SPACE CHARGE AS A CAUSE OF NEGATIVE RESISTANCE IN A TRIODE AND ITS BEARING ON SHORT WAVE GENERATION

By Lewi Tonks

Abstract

The mathematical theory of negative resistance in both plate and grid circuit of a triode has been worked out. Negative resistance is found when a virtual cathode is formed between grid and plate. In general this requires a plate voltage low compared to the grid voltage, a minimum electron current density depending upon the voltages used, and proper electrode spacing. For plane parallel construction the plate-grid distance must exceed the grid-cathode distance and for cylindrical construction the ratio of plate to grid diameter must exceed 2.15. Typical theoretical plate and grid characteristics are plotted. Failure of the experimental verification of these static characteristics through the occurrence of oscillations is not unexpected in view of the short relaxation time of the triode. *Applied to the short wave oscillations discovered by Barkhausen and Kurz* this theory in combination with that proposed by Gill and Morrell is able to explain their main features.

Gill[1] has reported oscillations in a triode which he has ascribed to unstable space charge, and he has given just sufficient mathematical treatment to substantiate the theory. Having been led to the same theory as a possible basis for explaining the type of high frequency oscillation first described by Barkhausen and Kurz,[2] I have developed it rather more completely.

Gill applied 40 v. to the grid and 8 v. to the anode of a triode in which the anode diameter was about five times the grid diameter, and then he gradually increased the electron current by increasing the filament temperature. At first anode and grid current increased in proportion. But soon a point was reached at which the anode current dropped and the grid current rose discontinuously. Further investigation showed that this phenomenon could be used to generate oscillations and these oscillations were observed. In triodes in which the ratio of anode diameter to grid diameter was less than five, neither current discontinuity nor oscillations of this type were found.

Before beginning the mathematical theory it will be well to obtain a qualitative idea of the space charge conditions which may exist between grid and anode. For simplicity consider a triode of plane parallel electrodes—cathode F at a distance a from grid G, and anode P at a distance c from G. The grid is supposed to be a continuous equipotential surface so that all electrons pass through it with the same speed (their thermal energy being neglected) along undeviated paths. Let P be maintained initially at cathode

[1] Gill, Phil. Mag. **49**, 993 (1925).

[2] Barkhausen and Kurz, Phys. Zeits. **21**, 1 (1920).

potential, i.e., zero, and let G be maintained at $+V_G$ volts. For small electron currents all electrons leaving F are accelerated until they pass through G. Then they enter a retarding field which reduces their velocity to zero at the surface of P. There they may be either reflected or collected. In either case the space charge in the GP space simply depends on the total electron current, that is, the sum of outward and return currents, since at any point an outbound and a return electron have the same speed and thus make equal contributions to the space charge.[1] Accordingly, the potential distribution between G and P is the same as that arising from an electron emission current originating at P and equal in intensity to the actual total current in the GP space.

This equivalence enables us to see what happens as the outbound current density i_1 is increased. The total current density I_1 increases at the same

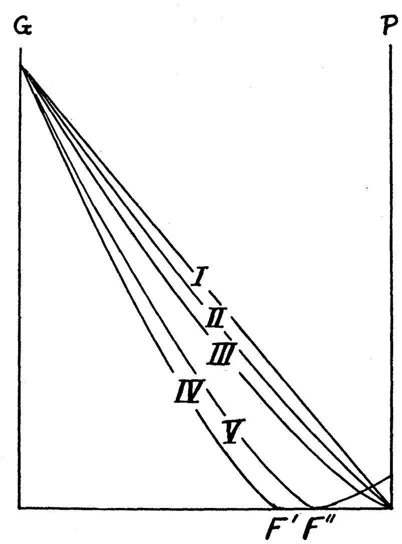

Fig. 1. Potential distribution between grid and plate for various current densities.

