Theory of Frequency Counting and Its Application to the Detection of Frequency-Modulated Waves*

EDOUARD LABIN†, SENIOR MEMBER, I.R.E.

Summary—Electronic circuits of the "frequency-counting" type furnish, in response to a sinusoidal signal of frequency $\omega/2\pi$, a continuous signal proportional to ω .

It may then be expected that, within certain limits, if the frequency ω is modulated, this "continuous" signal output will reproduce the modulation.

In this paper are studied, first, the validity of this principle of detection of frequency-modulated waves, with observations on the subject of detection in general, and second, the methods employed for carrying into effect the electronic counting.

I. GENERAL PRINCIPLES OF THE DETECTION OF F.M. WAVES BY COUNTING

ET US CONSIDER a nonsinusoidal, but periodic, current or voltage of period

$$T = \frac{1}{F} = \frac{2\pi}{\Omega} \,. \tag{1}$$

Its expression in Fourier series would be:

$$u = U_0 + \sum_{n} U_n \cos(n\Omega t + \phi_n)$$
 (2)

with the following value for the mean term:

$$U_0 = \frac{1}{T} \int_0^T u(t)dt = Fb = \frac{\Omega b}{2\pi}$$
 (3)

where b is the area covered by one period of the curve u(t). See Fig. 1.

We may consider that the coefficients U_0 , U_n , ϕ_n of the Fourier series are functions of the fundamental frequency $\Omega/2\pi$ and of other parameters which define the particular shape of the curve (for example, the crest A, or the slope of a wave front, etc.). Let us now formulate the following question:

What happens to the magnitude u(t) if we modulate the fundamental frequency, that is, if we make

$$\Omega = \Omega_0 + \Omega_v(t) \tag{4}$$

where the variable part Ω_v reproduces some intelligence, and the constant Ω_0 represents a central or carrier frequency? The simplest and most tempting answer would be to carry the new expression of Ω as a function of t into the expression of u as function of u. Supposing that the parameters of form do not vary due to the modulation, the following would be obtained:

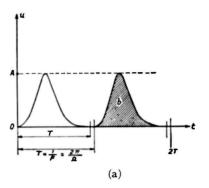
$$u = U_0[\Omega(t)]$$

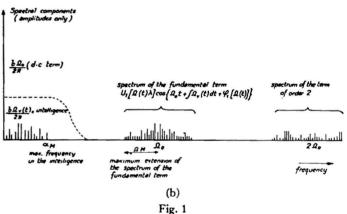
$$+ \sum_{n} U_{n}[\Omega(t)] \cos \left\{ n \int_{0}^{t} \Omega(t) dt + \phi_{n}[\Omega(t)] \right\}.$$
 (5)

* Decimal classification: R148.2×621.375.2. Original manuscript received by the Institute, August 28, 1946; revised manuscript received, July 11, 1947.

† Laboratoire d'Electronique et de Recherches Scientifiques Appliquées (LERSA) of the Philips Organization, Paris.

In the oscillatory terms, we have not replaced Ω by $\Omega(t)$, but Ωt by $\int_0^t \Omega dt$, for well-known reasons which relate to the physical signification of the frequency.





Let us suppose, first, that this operation of simple substitution of Ω by $\Omega(t)$ is legitimate; we shall come back to this point in a moment. It will be seen at once that, if the mean term U_0 is linear in Ω , a circumstance which, by (3), means that the area b is a constant independent of Ω , then the mean term, in the presence of the modulation, becomes

$$U_0[\Omega(t)] = \frac{C\Omega(t)}{2\pi} = \frac{b\Omega_0}{2\pi} + \frac{C\Omega_v(t)}{2\pi}$$
 (6)

and contains, in consequence, in its variable part, the intelligence completely separated.

More exactly, it is "separated" in the full sense of the word if its spectrum does not mix with that of one of the other terms of the whole wave modulated.

In Fig. 1 is shown, schematically, the spectral constitution of the signal u(t) modulated. From the viewpoint of the separation of the intelligence, the most dangerous oscillations are evidently the most extended ones, towards the beginning, of the spectrum of the fundamental term. Let Ω_{rM} be the maximum distance,

in frequency/ 2π , up to which said spectrum actually extends from Ω_0 , its center position. On the other hand, let α_M be the maximum frequency contained in the intelligence. For this intelligence to appear separated, the following must hold:

$$\alpha_M \ll \Omega_0 - \Omega v_M. \tag{7}$$

If this condition is satisfied, it will be sufficient, to recover the intelligence, to *filter* it out from the other components of Fig. 1 by means of a linear circuit. Of course, for said circuit to operate at ease, a sufficiently large interval must be arranged between α_M and $\Omega_0 - \Omega_{mM}$.

Now, it will appear clearly how the preceding facts can be used to solve the problem of detecting the intelligence in an ordinary frequency-modulated wave. First, as such a wave is sinusoidal (or very nearly so), it is deformed in a special step which is able to derive from it another wave such as the u(t) we considered above, and which presents an average value not zero and the same Ω as the incoming f.m. wave; more exactly, an average value which, in the case of the unmodulated incoming wave, is proportional to its frequency, $U_0 = b\Omega/2\pi$, where b is the area covered by one cycle of the deformed wave. If this coefficient remains constant in the presence of the modulation, and condition (7) is fulfilled, then said average value takes a variable part which reproduces the intelligence completely separated and, in addition, without any distortion!

If, on the other hand, the coefficient of area b is rendered independent of the amplitude of the received wave, a requisite which, as will be seen, is easily fulfilled, a system of detection of f.m. waves has been obtained, endowed with all the advantages which result from the classical "limitation of amplitude."

The production of a "continuous" magnitude directly proportional to the frequency of an incident oscillation is precisely the work done by the well-known direct-reading frequency meters based on the "counting" principle. Therefore, our f.m. detector is constituted by a frequency counter (self-"limited") followed by a low-pass filter.

We must now return to the hypothesis which we have made, according to which it is legitimate to consider that, if the distorted signal corresponding to the unmodulated wave of fixed frequency Ω is $u(t, \Omega)$, the signal corresponding to that "same" wave but modulated is obtained by replacing Ω by $\Omega(t)$ in the unchanged expression of $u:u[t, \Omega(t)]$. In other words, the question is to insure that the operation which we have described is quasi-stationary.

For it to be so, it is evidently necessary that the time required to establish the distorted signal be small compared to the duration of the most rapid variation contained in the modulation; i.e., $2\pi/\alpha_M$.

This condition will be amply satisfied by arranging matters so that the distorted signal consists of impulses whose full duration (of existence) τ is inferior to

the smallest period of the purely sinusoidal incoming wave which may appear, a period for whose value we may certainly adopt the expression $2\pi/\Omega_0 + \Omega_{vM}$. In fact, this last quantity is certainly inferior to $2\pi/\Omega_0 - \Omega_{vM}$, and this is greatly inferior to $2\pi/\alpha_M$ according to condition (7).

