THE DAVID W. TAYLOR MODEL BASIN

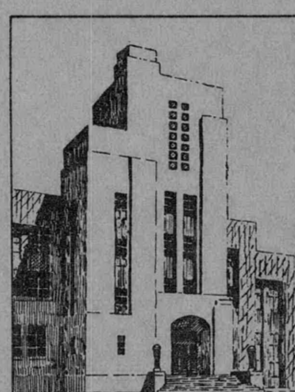
UNITED STATES NAVY

INTRODUCTION TO NON-LINEAR MECHANICS

PART IV

RELAXATION OSCILLATIONS

BY N. MINORSKY, Ph.D.



SEPTEMBER 1946 REPORT 564

NAVY DEPARTMENT

DAVID TAYLOR MODEL BASIN

WASHINGTON, D.C.

REPORT 564

INTRODUCTION TO NON-LINEAR MECHANICS

PART IV

RELAXATION OSCILLATIONS

BY N. MINORSKY, Ph.D.

SEPTEMBER 1946

FOREWORD

The report on the introduction to non-linear mechanics as a whole falls into four major divisions.

Part I, published as David Taylor Model Basin Report 534 under date of December 1944, is concerned with the topological methods; its presentation substantially follows the "Theory of Oscillations" by Andronow and Chaikin. The material is slightly rearranged, the text is condensed, and a number of figures in this report were taken from the book. Chapter V, concerning Liénard's analysis, was added since it constitutes an important generalization and establishes a connection between the topological and the analytical methods, which otherwise might appear as somewhat unrelated.

Part II, published as David Taylor Model Basin Report 546 under date of September 1945, gives an outline of the three principal analytical methods, those of Poincaré, Van der Pol, and Kryloff-Bogoliuboff.

Part III, published as David Taylor Model Basin Report 558 under date of May 1946, deals with the complicated phenomena of non-linear resonance with its numerous ramifications such as internal and external subharmonic resonance, entrainment of frequency, parametric excitation, and the like.

Part IV, published here, reviews the interesting developments of Mandelstam, Chaikin, and Lochakow in the theory of relaxation oscillations for large values of the parameter μ.

Part IV also contains a subject index to all four parts of this treatise.

TABLE OF CONTENTS

INTRODUCTION TO NON-LINEAR MECHANICS

PART IV

RELAXATION OSCILLATIONS*

127. INTRODUCTORY REMARKS

The term *relaxation oscillations*, introduced by Van der Pol (1)† (2) and commonly used at present, generally designates self-excited oscillations exhibiting quasi-discontinuous features. Because of the importance of such oscillations in applications in connection with the so-called "sweep circuits" in electronics, television, and allied fields, an extensive literature exists on this subject, References (3) through (9). Ph. LeCorbeiller (10) gives an interesting survey of various devices, both mechanical and electrical, by which these phenomena can be demonstrated; some of these devices have been known for centuries.

The characteristic feature of these phenomena is that a certain physical quantity (such as coordinate, velocity, etc., in mechanical problems, and charge, current, etc., in electrical problems) exists on two levels, remaining on each level alternately for a relatively long time but passing from one level to the other so rapidly that in the idealized representation the passage may be considered as instantaneous. A few examples taken from the paper by LeCorbeiller will illustrate these phenomena.

Figure 127.1 shows a device consisting of a container C, of the form shown, fastened to a support R capable of rotating about an axis A, perpendicular to the plane of the paper, and provided with a weight W sufficient to hold the system against the stop S. The container is slowly filled with water, and at the instant when the moment due to the weight of the water becomes greater than the moment due to the weight W, the system tumbles over against another stop S′. The container then empties, and the weight W brings the system back against the stop S, after which the filling period begins again, and so on. In this system the two levels, previously mentioned, are the angles θ and θ' at which the system is constrained by the stops

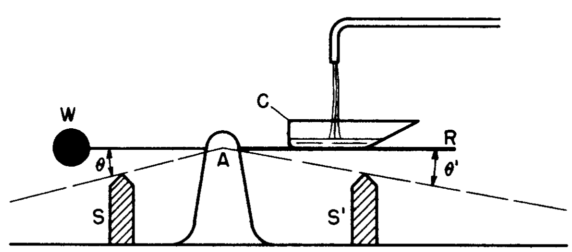


Figure 127.1

---

* The text of Part IV follows the presentation in "Theory of Oscillations," by Andronow and Chaikin, Moscow, (Russian), 1937. A complete bibliography on the subject of relaxation oscillations appears in this volume.

† Numbers in parentheses indicate references on page 66 of this report.

S and S$'$ respectively, and the representation of this system in the (θ, t)-plane appears as a periodic rectangular ripple with the length of its horizontal stretches determined by the rates of filling and of evacuation of C.

Another familiar example is the charging of a capacitor shunted by a gaseous conductor such as a neon lamp. During the charging period the capacitor's voltage gradually rises. At the point of ionization of the gaseous conductor, the neon lamp flashes and the capacitor is suddenly discharged, whereupon the gaseous conduction ceases abruptly and the charging period begins anew. Here, again, there are two levels, the voltage V_1 immediately before the discharge strikes and the voltage V_2 immediately after the extinction of the discharge. The transition from V_2 to V_1 is gradual, but the inverse transition from V_1 to V_2 is quasi-discontinuous. A phenomenon of this kind is represented in the (V, t)-plane by a so-called "saw-tooth" curve.

It was shown in Section 37, Part I, that a similar situation exists for large values of the parameter μ in the Van der Pol equation. Figure 37.1c shows quasi-discontinuous changes in the variable $x(t)$ between the two regions in which it changes but little.

When these periodic phenomena are represented in the phase plane of the variable undergoing rapid changes, they appear as closed curves with regions of very large curvature, such as the curves shown in Figures 37.2c and 37.3c. By idealizing these very rapid changes as discontinuous changes, closed trajectories of this kind become *piecewise analytic curves* "closed" by the discontinuous stretches.

Available analytical methods are inadequate for a rigorous treatment of these phenomena. In fact, all analytical methods presuppose that the parameter μ appearing in the basic quasi-linear equation

$$\ddot{x} + x = \mu f(x, \dot{x}) \qquad [127.1]$$

is very small. On the contrary, in some of these oscillations, which are expressible by Van der Pol's differential equation, this parameter is large. More specifically, in Figures 37.1c, 37.2c, and 37.3c, referred to above, the value of the parameter μ is 10.

Attempts have been made to extend the analytical methods to oscillations in which μ is large. In Section 36 it was shown that Liénard succeeded in obtaining certain conclusions regarding the qualitative aspect of the phase trajectories when μ was very large. N. Levinson (11) extended the proof of the existence of closed trajectories to cover oscillations in which μ is not small. In a recent publication (12) J.A. Shohat has indicated a form of series expansion formally satisfying the Van der Pol equation when μ is large. These various attempts, however, did not result in any complete analytical theory comparable to the one which has been studied in Part II in connection with oscillations in which μ is small.

Moreover, as will appear below, not all known relaxation oscillations seem to belong to the group of equations [127.1] of which the Van der Pol equation is a particular example. More specifically, it will be shown in Chapter XXI that relaxation oscillations are frequently observed in systems which are amenable to representation by differential equations of the *first* order which do not admit any analytic periodic solutions for the simple reason that these equations do not possess singularities, without which no closed analytic trajectories can exist; see Section 25. These difficulties led the school of physicists under the leadership of L. Mandelstam and N. Papalexi to evolve a theory, called by its authors the *discontinuous theory of relaxation oscillations*, whose exposition and applications will form the principal topic of Part IV.

The use of the concept of mathematical discontinuities for the purpose of describing a rapidly changing dynamical process, at least during certain instants of its evolution in time, is not new. It is recalled that the classical theory of mechanical impacts uses precisely the discontinuous method by assuming an infinitely small duration of the impact process during which the dynamics of the process is entirely ignored, and the "initial" and "terminal" conditions are correlated on the basis of certain *additional information* not contained in the differential equations themselves. This permits obtaining the correct overall effect of the impact without knowledge of its details. For an elastic impact, such additional information is supplied by the theorems of momentum and kinetic energy; to this information is added, for non-elastic impacts, the so-called *coefficient of restitution*, an empirical factor characterizing the loss of energy during the impact. This coefficient depends on the material of which the colliding bodies are composed.

One can imagine a discontinuous periodic motion generated by impacts from the following example given by Andronow and Chaikin (13). Let us assume that a perfectly elastic ball rolls without friction on a horizontal plane and strikes elastic walls WW perpendicular to the direction of its motion, as shown in Figure 127.2a. With the usual discontinuous treatment of mechanical impacts, the phase

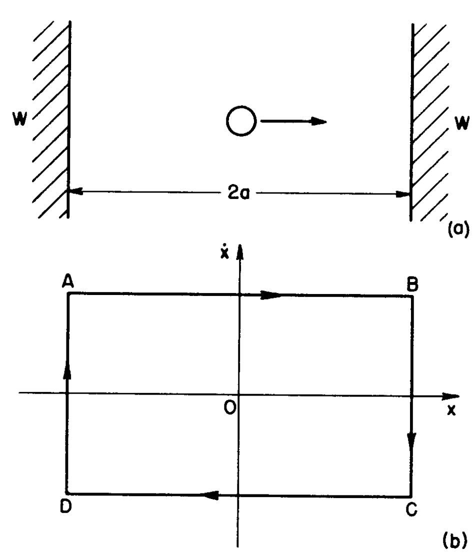


Figure 127.2

trajectory of such motion is represented by a "closed curve" ABCDA, a rectangle. On the branches AB and CD the motion is continuous with the constant velocities $\dot{x} = \pm v_0$, respectively; on the branches BC and DA, on the contrary, it is discontinuous.

In the systems of Figures 127.1 and 127.2 we encounter periodic phenomena having certain quasi-discontinuous features. The nature of these discontinuities in the two systems is, however, different. For the ball striking the walls there exists a definite external actuation, the reaction of the constraint, the wall, applied to the dynamical system, the ball; this actuation is properly "timed" by the distance $2a$ between the walls which determines the "period" of the motion. On the contrary, no external impact excitation exists during the discontinuities in the motion of the container shown in Figure 127.1. These discontinuities are due rather to a sudden loss of equilibrium between the moment M_W of the constant weight and that of the container M_C, occurring at a certain critical value, $M_C = M_W$. The change of equilibrium position from θ to θ' is not instantaneous, of course, but in comparison with the long periods of filling and evacuation of the container it may be considered as such in the idealized picture of the phenomenon. We can improve the idealization by making the moment of inertia of the system (W,C) relatively small; this will render the short time interval during which θ varies still shorter, which, in turn, makes the relative time intervals of filling and of evacuation still longer.

We find it expedient to define as *relaxation oscillations* those quasi-discontinuous oscillations in which the rapid changes between certain levels of a physical quantity occur as the result of the loss of a certain internal equilibrium in the system, and as *impulse-excited oscillations* those quasi-discontinuous oscillations in which these rapid changes are due to the action of certain external impulsive causes.

On this basis, the quasi-discontinuous oscillations of the container shown in Figure 127.1 are of a relaxation type, whereas the ball rebounding between the walls is an impulse-excited phenomenon. The essential difference between the two types of oscillations is that in relaxation oscillations the energy content stored in the system remains constant during the quasi-discontinuous changes of certain variables, whereas in impulse-excited systems, on the contrary, the energy content changes abruptly.

Impulse-excited oscillations do not require any particular additional information for their treatment, as will be seen in Chapter XXIV. For relaxation oscillations proper (see Chapters XXI, XXII, and XXIII) it is necessary to specify the conditions under which the discontinuities are bound to occur in a system; it will be shown that the basic assumption given in

Section 130 provides a criterion sufficiently broad to cover all known types of relaxation oscillations.