time and the potential distribution in the GP space passes successively through the states represented by Curves I and II, in Fig. 1, finally reaching Curve III representing the space charge saturated condition. Curve III is the graph of the space charge equation

$$V^{3/2} = I_1(\text{sat.})\ x^2/2.33\times10^{-6} \tag{1}$$

where

$$I_1(\text{sat.}) = 2.33\times10^{-6}V_G{}^{3/2}/c^2 \tag{2}$$

so that

$$V/V_G = (x/c)^{4/3} \tag{3}$$

It is clear that if I_1 be now increased above the value I_1 (*sat.*) we have a condition for which there is no corresponding simple emission state with P

as emitter. The equivalent state is space charge saturated emission from a fictitious plane electrode F' as shown by Curve IV. Such a fictitious electrode is called a *virtual cathode*[3] since it behaves theoretically and experimentally like a thermionic source of electrons. It is readily seen that the distance a_0 from G to F' is given by the ordinary space charge equation

$$I_1 = 2.33 \times 10^{-6} V_G^{3/2}/a_0^2, \qquad I_1 < I_1(\text{sat.}) \tag{4}$$

If, when the virtual cathode is at F', the potential V_P of P be made slightly positive, a small electron current reaches P and the virtual cathode shifts a small distance to the right into a new position F''. Denoting the current density i_P to P[4] by θi_1 and the separation of G and F'' by a, the space charge equations give

$$a = V_G^{3/4}/Bi_1^{1/2}(2-\theta)^{1/2} \tag{5}$$

$$c - a = V_P^{3/4}/Bi_1^{1/2}\theta^{1/2} \tag{6}$$

where $B = 656$ for practical units and $B = 3\pi^{1/2}(m/2e)^{1/4}$ for e.s.u. Adding, we find that

$$Bci_1^{1/2} = V_P^{3/4}/\theta^{1/2} + V_G^{3/4}/(2-\theta)^{1/2} \tag{7}$$

This equation relates θ, which is proportional to i_P, to both V_G and V_P. It shows that when V_P is small compared to V_G, θ, and, therefore, i_P vary as $V_P^{3/2}$, just as in the case of emission from a real cathode.

Eq. (7) may be transformed into another simpler form. Noting that $a = a_0$ when $\theta = 0$, Eq. (5) becomes

$$Bi_1^{1/2} = V_P^{3/4}/2^{1/2}a_0$$

When this is used to eliminate $Bi_1^{1/2}$ from Eq. (7) and we put V for V_P/V_G and η for a_0/c we obtain the reduced dimensionless equation,

$$V^{3/4} = \theta^{1/2}[1/2^{1/2}\eta - 1/(2-\theta)^{1/2}] \tag{8}$$

for the volt-ampere characteristic of P.

Instability is very apt to arise when an electrode possesses "negative resistance." In the present case the resistance of P is proportional to $\partial V/\partial\theta$. From Eq. (8)

$$\frac{\partial V}{\partial \theta} = \frac{4V^{1/4}}{3\theta^{1/2}}\left(\frac{1}{2^{3/2}\eta} - \frac{1}{(2-\theta)^{3/2}}\right) \tag{9}$$

[3] The first reference to the existence of a virtual cathode was made, so far as the author is aware, by Mr. D. C. Prince of this laboratory in an unpublished laboratory report dated June 8, 1922. There he calculated the position of the virtual cathode in the case of a grid with an axial filament.

[4] The hypothesis that all the electrons enter the GP space with the same velocity leaves no mechanism to select the electrons which are to pass on to P, but for mathematical handling it is simply sufficient to observe that some do pass on. Actually, of course, the electron beam is slightly non-homogeneous and the selection is accomplished by very small changes in the minimum voltage at the virtual cathode accompanying changes in the voltage of P.

and

$$\partial V/\partial \theta \leqq 0 \quad \text{when} \quad \theta \geqq 2(1-\eta^{2/3}) \tag{10}$$

Since θ is at most equal to unity it is necessary that

$$\eta \geq 1/2^{3/2} = 0.354 \tag{11}$$

That is, whenever with V_P zero a virtual cathode is formed more than 0.354 of the way from G to P, then, theoretically, P will possess negative resistance for some value of V_P.