We will impose, therefore, the following condition (amply sufficient for the legitimacy of the quasi-stationary treatment):

Each period of the distorted signal is an impulse terminated at the end of τ seconds, $\tau \leq \frac{2\pi}{\Omega_0 + \Omega_{nM}}. \tag{8}$

The two conditions (7) and (8) suffice to guarantee the correctness of the system.

It is to be noted that, as a consequence of (8), the coefficient of area b can be written:

$$b = \int_0^{\tau} u(t)dt, \quad \text{or, as well,} \quad \int_0^{p} u(t)dt$$
 (9)

with the upper limit τ , or ∞ , in place of T, when it is understood that for u(t) we take the expression of one single impulse.

II. GENERALIZATION

It will be instructive to consider the common detection of amplitude-modulated waves, from the same viewpoint as the preceding principle for f.m. detection.

In a.m., the beginning of the process is exactly the same: the wave to be handled is distorted so as to establish a mean value not zero.

The distorter is none other than the common diode, which (ideally) splits the wave along the axis, and only permits the subsistence of the semiwaves of the same polarity. As in the case of f.m., it is found that, when the wave handled is modulated, the intelligence appears in the mean value thus created.

Only, in the product bT which represents said mean value in the presence of a pure wave, it is now T which is constant and it is the factor of area b which is left proportional to the modulated parameter; i.e., the amplitude A of the incident wave. Condition (7) remains necessary, as well as (8), which is reduced to $\tau \leq T$, and which is verified by the fact that here τ is T/2.

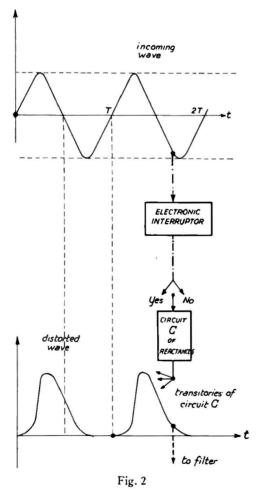
This observation seems to indicate that we are in the presence of a general principle which could be formulated as follows: to modulate a carrier wave is to vary one of its parameters without creating an average value; to detect said modulation is to distort the wave so as to create an average value not zero, a value in which the modulation is completely separated.

Both processes must be quasi-stationary, which presupposes, among other things, that the characteristic variations of the intelligence are slow compared with the carrier oscillations.

III. GENERAL OUTLINES OF THE COUNTING CIRCUITS

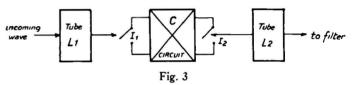
The classical idea for executing electronic counting consists in utilizing the wave, the frequency of which is to be measured, for commanding the electronic equivalent of an interrupter, so that the latter excites the transitory regime of a reactive circuit each time the wave being handled passes, for example, through zero in the "increasing" sense.

The transient thus provoked, which displays the role of what was called the "distorted signal" u(t), is, or rather should be, proper to the reactive circuit and independent of any parameter of the incident wave, except of its frequency; it must be terminated in a time delay shorter by a certain "reserve" than the period T (reserve for accommodating the smaller values which T may take as a consequence of the modulation). In other words, the incident wave is used only for "marking the cadence" of the distorted signal. This latter is then sent



to a filter which, when the incident wave is modulated, passes all the spectrum of the intelligence and eliminates the "high" frequencies, or, as is sometimes said in view of the impulsive character of the signal, the "peaks." See Fig. 2.

The electronic interrupter consists of a tube on the grid of which is applied the incident wave and which is adjusted so that the anode current varies from zero to maximum according to whether the grid is negative or positive. A little reasoning shows that, to realize a very cyclic regime, which means that the system returns to zero in all its parts (discharging again all that was discharged, and conversely) two interrupters are in general required, as shown in Fig. 3. One tube, L_1 , operated by the incoming wave, opens and closes I_1 on one side of circuit C. The "charge" (transient) of C when I_1 is opened is such that automatically tube L_2 , which flanks C on the other side, closes I_2 . When L_1 , under the influence of the incoming wave, closes I1, @ "discharges," in the opposite sense, this automatically causing I_2 to be opened by L_2 . The utilization branch may be situated, for instance, on the side of L_2 , in such a way that the unilateral impulse used is the discharge of Q through I_{2} - L_2 , the charge being made through L_1 . Practical circuits will soon illustrate these ideas.



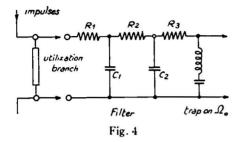
The use of *two* tubes naturally permits other methods for creating the unilateral impulses; for example, pushpull circuits. We will concern ourselves with systems of two tubes opening alternatively, of the type shown in Fig. 3, from which other systems may also be derived.

As circuit *C*, one may think of securing the simplest possible combinations with the minimum possible number of the elements "capacitor," "resistor," and "inductor." On low frequencies, it is, in fact, easy to find combinations of *R* and *C* alone which work well. On "high" frequencies—"high," for counters, is above 100 kc., for example—these combinations suffer from defects more and more inhibitory, and it becomes necessary to introduce the three elements *R*, *C*, and *L* at the same time.

Apart from the unilateral characteristics of the transient used, and of condition (8), we will impose the condition that said transient is not of an oscillatory character; it is then evident that it will embrace the maximum area b with the elements R, C, L given, and it is always interesting to make the factor b as large as possible (see (6)). Later, we will see a much more important reason for imposing this condition.

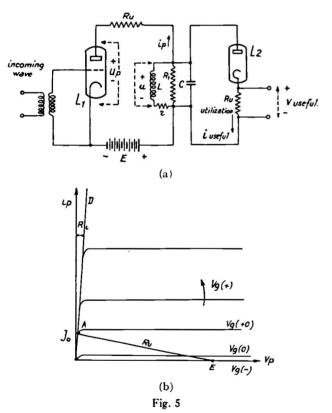
In the detailed study of practical counting circuits, we will consider as incident waves only the pure sinusoidal oscillations, for if, with such a wave, one has realized the preceding conditions, one is sure in advance that the detection will be correctly made on the modulated wave. On the other hand, to simplify, we will not show in our circuits the filter required for extracting the modulation. In this way, the reasoning is affected by an error which we will keep within acceptable limits

by agreeing once and for all to construct the filter on the model of Fig. 4, with $R_1 \ge Z$. This is the simplest model of audio-pass filter and the most efficient one (if it is desired to refine it, a trap may be interposed to specifically eliminate the frequencies near the fundamental Ω_0 of the "peaks").



