ELECTRICAL COMMUNICATION

APRIL 1940

Volume 18, Number 4

ELECTRICAL COMMUNICATION

A Journal of Progress in the Telephone, Telegraph and Radio Art

BERNARD C. HOLDING, EDITOR

EDITORIAL BOARD

E. A. Brofos G. Deakin E. M. Deloraine P. E. Erikson F. Gill
W. Hatton R. A. Mack H. M. Pease Kenneth E. Stockton C. E. Strong

Issued Quarterly by the

International Standard Electric Corporation

67 BROAD STREET, NEW YORK, N.Y., U.S.A.

Volume XVIII April, 1940 Number 4

Situated close to Malines Cathedral, this building houses the new Malines Rotary telephone exchange, which is equipped for automatic ticketing. Subscribers of this exchange have full-automatic service to Brussels and Antwerp. All toll connections are recorded on tickets, which provide valuable traffic data and give a clear indication of the operating efficiency of the exchange.

Ultra-High Frequency Loop Antennae*

By A. ALFORD, Associate A.I.E.E., and A. G. KANDOIAN,

International Telephone Development Company, New York, N.Y.

ANTENNAE which are discussed in this paper were developed for two applications.

(1) To serve as elements in radiating systems of localizers and radio ranges used to guide aircraft;

(2) To act as receiving antennae carried by aircraft.

These antennae are in some respects similar to the low frequency loop antennae, and for this reason will be referred to as U.H.F. loops. The unique property which makes them useful in the above applications is that they radiate only horizontally polarized waves. The directional characteristics of these antennae are similar to those of a vertical dipole except that the waves emitted are horizontally instead of vertically polarized.

CALCULATION OF THE FIELD

Let us consider an antenna system composed of four straight line radiators, 1, 2, 3 and 4, which are arranged in the form of a square as shown in Fig. 1.

We shall assume that the currents which flow in radiators 1, 2, 3 and 4 are all in phase, and that their directions at any given instant are either as shown by the arrows in Fig. 1 or the reverse. The amplitudes of the currents along each radiator will be assumed to vary sinusoidally with the crest of the current at the centre of each radiator.

The electromagnetic field produced by this antenna system at a distant point P may be obtained by adding the fields produced by the individual radiators.

Since the electric and magnetic fields at a distant point are always simply related, so that one may be obtained from the other, it is necessary to calculate only one of them, for example, the magnetic field.

MAGNETIC FIELD OF ONE RADIATOR

The magnetic field due to a linear radiator

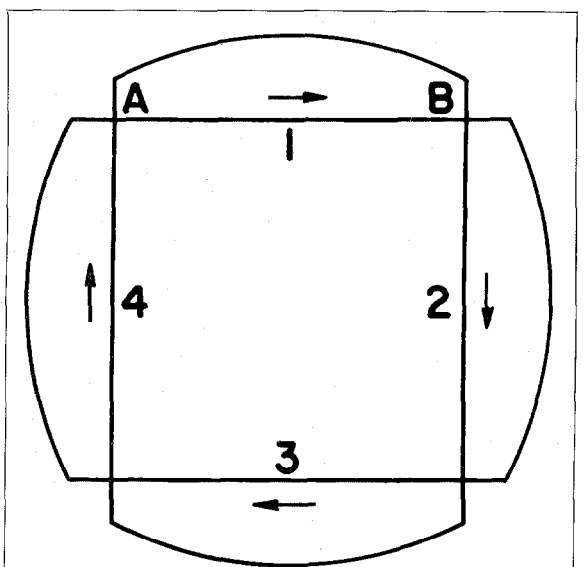


Fig. 1—Current distribution in U.H.F. loop antenna of Type 1.

with a chopped-off sine wave distribution of Fig. 1 may be readily calculated by the method described by one of the authors.† In fact, one may assume that the chopped-off sine wave distribution is the result of two travelling waves proceeding in opposite directions, and so timed that the currents are in phase at the centre of

* Paper presented at the A.I.E.E. Winter Convention, New York, N.Y., January 22nd–26th, 1940. Reprinted by permission of the A.I.E.E.

† A. Alford, *Electrical Communication*, July 1936, page 78

TABLE I

θ		0°	10°	20°	30°	40°	50°	60°	70°	80°	90°
$l/\lambda = \frac{1}{8}$	H From eq (6) ..	0.000	0.130	0.252	0.372	0.482	0.577	0.656	0.716	0.750	0.766
	H From eq (5) ..	0.000	0.137	0.270	0.393	0.506	0.602	0.681	0.739	0.774	0.786
$l/\lambda = \frac{1}{4}$	H From eq (5) ..	0.000	0.273	0.536	0.785	1.01	1.20	1.36	1.47	1.54	1.57
	H From eq (6) ..	0.000	0.221	0.436	0.654	0.857	1.039	1.194	1.313	1.386	1.414

the element. If the current in one of the two travelling waves is

$$i_i = i_0 \sin(\omega t - kz) \quad \ldots\ldots(1)$$

then the current in the other is

$$i' = i_0 \sin(\omega t + kz - lk) \quad \ldots\ldots(2)$$

The magnetic field due to the first travelling wave is

$$H = \frac{2i_0}{rc} \cot\frac{\theta_1}{2} \sin\left[\frac{\pi l}{\lambda}(1-\cos\theta_1)\right] \cos\left[\omega\left(t-\frac{r_A}{c}\right) - \frac{\pi l}{\lambda}(1-\cos\theta_1)\right] \quad \ldots\ldots(3)$$

while the magnetic field due to the second travelling wave is

$$H' = \frac{2i_0}{rc} \tan\frac{\theta_1}{2} \sin\left[\frac{\pi l}{\lambda}(1+\cos\theta_1)\right] \cos\left[\omega\left(t-\frac{r_A}{c}\right) - \frac{\pi l}{\lambda}(1-\cos\theta_1)\right] \quad \ldots\ldots(4)$$

The total magnetic field is

$$H_1 = H + H'$$

$$= \frac{2i_0}{rc}\left\{\tan\frac{\theta_1}{2} \sin\left[\frac{\pi l}{\lambda}(1+\cos\theta_1)\right] + \cot\frac{\theta_1}{2} \sin\left[\frac{\pi l}{\lambda}(1-\cos\theta_1)\right]\right\} \cos\left[\omega\left(t-\frac{r_A}{c}\right) - \frac{\pi l}{\lambda}(1-\cos\theta_1)\right] \quad \ldots\ldots(5)$$

In this equation r_A is the distance between one end A of the radiator and the distant point P; θ_1 is the angle between AP and radiator 1; l is the length of the radiator; c is the velocity of light; λ is the wave-length; and $2i_0$ is the loop current, that is, the current at the centre of the radiator.

If in place of r_A we use distance r_1 measured from the centre of radiator 1 to point P, then since

$$r_1 = r_A - l\cos\theta_1$$

equation (5) reduces to the following:

$$H_1 = \frac{2i_0}{r_1 c}\left\{\tan\frac{\theta_1}{2} \sin\left[\frac{\pi l}{\lambda}(1+\cos\theta_1)\right] + \cot\frac{\theta_1}{2} \sin\left[\frac{\pi l}{\lambda}(1-\cos\theta_1)\right]\right\} \cos\left(\omega t - kr_1 - \frac{\pi l}{\lambda}\right) \quad \ldots(6)$$

which shows that the phase of H_1 is the same at all points equidistant from the centre of the radiator.

When the length l of the radiator is small in comparison with λ, equation (5) reduces to a much simpler form, namely,

$$H_1 = \frac{2i_0}{rc}\left(\frac{2\pi l}{\lambda}\sin\theta_1\right)\cos(\omega t - r_1 k - \tfrac{1}{2}kl) \quad \ldots(7)$$

While this elementary equation is exact only when l/λ is very small, it remains a fair approximation up to $l/\lambda = \frac{1}{8}$. This may be seen from Table I, in which values of H obtained from equations (6) and (7) are given side by side for $l/\lambda = \frac{1}{8}$, and $l/\lambda = \frac{1}{4}$.

MAGNETIC FIELD OF THE LOOP

The magnetic field produced by the four radiators of the loop antenna is simply a vector sum of the individual fields. That is,

$$\bar{H} = H_1 + H_2 + H_3 + H_4 \quad \ldots\ldots(8)$$

where H_1, H_2, H_3, H_4 are the magnetic fields as given accurately by equation (6), or approximately by equation (7).

The magnetic vectors H_1 and H_3 due to radiators 1 and 3 are collinear and may be added directly. Similarly, vectors H_2 and H_4 may be added directly for the same reason. The resultant (H_1+H_3), however, in general makes an angle with the resultant (H_2+H_4) and this angle depends on the position of point P with respect to the loop; so that it is necessary either to add the two vectorially or to resolve them into mutually perpendicular components and add the latter. This second procedure is preferable and will be followed.

In any case, the first step is to add H_1 and H_3. In accordance with (7)

$$H_1 = \frac{2i_0}{rc}\left(\frac{2\pi l}{\lambda}\sin\theta_1\right)\cos\left(\omega t - r_1 k - \frac{lk}{2}\right) \quad \ldots\ldots\ldots\ldots (9)$$

$$H_3 = \frac{2i_0}{rc}\left(\frac{2\pi l}{\lambda}\sin\theta_3\right)\cos\left(\omega t - r_3 k - \frac{lk}{2}\right) \quad \ldots\ldots\ldots\ldots (10)$$

From Fig. 2 it follows, however, that $\theta_1 = \theta_3$, r_3 is not equal to r_1, but both may be readily expressed in terms of r_0, the distance from the centre of the loop to point P, and θ_2 the angle between element 2 and OP. From Fig. 2

$$r_1 = r_0 - l/2\cos\theta_2$$
$$r_3 = r_0 + l/2\cos\theta_2\,;$$

by substituting these values of r_1 and r_3 into (9) and (10), then adding H_1 and H_3 and finally simplifying the result of addition with the aid of the well-known formula

$$\cos A - \cos B = -2\sin\frac{A-B}{2}\sin\frac{A+B}{2}$$

we obtain

$$(H_1+H_3) = \frac{4i_0}{cr}\left(\frac{2\pi l}{\lambda}\sin\theta_1\right)\sin\left(\frac{\pi l}{\lambda}\cos\theta_2\right)\sin\left(\omega t - kr_0 - \frac{lk}{2}\right) \quad \ldots\ldots\ldots\ldots (11)$$

Similarly, since

$$(H_2+H_4) = \frac{4i_0}{cr}\left(\frac{2\pi l}{\lambda}\sin\theta_2\right)\sin\left(\frac{\pi l}{\lambda}\cos\theta_1\right)\sin\left(\omega t - kr_0 - \frac{lk}{2}\right) \quad \ldots\ldots\ldots\ldots (12)$$

If, in place of approximate equation (7), we had used the exact equation (6), the results would have been the same as (11) and (12), except that in place of $\left(\frac{2\pi l}{\lambda}\sin\theta_1\right)$, we would have

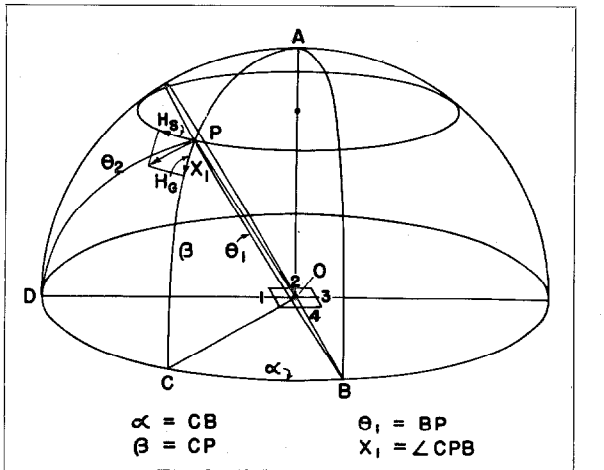


Fig. 2—Schematic diagram.

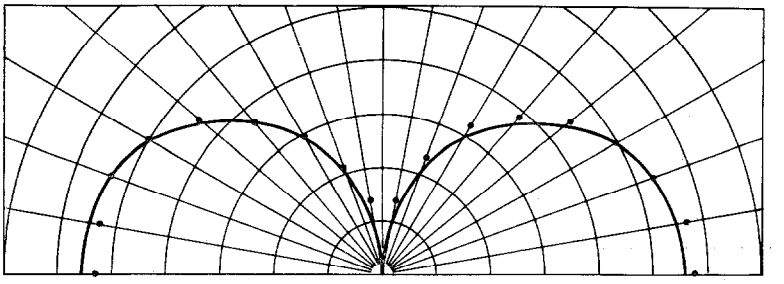


Fig. 2A (Left)—Vertical characteristic of U.H.F. loop antenna in plane α = 0° (at right angles to elements 1 and 3). Solid curve calculated. Points from measurements.

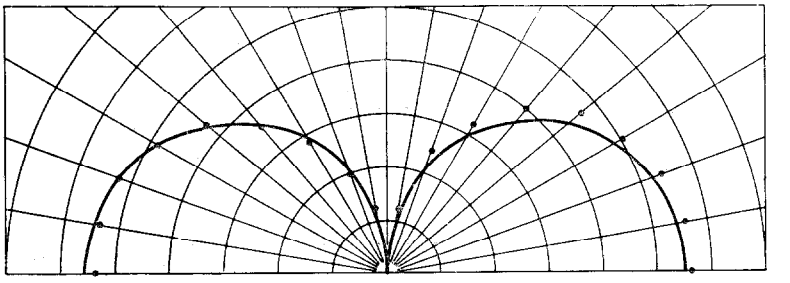


Fig. 2B (Right)—Vertical characteristic of U.H.F. loop antenna in plane α = 45° (through diagonal). Solid curve calculated. Points from measurements.

$$\left\{\tan\frac{\theta_1}{2}\sin\left[\frac{\pi l}{\lambda}(1+\cos\theta_1)\right] + \cot\frac{\theta_1}{2}\sin\left[\frac{\pi l}{\lambda}(1-\cos\theta_1)\right]\right\}$$

and a similar expression in place of

$$\left(\frac{2\pi l}{\lambda}\sin\theta_2\right).$$

The resultant vector (H_1+H_3) is at right angles to the plane defined by OP and OB. This vector may be resolved into two mutually perpendicular components, namely H_3 directed along the great circle APC and H_s at right angles to this great circle, that is, along the small circle PSQ through P. Both cosine and sine of angle Y_1, which vector (H_1+H_3) makes with H_g may be readily found from Fig. 2. In fact, since (H_1+H_3) is perpendicular to plane OPB it follows that $\angle CPB = X_1 = 90° - Y_1$. From spherical $\triangle CPB$ by the rule of sines

$$\frac{\sin X_1}{\sin\alpha} = \frac{\sin 90°}{\sin\theta_1}.$$

Hence

$$\sin(90° - Y_1) = \cos Y_1 = \frac{\sin\alpha}{\sin\theta_1}.$$

From the same triangle it follows that $\cos\theta_1 = \cos\beta \cdot \cos\alpha$.

With the aid of this relation $\sin Y_1$ may be expressed as follows:

$$\sin Y_1 = \sqrt{1-\cos^2 Y_1} = \frac{\sqrt{\sin^2\theta_1 - \sin^2\alpha}}{\sin\theta_1} = \frac{\sqrt{1-\cos^2\beta\sin^2\alpha-(1-\cos^2\alpha)}}{\sin\theta_1} = \frac{\cos\alpha\sin\beta}{\sin\theta_1} \quad \ldots\ldots\ldots\ldots (13)$$

Therefore,

$$H_{g_{1,3}} = (H_1+H_3)\cos Y_1 = (H_1+H_3)\frac{\sin\alpha}{\sin\theta_1} \quad \ldots\ldots\ldots\ldots (14)$$

and

$$H_{s_{1,3}} = (H_1+H_3)\sin Y_1 = (H_1+H_3)\frac{\cos\alpha\cdot\sin\beta}{\sin\theta_1} \quad \ldots\ldots\ldots\ldots (15)$$

Vector (H_2+H_4) is perpendicular to the plane OPD which is defined by OP and the line OD drawn from O, parallel to elements 2 and 4. Angle Y_2 which vector (H_2+H_4) makes with Meridian APC is equal to $90° - X_2$ where $X_2 = \angle CPD$. From right spherical $\triangle CPD$ it follows that

$$\frac{\sin X_2}{\sin(90°-\alpha)} = \frac{\sin 90°}{\sin\theta_2} \text{ and } \cos\theta_2 = \cos(90°-\alpha)\cos\beta.$$

Hence

$$\cos Y_2 = \sin X_2 = \frac{\cos\alpha}{\sin\theta_2}\ ;\ \sin Y_2 = \frac{\sqrt{\sin^2\theta_2-\cos^2\alpha}}{\sin\theta_2} = \frac{\sin\alpha\cdot\sin\beta}{\sin\theta_2} \quad \ldots\ldots (16)$$

Thus, the great circle component of (H_2+H_3)

$$H_{g_{2,4}} = \left(H_2+H_4\right)\frac{\cos\alpha}{\sin\theta_2} \quad \ldots\ldots\ldots (17)$$

and the small circle component is

$$H_{s_{2,4}} = -\left(H_2+H_4\right)\frac{\sin\alpha\cdot\sin\beta}{\sin\theta_2} \quad \ldots\ldots (18)$$

The negative sign used in the last equation is used to indicate that the small circle component of (H_2+H_4) points to the right when the small circle component of (H_1+H_3) points to the left. Thus, if we call one of them positive, the other has to be negative.

The total great circle component due to the antenna system of Fig. 1 is thus:

$$\bar{H}_g = \left(H_1+H_3\right)\frac{\sin\alpha}{\sin\theta_1} + \left(H_2+H_4\right)\frac{\cos\alpha}{\sin\theta_2} \quad \ldots\ldots\ldots\ldots (19)$$

while the small circle component is

$$\bar{H}_s = \left(H_1+H_3\right)\frac{\cos\alpha\sin\beta}{\sin\theta_1} - \left(H_2+H_4\right)\frac{\sin\alpha\sin\beta}{\sin\theta_2} \quad \ldots\ldots (20).$$

Both of these equations are accurate if the values of H_1, H_2, H_3, and H_4 are calculated from equation (6) and others like it for H_2, H_3, and H_4. When the values of H_1, H_2, H_3 and H_4 are taken from (5) and similar equations for H_2, H_3, H_4, then equations (19) and (20) will be only approximate. Such approximate equations, of course, are very much simpler. By using the approximate values for H_1, H_2, H_3, and H_4 we obtain

$$\bar{H}_g = \frac{8i_0}{rc}\frac{\pi l}{\lambda}\left[\sin\alpha \cdot \sin\left(\frac{\pi l}{\lambda}\cos\theta_2\right) + \cos\alpha \cdot \sin\left(\frac{\pi l}{\lambda}\cos\theta_2\right)\right] \quad \ldots\ldots\ldots(21)$$

$$\bar{H}_s = \frac{8i_0}{rc}\frac{\pi l}{\lambda}\left[\cos\alpha \cdot \sin\beta \cdot \sin\left(\frac{\pi l}{\lambda}\cos\theta_2\right) - \sin\alpha \sin\beta \sin\left(\frac{\pi l}{\lambda}\cos\theta_1\right)\right] \quad \ldots\ldots(22)$$

If we again assume that l/λ is small so that

$$\sin\left(\frac{\pi l}{\lambda}\cos\theta_2\right) = \frac{\pi l}{\lambda}\cos\theta_2\,,$$

$$\sin\left(\frac{\pi l}{\lambda}\cos\theta_1\right) = \frac{\pi l}{\lambda}\cos\theta_1\,,$$

then (21) and (22) reduce to a very simple result which is easy to visualize, namely

$$\bar{H}_g = \frac{8i_0}{rc}\left(\frac{\pi l}{\lambda}\right)^2\cos\beta \ldots\ldots\ldots(23)$$

$$\bar{H}_s = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(24)$$

While these equations are exact only when l/λ is very small, they remain reasonably accurate for most purposes even when $l/\lambda = 1/8$, and the horizontal characteristic is still substantially circular when $l = 1/4\lambda$, as may be seen from Table II. A comparison of calculated "vertical characteristic" with measurements is given in Figs. 2A and 2B. In these figures the solid curves are calculated from equation (19) using accurate equation (6) for H_1, H_2, H_3, H_4. The measured values are obtained by putting the plane of the loop antenna at right angles to ground, and then measuring the resulting horizontal characteristic.

At a distant point from any antenna, the small circle component E_s of the electric field is equal to the great circle component H_g of the magnetic field and the great circle component E_g of the electric field is equal to the small circle component H_s of the magnetic field. Accordingly, in view of (23) and (24)

$$E_g = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(25)$$

$$E_s = \frac{8i_0}{rc}\left(\frac{\pi l}{\lambda}\right)^2\cos\beta \ldots\ldots\ldots\ldots(26)$$

TABLE II

Values of $H_g \frac{rc}{2i_0}$ calculated from (19) using accurate equation (6) for H_1, H_2, H_3, H_4.

$\alpha = 0°$

l/λ	$\beta=0°$	$\beta=20°$	$\beta=40°$	$\beta=60°$	$\beta=80°$	$\beta=90°$
0.05	0.097	0.092	0.075	0.049	0.017	0.000
0.10	0.380	0.358	0.294	0.192	0.067	0.000
0.15	0.822	0.788	0.641	0.422	0.148	0.000
0.20	1.39	1.33	1.09	0.727	0.256	0.000
0.25	2.00	1.90	1.64	1.09	0.384	0.000

$\alpha = 45°$

l/λ	$\beta=0°$	$\beta=20°$	$\beta=40°$	$\beta=60°$	$\beta=80°$	$\beta=90°$
0.05	0.098	0.092	0.074	0.051	0.017	0.000
0.10	0.380	0.350	0.290	0.200	0.067	0.000
0.15	0.826	0.780	0.666	0.432	0.148	0.000
0.20	1.38	1.30	1.08	0.735	0.262	0.000
0.25	2.01	1.90	1.61	1.09	0.387	0.000

Values of $H_g \frac{rc}{2i_0}$ from approximate eq. (23).

For any value of α.

l/λ	$\beta=0°$	$\beta=20°$	$\beta=40°$	$\beta=60°$	$\beta=80°$	$\beta=90°$
0.05	0.099	0.093	0.076	0.049	0.017	0.000
0.10	0.402	0.378	0.309	0.201	0.070	0.000
0.15	0.888	0.835	0.680	0.444	0.155	0.000
0.20	1.58	1.49	1.21	0.790	0.275	0.000
0.25	2.46	2.31	1.87	1.23	0.428	0.000

Values of $H_g \frac{rc}{2i_0}$ from semi-empirical equation

$$H_g = \frac{8i_0}{rc}\left(\sin\frac{\pi l}{\lambda}\right)^2\cos\beta \ldots\ldots\ldots(23A)$$

which represents accurate values much better than (23).

For any value of α.

l/λ	$\beta=0°$	$\beta=20°$	$\beta=40°$	$\beta=60°$	$\beta=80°$	$\beta=90°$
0.05	0.097	0.091	0.074	0.048	0.017	0.000
0.10	0.382	0.359	0.292	0.191	0.066	0.000
0.15	0.824	0.775	0.630	0.412	0.143	0.000
0.20	1.38	1.30	1.06	0.690	0.240	0.000
0.25	2.00	1.88	1.53	1.00	0.348	0.000

Values of $H_g \frac{rc}{2i_0}$ from accurate equations (19) and (6) for $l/\lambda = 0.25$, $\beta = 0$ and variable α.

α	0°	10°	20°	30°	40°	50°
$H_g \frac{rc}{2i_0}$	2.00	2.00	2.00	2.00	2.00	2.00
α	60°	70°	80°	90°		
$H_g \frac{rc}{2i_0}$	2.00	2.00	2.00	2.00		

This result may be stated in more familiar language. Let us suppose that the loop of Fig. 1 is in a horizontal plane. Then E_s is the horizontal component of the electric field. The great circle component E_g may be resolved into two mutually perpendicular components, one of which is vertical while the other is " longitudinal." Both of these are equal to zero when $E_g = 0$, so that the total electrical field of the loop antenna may be said to be purely polarized. There is no vertical component present in any direction.

CALCULATION OF RADIATION RESISTANCE

Since the efficiency of the U.H.F. loop antenna depends primarily on the ratio of radiation resistance to the loss resistance of the conductors, it is important to know the value of the radiation resistance of the loop antenna, at least approximately. Equation (23) provides us with a very simple, and yet reasonably accurate, value of the magnetic field at any point in space, and permits us to apply the Poynting theorem for calculating the total radiated power. In accordance with this theorem the power which is radiated through element dS of a sphere of radius r around the antenna is given by

$$\frac{c}{8\pi} H^2 \, dS \ .$$

Since field H depends only on β and not on α, we may subdivide the sphere into narrow strips of width $rd\beta$ included between adjacent small circles, such as, for example, are defined by $\beta = \beta_1$, and $\beta = \beta_1 + d\beta_1$. The area of such a strip is equal to $rd\beta \ . \ 2\pi r \cos \beta = dS$. The total power radiated through the entire sphere is $P = \frac{c}{8\pi} \int H^2 dS$. The peak current at the centre of such element is $2i_0$ and the effective current is $\frac{2i_0}{\sqrt{2}}$. If we define as the radiation resistance of the loop antenna that quantity R which, when multiplied by the square of the effective peak current, gives us the total radiated power, then

$$R\left(\frac{2i_0}{\sqrt{2}}\right)^2 = P = \frac{c}{8\pi} \int_{-\pi/2}^{+\pi/2} H^2 \, 2\pi r^2 \cos \beta \, d\beta$$

$$= \frac{16 i_0^2}{c} \left(\frac{\pi l}{\lambda}\right)^4 \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \cos^3 \beta \, d\beta = \frac{64}{3c} i_0^2 \left(\frac{\pi l}{\lambda}\right)^4$$

and hence

$$R = 320 \left(\frac{\pi l}{\lambda}\right)^4 \text{ ohms} \quad \ldots\ldots\ldots\ldots\ldots\ldots (27)$$

Table III shows how R varies with l/λ .

TABLE III

l/λ	0.05	0.10	0.15	0.20
R	0.194	3.11	15.8	49.8

Equation (27) gives values of R which are too large, and this error increases with l/λ. The difficulty lies in the fact that equation (23) used in deriving (27) gives too large a value of H for the greater values of l/λ, as may be seen from Table II. When more accurate values of radiation resistance are desired, it is necessary to calculate R by numerical integration, using (6) in place of (7) for H_1. A fair approximation is obtained from the following semi-empirical equation

$$R = 320 \left(\sin \frac{\pi l}{\lambda}\right)^4 \quad \ldots\ldots\ldots\ldots\ldots\ldots (27a)$$

Equation (27) is a special case of a more general result which holds for loops of rectangular shapes and probably for loops of any shape, namely

$$R = 320 \left(\frac{\pi \sqrt{A}}{\lambda}\right)^4 \quad \ldots\ldots\ldots\ldots\ldots\ldots (28)^*$$

where A is the area of the loop. This more general equation, like (27), is also subject to the condition that l/λ is small.

Before closing the subject of radiation resistance of loop antennae, it may be of interest to note that if several loop antennae (for example N) were superposed and connected in some manner so that they were in series and yet the currents were in phase, then each antenna would produce a field H and total field produced by all working together would be NH and the power radiated would be proportional to $N^2 H^2$. It follows that the radiation resistance of a loop antenna made of N turns of wire is N^2 times that of a single turn loop, that is

$$R_N = 320 \, N^2 \left(\frac{\pi \sqrt{A}}{\lambda}\right)^4 \text{ ohms} \ .$$

* Equation for $\bar{H}$ for a rectangular loop may be derived by the same method used to obtain $\bar{H}$ for the square loop. Incidentally, the field of a small rectangular loop is circular.

DESIGN OF U.H.F. LOOP ANTENNAE

In practice the current distribution illustrated in Fig. 1 may be obtained by the use of the circuit illustrated in Fig. 3. In this arrangement, portions *BB'*, *DD'*, *AC* of conductors 1, 2, 3 and 4 are very close to each other, so that radiations from them cancel, and the effective current distribution, so far as radiation is concerned, is as in Fig. 1.

The impedance looking into terminals (*K*, *K'*) of network of Fig. 3 is usually largely reactive, and is thus not well suited for being fed directly

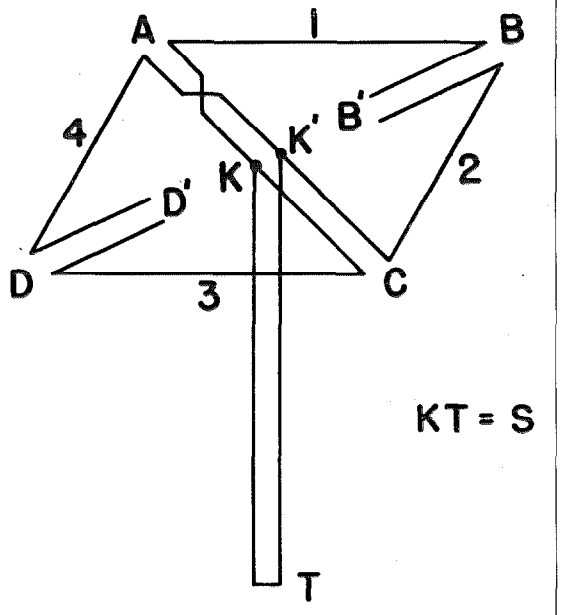


Fig. 3—U.H.F. loop antenna of Type 1.

by a shielded transmission line. In order to secure a more convenient input impedance an U.H.F. loop antenna is usually provided with what may be called a *resonant transmission line*. This feature is illustrated in Fig. 3, *KK'* – *T*.

The optimum design of the resonant transmission line, as well as the design of the radiators themselves, depends on factors which are seldom considered. For this reason, some of the details will be described.

First, we shall discuss the design of the resonant line. For a given field at a distant point it is necessary to maintain a certain value of current in the radiators. In turn, this value of current in the radiators corresponds to a cer-

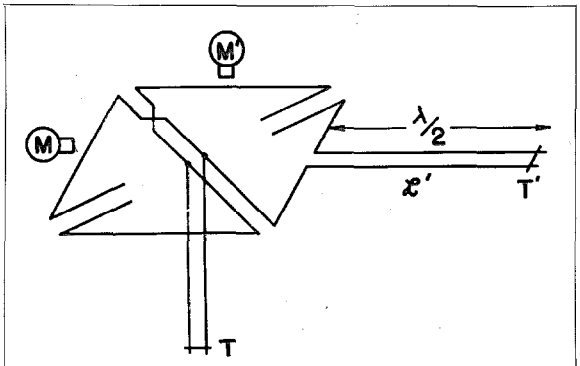


Fig. 4—Method 1.

tain fixed value of voltage across the terminals (*K K'*) of the resonant line. The power lost in this resonant line depends primarily on the current which flows in it and on the conductors which have to carry this current. It may be readily shown that the " loop current " *I* which flows through short-circuited end *T* of a transmission line section of length *s* (see Fig. 3) when the voltage is *V* at its terminals is given by the following relation :

$$I = \frac{V}{Z_0 \sin \frac{2\pi s}{\lambda}}$$

The power which is lost in the section of line is therefore

$$p = R \cdot \frac{V^2}{\left(Z_0 \sin \frac{2\pi s}{\lambda}\right)^2} .$$

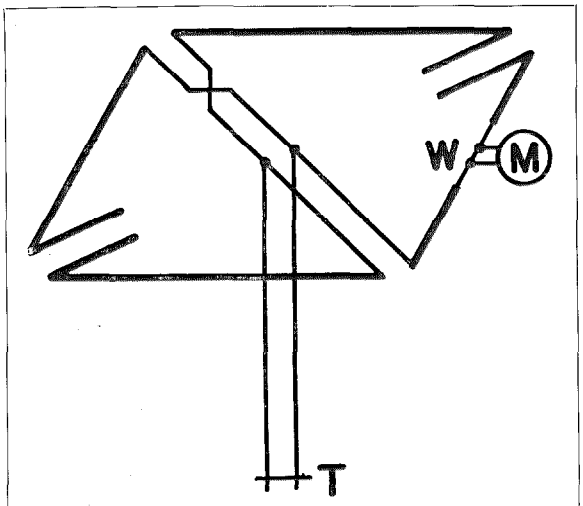


Fig. 5—Method 2.