Eq. (8) makes it possible to plot a single family of curves which will apply to any plane electrode P. Such a family is plotted in Fig. 2 with the θ-axis as axis of ordinates and V-axis as axis of abscissae. The characteristics

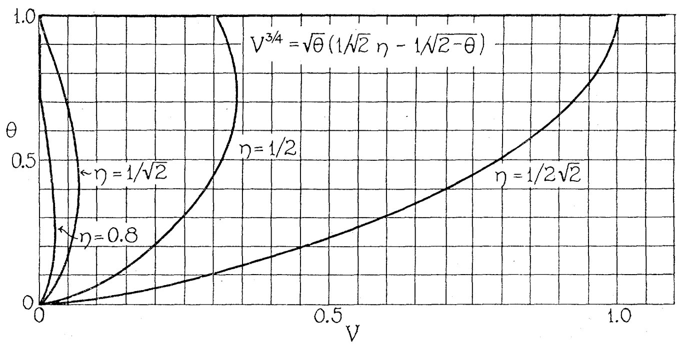

Fig. 2. Family of curves relating V and Θ.

for the smaller values of η (F' nearest G) follow the line $\theta=1$ after striking it, those for the larger values of η strike the θ-axis first, following it to $\theta=1$ and the.eafter following the latter.

These cu.ves show, as in every case of arc-type negative resistance, that for one applied voltage there may be three different currents, of which only two are stable. Qualitatively, how are these two possible? The answer is not difficult. When θ is small the GP space is carrying a total current of almost $2i_1$, and a virtual cathode is formed some distance from P on account of the large space charge. The large distance between F'' and P in turn, prevents more current from reaching P. But if, by any means, θ be increased sufficiently, even for a moment, then the space current falls to i_1 which, in some cases, will not be sufficient for the formation of a virtual cathode. Then the whole current i_1 will reach P and another stable state is attained. Between these two lies the unstable state in which the total space current exceeds i_1 only slightly, so that F'' is comparatively close to P and θ is comparatively large.

Thus far the whole argument has been based on the supposition that during the current and voltage changes which have been discussed, the electron current has remained constant, and the discussion has neglected the possibility of negative resistance in the grid circuit. The theory may be extended, however, so as to yield grid characteristics and to cover variations of i_1 by making some reasonable assumptions concerning the collection of electrons by G and the behavior of the electrons which have returned to the FG space through the grid. We shall assume that of the current i_0 leaving the cathode, the fraction ϕ is intercepted by G and the fraction ϕ' (equal to $1-\phi$) is passed by G. Thus

$$i_1=\phi' i_0 \tag{12}$$

Of the current i_1 a fraction θ reaches P so that $(1-\theta)i_1$ returns toward G. Of this we shall again suppose the fraction ϕ to be intercepted by G leaving the current i_2 to pass into the FG space. The double passage through G has by now given the electrons of i_2 some transverse velocity so that they do not approach F closely before returning toward the grid. Accordingly they do not make a full contribution to the space charge. We estimate their contribution at one-half normal. The space charge in the FG space corresponds, therefore, to a current

$$I_0=i_0+i_2/2+i_2/2=i_0+i_2 \tag{13}$$

Finally, we assume that the electrons of i_2 are all collected on their return to G, which is equivalent to neglecting their contribution to space charge after their third grid passage.

Before proceeding, it will be well to establish complete analytical relations between the currents i_0, i_1, i_2, i_G, i_P and I_0. Eqs. (12) and (13) give two of the necessary five equations. The three additional ones are readily seen to be

$$i_0=i_G+i_P$$
$$i_2=\phi'(1-\theta)i_1$$
$$i_P=\theta i_1$$

From the five the following useful equations can be derived by suitable eliminations:

$$i_G=(1-\phi'\theta)i_0 \quad =I_0(1-\phi'\theta)f(\theta) \tag{14A, B}$$
$$i_P=\phi'\theta i_0 \quad =I_0\phi'\theta f(\theta) \tag{15A, B}$$
$$i_1=\phi' i_0 \quad =I_0\phi' f(\theta) \tag{16A, B}$$

Where

$$f(\theta)=1/[1+\phi'^2(1-\theta)]$$

Examining the physics of the situation it is seen that when V_G is small and of the same order as V_P, c/α being not too great, then P will take the full current i_1 because any potential minimum between G and P lies above zero potential. This condition of the tube may be called State A. In this state i_G, and hence $J_G=B^2 i_G$ which is plotted instead in Fig. 3, is proportional to $V_G^{3/2}$ as shown by the line Oz. Here $\theta=1$ and I_0 is fixed by space charge according to the space charge equation

$$B^2 I_0=Z_G/\alpha^2 \tag{17}$$

where Z_G is put for $V_G^{3/2}$ for convenience. As V_G is increased State A persists until the potential minimum strikes zero potential, forming a virtual cathode. Referring to Fig. 3, for $V_P=9.7$ v. this occurs at w. At this point the virtual cathode begins to limit the current to P, θ begins to decrease and i_G begins to increase more rapidly. The tube is now in State B. I_0 is still fixed by Eq. (17). Increasing V_G further, the end of the space-charge-limited condition is reached at x where i_0 becomes equal to the cathode emission i_e. State C begins here. From this point on i_0 is constant and the virtual cathode