IV. FIRST CIRCUIT

The most simple circuit is shown in Fig. 5(a). Fig. 5(b) represents the characteristics $(i_p - v_p)$ of the pentode L_1 . Tube L_2 is a simple diode. We beg the reader not to concern himself for the moment with the particular mode of connection (low point of R_u not grounded).



When the incident wave is zero, the grid voltage, $v_{\varrho}(0)$, is such that tube L_1 is just at the cutoff point, without current. The full supply voltage +E is on the anode, the voltage u on the circuit is zero. During the postive half-period, the grid voltage increases to values under which the current i_{ϱ} flows. At first, capacitor C being neutral, it appears as a short circuit; that is, the current in the coil is initially zero and the tube current appears at first through the capacitive branch,

without encountering any appreciable impedance other than resistance R_u . In other words, the tube begins to operate on the load line of Fig. 5(b). Let $v_g(+0)$ be the grid voltage for which said line cuts the characteristic $(i_p - v_p)$ at the elbow A situated on the "limit characteristic" OD. If $v_q(+0)$ is small compared with the final voltage assumed by the grid of L_1 , this value $v_q(+0)$ will be reached in so small a fraction of the positive half-period that we may continue to suppose C as discharged, and consequently equivalent to a short circuit, up to that moment. We will then have placed the operating point of the tube from E to A by what is called in mechanics a "percussion." From this moment on, all the characteristics $(i_p - v_p)$ with $v_q \ge v_q(+0)$ will have the line OD in common, the operation point will remain on said line, whereof the equation is $v_p = R_i i_p$, with R_i of the order of 100 ohms in good pentodes. Let us call i_0 the current value which corresponds to point A.

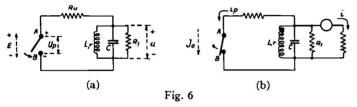
In fact, the percussion may be defined as a sudden short-circuiting of the space L_1 , so as to cancel the voltage +E which existed at its terminals; the small residual voltage which remains in the tube along the limit characteristic is negligible compared with E, and in any case it is possible to take account of it by incorporation of resistance R_i in R_u . After the short circuit is established, the current will gradually migrate from the capacitive branch, where it will pass from j_0 to 0, to the inductive branch, where it will pass from 0 to j_0 , while in the branch R_1 and in the tube L_p there will be an evolution which can be deduced from the preceding and of which it is possible to guess that, in R_1 , it will pass from zero to zero, and, in L_1 , from j_0 to j_0 .

If the circuit is adjusted so that the charge of the coil occurs in a regular manner, with the current i_L always increasing, the high point of the circuit will take a potential smaller than the low point; in other words, the voltage in the circuit, estimated in the sense indicated in Fig. 5(a), will be negative, and the diode L_2 will remain nonconductive and the branch in which it is placed will not intervene in the phenomenon. This shows, in passing, that the condition of a nonoscillating transient is not only favorable to sensitivity, but also necessary for the correct operation of closure and opening of the two complementary interrupters according to the model of Fig. 3.

In addition, we suppose that the circuit is adjusted in such a way that said unilateral charge is practically completed before the end of the semiperiod. Very close to said end, at the moment when the grid voltage of L_1 passes back through the value $v_g(+0)$, one will find the circuit in the following new "initial" state: current through the tube j_0 constant, circulating through R_u and coil L; branches C and R_1 neutral, in particular no voltage at the terminals of C; L_2 open. When v_g decreases below that value, the operating point of L_1 must leave the limit characteristic, and by reasoning of the same class as the preceding, it will be seen that it goes back from A to E by a "percussion" which cancels the current

 j_0 , and this quick annulment performs along the path constituted by the resistance R_u and the short circuit represented by the capacitor C in its discharge state. In other words, the interrupter L_1 reopens itself, reestablishing instantaneously a nonzero voltage at the terminals of the tube. This time, the voltage u on the circuit becomes positive, the diode L_2 closes, and the discharge of the coil takes place through the system of branches R_1 , C, and R_u (the R_u on the right of Fig. 5(a)). This is the useful period. Let us note that the tube L_1 is working along its limiting characteristics; i.e., practically in short circuit, a circumstance which is very convenient indeed for the power dissipation, which practically ceases to be a limiting factor.

Calculations will now make precise this qualitative analysis. We will write down the equations which govern the circuit in the hypothesis of unilateral charge and discharge without overlapping, and afterwards we shall get from them the necessary conditions to fulfill this hypothesis. In accordance with our preceding physical inquiry, the charging process consists of closing interrupter A, B of Fig. 6(a), canceling an initial voltage of +E. Therefore, the voltage can be calculated as the voltage which would exist permanently under voltage +E in (A, B) minus that other voltage which is provoked by a step-function impulse of height -E applied in (A, B).



Now, the first contribution to u, that of the permanent voltage +E, is obviously nil in our circuit, disregarding the d.c. voltage which could remain on the coil. The second contribution is obtained in operational form by using the operational transference between \mathcal{U}_p and \mathcal{U} (the script letters mean operational forms, or Laplacian transforms, of the magnitudes labelled by the corresponding common letters); i.e.,

$$U = \mathfrak{X}(p)U_p \tag{10}$$

and making, then:

$$U_p = -\frac{E}{p} \tag{11}$$

which is the well-known operational form of the stepimpulse mentioned above.

The generalized transference $\mathfrak{X}(p)$ can be taken from the ordinary symbolic calculus of the sinusoidal strata in which $i\omega$ is replaced by p. It will be found easily in our circuit:

$$3C(p) = \frac{1}{R_u C} \frac{p + \frac{r}{L}}{p^2 + p \left[\frac{r}{L} + \frac{1}{RC}\right] + \frac{1 + r/R}{LC}}$$
(12)

where we put:

$$R = \frac{R_1 R_u}{R_1 + R_u} = \text{value of } R_1 \text{ and } R_u \text{ in parallel.}$$
 (13)

Concerning the discharge period, the same is the outcome of a process which reduces itself to annul a constant current $i_p = j_0$ flowing in (A, B) by opening the interrupter. The current i which appears in the useful branch, as a consequence of this (see Fig. (6b)) is again the sum of the current—here 0—which corresponds to $i_p = \text{constant} = j_0$ supposed to flow permanently and that other current which corresponds to the application in (A, B) of a current-impulse of step-function form and of height $-j_0$. So, if we compute the operational transference of the currents:

$$\mathfrak{X}'(p) = \frac{\mathfrak{I}}{\mathfrak{I}_p}$$

we will have, in a completely analogic manner as before:

$$\mathfrak{I} = -\frac{j_0}{p}\mathfrak{K}'(p)$$

It is easily shown that both transfer functions \mathcal{K} and \mathcal{K}' , of voltages and currents, are identical (apart from the sign) if the two resistances R_u in series with the tubes L_1 and L_2 are identical, a condition we will satisfy.

