** pp. 441–445, May 1, 1957.

MORRISON, J.,[1] and LANDER, J. J.[1]

The Solution of Hydrogen in Nickel Under Hydrogen Ion Bombardment, Conference Notes M.I.T. Physical Electronics Conference, pp. 102–108, June, 1957.

NIELSEN, E. G.[1]

Behavior of Noise Figure in Junction Transistors, Proc. I.R.E., **45,** pp. 957–963, July, 1957.

PAYNE, R. M.[6]

Clemson Conducts School for Bell, Telephony, **152,** pp. 20–21, 50–51, June 15, 1957.

PERRY, A. D., JR.[1]

Pulse-Forming Networks Approximating Equal-Ripple Flat-Top Step Response, I.R.E. Convention Record, **2,** pp. 148–153, July, 1957.

PETERS, A. J., see Logan, R. A.

PFANN, W. G.,[1] and VOGEL, F. L., JR.[1]

Observations on the Dislocation Structure of Germanium Crystals, Acta Met., **5,** pp. 377–384, July, 1957.

PFANN, W. G.,[1] BENSON, K. E.,[1] and WERNICK, J. H.[1]

Some Aspects of Peltier Heating at Liquid-Solid Interferences in Germanium, J. Electronics, **2,** pp. 597–608, May, 1957.

[1] Bell Telephone Laboratories, Inc.
[6] Southern Bell Tel. & Tel. Co., Atlanta, Georgia.

PFANN, W. G.[1]

Zone Melting, Metallurgical Reviews, **2,** pp. 29–76, May, 1957.

PINES, D., see Huang, K.

REISS, H., see Fuller, C. S.

RUGGLES, D. M.[1]

A Miniaturized Quartz Crystal Unit for the Frequency Range 2-kc to 16-kc, Proc. 1957 Electronic Components Symposium, pp. 59–61, 1957.

SCAFF, J. H.[1]

Impurities in Semiconductors, Effect of Residual Elements on the Properties of Metals, A.S.M. Special Vol., pp. 88–132, 1957.

SHULMAN, R. G., see Jaccarino, V.

SLICHTER, W. P.,[1] and MCCALL, D. W.[1]

Note on the Degree of Crystallinity in Polymers as Found by Nuclear Magnetic Resonance, J. Poly. Sci., Letter to the Editor, **25,** pp. 230–324, July, 1957.

SMITS, F. M., see Miller, R. C.

STEELE, A. L.[7]

The 2–5 Numbering Plan and Selection of Exchange Names, Telephony, **153,** pp. 24–25, 48, July 20, 1957.

STONE, H. A., JR.[1]

Component Development for Microminiaturization, I.R.E. 1957 Convention Record, **6,** pp. 13–20, 1957.

STOUT, J. W., see Jaccarino, V.

THURMOND, C. D., see Dacey, G. C.

TURRELL, G. C.,[1] JONES, W. D.,[4] and MAKI, A.[4]

Infrared Spectra and Force Constants of Cyanoacetylene, J. Chem. Phys., **26,** pp. 1544–1548, June, 1957.

---

[1] Bell Telephone Laboratories, Inc.
[4] Oregon State College, Corvallis.
[7] Indiana Bell Telephone Company, Indianapolis.

VAN BERGEIJK, W. A.[1]

The Lung Volume of Amphibian Tadpoles, Science, **126,** p. 120, July 19, 1957.

VOGEL, F. L., JR., see Pfann, W. G.

WEINREICH, G.[1]

Ultrasonic Attenuation by Free Carriers in Germanium. Phys. Rev., Letter to the Editor, **107,** pp. 317–318, July 1957.

WEINREICH, G.,[1] and WHITE, H. G.[1]

Observation of the Acoustoelectric Effect, Phys. Rev., Letter to the Editor, **106,** pp. 1104–1106, June 1, 1957.

WEISBAUM, S.,[1] and BOYET, H.[1]

Field Displacement Isolators at 4-, 6-, 11- and 24-Kmc, Trans. I.R.E. PGMTT, **MTT-5,** pp. 194–198, July, 1957.

WEISS, M. T.[1]

Quantum Derivation of Energy Relations Analogous to Those for Nonlinear Reactances, Proc. I.R.E., Letter to the Editor, **45,** pp. 1012–1013, July, 1957.