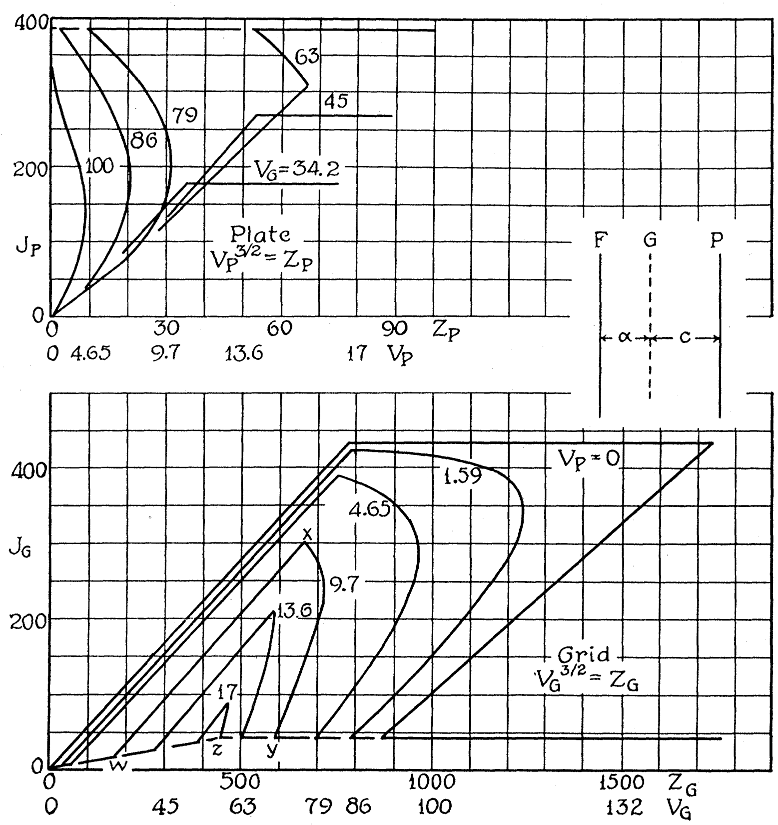

Fig. 3. Theoretical volt-ampere characteristics of triode for $\alpha=1$ and $c=1.5$ cm.

approaches P causing i_P to increase and i_G to decrease. The curves of Fig. (3) show that in State C V_G cannot be indefinitely increased if the tube conditions are to vary continuously. The reasons are those that have been discussed already in the simple case involving grid and plate only. Finally at y the virtual cathode again disappears through rising of the potential minimum, θ becomes equal to 1, and the tube is in State D. In this state i_G appears as the horizontal straight line at $J_G=43$ in the figure.

With this description in mind it is a simple matter to write the volt-ampere equations for each state.

State A. Substitution of Eq. (17) in Eqs. (14B) and (15B), noting that $\theta=1$ gives

$$J_G=(1-\phi')Z_G/\alpha^2 \qquad (18)$$

$$J_P=\phi' Z_G/\alpha^2 \qquad (19)$$

Where

$$J_G = B^2 i_G, \qquad J_P = B^2 i_P.$$

State B. Substituting Eq. (16B) in Eq. (7) to introduce I_0 in place of i_1, and then noting that I_0 is fixed by Eq. (17), we easily find

$$Z_P/Z_G=\theta\left[c/\alpha(\phi'(2-\theta)f(\theta))^{1/2}-1\right]^2/(2-\theta) \qquad (20)$$

Again, from Eqs. (17), (14B) and (15B)

$$J_G=(1-\phi'\theta)f(\theta)Z_G/\alpha^2 \qquad (21)$$

$$J_P=\phi'\theta f(\theta)Z_G/\alpha^2, \qquad (22)$$

parametric equations in θ which allow J_G and J_P to be plotted as functions of Z_G when Z_P is fixed, or vice versa.