Finally, inasmuch as the representation of a rapidly changing process by a mathematical discontinuity is always an idealization, it becomes necessary to analyze the conditions under which this idealization is justified in practice. In analyzing the behavior of the device shown in Figure 127.1, we have noted that the change from the angle θ to the angle θ' may be considered as quasi-discontinuous. The smaller the moment of inertia of the system, the more accurate is the approximation. In mechanical systems, such as the examples described above, it is obviously difficult to extend the hypothesis by assuming that the moment of inertia is zero, but in electrical systems neglect of one of the oscillatory parameters is a common practice. In both types of systems, instead of a "full" differential equation of the second order, the abbreviated or *degenerate* equation of the first order is frequently employed. By using degenerate equations, numerous problems can be treated as discontinuous in the phase plane; this appreciably simplifies their solution.

A simple example will show the application of degenerate equations for this purpose. Let us consider an oscillating circuit shown in Figure 127.3 comprising an inductance L, a capacity C, and resistors R and r as shown. The circuit may be closed on a source of d-c voltage by a switch S. It is useful to specify certain idealizations which appear somewhat trivial but which will be found to be of considerable importance in what follows.

We assume first that the left branch of the circuit ALRB has no capacity and that the right branch ACrB has no inductance. In other words, we neglect the effect of small, parasitic, distributed capacities in the inductance L and resistance R; likewise, we neglect the effect of a small inductance accompanying the flow of current in the branch ACrB.

The second assumption will be that the opening and closing of S is instantaneous, that is, occurs in an infinitely small time interval (t - 0, t + 0).

Let us assume that we open and close the switch S at some frequency. The process will then be represented by a sequence of very long intervals, when S is either closed or opened, separated by infinitely short intervals of closing and opening. During the long intervals there will generally be a

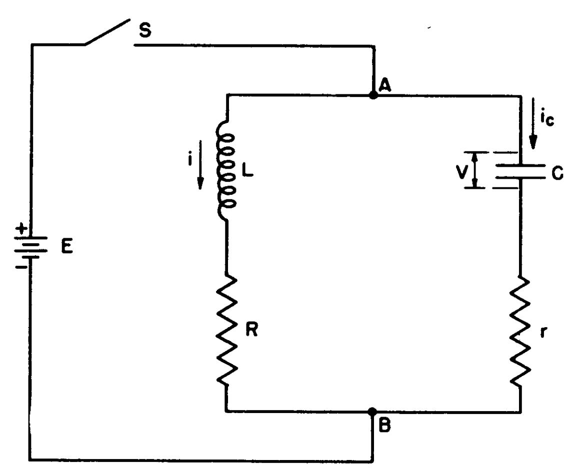


Figure 127.3

certain oscillatory process in the circuit $(L,C,R+r)$ describable by a differential equation of the second order whose phase trajectories are spirals converging toward a focal point, as we know from Section 5. The closing or opening of S will disturb this process by introducing a certain transient. Without any loss of generality we may consider the first closing of S at $t_0 = 0$ and assume that for $t < t_0$ the circuit was "dead." It is noted that, at the instant of closing, the two circuits ALRB and ACrB are *in parallel*, and the differential equations are

$$L\frac{di}{dt} + Ri = E; \qquad rC\frac{dV}{dt} + V = E \qquad [127.2]$$

where $E = 0$ for $t \leqq t_0$ and $E =$ constant for $t \geqq t_0$. With these assumptions one finds that

$$i = \frac{E}{R}\left(1 - e^{-\frac{R}{L}t}\right); \qquad \frac{di}{dt} = \frac{E}{L}e^{-\frac{R}{L}t}$$
$$V = E\left(1 - e^{-\frac{1}{rC}t}\right); \qquad \frac{dV}{dt} = \frac{E}{rC} \qquad [127.3]$$

It is seen that for $t = t_0$ the solutions of Equations [127.2], $i(t)$ and $V(t)$, are continuous but not analytic in the sense that their first derivatives undergo discontinuous jumps $\left(\frac{di}{dt}\right)_0 = \frac{E}{L}$; $\left(\frac{dV}{dt}\right)_0 = \frac{E}{rC}$. Noting that $i_c = C\frac{dV}{dt}$ we can take instead of $\frac{dV}{dt}$ the variable i_c and state that under the assumed idealizations the functions $\frac{di}{dt}$ and i_c undergo discontinuities $\frac{E}{L}$ and $\frac{E}{r}$, respectively. If one takes the plane of the variables $\left(\frac{di}{dt}, i_c\right)$, the process occurring at $t = t_0$ is represented by a discontinuous jump of the representative point from the origin to the point A whose coordinates are $\frac{E}{r}$ and $\frac{E}{L}$; see Figure 127.4. After the initial discontinuity the subsequent motion of the representative point will follow a continuous trajectory AB which, as the transient dies out, will eventually approach a damped oscillatory motion represented by a convergent spiral. If, at a later instant $t = t_1$, the switch is opened, another jump will occur, but this jump will generally not bring the representative point back to the origin but to some other point A″, and so on. It is thus seen that in the phase plane of these particular variables $\left(\frac{di}{dt}, i_c\right)$ the history of the system will be represented by a sequence of disconnected, spiral arcs "joined" by discontinuous stretches such as OA, BA′, ··· . We thus obtain a *piecewise analytic representation* of such a phenomenon.

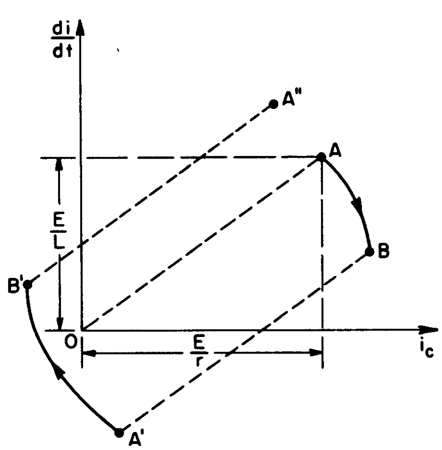


Figure 127.4

It is important to emphasize once more that such a representation of a quasi-discontinuous phenomenon by discontinuities in the phase plane of certain variables is possible only because we have introduced certain idealizations into the problem.

a. We consider the time interval during which rapid changes occur as an infinitely short interval.

b. The effect of the parasitic parameters is neglected, which enables us to deal with the *degenerate* equations [127.2] of the first order instead of the full equations of the second order.

c. We have selected the dynamical variables $\frac{di}{dt}$ and i_c which are capable of undergoing discontinuities under Assumption b.

The necessity for Assumptions a and b is obvious. As for Assumption c, it is clear that if one selected some other variables, for example, i and V instead of $\frac{di}{dt}$ and i_c, discontinuous representation in such a phase plane would be lost since these variables are continuous.

In the preceding discussion we have tacitly assumed an impulse-excited phenomenon as previously defined. For a pure relaxation phenomenon we must answer an additional question, namely, how to determine the instants (in the time representation) or the points (in the phase-plane representation) at which the discontinuity occurs in a system of this kind.

It is impossible to go beyond this point without formulating some kind of *a priori* assumption, as will be explained in Section 130.

CHAPTER XX

FUNDAMENTALS OF THE DISCONTINUOUS THEORY OF RELAXATION OSCILLATIONS

128. SOLUTIONS OF A DIFFERENTIAL EQUATION IN THE NEIGHBORHOOD OF A POINT OF DEGENERATION

We now propose to investigate the nature of the solutions of a differential equation of the second order with constant coefficients, for example,

$$a\ddot{x} + b\dot{x} + kx = 0 \qquad [128.1]$$

when one of the coefficients approaches zero. In an electrical problem, $a = L$, $b = R$, and $k = 1/C$; in a mechanical one, $a = m$ (mass), b is the coefficient of "velocity damping," and k is the spring constant.

First, if b approaches zero, one readily sees that the oscillatory damped motion approaches the oscillatory undamped motion. We saw that in the phase plane the solutions of [128.1] with $b \neq 0$ but small are spirals approaching a stable focal point; this remains true as b approaches zero. For $b = 0$, the origin is a vortex point and the trajectories are closed. It is thus seen that there is a definite difference between the qualitative aspect of trajectories when b is very small and that when b is equal to zero. From a practical standpoint, however, there is hardly any difference between the two cases; no discontinuities of any kind exist in the solutions.

Of greater practical interest are the cases when either $a \to 0$ or $k \to 0$. We shall examine first the case when $a \to 0$. It is apparent that when $a = 0$ Equation [128.1] becomes an equation of the first order and its solution is given in terms of *one* constant of integration, namely,

$$x = x_0 e^{-\frac{k}{b}t} \qquad [128.2]$$

where x_0 is that constant. By differentiating [128.2] we obtain

$$\dot{x} = -\frac{k}{b} x_0 e^{-\frac{k}{b}t} = -\frac{k}{b} x \qquad [128.3]$$

It is seen that the coordinate x and the velocity $\dot{x}$ are not independent but are related by Equation [128.3]. In other words, in the phase plane the trajectories of Equation [128.2] are reduced to a single line $y = -\frac{k}{b}x$, and the rest of the plane is not involved. This fact can be expressed by stating that the phase space of a differential equation of the first order is *unidimensional*, that is, it is a *phase line* instead of a phase plane.

The limit case, $a = 0$, never actually occurs in practice since in any electrical system containing resistance and capacity there is always a small residual or "parasitic" inductance. Likewise, mechanical systems without inertia are only idealizations. For these reasons it is preferable to investigate the effect of a small coefficient a in the solution of Equation

[128.1] rather than to drop this coefficient in the differential equation itself.

The solution of [128.1] is

$$x = C_1 e^{r_1 t} + C_2 e^{r_2 t} \qquad [128.4]$$

where C_1 and C_2 are the constants of integration and r_1 and r_2 are the roots of the characteristic equation

$$ar^2 + br + k = 0$$

If the initial conditions $t = 0$, $x = x_0$, and $\dot{x} = \dot{x}_0$ are given, one obtains

$$x_0 = C_1 + C_2; \qquad \dot{x}_0 = C_1 r_1 + C_2 r_2 \qquad [128.5]$$

from which one obtains the values of C_1 and C_2:

$$C_1 = \frac{x_0 r_2 - \dot{x}_0}{r_2 - r_1} \quad \text{and} \quad C_2 = \frac{x_0 r_1 - \dot{x}_0}{r_1 - r_2} \qquad [128.6]$$

where

$$r_{1,2} = -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{k}{a}} \approx -\frac{b}{2a} \pm \frac{b}{2a}\left(1 - \frac{2ak}{b^2}\right) \qquad [128.7]$$

In this expression only one term is retained in the expansion of the square root, since a is small. This gives

$$r_1 = -\frac{k}{b} \quad \text{and} \quad r_2 = -\frac{b}{a} + \frac{k}{b} \approx -\frac{b}{a} \qquad [128.8]$$

If the values [128.6] of the constants and [128.8] of the approximate expressions for the roots r_1 and r_2 are substituted in Equation [128.4], the approximate solution $x_1(t)$ of [128.1] is in the form

$$x_1(t) = x_0\left[e^{-\frac{k}{b}t} - \frac{ak}{b^2} e^{-\frac{b}{a}t}\right] + \frac{a}{b}\dot{x}_0\left[e^{-\frac{k}{b}t} - e^{-\frac{b}{a}t}\right] \qquad [128.9]$$

It is to be noted that the solution $x_1(t)$ is an approximate one because the expansion of the square root has been limited to the first two terms; this is justified by the assumed smallness of a.