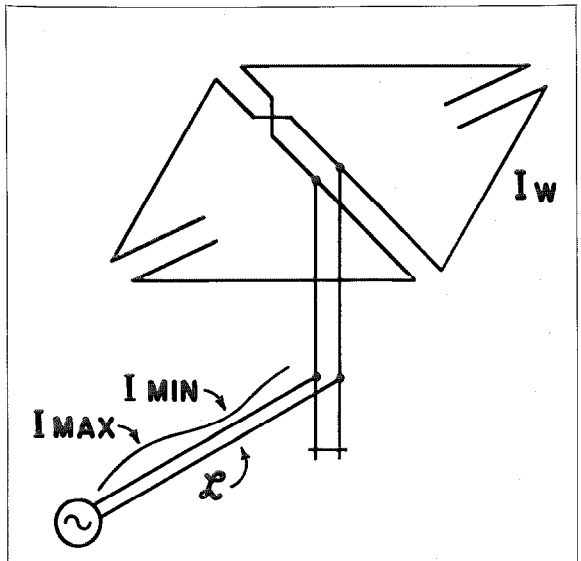


Fig. 6—Method 3.

From this equation it may be readily seen that in order that the losses may be kept at a minimum, it is desirable

(1) To keep V, which corresponds to a given current in the radiators, as low as possible.
(2) To make R as low as possible, and
(3) To keep Z_0 as high as possible.

Voltage V for a given current in the radiators can be kept down by using conductors of low surge impedance, that is, by making the radiators of wide strips of metal. The second and third requirements have to be met by making R/Z_0^2 a minimum, because the two quantities R and Z_0 are mutually interrelated. It may be pointed out that the requirement that R/Z_0^2 be a minimum is different from the usual requirement that " attenuation " in a matched line be a minimum, because the latter requirement demands that R/Z_0 and not R/Z_0^2 be a minimum. Thus, for example, in case of a coaxial line the requirement that R/Z_0 be minimum leads to the well-known ratio of a diameter of the outer to the inner conductors be around 3.8, while the requirement that R/Z_0^2 be minimum results in an optimum ratio around 9.0.

While the proportions of the conductor used in the " resonant line " are not critical, the controlling relations must be kept in mind if an efficient design is to be achieved, particularly when the dimensions of the loop antenna, and hence its radiation resistance, are relatively small.

MEASUREMENT OF RADIATION RESISTANCE

The radiation resistance of a loop antenna may be measured by several different methods.

Method 1 : The loop antenna is placed in the field of another radiator (for example, a half wave) which is a number of wavelengths away. One of the radiators of the loop antenna is cut in two as to its centre, and the circuit is completed by means of a half wavelength of transmission line as illustrated in Fig. 4. The complete circuit is resonated by moving short-circuiting bar T of the " resonant line " of the loop. The current which is induced in the loop antenna is measured by a very sensitive receiver M, so designed that it does not couple-in any appreciable resistance into the loop circuit. This freedom from coupled-in resistance may be established by using another control receiver or very sensitive meter M', while varying the coupling between receiver M and the loop.

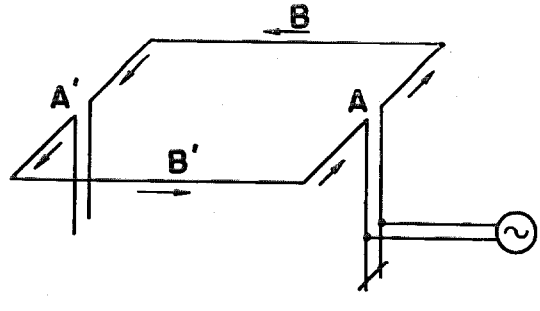


Fig. 7—U.H.F. antenna of Type II.

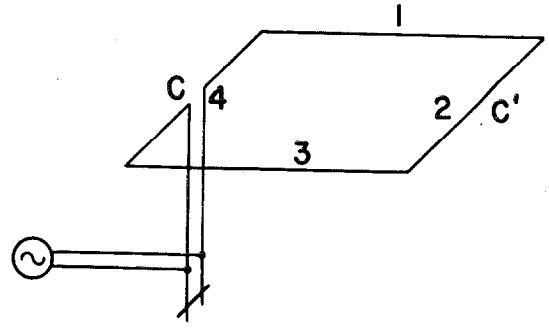


Fig. 8—U.H.F. antenna of Type III.

Receiver M must be calibrated against a linear current meter so that the reading may be translated into relative values of current.

The measurements consist of introducing known amounts of reactance in series with one of the radiators of the loop by moving short-circuiting bar T' of half wave line L' and measuring the corresponding decrease in current in the radiators by means of receiver M. The result of such measurements is a resonance curve from which the total resistance, that is, the radiation resistance plus loss resistance, can be calculated.

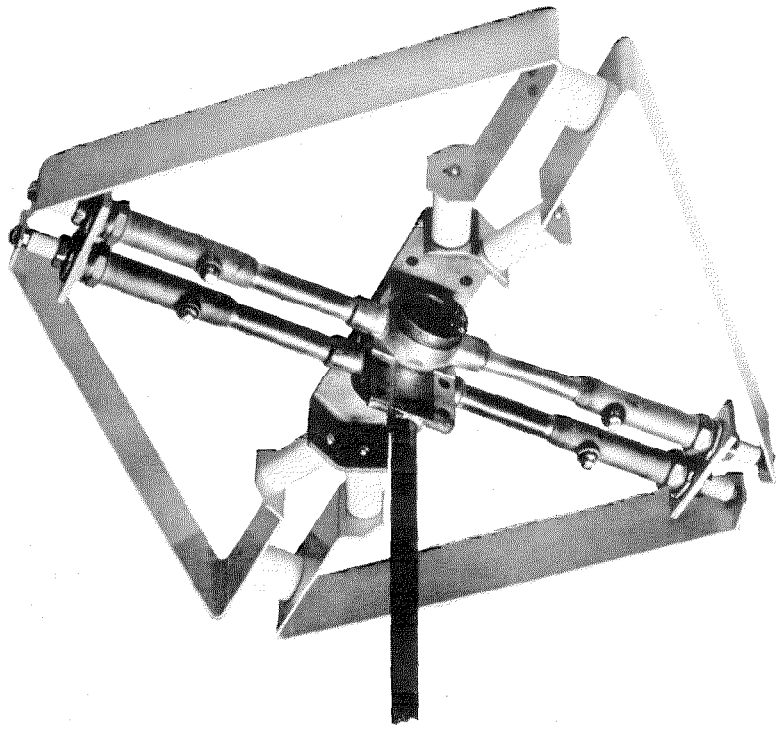


Fig. 9—U.H.F. loop antenna of Type I.

In order to separate the radiator resistance from the loss resistance, the radiators of the loop are replaced by others of similar construction, but which are bent in such a way that the area of the loop is decreased to negligible size so that the radiation resistance practically vanishes. Then another resonance curve is obtained from which the loss resistance can be calculated.

Method 2 : One of the radiators of the loop antenna is cut as in Method 1, but the gap bridged by a short piece of wire W which is used as a shunt for measuring current. The loop antenna placed in the field of another radiator as in Method 1 is resonated by means of short-circuiting bar T of the resonant line, and current I_W in wire W is compared with current I_T in short-circuiting bar T. Then short-circuiting bar T is translated by known amounts and a resonance curve is obtained. The resistance R_T which is calculated from this curve is not the sum R_W of radiation resistance plus loss resistance, but is the apparent resistance facing the short-circuiting bar T. The two resistances, however, are quite simply related because the dissipated power must be equal to

$I_T^2 R_T$ as well as to $I_W^2 R_W$

so that

$$I_T^2 R_T = I_W^2 R_W$$

and

$$R_W = R_T \frac{I_T^2}{I_W^2} \quad .$$

Since the ratio of these two currents is measured, the value of R_W can be calculated from R_T.

Method 3 : The loop antenna is supplied with power in the normal way by a transmission line L which is very nearly matched, so that the ratio of maximum to minimum of current in the standing waves is very nearly unity. The power which is supplied to the antenna can then be accurately measured by measuring the maximum and minimum currents or voltages along the transmission line. This delivered power is

$$P = Z_0 \, i_{max} \cdot i_{min}$$

where Z_0 is surge impedance of the transmission line. Current I_W which this known power produces in the radiator of the loop antenna is also

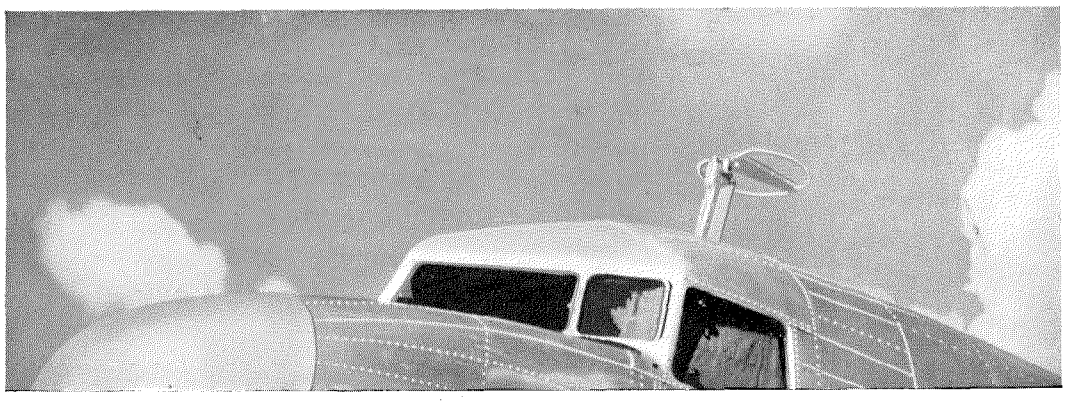


Fig. 10 —U.H.F. loop antenna of Type II installed on a transport 'plane.

measured. Then the total resistance R_W of the antenna (that is, radiation resistance + loss resistance) is calculated from the following equation

$$RI_W^2 = Z_0\, i_{max} \cdot i_{min}.$$

In order to separate the radiation resistance from the loss resistance, the radiators of the loop are then bent so as to reduce the radiation resistance to a negligible value, and the process is repeated, this time yielding only the loss resistance.

All three of these methods, plus separate measurements of resistance of the resonant line, were used to check the radiation calculated from equation (27). The results of some of these measurements are summarized in Table IV.

TABLE IV

Method 1	23 ohms
Method 2	20 ohms
Theory, Equation (27)	44 ohms
Theory, Equation (27A)	35 ohms

Frequency used 109.9 Mc.

LOOP ANTENNAE OF TYPES II AND III

It is sometimes desirable to make a loop antenna as compact and as simple as possible. Under such circumstances the design illustrated in Fig. 7 may be adopted. The loop antenna of Fig. 7 has a current distribution which results in somewhat greater radiation in direction AA' than in direction BB'. This deviation from a circular pattern, however, is rather small when the antenna dimensions are small in comparison with λ, as is usually the case. Antennae of this design will be referred to as U.H.F. loop antennae of type II.

The loop antenna of Fig. 8 is probably the simplest one of all loop antennae. The main difficulty with this type of loop is that when this antenna is made large enough to have sufficient radiation resistance to be efficient, the current distribution becomes quite asymmetric, with most of the current in radiator (2) and little current in radiator (4). This current distribution results in maximum radiation in direction CC' and some vertically polarized radiation in directions other than CC'. Such antennae will be referred to as U.H.F. loop antennae of type III.

GAIN OF LOOP ANTENNAE AND RADIATION EFFICIENCY

Since the field pattern of a small loop antenna is identical except for polarization with that of a small dipole, it is manifest that for the same amount of input power, the fields at a distant point should be equal. The difference, if any, will be due to losses in associated feeding arrangements.

The loop antennae of type I with radiation resistance around 25 ohms, if carefully designed, have radiation efficiencies somewhat greater than 95 per cent.

A direct comparison of the fields produced by a loop of type I and a horizontal half wave when the two were fed with exactly the same power showed that the fields were very nearly equal with a slight preference (about $\frac{1}{4}$ db.) for the field due to the loop.

With the smaller loop antenna having radiation resistances of the order of 5 ohms, the radiation efficiency is naturally below 95 per cent. because the loss resistance of the associated resonant circuit is usually of the order of 1 or 2 ohms.

In general, the radiation efficiency is equal to

$$\eta = \frac{\text{radiation resistance}}{\text{radiation resistance} + \text{loss resistance}}.$$

When the radiation resistance is greater than loss resistance, the efficiency is high, but when it is less than the loss resistance, the efficiency is low. When the radiation resistance is low, the radiation efficiency is proportional to the radiation resistance.

APPLICATIONS OF LOOP ANTENNAE OF THE THREE TYPES

U.H.F. antennae of type I illustrated in Figs. 3 and 9, because of their higher efficiency, are particularly well suited for use as transmitting antennae. Thus, for example, an array of these antennae may be used as the radiating system of a localizer for guiding aircraft to an airport.

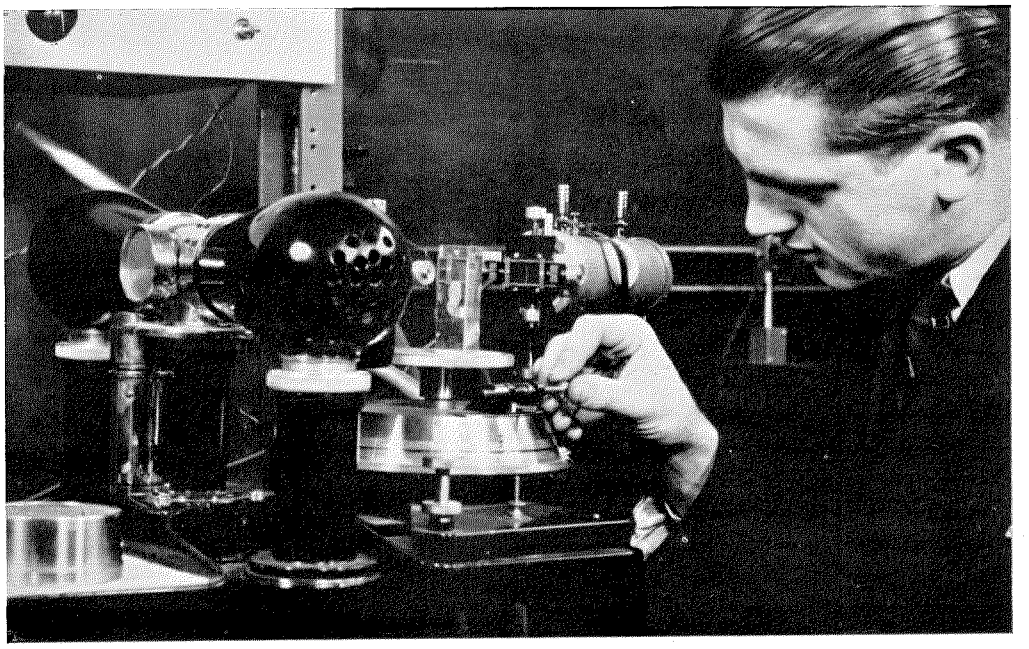
Antennae of type II are particularly well adapted for use as receiving antennae carried by aircraft.

These antennae consist of fewer parts and are therefore lighter than antennae of type I. When installed on aircraft about two feet above the fuselage, their directional characteristics in the horizontal plane are nearly circular, which is exactly what is desired in such applications.

Figure 10 illustrates an experimental antenna of this type installed on a Douglas Transport 'plane.

Antennae of type III are best suited for use in portable instruments such as, for example, field exploring devices.

U.H.F. loop antennae can be built for frequencies of the order of 10 megacycles. At these frequencies the loop antennae make excellent "omnidirectional" horizontally polarized standard antennae for measuring gain of horizontally polarized directional arrays.



Measurement of the orientation of the principal planes and axis of a piezo-electric plate for frequency stabilization, at the Laboratoires L.M.T., Paris.

High Fidelity Public Address and Translating System

By S. D. WILBURN, and S. C. TENAC,

Cía Unión Telefónica del Rio de la Plata

This paper describes a combined high fidelity public address and translating system which was provided for the Inter-American Congress for the Consolidation of Peace, Buenos Aires, and which has since been used with satisfactory results for other important multi-language conferences in South America during the past three years. The system was designed to accomplish two purposes: First, to overcome adverse acoustical conditions and permit the voice of the speakers to be heard comfortably at any point in the auditorium, with the absence of "loud speaker effect" which is generally present when the quality of the system as a whole is inferior, or when the loud speakers are visible and improperly placed with respect to the person speaking. Second, to enable the delegates to listen to nearly instantaneous translations of the speeches in any three of the four languages used by the speakers in addressing the assembly.

THE proceedings of the Inter-American Congress for the Consolidation of Peace were held in the National Hall of Congress, Buenos Aires, where acoustic conditions are very unfavourable for normal speech propagation. The ceiling height is about 25 metres, and, owing to heavily upholstered seats, a completely carpeted floor and heavy curtains of soft material over the three tiers of balconies which occupy the greater portion of the semi-circular wall space to a height of about 17 metres, there is an almost total absence of deflecting surfaces in the lower part of the auditorium. The delegates and officials at the Congress represented countries having four different national languages: Spanish, Portuguese, French and English, and as no one language was officially recognized by the Congress, it was permissible to address the conference in any one of these four languages.

In order to facilitate the proceedings, the Argentine Government instructed the Cía. Unión Telefónica del Rio de la Plata to design and install a suitable public address and translating system. The equipment was supplied by the Cía. Standard Electric Argentina, a licensee company of the International Standard Electric Corporation.

With some modifications the system was later used by the Cía. Peruana de Teléfones for the Eighth Pan-American Congress in Lima, Peru, in 1938, where four languages—Spanish, Portuguese, French and English—were used. In 1939 the system was again used by the Cía. Unión Telefónica del Rio de la Plata to provide facilities for the Universal Postal Congress which held a conference in Buenos Aires during April and May. In the latter conference all speeches were made in the official language, which is French, but for the convenience of certain groups of delegates, simultaneous translations were made in Spanish, German and English.

HIGH FIDELITY PUBLIC ADDRESS SYSTEM

It is generally recognized that for speech only, a public address system with a reasonably horizontal frequency characteristic from, say, about 200 to 3 600 p:s will give excellent results as far as intelligibility, i.e. articulation, is concerned. On the other hand, to reproduce a variety of speaking voices with complete naturalness probably requires a frequency characteristic equal to or approaching that necessary for the reproduction of orchestral music. At least, practical tests seem to indicate this.

In the particular case of the Inter-American Congress the primary prerequisite of the public address system was, in practical language, naturalness. A number of internationally-known statesmen were to use the system, some of them possessing equally familiar speaking voices of which they might well be proud, and it was especially desired to deliver these voices to the

audience with the minimum perceptible distortion from their natural timbre, such as would be brought about by an inadequate frequency characteristic or bad phase relations between the voice of the speaker and the output of the loud speakers. For this purpose the arrangement shown in schematic block form in the upper part of Fig. 1 was adopted. From the microphone to the output of the loud speakers this system had a practically smooth horizontal frequency characteristic from about 30 to more than 10 000 p : s.

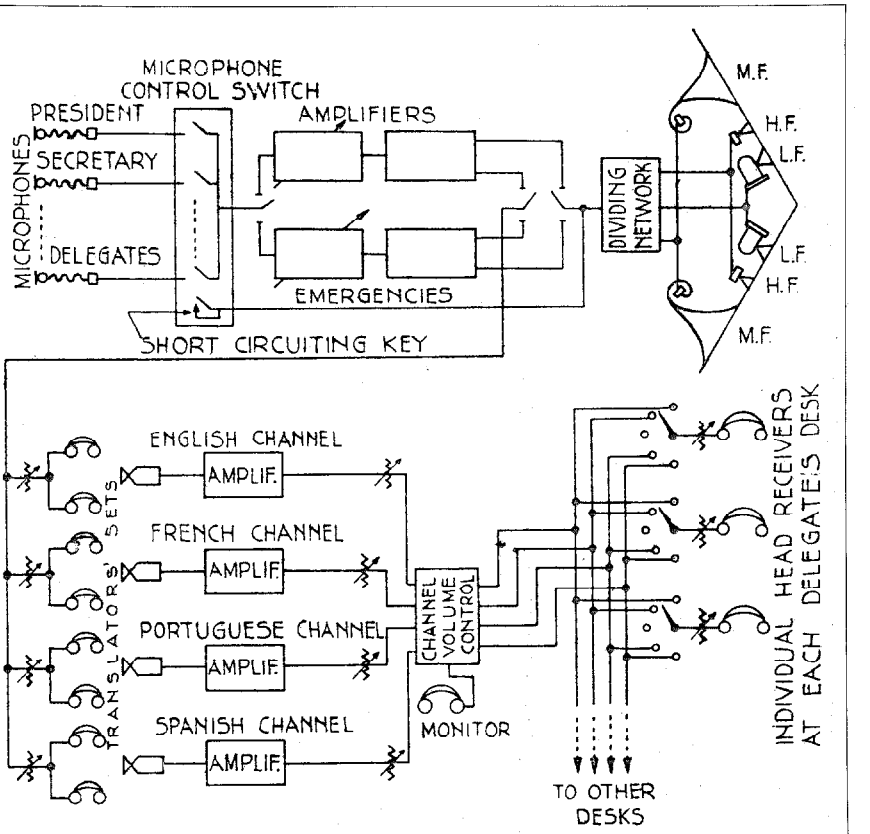


Fig. 1—Schematic of circuit arrangement for public address and translating system.

As all speeches were made from a speakers' platform facing the audience, a special reading desk was located immediately before the presiding officer's chair. Figure 2 shows the front of the desk which consisted of two latticed diagonal faces, each 88 by 140 centimetres and set at an angle of 240 degrees. Behind each latticed diagonal and inside the desk was located a tier of loud speakers, for low, medium and high frequencies. The inside of the desk where the loud speakers were located was divided into two compartments, corresponding to the two sets of loud speakers. The amplifier output was passed through a dividing network (see Fig. 3), for separating the total frequency range transmitted into three bands, corresponding to the low, medium and high frequency loud speakers. The angle of 240 degrees between the diagonal faces was adopted as tests indicated that at this angle the two sets of loud speakers inside the desk would cover uniformly the entire auditorium.

On top of the reading desk (Fig. 2) were located a moving coil microphone for the public address system serving the auditorium, a velocity microphone to serve as a pickup for an extensive national and international broadcast network, a raised shelf for the speaker's manuscript, and a tubular reading light attached to the microphone stand. The other four objects seen on top of the desk are holders in which were placed notices in four languages, requesting delegates to speak reasonably close to the microphone.

Mechanical coupling between the loud speakers inside the desk and the microphone on top of it was reduced to a minimum by completely lining with thick, soft felt the two inside compartments of the desk where the loud speakers were located. Also, both microphones on top of the desk were placed on sponge rubber pads. With this arrangement no difficulty was experienced in securing ample volume before reaching a point where energy circulated between the microphone of the public address system and the loud speakers.

The lower rear portion of the desk formed a raised speakers' platform, corresponding in height to the lower baseboard of the front of the desk, as shown in Fig. 2. To avoid vibration from the feet of the speaker, the platform was covered with a carpet over thick felt. The desk stand carrying the microphone for the public address system was set forward about 12 centimetres towards the face of the speaker. This arrangement is shown indistinctly in Fig. 2.

For the public address system a Standard Electric moving coil microphone was used in conjunction with a pre-amplifier feeding a main amplifier whose output of 30 watts was connected to the loud speakers through a dividing network. Each group of loud speakers comprised moving coil units for the high and medium frequencies, and a 12 in. paper cone dynamic speaker for the low frequencies. The latter was mounted on a hinged baffle in the lower part of the desk, the baffle permitting adjustment of the angle of the dynamic speaker to secure proper acoustic phase relations with the other loud speakers.

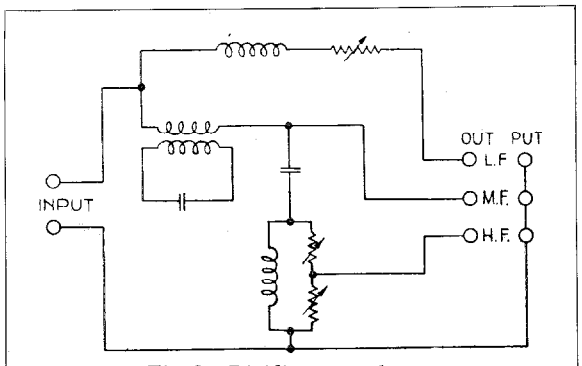


Fig. 3—Dividing network.

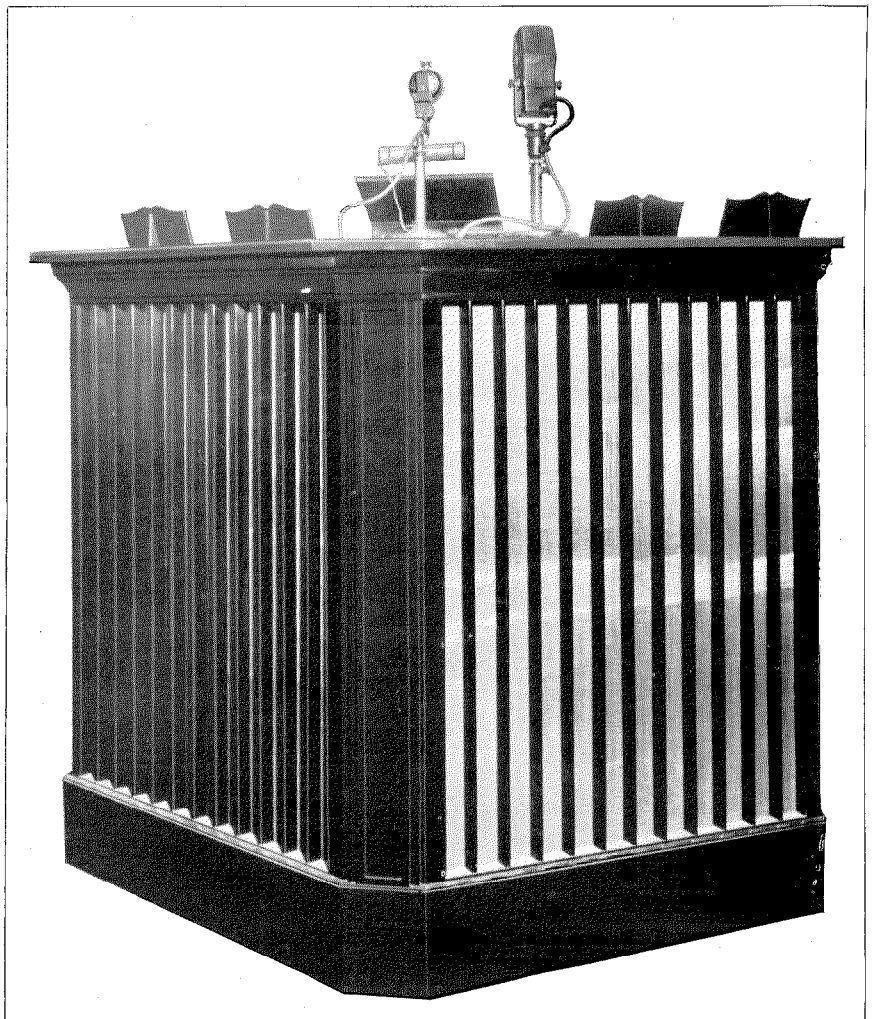


Fig. 2—Front view of special speaker's desk.

The refinement of quality coming from the loud speaker group depended upon the correct poling of the field and moving coils of all speakers, the adjustment of the angle of the hinged baffle carrying the dynamic speakers, the relative distance of each loud speaker from the front of the desk and the adjustment of the resistances in the dividing network (see Figure 3), to secure the proper relation between the output volume of the low, medium and high frequency loud speakers. Without these refinements the quality from the whole group was what would ordinarily be called very

good, but after these adjustments had been carried out with the maximum precision obtainable, the naturalness and timbre were truly impressive.

TRANSLATING SYSTEM

Obviously, the success of a project for the instantaneous translation of speeches for an audience depends upon the proficiency of the translators, whose task is difficult, even under the most favourable conditions. The engineer can do no more than facilitate the work of the translator and the listening audience by providing suitable apparatus, properly arranged.

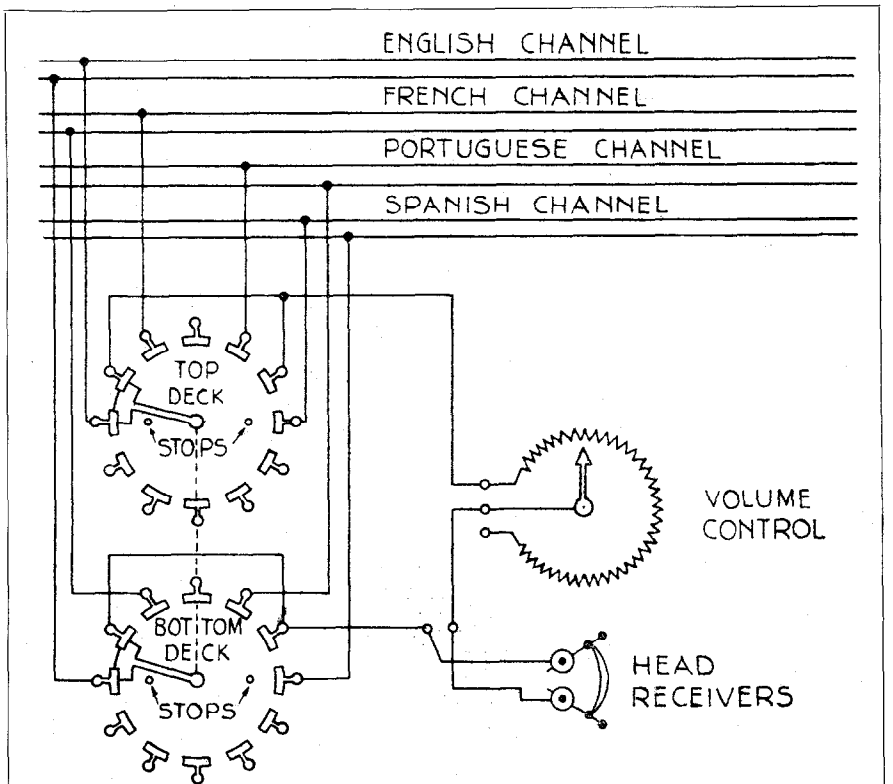


Fig. 5—Circuit diagram of language-selecting set.

The principal requirements of the apparatus are: high quality of speech reaching the translator in comfortable volume, a high fidelity transmission channel between the translator and each seat in the auditorium, a simple, plainly designated device at each seat to permit the selection of the desired language without confusion, and a means of controlling the volume of speech. Instantaneity of translations is, of course, impossible, but it can be closely approached with a good staff of translators who have had sufficient practice with the apparatus.

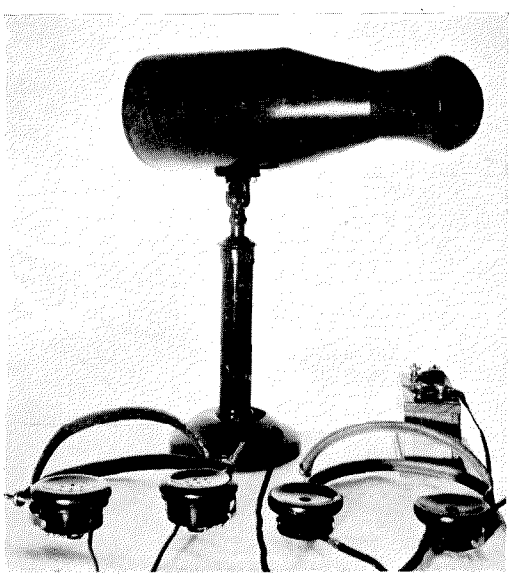


Fig. 4—Special transmitter mounting.