WEISS, M. T.[1]

A Solid State Microwave Amplifier and Oscillator Using Ferrites, Phys. Rev., Letter to the Editor, **107,** pg. 317, July 1, 1957.

WERNICK, J. H.,[1] and BENSON, K. E.[1]

Zone Refining of Bismuth, J. Metals, **9,** p. 996, July, 1957.

WERNICK, J. H., see Pfann, W. G.

WHITE, H. G., see Weinreich, G.

[1] Bell Telephone Laboratories, Inc.

Recent Monographs of Bell System Technical Papers Not Published in This Journal*

AARON, M. R.

Use of Least Squares in System Design, Monograph 2828.

ANDREATCH, P., JR. and THURSTON, R. N.

Disk-Loaded Torsional Wave Delay Line, Monograph 2827.

BOND, W. L., see McSkimin, H. J.

BOOTHBY, O. L., see Williams, H. J.

BOYET, H. see Seidel, H.

BOYET, H. and SEIDEL, H.

Analysis of Nonreciprocal Effects in N-wire, Ferrite-Loaded Transmission Line, Monograph 2829.

BUCK, T. M. and McKIM, F. S.

Depth of Surface Damage Due to Abrasion on Germanium, Monograph 2805.

BURKE, P. J.

Output of a Queuing System, Monograph 2766.

BURNS, F. P.

Piezoresistive Semiconductor Microphone, Monograph 2830.

CIOFFI, P. P.

Rectilinearity of Electron-Beam Focusing Fields from Transverse Determinations, Monograph 2844.

DAVID, E. E., JR.

Signal Theory in Speech Transmission, Monograph 2831.

* Copies of these monographs may be obtained on request to the Publication Department, Bell Telephone Laboratories, Inc., 463 West Street, New York 14, N. Y. The numbers of the monographs should be given in all requests.

Seidel, H. and Boyet, H.

Polder Tensor for Single-Crystal Ferrite with Small Cubic Symmetry Anisotropy, Monograph 2843.

Sherwood, R. C., see Williams, H. J.

Shockley, W.

Transistor Physics, Monograph 2836.

Slichter, W. P.

Nuclear Magnetic Resonance in Some Fluorine Derivatives of Polyethylene, Monograph 2819.

Swanekamp, F. W., see Van Uitert, L. G.

Thurston, R. N., see Andreatch, P., Jr.

Van Uitert, L. G.

Magnesium-Copper-Manganese-Aluminum Ferrites for Microwave Uses, Monograph 2799.

Van Uitert, L. G.

Magnetic Induction and Coercive Force Data on Members of Series $BaAl_xFe_{12-x}O_{19}$, Monograph 2801.

Van Uitert, L. G.

Effects of Annealing on Saturation Induction of Ferrites With Nickel or Copper, Monograph 2845.

Van Uitert, L. G., Read, M. H., Schnettler, F. J., and Swanekamp, F. W.

Permanent Magnet Oxides Containing Divalent Metal Ions, Monograph 2841.

Walker, L. R.

Magnetostatic Modes in Ferromagnetic Resonance, Monograph 2800,

Wertheim, G. K.

Energy Levels in Electron-Bombarded Silicon, Monograph 2840.

Williams, H. J., Sherwood, R. C., and Boothby, O. L.

Magnetostriction and Magnetic Anisotropy of MnBi, Monograph 2838.

Willis, F. H.

Some Results with Frequency Diversity in a Microwave Radio System, Monograph 2842.

Contributors to This Issue

Andrew H. Bobeck, B.S.E.E., 1948; M.S.E.E., 1949, Purdue University; Bell Telephone Laboratories, 1949–. Since completing the Laboratories' Communications Development Training Program in 1952 Mr. Bobeck has been engaged in the design of both communications and pulse transformers and more recently in the design of solid state memory devices. Member I.R.E., Eta Kappa Nu and Tau Beta Pi.

B. C. Bellows, Jr., B.S., Cornell University, 1936; General Electric Co., 1936–39; Bell Telephone Laboratories, 1939–. From 1939 to 1941 Mr. Bellows was engaged in engineering trial installations of telephone equipment, particularly multi-channel coaxial cable equipment. During World War II he specialized in the mechanical design and engineering of airborne radars. From 1945 to 1957 he was engaged in the design of circuits and equipment for point-to-point microwave radio relay systems for telephone and television transmission. On May 1, 1957 he was named Transmission Measurement Engineer. Member Eta Kappa Nu and Phi Kappa Phi.