Therefore, it will be sufficient to study the $\mathfrak{R}(p)$ of (12): the conditions which insure the correctness of the charge are identically the same as those which insure the correctness of the discharge. This symmetry between the two behaviors, obviously desirable, completely identifies the whole operation with the ideal of the two interrupters as represented by Fig. 3.

The formula for $\mathfrak{X}(p)$ conduces to an unilateral non-oscillating time curve for the generating function of $-E/p \cdot \mathfrak{X}$ only if the denominator has two real roots, which will be negative. From this we extract the following condition:

$$\left(\frac{r}{L} + \frac{1}{RC}\right)^2 \ge \frac{4\left(1 + \frac{r}{R}\right)}{LC} \tag{14}$$

We know that the most rapid transitory response is obtained when this condition is fulfilled with the = sign (critical damping). We have then

$$p_0 = \frac{1}{2} \left[\frac{r}{L} + \frac{1}{RC} \right] = \frac{\sqrt{1 + \frac{r}{R}}}{LC} \text{ (or >, but just)} \quad (15)$$

and the denominator of $\mathfrak{X}(p)$ in (12) has this $-p_0$ as a double root. The generating function of $-E/p(\mathfrak{X}(p))$ can then be written:

$$u = -\frac{E}{R_u C} \left[\frac{rC}{1 + \frac{r}{R}} + \left\{ \frac{\frac{1}{2} \left(\frac{1}{RC} - \frac{r}{L} \right) \sqrt{LC}}{\sqrt{1 + \frac{r}{R}}} t - \frac{rC}{1 + \frac{r}{R}} \right\} e^{-p_0 t} \right]$$
(16)

(these results can be found, in what concerns their mathematical aspects, in any of the current operational calculations treatises).

At t = 0, we have in fact u = 0. At $t = \infty$, the right-hand member tends towards

$$u_{\infty} = -\frac{Er}{R_u \left(1 + \frac{r}{R}\right)}$$

which is obviously the d.c. voltage reigning at the terminals of the circuit owing to the ohmic loss in the resistance r of the coil. If this resistance is sufficiently small with respect to R_u (and R), this residual voltage is practically zero. If its final value plays any part whatsoever, it is only to delay the closing of the diode for the discharge a little, but some little delay is not disagreeable to us, as in the qualitative theory, where we had disregarded this effect, it appeared that the beginning of the discharge was the moment where the signal passed a definite, yet positive, value $(v_q = v_q(+0))$ and not v_{q0}); i.e., a moment slightly in advance with respect to the very end of the half-period. This observation shows that it is important not to exaggerate the value of r (in other words, to work with coils of reasonably good quality) but once this condition is fulfilled, it can be dismissed from our thoughts, the r-dependent terms can be disregarded, and the impedance of the coil can be written simply Lp.

If we keep in mind this convention, to neglect terms in r which we are going to accept throughout the whole paper, our solution gets the following aspect:

$$p_0 = \frac{1}{2} \frac{1}{RC} = \frac{1}{\sqrt{LC}}; \quad \frac{1}{R} \sqrt{\frac{L}{C}} = 2$$
 (17)

$$u = -\frac{Et}{R_u C} c^{-p_0 t}. \tag{18}$$

It is easy to get a universal drawing for the u(t) by introducing dimensionless relative variables:

$$p_0 t = \frac{t}{2RC} = \frac{t}{\theta} = x \left(\theta = 2RC = \frac{1}{2} \frac{L}{R}\right). \quad (19)$$

$$\frac{R_1}{R_u} = \nu \tag{20}$$

$$\frac{u}{E} = y. (21)$$

With these parameters, we have

$$y = -\frac{2\nu}{1 + \nu} x e^{-x}.$$
 (22)

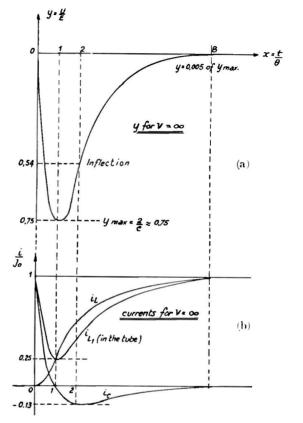


Fig. 7

This function y(x) is drawn in Fig. 7 for $v = \infty$. As we wished, it is negative and nonoscillating. In x = 1, i.e., $t = \theta$, it presents a maximum $y_{\text{max}} = -2/e \approx 0.75$, which means that the voltage on the circuit, in this case of $v = \infty$, rises (negatively) up to the three-quarter parts of the feeding voltage E at our disposal. In x = 8, y has fallen down to a few thousandths of its maximum value, so we will be generous if we take for "the duration" τ of the impulse (we consider the whole function as "an impulse"), the value

$$\tau = 8\theta = 16RC = 4\frac{L}{R} {23}$$

As the integral $\int_0^\infty xe^{-x}dx$ is equal to 1, and we have $dt = \theta dx$, the area under the impulse is equal, in absolute value, to

$$b = 2E\theta = Li_0. \tag{24}$$

We remain here in the case $v = \infty$; i.e., $R_1 = \infty$, so that R is the same as R_u ; and accordingly $\theta = (L/2R_u)$ and $j_0 = (E/R_u)$.

The result (24) can also be derived directly by noting that, in the branch L, we have

$$-u=L\frac{di}{dt};$$

therefore,

$$\int_{0}^{\infty} - u dt = L[i_{\text{final}} - i_{\text{initial}}]$$

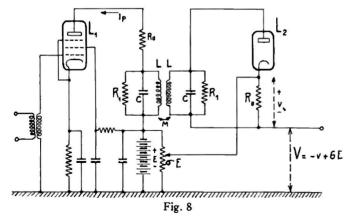
and, as i (final) in the coil L is j_0 (end of the charge), and i (initial) $\equiv 0$, we find (24) again. This reasoning has the virtue of making the result (24) a very intuitive one. From u(t), the values of $i_L = -(1/L)\int^t u dt$, $i_c = -C(du/dt)$, $i_{R_1} = (u/R_1)$, and i_{L_1} (in the tube $= i_L + i_c + i_{R_1}$) are easily deduced; they are represented in Fig. 7(b). If ν is not ∞ , meaning R_1 not infinite, all the values of u are reduced in the factor v/1 + v. Let us note by the way, that according to a former remark, the ratio u/E of the charge period is exactly equal, apart from the sign, to the ratio $i(useful)/j_0$ of the discharge period. As the utilization resistance in series with L_2 is the same as the one which determines the operating point in series with L_1 , we have

$$i_{useful} = (useful \ voltage)/R_u = \frac{v}{R_u}$$
 (25)

and on the other hand $j_0 = (E/R_u)$; so we will have, during the discharge,

$$\frac{v}{E} = \left| \frac{u}{E} \right|_{\text{of charge}} = \text{function } y(x) \text{ of (24) and Fig. 8.} \quad (26)$$

This shows that in order to get the best possible sensitivity, it is recommended to make $\nu = \infty$; i.e., $R_1 = \infty$. Nevertheless, it is still convenient to load the circuit, for instance, by $R_1 = 10~R_u$, in order to avoid the possibility that the internal resistance of the diode, which reaches very high values while the diode is closing, may extend the transitory period unduly. In the case when sensitivity is a less important factor than some other quality (see below), we will be able to make v finite and eventually use its value to take care of special requirements.