The arrangement of the translating circuit which was designed and provided for the Inter-American Congress is shown in schematic block form in the lower portion of Fig. 1. As indicated in this figure, a tap was taken from the output of the high fidelity public address system amplifier and connected to a group of eight high quality double head receivers for the use of the translators in listening to the speeches. Two sets of receivers were provided for each language, to permit a relief translator to follow the proceedings. This enabled the translators

to relieve each other from time to time without losing any part of the speech.

The translators were located in a small room at the back of the auditorium, and to avoid confusion of voices in the room, special transmitter mountings were provided to enable each translator to speak into his own transmitter without being heard by the other translators near by. This special transmitter mounting, shown in Fig. 4, consists of a felt lined torpedo-shaped wooden tube about 25 centimetres long, closed at one end and provided with a mouthpiece at the other end. The mouthpiece opening is sufficiently large and deep to permit the translator's lips to be well inside the flared opening. A high quality transmitter capsule was suspended by rubber bands inside the tube, and the tube was provided with three breathing holes to equalize internal air pressure. A volume control, seen at the right of Fig. 4, was provided for each translator to permit volume adjustment to suit individual requirements.

The output from each special transmitter was connected to the input of a small high grade amplifier, the output of which was connected to a monitoring set and volume control, provided to regulate the volume on the different language channels; this was necessary as the volume in the receivers at the delegates' desks was substantially influenced by variations in the number of listeners. This variation in volume was too wide to be taken care of with the individual volume control on the desk of each delegate.

From the common monitoring and volume control set the language channels were multipled to individual language-selecting sets provided at each seat in the auditorium. Figure 5 shows the detailed circuit arrangement of the individual language-selecting sets, and Fig. 6 is from a photograph of the language-selecting cabinet located on each delegate's desk. The semi-circular white disc in the lower left hand corner of the cabinet is the dial of a positive acting rotary switch used for connecting the delegate's head receiver to any one of the four language channels. The four lines on the dial correspond to the four languages. The language designation which appeared on each line of the dial is not visible in the illustration. The white

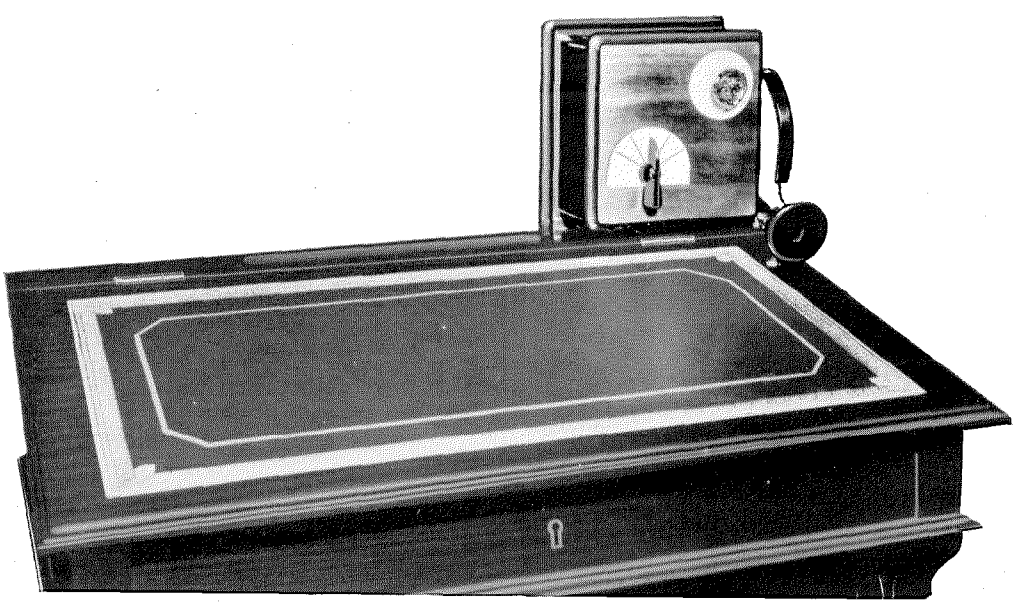


Fig. 6—Language-selecting cabinet.

dial in the upper right hand corner of the cabinet is the individual volume control.

CONCLUSION

One of the most interesting features of this public address and translating service was the general and enthusiastic comments made by delegates and others present at the Congress. This may be taken as an indication that an audience of this kind readily recognizes and keenly appreciates first-class quality in a public address system. Throughout the period of the Congress there were favourable comments from delegates and visitors of different professions, including engineers who were familiar with the subject, and statesmen who are not supposed to be particularly interested in such matters. The public address system was especially appreciated by members of the Argentine National Congress who were present, and who were familiar with the unusually unfavourable acoustic characteristics of the auditorium.

Harmonic Voltage Generation in Polyphase Rectifier Circuits*

By H. RISSIK, Hons. B.Sc. (Eng.), A.M.I.E.E., M.A.I.E.E.,

International Standard Electric Corporation

IT is well known that the D.C. output voltage of a polyphase rectifier system is not constant, but oscillates about a mean value. This value is, in fact, the mean voltage of the successively operating rectifier phases during their respective current-conducting periods. Referring to Fig. 1, the contribution of each rectifier phase to the instantaneous output voltage v_d is the same and is, in effect, the cusp of a sinusoid of maximum value $\sqrt{2}E$ (where E is the R.M.S. phase voltage of the rectifier system) extending from $\theta = -\frac{\pi}{p}$ to $\theta = +\frac{\pi}{p}$. Here p is the number of rectifier phases. Thus each phase voltage in turn has the instantaneous value $e = \sqrt{2}E \cos\theta$ between these limits.

The D.C. output voltage of a polyphase rectifier system may thus be represented by a rectilinear voltage V_d upon which is superimposed an alternating voltage "ripple," *i.e.*, a combination of A.C. voltage harmonics of varying amplitude and frequency. The amplitudes of the individual harmonics as well as of the resultant ripple depend only on the number of rectifier phases p, but the frequencies of the individual harmonics are determined also by the periodicity f of the A.C. supply. The instantaneous output voltage may be represented by the Fourier series

$$v_d = A_0 + A_1\cos p\theta + A_2\cos 2p\theta + \ldots + A_r \cos rp\theta + \ldots + B_1 \sin p\theta + B_2 \sin 2p\theta + \ldots + B_r \sin rp\theta + \ldots .$$

The constant term A_0 here represents the mean output voltage V_d which is, therefore, given by

$$V_d = \frac{p}{2\pi}\int_{-\pi/p}^{+\pi/p} \sqrt{2}E \cos\theta \, . \, d\theta$$

$$= \sqrt{2}E \frac{p}{\pi} \sin\frac{\pi}{p} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

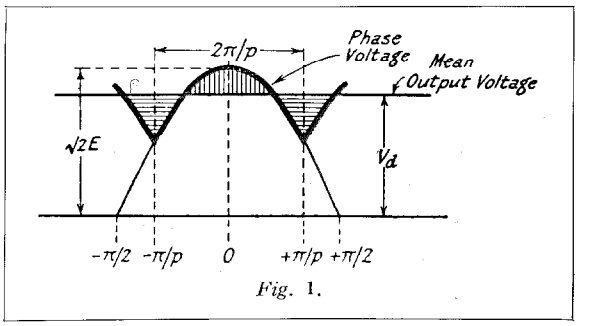


Fig. 1.

The harmonic coefficients are given by

$$A_n = \frac{p}{\pi}\int_{-\pi/p}^{+\pi/p} \sqrt{2}E \cos\theta \cos n\theta \, . \, d\theta$$

and

$$B_n = \frac{p}{\pi}\int_{-\pi/p}^{+\pi/p} \sqrt{2}E \cos\theta \sin n\theta \, . \, d\theta$$

---

* Reprinted from *The Electrician*, August 11th, 1939.

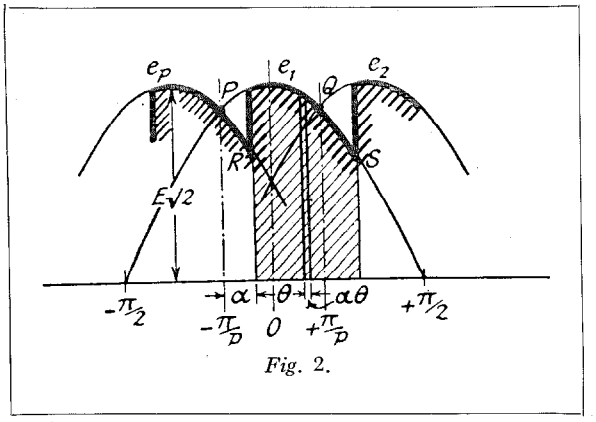


Fig. 2.

where $n = mp$ and m is any integer. The nth cosine coefficient is thus

$$A_n = \sqrt{2}E\frac{p}{\pi}\int_{-\pi/p}^{+\pi/p}\cos\theta\ \cos\ mp\theta\ .\ d\theta$$

$$= \frac{\sqrt{2}Ep}{\pi(m^2p^2-1)}$$

$$\left[2mp\ \sin\ m\pi\ \cos\frac{\pi}{p} - 2\ \cos\ m\pi\ \sin\frac{\pi}{p}\right]$$

$$= \frac{2\sqrt{2}Ep}{\pi\ (m^2p^2-1)}\left[\pm\ \sin\frac{\pi}{p}\right]$$

$$= \pm\ \frac{2}{n^2-1}\left[\sqrt{2}E\frac{p}{\pi}\sin\frac{\pi}{p}\right]$$

$$= \pm\ \frac{2V_d}{n^2-1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2a)$$

Similarly the nth sine coefficient is

$$B_n = \sqrt{2}E\frac{p}{\pi}\int_{-\pi/p}^{+\pi/p}\cos\ \theta\ \sin\ mp\theta\ .\ d\theta$$

$$= 0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2b)$$

The amplitude of the output voltage harmonic of frequency $nf = mpf$, where f is the A.C. supply frequency, is given by $\sqrt{(A_n^2 + B_n^2)}$. The R.M.S. value of the nth harmonic is thus:

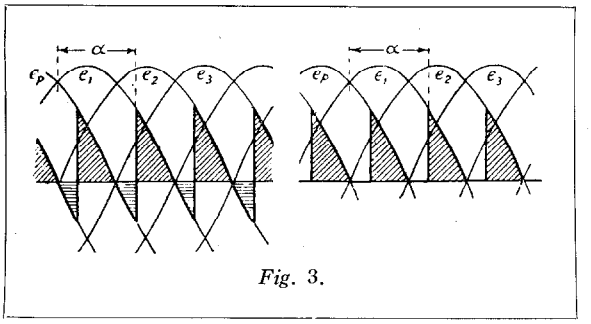


Fig. 3.

$$V_n = \frac{\sqrt{2}}{n^2-1}V_d \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Equation (3) indicates that the harmonic ripple in the output voltage of a polyphase rectifier system contains only those harmonics having frequencies which are multiples of both the supply frequency f and the number of rectifier phases p, and that the magnitude of any harmonic depends only upon its order $n = mp$.

Retarded Commutation (Grid-control)

In the case of a grid-controlled† rectifier system in which the instant of commutation is artificially delayed by the angle α—the so-called "ignition angle"—the D.C. output voltage is as shown in Fig. 2. It is clear that its mean value is less than in the case of Fig. 1. Actually it is

$$V_d = \frac{p}{2\pi}\int_{\alpha-\pi/p}^{\alpha+\pi/p}\sqrt{2}E\ \cos\ \theta\ .\ d\theta$$

$$= \frac{\sqrt{2}Ep}{2\pi}\left[\ \sin\left(\alpha+\frac{\pi}{p}\right) - \sin\left(d-\frac{\pi}{p}\right)\right]$$

$$= \sqrt{2}E\frac{p}{\pi}\sin\frac{\pi}{p}\cos\alpha$$

$$= V_{d0}\ .\ \cos\ \alpha \qquad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

where V_{d0} is the normal value of the output voltage, when $\alpha = 0$. The harmonic coefficients A_n and B_n are in this case given by

$$A_n = \sqrt{2}E\frac{p}{\pi}\int_{\alpha\mp\pi/p}^{\alpha+\pi/p}\cos\ \theta\ mp\theta\ .\ d\theta$$

$$= \sqrt{2}E\frac{p}{\pi}\ \sin\frac{\pi}{p}$$

$$\left[\frac{\cos\ (mp+1)\alpha}{mp+1} - \frac{\cos\ (mp-1)\alpha}{mp-1}\right]$$

$$= -\frac{2V_{d0}}{n^2-1}\ (\cos\ n\alpha\ \cos\ \alpha + n\ \sin\ n\alpha\ \sin\ \alpha) \qquad \ldots\ldots (5a)$$

and

$$B_n = \sqrt{2}E\frac{p}{\pi}\int_{\alpha-\pi/p}^{\alpha+\pi/p}\cos\ \theta\ \sin\ mp\theta\ .\ d\theta$$

$$= \sqrt{2}E\frac{p}{\pi}\sin\frac{\pi}{p}$$

† Cf. G. Rabuteau: "Hot-Cathode Mercury-Vapour High-Tension Supply Equipment for Broadcasting Stations," *Electrical Communication*, October, 1936, p. 153.

$$\left[\frac{\sin(mp+1)\alpha}{mp+1} - \frac{\sin(mp+1)\alpha}{mp-1}\right]$$

$$= -\frac{2V_{d0}}{n^2-1}(\sin n\alpha \cos\alpha - n\cos n\alpha \sin\alpha) \quad \ldots\ldots\ldots(5b)$$

The amplitude of the output voltage harmonic of frequency $nf = mpf$ is given by $\sqrt{(A_n^2 + B_n^2)}$. The R.M.S. value of the nth voltage harmonic is thus :

$$V_n = \frac{\sqrt{2}\,V_{d0}}{n^2-1} \cdot \sqrt{(\cos^2\alpha + n^2 \sin^2\alpha)} \quad \ldots(6a)$$

$$= \frac{\sqrt{2}\,V_d}{n^2-1} \cdot \sqrt{(1 + n^2\tan^2\alpha)} \quad \ldots\ldots(6b)$$

$$= V_{n0}\sqrt{(1 + n^2\tan^2\alpha)} \quad \ldots\ldots\ldots(6c)$$

where V_{n0} is the value of V_n in the case of normal systems, where $\alpha = 0$.

This fundamental relation is shown plotted in terms of cos α in Fig. 175 on p. 313 of the author's *Mercury-Arc Current Convertors*. (See also Table V on p. 314 of the same work.)

Equations (4) to (6c) are valid, strictly speaking, only in the case of an infinitely inductive rectifier load. The D.C. output voltage wave-form for large values of α is then as shown in Fig. 3 (a). It is seen that the induced e.m.f. of the rectifier load maintains the current flow after the phase voltage has fallen to zero. A little consideration will show that under these circumstances the output voltage wave becomes symmetrical about the zero axis when $\alpha = \pi/2$, so that the mean output voltage V_d is then zero. That this is so is also evident from equation (4). The conditions are quite different, however, in the case of a non-inductive rectifier load. Obviously, the output voltage wave cannot then continue below the zero axis ; that is to say, each rectifier phase ceases to carry current as soon as its voltage has fallen to zero. The output voltage wave is accordingly cut off at the zero axis as shown in Fig. 3 (b), the point of cut-off occurring where $\alpha = (\pi/2 - \pi/p)$. For values of α less than this, equation (4) is valid even in the case of a non-inductive rectifier load. Beyond this critical value, however, the upper limit of integration in the expression for V_d then becomes $\pi/2$ instead of $(\alpha + \pi/p)$ so that the mean output voltage is given by

$$V_d = \frac{p}{2\pi}\int_{\alpha-\pi/p}^{\pi/2} \sqrt{2}E\cos\theta \,.\, d\theta$$

$$= \frac{\sqrt{2}Ep}{2\pi}\left[1 - \sin\left(\alpha - \frac{\pi}{p}\right)\right]$$

$$= V_{d0}\left[\frac{1 - \sin(\alpha - \pi/p)}{2\sin(\pi/p)}\right] \quad \ldots\ldots\ldots(7a)$$

It is seen that in this case V_d becomes zero when $\alpha = (\pi/2 + \pi/p)$. Also, when the ignition angle exceeds the critical value $\alpha = (\pi/2 - \pi/p)$, the current-conducting period per rectifier phase is no longer $2\pi/p$ but $(\pi/2 + \pi/p - \alpha)$. Denoting this latter angle by β, we have $(\alpha - \pi/p) = (\pi/2 - \beta)$. In terms of the current-conducting period per rectifier phase equation (7a) thus becomes

$$V_d = V_{d0}\left[\frac{1 - \cos\beta}{2\sin(\pi/p)}\right]$$

$$= V_{d0}\left[\frac{1 - \cos\beta}{\sqrt{2\,(1 - \cos\beta_0)}}\right] \quad \ldots\ldots\ldots(7b)$$

where β_0 is the normal current-conducting period per phase, when $\alpha = 0$. It is seen that, as before, V_d is zero when $\beta = 0$, i.e., when $\alpha = (\pi/2 + \pi/p)$. It is evident also that when the ignition angle reaches the critical value $\alpha = (\pi/2 - \pi/p)$, and hence $\beta = \beta_0$, equations (4), (7a) and (7b) must all give the same value for the D.C. output voltage, viz., $V_d = V_{d0}\sin(\pi/p)$.

As the result of the change in wave-form of the D.C. output voltage illustrated in Fig. 3, the several voltage harmonics are also altered in value. In the case of non-inductive rectifier loading the harmonic coefficients A_n and B_n become

$$A_n = \sqrt{2}E\frac{p}{\pi}\int_{\alpha-\pi/p}^{\pi/2} \cos\theta\cos mp\theta \,.\, d\theta$$

$$= \sqrt{2}E\frac{p}{\pi}\int_{\pi/2-\beta}^{\pi/2} \cos\theta\cos mp\theta \,.\, d\theta$$

$$= \frac{\sqrt{2}Ep}{\pi(m^2p^2-1)}$$

$$\left[mp\sin mp\theta\cos\theta - \cos mp\theta\sin\theta\right]_{\pi/2-\beta}^{\pi/2}$$

$$= \frac{\sqrt{2}Ep}{\pi(n^2-1)} \cdot \sin\frac{n\pi}{2}$$

$$\left(\sin n\beta \cos\beta - n\cos n\beta \sin\beta\right) \ldots\ldots(8a)$$

and $$B_n = \sqrt{2}E\frac{p}{\pi}\int_{\alpha-\pi/p}^{\pi/2} \cos\theta \sin mp\theta \cdot d\theta$$

$$= \sqrt{2}E\frac{p}{\pi}\int_{\pi/2-\beta}^{\pi/2} \cos\theta \sin mp\theta \cdot d\theta$$

$$= \frac{\sqrt{2}Ep}{\pi(m^2p^2-1)}$$

$$\left[-mp\cos mp\theta\cos\theta - \sin mp\theta\sin\theta\right]_{\pi/2-\beta}^{\pi/2}$$

$$= \frac{\sqrt{2}Ep}{\pi(n^2-1)} \cdot \sin\frac{n\pi}{2}$$

$$\left(\cos n\beta\cos\beta + n\sin n\beta\sin\beta - 1\right) \ldots(8b)$$

As before, the amplitude of the output voltage harmonic of frequency $nf = mpf$ is given by $\sqrt{(A_n^2 + B_n^2)}$. In the case of non-inductive rectifier loading, and with $\alpha > (\pi/2 - \pi/p)$, equations (8a) and (8b) thus give for the R.M.S. value of the nth harmonic

$$V_n = \frac{\sqrt{2}Ep}{\pi(n^2-1)} \cdot \sqrt{\left[\cos^2\beta + n^2\sin^2\beta - 2n\sin n\beta\sin\beta - 2n\cos n\beta\cos\beta + 1\right]}$$

$$= \frac{\sqrt{2}Ep}{\pi(n^2-1)} \cdot \sqrt{\left[(\cos n\beta - \cos\beta)^2 + (\sin n\beta - n\sin\beta)^2\right]} \quad \ldots\ldots(9a)$$

$$= V_{n0} \cdot \sqrt{\left[\frac{(\cos n\beta - \cos\beta)^2 + (\sin n\beta - \sin\beta)^2}{(1-\cos\beta_0)}\right]} \quad \ldots\ldots(9b)$$

which last expression is of the same form as (6c).* The validity of equation (9b) may be verified by considering the limiting case where $\alpha = (\pi/2 - \pi/p)$, and thus $\beta = \beta_0 = 2\pi/p$. We then have

$$V_n = V_{n0} \cdot \frac{\sqrt{[(1-\cos 2\pi/p)^2 + (0 - n\sin 2\pi/p)^2]}}{(1-\cos 2\pi/p)}$$

$$= V_{n0} \cdot \sqrt{(1 + n^2\cot^2\frac{\pi}{p})}$$

$$= V_{n0} \cdot \sqrt{(1 + n^2\tan^2\alpha)}$$

which is the same result as (6c), as is to be expected.

Finite Commutation (Overlapping)

So far consideration has been given only to the ideal case of polyphase rectification in which commutation of the load current from one rectifier phase to the next is assumed to occur instantaneously—in fact, at the instants corresponding to the points of intersection of the phase voltage waves (P and Q in Fig. 2) in the case of a *normal* rectifier system; and at the instants when the grid potentials are made positive (R and S in Fig. 2) in the case of a *grid-controlled* rectifier system.

Consider now the actual case in which, due to the inevitable presence of circuit inductance, commutation does not take place instantaneously but occupies a finite time represented in Fig. 4 by the angle of overlap, u—so-called because during this interval the two successive phase currents overlap, and in such a way that their instantaneous sum remains equal to the steady D.C. load current carried by each rectifier phase in turn.† During the overlap interval the instantaneous output voltage v_d is the mean of the instantaneous voltages of the two overlapping phases; during the remainder of the current-conducting period $(2\pi/p - u)$, v_d is equal to the instantaneous voltage of the working phase. Taking the point of intersection of the two phase-voltage waves as the origin of reference, the mean output voltage is given by

$$v_d = \frac{p}{\pi}\left[\int_0^u \frac{1}{2}(e_1 + e_2)\,d\theta + \int_u^{2\pi/p} e_2 d\theta\right]$$

$$= \frac{\sqrt{2}Ep}{2\pi}\left[\int_0^u \frac{1}{2}(\cos\overline{\theta + \pi/p} + \cos\overline{\theta - \pi/p})\,d\theta\right.$$

* Cf. Dr. Teago's analysis on p. 420 of *The Electrician*, April 1, 1938, where the above equation (9b) is derived in a somewhat different form. The analysis there established appears to be faulty in that it leads to incorrect expressions corresponding to the preceding equations (8a) and (8b).

† Cf. The author's *Mercury-Arc Current Convertors*, pp. 26–34 (Sir Isaac Pitman & Sons, 2nd edn., 1940).

$$+\int_u^{2\pi/p}\cos(\theta-\pi/p)d\theta\Big]$$

$$=\frac{\sqrt{2}Ep}{2\pi}\sin\frac{\pi}{p}.(1+\cos u)$$

$$=V_{d0}\left[\frac{1+\cos u}{2}\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots(10)$$

A comparison of equations (4) and (10) shows that for a given angle of delayed commutation the reduction in D.C. output voltage is less (actually one-half as great) if the delay is *natural* (*i.e.*, due to the presence of circuit inductance) than if the delay is *forced* (*i.e.*, due to the action of grid-control). This is clearly shown by Fig. 5, for which $p = 6$ and in which the angle of delay is π/p : diagram (a) shows the output voltage for $\alpha = 30°$, and diagram (b) that for $u = 30°$.

Here, again, the difference in output voltage wave-form makes for a difference in magnitude of the several output-voltage harmonics. In the case of normal commutation and current overlapping, as in Fig. 4, the instantaneous D.C. output voltage may be written

$$v_d = \left[\frac{1}{2}(e_1+e_2)\right]_0^u + \left[e_2\right]_u^{2\pi/p}$$

$$= \left[e_2\right]_0^{2\pi/p} - \left[\frac{1}{2}(e_2-e_1)\right]_0^u$$

$$= \left[\sqrt{2}E\cos(\theta-\pi/p)\right]_0^{2\pi/p}$$

$$-\left[\sqrt{2}E\sin(\pi/p)\sin\theta\right]_0^u$$

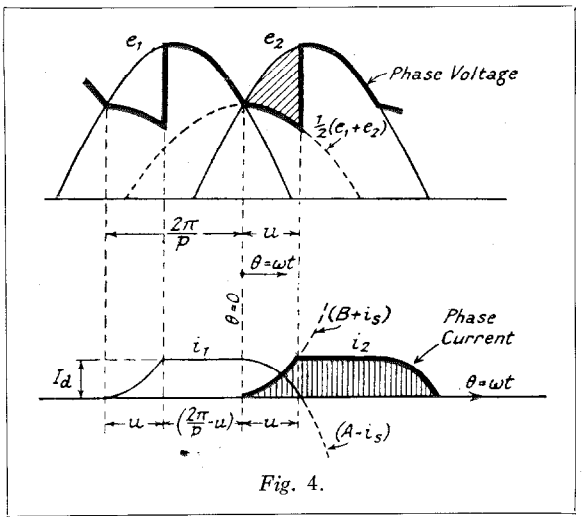


Fig. 4.

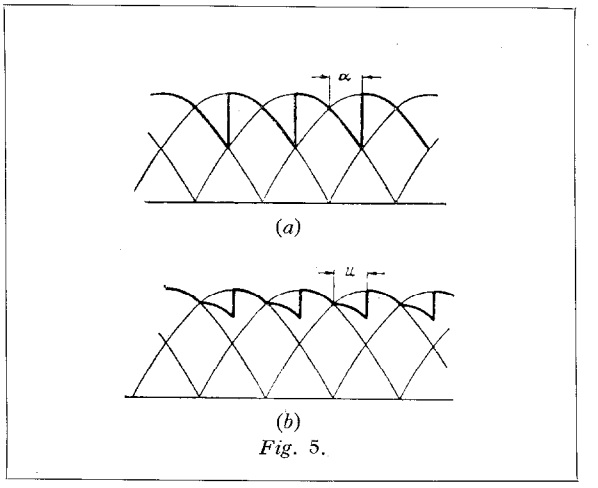


Fig. 5.

so that the harmonic coefficients A_n and B_n become

$$A_n = \frac{p}{\pi}\left[\sqrt{2}E\int_0^{2\pi/p}\cos(\theta-\frac{\pi}{p})\cos mp\theta\,.\,d\theta\right.$$

$$\left.-\sqrt{2}E\sin\frac{\pi}{p}\int_0^u\sin\theta\cos mp\theta\,.\,d\theta\right]$$

$$=\frac{\sqrt{2}Ep}{2\pi}\sin\frac{\pi}{p}\left[\frac{1+\cos(mp+1)u}{mp+1}\right.$$

$$\left.-\frac{1+\cos(mp-1)u}{mp-1}\right]$$

$$=\frac{V_{d0}}{2}\left[\frac{\cos(n+1)u}{n+1}-\frac{\cos(n-1)u}{n-1}-\frac{2}{n^2-1}\right]$$

$$=\frac{-V_{d0}}{n^2-1}(\cos nu\cos u+n\sin nu\sin u+1)$$

$$\ldots\ldots\ldots\ldots(11a)$$

and

$$B_n = \frac{p}{\pi}\left[\sqrt{2}E\int_0^{2\pi/p}\cos(\theta-\frac{\pi}{p})\sin mp\theta\,.\,d\theta\right.$$

$$\left.-\sqrt{2}E\sin\frac{\pi}{p}\int_0^u\sin\theta\sin mp\theta\,.\,d\theta\right]$$

$$=\frac{\sqrt{2}Ep}{2\pi}\sin\frac{\pi}{p}\left[\frac{\sin(mp+1)u}{mp+1}\right.$$

$$\left.-\frac{\sin(mp-1)u}{mp-1}\right]$$

$$=\frac{V_{d0}}{2}\left[\frac{\sin(n+1)u}{n+1}-\frac{\sin(n-1)u}{n-1}\right]$$

$$=-\frac{V_{d0}}{n^2-1}(\sin nu\cos u-n\cos nu\sin u) \quad \ldots(11b)$$

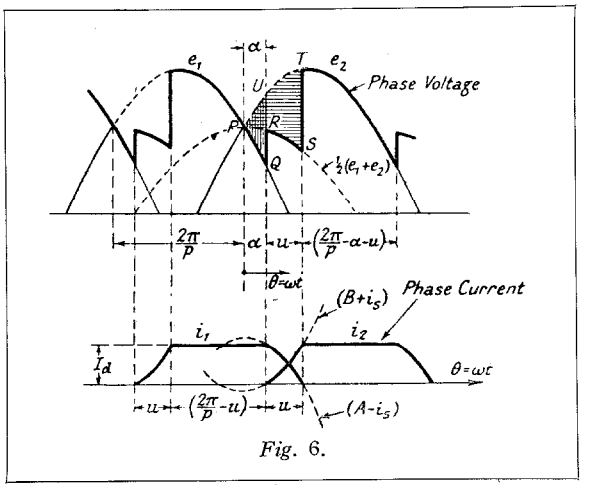


Fig. 6.

As is to be expected from a comparison of Figs. (5a) and (5b) the above two equations (11a) and (11b) lead, for a given angle of commutation delay (u or α), to lower values for A_n and B_n than do the corresponding equations (5a) and (5b). In this case the R.M.S. value of the nth harmonic becomes

$$V_n = \sqrt{\frac{1}{2}(A_n^2 + B_n^2)}$$

$$= \frac{\sqrt{2}\,V_{d0}}{n^2-1} \cdot \frac{1}{2}\sqrt{\Big[1+\cos u\,(\cos u + 2\cos nu) + n\sin u\,(n\sin u + 2\sin nu)\Big]}$$

$$= V_{n0} \cdot \frac{1}{2}\sqrt{\Big[(\cos nu + \cos u)^2 + (\sin nu + n\sin u)^2\Big]} \quad \ldots\ldots(12a)$$

which approximates to

$$V_n = V_{n0} \cdot \sqrt{\Big[\frac{1}{2}(1+\cos nu) + \frac{nu}{4}(nu + \sin nu)\Big]} \quad \ldots\ldots\ldots(12b)$$

Equation (12a) is shown plotted in terms of u and for several values of $n = mp$ in Figs. 38, 39 and 40 on pp. 92 and 93 of Marti and Winograd's *Mercury-Arc Power Rectifiers* (McGraw Hill Book Co., 1930); whilst the universal relation of equation (12b) is shown plotted in terms of the product $nu = mpu$ in Fig. 10 on p. 52 of the author's *Fundamental Theory of Arc Convertors* (Chapman & Hall, 1939).