On the other hand, for $a = 0$ the solution of Equation [128.1] of the second order becomes the same as that of the equation of the first order given by [128.2]. To emphasize the fact that the solution [128.2] is the same as that of Equation [128.1] when *complete degeneration* occurs, that is, when $a = 0$, we will write it as

$$\bar{x}(t) = x_0 e^{-\frac{k}{b}t} \qquad [128.2]$$

Consider now the function

$$\phi(a,t) = x_1(a,t) - \bar{x}(a,t) = -x_0 \frac{ak}{b^2} e^{-\frac{b}{a}t} + \frac{a}{b}\dot{x}_0\left[e^{-\frac{k}{b}t} - e^{-\frac{b}{a}t}\right] \qquad [128.10]$$

This function represents the difference between the approximate solution $x_1(t)$ of [128.1] in the neighborhood of the point of degeneration, where a is very small, and the solution $\bar{x}(t)$ of a completely degenerated equation [128.1] of the first order. The function $\phi(t)$ approaches zero uniformly in the interval $0 < t < \infty$ when $a \to 0$.

The expression for the derivative of this function is

$$\dot{\phi}(a,t) = \dot{x}_1(a,t) - \dot{\bar{x}}(a,t) = \left(x_0 \frac{k}{b} + \dot{x}_0\right) e^{-\frac{b}{a}t} - \frac{ak}{b^2} \dot{x}_0 e^{-\frac{k}{b}t} \quad [128.11]$$

For very small values of t the function $\dot{\phi} \approx x_0 \frac{k}{b} + \dot{x}_0$, and it is impossible to reduce it by reducing the coefficient a. However, for a sufficiently large t, which is supposed to be fixed, one can always find a value of a small enough so that the value of $\dot{\phi}(t)$ is smaller than a given positive number ϵ.

We can express this by saying that whereas the function $\phi(a,t)$, considered as a function of a, approaches zero *uniformly* in the interval $0 < t < \infty$ when $a \to 0$, the function $\dot{\phi}(a,t)$ behaves in a like manner only when the values of t are sufficiently large. For $t = 0$ the convergence of the function $\dot{x}_1(t)$ to the function $\dot{\bar{x}}(t)$ when $a \to 0$ is not uniform. In other words, to a given a, however small, one can always assign a value of $t = t_1$ such that $\dot{\phi} \approx x_0 \frac{k}{b} + \dot{x}_0$. Only in a very special case, when $x_0 \frac{k}{b} + \dot{x}_0 = 0$, does this non-uniformity of convergence disappear, but this case is of no practical interest.

One can also state that the difference between the approximate solution $x_1(t)$ of a quasi-degenerate system when a is very small and the corresponding solution $\bar{x}(t)$ of a completely degenerate system when $a = 0$ approaches zero in the whole interval $0 < t < \infty$ when $a \to 0$ except in a very small neighborhood around the point $t = 0$; this neighborhood is smaller as a is smaller and in it the difference $\dot{x}_1(t) - \dot{\bar{x}}(t)$ of the slopes of the two curves $x_1(t)$ and $\bar{x}(t)$ cannot be reduced. This means that the function $x_1(t)$ undergoes a quasi-discontinuous jump in this neighborhood.

When the parameter $k \to 0$, the problem is treated in a similar manner. First, for a completely degenerate equation with $k = 0$, Equation [128.1] becomes

$$a\ddot{x} + b\dot{x} = 0 \quad [128.12]$$

Integrating it, one obtains

$$a\dot{x} + bx = M \quad [128.13]$$

where M is the constant of integration. The value of M is determined by the initial conditions, namely,

$$a\dot{x}_0 + bx_0 = M \quad [128.14]$$

The solution of Equation [128.13] is

$$x = \frac{M}{b} + Ce^{-\frac{b}{a}t} \qquad [128.15]$$

where C is a constant of integration. One obtains finally

$$\bar{x}(t) = x_0 + \dot{x}_0 \frac{a}{b}\left(1 - e^{-\frac{b}{a}t}\right) \qquad [128.16]$$

If, however, one proceeds with the solution of Equation [128.1] in the neighborhood of its degeneration, where k is very small, the approximate solution is

$$x_1(t) = x_0 e^{-\frac{k}{b}t} + \frac{a}{b}\dot{x}_0\left(1 - e^{-\frac{b}{a}t}\right) \qquad [128.17]$$

By forming the functions

$$\psi(k,t) = x_1(k,t) - \bar{x}(k,t) \quad \text{and} \quad \dot{\psi}(k,t) = \dot{x}_1(k,t) - \dot{\bar{x}}(k,t) \qquad [128.18]$$

one ascertains by an argument similar to that given in connection with the functions $\phi(a,t)$ and $\dot{\phi}(a,t)$ that for a sufficiently small k the function $\dot{\psi}(k,t)$ approaches zero when $k \to 0$ uniformly in the interval $0 < t < \infty$, whereas $\psi(k,t)$ approaches zero when $k \to 0$ for all values of t except when $t \to \infty$, for which value $\psi(k,t)$ approaches the value x_0.

129. IDEALIZATIONS IN PHYSICAL PROBLEMS

In applications, idealizations of quasi-degenerate systems as absolutely degenerate systems are frequently made, as was mentioned in Section 127. Thus, for example, the differential equation of the so-called (R,C)-circuit is usually written in the form

$$R\frac{di}{dt} + \frac{1}{C}i = 0 \qquad [129.1]$$

where i is the current in the idealized circuit ($L = 0$). The corresponding quasi-degenerate equation, with L very small, is

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{C}i = 0 \qquad [129.2]$$

With the solution of the absolutely degenerate equation [129.1] designated by $\bar{i}(t)$ and that of the quasi-degenerate equation [129.2] designated by $i_1(t)$, the functions $\bar{i}(t)$ and $i_1(t)$ have the appearance shown in Figure 129.1. For Equation [129.1], with $L = 0$, the current starts from the point A and decays exponentially thereafter. For the quasi-degenerate equation [129.2], with L very small, the current starts from zero, increases very rapidly, and follows the curve $i_1(t)$ which becomes practically identical with the curve $\bar{i}(t)$ after a very short time; this time is shorter as L is smaller. According to the absolutely degenerate equation [129.1], with $L = 0$, the current $\bar{i}(t)$ undergoes a true mathematical discontinuity OA at $t = 0$, whereas the current $i_1(t)$ of the quasi-degenerate equation [129.2], with $L \approx 0$, has a quasi-discontinuity OA′ which approaches OA when L approaches zero.

Although these facts are well known if considered independently of the previous history of the system, as in the above discussion, they appear in a somewhat different form in a quasi-discontinuous stationary relaxation oscillation. Thus, for example, the oscillation depicted in Figure 127.4 is governed by the full differential equation of the second order on the analytic branches AB, A′B′, ⋯ of the trajectories; these branches are determined in terms of the *two* constants of integration corresponding to the initial conditions represented by the points A and A′ of the phase plane. During the jumps BA′, B′A″, ⋯ the transition is governed by two differential equations of the first order. Since the solutions of each of these equations are determined by one constant of integration, there appears a relation between the variables $\frac{di}{dt}$ and i_c of the phase plane which does not exist on the analytic branches AB, A′B′, ⋯ . As a result of this, the jumps must occur along a certain direction in the phase plane.

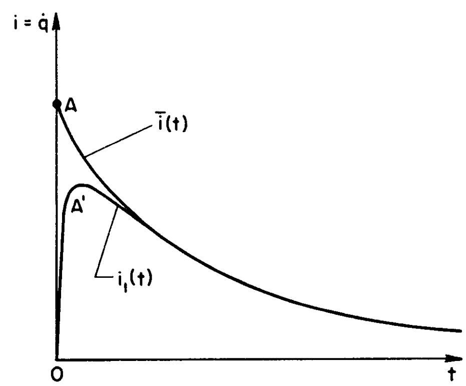


Figure 129.1

Thus the assumption of certain idealizations specified in connection with Figure 127.3 not only permits a discontinuous treatment of the problem but also indicates the direction of the jumps in the phase plane.

It is impossible, however, to proceed beyond this point if one attempts to apply these idealizations to the problem of discontinuous stationary relaxation oscillations. What is lacking in the system of Figure 127.4 is the mechanism by which the phenomenon of "closing" or "opening" the switch is produced *spontaneously by the internal reactions* of the circuit itself.

In order to be able to formulate this condition and thus to complete the discontinuous theory it will be necessary to introduce an *a priori* proposition whose validity is justified only by its agreement with the observed facts. This emphasizes once more the physical nature of this theory as distinguished from the contents of Parts I and II in which the argument was purely analytical.

130. CRITICAL POINTS OF DIFFERENTIAL EQUATIONS; BASIC ASSUMPTION

It was shown that, as a result of certain idealizations, discontinuities appear in the mathematical treatment of physical phenomena which exhibit rapid changes at certain points of their cycles. The use of discontinuities is convenient in some respects but inevitably introduces certain complications. These facts are too well known from the theory of mechanical impacts to need any emphasis here.

It is obvious that similar difficulties are to be expected if a discontinuous treatment of relaxation oscillations is adopted. For example, in an electronic "sweep circuit" some variables change so rapidly at certain points in the cycle that it is natural to attempt to idealize these changes by mathematical discontinuities. Obviously, any attempt to explain these changes on the basis of some kind of impact is difficult because the energy delivered by the external source, a battery, remains constant, and one cannot very well correlate the apparent continuity of the energy input into the system with the quasi-discontinuous changes of some of its variables. Very frequently a slight change of a parameter causes a disappearance of the phenomenon,and vice versa. In some particularly simple circuits in which the effect is known to exist, one succeeds in "explaining" it by a more or less elementary physical argument. In more complicated circuits it is impossible to give an account of what actually happens and, still less, to predict theoretically the existence, or non-existence, of such effects. There exists no analytical theory of these oscillations which would permit a treatment of these phenomena on a uniform basis as was possible for the quasi-linear oscillations with which we were concerned in Parts I, II, and III.

In order to be able to find a solution and to correlate the numerous experimental phenomena on a common basis, it becomes necessary to define terms and to introduce some kind of basic assumption, the value of which is to be justified by its agreement with the observed facts.

Definition: Critical points are the points at which the differential equation describing a phenomenon in a certain domain ceases to describe it.

Basic Assumption: Whenever the representative point following a trajectory of the differential equations describing a phenomenon reaches a critical point, a discontinuity occurs in some variable of the system.

Since, by virtue of the basic assumption, the occurrence of discontinuities depends on the existence of critical points, it is necessary to specify certain criteria by which their existence can be ascertained. In what follows we will encounter two principal criteria.

1. Let us consider a system of differential equations of the form

$$\frac{dx}{dt} = \frac{P(x,y)}{R(x,y)}; \qquad \frac{dy}{dt} = \frac{Q(x,y)}{R(x,y)} \qquad [130.1]$$

Obviously these differential equations become meaningless and, hence, cease to describe a physical phenomenon for the points (x_i, y_i) for which $R(x_i, y_i) = 0$. This equation represents a certain locus of critical points, and by virtue of the basic assumption a discontinuity occurs each time the representative point reaches the curve $R(x_i, y_i) = 0$. It is important to note that, as far as the *trajectory* is concerned, the passage through a critical point does not in any way affect its determinateness since R cancels out in the expression $\frac{dy}{dx} = \frac{Q}{P}$. It is impossible, however, to determine the *motion* on the trajectory in the neighborhood of a critical point. In this respect the local properties of a critical point are opposite to those of a singular point where the trajectory is indeterminate but the motion is determinate.