The design of a circuit intended to operate a wave of a maximum frequency Ω_M (this is the $\Omega_0 + \Omega_{rM}$ of the general theory of §3); i.e., of a minimum half-period of π/Ω_M , is readily accomplished by writing the condition (8) of the quasi-stationary regime, which contains in itself the condition of correctness (unilaterality) of the transitory:

$$\frac{\pi}{\Omega_M} = 4 \frac{L}{R} = 16RC$$

or, making conspicuous the R_u as a term of comparison:

$$\frac{L\Omega_M}{R_n} = 0.75 \frac{\nu}{1+\nu} \tag{27}$$

$$\frac{1/C\Omega_M}{R_u} = 5 \frac{\nu}{1+\nu}$$
 (28)

To this we add the formula giving the sensitivity, or still better, the full equation of the output voltage v as a function of the current frequency ω :

$$v = \frac{Lj_0\omega}{2\pi} = \frac{EL\omega}{2\pi R_u} = \frac{E}{2\pi} \frac{L\Omega_M}{R_u} \frac{\omega}{\Omega_M}$$

$$\approx \frac{0.12\nu}{1+\nu} E \frac{\omega}{\Omega_M} \text{ (for } \omega \leq \Omega_M). \tag{29}$$

With a given tube L_1 and a given feeding d.c. voltage E, R_u is imposed by the characteristic curves of the tube, and from there, all the design comes out in a very convenient form by rendering C in (28) equal to the smallest possible value; i.e., the parasitic capacitance of the tube output and connections. For instance, with the EL3 as tube L_1 , and E=320 volts, it results $R_u\approx 3200$ ohms and if we adopt (generously) $C_{\min}=20$ $\mu\mu fd.$ we have, with the choice $\nu=\infty$; i.e., $R_1\gg R_u$ (practically, R_1 of the order of 50,000 ohms), a limiting frequency of 500 kc. The grid voltages $v_{\sigma 0}$ for cutoff and $v_{\sigma(+0)}$ for the bending point A (see Fig. 5) are then of more or less -12 and 0 volts, so that the incident signal has to be at least 40 volts. The residual voltage in A on the anode is no more than 3 to 5 volts.

If the problem reduces itself to the construction of a simple frequency meter for pure sinusoidal waves up to 500 kc., such a design is perfectly convenient. Then it is sufficient to connect, in series with R_u and L_1 , a common d.c. ammeter of maximum sensitivity equal to $6.12j_0$ (in the preceding example, 12 ma.), which corresponds to the maximum incident frequency Ω_M , and according to (29) the deflection δ of the ammeter pointer for the frequency ω will be given by

$$\frac{\delta}{\delta_M} = \frac{\omega}{\Omega_M},\tag{30}$$

which means we have a perfect linear scale.

If the pursued aim is a detector of frequency-modulated waves, the situation is not quite so comfortable. As a matter of fact, it will then be necessary to send the useful voltage u which appears at the terminals of the circuit to the grid of another tube, and Fig. 5 shows that none of said terminals are grounded.

Should one be exclusively concerned with the sole task of detecting an ordinary modulation, the difficulty could be overcome by the use of a simple (but big) separating capacitor, which cuts the d.c. only and transmits everything else unchanged down to the lowest frequency contained in the intelligence. We have then, in (29),

$$\omega = \Omega_0 + \Omega_\nu(t),$$

and the useful voltage becomes,

$$v = v_{00} + v_{0u} \begin{cases} v_{00} = 0.12 \frac{\nu}{1+\nu} E \frac{\Omega_0}{\Omega_M} \\ = 0.12 \frac{\nu}{1+\nu} E \frac{\Omega_0}{\Omega_0 + \Omega_{vM}} \\ v_{0u} = 0.12 \frac{\nu}{1+\nu} E \frac{\Omega_v}{\Omega_M} \\ = 0.12 \frac{\nu}{1+\nu} E \frac{\Omega_v(t)}{\Omega_0 + \Omega_{vM}} \end{cases}$$
(31)

The part v_{00} , of d.c., does not interest us. The very useful part, v_{0u} , which reproduces the intelligence, can be applied to a grid. As, in this case, we will try to use a receiving tube as L_1 , we will find an important value of R_u which, by (28) would conduce, for $v = \infty$, to a much too low capacitance at $\Omega_M \approx 500$ kc. So we shall manipulate v. For instance, with an EF9 tube, E = 150 volts, $i_0 = 5$ ma. $R_u = 30,000$ ohms, $\Omega_0 = 465$ kc. (second intermediate frequency of normal receivers), $\Omega_{rM} = 75$ kc., $\Omega_M = 540$ kc., one gets an acceptable value of $20 \mu \mu fd$. for C by choosing v = 0, 1; i.e., $R_1 = 3.000$ and, by (31), v_{0u} , for the maximum excursion $\Omega_v(t) = \Omega_{rM}$, happens to be:

 $v_{0uM} = \text{max. useful voltage} = 0.0014E \approx 0.21 \text{ volt.}$

V. SECOND CIRCUIT

But the preceding solution of the connection problem is impossible when we wish to transmit the d.c. term voo as well, or at least very slow (infra-audio) variations of it. This case occurs in those f.m. transmitting systems where the frequency modulation is performed by controlling directly the parameters of the circuit of an auto oscillator. It is well known that, in this case, the necessary stability of the central frequency cannot be achieved, but by the use of an automatic-controlling link we are able to supply an "infra-audio" voltage v₀₀ which reproduces the slow variations of said central frequency and applying it back to the system which controls the auto-oscillating circuit, in order to counteract said slow variations. Therefore, it is clear that the f.m. detector placed in the heart of this control link must be able to transmit an infra-audio output.

One means to secure this is to separate the discharge circuit or utilization circuit L_2 from the charge circuit or excitation circuit, by a transformer. This solution was suggested and experimentally studied by Ziegler, whose name should be given to the novel circuit.

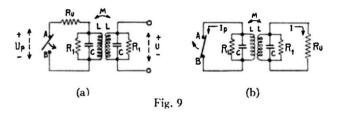
The complete layout is represented by Fig. 8, including a connection derived from the battery and whose role is to define a "zero," as will be explained below.