Charles A. Desoer, Dipl. Ing., University of Liege (Belgium), 1949; Sc.D. Massachusetts Institute of Technology, 1953; Bell Telephone Laboratories, 1953–. Since joining the Laboratories Mr. Desoer has specialized in linear and transistor network development in the Transmission Networks Development Department. Senior Member I.R.E.

R. Shiels Graham, B.S., University of Pennsylvania, 1937; Bell Telephone Laboratories, 1937–. His work has been with the design of equalizers, electrical wave filters and similar apparatus for use on long-distance coaxial cable circuits, microwave systems and both telephone and television transmission. During World War II, Mr. Graham designed circuits for electronic fire control computers for military use. He also developed methods for computing network and similar problems on a digital relay computer. He presently supervises the video and intermediate frequency network group. He is a senior member of the I.R.E., and a member of Tau Beta Pi, and Pi Mu Epsilon.

Gerald Kronacher, Dipl. Eng., Federal Institute of Technology, Zurich, Switzerland, 1937; Assistant Professor, Federal Institute of Technology, 1938; mining engineer, Bolivia, 1939–1946; General Electric Company, 1946–1948; Air Associates, Inc., 1948–1951; Arma Corporation, 1951–1953; Bell Telephone Laboratories, 1953–. Since joining the Laboratories Mr. Kronacher has been associated with the Military Systems Engineering Department studying input and output problems for digital computers. He is the author of many published technical articles.

Pierre Mertz, A. B., 1918; Ph.D., 1926, Cornell University; American Telephone and Telegraph Company, 1919–1921, 1926–1934; Bell Telephone Laboratories, 1935–. Mr. Mertz's work with the Bell System has been concerned primarily with transmission problems relating to telephotography and television. Since 1950 Mr. Mertz has acted as a consultant in the Systems Engineering Department on such projects as micro-image readers and commercial and military data transmission problems. Fellow of the I.R.E. and the Society of Motion Picture and Television Engineers; member, American Physical Society, Optical Society of America and the Inter-Society Color Council.

Doren Mitchell, B.S., Princeton University, 1925; American Telephone and Telegraph Company, 1925–1934; Bell Telephone Laboratories, 1934–. Mr. Mitchell's early work with the Bell System was concerned with field studies of transmission on long telephone circuits and radio circuits, including supervision of the initial operation of the New York to Buenos Aires radio-telephone circuit. Until 1942 Mr. Mitchell worked on voice operated devices of various kinds including compandors, echo suppressors and automatic switching devices. During World War II he participated in military projects involving transmission systems and problems of laying wire from airplanes. Since the war Mr. Mitchell has been primarily concerned with radio systems. In 1955 he was appointed a Special Systems Engineer supervising a data transmission system for the SAGE project, and planning other special services involving radio. Mr. Mitchell has been granted over seventy patents. Member I.R.E.

R. C. Prim, B.S. in E.E., University of Texas, 1941; A.M., Ph.D., Princeton University, 1949; General Electric Company, 1941–1944; Naval Ordnance Laboratory, 1944–1948; Bell Telephone Laboratories, 1949–. Since joining the Laboratories Mr. Prim has been a member of the Mathematical Research Department engaged in research and consultation in the fields of theoretical mechanics, solid state electronics, aerial warfare and activities analysis. In 1955 he was placed in charge of a sub-department concerned with Computing and Theoretical Mechanics, and is presently in charge of the Communications Fundamentals sub-department. Member American Mathematical Society, American Physical Society, Tau Beta Pi and Sigma Xi.

WERNER ULRICH, B.S., 1952; M.S., 1953; Eng. Sc.D., 1957, Columbia School of Engineering; Bell Telephone Laboratories, 1953–. Mr. Ulrich's first assignment was on the design of an input circuit for an electronic memory and control device. Subsequently he was engaged in the design of logical circuits for electronic controls. Since 1954, he has been working on automatic testing and maintenance facilities for electronic switching systems. Mr. Ulrich is a member of the I.R.E. and Tau Beta Pi.