2. The existence of critical points or of a locus of such points can sometimes be revealed from the study of trajectories in a certain domain of the phase plane. A typical example in which this can be done is shown in Figure 130.1. The trajectories arrive at, or depart from, a certain threshold L from both sides, as shown. *If no singular points, that is, postions of equilibrium, exist in a narrow domain surrounding* L, *one can assert that the line* L *is a locus of critical points.*

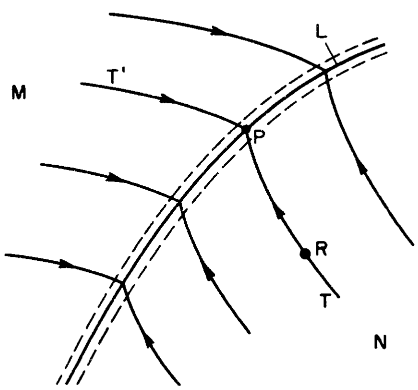


Figure 130.1

It is apparent that the trajectories situated in the regions M and N belong to two different differential equations. Let us assume that the phenomenon is represented by the motion of the representative point R on a trajectory T of the region N. Since the singular points are absent by our assumption, R will reach a point P on L in a finite time. Having reached this point, the representative point finds itself in a kind of *analytical impasse* from which there is no normal issue, that is, along the integral curves. In fact R cannot pass onto the trajectory T′ passing through P nor can it turn back on T since, in both cases, this would be inconsistent with the differential equations prescribing a definite direction on the trajectories of the two regions M and N. Nor can the representative point remain at the point P which is not a position of equilibrium. The differential equations cease to have any meaning at the point P and therefore cease to represent the physical phenomenon. Hence the point P is a critical point, and the line L is a locus of such points. By our basic assumption the discontinuities necessarily occur once the representative point has reached some point on L.

We are going to use the basic assumption extensively in the investigation of relaxation oscillations in relatively complicated circuits in which it is impossible to predict the nature of the phenomenon on a basis of elementary intuitive reasoning. It is useful to illustrate the application of the basic assumption to the simple example given previously, Figures 127.3 and 127.4. It is apparent that at the instant when the switch is closed, or opened, the differential equation of the second order ceases to describe the phenomenon since the right and the left portions of the circuit, instead of being in series, become in parallel. Hence this instant corresponds to a critical point, and a discontinuity is to be expected as we have ascertained by the elementary argument. The same argument applies to the ball striking the walls. In these two examples the oscillations are of the impulse-excited type, and the application of the basic assumption does not yield anything of interest.

We shall see that in connection with relaxation oscillations proper the basic assumption will be a useful tool by which the possibility of relaxation oscillations can be ascertained.

131. CONDITIONS OF MANDELSTAM

At the end of Section 127 certain idealizations and a choice of variables were specified so as to be able to introduce a discontinuous treatment of certain problems. It was shown that the necessary condition for such a treatment is the degeneracy of differential equations from the second to the first order if the variables $\frac{di}{dt}$ and $\frac{dV}{dt}$ are selected. In the preceding section we have formulated a sufficient condition for the occurrence of a discontinuity on the basis of a certain basic assumption.

There still remains one question to be settled, namely, the determination of the discontinuities once we have ascertained by the assumption that the discontinuity has to occur. Using the terminology of the phase plane, we can specify this last part of the problem as follows. Let us assume that the representative point R has reached a critical point $A(x_1, y_1)$. We may question into which other point $B(x_2, y_2)$ the representative point will jump from the point A. In discussing the solutions [127.3] of Equations [127.2] we have already touched this subject and found that in the very special case considered there the jump is from A(0,0) to $B\left(\frac{E}{r}, \frac{E}{L}\right)$.

L. Mandelstam formulated the conditions of a jump on the basis of certain plausible assumptions regarding the continuity of energy during the infinitely short time interval of the discontinuity. It is to be noted that these conditions of Mandelstam are useful for relaxation oscillations and not for impulse-excited oscillations for reasons which will appear later. The argument of Mandelstam is based on the continuity of the functions $i(t)$, the current through an inductance L, and $V(t)$, the voltage across the capacitor, as was previously mentioned in connection with the expressions [127.3] representing solutions of the degenerate equations of the first order. Since

$i(t)$ and $V(t)$ are continuous, clearly the electromagnetic energy $\frac{Li^2}{2}$ stored in an inductance and the electrostatic energy stored in a capacitor are also continuous functions of time. One obtains the conditions of Mandelstam by writing

$$\Delta i\Big|_{t_o - 0}^{t_o + 0} = 0; \qquad \Delta V\Big|_{t_o - 0}^{t_o + 0} = 0 \qquad [131.1]$$

where $(t_0 - 0,\ t_0 + 0)$ is the infinitely small time interval during which the discontinuity occurs. The important point to be noted in connection with these conditions is that they are applicable to an infinitely small time interval and to circuits with finite dissipative parameters. The first restriction is trivial and is nothing but the expression of a continuity of functions $i(t)$ and $V(t)$. As to the second, it requires an additional remark. One could formulate the following case in which the conditions of Mandelstam apparently do not hold. Let us assume that a charged capacitor is suddenly short-circuited so that its energy $\frac{CV^2}{2}$ disappears instantly; this seems to contradict the second condition [131.1] of Mandelstam. The fallacy of this reasoning lies in the fact that the only way in which the energy can disappear suddenly is to be totally converted into heat. But in order that this may occur, a finite dissipative parameter must be present. If, however, such a dissipative parameter exists in the circuit, there exists also a finite time constant so that the disappearance of the charge, and, hence, of the energy, cannot be instantaneous, and it is sufficient to define a small time interval consistent with the time constant of the circuit to ensure the validity of the conditions of Mandelstam.

The usefulness of the conditions of Mandelstam is limited only to relaxation oscillations proper. In fact, in impulse-excited oscillations the idealization employed is of an entirely different kind, and it is assumed that the energy exchanges between the system and an external source occur instantaneously. Summing up the result of this and of the preceding sections, it can be stated that the basic assumption and the conditions of Mandelstam are useful in studies of relaxation oscillations but are unnecessary for impulse-excited oscillations.

132. REMARKS CONCERNING SYSTEMS OF DEGENERATE DIFFERENTIAL EQUATIONS

In Section 3 it was mentioned that a differential equation of the second order can always be represented by a system of two differential equations of the first order if a new variable $y = \frac{dx}{dt}$ is introduced. Likewise, a differential equation of the n^{th} order can be reduced to a system of n differential equations of the first order by introducing the variables $\frac{dx}{dt} = y_1$, $\frac{dy_1}{dt} = y_2, \cdots$.

With degenerate differential equations the situation is somewhat different. Thus a completely degenerate equation of the second order is, in fact, a differential equation of the first order. As a result of this, the phase space, instead of being two-dimensional, that is, a phase plane, becomes uni-dimensional, that is, a phase line. Moreover, instead of analytic trajectories, piecewise analytic trajectories become possible.

A system of two differential equations of the second order can generally be reduced to a system of four differential equations of the first order, which means a system of the fourth order. If, however, each of the original differential equations of the second order degenerates into one equation of the first order, the system of the fourth order reduces to one equation of the second order, and its solutions can be represented by trajectories in a phase plane. This resultant equation of the second order, however, represents the result of degeneration of the system of the fourth order. We can express this by saying that we have a *doubly degenerate* system. Since each of the two differential equations of the first order admits discontinuous solutions, the doubly degenerate system of the second order will also possess certain discontinuous stretches in its phase plane so that its trajectories, in general, will be composed of certain analytic arcs joined by these stretches. Under certain conditions a doubly degenerate system of the second order may degenerate into a single differential equation of the first order; we can call such a case a *triply degenerate* system. We shall encounter one such system in what follows. In a triply degenerate system one differential equation of the first order represents the result of the degeneration of the system of the fourth order.

Although the use of degenerate equations extends the application of topological methods to a series of important practical cases which could not otherwise be represented in a phase plane, it should be noted that piecewise analytic trajectories of this kind "closed" by discontinuous stretches do not exhibit the features common to the regular analytic trajectories studied in Part I. We shall encounter, for example, systems in which no singularities exist inside such "closed" trajectories; in some other systems such trajectories include singularities alternating in the course of time between a focal, or nodal, point and a saddle point. Moreover, certain difficulties arise in the formulation of conditions for stability of such degenerate systems, as will be specified in Chapter XXV. These peculiarities are, of course, to be expected. They reflect to some extent the fact that the discontinuous theory of relaxation oscillations is based on certain idealizations; this fact makes it difficult to compare it directly with classical methods making use of analytic functions. The principal usefulness of this theory at present is that it permits obtaining satisfactory qualitative information in a great majority of important practical problems whose solutions are beyond the reach of existing analytical methods.

CHAPTER XXI

DEGENERATE SYSTEMS WITH ONE DEGREE OF FREEDOM

133. PERIODIC SOLUTIONS OF DEGENERATE SYSTEMS OF THE FIRST ORDER

A differential equation of the first order

$$\frac{dx}{dt} = f(x) \qquad [133.1]$$

obviously does not possess continuous analytic periodic solutions. Moreover, one can assert that if the function $f(x)$ is single-valued, no continuous, although not necessarily analytic, periodic solutions are possible. In fact, in order that *some* sort of periodicity may exist, it is necessary that the system traverse the same line $x = x_1$ with two oppositely directed velocities; this, however, is impossible if $f(x)$ is single-valued. In the example illustrated by Figure 127.2b, we saw that the "closed" periodic trajectory ABCDA is characterized by the fact that the function $f(x) = \pm v_0$ is actually double-valued. It so happens that, in the example referred to, the function $f(x)$ is a constant. It is easy, however, to waive this restriction by assuming that the plane on which the ball rolls, instead of being horizontal, rises toward the right wall. Then the trajectory will be a trapezoid ABCDA, as shown in Figure 133.1. One could imagine still other cases by assuming that, instead of rolling on a plane, the ball rolls on some kind of cylindrical surface whose generating lines are parallel to the $\dot{x}$-axis. In this case the trajectory would be formed by stretches AB and CD of analytic arcs "closed" by the discontinuous stretches BC and DA.

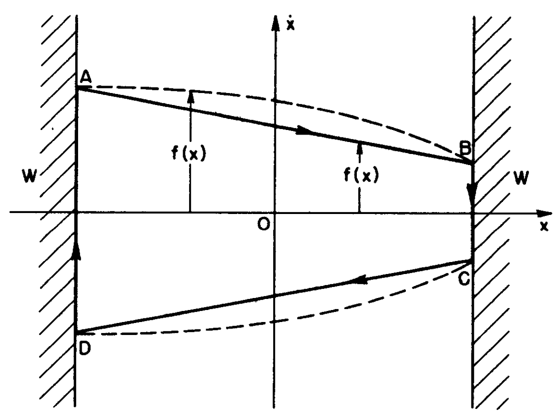


Figure 133.1

The change from one branch of the function $f(x)$ to the other one generally occurs at critical points and is discontinuous. Very frequently this is equivalent to saying that the phenomenon is governed by two distinct differential equations during its cycle. During one fraction of the cycle the phenomenon is described by one differential equation and during the other fraction by the other equation. The change from one differential equation to the other occurs at the critical points. In the example illustrated by Figure 133.1 the function $f(x)$ happens to have branches which are symmetrical with respect to the x-axis. This is not always so, as will be shown later. We will now illustrate this matter by the following two well-known examples.

134. RELAXATION OSCILLATIONS IN A CIRCUIT CONTAINING A GASEOUS CONDUCTOR

In view of the fact that this subject has been explored, we shall omit the familiar details and will endeavor primarily to show the application of the discontinuous theory in order to prepare the groundwork for more complicated cases beyond the reach of elementary theory.