State C. Under emission limitation i_0 is constant and equal to i_e. Using Eq. (16A) to replace i_1 by i_0 in Eq. (7) we then have

$$c(\phi' J_e)^{1/2}=(Z_P/\theta)^{1/2}+(Z_G/(2-\theta))^{1/2} \qquad (23)$$

where $J_e=B^2 i_e$. From Eqs. (14A) and (15A) we have for the currents,

$$J_G=(1-\phi'\theta)J_e, \qquad J_P=\phi'\theta J_e, \qquad (24) \quad (25)$$

another set of parametric equations.

State D. From Eqs. (24) and (25) for $\theta=1$

$$J_G=(1-\phi')J_e, \qquad J_P=\phi' J_e \qquad (26) \quad (27)$$

independent of Z_G and Z_P.

Characteristics calculated from these equations show the existence of negative resistance in both grid and plate circuits. Fig. (3) shows J_G plotted against V_G (Z_G) for various plate voltages for the following set of constants:

Cathode-grid separation $=\alpha=1$ cm.
Grid-plate separation $=c=1.5$ cm.
Electron emission $=i_e=10^{-3}$ amp. cm^{-2}
Grid capture fraction $=\phi=0.1$

The plate current against plate voltage curves for the same constants also appear in Fig. (3). The effect of increasing i_e is to change the J_G, J_P, Z_G, and Z_P scales in the same proportion, so that not only would the marginal scale numbers have to be increased but the individual voltages applying to each curve would also have to be increased.

Figure (4) is a plot similar to Fig. (3) and for the same constants, except α, which in this case is taken to be 1.363 cm. It is noted that for V_P greater than about 1.35 v. the characteristic is composed of two unconnected branches, the two straight lines Oz and zs forming one, and a closed loop the other. If α be increased until $\alpha^2/c^2=\phi'$ it is found that point z moves to point v and all characteristics for $V_P>0$ consist of two branches; and if α be further increased until $\alpha^2/c^2=2\phi'/(1+\phi'^2)$ the point t coincides with point u, and the loop system vanishes. It is doubtful whether any significance can be attached to these mathematical peculiarities. They are mentioned simply to give a more complete picture of the degeneration of the curve systems.

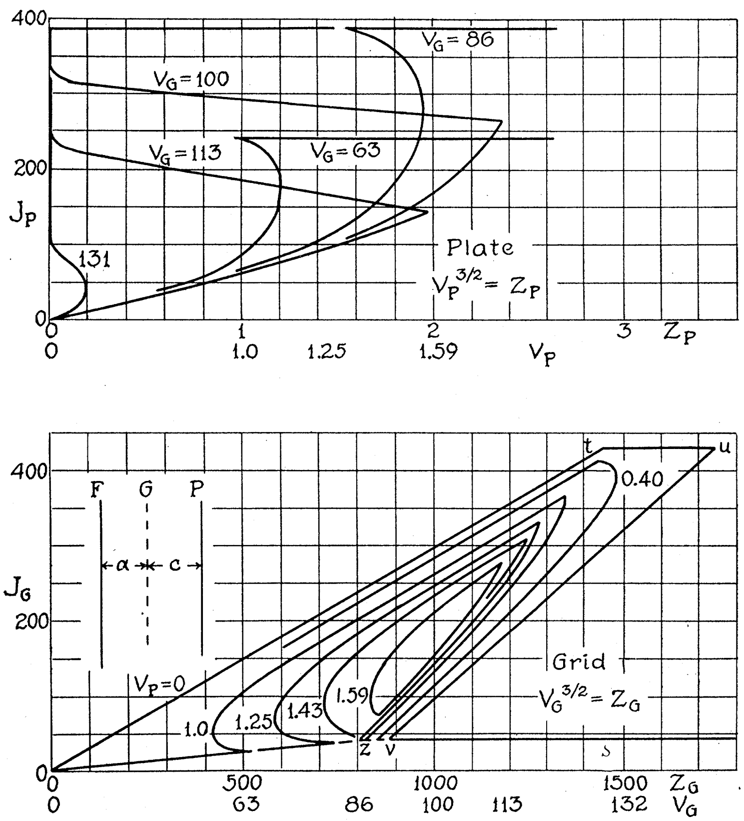

Fig. 4. Theoretical volt-ampere characteristics of triode for $\alpha=1.36$ and $c=1.5$ cm.