The qualitative theory is very much the same as for

the preceding circuit. It is necessary to give the coils of the transformer such a sense that, during the charge, when the current from low to high in the primary is increasing, the voltage from high to low in the secondary is negative. Then, we have a charge again through the primary coil suddenly short-circuited in the presence of the secondary, but with the branch L_2 passive, lasting until the value of current j_0 of point A of Fig. 5(b) is reached; and thereafter, a discharge of the primary through the secondary, from j_0 to 0, the branch L_1 being suddenly opened, and the branch L_2 closed, being now active in the secondary. In the same way as before, the voltage u during the charge is obtained operationally by

$$U = -\frac{E}{p} \Re(p) \tag{32}$$

where $\mathfrak{C}(p)$ is the voltage-transfer function (U/U_p) of the circuit in Fig. 9(a), being -E/p, the operational form of the step impulse of voltage of value E equivalent



to closing the interrupter (A, B) through u_p . Analogously, the current i during the charge is obtained by

$$3 = -\frac{j_0}{p} \, \mathfrak{F}'(p) \tag{33}$$

where \mathfrak{R}' is the current's transfer function $(5/5_p)$ of the circuit of Fig. 9(b), being $-j_0/p$, the operational form of the step-impulse of current of height j_0 which is equivalent to opening the interrupter (A, B) through i_p .

The following theorem can be demonstrated: Given a quadripole terminated by a definite branch, on the terminals of which the voltage is u, let us excite it by a voltage u_p from a constant-voltage source through a resistance R_u ; and let $\mathfrak{C}(p)$ be the generalized sinusoidal voltage transfer function $\mathcal{U}/\mathcal{U}_p$. At the terminals of said branch, let us now connect a resistance R_u' , and let i be the current which flows through R_u' when the quadripole is excited by a source of constant current i_p . Let us call, under this circumstance, $\mathfrak{K}'(p)$ the generalized current-transfer-function $3/3_p$. Theorem: \mathfrak{K}' is identical with \mathfrak{K} if the two following conditions are fulfilled: (a) quadripole symmetrical; (b) $R_u' = R_u$. (The preceding case of a single antiresonant circuit is a particular one of this theorem.)

This result obliges us to make the transformer symmetrical, which means equalizing all the elements of a same nature on both sides, a desirable feature from the constructional point of view. Once this has been done (it was supposed so in Figs. 8 and 9), we are sure that the

discharge will be exactly symmetrical to the charge (in our case, we do not have $\mathcal{K}' \equiv \mathcal{K}$, but $\mathcal{K}' = -\mathcal{K}$, because of our sense conventions), and it is sufficient, as in the precedent case, to study the voltage u during the charge.

The generalized sinusoidal transference function \Re of voltages in Fig. 9 is calculated in Appendix A, from which the following formula is obtained:

$$\mathcal{K} = \frac{kg_{\mathbf{u}}Lp}{\left[1 + (1-k)Lp(g_1 + Cp)\right]\left[1 + (1+k)Lp(g_1 + Cp)\right] + g_{\mathbf{u}}Lp\left[1 + (1-k^2)Lp(g + Cp)\right]}.$$
 (34)

where we put:

$$M = kL. (35)$$

The operational quantity whose generating function is the voltage u(t) looked for, being $E-3\mathcal{C}/p$, or, here $-Eg_uL/D(p)$, it is obvious that, in order that u(t) should not be oscillatory, it is necessary that the denominator D(p) has only real roots, which evidently will be negative.

The discussion of the 4th-degree polynomial D(p) can be made completely, with some graphic aid, a feature which is worth while to stress in order to sustain the confidence in operational methods. But as this discussion is too lengthy to be reproducible in the frame of a normal paper, we must content ourselves with giving a short outline of the method in Appendix B, and picking out, here, the following end results.

An over-abundant condition that the transitory will be unilateral is:

$$\frac{1}{R_1^2} \frac{L}{C} = \frac{4}{1-k}.$$
 (36)

Once this is fulfilled, the time of duration of the transitory is of the order of magnitude of the inverse of that (negative real) root of D(p) which is located nearest to the origin. A little consideration shows that this time delay can be taken as:

$$\tau = 6(2 + \nu) \frac{L}{R_1} \qquad \left(\nu = \frac{R_1}{R_u} = \frac{g_u}{g_1}\right).$$
 (37)

If we use this value to write down condition (8), and introduce, as before, resistance R_u as a comparison term, we find, grasping also (36), the following two equations:

$$\frac{L\Omega_M}{R_u} = \frac{\pi}{6} \frac{\nu}{2+\nu} \le \frac{0.5\nu}{2+\nu} \tag{38}$$

$$\frac{1/C\Omega_M}{R_u} = \frac{24}{\pi} \frac{\nu(2+\nu)}{1-k} \le \frac{6\nu(2+\nu)}{1-k} . \tag{39}$$

These are the resulting design equations of our problem. In order to obtain a practical discussion from these, let us add to them the sensitivity equation. For this, we observe that the *mean value* of the function $i(t) = \text{generating function of the operational } j_0 \mathfrak{C}(p)/p$ $= j_0 k g_u L/D(p)$ (see (34)), which is what we need to put in (6), can be gained without writing down i(t). In fact,

the operational form of i(t) taken at p=0, which by (32(b)) and (34) amounts to Mg_uj_0 . As the voltage v on the useful resistance R_u is R_u times this multiplied by $\Omega/2\pi$ (see (6)), and j_0 is E/R_u ; and if finally we take into account the special connection of Fig. 9 which combines -v with a fraction σ of the feeding voltage E, we have as net result:

by a well-known result, the quantity $\int_0^\infty i(t)dt$ is equal to

output =
$$\sigma E - v$$
, $v = M j_0 \frac{\Omega}{2\pi} = \frac{EM\Omega}{2\pi R_u}$, or, too, by (48):

$$v = \frac{E}{12} \frac{\kappa \nu}{1 + \nu} \frac{\Omega}{\Omega_M} \,. \tag{40}$$

As in the preceding case, the particular result $v=M\Omega j_0/2\pi$ could have been obtained by integrating directly the equation of the current in the coil, and taking into account the hypothesis that the discharge begins and ends at the two limits of the integral.

As has been stated, v is proportional to E, so that the whole arrangement can be "compensated" by making, for a predeterminated value Ω_0 on which it is desired to have output zero, $\sigma = M\Omega_0/2\pi R_u$. It is to be noted that the two parameters M and R_u which define this value can be made stable much more easily than the tuned circuits of common discriminators. In this fact lies an important advantage of frequency-counting detectors in frequency-stabilizing links.