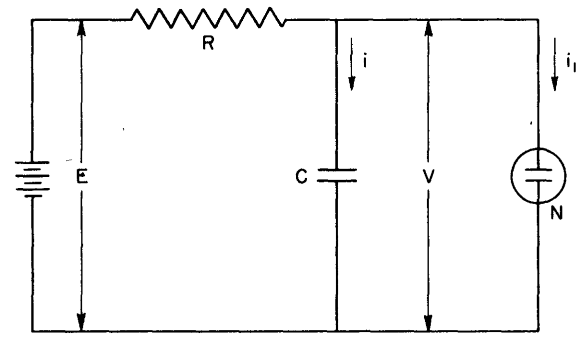


Figure 134.1

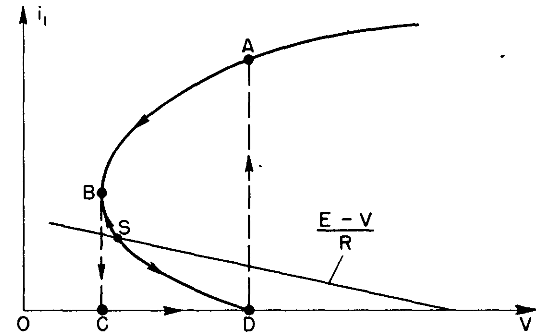


Figure 134.2

Figure 134.1 shows a circuit with the usual notations and with the positive directions as indicated; N is a gaseous conductor, such as a neon tube.

Figure 134.2 represents the characteristic* ABD of the gaseous conductor N, which can be represented by a non-linear empirical relation $i_1 = \phi(V)$.

The differential equations of the circuit obviously are

$$R(i + i_1) + V = E; \qquad i = C\frac{dV}{dt} \qquad [134.1]$$

These equations reduce to the following differential equation of the first order

$$\frac{dV}{dt} = \frac{1}{RC}\left[E - V - R\phi(V)\right] \qquad [134.2]$$

This equation is valid only when the discharge exists; during its extinction $i_1 = \phi(V) = 0$, and we have

$$\frac{dV}{dt} = \frac{1}{RC}(E - V) \qquad [134.3]$$

Since we know that the phenomenon is characterized by alternate striking and extinction of the discharge, it is apparent that it is represented

---

* In order to avoid any misunderstanding, we shall use the term *characteristic* in the engineering sense, that is, to designate a certain experimental curve connecting the values of certain physical quantities, such as current i_1 through the non-linear conductor N and voltage V across it. We will reserve the term *trajectory* to mean an *integral curve* of a differential equation, as we have done previously.

alternately by two distinct differential equations, [134.2] and [134.3]. Since these differential equations are of the first order, only one constant of integration is involved in their solutions. But as there are two dynamic variables $i_1 = \phi(V)$ and V, it is apparent that there must exist a definite relation between these variables, as has been mentioned in connection with Equation [128.3]. During the time intervals when the discharge exists, this relation is obviously the characteristic ABD.

Since the phase space is uni-dimensional for this system, it is apparent that during the intervals when the discharge exists the characteristic appears as the *phase line* and during the intervals of extinction the phase line is the V-axis. Aside from these phase lines, the plane is not involved.

As is well known from elementary theory, the points of equilibrium are given by the points of intersection of the characteristic and the straight line $i = \frac{E - V}{R}$. On the upper branch AB of the characteristic the equilibrium is stable; on the lower branch BD it is unstable, as is easy to ascertain by an elementary procedure. The case when the straight line $i = \frac{E - V}{R}$ cuts the upper branch of the characteristic is obviously of no interest from the standpoint of oscillations. We shall confine our attention, therefore, to the case when the resistance R has been adjusted to a value at which the straight line and the characteristic intersect at some point S situated on the lower unstable branch BD.

Assume that we start the investigation of the phenomenon at the instant when the discharge has just appeared; this instant is represented by the point A in Figure 134.2. Since $\frac{dV}{dt} < 0$ on the upper branch in this case, the representative point will move from A to B on the upper branch. The arrows on the characteristic, which, as was just explained, is also the phase line, indicate the positive directions consistent with the differential equation. It follows, therefore, that having reached the point B, the representative point finds itself in a situation which was specified in connection with Figure 130.1, so that the point B is a critical point of the differential equation [134.2], and by our basic assumption we can assert that a jump must occur at this point. In order to determine the character of the jump we have to apply the condition of Mandelstam. The only form of stored energy is the electrostatic energy $\frac{CV^2}{2}$ stored in the capacitor. Thus the condition of the jump is

$$\Delta V \Big|_{t_0 - 0}^{t_0 + 0} = 0$$

which means that "during" the jump the voltage V across the capacitor remains constant, that is, the jump occurs parallel to the i_1-axis along the stretch BC. The process of extinction is thus "explained" on the basis of the discontinuous theory.

Beginning with the point C, Equation [134.3] describes the process of charging the capacitor with no discharge. During that time interval the representative point moves along the V-axis with a finite velocity until the point D is reached. Here the discharge strikes again, and the representative point is transferred discontinuously to the point A, after which the cycle is repeated. Since at the point D Equation [134.3] ceases to represent the phenomenon, we conclude that D is also a critical point.

It is thus seen that the familiar phenomenon of relaxation oscillations of a circuit containing a gaseous conductor can be treated consistently on the basis of the discontinuous theory and can be represented by a piecewise analytic cycle ABCDA formed by two analytic branches AB and CD "closed" by the discontinuous stretches BC and DA. We note also certain differences between the discontinuous oscillations just studied and the continuous oscillations treated by the classical theory in Part I. The closed trajectory ABCDA is not of a limit-cycle type in that the phenomenon jumps, so to say, directly into its cycle without a gradual asymptotic spiraling around it. This cycle resembles in all respects a similar cycle shown in Figures 127.2b and 133.1 describing the idealized behavior of a ball undergoing reflections from constraining walls.

So far, study of such oscillations does not reveal anything new, and the main purpose of this discussion is to illustrate the application of the discontinuous theory of relaxation oscillations to a familiar phenomenon generally treated by elementary methods. As we proceed further, the usefulness of this theory will become more manifest. In the example treated in the following section, for instance, it would be relatively difficult to apply the semi-intuitive physical argument. In still further examples it would be altogether impossible.

135. RC-MULTIVIBRATOR

As a second example of a system which can be described by one differential equation of the first order, we shall investigate the so-called RC-multivibrator circuit shown in Figure 135.1. The electron tube V_2 is a nonlinear conductor characterized by the equation $I_a = \phi(e_g)$. The tube V_1 appears here merely as a linear amplifier amplifying the potential difference ri, the feed-back voltage, between the points B and D and applying the amplified voltage $e_g = kri$ to the grid of V_2, k being the amplification factor of V_1. The circuit is idealized in the customary manner, that is, the effects of the parasitic inductance, the grid current, and the anode reaction are neglected.

The differential equations of the circuit with the positive directions shown are

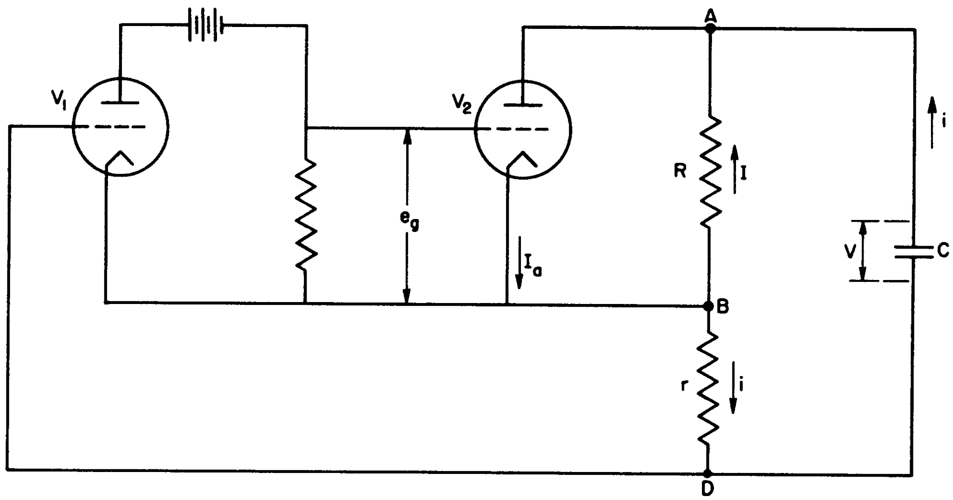


Figure 135.1

$$(R + r)i + V = R\phi(kri); \qquad i = C\frac{dV}{dt} \qquad [135.1]$$

These equations reduce to the equation

$$\left[krR\phi'(kri) - (R + r)\right]\frac{di}{dt} = \frac{1}{C}i \qquad [135.2]$$

The critical point $i = i_1$ is given by the equation

$$f(i_1) = \left[krR\phi'(kri_1) - (R + r)\right] = 0 \qquad [135.3]$$

where $\phi'(kri_1)$ designates the slope of the characteristic at the point $i = i_1$. One can also apply the argument given in connection with Figure 130.1. For this purpose the graphical procedure shown in Figures 135.2a and b will be useful. Figure 135.2a shows the characteristic $I_a = \phi(kri)$ of V_2 multiplied by a constant factor R. The tube V_2 is supposed to be biased at a point O in the middle of the rectilinear part of its characteristic. From the usual form of the characteristic $\phi(kri)$ it is apparent that its slope $\frac{d\phi}{di} = \phi'$ is maximum when $i = 0$ and approaches zero monotonically as $|i| \to \infty$. Let $\phi'(0) = S$ and assume that $RS > R + r$, without which, obviously, no self-excitation is possible. The curve $V(i)$ represents the function $R\phi(kri) - (R + r)i$. The curve shown in Figure 135.2b is the slope curve of the function $R\phi(kri)$ multiplied by a constant factor kr and referred to the axis M′N′. It is apparent that if this curve is referred to the MN-axis parallel to M′N′ and at a distance $(R + r)$ from the origin O, the ordinates $f(i)$ of this curve represent the left-hand term of the expression [135.3]. Hence, the points P and Q are critical points, and by transferring these points on the diagram of Figure 135.2a one obtains the critical points B and D situated on the $V(i)$-curve and symmetrical with respect to the origin under the assumed idealization of the