Finite and Retarded Commutation (Overlapping with Grid-control)

Lastly, Fig. 6 shows the general case in which both a forced and a natural delay take place in the commutation process, represented by the ignition angle α and the succeeding angle of overlap u. In this case the instantaneous D.C. output voltage may be written

$$v_d = \Big[e_1\Big]_0^{\alpha} + \Big[\frac{1}{2}(e_1+e_2)\Big]_{\alpha}^{\alpha+u} + \Big[e_2\Big]_{\alpha+u}^{2\pi/p}$$

$$= \Big[e_2\Big]_0^{2\pi/p} - \Big[e_2-e_1\Big]_0^{\alpha} - \Big[\frac{1}{2}(e_2-e_1)\Big]_{\alpha}^{\alpha+u}$$

$$= \Big[\sqrt{2}E\cos(\theta-\pi/p)\Big]_0^{2\pi/p} - \Big[2\sqrt{2}E\sin(\pi/p)\sin\theta\Big]_0^{\alpha} - \Big[\sqrt{2}E\sin(\pi/p)\sin\theta\Big]_{\alpha}^{\alpha+u}.$$

The first term in the above expression is the output voltage in the ideal case, and leads to equation (1). The second term takes account of the loss in output voltage due to retarded commutation (grid-control). The third term similarly takes into account the further loss associated with finite commutation (current overlapping). The mean D.C. output voltage is thus

$$V_d = \frac{p}{2\pi}\int v_d \,.\, d\theta = V_{d0} - V_{d0}(1-\cos\alpha) - V_{d0}\Big(\frac{\cos\alpha - \cos\overline{\alpha+u}}{2}\Big)$$

$$= V_{d0}\Big[\frac{\cos\alpha + \cos(\alpha+u)}{2}\Big] \quad \ldots\ldots(13)$$

This is the fundamental output-voltage equation in its most general form and reduces to (10) when $\alpha = 0$, to (4) when $u = 0$, and to (1) when both $\alpha = 0$ and $u = 0$. The several voltage harmonics may be evaluated in the usual way. The Fourier coefficients are, as before, given by

$$A_n = \frac{p}{\pi}\int v_d \cos n\theta \,.\, d\theta \text{ and}$$

$$B_n = \frac{p}{\pi}\int v_d \sin n\theta \,.\, d\theta$$

where $n = mp$ and m is any integer. The integration gives

$$A_n = \frac{V_{d0}}{2}\left[\frac{\cos(mp+1)\alpha+\cos(mp+1)(\alpha+u)}{mp+1} - \frac{\cos(mp-1)\alpha+\cos(mp-1)(\alpha+u)}{mp-1}\right]$$

and

$$B_n = \frac{V_{d0}}{2}\left[\frac{\sin(mp+1)\alpha+\sin(mp+1)(\alpha+u)}{mp+1} - \frac{\sin(mp-1)\alpha+\sin(mp-1)(\alpha+u)}{mp-1}\right].$$

The amplitude of the harmonic voltage of frequency $nf = mpf$ is, as before, $\sqrt{(A_n^2 + B_n^2)}$ so that the R.M.S. value of the nth output-voltage harmonic is

$$V_n = V_{n0} \cdot \frac{1}{2}\sqrt{\left[2(\cos^2\alpha + n^2\sin^2\alpha) + (n^2-1)(\sin^2\overline{\alpha+u} - \sin^2\alpha) + 2\cos nu(\cos\overline{\alpha+u}\cos\alpha + n^2\sin\overline{\alpha+u}\sin\alpha) + 2n\sin nu\sin u\right]}$$

$$= V_{n0} \cdot \frac{1}{2}\sqrt{\left[2(1+\cos nu\cos u)(\cos^2\alpha+n^2\sin^2\alpha)+2n\sin nu\sin u + (n^2-1)\sin u(\cos nu\sin 2\alpha+\sin 2\alpha+u)\right]} \quad \ldots\ldots\ldots\ldots (14a)$$

This general expression for V_n reduces to (12a) in the case of normal rectifier operation, where $\alpha = 0$; and to (6a) in the case of the ideal grid-controlled rectifier, for which $u = 0$. For values of $\alpha < 45°$ equation (14a) approximates to

$$V_n = V_{n0} \cdot \sqrt{\left[(\cos^2\alpha + n^2\sin^2\alpha)(1+\cos nu)\right]} \quad \ldots\ldots\ldots\ldots (14b)$$

which is of the same form as (6a) and (12b).

If the rectifier load is non-inductive, equations (13) and (14) remain valid for values of α up to the critical value $\alpha = (\pi/2 - \pi/p)$. Thereafter overlapping can no longer occur, for the current in one phase falls to zero before the next phase becomes operative. Hence equations (7) to (9) then apply.

An Impulse Measuring Set*

By A. S. GRANT, B.Sc., and D. H. MACNEE, B.Sc., A.M.I.E.E.,

Valve Laboratory, Standard Telephones and Cables, Limited, North Woolwich, England

The paper describes an instrument designed to measure the crest amplitude of an electrical impulse, and at the same time the decay of that impulse to a specified level below the initial peak. The instrument was designed primarily to measure the microphonic noise experienced in valves, but could be equally well applied to the measurement of the crest and decrement of any similar wave form. The instrument is direct reading and can be made readily applicable to a wide range of values.

INTRODUCTION

THE measurement of a complex wave form of a transient nature characterized by an initial peak followed by a more or less heavily damped wave, of average amplitude appreciably smaller than the initial peak, presents certain fundamental difficulties. The method most commonly employed is that afforded by the oscillograph, and while this undoubtedly provides the maximum of information of the nature of the wave form under examination, it is only applicable in practice to those cases in which a limited number of measurements are required, on account of the expense and time involved in taking individual oscillograms.

Since for many of the purposes for which the instrument, herein described, was required it was necessary to differentiate between the type of wave characterized by a high crest and rapid decrement and that characterized by relatively low crest and slow decrement, facilities were provided for the simultaneous measurement of the crest value and decrement, the latter in terms of the time taken to reach a predetermined level below the crest.

The crest impulse is registered by a diode-condenser combination. The signal is rectified and caused to charge a condenser connected between cathode and grid of an amplifying valve, the registering meter is connected in the cathode-anode circuit of this valve and deflects in proportion to the crest applied signal. For this measurement the diode-condenser combination must have a leakage sufficiently small to hold its charge substantially constant over the period of measurement, and the condenser must reach a condition of full charge during the extremely brief period of the crest signal. This was attained by the use of a cascade arrangement of rectifier and progressively increasing capacities such that the first met the requirements of charging time, and the last the requirement of low leakage. Fuller details of this arrangement are given later in this paper. It was further necessary to devise a timing device which could be made to start operation on the application of the applied signal and to cease when that signal amplitude had decayed to a specified level below the initial amplitude. To this effect, use is made of the well-known valve bridge circuit shown schematically in Fig. 1, a sensitive relay R being bridged between the plates of the two similar valves V_1 and V_2. Assuming the applied voltages E_1 and KE_1 and resistive loads R_1 and R_2 to be equal, there will be no resultant current in the relay R. Voltage E_1 is derived from the crest signal to which it is directly proportional, while voltage KE_1 is obtained from the output of an amplifier (of gain K), which is also energized by the applied signal. Hence at the moment of application of the signal there will exist a difference of level between the grid voltages of V_1 and V_2 and a consequent unbalance current in their plate

* Reprinted from the *Journal of Scientific Instruments*, Vol. xvi, No. 9, September, 1939.

circuits which can be used to operate relay R. E_1, the voltage applied to the grid of V_1, is held substantially constant during measurement as previously explained; the voltage KE_1 will however become progressively smaller as the signal decays, until finally it reaches the same value as E_1; the bridge is then once more in balance and there is no resultant current through R. Relay R is used to effect the starting and stopping of the timing device, which will thus start on application of the signal, and stop when the signal level has decayed by a level difference K which is afforded by the gain of the amplifier. The operation of this circuit is explained in somewhat greater detail in a subsequent paragraph.

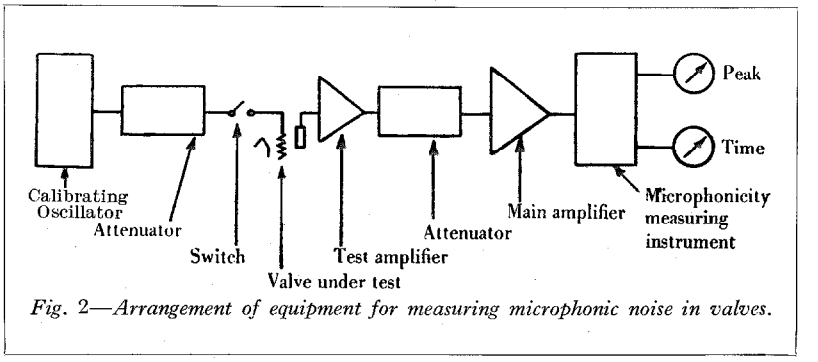


Fig. 2—Arrangement of equipment for measuring microphonic noise in valves.

DESCRIPTION OF INSTRUMENT

In what follows the instrument is described in a form suitable for the measurement of valve microphonicity. It will, however, be readily appreciated that, by substituting the impulse voltage it is desired to measure for the valve under test in the following description, the instrument could equally well be utilized to measure the peak and decrement of any impulse wave within its range.

A schematic diagram indicating the arrangement of the equipment necessary for the measurement of microphonic noise in valves is shown in Fig. 2. The test amplifier which incorporates the valve under test is followed by a variable attenuator and a main amplifier. The main amplifier, which has a gain of 75 db. and a frequency characteristic which is linear to within ± 2.5 db. between 20 and 10 000 p : s, and to within 0.5 db. between 90 and 8 000 p : s, feeds the impulse measuring set.

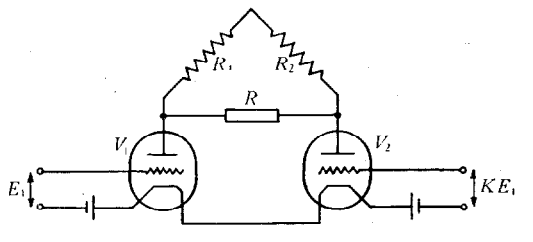


Fig. 1—Balanced valve bridge circuit.

A simplified schematic diagram of the impulse measuring set is shown in Fig. 3. Consider this circuit with no signal voltage impressed between the points a, a. The push switch P_1 is depressed and released, bringing the grid of valve V_5 to its cathode potential and allowing the meter M_1 to read its maximum value. As the current through V_5 is inversely proportional to the grid voltage or signal, the maximum reading on the meter is considered as the zero. The grid of the similar valve V_6 is also at cathode potential when no signal is present, thus the valves V_5 and V_6 are taking equal anode currents through the equal resistances R_{17} and R_{22} respectively. Hence the bridge circuit formed by V_5, V_6, R_{17} and R_{22} is in balance, and no current flows through the relay R, the contacts of which are open. At the same time as P_1 is operated the push switch P_2 is depressed and released, bringing the grid of the valve V_9 to a potential, negative with respect to its cathode, due to the current through R_{14}. Now the valve and resistances V_9, R_{19}, R_{13} and R_{12} form the arms of a bridge circuit with the meter M_2 connected to opposite junctions. The values are such that when the grid of V_9 is at cathode potential the voltage drop across R_{19} is equal to that across R_{12} and no current passes through the meter; also, when the push switch P_2 is depressed the current through the valve and R_9 decreases, the voltage drops across R_{19} and R_{12} are no longer equal, and the out of balance current causes the meter to read full scale. Again, this position of the meter is considered as the zero. When a signal appears at aa it is simultaneously applied

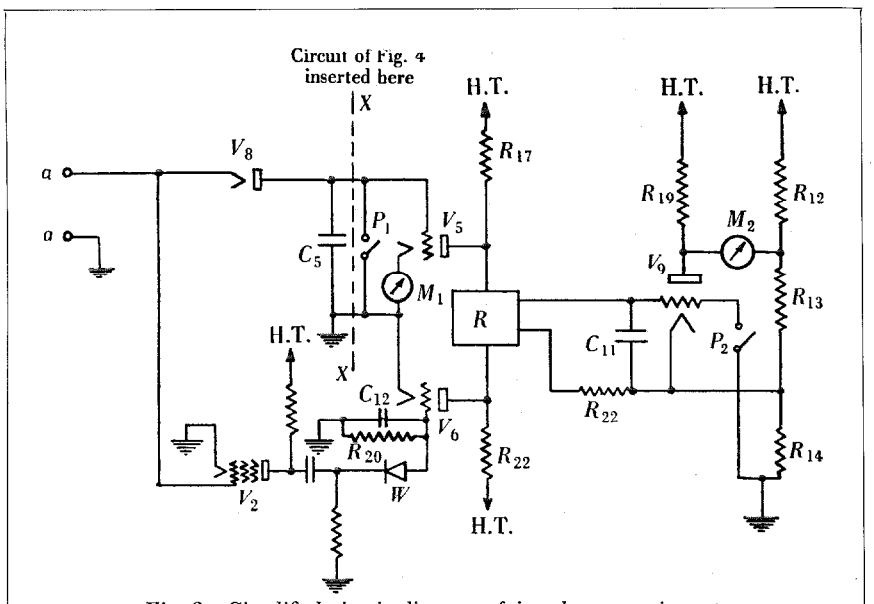


Fig. 3—Simplified circuit diagram of impulse measuring set.

between grid and cathode of V_2 and between the cathodes of the diode V_8 and of the triode V_5. The effect of the signal applied between the cathodes of the diode and triode is that a rectified negative potential is built up across the condenser C_5 proportional to the crest value of the applied signal. This biases back the valve V_5 and the meter M_1 deflects in proportion to the crest value. It may be assumed at this stage that the charging time of the condenser C_5 through the diode is very small compared with the time for which the signal is at its crest, and that the back resistance of the diode and the grid cathode resistance of V_5 is so large that once the potential has been built up on the condenser, it discharges very slowly. In practice, using commercial valves, these conditions are unobtainable with a single stage; however, in a later paragraph it is explained how this effect is produced.

It is assumed, then, that immediately the signal is applied the current through V_5 is reduced by a proportionate amount and remains indefinitely at this figure.

The circuit immediately associated with V_2 is specially designed to have a linear amplification of 30 db. between 0 and 20 V output; above this level the gain falls off so that the output volts remain approximately constant with increasing input. Now the signal, amplified by V_2, will be impressed between grid and cathode of V_6 through the metal rectifier W. This results in the condenser C_{12} receiving a rectified negative potential which will be proportional to the crest value of the applied voltage. The time constant of C_{12} and R_{20} is designed to be large compared with the frequency period of the applied signal, but is sufficiently small for the voltage on the grid of V_6 to approximate closely to the crest value of the output of V_2 at any instant during the time for which the signal is present.

It may now be seen that when the signal first appears at its maximum value, the current in V_5 falls to a certain figure and the current in V_6 falls, if not to zero, to a value 30 db. below that of V_5. The bridge formed by V_5, V_6, R_{17} and R_{22} is unbalanced and relay R will operate, the contacts will close, shorting the condenser C_{11} through the resistance R_{22}. The valve V_9 will commence to take more and more current, tending to bring the associated bridge circuit V_9, R_{19}, R_{13} and R_{12} into balance. The meter M_2 will thus start deflecting at a rate which will depend on the values of C_{11} and R_{22}. When the signal has fallen to a value of 30 db. below its original value the input to V_6 will be equal to the original input voltage; hence the current in V_6 will be equal to that in V_5. The relay bridge circuit will again be in balance and the relay contacts will open, stopping any further movement of M_2. The condenser C_{11} is large enough to hold its charge for a considerable period, and the deflection of M_2 will thus be proportional to the time taken for the signal to fall 30 db. below its original level.

It should be pointed out that the difference in level of 30 db. which is set up across "the relay bridge circuit" has been arbitrarily chosen. By suitable modification of the circuit associated with V_2 this difference in level may be set at

any desired figure, and hence a measurement of the decay of the applied signal to any specified level below its peak obtained. In the instrument as utilized, 30 db. was selected as the maximum value of difference in level which would be required for the measurements contemplated, and a potentiometer (not shown in Fig. 3) was incorporated in the circuit of V_2 to reduce the gain of the circuit associated with V_6 as desired.

In a previous paragraph it was stated that one of the principal difficulties encountered was the impracticability of obtaining a diode-condenser combination in the position of C_5 and V_8 (Fig. 3) which would fulfil the following two conditions simultaneously: (*a*) that the charging time of the combination should be small enough to permit of C_5 charging to a voltage closely simulating the crest signal applied, and (*b*) that the leakage associated with the combination should be low enough to permit of the charge when received being held, with no appreciable leakage, over the period required for the completion of the measurement.

This difficulty was overcome in the following manner. A three-valve circuit was constructed in which the charging circuit of each succeeding valve had increasingly longer charging time and correspondingly longer discharging time; this is inserted at *XX* (Fig. 3). Fig. 4 is a simplified drawing showing the operation of this circuit. The condenser C_5 is made so small that its charging time through V_8 is comparable to the duration of the crest value of the signal. The effect of the signal appearing at *aa* is that a negative voltage proportional to the crest value of the signal is built up on the condenser C_5. This voltage biases back the valve V_3 and consequently there is a fall in voltage across R_3. The potentiometer R_7 will have been adjusted so that it counterbalances the voltage across R_3 to such an extent that the grid of valve V_4 is at its cathode potential. The decrease in voltage across R_3 will result, therefore, in a negative potential being built up across C_9, which is of larger capacity than C_5. The charge on C_5 will remain approximately constant for a long enough time for C_9 to be charged through V_9 to a potential equal in magnitude to the decrease in potential across R_3. C_9 will, in turn, hold its charge for a very much longer period and allow C_{10} to be charged in exactly the same manner to a potential equal to the decrease in potential across R_{11}.

To obtain accurate timing results it was found necessary to use a very sensitive relay, and, to protect this from excessive out of balance currents, two mercury-vapour rectifiers have been connected back to back across it. In order to prevent the relay out of balance current from affecting the reading of the peak meter M_1, the V_5 valve circuit was duplicated, V_5 being associated with the relay, and V_{5a} being associated with the peak meter. Neither of these modifications is shown on Figs. 3 and 4.

Under operating conditions C_{11} will discharge according to the well-known equation $V_t/V_0 = e^{-t/RC}$. The calibration of the meter scale then proceeds from the use of this equation and a knowledge of the relationship between meter deflection and the ratio V_t/V_0.

TESTING

The method utilized to produce the effect of brief charging time and negligible leakage for the diode-condenser combination C_5, V_8 (Fig. 3) has been referred to above. In the early stages

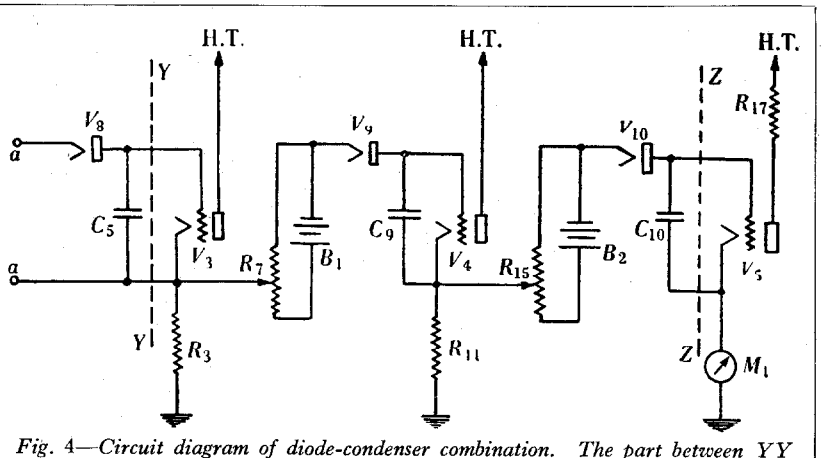


Fig. 4—Circuit diagram of diode-condenser combination. The part between YY and ZZ is inserted at XX (Fig. 3).

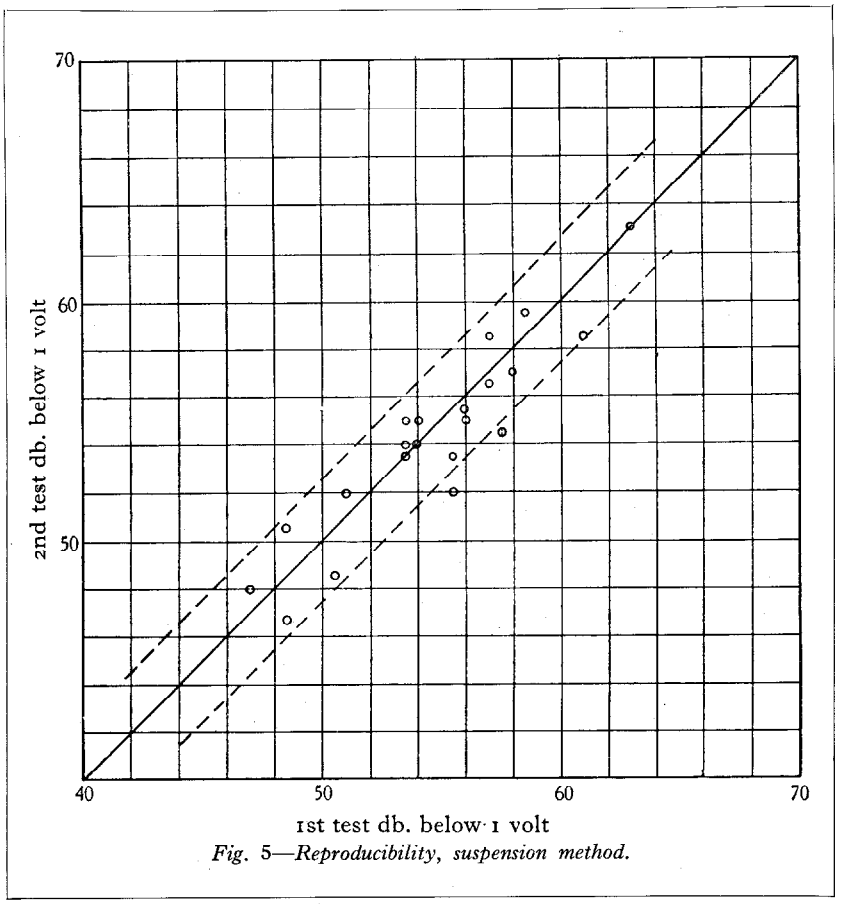


Fig. 5—Reproducibility, suspension method.

of the development of this instrument, operation was attempted utilizing only one stage of diode capacity. As is well known the tail of the voltage-current characteristic of the average diode does not pass through zero, but tails off to some negative value. When this small negative value is negligible in comparison with the amplitude of the applied signal, the effective impedance of the diode is low and the charging time of the condenser short, so that it is then possible to measure the crest value with reasonable accuracy. However, when the amplitude of the signal is small, the negative tail of the diode becomes no longer negligible in relation to the signal amplitude, the diode impedance increases rapidly as the condenser approaches full charge and the charging time is appreciably lengthened, resulting in a crest reading lower than the crest amplitude of the applied signal.

In order to check that the inclusion of the cascade arrangement of diodes and increasing capacities had effectively eliminated this effect, a series of crest amplitude readings of microphonicity were taken on a single valve with varying attenuation between the valve and the instrument. Every effort was made to reproduce as exactly as possible the strength of the impact and other conditions for each test.

TABLE I

CREST AMPLITUDE READINGS OF MICROPHONICITY

Attenuation in circuit	Meter reading	Crest microphonicity in db. below 1 V
40	1.28	45
37.5	1.14	45.5
35	1.08	45.5
32.5	0.94	45.5
30	0.84	45
27.5	0.68	45.5
25	0.53	46
22.5	0.32	46
20	0.20	46.5

The results obtained are tabulated in Table I. The first column represents the attenuation introduced into the circuit on the attenuator shown immediately before the main amplifier in Fig. 2. The second column comprises the readings on the crest meter scale, and the third the crest value of microphonic noise obtained by calibration of the meter reading by means of the calibrating oscillator (Fig. 2). The error, over a range of 20 db., does not exceed ± 0.75 db.

It will be seen from an inspection of Fig. 3 that the amplitude to which the applied signal will decay before relay R operates and the timing circuit ceases to function is determined by the difference in gain between the arm of the circuit containing V_2 and the cascade arrangement of rectifiers and condensers represented in Fig. 2 by V_8 and C_5. Furthermore, if this gain difference is not substantially linear with

TABLE II

TYPICAL RANGE OF MEASUREMENTS OF MICROPHONICITY

Valve No.	Crest mV	Time (sec.) to decay to specified level below crest value		
		9 db.	19 db.	34 db.
1	0.32	0.45	1.31	4.10
2	0.56	0.23	0.34	0.55
3	0.67	0.11	0.13	0.36
4	2.38	0.88	1.80	2.75
5	2.98	0.13	0.29	0.5
6	3.16	0.31	0.86	1.82
7	3.76	0.14	0.22	0.4
8	6.34	0.37	0.89	1.31
9	6.70	0.25	0.38	0.76
10	8.91	1.0	3.5	9.1

amplitude of the crest signal, the operation of the timing device will be itself dependent upon the amplitude of the crest signal. A check test was therefore taken to determine the variation in this gain difference between the two arms of the circuit with signal amplitude. A variation of slightly less than ± 1 db. in gain over the whole range of the crest meter M_1 (about 20 db.) was obtained.

Figure 5 represents a reproducibility curve for peak values of microphonicity obtained on several valves of one type. In obtaining this curve each valve was suspended by a light string in free air and the electrode supplies connected to the valve pins through flexible leads and light clips. The impact was produced by a small rubber-covered hammer pendulum which was allowed to strike the base of each valve once per test after it had described a given arc. Each valve was tested twice, the values in the first test being plotted horizontally, and those in the second vertically. The reproducibility figure of ± 2.5 db. compares with the figure of ± 1.0 db. obtained by Penick,[1] utilizing a more elaborate method of supporting the valve under test in a soft rubber clamp and making connections via mercury cups. It is probable that an appreciable proportion of variation between readings obtained is traceable to noise introduced by movement of the clips on the valve pins, and had mercury cups been used it is possible that a figure of reproducibility approximating closely to Penick's ± 1.0 db. would have been obtained.

As an example of one of the uses to which this instrument can be put, Fig. 6 shows microphonicity decay curves for three different valves of the same type; the vertical scale represents the amplitude of the signal in fractions of the initial crest and the horizontal scale the time in seconds.

Table II has been drawn up to indicate a typical range of measurements of microphonicity taken with this instrument. The valves measured include both filamentary and indirectly heated types and have been selected primarily with the object of indicating the range of values involved.

CONCLUSION

As has been previously stated, this instrument has been designed primarily for use in measuring the microphonicity in small valves, and such

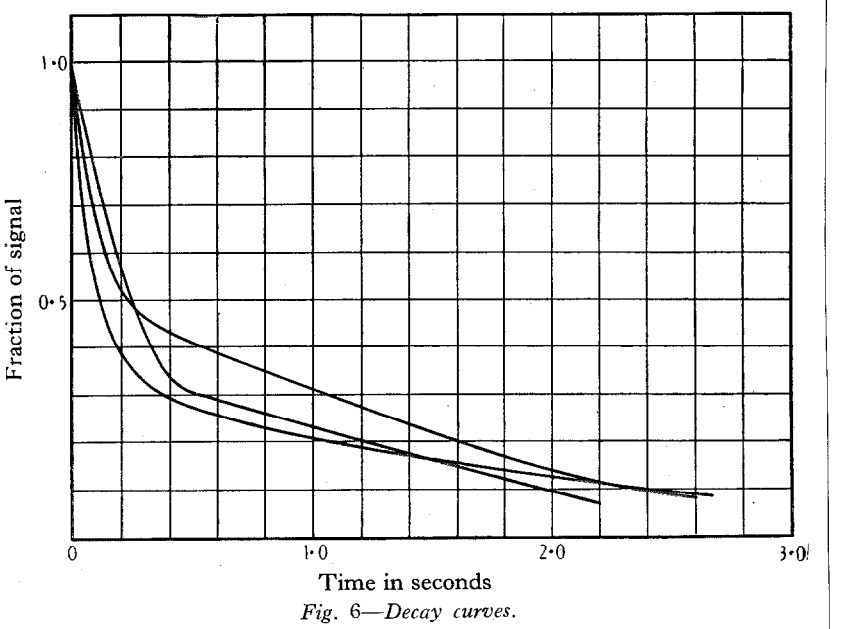


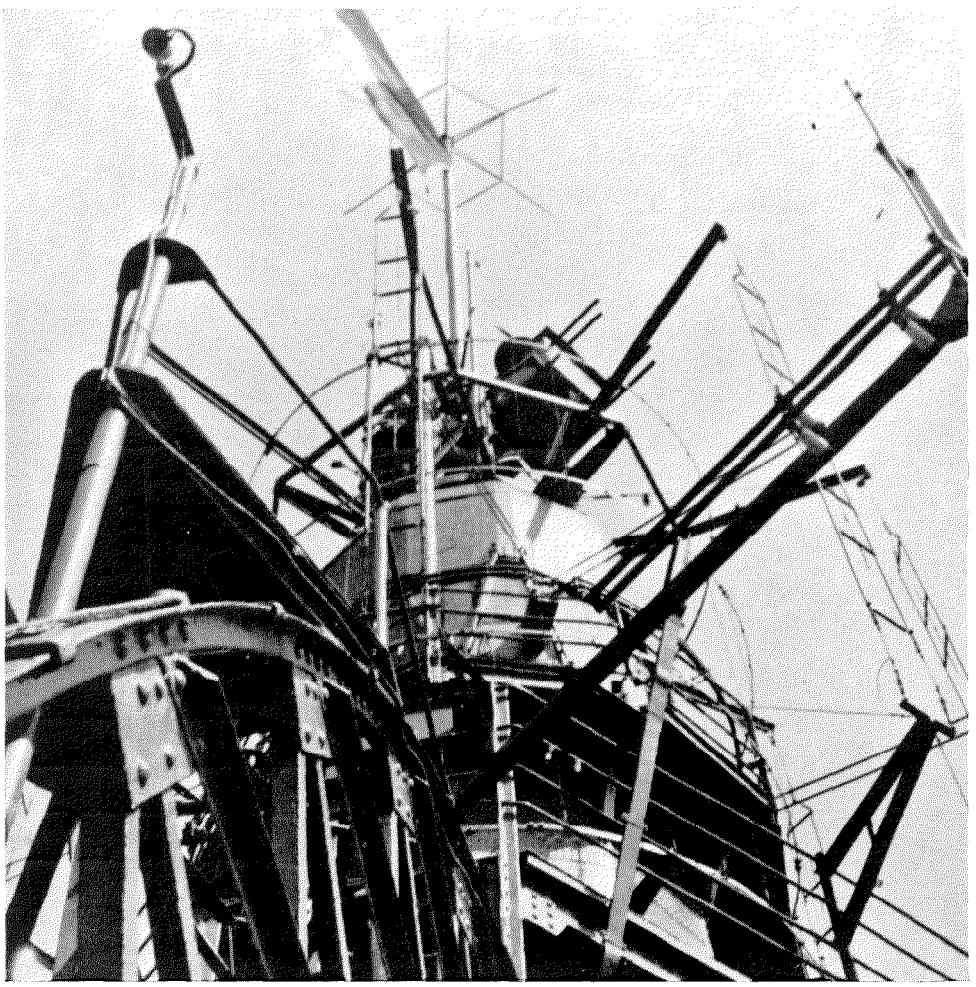
Fig. 6—Decay curves.

[1] Penick, *Bell System Tech. Jour.* 13, p. 614 (Oct. 1934).

tests as have been made have been related to this particular function of the instrument. It is felt, however, that a wider application can be claimed for it, and that it should prove of general use for the measurement of impulse phenomena whenever numerous measurements are necessary and details as to frequency are not of importance. The method described has the outstanding advantage over the amplifier-indicator methods more commonly employed,[2] in that the results are not dependent upon the time response characteristic of the indicator utilized. The crest value and decay of the wave are registered as two distinct readings.

[2] See ref. 1.

Thanks are due to Standard Telephones and Cables, Ltd., for permission to publish the above description. It is also desired to acknowledge the valuable help and advice given by Dr. D. H. Black of the Valve Laboratory of this firm, and the assistance of Mr. A. H. Hooke of the same laboratory.



[*Photo : L. M. T., Paris, France.*

The television antenna and reflectors at the Eiffel Tower, Paris. Until the outbreak of war the sound accompanying the television programme was radiated from the spider antenna above the flag.