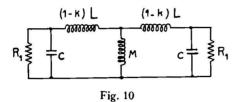
Equations (38), (39), and (40) allow discussing and designing any planned circuit under all circumstances, whatever the imposed data may be: tube (i.e., R_u), coupling factor k, minimum parasitic capacitance (i.e., (1/C) max), maximum frequency Ω_M , and so on. The explicit calculus of any design is so easy that it would be redundant to insist upon the matter. Orders of magnitude are a little smaller than in the case of one tube but can be illustrated roughly by the same figures.

A drastic feature of the preceding results is that the parameter $\nu = R_1/R_u$ plays an important part as a limitating factor, whereas in the case of a single circuit its value could well be infinite. Here, the value of ν only can be ∞ if, simultaneously, k=1, and then all will degenerate into the preceding case (the mathematical verification of this statement is left to the reader). As soon as k < 1, the transformer presents, by its leakage inductances, parasitic reactances which must be specially damped, if it is desired to avoid oscillations. As a matter of fact, it can be verified that the condition (36), if written:

$$\frac{(1-k)L}{R_1}=4R_1C,$$

represente the condition of critical damping, by the

resistance R_1 , of the resonant parasitic circuit formed by the total leakage inductance of the transformer (1-k)L and the capacitors C (see Fig. 10). Therefore, in the present case, it is necessary to anchor down the resistances R_1 immediately at the terminals of the coils, and the values of said parallel resistances R_1 become very rapidly small as soon as the coupling coefficient has those values which are common even in the most refined high-frequency transformers.



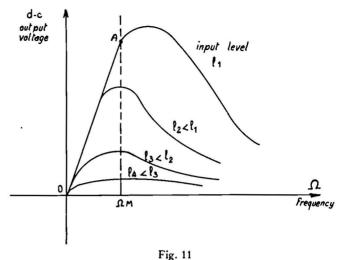
It is, moreover, obvious that the part played by this new factor makes it so that it is not always desirable to have the k as near to 1 as possible. See especially equation (39) which conduces to C nil when $k\rightarrow 1$. Here lies an important trap of which one cannot be aware until having completed the exact theory of the circuit.

VI. COMPLETE CHARACTERISTICS

It is interesting to consider in their full extension the family of curves which, in a simple frequency counter, not compensated, give the d.c. output voltage as a function of the frequency of the applied signal for various values of this signal's level.

For a sufficiently high level e_1 of the applied signal, and a circuit adjusted upon the maximum frequency Ω_{M} , we comply exactly with the foregoing theory; i.e., we have first, for $\Omega < \Omega_M$, a characteristic $v(\Omega)$ strictly straight, 0A, whose slope depends only from the circuit and not from e1. So we have, in this zone, an ideal f.m. detector; i.e., linear and auto-limited at the time. If, under a constant input level, we push the frequency beyond Ω_M , the condition for the complete fading out of the impulses before the end of the half-periods, is fulfilled worse and worse. The mean value of the areas of impulses, cut prematurely by the applied wave going back to zero, cannot increase any more and finally begins to decrease very fast, as soon as the areas happen to be cut in the initial zone of the impulses, where the major part of their value is concentrated. Thus we will get a curve such as the one labelled e₁ on Fig. 11.

For a level $e_2 < e_1$, but yet greater than the threshold, we have a similar curve, but the separation from the straight line 0A occurs before, because the duration of the initial percussion which closes the interrupters equivalent to the tubes is becoming significant as the time taken by the grid voltage to go from one level, v_{g0} , to another, $v_{g(+0)}$, is a fraction of the half-period so much the greater as the final height to which the signal raises within the half-period is lower. This is the same as saying that the independence of the area with respect to the period of the signal does not extend up to periods as short as before.



. .8. . .

For a level e3 still smaller, this independence ceases practically to exist for any frequency whatsoever, and the circuit ceases to behave as a counter. Finally, for levels such as e4 so small that the plate current does not limit itself at all, there does not remain any vestige of the impulses, the plate current reproduces the grid voltage with its smooth sinusoidal form, as in ordinary amplifier; i.e., the circuit becomes linear and the characteristic $v(\Omega)$ tends simply toward the ordinary "selectivity" curve, or "frequency-response" characteristic of the plate circuit, with exact proportionality of the levels, v = Ae. If the plate circuit is heavily damped, as it is in our case for the convenience of the frequency counting under higher levels, then said selectivity curve does not present the ordinary resonant peak. But this circumstance does not prevent the whole of the Ziegler circuit from transmitting without modification, and further rectifying by the diode in the ordinary form, oscillations of a level $e \le e_4$; the only difference is that it does it without appreciable frequency discrimination. This fact inspired in M. Ciancaglini (of the same Laboratory) the idea of using a counter as an unique step for detection in an universal a.m. and f.m. receiver. In that arrangement, we pass from the position f.m. to the position a.m. simply by dividing the input level to the detector stage by, say, 20 or 30, which is a very easy matter to achieve by an ordinary knob or push-button. In the simplicity with which the service of a receiver can so be changed lies a useful advantage indeed of f.m. detection by counting.

Over-All Study

(1) The preceding analysis is a schematic one only, owing to the fact that the two tubes were treated as mere interrupters. In a more complete study, it would become necessary to introduce the real curved characteristics of the tubes in the zones where they are working; i.e., the vicinity of the "limiting" straight portion in what concerns L_1 , and of the cutoff in what concerns L_2 .

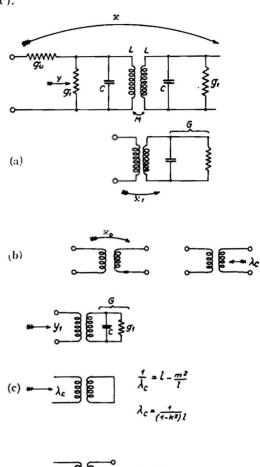
This will raise, of course, a nonlinear problem. But

experience shows that the results of the schematic analysis are not so far from reality as to make us worry about such a refinement. The difference between the measured facts and those predicted by the preceding schematic analysis comes much more from the difficulty of constructing elements L, C, M with exactly pre-established values, than from the residue of nonlinearity. If one decides first to construct a symmetrical transformer, measuring afterwards the values of its parameters L, M, C, and then adjust the working frequency and the load g₁ in order to fulfill the conditions of the preceding theory, one obtains an agreement between experiment and theory up to 10 per cent. This will be shown with more details in a further paper dealing with the practical side of the question, in which our results will be applied specially to the case of a link of automatic control of central frequency.