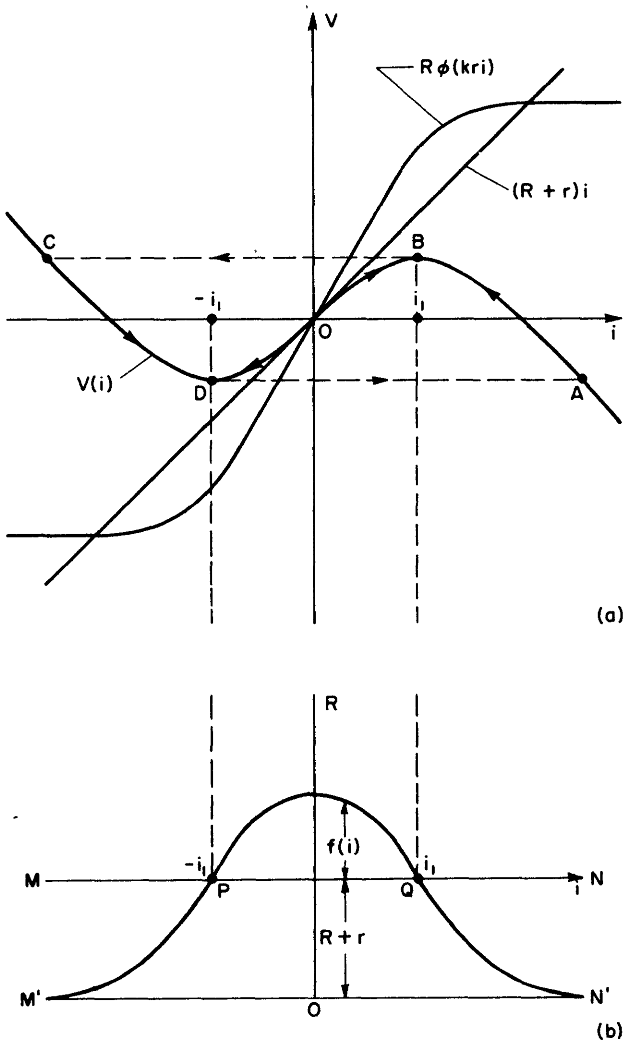


Figure 135.2

characteristic. It is apparent also that the inner interval $(-i_1 < i < +i_1)$ is unstable since $f(i) > 0$ in that interval. On the contrary, the outer interval $(i > +i_1;\ i < -i_1)$ is stable since $f(i) < 0$ in that interval, as is seen at once from Equation [135.2]. The positive directions on the phase line, the $V(i)$ curve, are shown in Figure 135.2a. The argument specified in Section 130 is therefore applicable. Thus, for example, when the representative point has reached the point B on the stretch AB, it has to jump discontinuously, and the second condition of Mandelstam specifies that the jump must occur parallel to the i-axis, since in such a case the energy stored in the capacitor does not change during the jump. The discontinuous stretch BC so traversed ends at the point C. Here the phenomenon is again governed by the differential equation, and the analytic stretch CD is traversed with a finite velocity. At the point D another jump occurs along the stretch DA, so that a closed piecewise analytic cycle ABCDA results. The phenomenon therefore follows a pattern substantially the same as that which has previously been investigated in connection with the ball striking the walls and also with the neon-tube oscillator. The difference between the neon-tube circuit and the multivibrator circuit is that for the former the phenomenon is described alternately by two separate differential equations [134.2] and [134.3], whereas for the latter the phenomenon is governed by only one differential equation [135.2]. Because of the symmetry of the curve $V(i)$ and that of the critical points, however, the closed cycle is obtained with the origin as the center of symmetry.

136. SYSTEM WITH ONE DEGREE OF FREEDOM DESCRIBABLE BY TWO DIFFERENTIAL EQUATIONS OF THE FIRST ORDER

The systems with one degree of freedom so far considered could be expressed by one differential equation of the first order. Chaikin and Lochakow (14) have investigated an interesting case when a dynamical system with one degree of freedom is describable by *two* differential equations of the second order possessing critical points. We shall outline briefly the principal features of the relaxation oscillations appearing in this case.

The fact that the system of differential equations is now of the second order will account for the representation of the phenomenon in the phase plane instead of its uni-dimensional representation by a phase line, as in previous cases. It will be shown that the discontinuous theory permits establishing the principal features of the phenomenon, although a direct intuitive argument, such as is applicable in the simple systems of Sections 133, 134, and 135, is insufficient here.

The circuit investigated by Chaikin and Lochakow is shown in Figure 136.1. It is observed that the only difference between this circuit and that shown in Figure 135.1 is that here the inductance L replaces the resistance R of the circuit of Figure 135.1. It will be shown now that the behavior of this circuit exhibits features entirely different from those of the circuit of Figure 135.1.

Using the previous notations, we obtain, by Kirchhoff's laws,

$$I_a = \phi(kri) = I + i; \qquad L\frac{dI}{dt} - ri - \frac{1}{C}\int i\,dt = 0 \qquad [136.1]$$

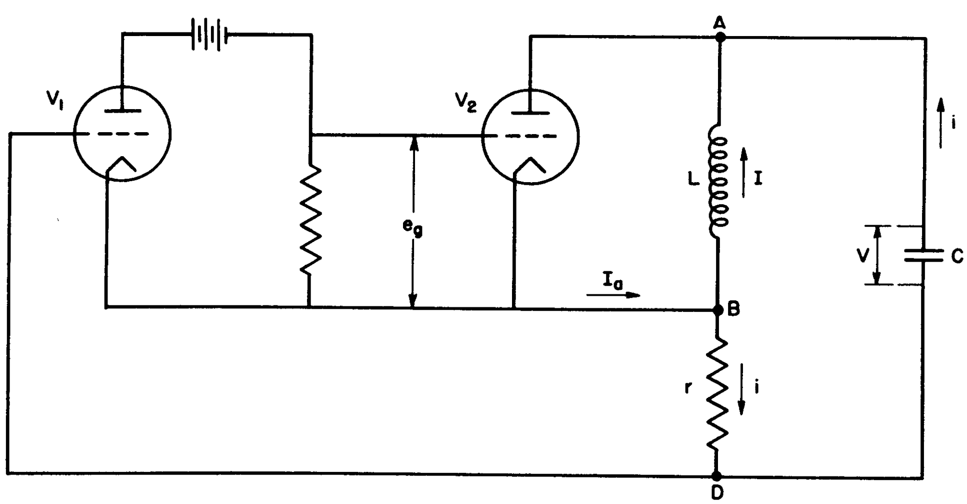


Figure 136.1

Introducing the variables $x = kri$ and $y = \frac{dI}{dt}$ and differentiating these equations, one gets

$$\frac{dx}{dt} = \frac{y}{\psi(x)}; \quad \frac{dy}{dt} = \frac{x}{krLC} + \frac{1}{kL}\frac{y}{\psi(x)} \qquad [136.2]$$

where

$$\psi(x) = \phi'(x) - \frac{1}{kr} \qquad [136.3]$$

The phase trajectories are given by the equation

$$\frac{dy}{dx} = \frac{x\psi(x)}{krLCy} + \frac{1}{kL} \qquad [136.4]$$

The point $x = y = 0$ is clearly a singular point. It is apparent from the form of the characteristic of electron tubes that the function $\psi(x)$ decreases monotonically from a positive value $\psi(0) = S - \frac{1}{kr}$ to a negative value $\psi(x) = -\frac{1}{kr}$ when $|x| \to \infty$. Hence there exist two roots $x = \pm x_1$, for which both $\frac{dx}{dt}$ and $\frac{dy}{dt}$ become infinite. According to our definition, these roots $x = \pm x_1$ are the *critical points* of the system [136.2]. In the phase plane the values $x = \pm x_1$ determine two *critical thresholds*; that is, they are loci of the critical points of this system. It is to be noted that at these thresholds the tangent to the phase trajectories

$$\left(\frac{dy}{dx}\right)_{x = \pm x_1} = \frac{1}{kL} \qquad [136.5]$$

is determinate. It is impossible, however, to determine the *motion* of the representative point since the differential equations become meaningless at these points.

In order to show that these thresholds $x = \pm x_1$ separate the regions of the phase plane in which the topological structure of trajectories is radically different, we shall simplify the problem slightly without, however, introducing any qualitative changes. Figure 136.2 represents the characteristic of the electron tube V_2 shown in Figure 136.1. We shall exclude the critical thresholds $x = \pm x_1$ by drawing parallels to the y-axis on both sides of each threshold so as to obtain narrow strips of width Δx. The function $\psi(x)$ will then be $S(x) - \frac{1}{kr} > 0$ and $-\frac{1}{kr}$ in the inner and outer intervals, respectively. This will cause no qualitative changes but will change slightly the form of the

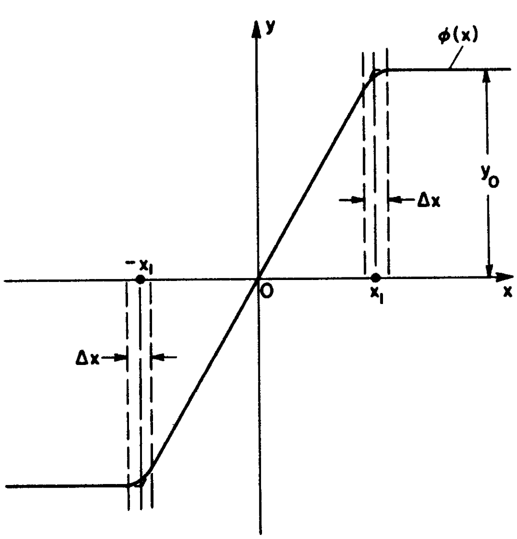


Figure 136.2

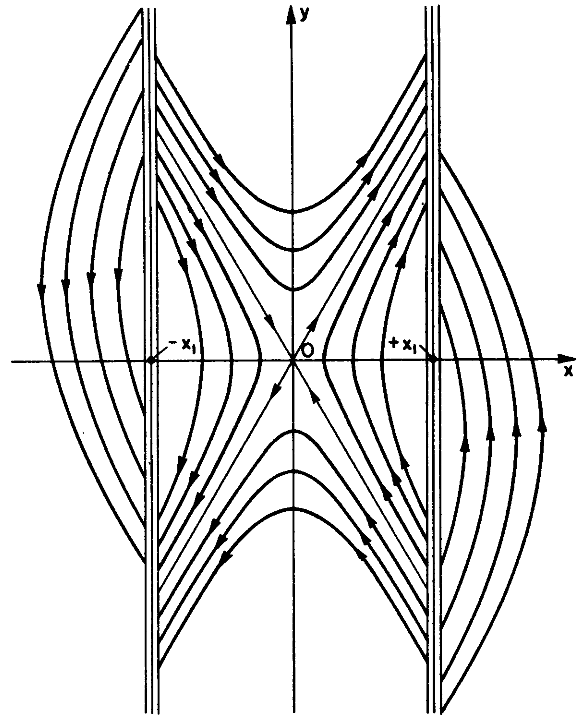


Figure 136.3

trajectories in the neighborhood of the strips; see Section 9, Part I. Since the equations then become linear, the standard procedure, see Section 18, Part I, shows that the origin O appears as a saddle point for the trajectories of the inner interval ($-x_1 < x < +x_1$) and as either a focal point (if $r < 2\sqrt{\frac{L}{C}}$) or a nodal point (if $r > 2\sqrt{\frac{L}{C}}$) in the outer interval ($x < -x_1$; $x > +x_1$). We shall assume the existence of a focal point in the outer interval since this is commonly encountered in applications.

The picture of trajectories in this case is shown in Figure 136.3. A continuum of hyperbolic trajectories corresponding to the saddle point fills the inner interval, and a continuum of spirals corresponding to the focal point appears in the outer interval. By an elementary discussion of the sign of $\frac{dy}{dx}$ in Equation [136.4] for different quadrants of both intervals, it can be ascertained that the positive directions on the trajectories are oriented as shown by the arrows. It is apparent that these thresholds $x = \pm x_1$ are loci of critical points, and discontinuities occur whenever the representative point following a trajectory reaches one of the lines $x = \pm x_1$. Since the differential equations become meaningless at the critical points, the determination of discontinuities is made by the conditions [131.1] of Mandelstam, which, in this case, are

$$\Delta I \Big|_{t_0 - 0}^{t_0 + 0} = \int_{t_0 - 0}^{t_0 + 0} y\,dt = 0; \qquad \Delta V \Big|_{t_0' - 0}^{t_0 + 0} = \frac{1}{krC}\int_{t_0 - 0}^{t_0 + 0} x\,dt = 0 \quad [136.6]$$

Applying these conditions to Equations [136.1], one obtains

$$\phi(x_1) - \frac{x_1}{kr} = \phi(x_2) - \frac{x_2}{kr}; \qquad y_1 - y_2 = \frac{x_1 - x_2}{kL} \qquad [136.7]$$

where x_2 and y_2 are the coordinates of the representative point immediately after the discontinuity. Since the function $\phi(x)$ is an empirical curve, it is preferable to plot first the function $\theta(x) = \phi(x) - \frac{x}{kr}$ as was done in Figure 135.2a. One obtains for $\theta(x)$ a curve similar to the curve $V(i)$ of Figure

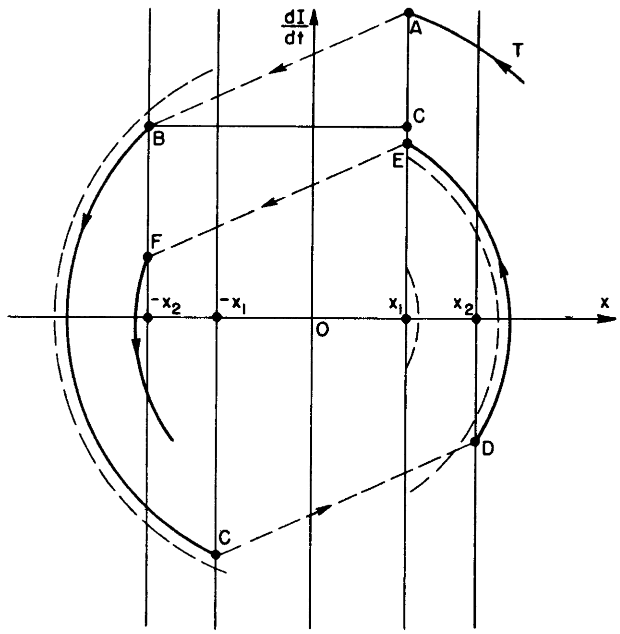


Figure 136.4

135.2a. By expressing the condition $\theta(x_1) = \theta(x_2)$, one can find x_2 if x_1 is given. We omit this graphical construction and merely mention that for characteristics of electron tubes commonly encountered the jump occurs from $+x_1$ to $-x_2$ (with $|x_2| > |x_1|$) and from $-x_1$ to $+x_2$. When x_2 is known, the second equation [136.7] permits determining y_2 for given x_1 and y_1.