The Development of the Civil Aeronautics Authority Instrument Landing System at Indianapolis*

By W. E. JACKSON,
Civil Aeronautics Authority, Washington, D.C.,
A. ALFORD, Associate A.I.E.E., and P. F. BYRNE,
International Telephone Development Company, New York, N.Y.,
and
H. B. FISCHER,
Bell Telephone Laboratories, Inc., New York, N.Y.

I.—HISTORICAL AND GENERAL OBSERVATIONS

Introduction

THE most important problem in flying to-day is that of landing a 'plane under adverse weather conditions. A modern transport 'plane can take off and fly to the vicinity of its destination through nearly any type of weather by means of modern methods of air navigation. The use of the radio range, together with flight and navigation instruments, has made this possible. However, it is extremely hazardous to land aircraft with standard equipment under conditions of low ceiling and poor visibility. This limitation is the reason for the majority of flight cancellations of scheduled air transport.

Under poor weather conditions when landing is permitted by present regulations, the difficulties in landing are still acute. It usually takes a 'plane as much as 15 or 20 minutes to make a landing at a busy airport after being cleared to land by the airport traffic control tower. Concurrently, other 'planes will arrive at the cone of silence over the range station and will have to circle at various altitudes above the radio range and each await its turn to land. Thus it is possible that a 'plane may have to wait an hour or more before it is able to make a landing. When there is little fuel left, the problem becomes even more critical.

Various methods of instrument landing have been attempted; however, the most promising method is one which utilizes radio waves. These waves are used to give lateral guidance, vertical guidance and position "fixes." With the aid of such a system, the entire process of radio instrument landing would take from 4 to 5 minutes. This would help to reduce the traffic congestion at an air terminal and avoid delay in landing.

If air transportation is to compete with other established modes of transportation, it must continue its safety record, and at the same time increase the regularity of scheduled flight operations. This can best be done by the adoption of a standard reliable instrument landing system.

The Civil Aeronautics Authority, realizing this to be the bottleneck of safe flying under conditions of low ceiling and poor visibility, has endeavoured to overcome this difficulty by fostering the development of a suitable instrument landing system. In May, 1938, bids were issued on specifications covering the development of a system which would comply with the rigid requirements of airline service. The specifications were based on the development of the art of instrument landing prior to this date.

Development of Instrument Landing

The history of the development of instrument landing dates back approximately ten years. The first serious work on instrument landing was done by the Bureau of Standards for the Aeronautics Branch of the Department of Commerce in 1928. This system used a visual radio range beacon as a runway localizer, and marker beacons located on the course to inform the pilot of his location with respect to the point of landing. The altitude was controlled by means of a barometric altimeter.

* Paper presented at the A.I.E.E. Winter Convention, New York, N.Y., January 22nd–26th, 1940. Reprinted by permission of the A.I.E.E.

Using this system without the markers, Lieutenant J. H. Doolittle of the Daniel Guggenheim fund made the first successful instrument landing in history on September 24th, 1929, at Mitchell Field.

In 1929, Diamond and Dunmore of the Bureau of Standards proposed a method providing guidance in the vertical plane using a constant intensity glide path. This, along with the runway localizer, gave a three-dimensional system. The runway localizer operated on 278 kilocycles and utilized small multi-turn loops. The glide path operated on 90.8 megacycles and utilized a horizontally polarized array. The marker beacons used long low transmission line antennae operating on 3 105 kilocycles. Installations were made at College Park in 1931; at Newark, New Jersey, in 1933; and at Oakland, California, in 1934. The Newark and Oakland installations used two marker beacons, one at the airport boundary and one along the line of approach, some distance out from the airport. The system was designed to simplify the interpretation of the radio signals by the pilot. A single cross-pointer instrument provided visual runway-localizer and landing-beam course indications and distinctive modulation of the approach and boundary marker beacons was employed. On March 20th, 1933, a completely blind flight was made from College Park to Newark (except for a slight break over Baltimore), on a day when the weather grounded all other aircraft on the Atlantic seaboard. Altogether over a hundred blind landings were made at Newark.

Test flights of the Newark installation indicated the practicability of the fundamental principles of the system and pointed toward further development. Reduction of cost of ground station equipment, elimination of bends in the localizer course due to presence of railroad tracks, power lines, etc., and increased slope of the landing path were desired. Since the localizer course is more susceptible to bending at low frequencies, it was decided to use ultra-high frequencies for the runway localizer as well as the glide path. Tests were made on a combined runway localizer and landing beam operating on a single ultra-high frequency. Further study indicated that horizontal polarization was preferable for the glide path.

In 1933, the Airways Division of the Department of Commerce installed an instrument landing system using the conventional aural type radio range augmented by a radio marker and a sensitive Kollsman altimeter on the 'plane. The 'plane is guided over the radio range cone of silence which is approximately two miles from the field at a specified altitude, and from this point the pilot glides along the radio range course and over the marker to the field at a rate of descent of 400 feet per minute. This system was not considered feasible as an instrument landing system since the accuracy of the barometric altimeter is not great enough to prevent the possibility of undershooting or overshooting the airport. However, this is the system which is used at present for coming into airports under low ceilings.

In 1932 and 1933, the Army Air Corps at Wright Field developed and tested a system utilizing a radio compass for guidance and a sensitive altimeter for checking the altitude over ultra-high frequency markers. Although installations were begun at a number of airports throughout the United States, it was abandoned since a majority of the airlines felt that it did not give sufficiently precise indications for safe commercial use. Its disadvantages were that a barometric altimeter cannot be relied upon, and that a heading established by radio compass is subject to error when cross winds are present. The major advantage was that the system was simple to fly and required a relatively short pilot-training period.

In 1935, the Washington Institute of Technology completed an instrument landing system with elements similar to those of the original Bureau of Standards system. The ground equipment, except the marker beacon, was mounted in an automobile trailer to permit the operation of the system in any direction. The glide path transmitter operated at a frequency of approximately 93 megacycles, its antenna array utilizing horizontal polarization. The localizer operated on 278 kilocycles and transmitted into small multi-turn loops, keying *I* in one loop and *A* in the other. The *A*—*I* visual indication method was discarded in favour of a double modulator type of localizer; that is, modulating one loop at 65 cycles and the other at 86.7 cycles. Originally, two separate power

amplifiers were used, each of which was modulated at one of the two frequencies mentioned above. Each power amplifier independently excited one of the loops. Any change of output in either amplifier caused localizer course variations. To overcome this difficulty, a single power supply was used, and a mechanical modulator connected to the output of the tank circuit was used to supply each of the loops, one at 65 cycles and the other at 86.7 cycles. Thus any change in amplifier output varied both patterns, and the course alignment was not affected. This system was found to work quite effectively. The major disadvantage with this and all other systems using relatively low radio frequencies for the runway localizer is that bends and multiple courses may be obtained when operated in the vicinity of transmission lines, railroad tracks, mountains, and broken terrain.

Transcontinental and Western Air, Inc., in 1935 developed a combination glide path and localizer unit at the Kansas City airport, using a frequency of 85 megacycles. Considerable difficulty was encountered due to variations in the glide path when crossing over a river. It was not until later that it was determined that vertically polarized waves were responsible for the discontinuity.

In 1935, the Bureau of Air Commerce conducted further tests on several modifications of the basic Bureau of Standards system. It was decided to abandon keying systems for course identification in favour of a double modulation system, using 65 cycles in one loop and 86.7 cycles in the other, giving a visual indication. Later it was found that bent and multiple courses which prevailed at the low frequencies could be eliminated by the use of an ultra-high frequency localizer, provided care was exercised in locating the localizer with respect to reflecting objects such as hangars, gas tanks, towers, and buildings.

In 1934 and 1935, Dr. E. Kramar of the Lorenz Company in Germany developed a blind landing system, using the same elements originally used by the Bureau of Standards, that is, glide path, runway localizer, marker beacons, and monitor system. The transmitter was operated on 33.3 megacycles, and excited a vertical half-wave radiator. On each side of the vertical radiator, a reflector was located with a relay at its centre. One reflector was keyed by dots and the other with dashes. By interlocking the dots and dashes, two elliptical patterns were obtained, producing an equi-signal zone which gave two courses. A visual indicator was obtained by means of a rectifier circuit and amplifier. However, the indicator gave a kicking needle indication. The glide path was obtained by following a constant field intensity path. The outer marker was located about 1.9 miles from the airport and the inner marker about 0.19 mile from the airport. Each of the markers transmitted on 38 megacycles, the outer marker modulated at 700 cycles and keyed with dashes four-tenths of a second long, and the inner marker modulated at 1 700 cycles and keyed with dots one-fifteenth of a second long. Each marker gave an aural indication as well as lighting an individual light on the instrument panel. A complete monitoring apparatus was provided.

Through the courtesy of the International Telephone and Telegraph Corporation, one complete set of equipment was installed at the Indianapolis airport for test purposes. These tests indicated that approaches could be made to the field under conditions of poor visibility and low ceilings with good reliability. However, several limitations to this system were found. The radio range could be flown better aurally than visually. Originally, the course was about 6° broad, the width desired in Europe. The pilots in the U.S.A. who flew this equipment felt that the courses were too wide. The course was reduced to 2°, but even further reduction was desired.

Another difficulty observed was the fact that the glide path was smooth down to a point just beyond the end of the runway, at which point the glide path took a definite dive into the ground. The reinforcing steel in the runway was probably the cause for this abrupt change in the glide path. It was later determined that the use of horizontally polarized waves instead of vertically polarized waves could be used to eliminate this difficulty.

The frequencies used, namely 33.3 and 38 megacycles, are not believed satisfactory for a system which would be used universally, since reflection from the Heaviside layer would

occasionally cause interference between stations several thousand miles apart. The Lorenz system, however, had the advantage that the entire system operated on ultra-high frequencies which were free from atmospheric disturbances.

In 1936 and 1937, United Air Lines and the Bendix Corporation constructed and tested a system having the same elements as the Bureau of Standards system, using a combined glide path and localizer transmitter operating on 91 megacycles. Horizontally polarized waves were utilized, and to avoid directional effects a horizontal loop which had a non-directional characteristic was used for the receiving antenna. An automatic pilot was occasionally utilized in making instrument landings. This enabled the pilot to concentrate on the indications of the cross-pointer instrument and follow them more accurately. During 1936 and 1937 approximately 3 000 hooded landings were made on this system.

It is evident that a considerable fund of knowledge has been gained from the numerous systems which have been described. Many notable contributions have been made by Lorenz, the Washington Institute of Technology, Bendix, United Air Lines, Transcontinental and Western Air, Inc., and the various Government agencies which have worked on the problem. Each of the systems has its limitations, although some are better than others for making instrument landings consistently under service conditions.

It was desired to set up specifications which could be used as the basis of a standard instrument landing system for installation on a national scale. The use of ultra-high frequencies was felt desirable, not only because of its freedom from atmospheric and inter-station interference, but because the present overcrowding of the low-frequency spectrum makes it impossible to procure frequencies for use in instrument landing on a national scale. This is in accordance with the general plan of the C.A.A. eventually to operate the majority of air navigation aids on the ultra-high frequencies.

Recommendations of the Radio Technical Committee for Aeronautics

The matter was considered on December 17th, 1937, at the Sixth Meeting of the Radio Technical Committee for Aeronautics, which is a co-ordinating group to recommend policies in regard to radio development as applied to aeronautics. The R.T.C.A. is made up of representatives from the airlines, manufacturers, and various Government agencies. The R.T.C.A. sub-committee on instrument landing made the following recommendations which are quoted here.

(1) *Runway Localizer*

(*a*) The runway localizer shall operate on an ultra-high frequency, preferably in the band 92–96 megacycles, or, if the localizer transmitter is operated as a separate unit, in the band 108–112 megacycles.

(*b*) Straight course, i.e., one which has no bends or multiple courses perceptible to a pilot flying in still air.

(*c*) The difference in the magnitude of the two patterns of the localizer shall be 0.5 db. at 1.5° either side of the centre line as measured with a linear detector.

(*d*) The vertical needle of the cross-pointer indicator shall give a 10° deflection indication for a 1.9° angular deviation from the centre line of the runway.

(*e*) The range of use as a runway localizer should be twenty miles at 3 000 feet.

(*f*) Freedom from interference pattern effects perceptible to the pilot both in elevation and azimuth.

(2) *Glide Path*

(*a*) The glide path shall operate on an ultra-high frequency, preferably in the band 92–96 megacycles.

(*b*) A smooth glide path shall be provided, i.e., one which is free from interference pattern effects perceptible to the pilot when on the localizer course.

(*c*) The system shall be capable of adjustment to provide a suitable glide path.

(3) *Markers*

(*a*) The markers shall operate on 75 megacycles.

(*b*) It shall be possible positively to identify each marker both aurally and visually by modulation and keying. Modulation frequency of the inner marker shall be 1 300 cycles, and that of the outer marker shall be 400 cycles.

(*c*) A normal arrangement of markers would be:

(1) At the normal intersection with the glide path.

(2) Near the boundary of the airport, the exact location to be determined by local conditions.

(*d*) The marker beacons shall have an array adjustable so that when installed in the boundary position the beam will cause useful indications of a visual device within 700 feet either side of the on-course path and for 300 feet along the glide path trajectory. Indications from this marker shall be receivable to an altitude of 2 000 feet.

(*e*) The outer markers shall have sufficient power to accomplish a similar visual indication with the same beam pattern at 2 000 feet.

(4) *Monitor System*

(*a*) Satisfactory means for indicating visually the operation of all equipment shall be provided at a central point.

(*b*) Whatever form of visual indication may be employed shall be smooth in performance and have no irregular characteristics.

(5) *General Characteristics*

(*a*) Frequency of emission of all the elements of the system shall be equivalent to that obtained with a low temperature coefficient quartz crystal.

(*b*) The number of fixed or portable equipments required will depend on conditions prevailing at individual airports.

(*c*) The installation of the foregoing equipment shall not constitute an obstruction to a normal approach to a runway.

(6) The installation of the best known type of approach and runway lights appears to be a most desirable measure in combination with instrument landing facilities.

These are the basic recommendations which were used by the Bureau of Air Commerce in preparing specifications for the development of an instrument landing system. The Bureau of Air Commerce advertised bids, and in June, 1938, awarded a contract to International Telephone Development Company, Inc., to carry on this development. The complete experimental landing system provides facilities for landing in four directions. The equipment was installed at Indianapolis and was accepted by the Civil Aeronautics Authority in October, 1939.

II.—INDIANAPOLIS INSTRUMENT LANDING INSTALLATION

Airport Facilities

Four complete groups of instrument landing

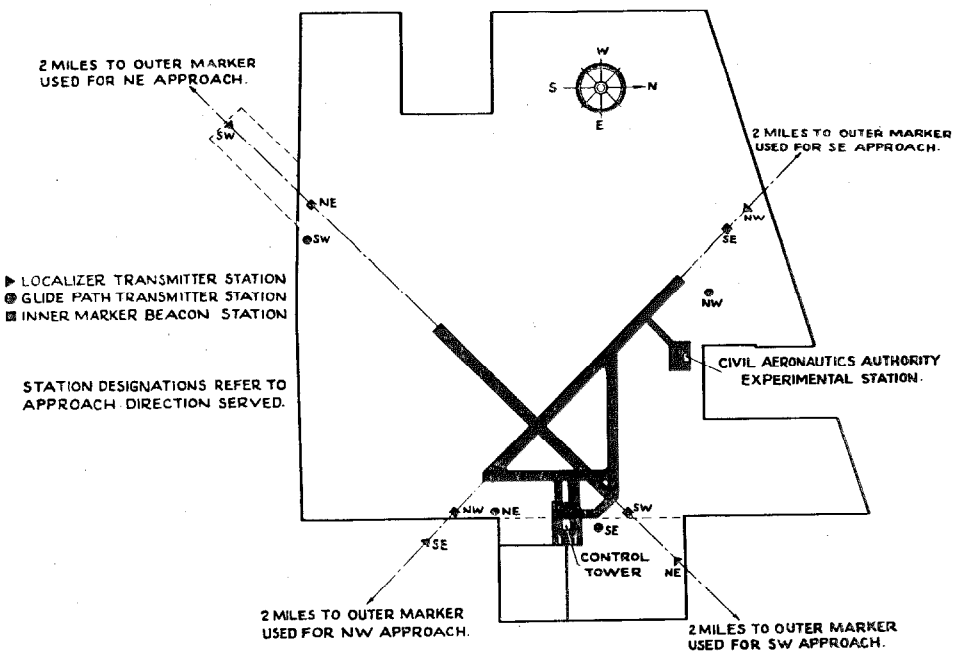


Fig. 1—Indianapolis airport and instrument landing layout.

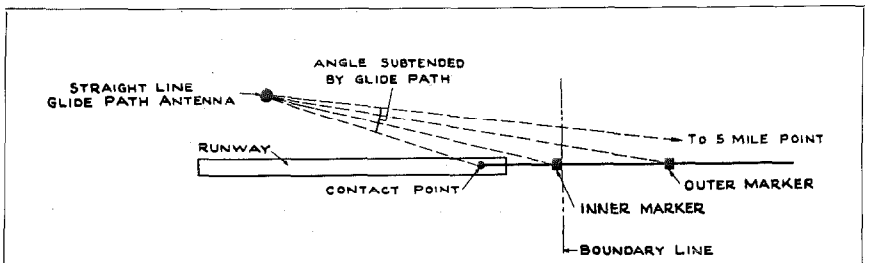


Fig. 2—Location of straight line glide path array.

transmitters have been installed at the Indianapolis Municipal Airport permitting landings in any of four wind directions; north-east, south-east, south-west and north-west (see Fig. 1). Each landing direction requires the installation of four separate transmitters; localizer, glide path, inner marker and outer marker. Each of these installations is connected to the airport control tower by telephone lines, so that the group of four transmitters for any of the landing directions can be turned on to meet local wind conditions.

At the outset of the airport installations, the plans called for four installations of constant intensity glide paths of parabolic shape. This is a well-known type of path obtained by using a simple antenna with, preferably, a broad radiation pattern. The transmitter and radiators may be placed either off the end of the runway or to the side of the runway, but the latter position, if opposite or slightly forward from the far end of the runway, will usually place the contact point at a more favourable location near the beginning of the runway and give the 'plane a longer space in which to land and roll to a stop.

Glide Path Development

Before these installations were complete, however, a system for producing a glide path of the constant intensity type, which is substantially straight from the outer marker to the point of contact, had been developed, and in order to test the new type of path an installation was made at Indianapolis in place of one of the parabolic glide path installations.

Previous to the time of this test installation, belief had been expressed by a number of authorities on the subject that a straight line glide path, as described above, would be ideal because it could be flown with a constant throttle setting, and because the path would clear local obstacles with a greater margin. After a number of flight tests by competent pilots, including air-line pilots, it developed that a path of this shape had too great a rate of descent at the contact point causing the 'plane to land excessively hard, which, if adopted, would cause the passengers considerable alarm and discomfort. After these and subsequent tests, it developed that the ideal shape for a glide path would be a substantially straight line from a point 1 500 feet at five miles from the airport to the airport boundary, whence the path would become slightly parabolic in shape, intersecting the runway surface at approximately one degree. The path would pass over the outer marker, two miles from the airport boundary, between 500 and 700 feet elevation. A path of approximately this shape has since been produced at Indianapolis.

The method by which constant intensity paths of various shapes, including the straight line and improved arrangements, have been produced involves locating the glide path antennae at a considerable distance to one side of the runway, and forward along the runway

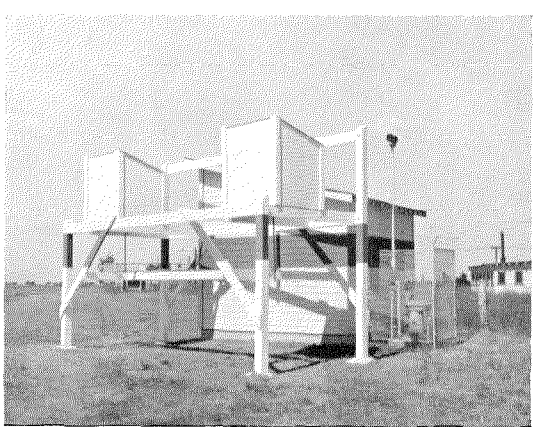


Fig. 3—Straight line glide path building and antenna array.

so that the various points along the glide path appear in different directions as viewed from the glide path antenna as shown in Fig. 2. In this way, the angle subtended by the glide path, as viewed from the glide path antenna, is opened up so that it becomes possible, by making the antenna directive in the horizontal plane, to proportion the radiation along the extent of the glide path and control the height of the path at the various points. Thus, for example, by sending less energy towards the outer marker the glide path may be made higher there, and by sending more energy towards the five-mile point, the path may be lowered at this point.

In Fig. 3 is illustrated the antenna system which produced an approximately straight line glide path from the outer marker to the point of contact. This system consisted of two ultra-high frequency loop antennae spaced approximately one wave-length apart, fed 180 degrees out of phase with unequal amounts of energy, and located in front of a reflecting screen. The antenna system which produces the more suitable type of path involves the use of a greater number of elements to produce the required field pattern shape.

The installation for the "straight line" glide path is located 1 350 feet off the centre line of the runway and 1 050 feet forward from the end. The parabolic curve glide path installations are located 400 feet off the centre of the runway and approximately opposite the far end of the runway.

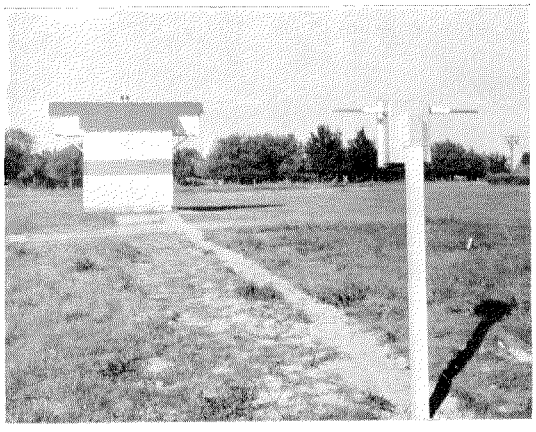


Fig. 4—Localizer building and field monitor.

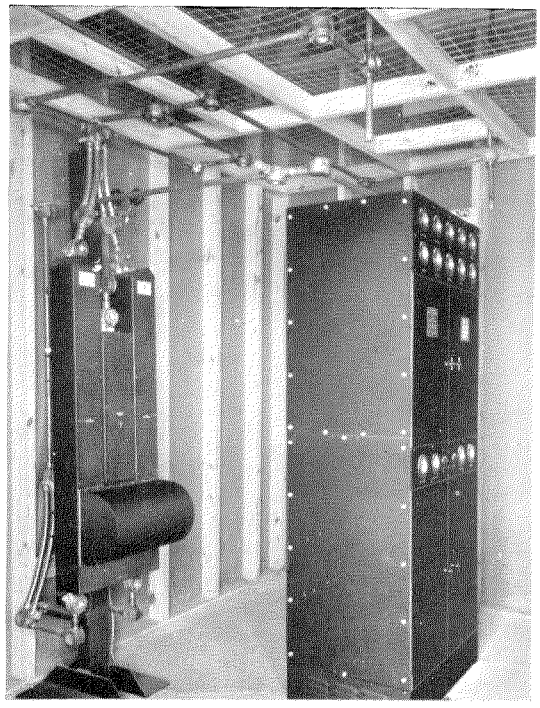


Fig. 5—Interior of localizer building showing transmitter and mechanical modulator.

Localizer Installation

The localizer installation consists of a wooden building approximately 12 feet square located off the end of the runway, and a localizer monitor unit situated about 100 feet in front of the building on a line with the runway (see Fig. 4). Inside the building are the localizer antennae mounted under and close to the apex of the roof, the localizer transmitter, the modulator unit, the voltage regulator, the service switches and telephone (see Fig. 5). A large wire screen is placed horizontally across the inside of the building just below the antennae primarily to minimize the radio frequency field in the lower part of the building, in order to avoid induction of currents in, and consequent re-radiation from, the irregular objects below the screen. The transmitter and modulator units are mounted on the floor underneath the wire netting.

Localizer Antennae

Early work has shown that fewer common objects around an airport reflect horizontally polarized waves to the same extent as vertically polarized waves. Consequently, it is to be expected that, in general, cleaner patterns can

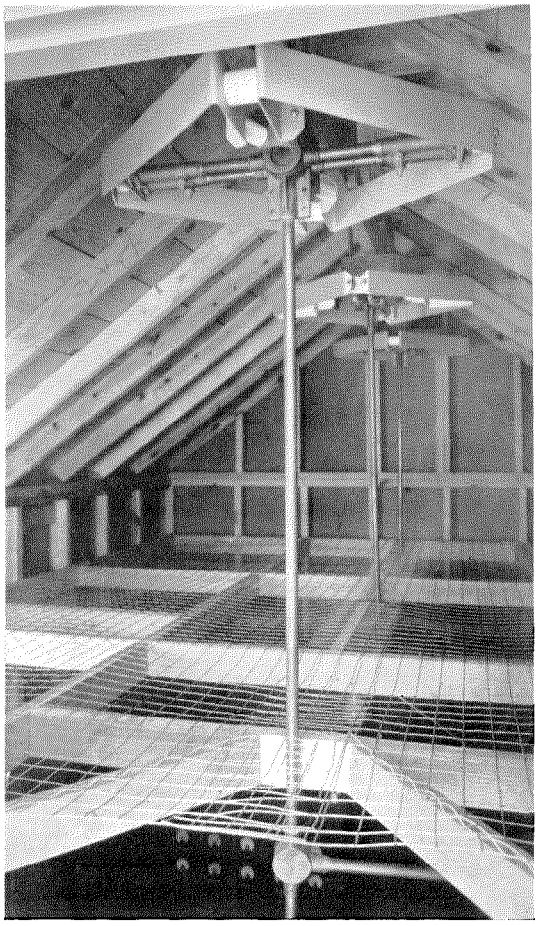


Fig. 6—Localizer array of U.H.F. loop antennae.

be obtained with horizontally polarized waves than with vertically polarized waves. This point has been strikingly demonstrated at Indianapolis by making direct comparison between two localizers placed in the same location and having identical radiation patterns, the only difference being that one was horizontally polarized and the other vertically polarized. Both of these patterns were fairly broad, so that radiation was sent out to various surrounding objects which could re-radiate and superimpose their re-radiated fields on to the primary field. As a result of the addition of these fields, bends appeared in the course. The amplitude of the bends was considerably greater with vertically polarized radiation than with horizontally polarized radiation. In addition, recordings made of the field radiated in other directions led to the same conclusions.

These tests showed that a greater number of objects in the field were more effectively reflecting the vertically polarized waves than the horizontal waves. These objects included a hangar, a power line, several buildings, a group of trees and several fences. While it is, of course, possible to find objects, for example horizontal wires, which will reflect horizontally polarized waves more effectively than vertically polarized waves, the fact remains that fewer of the common objects which are confronted in practice reflect the horizontally polarized waves to the same extent that they reflect the vertically polarized waves.

In addition to the use of horizontal polarization, it is very important to use pure polarization. If both vertical and horizontal components are present it is possible, with certain arrangements of the 'plane's receiving antenna, that either the main course appears displaced from its true location or, which is even worse, a new apparent course is obtained which would not be present if the antenna were operating properly. This situation might arise if the receiving antenna were injured or improperly installed. It is apparent that the presence of vertical components with the horizontal components may actually be dangerous.

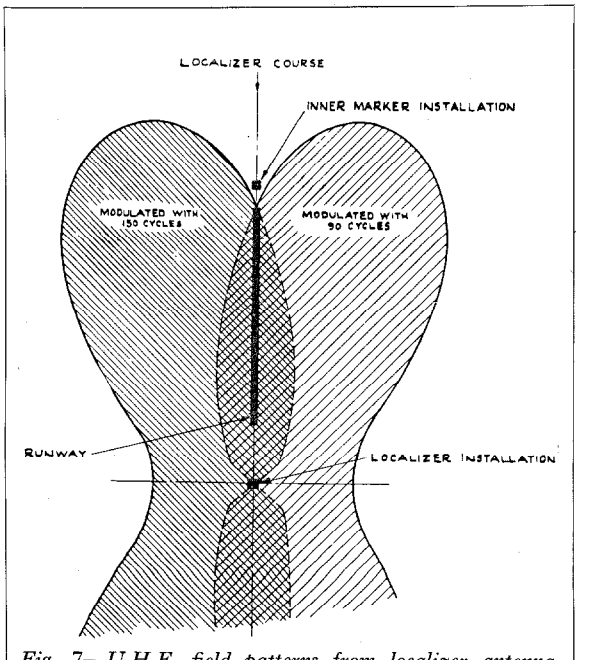


Fig. 7—U.H.F. field patterns from localizer antenna.

In addition to this, the presence of vertical components leads to an effect which has been referred to as "pushing" of the course. This is noticed when the 'plane is manœuvring to locate the course. The apparent course appears to be displaced depending on the orientation of the 'plane and when the pilot attempts to turn into the course, he finds that it has disappeared. This effect, under some conditions, may amount to pushing the course as much as one mile off the true course at a distance of ten miles from the localizer, while with pure polarization this effect does not appear.

In order to realize the advantages of horizontal polarization, a special antenna element was developed primarily for localizer use which radiates pure horizontally polarized waves in all directions. The field pattern in the horizontal plane is circular and in the vertical a figure of eight with zero radiation upward. This element is known as the ultra-high frequency loop antenna, three of which are shown in Fig. 6 mounted under the roof of a localizer house. When these loop antennae are used in the localizer antenna system, the maximum vertical component throughout the entire field pattern does not exceed 5% of the horizontal component at that point.

By using three of these elements arranged in line at right angles to the axis of the runway, it is possible to obtain a two-course radio range with patterns similar to those illustrated in Fig. 7. The course attained will be relatively broad, and there will be a substantial amount of radiation directed toward surrounding objects. While this pattern is suitable for relatively clear locations, and would be excellent for use as a radio range pattern where the antenna is located well above surrounding objects, such a pattern may not be entirely satisfactory in an airport in the presence of hangars, buildings, and other reflecting objects. For this application, a pattern which results in much sharper courses and much less off-course radiation is desirable in order to secure freedom from course irregularities. In order to produce a more suitable pattern, more radiators are required. Radiating systems with 5, 6 and 7 elements have been used at Indianapolis, and a localizer building with a 5-element array is shown in Fig. 4. The fourth and fifth elements, placed in line with the centre three, are located in the extensions built on to the ends of the buildings. These extra elements are parasitically driven.

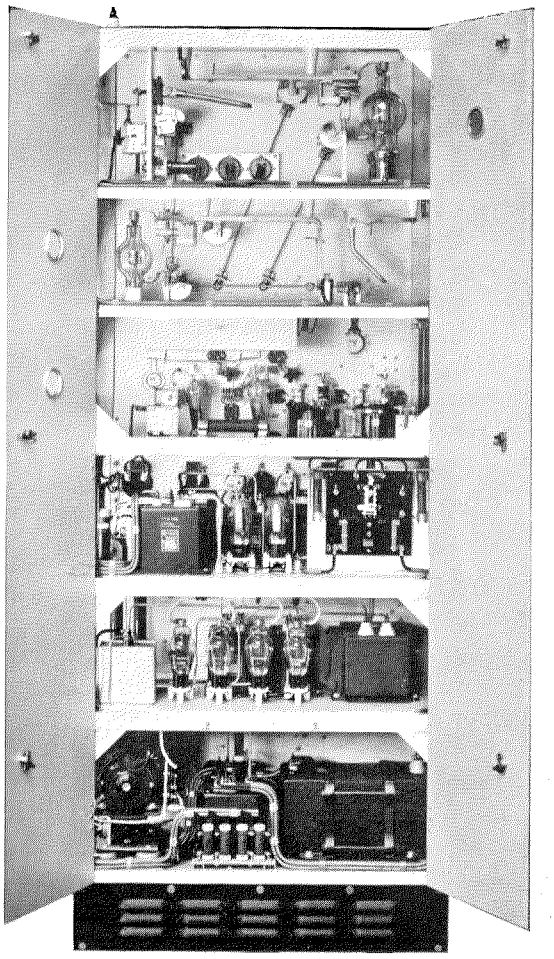


Fig. 8—Interior of localizer transmitter from rear.