There are two factors which may affect the shapes of these characteristics, electron reflection (or secondary emission) from the plate and the transverse velocities imparted to the electrons by the grid. As long as there is a virtual cathode between G and P, or even a potential minimum at which the potential is not more than $\frac{1}{2}V_P$, perhaps, few electrons are able to leave P. Therefore this factor does not affect the xy portion of the $V_P=9.7$ v. grid characteristic in Fig. 3, for example, but may well result in a turning upward of the characteristic somewhat to the right of y. This is probably enhanced by the second factor which, in the first place, makes the whole virtual cathode phenomenon less distinct, and secondly allows less electrons to reach P

the higher V_G on account of the greater transverse velocities imparted. The effect of the two factors on the plate characteristics is possibly a downward slant in their $J_P=387$ portions. The essential features of the curves are, however, unaffected.

The present theory has been developed for a triode of plane parallel structure for mathematical reasons. But any tests must almost of necessity be made with cylindrical tubes because of the difficulty of obtaining a plane electron emitter. It becomes necessary, then, to adapt the present theory to the cylindrical case.

The condition for the existence of a virtual cathode is of primary importance. A virtual cathode is possible when in any cross-section of the triode (and except for a small correction due to grid captures) the space charge limited current, j_0, from the cathode to the grid exceeds the space charge limited current, j_1, which the grid could draw from the plate as emitter, that is, when

$$j_0 > j_1$$

Now, in the cylindrical case[3]

$$\left.\begin{aligned} j_0 &= V_G^{3/2}/A_c r_G \beta_{FG}^2 \\ j_1 &= V_G^{3/2}/A_c r_G \beta_{PG}^2 \end{aligned}\right\} A_c = (9/2^{3/2})(m/e)^{1/2}$$

where β^2 denotes the function of the ratio of collector to emitter radius which has been tabulated by Langmuir and Blodgett.[5] For a virtual cathode, then, we must have $\beta_{PG}^2 > \beta_{FG}^2$. In the usual single wire filament tube the grid is about 100 times the filament diameter and β_{FG}^2 is not very different from 1.08. Accordingly $\beta_{PG}^2 > 1.08$, and from the table, $r_P/r_G > 2.15$. This is the condition analogous to $c > \alpha$ in the case of the plane tube.

Beyond this single quantitative feature, the qualitative results for the plane case can be bodily transferred to the cylindrical one.

It is of utmost importance to reproduce the theoretical volt-ampere characteristics experimentally, if possible. In determining the volt-ampere characteristic of a device possessing negative resistance oscillations are usually prevented by loading the oscillating circuit with an excess of ohmic resistance. The possibility still remains, however, that the circuit may oscillate in some higher frequency mode. In most cases this can be prevented by making the possible parasitic frequency high compared to the relaxation time of the device, so that were the device to oscillate, it could no longer follow its static characteristic and the effective negative resistance would be very much reduced or obliterated.

In the present case there is the peculiar difficulty that the relaxation time of the triode is extremely short. The time required for an electron to pass from grid to virtual cathode and back to grid gives its order of magnitude. For a 0.5 cm diameter grid at 40 v. operating under space charge limited conditions, this is about 5.4×10^{-9} sec. The tube leads themselves form possible oscillating circuits having natural periods comparable with this.

[5] Langmuir and Blodgett, Phys. Rev. **22**, 353 (1923).

For example an oscillating circuit consisting of two concentric cylinders of 0.5 and 1.0 cm diameter, respectively, and 2.5 cm long joined by a 4 cm loop of 0.05 cm wire has a natural period of 3.1×10^{-9} sec. Accordingly no matter how heavily the external circuits of the triode are loaded with resistance, parasitic oscillations confined practically to the tube itself may arise. Thus it may well be extremely difficult if not impossible to prevent the triode from oscillating in the range where it shows negative resistance. As a matter of fact, I have not been able to stop the very high frequency oscillations by any system of loading the circuit which I have tried.

Turning now to a possible relation between the present theory and short wave oscillations of the type discovered by Barkhausen and Kurz, we find that several important features of those oscillations can be accounted for.

(1) According to the theory, negative resistance is only to be found when a virtual cathode exists in the tube. In a cylindrical triode this can occur only if $r_P/r_G>2.15$. Experimentally it has been found that no high vacuum triode oscillates if $r_P/r_G\leqq 2.0$.[6]

(2) The theory calls for an intimate dependence of oscillations on space charge conditions. Several of the tubes tested in the course of an experimental investigation showed this dependence clearly in that the three-halves power law relating current to voltage was brought out in the oscillation behavior.