The representation of such a quasi-discontinuous oscillation in the phase plane of the variables (x,y) is shown in Figure 136.4. It is noted that the jump from (x_1,y_1) to (x_2,y_2) is the same for all trajectories. Assume that the representative point, following a trajectory T of the outer interval, has reached a critical point A of coordinates (x_1,y_1). The second equation [136.7] then permits determining the "direction" of the jump AB in the phase plane. At the point B another spiral trajectory begins which encounters at C another critical point resulting in the jump CD, and so on. It can be shown, although we omit the proof, that after one turn of the radius vector the piecewise analytic spiral ABCDE will approach the origin, which means that the point E is below the point A if the latter is sufficiently distant from the origin O. If, however, one applies the same reasoning to a point A′ near the x-axis on the critical line $x = \pm x_1$, one finds, on the contrary, that E′ is above A′. Thus the large spirals shrink and the small ones grow with each turn of the radius vector. Hence there exists one and only one piecewise analytic spiral for which the points A and E coincide so that the trajectory becomes closed. Such a closed trajectory may be termed a *piecewise analytic limit cycle*, and it is stable. One notes the difference between such a trajectory and the piecewise analytic cycles described in Sections 133, 134, and 135, which are not of the *limit-cycle* type. The difference is due to the fact that the present phenomenon is expressible by *two* differential equations of the first order and is represented in the *phase plane*, whereas in the sections mentioned above it was determined by *one* differential equation of the first order and was represented by *phase lines* without involving any other points of the phase plane.

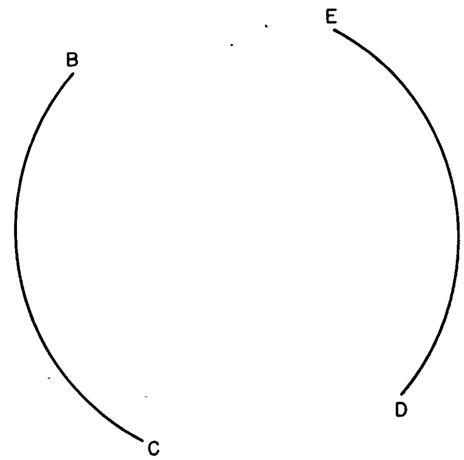


Figure 136.5

Another remarkable feature of this analysis is that the unstable hyperbolic trajectories of the inner interval do not appear at all in this representation because that interval is traversed discontinuously. Experiments made by Chàikin and Lochakow corroborate these theoretical conclusions. In their experiments a cathode-ray oscillograph was adapted to record the phase trajectories of the system, as was explained in connection with Figure 24.7, Part I. The record has the appearance shown in Figure 136.5. There are two spiral arcs BC and DE corresponding to a relatively slow motion of the electronic beam on the analytic trajectories of the outer interval. The quasi-discontinuous jumps EB and CD remain unrecorded because of the much higher speed of the beam in the region of the inner interval. It seems logical to assume that when the phenomenon starts from rest, *one* hyperbolic trajectory of the inner interval is actually traversed, but once the first critical point has been reached, the phenomenon begins to "skip" the inner interval and continues to do so thereafter.

Once the essential features of this phenomenon have been ascertained by this method, it is easy to give a corresponding physical interpretation. The fact that the jumps occur obliquely in the phase plane of the variables (x,y) means that at these instants the voltage $L\frac{dI}{dt}$ across the inductance and the current $i = C\frac{dV}{dt}$ in the capacitor undergo quasi-discontinuities, although the current I through the inductance and the voltage V across the capacitor remain continuous. In other words, the functions $I(t)$ and $V(t)$ are continuous but are not analytic at the jumps in the sense that they have discontinuous first derivatives. This is in agreement with the elementary theory of an impulsive excitation of the idealized (L,R)- and (R,C)-circuits. It is apparent, however, that mere knowledge of these well-known facts would be entirely insufficient for the purpose of establishing a qualitative picture of a complicated phenomenon of this nature if no general method, such as that offered by the discontinuous theory, were available.

CHAPTER XXII

MULTIPLY DEGENERATE SYSTEMS

137. MULTIVIBRATOR OF ABRAHAM-BLOCH*

As an example of a system with two degrees of freedom describable by two differential equations of the first order we shall investigate the behavior of the circuit shown in Figure 137.1. Neglecting the effect of grid current and anode reaction and using the notations and positive directions shown, we obtain the following equations:

$$I_1 = I_{a1} + i_1; \qquad I_2 = I_{a2} + i_2$$

$$RI_1 + \frac{1}{C}\int i_1 dt + ri_1 = E; \qquad RI_2 + \frac{1}{C}\int i_2 dt + ri_2 = E \qquad [137.1]$$

$$I_{a1} = \phi(e_{g1}) = \phi(ri_2); \qquad I_{a2} = \phi(e_{g2}) = \phi(ri_1)$$

where $I_a = \phi(e_g)$ is the non-linear characteristic of the electron tubes V_1 and V_2. Differentiating the second group of equations [137.1] and making use of the other two groups, we obtain the following system of differential equations:

$$(R + r)\frac{di_1}{dt} + \frac{1}{C}i_1 + Rr\phi'(ri_2)\frac{di_2}{dt} = 0$$

$$Rr\phi'(ri_1)\frac{di_1}{dt} + (R + r)\frac{di_2}{dt} + \frac{1}{C}i_2 = 0 \qquad [137.2]$$

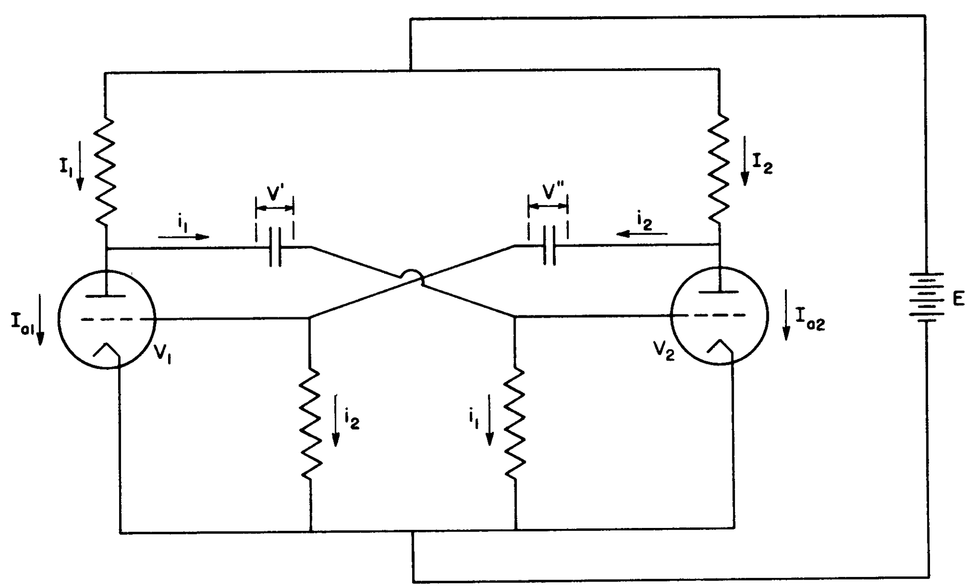


Figure 137.1

* This device is described in Reference (15).

This system reduces to the following equations:

$$\frac{di_1}{dt} = \frac{(R + r)\frac{i_1}{C} - Rr\phi'(ri_2)\frac{i_2}{C}}{R^2r^2\phi'(ri_1)\phi'(ri_2) - (R + r)^2}$$

$$\frac{di_2}{dt} = \frac{(R + r)\frac{i_2}{C} - Rr\phi'(ri_1)\frac{i_1}{C}}{R^2r^2\phi'(ri_1)\phi'(ri_2) - (R + r)^2} \qquad [137.3]$$

which are of the form

$$\frac{di_1}{dt} = \frac{P(i_1, i_2)}{U(i_1, i_2)}; \qquad \frac{di_2}{dt} = \frac{Q(i_1, i_2)}{U(i_1, i_2)} \qquad [137.4]$$

One notes the symmetry of Equations [137.3] with respect to the variables i_1 and i_2; this symmetry is due, of course, to the symmetry of the circuit.

The phase trajectories in the (i_1, i_2)-plane are given by the equation

$$\frac{di_2}{di_1} = \frac{Q(i_1, i_2)}{P(i_1, i_2)} \qquad [137.5]$$

From the explicit values of the functions P and Q, Equations [137.3], one observes that the origin $i_1 = i_2 = 0$ is a singular point, that is, an equilibrium point of the circuit.