With the antenna systems outlined above, the courses produced at Indianapolis are entirely free of multiple courses; the bends in any of the courses do not exceed 0.15°, and these bends could be further reduced if not completely eliminated by the results of more recent localizer development work. However, since these bends are so small that they are barely perceptible on the cross-pointer instrument, further improvement is not warranted. Experience has shown that unless bends exceed 0.18° or more in amplitude, they are not objectionable.

The two overlapping patterns in the radiated field give side indications of the guiding path which may be easily located at distances of

twenty miles from the airport boundary. The course can be picked up at a distance of seventy miles from the airport while flying at an elevation of 4 000 feet or over.

The four localizers installed at Indianapolis have various degrees of sharpness; the sharpness being the ratio in db. of the field strength of the 90-cycle and 150-cycle patterns at 1.5° off course. The sharpest of the patterns exceeds 6 db., while others range in the neighbourhood of 2 db.

The courses are straight and reliably stable; maintaining their alignment to within one-tenth of one degree under all normal weather conditions.

Localizer Modulator

The localizer modulator is a mechanical modulator constructed to furnish modulated radio frequency energy for the two overlapping localizer field patterns (see Fig. 7). Inasmuch as the position of the localizer course is largely dependent on the relative amplitude of the two field patterns, the mechanical type of modulator is used because of its inherent stability and freedom from ageing such as would be found in vacuum tube modulators. The modulator divides the output of the localizer transmitter equally, modulating one-half of the power at a 90-cycle rate and the other half at a 150-cycle rate. Bridge networks, built into the modulator, divide the transmitter output and prevent crosstalk or inter-action between the 90- and 150-cycle modulator sections, as well as feed the carrier and side bands with proper amplitude and phase to the three main localizer antenna radiators. Controls are provided in the unit for adjusting the position of the course with respect to the runway.

The normal percentage of modulation obtained with the mechanical modulator ranges between 95 and 100%, although any percentage below this range may be obtained. The radio frequency modulated envelope shape from either the 90 or the 150 modulator section, independently, is such that the audio frequency distortion as measured in the output of a non-distorting detector does not exceed 10%, and normally ranges close to 6%. Cross modulation between channels does not exceed 4%; this percentage being the per cent. of 150-cycle modulation appearing in the 90-cycle channel or the per cent. of 90-cycle modulation in the 150-cycle channel.

Localizer Transmitter

The rated power output of the localizer transmitter is 300 watts, unmodulated, on a frequency of 109.9 megacycles (see Figs. 5 and 8). The transmitter operates from 220 volts, single-phase, 60 cycles, and receives its entire power of approximately 2 000 watts through a voltage regulator which holds the voltage within plus or minus 1% for all normal feeder variations.

The transmitter is crystal-controlled from a 4 579.17 kilocycle temperature-controlled crystal, and the crystal oscillator tube functions also as a frequency multiplier. Three frequency multiplier stages follow the crystal oscillator tube to produce the output operating frequency. In order to raise the output of the final multiplier to the rated transmitter output level, the last two stages are push-pull amplifiers operating on the output frequency.

In order to ensure reliable operation of the transmitter during cold weather, the

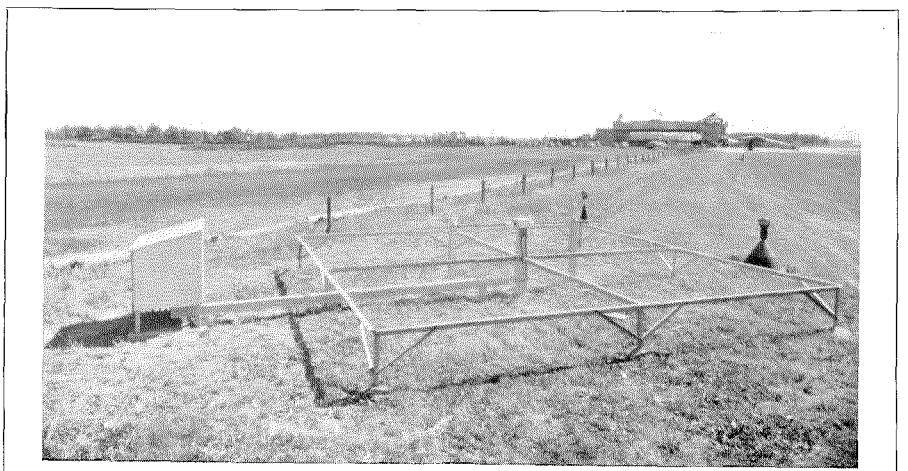


Fig. 9—Inner marker installation showing transmitter house, counterpoise and antennae.

ventilating and modulator motors are electrically heated to maintain their bearing temperature above the freezing point. Special lubricating oils which remain viscous to very low temperatures are used in the bearings of contactors and other mechanical parts. The circuit breaker dash pot also contains a special fluid which remains viscous down to the low temperatures in which this unattended equipment will be required to operate.

The transmitter contains control relays so that it can be operated from the tower. A monitoring arrangement is provided which operates in conjunction with the field monitor (described below) so that the position of the course with respect to the runway, as well as the strength of the radiated patterns, may be ascertained from meters located on the front of the transmitter.

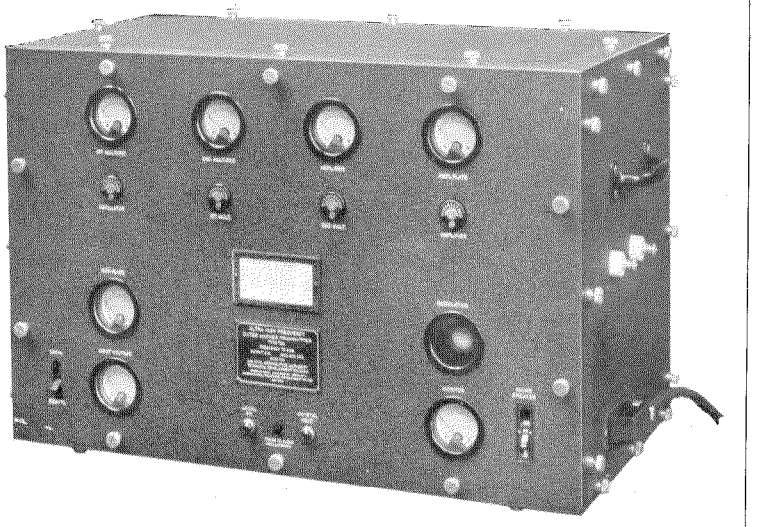


Fig. 10—Marker transmitter.

Localizer Monitor

In order to monitor the operation of the localizer installation from the control tower, a "localizer monitor" is installed about 100 feet from the localizer building on a line between the latter and the end of the runway, as shown in Fig. 4. The monitor contains an insensitive receiver which operates a course-indicating instrument in the localizer transmitter and a duplicate instrument in the control tower desk. The indication is the same as would be obtained from the vertical needle in the cross-pointer instrument of an airplane were it stationed at the monitor site. The monitor facilities are adjusted to indicate the extent of off-course deviation in fractions of one degree.

Underground Localizer Installation

In one localizer installation where insufficient space was available off the end of the runway, making it necessary to locate the building close to the runway, the equipment was installed in a concrete pit below ground level to lower the building height and reduce the hazard. The floor of the pit is approximately 8 feet below ground level, and accommodates the same equipment as the normal building with the addition of the high tension distribution transformer. The pit is covered by a wooden structure shaped like the roof of the normal localizer building. Under the roof are the localizer antennae situated about 6 feet above ground level (see Fig. 6). The only additional equipment required for this installation is a sump pump to protect against the possibility of water gathering in the pit.

The lowering of the antenna system close to the ground in this installation in no way affected the operation of the localizer, except that the distance from the airport at which the course could be picked up was somewhat reduced.

Glide Path Installation

There are two types of glide path installation at Indianapolis; the installation for producing a parabolic path, and the installation for producing a straight line path. In both installations, the transmitter is located inside a wooden building approximately the same size as the localizer house. The antennae for the straight line installation are erected outside the building, as shown in Fig. 3, while the antennae for the parabolic installation are erected inside the building. The antennae for the latter installation consist of two half-wave radiators fed in

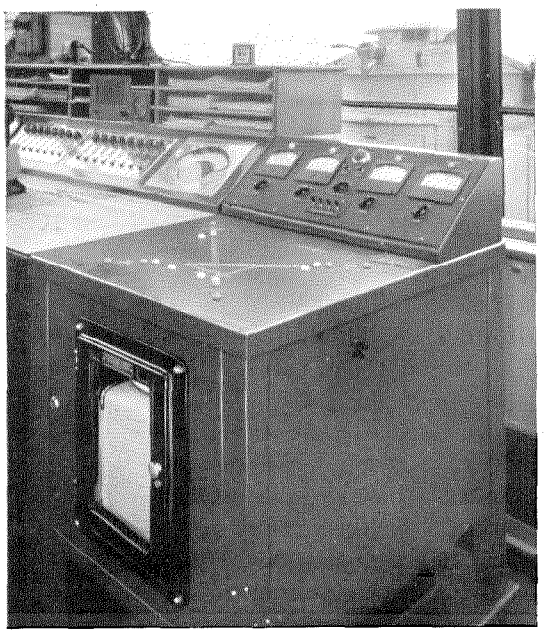


Fig. 11—*Monitor and control desk in airport control tower.*

phase and mounted horizontally end to end along one wall about 10 feet above the floor. A vertical wire screen the length of the building, and extending from the roof to about 5 feet above the floor, is located one-quarter wavelength behind the antennae. The building is so oriented that the maximum of radiation is directed towards the approach end of the runway served by the installation.

Glide Path Transmitter

The glide path transmitter is identical in construction and size with the localizer transmitter, except that it operates on a frequency of 93.9 megacycles; requires a crystal with a frequency of 3 912.5 kilocycles; the output is modulated at 60 cycles and no monitoring facilities for use with a field monitor are required. The 60-cycle modulation is accomplished by applying 60-cycle plate voltage to the two tubes in parallel in the output stage. Means for adjusting this voltage are provided in the transmitter, so that the field strength along the constant intensity path can be adjusted. In general, the field strength for all constant intensity paths must be identical at a given height over the outer markers.

Marker Installation

Marker beacon installations consist, mainly, of a small waterproof aluminium house situated alongside a wire screen counterpoise 20 feet square, over the centre of which are located two half-wave radiators mounted end to end (see Fig. 9). The waterproof house is approximately $3\frac{1}{2}$ feet long, $2\frac{1}{2}$ feet wide and 3 feet high, with a sloping roof, mounted above ground on angle legs. The wire screen is supported by an angle framework and also mounted above ground on angle legs. Concentric gas-filled transmission lines joining the waterproof house

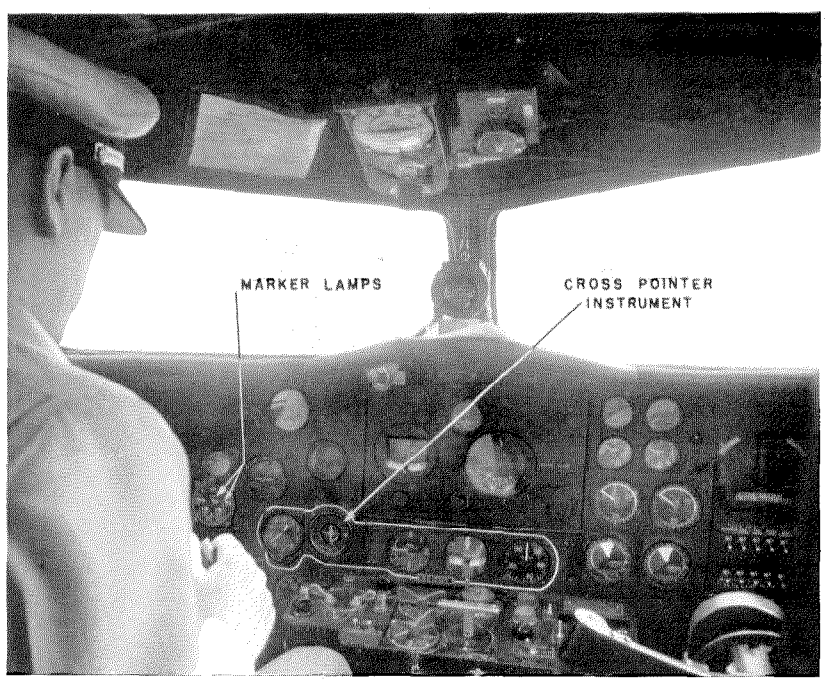


Fig. 12—*Airplane instrument panel showing cross-pointer instrument and marker lamps.*

with a matching box located between the ends of the radiators, are used for feeding the marker transmitter output to the radiators.

The marker transmitter and its associated line voltage regulator, along with a telephone box and cable terminal, are mounted inside the waterproof house. The transmitter is mounted on a sliding shelf so that it can be easily pulled out of the small house for servicing while in operation. Whilst in this position, complete access may be had to all parts of the transmitter.

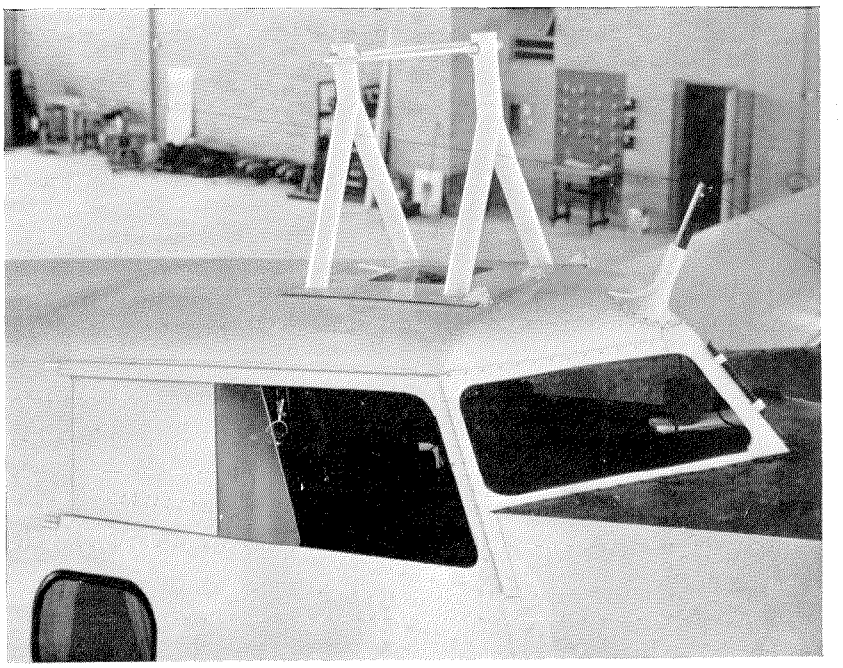


Fig. 14—U.H.F. loop antenna for reception of localizer and glide path indications mounted on top of airplane.

The installation is oriented so that the end-on radiators are parallel to the centre line of the runway. In this way, the major axis of the elliptically-shaped pattern in the horizontal plane is at right angles and situated over the runway centre line. In outer marker installations, the radiators are located one-quarter wavelength over the counterpoise, whereas the radiators in the inner marker installation are only one-eighth wavelength above the counterpoise. In the latter installation, where the equipment is located close to the end of the runway, the maximum overall height of any piece of equipment does not exceed 4 feet.

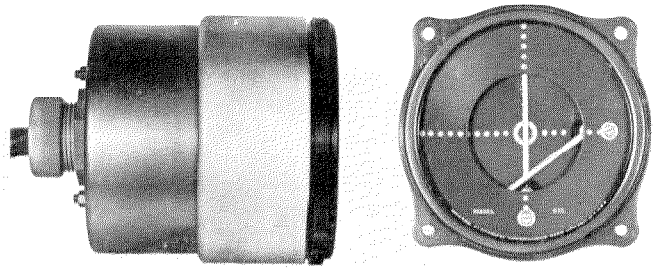


Fig. 13—Cross-pointer instrument.

Marker Transmitters

The marker transmitters are approximately 18 inches high, 27 inches wide, and 15 inches deep (see Fig. 10). The transmitter is built as a single portable unit, containing all radio frequency and power supply circuits, audio oscillator and modulator, keying and monitoring equipment. All electrical connections are made to the transmitter through easily-removed plugs. The rated output of the transmitter is 5 watts on 75 megacycles, and the carrier frequency is stabilized by a crystal on 4 166.7 kilocycles.

In the R.F. stage line-up, the crystal oscillator

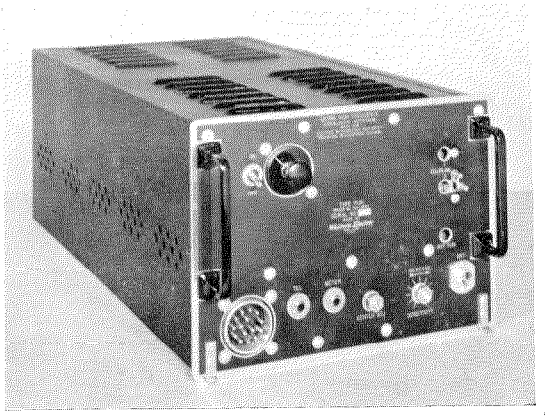


Fig. 15—Runway localizer receiver.

tube serves as a frequency multiplier, in addition to its duty as an oscillator. Following the oscillator stage are two frequency multiplier stages and one amplifier stage operating on carrier frequency. The radio frequency carrier in the outer marker transmitter is modulated with 400 cycles, keyed so as to produce equal length pulses at the rate of two per second. The inner marker carrier is modulated with 1 300 cycles, and keyed so as to produce equal length pulses at the rate of six per second. Keying of the modulating voltage is accomplished with electrical contacts mounted on a small electric motor located within the transmitter. The motor operates on a low D.C. potential, and requires a relatively small amount of power from the D.C. plate supply source. The motor is especially designed and constructed for reliable operation in the extreme temperature ranges of minus 40° C. to plus 75° C. in which the marker equipment is required to operate.

The transmitters are intended for remote operation and monitoring from the airport control tower. A monitor unit is built into the transmitter which returns a signal to the tower, where correct operation may be observed. The monitor signal returns over one pair of telephone wires, which also carries D.C. for controlling the transmitter, as well as voice and ringing frequencies for telephone service between tower and marker.

Equipment Control

The airport tower is equipped with a monitor and control desk by which the tower operator can select landing directions at will (see Fig. 11). Calibrated instruments and signal lamps give quantitative and qualitative indications of the various transmitters. These indicators are operated by the transmitter monitors and field monitors, and indicate the position of the glide path, alignment of the runway localizer courses and correct operation of the marker transmitters. If the output of any transmitter falls below a predetermined level, visual and aural alarms indicate the trouble. The alarms also operate if the runway localizer course should shift sufficiently to cause an airplane following the course to land off the runway.

The top of the desk is a pictorial map of the airport showing clearly the position of the various runways and transmitters. Coloured signal lamps represent each transmitter, and are lighted when the corresponding transmitters are in operation. At the head of each runway is located a translucent miniature airplane which is lighted to show the direction of approach.

A pen recorder incorporated in the unit supplies a permanent record of the operation of the system. This instrument records the elapsed time a runway has been in operation, the output of each transmitter, and the localizer course alignment. This multiple function is accomplished with a single pen by recording successive short intervals of the operation of each unit.

Telephone wires emanating from this desk and terminating at the field installations provide means for starting and stopping the equipment, as well as monitoring the operation of the equipment. Telephone type cables were laid on the airport property to provide wire connections between the control tower and the localizer installations, glide path installations and the inner marker installations. Telephone lines from the control tower to the outer marker positions are provided by the local telephone company.

The outer markers, which operate from the 115-volt mains, receive their power from the local power company mains closest to the installation site. Power for operation of the localizer, glide path and inner marker equipment is fed from the control tower at 2 300 volts in high tension cables which are also laid on the airport property.

III.—AIRPLANE INSTRUMENT LANDING RECEIVING INSTRUMENT

Airplane Facilities

In order to utilize the various signals sent by the transmitters on the ground, the pilot observes the position of the two needles of a cross-pointer meter, and also notes when either of two marker indicator lamps is illuminated. A view of a typical airplane instrument panel, showing the location of the meter and the lights, is shown in Fig. 12, while a view of the cross-pointer meter is shown in Fig. 13. The two needles of the meter are operated by the outputs of two special radio receivers which receive the transmissions from the glide path and the runway localizer transmitters. Three indicator lights are operated from the output of the marker beacon receiver. Two of these are used in connection with the instrument landing system, and show when the 'plane is over the outer marker and the inner marker. The other light shows the pilot when he is over one of the fan markers located at important points along the airway, or over the towers of the radio range transmitting station which are equipped with cone of silence markers. The marker beacon receiver is in continuous operation during the course of a flight as it has other important uses in addition to its part in the instrument landing procedure; while the glide path and the localizer receiver are used only while making an instrument landing.

One antenna, shown in Fig. 14, is used to receive both the runway localizer and the glide path signals. This antenna is of the horizontal ultra-high frequency loop type tuned to 93.9 and 109.9 megacycles, and is designed to receive horizontally polarized waves. As mounted on a 'plane it has an essentially circular horizontal reception pattern. A horizontal dipole antenna mounted underneath the 'plane is used to receive the marker beacon signals.

In common with all receivers designed for use in airplanes, the receivers for this service must be compact, lightweight and rugged. In addition to these basic requirements, the demands of this new service impose a very rigorous requirement for stability of operation. In spite of wide variations in ambient temperature, humidity and primary power supply voltage, the characteristics of the receivers must remain substantially constant, as any variation during the landing manœuvre will result in an apparent change in the path. The principal characteristics of these three receivers are described in the following sections.

Runway Localizer Receiver

This receiver, a front view of which is shown in Fig. 15, employs a superheterodyne circuit, and has a crystal-controlled beating oscillator for operation at a fixed frequency of 109.9 megacycles. As explained in a previous section of this paper, the runway localizer transmitter is continuously modulated with 90-cycle and 150-cycle tones. In order to ensure that there is no apparent error in the course, it is essential that the relative gain for these two modulation frequencies remains constant.

A simplified schematic of the circuit arrangement is shown in Fig. 16. The radio frequency input is applied to the control grid of the first detector *V*1, while the beating oscillator voltage

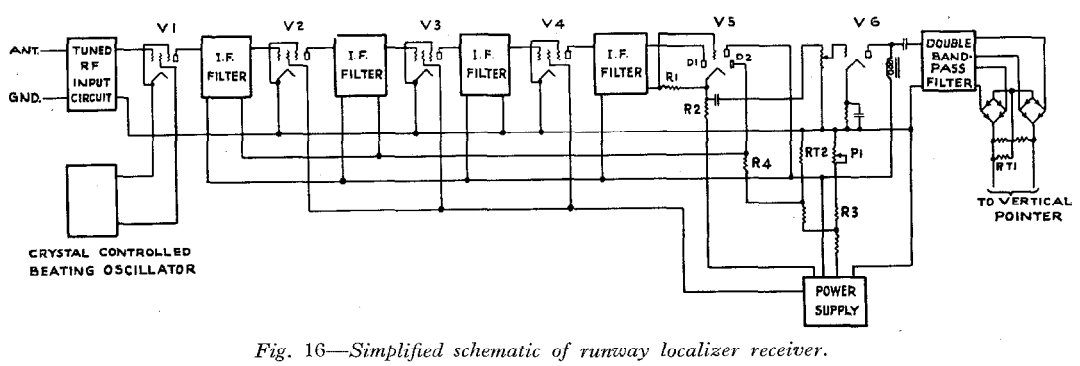


Fig. 16—Simplified schematic of runway localizer receiver.

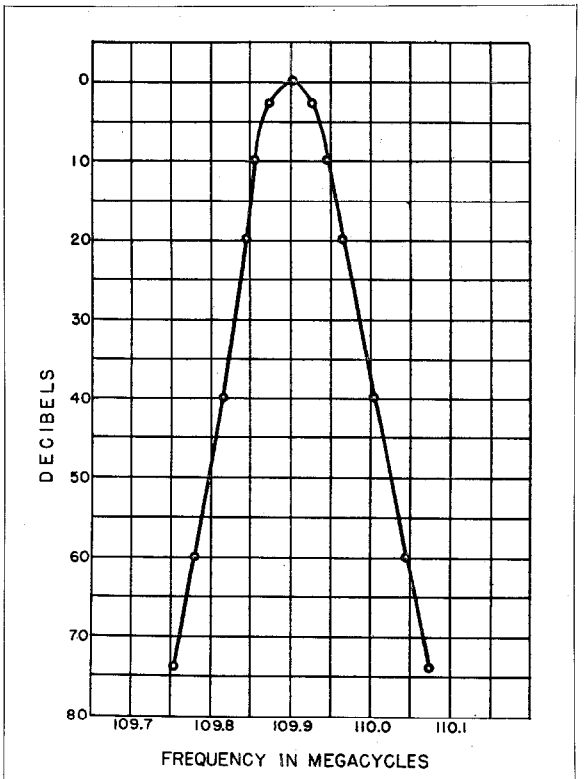


Fig. 17—Selectivity of runway receiver.

is impressed on the screen grid. The output of the tube is fed to a three-stage intermediate frequency amplifier system operating at a frequency of 10 megacycles, and having four inductively - coupled, double - tuned circuits. These highly selective circuits, in connection with the tuned input circuit, provide the radio frequency selectivity shown in Fig. 17. The output of the intermediate frequency system is fed to diode $D1$ of the detector tube $V5$ and the audio frequency voltage developed across the diode load resistance $R1$ is impressed upon the grid of the triode section of this tube. The amplified audio frequency output appearing across the cathode resistance $R2$ is applied to the grid of the second audio frequency amplifier tube $V6$. The plate of this tube is connected to the double band pass audio frequency filter which separates the 90-cycle and the 150-cycle components of the signal modulation. Two copper oxide rectifiers containing selected and balanced units are used to rectify the two audio voltages for operating the vertical needle of the cross-pointer meter. In order to compensate for any unbalance which exists in the audio frequency filter due to temperature variation, a temperature-sensitive resistance element $RT1$ having a high negative temperature coefficient is connected across one of the rectifier load resistors.

Signals received from the localizer transmitter vary over a wide range of intensity as the 'plane approaches the airport, and therefore a wide range automatic volume control system is required to maintain a nearly constant audio output. The control grids of the intermediate frequency amplifier tubes $V2$ and $V3$ obtain their normal minimum bias from the drop across resistance $R3$ and the gain adjusting potentiometer $P1$. When a signal is received, the automatic volume control circuit increases the bias applied to the control grids in proportion to the signal level applied to the receiver. An increase in signal results in an increased negative bias on the grid of $V5$ which reduces the plate current of the triode section flowing through the cathode load resistance $R2$. The voltage drop across the resistance is reduced and the potential of the cathode of $V5$ is lowered until it becomes more negative than the normal bias applied to diode plate $D2$ of the $V5$. At that point, diode $D2$ draws current which flows through $R4$, and the voltage drop across that resistor is added to the normal bias to secure automatic volume control on the grids of $V2$ and $V3$. Bridged across the minimum bias resistor is another temperature-sensitive element $RT2$ which causes the bias voltage appearing across $R4$ to decrease with increasing ambient temperature in order to maintain nearly constant receiver gain.

Glide Path Receiver

This receiver, which is quite similar in appearance to the localizer receiver, is a crystal-controlled superheterodyne set operating at a fixed frequency of 93.9 megacycles, intended for the reception of tone-modulated continuous wave transmission from the glide path transmitter located at the airport. It is essential that the overall receiver gain remain practically constant for a wide range of variation of temperature, battery voltage and humidity, as the determination of the path of descent of the airplane is based upon a fixed receiver sensi-

tivity. In order to meet these requirements, voltage regulators, ballast lamps and temperature-sensitive resistance elements are used within the unit.

Marker Beacon Receiver

This receiver, a photograph of which is reproduced in Fig. 18, is a crystal-controlled superheterodyne unit operating on a fixed frequency of 75 megacycles, and arranged to provide both visual and aural indication of the signal being received.

Visual indications are provided by three indicator lamps mounted on the airplane instrument panel. As previously pointed out, two of the lamps are used for instrument landing purposes, and the third is used to indicate the position of fan markers and cone of silence markers. The modulation frequency for the latter is 3 000 cycles; the inner and outer markers employ 1 300 and 400 cycles respectively.

Three filters are provided in the receiver to separate the 400, 1 300 and 3 000-cycle modulating tones. The output of each filter is applied to a copper oxide rectifier, and the resulting current flows through a direct current winding on a saturable core reactor. An alternating current winding, also on this core, is in series with the indicating lamp and the lamp voltage supply source. When the impedance of the reactor is lowered by saturation due to the direct current, sufficient alternating current flows to operate the lamp.

IV.—RESULTS

After the equipment was installed, a meeting of the Radio Technical Committee for Aeronautics was held at Indianapolis in September, 1939, for the purpose of demonstrating the equipment to the members, and to determine if the system developed was suitable for airline use. The committee unanimously agreed that the General Radio Technical Committee for Aeronautics specifications had been complied with fully. However, in view of the improvement in the glide path which was demonstrated as being entirely feasible, the airline pilots present felt that neither a straight line nor a curved glide path was entirely satisfactory, and made the following recommendations as to the most desirable shape of the glide path when on the localizer course:

> "A glide path intersection shall be obtained at 1 500 feet altitude at a distance of six miles from the transmitter end of the runway. The

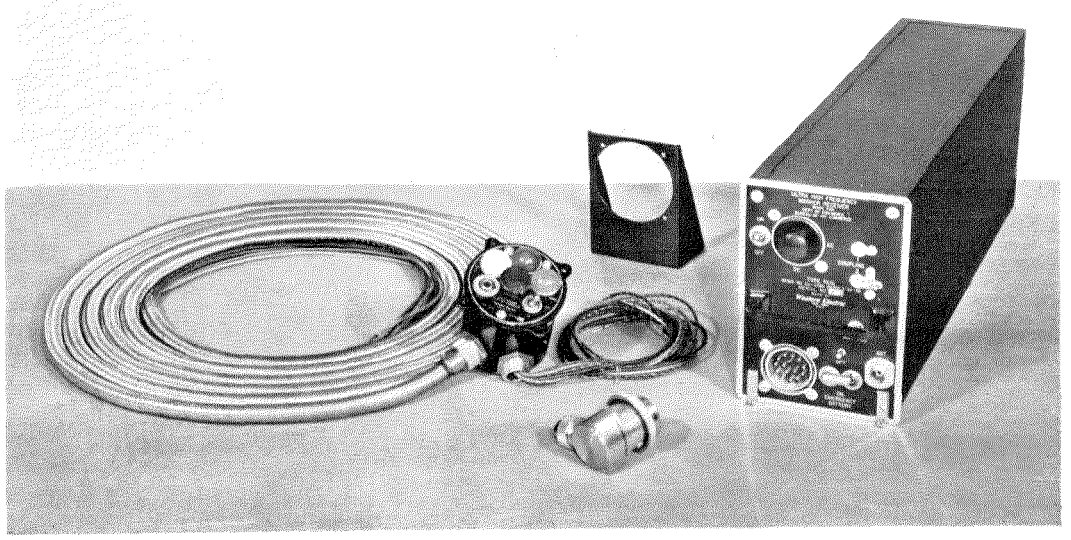


Fig. 18—Marker beacon receiver.

point of contact shall be not closer than 3 000 feet from the far end of the runway. At the point of contact the glide path shall have an angle with the runway not less than one degree and not more than two degrees. The glide path shall pass through the following points not less than 500 feet nor more than 700 feet altitude at a distance of three miles from far end of runway, and 1 500 feet altitude at six miles. The rate of descent shall not exceed 800 feet per minute."