(3) The characteristics of Fig. 3 lead one to expect oscillations at the point where emission limitation of current sets in. This was found to be the case in many instances. According to Fig. (4), on the other hand, when corresponding to c but slightly greater than α, r_P is but slightly greater than $2r_G$, then oscillations may be expected even when the tube is in the space-charge-limited condition. This also, has been verified. But not all the experimental results can be handled so easily.

Many tubes fail to exhibit the space charge dependence of the oscillations through the three-halves power law. Another difficulty lies in the fact that the strongest oscillations occur with the plate at a negative voltage and only weak oscillations occur when the plate is positive as required by the theory. Once oscillations have begun one can understand how current can reach a negative plate and oscillations continue, but one would expect this voltage arrangement to be somewhat unfavorable. However, the outstanding phenomenon which is inexplicable on this theory is the frequency behavior of the oscillations.

In general, experimenters have found that tuning of the triode circuits affects the frequency but little, if at all. The frequency depends apparently on the time an electron remains in the grid-cylinder space. Thus the period of the oscillation is seldom very different from this time as calculated and the variation of this time with grid and cylinder voltages gives the approximate variation of the oscillation period.[7] To explain this behavior Gill and

[6] The experiments of Kapzov, Zeits. f. Physik **35** (2), 129 (1925) are of particular interest in this connection. Several tubes having $r_P/r_G\leqq 2.0$ oscillated with mercury vapor present but stopped oscillating as the pressure was reduced. The vapor may well cause a redistribution of potential that makes a virtual cathode possible.

[7] A. Scheibe, Ann. d. Physik **73**, 54 (1924).

Morrell[8] have advanced a dynamic theory. They considered a plane triode with plate at filament potential and, supposing electron currents to be so small that space charge played no part, they studied the energy interchanges between each electron in the grid-plate space and a small alternating electric field superposed on the static field between grid and plate. They found that as the impressed frequency is increased continuously, the a.c. resistance between grid and plate decreases to a maximum negative value at a period 1.33 times the time taken by an electron to travel from grid to plate and back to grid. Thereafter the resistance oscillates indefinitely as higher and higher frequencies are attained.[9]

Blurring effects probably prevent negative resistance in the overtones, leaving the fundamental as the only possible self-sustaining oscillation. This theory is thus able to explain the frequency behavior. Incidentally it requires zero voltage on the plate, which agrees with experiment better than does the static theory.

We have, then, two independent theories, each of which is capable of explaining those characteristics of the oscillations which the other cannot, and each supplements the other in its weakest feature. For, the dynamic theory applies directly to the high frequencies at which the static theory approaches the limit of its qualitative applicability, and the static theory handles space charge of such amount as to seriously complicate the dynamic theory. Accordingly, it appears highly probable that the Barkhausen-Kurz type of oscillation arises from a joint action of the two causes.[10]

In the near future I hope to publish the experimental results on which some of the general conclusions cited here are based.

Research Laboratory,
General Electric Co.,
June 9, 1927.

[8] Gill and Morrell, Phil. Mag. **44,** 161 (1922).

[9] Through an oversight the authors have missed this feature in the behavior of the resistance. Referring to their table, p. 174, there is a region of positive work between $p=\pi/T$ and $p=2\pi/T$.

[10] Other theories have been proposed to account for these oscillations. Van der Pol, Physica **5,** 1 (1925) suggests that an excess of electrons returning through the grid causes a decrease in emission which later causes an excess emission, etc., thus setting up an oscillating condition. Several objections to this theory immediately suggest themselves. First, the frequency of oscillation calculates to one-half the actual value. Second, the electron excess combined with the deficiency would constitute a more nearly normal space charge, tending to decrease rather than accentuate the irregularity. Third, no oscillations could occur when the electron current is emission limited. J. Sahánek, Phys. Zeits. **26,** 368 (1925) has observed certain oscillations and proposes a theory which he believes may, with slight modification, be applicable to the Barkhausen-Kurz type. The theory is difficult to follow. It depends, apparently, upon the voltage drop along the filament. According to it oscillations are possible only when $2<r_C/r_G<5$, and it may give the correct value for the frequency.