- (2) The frequency counter is intrinsically a detector of low over-all sensitivity. If we wish to illustrate sensitivity by the following figure, with a common receiver tube as L_1 and 75 kc. useful deviation, we have a few tenths of a volt of useful output for several ten volts of input. But we have already seen that the 0.2 volt of useful output is more than sufficient for the following audio steps, and that the several tens of volts of input are unavoidable in all detection systems which claim limiting action. As a matter of fact, in the detection of f.m. the incoming voltage displays much more the role of the local oscillation in a mixer than that of a signal to be reproduced. So the ratio of audio level to applied signal level is not a significant figure.
- (3) As the tube is used merely as an interrupter, it could well be replaced by a thyratron if the working frequency would fall to the 100-kc. zone where the modern gas tubes are still able to oscillate. Then the j_0 of the preceding theory can rise to more than ten times higher values than stated before, R_u becomes more than ten times lower, and the damping resistor R_1 can be a much higher fraction thereof. So, the irrelevant "sensitivity figure" (ratio of levels) rises up to values equal or greater than those encountered with discriminator detectors. This solution can be used specially in frequency stabilizing links, but requires some further circuit techniques which can not be dealt with in the present paper.
- (4) A very important point is the following: the common discriminator with resonant circuits is a very delicate device because it adds to the ordinary worries of double tuning, those much more ticklish which are presented, as it is well known, by all differential sets. Its tolerances lay at the extreme limit of actual mass production, and in particular, it is completely out of question to avoid the post-fabrication adjusting of the discriminator in the receivers, one by one. On the contrary, the frequency counter is a very strong structure, demanding but the easiest tolerance and no adjusting whatsoever is necessary once put in a receiver. This is an important feature from the economical standpoint.

APPENDIX I.

To derive the voltage transfer function of the circuit of Fig. 9 of the paper, it is convenient to reason systematically with the admittances. See Fig. 12(a), (b), and (c).



If we call y the admittance seen in Fig. 12(a), and \mathfrak{X}_1 the transference shown there, we have first:

$$3C = 3C_1 \frac{g_u}{g_u + Y}$$
 (41)

Let us begin to study \mathfrak{K}_1 . If we call G, provisionally, the total conductance $g_1+\gamma$ which loads the secondary $(\gamma=Cp, \text{ as } \lambda=1/Lp)$, we can discompose \mathfrak{K}_1 in the product of the transference \mathfrak{K}_0 by λ_c/λ_c+G , where λ_c is the admittance seen from the secondary backwards with source passivated; i.e., here by short-circuiting the primary. We have immediately:

$$3C_0 = \frac{m}{l} = k \tag{42}$$

(m and l are impedances Mp and Lp) so that:

$$3C_1 = 3C_0 \frac{\lambda_c}{\lambda_c + G} = \frac{k\lambda_c}{g_1 + \gamma + \lambda_c}$$
 (43)

and, putting in (41)

$$3C = \frac{g_u}{g_u + Y} \cdot \frac{k\lambda_c}{g_1 + \gamma + \lambda_c}$$
 (44)

Let us now calculate Y. We have, in Fig. 12(a) and 12(c):

$$Y = g + \gamma + Y_1 \tag{45}$$

and Y_1 can be written in function of the "load" G and the admittances λ_c and λ_0 at short-circuit and open-circuit, by a general formula of the theory of quadripoles:

$$Y_1 = \lambda_e + \frac{\lambda_e(\lambda_0 - \lambda_e)}{G + \lambda_e}$$
 (G, we recall, $= g + \gamma$)

from where, in Y:

$$Y = g_1 + \gamma + \lambda_c + \frac{\lambda_c(\lambda_0 - \lambda_c)}{g_1 + \gamma + \lambda_c}$$
 (46)

and putting this value in (44):

$$3C = k \frac{\lambda_c}{g_1 + \gamma + \lambda_c} \frac{g_u}{g_u + g_1 + \gamma + \lambda_c + \frac{\lambda_c(\lambda_0 - \lambda_c)}{g_1 + \gamma + \lambda_c}}.$$

Rearranging and substituting λ_c and λ_0 by their values, (see Fig. 12(c) we have:

$$3C = \frac{kg_u l}{1 + l(g_u + 2g_1 + 2\gamma) + (1 - k^2)l^2(g_1 + \gamma)(g_u + g_1 + \gamma)}$$

We can rearrange the denominator and finally get:

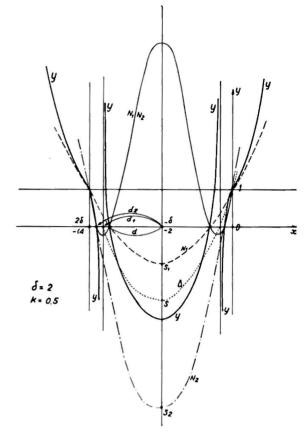


Fig. 13

$$\mathfrak{AC} = \frac{kg_u l}{\left[1 + (1-k)l(g_1+\gamma)\right]\left[1 + (1+k)l(g_1+\gamma)\right] + g_u l\left[1 + (1-k^2)l(g_1+\gamma)\right]}$$

If, in this formula, we make l = Lp, $\gamma = Cp$, we obtain the formula of the paper.

APPENDIX II

By using the dimensionless parameters.

$$x = p\sqrt{LC}$$

$$2\delta = g_1 \sqrt{\frac{L}{C}}$$

$$\frac{g_u}{g_1} = \frac{R_1}{R_u} = \nu,$$

the polynomial to be studied takes the form:

$$D(x) = N_1 N_2 + 2\nu \delta x \Delta$$

with:

$$N_1 = 1 + (1 - k)x(x + 2\delta)$$

$$N_2 = 1 + (1 + k)x(x + 2\delta)$$

$$\Delta = 1 + (1 + k^2)x(x + 2\delta).$$

The roots of D(x) = 0 are discussed by investigating the cutting of the curve $y(x) = \frac{N_1(x)N_2(x)}{\Delta(x)}$ by the straight

line $-2\nu\delta x$. The curve y can be constructed and discussed for all values of the two parameters k and δ . A typical example is given by Fig. 13, where we inscribe successively the three parabolas, $N_1(x)$, $N_2(x)$, and $\Delta(x)$, the product curve, $N_1 \cdot N_2$ and, finally, the gradient $y = N_1 N_2 / \Delta$. All the curves which serve as steps to construct y(x) are very easy to locate with the aid of their peaks S_1 , S_2 , S_3 , and their cutting points with the axis. All the curves have the vertical $(-\delta)$ as axis of symmetry. In the case represented, $(k=0, 5, \delta=2)$, we have obviously four cutting points with any straight line— $(2\nu\delta)x$. By keeping δ constant and raising k up to its limiting value 1, one sees that the evolution of the curve is such that for $1-k>1/\delta^2$, the central branch is above the -0x axis, so it is no more sure that the straight lines cut it in four points, and the over-abundant-condition of no oscillation becomes $1-k=1/\delta^2$, which is (36) of the paper. Once this is insured, an approximative value of the root which is located nearest the origin can be derived by reducing D(x) to its linear term, from which we get $x_0 = -1/2\delta(2+\nu)$. Going back to p by $x = p\sqrt{LC}$, and taking as duration of the transitory $6/p_0$, we obtain the value (37) of the paper.