Further work has been conducted by the International Telephone Development Company, which has succeeded in demonstrating a glide path meeting the pilots' recommendations.

Approximately fifty hooded landings have been made in various types of aircraft such as Douglas DC-3, Boeing 247 and Waco N. Numerous instrument approaches have been made under conditions of heavy rain, snow, low ceilings and poor visibility. A programme of flight training and testing is being organized in order that the airlines, army and navy may flight-test the equipment, either under the hood or under bad weather conditions.

The Civil Aeronautics Authority is planning to instal a number of similar systems throughout the United States in the near future for service tests and pilot training.

V.—CONCLUSION

(1) To effect precise landings with a minimum of effort on the pilot, experience has shown that four separate elements are advantageous, a localizer to furnish lateral guidance, a glide path to provide the path of descent, and two markers to indicate the progress along the course determined by the localizer and glide path.

(2) Ultra-short waves may be successfully employed to provide both horizontal and vertical guidance as well as marker indications.

(3) Localizer courses must be much sharper than heretofore employed, to avoid mutilation by reflections. The difference in amplitude of the patterns should be at least 2.2 db., 1.5° off course in the majority of installations. This is approximately four times sharper than stipulated in specifications of the Radio Technical Committee for Aeronautics. In some instances it is necessary to have a sharpness as great as 6 db. in order to eliminate bends in the course.

(4) Localizer polarization must be pure to secure independence of the aircraft antenna characteristics and direction of approach of the airplane to the localizer course.

(5) The localizer should be independent from the glide path in order to permit its installation well beyond the end of the runway, so that it is not a hazard to aircraft landing or taking off.

(6) Experience has indicated that the two-course localizer is inherently more simple to use than localizers with four or more courses.

(7) It is advantageous for the glide path to be of a controllable type, so that path shape can be adjusted during installation to accommodate the landing characteristics of aircraft.

REFERENCES

"Status of Instrument Landing Systems," by W. E. Jackson, *Proceedings of Institute of Radio Engineers*, June, 1938.

Report of Sixth Meeting of Radio Technical Committee for Aeronautics, Washington, D.C., December 17th, 1937.

"Ultra-High Frequency Loop Antennae," by A. Alford and A. G. Kandoian, presented during A.I.E.E. Winter Convention, January, 1940. Reprinted in this issue, p. 255.

The Polarization of Molecules

Application to the Study of Adsorption and Greasiness

By JEAN J. TRILLAT,

Professor of the Faculty of Sciences of Besançon, France

INTRODUCTION

THE structure of molecules plays a fundamental rôle in a great number of essential phenomena in Chemistry, Physics and Biology, such as adsorption, catalysis, osmosis and also certain associated phenomena such as skin effect, surface tension, adherence and molecular association.

Investigation of the important effects of molecular structure has become possible only recently. For the study of this problem, methods must be available for determining precisely the arrangement of the atoms in the molecule and the more or less symmetrical form of the latter.

Such methods, now well known, have been developed to a high degree of perfection. Among them may be cited: spectroscopy yielding indications on symmetry, molecular vibrations and moments of inertia, as well as the Raman effect whereby certain groups in the molecule may be characterized; X-ray spectroscopy disclosing the crystalline structure and permitting calculation of the length of molecules, their disposition in the crystal and frequently even the disposition of the atoms in the molecule, as well as the molecular arrangement within a liquid; electronic diffraction, making possible the study of very thin layers, the atomic arrangement and—in the case of a gas—the structure of isolated molecules; the variation of molecular heat as a function of temperature (disclosing information on nuclear deflections), symmetry and vibrations; and, finally, providing means for research on molecular polarity, the angle of valency, symmetry, mobility and deformation. It is the latter phenomenon with which the present article is concerned.

It is known that every molecule is an aggregation of atoms consisting of a heavy, positively charged nucleus and of an exterior framework of negatively charged electrons gravitating around this nucleus, the total of the positive and negative charges, respectively, being exactly equal for one atom or a neutral molecule. All the properties of atoms and molecules, both physical and chemical, originate from these elementary electrified particles which, as a first approximation, may be reduced to two: the *proton*, the unit of positive electricity, which is the nucleus of the hydrogen molecule; and the *electron*, unit of negative electricity.

The structure of the molecule will be known when the relative position of the different particles constituting the molecule are determined; that is, their distances, and the angles of valency uniting them. *Dielectric polarization* is precisely one of the most important methods whereby this result may be achieved.

What then is dielectric polarization? It is the behaviour of the molecule, formed of electrified particles, in an electric field—a behaviour which is only disclosed to us by a single experimental quantity, namely, the *specific inductive capacity* or *dielectric constant* which we shall now define.

It is known that it is necessary to introduce the dielectric constant ε of a medium in the macroscopic study of the field as the reduction factor in electrostatic actions, or as the coefficient of increase in capacities resulting from the interposition of an insulating medium between the charges or conductors which initially are assumed to be in a vacuum.

In other words, if B is the electric induction and H the inductive field, ε is defined as the ratio $\frac{B}{H}$. An attempt will be made to link up the dielectric constant ε with the molecular quantities.

The first attempt along these lines was made

by Clausius and Mosotti, who assumed that any molecule, from an electric viewpoint, could be regarded as a minute *perfectly conducting* sphere of radius a.

They further formulated the following hypothesis: If a sphere be traced around a given molecule as a centre, large with respect to the average distance of separation of the molecules but small compared to the apparatus containing the test substances, the resultant reaction of the centrally placed molecule on the other molecules within the sphere will be zero.

This hypothesis is probably only justified in the case of gases and very diluted solutions in certain solvents.

If this hypothesis and the spherical model molecule be accepted, it can be shown that the dielectric constant ε, the molecular mass M, the density ρ and the radius a of the molecule are related as follows:

$$\frac{\varepsilon-1}{\varepsilon+2}\cdot\frac{M}{\rho}=\frac{4}{3}\pi Na^3 \quad (N=6.06\times10^{23}).$$

The left hand member, all the quantities of which are susceptible of direct measurement, is defined as representing *polarization*. The radius of the molecule can thus be obtained.

Unfortunately, this equation contains erroneous assumptions for the majority of cases; for example, that the volume of a molecule being equal to the sum of the volume of the constituent atoms, the polarization of an aggregate must be equal to the sum of the polarizations of the components.

In reality, the comparison of a molecule to a conducting sphere is a rather simplified conception. Since a molecule consists of atomic nuclei around which the electrons circulate, an exterior electric field must cause distortion of the electronic orbits. The existence of a dielectric constant must be attributed to such distortion.

This distortion of the orbits, moreover, has been proved experimentally by the modification of the spectra in an electric field or the Stark effect. This phenomenon permits interpretation of the dielectric constant; and, utilizing Bohr's radius of the first orbit, the polarization P for hydrogen is given by:

$$P_0=\frac{9}{2}\times\frac{4}{3}\pi Na^3.$$

With the exception of the factor $\frac{9}{2}$ approximately, the equation is the same as given by the Clausius-Mosotti theory. While the conclusions to be drawn are valid, some reason for the disagreement obviously must exist. In the foregoing it is, in effect, assumed that polarization is due solely to the relative displacement of the constituent charges under the influence of the exterior electric field, and without regard to any pre-existing polarization. Debye has given the explanation: the molecule is not in effect necessarily symmetrical and, consequently, the centre of gravity of the positive and negative charges may not coincide.

In other words, from the viewpoint of the reactions, the positive and negative charges of the molecule may be replaced by two:

(1) A positive charge, equal to the sum of the positive charges of the molecule, and applicable to the centre of gravity of the positive charges.

(2) A negative charge, equal to the sum of the negative charges (and consequently of the positive charges if the molecule be neutral), and applicable to the centre of gravity of the negative charges.

In the absence of any exterior electric field, the molecule thus may be compared to a small permanent magnet with a positive pole and a negative pole separated by a certain distance d. This unit of two charges, $+q$ and $-q$, equal and of contrary sign, located at two points separated by the distance from d, constitutes an *electric doublet* or a *dipole*.

The electric moment μ of this dipole, by definition, is:

$$\mu=q\,.\,d\,.$$

It is the electric moment of the molecule, which is therefore called a *polar molecule*, possessing a *permanent moment*.

If the two charges $\pm q$ coincide, the electric moment of the molecule is obviously zero; i.e., molecules with completely symmetrical atomic distribution. Such molecules, accordingly, are *non-polar*.

A polar molecule, with respect to the electrical field, is comparable to a *magnetized needle* with respect to the magnetic field. The magnetic needle is oriented in the magnetic field (com-

pass); the dipole is similarly oriented in an ambient electric field of exterior origin.

This permanent moment is variable according to the chemical nature of the molecule; but, for a molecule of specific type, it is constant inasmuch as it does not change in form. This is why it is designated permanent moment.

We have seen that the electric field is capable of distorting the electronic orbits and, consequently, causing displacement of the molecular charges, all the positive charges being " drawn " to one side and the negative charges to the other, thus producing an electric moment which is *non-permanent* since it will disappear instantaneously with the field. This new moment, therefore, is an *induced moment* which results from the temporary distortion of the molecule by the field. If the field is not very strong, the moment induced is proportional to it; the coefficient of proportionality is a measure of the distortion of the molecule. Actually, the Clausius-Mosotti theory only takes account of the induced moment.

A molecule without a permanent moment—a non-polar molecule—is oriented in the electric field as a result of the induced moment.

A polar molecule—one with a permanent moment—moves both under the influence of its permanent moment and of the induced moment to which it is temporarily subjected. These two moments are geometrically superposed in accordance with vector laws.

But it is known that, in gases or liquids, molecules show an irregular movement which increases in velocity with rising temperature. The incessant impacts of the polar molecules or dipoles obviously are in opposition to their orientation in the electric field. As the temperature rises, the force and frequency of the impacts increase. Thus, the orientation due to the field on the dipoles is opposed to the de-orientation due to temperature. If, then, a system of molecules has a permanent moment (for example, a drop of a liquid) outside an electric field or a beam of light, the total moment will be zero because of the neutralizing effect of the combined orientations. In a very powerful field, all the molecules will be arranged in parallel, at least theoretically, and the resultant moment will reach a maximum which will diminish to zero as the temperature becomes elevated.

The theory of this phenomenon is quite similar to Langevin's theory of *paramagnetic bodies*. It was originated by Debye, who showed that the polarization P, as is clear from the preceding, is the sum of two terms: one, *independent of the temperature*, depends on the distortion of the electronic orbits or, if preferred, on the displacement of the charges by induction; the other, which decreases when the temperature increases, is due to the *partial orientation* of the molecules.

The Debye equation is as follows:

$$P = \frac{\varepsilon - 1}{\varepsilon + 2} \cdot \frac{M}{\rho} = \frac{4}{3}\pi N \left\{ \alpha + \frac{\mu^2}{3kT} \right\} = P_0 + P_e.$$

α is an order of magnitude characterizing the distortion; it is proportional to the induced moment found in the Clausius-Mosotti theory and is equal to $\frac{9}{2} a^3$ in the case of the hydrogen atom. μ is the permanent electric moment already defined: T the absolute temperature, k a constant, termed the Boltzmann constant, equal to $\frac{R}{N}$ (R = the constant of perfect gases; and N = Avogado's number = 6.06×10^{23}).

If $\mu = 0$ and the molecule is non-polar, only the term $\frac{4}{3}\pi N \alpha$ remains, thus reverting to the particular case to which the Clausius-Mosotti theory applies.

***Determination of* μ**

What is now of interest is the possibility of calculating this permanent electric moment μ. Several methods are available, consisting essentially in determining the dielectric constant and the distortion (or deformation) α.

(1) Recourse may be had to the action of the temperature on the polarization P. The variation of P with temperature being of the form $P = A + \frac{B}{T}$, it will be seen that the curve $T.P$ will be a straight line with an angular coefficient $A = \frac{4}{3}\pi N \alpha$ (whence the distortion α); from the ordinate at the origin, $B = \frac{4}{3}\pi N \frac{\mu^2}{3k}$, μ may be derived.

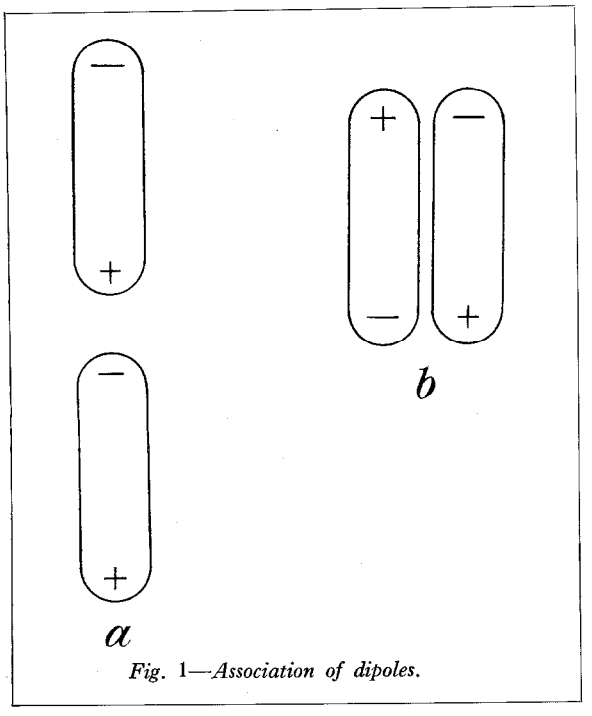


Fig. 1—Association of dipoles.

As to the measurement of P, which is also equal to $\frac{\varepsilon-1}{\varepsilon+2}\cdot\frac{M}{\rho}$, it consists in making measurements of specific inductive capacity at different temperatures and densities ρ. For the measurement of ε, measurements of capacity or of attraction or repulsion forces in the medium under consideration, are utilized. This method is particularly applicable to polarized gases or vapours where the molecules are not united.

(2) Other methods are applicable, involving the introduction of high frequency alternating fields.

Consider as an example the case of a *field of very high frequency*. Here the polarized molecules have no time to orient themselves between alternations, and orientation of the dipoles no longer plays any part in the dielectric phenomena. Then only the distortion effects of the electronic orbits occur, instantaneous and characterized by α.

Measurement of the dielectric constant ε under these conditions gives :

$$P_0 = \frac{\varepsilon-1}{\varepsilon+2}\,\frac{M}{\rho} = \frac{4}{3}\pi N\alpha\,;$$

hence, α may be derived.

Let us now take a similar measurement with a field of lower frequency. Re-arrangement of the orientations then takes place at a far higher speed than the field variations ; orientation at each instant is practically always equal to that in a constant field, and polarization is given by the complete equation. Use may accordingly be made of Hertzian waves of fairly long wavelength.

The two equations :

$$\begin{cases} P = \frac{4}{3}\pi N\left(\alpha + \frac{\mu^2}{3kT}\right) = P_0 + P_e \\ P_0 = \frac{4}{3}\pi N\alpha \end{cases}$$

give the value of μ through the elimination of α.

(3) Finally, the electric field of a light wave may be taken as a high frequency electric field. According to Maxwell's theory, the specific inductive capacity ε is related to the *index of refraction* n by the equation : $\varepsilon = n^2$.

The polarization P_0 then becomes :

$$P_0 = \frac{n^2-1}{n^2+2}\,\frac{M}{\rho} = \frac{4}{3}\pi N\alpha.$$

This quantity P_0 is well-known to chemists ; it is the *molecular refraction* which, according to Debye, is proportional to the distortion α and is independent of temperature.

Similarly, the molecular refraction of a compound should be equal to the sum of the atomic refractions of the components. The measurement of P_0 then reduces to a measurement of index and density.

The total polarization P can be determined either as above, by a dielectric constant measurement in a low frequency alternating current field or in a constant field, or by the measurement of indices in the infra-red spectrum. The final result is :

$$\mu = \sqrt{\frac{9kT}{4\pi N}\,(P-P_0)}\,.$$

(4) Still another method exists, attributable to Estermann. When a molecular jet is introduced into a non-homogeneous electric field, a distortion of the jet is produced ; hence, the electric moment may be deduced. Experience in the electric field has been analogous to that of Stern and Gerlach in the magnetic field

domain. It does not permit of highly precise determinations.

Polarization and Molecular Structure

What applications can be drawn from these theoretical conceptions, and what are the properties of polarized or non-polarized molecules ?

In the first instance, a polarized molecule produces round itself an electric field, analogous to a magnet and its magnetic field. If two polarized molecules approach each other they exercise a mutually orienting action and, according to their distance and relative positions, can either attract or repel each other. If they are free, attractions will result ; moreover, each of the two molecules distorts the other and an induced moment is superposed on their respective permanent moments. There will be *association*, such as occurs in the case of two permanent magnets (Fig. 1).

A non-polarized molecule also produces an electric field round itself by reason of the unequal distance at any point in space of the elementary positive and negative charges. This field, however, is very much weaker, particularly at some distance. The orienting action and mutual distortion are considerably less than in the case of polar molecules. These molecules, without a permanent moment, behave like pieces of soft iron with regard to one another.

Thus, given equivalent conditions, an assembly of polarized molecules will have *greater cohesion* than a non-polarized assembly.

The permanent electric moment also indicates the *structure of the molecules*, sometimes leading to quite unexpected results.

Consider the case of *water* as an example. It might be thought that the atoms of oxygen and hydrogen would take up their positions symmetrically on a straight line. In this case, the electric moment evidently would be zero. Experience, however, has shown that water in a state of vapour possesses a permanent moment equal to 1.80×10^{18}. The molecule of water is, then, strongly polarized and accordingly asymmetrical. Allowing for stability, it will be found that the only configuration possible is the triangular molecule (Fig. 2) with an angle of valence $\theta = 55°$, a distance $O - H = 1.02$ *A* and $H - H = 1.6\text{Å}$. This structure is confirmed by the infra-red absorption spectrum of water.

It is also found that ammonia has a blunt pyramidal molecule with $H - H = 1.73$ Å, $H - N = 1.04$ Å, pyramidal height 0.3 Å, and angle of height on the sides = 73°. On the other hand, the molecule of carbon dioxide gas CO_2 has an electric moment equal to zero and is, therefore, symmetrical.

Passing now to the case of organic molecules ; the following are some results which represent good examples :

Methane	CH_4	$\mu = 0$
Methyl Chloride	CH_3Cl	1.97×10^{-18}
Methylene Chloride	CH_2Cl_2	1.59×10^{-18}
Chloroform	$CHCl_3$	0.95×10^{-18}
Carbon Tetrachloride	CCl_4	0

The above shows that molecules of methane and carbon tetrachloride are perfectly symmetrical, which conforms to chemical theory, the four atoms *H* or *Cl* being disposed at the summit of a regular tetrahedon, the centre of which is occupied by the carbon atom (Fig. 3).

On the other hand, chlorides of methyl and methylene, also chloroform, are asymmetrical. Similarly, it will be found that ethane, ethylene and acetylene have an electric moment equal to zero, which conforms to derived formulae, whilst butylene $CH_2 = CH - CH_2 - CH_3$ is polarized with a moment $\mu = 0.4 \times 10^{-18}$. This

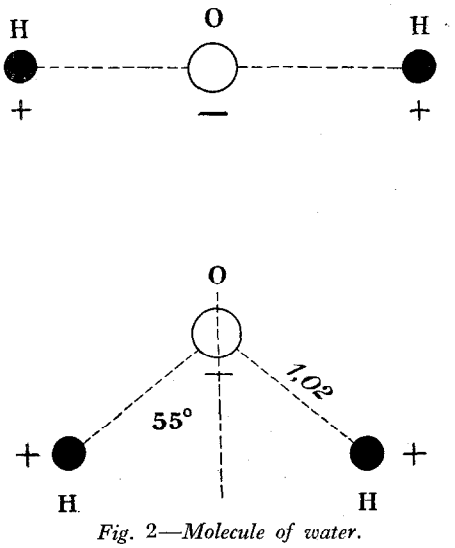


Fig. 2—Molecule of water.

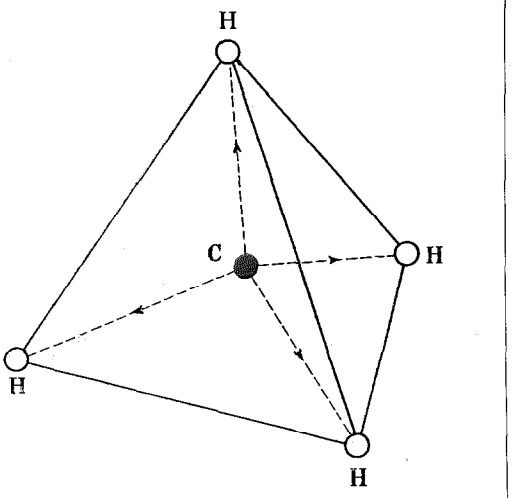


Fig. 3—Molecule of methane.

latter result shows that *a double link* in an asymmetrical position also produces electric asymmetry.

Lastly, certain isomers are immediately distinguishable by their moments. For example, dichlorethylene trans has a moment equal to zero and is, therefore, symmetrical, whilst the cis is polarized. Hence the possibility of distinguishing between these two bodies immediately (Fig. 4).

The permanent moment measured in polarized molecules is actually an average moment. It results from the *vectorial sum* of the partial moments corresponding to groups of atoms or to double or triple links.

The variable electro-affinity of the atoms comprised in the molecule sometimes makes it possible to localize, in *certain privileged regions*, the electronic asymmetry which contributes towards the permanent average moment and which, with the derived formulae, may be represented as lines of valence similar to dipoles initiating local ionization.

The vectorial composition of these local moments supplies the total moment μ of the molecule. Thus, benzene, according to Kékule's formula, has a zero total moment, the resultant of the partial moments being nil; but, if an NO_2 be substituted for an H, polarization will be produced, characterized by a moment $\mu = 3.6 \times 10^{-18}$.

In the dinitro derivations, para-dinitrobenzene has a zero moment in consequence of the symmetrical arrangement of the two NO_2 groups. Meta- and orthodinitro-benzene, on the other hand, have a permanent moment (Fig. 5).

All these facts, and others, point towards a generalization of the theory of the vectorial composition of partial moments, as suggested by Eucken. A specific electric moment is attributed, not to each substitute group in a polarized molecule, but rather to each *link* between two atoms, this partial moment being supposed to coincide with the direction of the link. It is assumed, accordingly, that the different valences of a given atom have specific directions forming a constant angle; for C and H, the angle of valences is assumed as 110°.

The success of this theory has been quite remarkable, in a number of cases agreement between the experimental and calculated values being surprisingly close. The moment μ of the polarized groups such as $H-C$, $C=O$, $C-Cl$, $C-C$, $H-O$, $C-Br$, $C-C$, etc., may thus be calculated.

Influence of the Physical State

We have seen that, under comparable conditions, an aggregate of dipolar molecules possesses greater cohesion than an aggregate of non-polarized molecules. This is due to the mutual attraction of dipolar molecules; this attraction naturally depends on the *physical state* of the substance, whether gaseous, liquid or solid.

In the gaseous state the molecules, separated at large average distances, have little or no tendency to become associated, and the orientation of the permanent moments in the electric field is not opposed by mutual interactions. Debye's

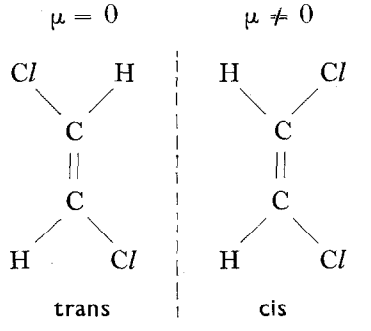


Fig. 4.—Dichlorethylene cis-trans.

theory is valid in this case only if the matter is in a diluted state, similar to the case of molecules diluted by a *non-polar* solvent with which they have no tendency to become associated.

Consequently, for the calculation of μ corresponding to the normal value of polarization, measurement is necessary, either in the gaseous state or after simple dilution of the polar body in a non-polar liquid, the value of the molecular polarization for an infinite dilution being extrapolated. For zero dilution, i.e., pure liquid of the polar body, the resulting value will be lower than the preceding one, due to the association phenomenon. The difference between these two defines a *co-efficient of association.* Thus, for water in the vaporized state, $\mu = 1.8\ D$; and, for water in the liquid state, $\mu = 1D$ (1 Debye $= 1D = 10^{-18}$). In general, therefore, it may be said that values of permanent moment and molecular association are definitely related.

APPLICATION TO ADSORPTION PHENOMENA

This conception of molecular polarization, just defined, plays a highly important rôle in connection with numerous phenomena in Physics, Chemistry and Biology. They include solubility, hydration of ions and molecules, swelling and colloidal stability as well as *adsorption.* The latter will now be considered.

What, essentially, is adsorption? Its etymology implies fixation of a body by the surface of another. More specifically, consider two different media, A and B, separated by an intervening surface; A and B may be solids, liquids or gases. The properties of the two media A and B are discontinuous at the intervening surface, termed the adsorption layer. To illustrate further:

(1) Consider a glass bulb evacuated at ordinary room temperature. If it be heated to approximately 400° C. additional gas may be extracted. Langmuir demonstrated that this gas contained air, hydrogen and, especially, water vapour, which remain fixed on the internal wall of the bulb with a minute thickness of molecular dimensions.

This fixation or adsorption of the water vapour by glass has been known for some time. It explains the surface electric conductivity of glass, preventing its use as an insulator in precise experimentation.

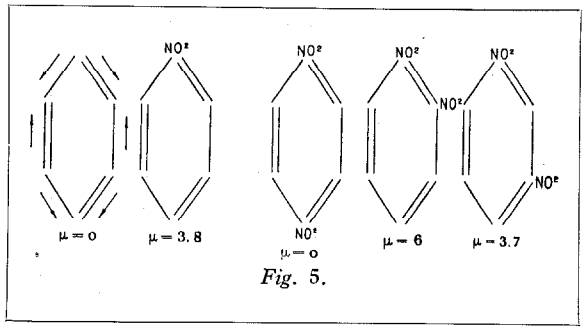


Fig. 5.

Obviously, the larger the glass surface the greater the quantity of gas adsorbed. It is well known that certain porous materials, such as charcoal, adsorb gases which liquefy easily.

(2) The surface of separation between air and a saline solution is constituted by pure water. The intervening layer composed of the solution and air, therefore, possesses properties different from the solution; its salt concentration is nil here. This phenomenon is termed negative adsorption.

Contrarily, if the first terms of the fatty acid series (e.g. acetic acid, butyric acid) be dissolved in water, it will be found that the inter-surface concentration is increased; it is an example of *positive adsorption.* These two cases are related as indicated in Gibb's thermodynamic formula:

$$u = \frac{-C}{RT} \times \frac{d\sigma}{dc},$$

in which C is the concentration of the material dissolved in the solution, u the excess concentration in gr/cm³ in the intervening layer, σ the superficial tension, T the absolute temperature and R the gas constant. It will be seen that u and $\frac{d\sigma}{dC}$ are of contrary signs. If, therefore, the superficial tension varies inversely as the concentration, there will be more foreign matter in the superficial layer than in the mass. Thus, the superficial tension in the intervening layer is dependent on positive or negative adsorption.

Rayleigh's, Langmuir's and Devaux's experiments have demonstrated the foregoing results by dissolving small quantities of fatty molecules endowed with permanent moment (acids, alcohols, etc.) in water; the molecules are

attracted to the surface of the water, there forming a kind of paving of fatty molecules. The same applies if a trace of fatty substance be placed on the surface of water, even though the former be barely soluble : the attraction between the dipoles of the water and the dipoles of the fatty acid, for example, causes the latter simultaneously to fix and orient themselves on the surface. This observation will be very useful hereinafter.

(3) A final example of adsorption is the *passivity* of metals in air or in oxidizing media. It is explained by the superficial fixation of an invisible metallic oxide film, more precious than the metal. Although the metal appears unaltered to the eye, its surface is covered by a kind of protecting coating ; thus chromium, nickel, zinc and aluminium are passive or inactive when exposed to the atmosphere.

The extreme importance of adsorption phenomena is evident inasmuch as substances cannot enter into contact with or re-act on others except through their surfaces ; and, therefore, the condition or composition of their surfaces or coating, i.e., the presence or absence of adsorption molecules, definitely controls a great number of their characteristics, both physical and chemical.

Molecular polarity theories summarized at the beginning of this article show that adsorption must be considered largely as a *manifestation of this polarity*, i.e., whether permanent electric moments or dipoles are present or not. In effect, the presence of an electric doublet in the molecule means that the latter is endowed with peculiar properties ; in particular, mutual attraction of oppositely charged particles.

The most highly polarized molecules are those having the greatest tendency to join on to other molecules, like or unlike. Adsorption as a superficial fixation of adsorbed molecules is thus easily explained.

To consider the phenomenon somewhat more closely : it has just been seen that association in liquids depends on the permanent electric moment, substances with large moments, such as water, alcohol and acids, being strongly associated. Similarly, solubility phenomena depend largely on molecular polarity. In general, hydrophilic groups (*HO—COOH – COOH – CONH*$_2$, etc.) the presence of which endows organic matter with the property of being soluble in water, are also polar groups, and liquids containing them are associated. They are, moreover, mutually soluble. It may be said, furthermore, that non-polar liquids and molecules are not generally soluble except in non-polar substances (liquids). Thus, sulphur is soluble in carbon disulphide but not in water or alcohol. Paraffin hydrocarbons are miscible with benzene, and are mutually miscible provided their molecular weight is not too great ; they are not soluble in water or in the lower aliphatic alcohols.

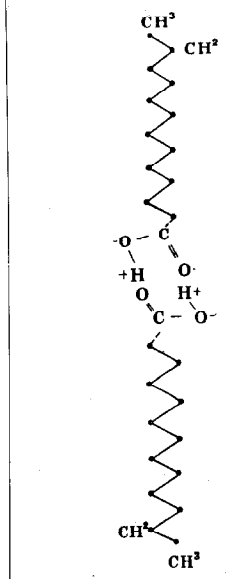


Fig. 6—Diagram of fatty acid molecule.

This must not, nevertheless, be taken as absolute. An example will be given hereinafter involving a series of fatty aliphatic acids in connection with further consideration of adsorption.

The series of saturated fatty acids begins with formic and acetic acid. Passing from one to another by the addition of a CH_2 group, acids of high molecular weight are reached, such as palmitic (C_{16}) and stearic acid (C_{18}).

Acids in the first portion of the series are soluble in water ; the contrary is true of acids in the latter part of the series. This may be explained by the polar character of the molecule as well as by its length.

Actually, all these molecules have a substantially constant permanent electric moment due to the *COOH* group. A fatty acid molecule, accordingly, may be represented as indicated in Fig. 6 ; it is at the extremity *COOH* that the "activity" of the molecule is localized.

The structure (Fig. 6) shows on the one hand, that the molecules tend (which will be reverted

to subsequently) to group themselves in twos with their extremities opposed. On the other hand, the *COOH* group, being polar, is attracted by water, which is equally polar, the attraction being greater the lighter the molecule. This pertains to the first portion of the series, thus explaining the solubility of the groups. With greater elongation of the carbon chain *R* and masking of the electric moment in the neutral character of the large molecule, the end molecules *COOH* will tend to dissolve due to the attraction of the water; the residue of the molecule, however, will be resistant to solution: solubility in water will decrease whilst solubility in a non-polar liquid, such as saturated hydrocarbon, will increase. Simultaneously, lateral associations will appear, due to the forces radiated laterally by the carbon atoms of the chain (Van der Waals forces).

Another very important result will be the following: if such dipolar molecules come in contact with a surface in the neighbourhood of which an electric field exists due, for example, to atoms, they will be attracted by this surface and will become rooted by their active group (*COOH*).

This is what occurs if a very small quantity of fatty acid be placed on the *surface of water*. Langmuir (1917) demonstrated that the molecules, thus held and adsorbed at the surface, are oriented with their electrified portions towards the water, i.e., the portions containing the dipole—the *COOH* acid group—or, applying his terminology for the active group, "liophile." This adsorption will be accompanied by the side-to-side orientation of all the molecules which, in some kind, take up the position of the floats on fishing rods. In this way the terminal surface will consist entirely of radicals CH_3 (Fig. 7).

Accordingly, orientation in the form of a thin film on the surface of the water is confined to the polar molecules; the others, which do not attract water, contribute to the formation of globules and represent the case of saturated hydrocarbons. It is a well-known fact that saturated mineral oils do not spread themselves over water, whereas vegetable oils, such as olive oil, which consists of highly polarized molecules, form a mono-molecular film.

Similarly, if fatty acid molecules contact *a metallic surface*, they are adsorbed and orient themselves in consequence of the electric field radiated by the superficial atoms of the metal. It is again the polar group *COOH* of the fatty molecule which fixes itself but, inasmuch as the atoms of the metal are immobile, the fatty film does not have a tendency to spread. It adheres strongly to the metal and this is definitely the origin of *greasiness*, which is consequently electric in origin.

Furthermore, this orientation is not necessarily confined to a single mono-molecular formation. Thus, a large number of fatty molecules are able to assume a chain formation, resulting in *stratifications* with an individual thickness equal to that of a molecule (Fig. 8).

While the preceding is based on hypotheses, the author has succeeded in verifying the majority by X-ray or electronic diffraction methods. By way of illustration, it will next be shown how they can be applied to a particular case, i.e., lubrication and greasiness.

APPLICATION TO LUBRICATION PHENOMENA

When a film of lubricating oil is interposed between two metallic surfaces, two cases must be considered:

(1) A relatively thick liquid film. It represents a case of fluid friction due to hydrodynamic reactions with which the present article is not concerned.

(2) Lubrication, or oily greasing, by means of a film not greater than a few molecules in thickness.

This method of lubrication is most important, since it is applied almost invariably in the

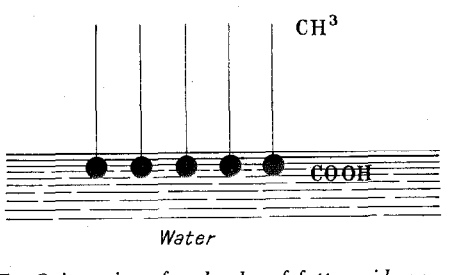


Fig. 7—Orientation of molecules of fatty acids on water.

starting of machines, and also in the operation of heavily loaded machinery, gearing, etc.

Lubrication limits, coupled with associated problems of adherence, structure and molecular orientation, are thus most likely to throw light on the nature of lubrication phenomena. An attempt will, therefore, be made to link up the greasiness of oils with the above-mentioned polarity and adsorption phenomena.

It is necessary first to stress the fact that greasiness is not an intrinsic property of oil itself. It is rather a property due to the intervening surface formed by the oil with the metallic or any other solid bodies. The nature of the two metallic surfaces and the molecular character of the oil film separating them all influence the resistance opposed to the stress, i.e., the *coefficient of static friction.*

The most important factor in the study of greasiness is probably the fact that the oil molecules take up a well-defined direction when they are very close to or in contact with a metallic surface. In other words they orient themselves. It is now known that this orientation is due to the presence at one of the extremities of the molecule of a *permanent electric moment.* It is here that the nature of the molecule wields its influence.

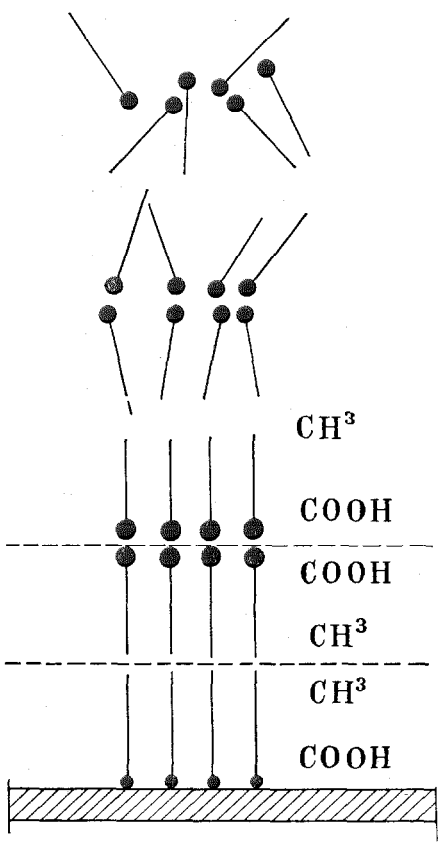


Fig. 8.

Assume that the film spread over the metallic surface is formed entirely of molecules of a specific oil, either vegetable or animal, or of a fatty acid such as stearic or oleic acid. The molecule of one or the other assumes the form of an elongated rod, terminated at one end by the group CH_3 and at the other end by the active group *COOH* comprising the electric doublet.

In the case of a triglyceride, the active end will be the "glyceryl radical."

Upon contact with the metal, these polar molecules are adsorbed, as already indicated, and form a mono-molecularly oriented base. Additional molecules are drawn across this base and form a second oriented layer, etc. The electric field, however, falls very rapidly—orientation becomes less and less definite ; and, after a distance corresponding to a few molecular thicknesses, the molecules are disposed in an entirely irregular manner. The limiting film is then reached (Fig. 8).

The total thickness varies according to the activity of the liquid molecules and that of the superficial field of the metal. Thus the very different degrees of greasiness observed with different combinations of metal and oil are largely explained.

It is, therefore, logical to anticipate that, with increasing size, molecules will grip more firmly laterally by Van der Waals forces, and orientation will be more pronounced. Greasiness is, therefore, also a function of the average dimensions of the molecules of the lubricant—a fact which has been established by Woog's ingenious experiments.

It follows that polar molecules, endowed with permanent electric moment, yield this very desirable feature of greasiness. Does it, however, imply that paraffin hydrocarbons, such as are contained in paraffin oil, are incapable of orientation or adsorption on contact with a metal ?

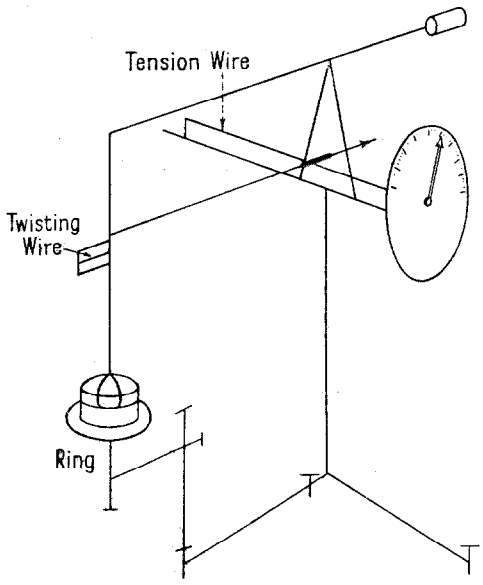


Fig. 9—Diagram of Lecomte du Noüy's interfacial balance.

Molecules of the saturated hydrocarbon type behave differently from those of the corresponding fatty acids inasmuch as their electric charges are distributed in a perfectly symmetrical manner and their permanent electric moment is zero. It has, however, been seen that, even if a molecule does not possess a permanent electric moment, its electric symmetry may be unbalanced by *influence* in an electric field. An electric moment appears temporarily, due to distortion of the molecule.

Under these conditions, when a paraffinic hydrocarbon, such as hexane or tetradecane, comes into contact with a metallic surface the following phenomenon occurs: the molecules move quite haphazardly and without any orientation at a small distance from the surface; the molecules in direct contact with the metallic surface are adsorbed by induction and move perpendicularly to the surface. This adsorption will, however, be *very much less energetic* than that of the polar molecules and fatty acids or of triglycerides, and this adsorption layer of non-polar hydrocarbons may be easily destroyed by pressure so that the metallic surface can come into direct contact and grip. In this case, therefore, there will be very little or no greasiness. If greasiness is desired such an oil will be unsuitable due to its tendency to be driven off by pressure. On the other hand, with its mobility, and its wetting power, it will be very suitable for hydrodynamic greasing in the form of a thick film whose principal function is to separate the surfaces and dissipate heat. This will, for example, be the case of very rapidly moving mechanisms.

From the foregoing it seems that it should be possible to produce lubricating oils combining the two properties; if a trace of polar molecules such as fatty acids, glycerides, or saturated alcohols is added to an oil containing only paraffinic hydrocarbons, these molecules, having a strong permanent moment, will be *selectively* attracted by the metal, where they will form strongly retained oriented strata. The saturated hydrocarbon molecules will roll on these strata.

The practicability of this theory was demonstrated by Wells and Southcombe in 1920, and later by Woog; it will be seen, therefore, that the greasiness does not depend on the mass of the oil but on the thin layer adsorbed on the frictional surface.

In studying these phenomena I found that the application of diffraction of X-rays and the diffraction of electrons disclosed a like monomolecular layer of oriented grease on the surface of a metal. This latter method enabled me to calculate directly the distance between two adjacent atoms of carbon, as being equal to 1.25 Angstroms. It is possible to draw practical conclusions on the value of an oil, as has been shown recently in England in regard to the lubrication of airplane motors. Similar investigations have been made in the laboratories of the French Air Ministry.

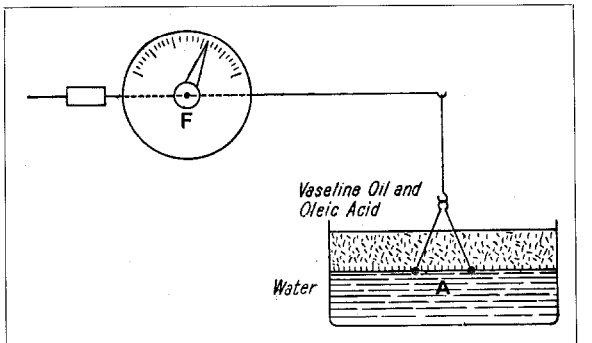


Fig. 10—Selective adsorption of oleic acid molecules at the oil/water separation surface.

In continuing these investigations, however, I adopted an entirely different method, and sought answers to the questions : is it possible definitely to demonstrate the *selective adsorption of polar molecules* in a mixture as complex as oil, and, if so, is it possible to measure the thickness of the adsorbed layer and to draw practical conclusions therefrom ?

For this purpose I turned to the measurement of the interfacial tension between an oil and pure water. This form of measurement consists in the pulling of a platinum ring arranged at the separating interface between the oil and the water, the pull being indicated by the twisting of a steel wire.

The most suitable instrument is Lecomte du Noüy's interfacial balance, which directly measures in dynes/cm the force which is required to tear the film of water brought up by the ring into the middle of the oil (Fig. 9).

Accordingly :

$$2\ l.\ \sigma = f = 4\ \pi\ r.\ \sigma$$

l = periphery of the ring, r = radius, while σ = interfacial tension.

The interfacial tension, like the surface tension, is very sensitive to the slightest modifications in the composition of the two contacted liquids. This observation was, in particular, used by Dubrisay, who deduced from it a very sensitive method of analysis (capillary analysis).

Supposing, therefore, that we have a layer of pure water and a layer of oil assumed in the first instance to consist only of saturated hydrocarbons, i.e., non-polar molecules. As a result

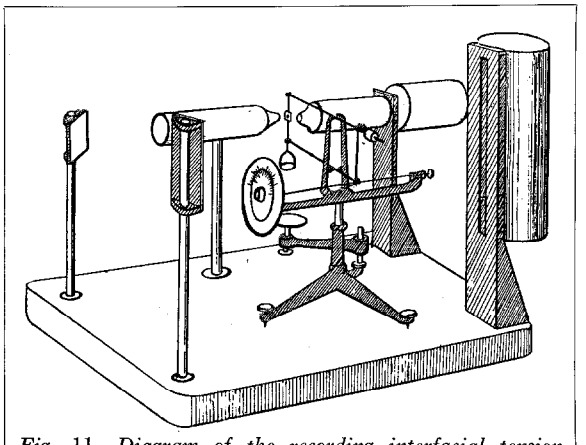


Fig. 11—Diagram of the recording interfacial tension balance (J. J. Trillat).

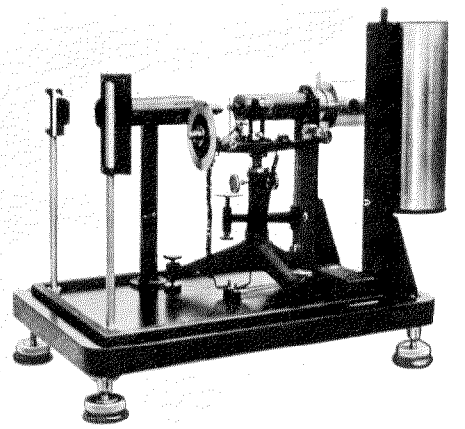


Fig. 12—Recording interfacial balance (J. J. Trillat).

of thermal agitation the hydrocarbon molecules will pass in the neighbourhood of the highly polarized molecules of the water, but they will hardly be retained, if at all ; the interface layer will, therefore, have a composition similar to that of oil, and this irrespective of the time of contact (Fig. 10).

Supposing, on the other hand, that inert oil has had added to it a small quantity of *polar molecules* like fatty acids, for example, in the proportion of one polar molecule to 10 000 non-polar.

Molecular and thermal impacts cause these molecules to pass in close proximity to the polar molecules of the water ; they are at once attached to the latter by electrostatic attraction, thus being adsorbed on the separating surface.

Progressively, as a function of the time, the number of adsorbed polar molecules will increase until the whole of the interface is saturated. There will, therefore, be an *increase in the concentration* of the polar molecules in the transition layer which, in practice, consists solely of these latter. This increase in concentration is regulated by Gibbs's formula.

This adsorption is, thus, *selective* since it results in the sorting out of the polar and non-polar molecules ; as this adsorption is accompanied by a *decrease in the interfacial tension* as a function of time, we can follow the phenomenon by repeated measurements of interfacial tension, made by means of the interfacial balance.

On making these measurements I found the variation which I expected. The nature of the

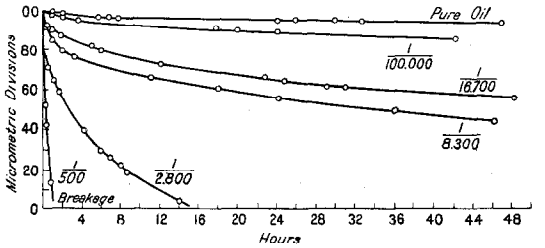


Fig. 13—Paraffin oil + oleic acid on distilled water.

added molecules and the value of their electric moment naturally play an essential part; a very clear proof is obtained by dissolving either cyclohexanol (possessing a permanent electric moment) or cyclohexane (electric moment nil) in paraffin oil. In the first case, there will be a considerable drop in the inter-surface tension, whilst nothing happens in the second case.

Finally, it should be noted that the method of measuring the inter-facial tension is of extraordinary sensitivity, revealing as little as 1/100 000 of active impurity admixed with the paraffin oil.

As these absolute measurements are delicate and occupy considerable time, I considered replacing them by relative measurements. Instead of drawing out the ring, only a known and constant quantity is taken, in such a way as to draw a film of oil on the water. For this purpose, as soon as the oil has been poured on the water, the platinum ring is placed exactly on the oil/water separation surface and a vertical force applied thereto which is determined by a known torsion of the wire corresponding, for example, to 20 dynes/cm (Fig. 10).

Since the interfacial tension diminishes as a function of the time as a result of the increase in the number of active adsorbed molecules, the force remaining constant, the ring will rise progressively. This displacement may be followed by means of a micrometer; recently, I succeeded in *photographically* recording this displacement as a function of time, and in constructing an apparatus which is absolutely automatic and does not require an operator (Figs. 11 and 12).

By means of this apparatus I was able to make a number of determinations, both theoretical and practical.

In the first instance, let us consider a pure paraffinic oil in contact with the water; the curve registered shows no modification as a function of time, which confirms that there is no adsorption of these saturated non-polar hydrocarbon molecules (Fig. 13).

Let us add variable and known quantities of oleic acid which is strongly polar, so as to obtain a scale of concentrations comprised between 1/500 and $1/10^6$. Each of these oils is characterized by a particular curve tending to flatten towards its end, and which to some extent measures the speed of adsorption (Fig. 13).

Each curve thus characterizes a concentration, and it is, therefore, easy to determine an unknown concentration by extrapolation, for example, or by means of certain calculations. In this way, it is possible to establish an actual table of grades of the activity of the oil expressed in oleic acid.

It will be found, for example, that an active oil kept in a glass bottle for some months gives a curve different from that of new oil; this results from the adsorption of a certain quantity of oleic acid molecules by the inner sides of the glass, and the number may be measured by a graded table.

It is, however, possible to go very much further. Consider, for example, a paraffin oil

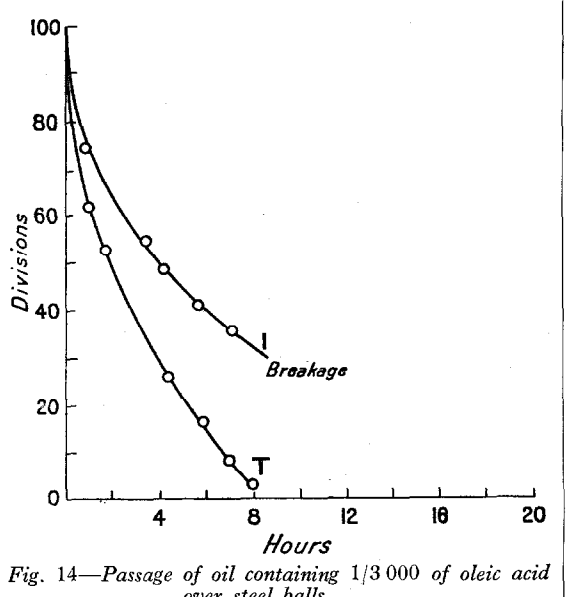


Fig. 14—Passage of oil containing 1/3 000 of oleic acid over steel balls.
T, paraffin oil + 1/3 000 oleic acid (proof);
I, same oil after passage over steel balls.

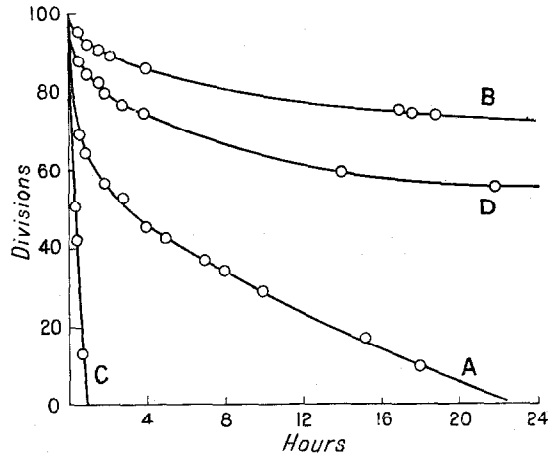


Fig. 15—Passage over filter paper.
A, oil containing 1/2 800 oleic acid; B, oil (A) after passing over filter paper; C, oil containing 1/500 oleic acid; D, oil (C) after contact 1 hour with 2 filter papers and filtration through a third.

to which 1/3 000 of oleic acid has been added. Let us put it into contact with any metallic surface, say, calibrated balls of steel. Then let us take this oil and again submit it to the tension measurement of oil/pure water. If previous postulations are correct, the oil, on contact with the steel balls, will abandon some of its molecules of oleic acid, which will selectively become fixed on the metal to form the limiting film. Consequently, concentration of oleic acid will have diminished in the oil, and we should, on contact with the water, obtain a curve which is different from the first (Fig. 14).

This new concentration can be determined by reference to our grading curves. Thence it is easy to calculate the number of oleic acid molecules adsorbed, and also the thickness of the limiting films or the section of the molecules of oleic acid.

I have thus found that this layer of adsorption has a thickness of about 28 μμ, and comprises *a dozen* molecules of oleic acid placed end to end. The origin of greasiness lies in this adsorbed film, to which we have for the first time had direct access.

The measurement of interfacial water/oil tension, therefore, only serves as an *intermediary* of a very definite proportioning method, as it is useless to say that this result could not be obtained by any chemical method.

Similar experience has been gained in filtering in a single operation a paraffin oil which is activated by oleic acid through *filter paper or cotton.* The adsorption is then considerable, the oil, after filtration, has lost a large proportion of its oleic molecules and, accordingly, its greasiness. Curve Fig. 15 is an example of this, and shows the importance of the method with regard to the influence of filtration on oil qualities.

The same experiments can, of course, be made with technical commercial oils, and work in this field is proceeding in various technical laboratories.

The variability of an oil can thus be studied by its operation in machinery or under heat, its modifications as a result of filtration, its degree of refinement, etc. The two following curves in Fig. 15 have been obtained with a well-known automobile oil.

With the recording apparatus these tests are very simple, and the measurements are very accurate and true. It should be understood that the field of application for this method is not limited merely to lubrication; many of the adsorption phenomena in liquid phase can be dealt with in this way, such as, for example, the fixation of dyes.

The object of this article has been principally to show how, proceeding from entirely theoretical conceptions of molecular structure and permanent or induced electric moments, a satisfactory explanation may be given for a considerable number of little known facts, and how these investigations could be extended into the purely practical field.

Comparative Telephone Development of the World by Principal Countries

JANUARY 1st, 1928 TO JANUARY 1st, 1938

Compiled by the Comptroller's Department, International Telephone and Telegraph Corporation

FROM the viewpoint of telephone development, the principal countries of the world may be classified under :—

(1) Absolute number of stations.
(2) Number of stations per 100 of population.
(3) Average compound rate of growth.

In the tabulation on this page, twenty countries are listed sequentially according to each of these classifications. As might be expected, the highest rate of growth occurred in certain of the least intensely developed countries.

Amongst the industrialized European countries, in terms of development density (telephones per 100 of population), Sweden, Denmark and Switzerland, respectively, rank highest, followed by Norway and Great Britain. Outside Europe and the U.S.A., the development densities of New Zealand, Canada and Australia are impressive.

The table on page 320 contains statistics for the 10-year period. It will be noted that in general the countries in which the International Telephone and Telegraph group operates compare very favourably in rate of growth with most of the other countries.

COMPARATIVE TELEPHONE DEVELOPMENT

THE 20 COUNTRIES LEADING IN RESPECT OF : (1) GREATEST NUMBER OF TELEPHONES ON JANUARY 1, 1938 ; (2) MOST TELEPHONES PER 100 POPULATION ON JANUARY 1, 1938 ; (3) FASTEST RATE OF GROWTH IN THE TEN-YEAR PERIOD JANUARY 1, 1928 TO DECEMBER 31, 1937

	(1) Greatest Number of Telephones		(2) Most Telephones per 100 Population		(3) Fastest Rate of Growth in Ten Years	
Rank	Country	Telephones	Country	Telephones per 100 Population	Country	Average Compound Rate of Gain
1	U.S.A.	19 453 401	U.S.A.	15.1	Russia	13.3%
2	British Commonwealth	5 590 000	New Zealand	11.9	Brazil	8.4
	Great Britain and North Ireland	3 029 456	—	—	—	—
	Canada	1 322 794	—	—	—	—
	Australia	594 855	—	—	—	—
	New Zealand	192 020	—	—	—	—
	Union of South Africa	189 601	—	—	—	—
	Asiatic Possessions	147 400	—	—	—	—
	Other Possessions	104 850	—	—	—	—
3	Germany	3 623 697	Canada	11.9	Spain	7.8
4	France	1 552 618	Sweden	11.8	Mexico	7.8
5	Japan	1 304 693	Denmark	11.3	Italy	7.4
6	Russia	950 000	Switzerland	10.3	Belgium	7.2
7	Sweden	738 698	Australia	8.7	Switzerland	6.8
8	Italy	600 501	Norway	7.6	British Africa	6.7
9	Switzerland	430 877	Great Britain	6.4	Peru	6.6
10	Denmark	425 951	Germany	5.3	Great Britain	6.4
11	Netherlands	401 484	Belgium	4.7	British Asiatic Possessions	6.3
12	Belgium	393 528	Netherlands	4.7	Chile	6.3
13	Argentina	377 473	Finland	4.5	Colombia	6.3
14	Spain	300 000	Austria	4.1	France	5.8
15	Austria	281 790	France	3.7	Japan	5.7
16	Poland	272 300	Argentina	3.0	Poland	5.6
17	Brazil	241 561	South Africa	1.9	Netherlands	5.3
18	Norway	222 010	Japan	1.8	Uruguay	5.1
19	Czechoslovakia	220 510	Hungary	1.7	Argentina	5.0
20	Finland	171 741	Chile	1.5	Venezuela	4.9

Compiled by I.T. & T. Comptroller's Department, from "Telephone and Telegraph Statistics of the World" issued by the American Telephone and Telegraph Company.

COMPARATIVE TELEPHONE DEVELOPMENT OF COUNTRIES, 1928–1938

	Number of Telephones, January 1			Growth 1928–38			Per Cent. of World			Telephones Per 100 Population		
					Per Cent.							
	1928	1937	1938	Number	10 Years	Yearly	1928	1937	1938	1928	1937	1938
U.S.A.	18 522 767	18 433 400	19 453 401	930 634	5.02	0.49	59.77	49.69	49.57	15.8	14.37	15.09
Canada	1 259 987	1 266 228	1 322 794	62 807	4.98	0.49	4.07	3.41	3.37	13.2	11.48	11.90
Australia	442 362	562 868	594 855	152 493	34.47	3.0	1.43	1.52	1.51	7.2	8.31	8.71
New Zealand	144 552	178 599	192 020	47 468	32.84	2.9	0.47	0.48	0.49	10.0	11.25	11.97
South Africa	97 155	169 419	189 601	92 446	95.15	6.9	0.31	0.46	0.48	1.3	1.75	1.90
Total Dominions	1 944 056	2 177 114	2 299 270	355 214	18.27	1.7	6.28	5.87	5.85	7.5	7.5	7.8
Austria	165 231	279 595	281 790	116 559	70.54	5.5	0.53	0.76	0.72	2.4	4.10	4.12
Belgium	196 691	361 685	393 528	196 837	100.07	7.2	0.63	0.98	1.00	2.5	4.34	4.70
Czechoslovakia	134 123	207 287	220 510	86 387	64.41	5.1	0.43	0.56	0.56	0.9	1.36	1.43
Denmark	324 232	408 875	425 951	101 719	31.37	2.8	1.05	1.10	1.09	9.3	10.89	11.25
Finland	108 973	160 469	171 741	62 768	57.60	4.6	0.35	0.43	0.44	3.0	4.20	4.48
France	883 406	1 481 788	1 552 618	669 212	75.75	5.8	2.85	4.00	3.96	2.2	3.51	3.70
Germany	2 814 996	3 431 074	3 623 697	808 701	28.73	2.55	9.08	9.25	9.23	4.4	5.08	5.31
Great Britain and Northern Ireland	1 633 802	2 791 597	3 029 456	1 395 654	85.42	6.4	5.27	7.53	7.72	3.6	5.93	6.41
Hungary	120 000	137 651	149 339	29 339	24.45	2.2	0.39	0.37	0.38	1.4	1.53	1.65
Italy	292 867	560 660	600 501	307 634	105.04	7.4	0.95	1.51	1.53	0.7	1.31	1.38
Netherlands	238 602	382 173	401 484	162 882	68.27	5.3	0.77	1.03	1.02	3.1	4.47	4.65
Norway	179 484	210 608	222 010	42 526	23.69	2.1	0.58	0.57	0.57	6.4	7.26	7.61
Poland	157 425	244 924	272 300	114 875	72.97	5.6	0.51	0.66	0.69	0.5	0.71	0.79
Rumania	56 024	70 678	81 205	25 181	44.95	3.8	0.18	0.19	0.21	0.3	0.36	0.41
Russia	260 000	950 000	950 000	690 000	265 38	13.3	0.84	2.56	2.42	0.2	0.55	0.53
Spain	141 214	341 390	300 000	158 786	112.44	7.8	0.46	0.92	0.76	0.6	1.38	1.20
Sweden	466 787	687 566	738 698	271 911	58.25	4.7	1.51	1.85	1.88	7.7	10.97	11.75
Switzerland	223 597	412 324	430 877	207 280	92.70	6.8	0.72	1.12	1.10	5.6	9.86	10.26
Other Europe	225 953	392 808	423 525	197 572	87.44	6.5						
Total Europe	8 623 407	13 513 152	14 269 230	5 645 823	65.47	5.2	27.83	36.43	36.36	1.6	2.35	2.47
Japan	750 561	1 197 129	1 304 693	554 132	73.83	5.7	2.42	3.23	3.33	1.2	1.70	1.82
Other Asia	291 838	493 849	516 498	224 660	76.98	5.9	0.94	1.33	1.31	0.04	0.05	0.05
Total Asia	1 042 399	1 690 978	1 821 191	778 792	74.71	5.75	3.36	4.56	4.64	0.1	0.16	0.17
Africa (1)	106 202	165 797	183 654	77 452	72.93	5.6	0.34	0.44	0.47	0.08	0.12	0.13
Oceania (1)	90 813	99 413	109 090	18 277	20.13	1.85	0.29	0.27	0.28	4.8	4.8	5.2
Central America	22 511	27 036	28 553	6 042	26.84	2.4	0.07	0.07	0.07	0.3	0.38	0.40
Cuba	68 828	44 582	51 163	(17 665)	(25.67)	(3.2)	0.22	0.12	0.13	1.9	1.02	1.15
Puerto Rico	13 222	14 803	15 171	1 949	14.74	1.4	0.04	0.04	0.04	0.9	0.84	0.85
Mexico	64 916	124 779	138 081	73 165	112.71	7.8	0.21	0.34	0.35	0.4	0.66	0.72
Other North America and W. I.	27 373	41 595	40 858	13 485	49.26	4.1	0.09	0.11	0.11			
Total North America (2)	196 850	252 795	273 826	76 976	39.10	3.3	0.63	0.68	0.70	0.55	0.63	0.68
Argentina	232 041	348 184	377 473	145 432	62.68	5.0	0.75	0.94	0.96	2.2	2.77	2.96
Bolivia	2 612	2 450	2 555	(57)	(2.18)	(0.2)	0.01	0.007	0.007	0.1	0.08	0.08
Brazil	108 189	221 982	241 561	133 372	123.28	8.4	0.35	0.60	0.62	0.3	0.45	0.47
Chile	38 573	63 055	70 867	32 294	83.72	6.3	0.13	0.17	0.18	0.7	1.38	1.54
Colombia	21 110	41 400	38 759	17 649	83.60	6.3	0.07	0.11	0.10	0.3	0.42	0.39
Ecuador	4 407	6 750	7 083	2 676	60.72	4.9	0.01	0.02	0.02	0.2	0.27	0.28
Peru	13 695	23 883	25 981	12 286	89.71	6.6	0.04	0.06	0.07	0.2	0.36	0.39
Uruguay	26 934	34 500	44 068	17 134	63.61	5.1	0.09	0.09	0.11	1.5	1.68	2.11
Venezuela	13 098	17 388	21 058	7 960	60.77	4.9	0.04	0.05	0.05	0.4	0.51	0.61
Other South America	3 151	5 843	6 002	2 851	90.48	6.7	0.01	0.016	0.017			
Total South America	463 810	765 435	835 407	371 597	80.12	6.1	1.50	2.06	2.13	0.6	0.80	0.86
World excluding U.S.A.	12 467 537	18 664 684	19 791 668	7 324 131	58.75	4.7	40.23	50.31	50.43	0.69	0.94	0.99
World including U.S.A.	30 990 304	37 098 084	39 245 069	8 254 765	26.64	2.4	100.00	100.00	100.00	1.6	1.71	1.79
I.T. & T. System	284 722	992 461	1 042 489	757 767	266.14	13.4	0.92	2.68	2.65			

(1) Excluding British Dominions. (2) Excluding U.S.A. and Canada.