We first inquire whether closed analytic trajectories are possible in the system. Applying the negative criterion of Bendixson, see Section 25, one ascertains that

$$\frac{dP}{di_1} + \frac{dQ}{di_2} = \frac{2(R + r)}{C} = \text{constant} \qquad [137.6]$$

Hence no closed analytic trajectories exist here. The nature of the singular point $i_1 = i_2 = 0$ is determined by the equations of the first approximation. If we let $\phi'(ri_1)_{i_1 = 0} = \phi'(ri_2)_{i_2 = 0} = S$ and if we assume that $RrS > R + r$, Equations [137.3] become

$$\frac{di_1}{dt} = \frac{R + r}{M} i_1 - \frac{RrS}{M} i_2; \qquad \frac{di_2}{dt} = -\frac{RrS}{M} i_1 + \frac{R + r}{M} i_2 \qquad [137.7]$$

where

$$M = C[R^2r^2S^2 - (R + r)^2] > 0$$

The characteristic equation of this system is

$$\lambda^2 - \frac{2(R + r)}{M}\lambda + \frac{[(R + r) + RrS][(R + r) - RrS]}{M^2} = 0 \qquad [137.8]$$

The origin $(i_1 = i_2 = 0)$ is therefore a saddle point. One ascertains easily that, when $RrS < R + r$, the origin is a stable nodal point, but this case is of no interest from the standpoint of relaxation oscillations. We will assume the existence of a saddle point. Since, initially, $RrS > R + r$ and the origin is unstable, the variables i_1 and i_2 begin to increase. On the other hand, from the form of the characteristic $I_a = \phi(ri)$, we know, see Section 135, that $|\phi'(ri)| \to 0$ when $i \to \infty$. The function

$$U(i_1, i_2) = C[R^2 r^2 \phi'(ri_1)\phi'(ri_2) - (R + r)^2] \qquad [137.9]$$

which is initially positive, decreases monotonically when i_1 and i_2 increase and is negative when i_1 and i_2 are very large and are equal. Hence there are certainly some values of i_1 and i_2 for which $U = 0$. This means that the system [137.3] has critical points so that by virtue of the basic assumption, Section 130, discontinuities must occur at these points. The expression $U(i_1, i_2) = 0$ defines a certain curve Γ_1 in the plane of the variables i_1 and i_2. In order to simplify the notation, let us designate by x_1 and y_1 the values of i_1 and i_2, respectively, situated on the curve Γ_1 shown in Figure 137.2. The equation of this curve is then

$$R^2 r^2 \phi'(rx_1)\phi'(ry_1) - (R + r)^2 = 0 \qquad [137.10]$$

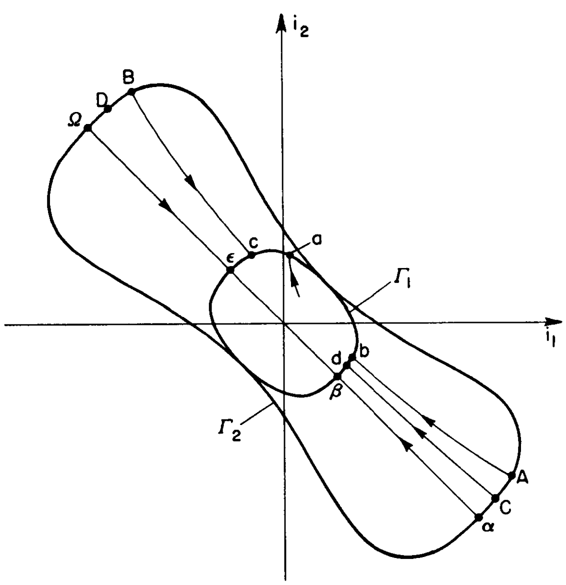


Figure 137.2

From the form of the functions $\phi'(rx_1)$ and $\phi'(ry_1)$, approximating the slope to the characteristic $I_a = \phi(e_g)$ of the electron tube, one ascertains that the curve Γ_1 is a symmetrical closed curve with the origin at its center and has the form shown in Figure 137.2. As soon as the representative point following a trajectory has reached a point (x_1, y_1) on the curve Γ_1, a jump must occur, and the point (x_2, y_2) into which the representative point jumps is determined by the conditions of Mandelstam which we now propose to apply. Since the only form of stored energy in this case is purely electrostatic, namely, the energy stored in capacitors, the conditions of Mandelstam require that the potential differences V' and V'' across the capacitors remain constant during the discontinuity. This gives the conditions

$$V' = E - R\phi(ri_2) - (R + r)i_1; \qquad V'' = E - R\phi(ri_1) - (R + r)i_2 \qquad [137.11]$$

If we designate now the terminal coordinates of the jump by (x_2, y_2), the conditions of Mandelstam obviously are

$$R\phi(ry_1) - (R + r)x_1 = R\phi(ry_2) - (R + r)x_2$$
$$R\phi(rx_1) - (R + r)y_1 = R\phi(rx_2) - (R + r)y_2 \qquad [137.12]$$

In these notations (x_1, y_1) represents the critical point on the curve Γ_1 from which the representative point is transferred discontinuously to some other point (x_2, y_2) determined by these equations. Since there is generally a one-to-one correspondence between the points (x_1, y_1) and (x_2, y_2), we conclude that the locus of the points (x_2, y_2) is another curve Γ_2 which is also a closed curve symmetrical with respect to both variables i_1 and i_2 since the circuit is entirely symmetrical.

The phase-plane representation of the behavior of the Abraham-Bloch multivibrator can then be described as follows. Assume that the representative point has reached a critical point a on the curve Γ_1. From this point it jumps into a *corresponding point* A on the curve Γ_2. At the point A begins a stretch of a continuous analytic trajectory Ab. The point b is another critical point where another jump bB occurs. At B a second continuous trajectory begins which ends at another critical point c, from which a new jump cC takes place; at C a continuous stretch Cd begins which ends at d, and so on.

By a more detailed analysis of such piecewise analytic trajectories interrupted by discontinuities, it can be shown that the transient behavior of the Abraham-Bloch multivibrator approaches a stationary symmetrical state, which is oriented along the bisector of the i_1- and i_2-axes and which consists of a continuous stretch $\alpha\beta$ followed by a jump $\beta\Omega$, then by another continuous trajectory $\Omega\epsilon$ followed by another jump $\epsilon\alpha$ which closes the "cycle." In this case, the cycle is uni-dimensional, that is, it is the phase line $\Omega\alpha$.

It is apparent that if, instead of assuming the asymmetry of the initial conditions as we did, only a steady-state condition was aimed at, a simplification could be made in the differential equations by introducing the conditions of symmetry, that is, $i_1 = -i_2$ and $\phi'(ri_1) = -\phi'(ri_2)$. With this simplification, one obtains a single differential equation of the first order, namely,

$$\frac{di}{dt} = \frac{(R + r) + Rr\phi'(ri)}{R^2r^2[\phi'(ri)]^2 - (R + r)^2}\ \frac{i}{C} \qquad [137.13]$$

which reduces to the form

$$[Rr\phi'(ri) - (R + r)]\frac{di}{dt} = \frac{i}{C} \qquad [137.14]$$

This is identical with Equation [135.2], the equation of an asymmetrical relaxation oscillation. We conclude, therefore, that the transient behavior of the Abraham-Bloch multivibrator is represented by a doubly degenerate system. As far as the stationary condition described by one differential equation

[137.14] of the first order is concerned, the multivibrator is represented by a triply degenerate system, to use the terminology of Section 132.

138. HEEGNER'S CIRCUIT; ANALYTIC TRAJECTORIES

As a second example of a doubly degenerate system, we shall investigate the so-called Heegner circuit (16) shown in Figure 138.1, which is the same as the circuit of an asymmetrical RC relaxation oscillation shown in Figure 135.1, the only difference being that there is now an additional capacitor C_2 shunting the resistor R. We propose to show now that the addition of the capacitor C_2 radically modifies the behavior of the circuit. It would be impossible to see this without the criteria of Section 130.

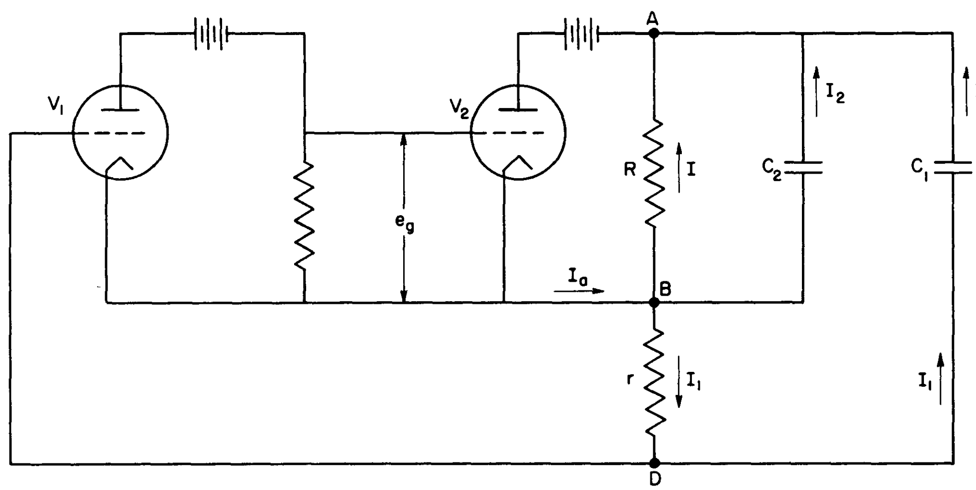


Figure 138.1

With the positive directions indicated by arrows, the application of Kirchhoff's laws gives

$$I_a = I + I_1 + I_2; \qquad I_1 = C_1\frac{d}{dt}(RI - rI_1); \qquad I_2 = C_2R\frac{dI}{dt} \quad [138.1]$$

Putting $I_a = \phi(krI_1)$, as we have previously, and eliminating I, we obtain the following system of the second order:

$$\frac{dI_1}{dt} = -\frac{1}{rC_1}I_1 + \frac{1}{rC_2}I_2$$
$$\frac{dI_2}{dt} = \frac{R - Rrk\phi'(krI_1)}{RrC_1}I_1 - \frac{R + r - Rrk\phi'(krI_1)}{RrC_2}I_2 \quad [138.2]$$

It is observed that this system of differential equations has no critical points, and hence no discontinuities are to be expected. The only singular point is $I_1 = I_2 = 0$. Noting that $\phi'(0) = S$ is a maximum and that the function $|\phi'(krI_1)| \rightarrow 0$ when $|I_1| \rightarrow \infty$, the characteristic equation of the system

[138.2] is

$$\lambda^2 + \left[\frac{C_2}{C_1} + \frac{R + r - RrkS}{R}\right]\frac{1}{rC_2}\lambda + \frac{1}{RrC_1C_2} = 0 \qquad [138.3]$$

It is noted that the singularity here is *not* a saddle point. Hence, it is either a nodal point, if the roots λ_1 and λ_2 are real, or a focal point, if they are conjugate complex. In both cases the singularity is unstable if

$$\frac{R + r - RrkS}{R} \geqq \frac{C_2}{C_1} \qquad [138.4]$$

If the singularity is unstable, self-excitation from rest is possible. By approximating the experimental function $I_a = \phi(krI_1)$ by a polynomial and by applying the bifurcation theory, it is possible to establish the existence of a stable limit cycle. We shall omit this calculation since the fact that the Heegner circuit is capable of producing continuous self-excited oscillations is well known.

Although the Heegner circuit does not produce quasi-discontinuous relaxation oscillations, it is mentioned here as an example of the difficulties which appear if an analysis is attempted on the basis of a more or less intuitive physical argument. In the meantime, this circuit provides an additional example for testing the validity of the basic assumption.

139. TRANSITION BETWEEN CONTINUOUS AND DISCONTINUOUS SOLUTIONS OF DEGENERATE SYSTEMS

On the basis of the preceding argument, one might ask whether a gradual modification of a given dynamical system, for example, an electric circuit, might cause a transition from continuous performance to quasi-discontinuous performance, or vice versa. It will now be shown that such a transition is indeed possible and that its mathematical formulation can be reduced to the question of the appearance, or disappearance, of critical points as a result of the variation of a certain parameter in the differential equations. In order to show this, let us consider a slightly modified Heegner circuit such as that shown in Figure 139.1. The modification in question is that the capacitor C_2, instead of being connected to the point B, is now connected by an adjustable sliding contact to some point E along the resistor r, as shown. Let r_1 be the resistance between the points B and E and r_2 be that between E and D, where $r_1 + r_2 = r$ and $r_1/r = \beta$. Proceeding as we did in Section 138, we obtain the following system of equations:

$$\frac{dI_1}{dt} = -\frac{1}{(1-\beta)rC_1}I_1 + \frac{1}{(1-\beta)rC_2}I_2$$

$$[139.1]$$

$$\frac{dI_2}{dt} = \frac{[\beta r + R - rR\phi' kr(I_1 + \beta I_2)]\frac{I_1}{C_1} - [r + R - rR\phi' kr(I_1 + \beta I_2)]\frac{I_2}{C_2}}{(1-\beta)[R + \beta r - \beta rR\phi' kr(I_1 + \beta I_2)]}$$

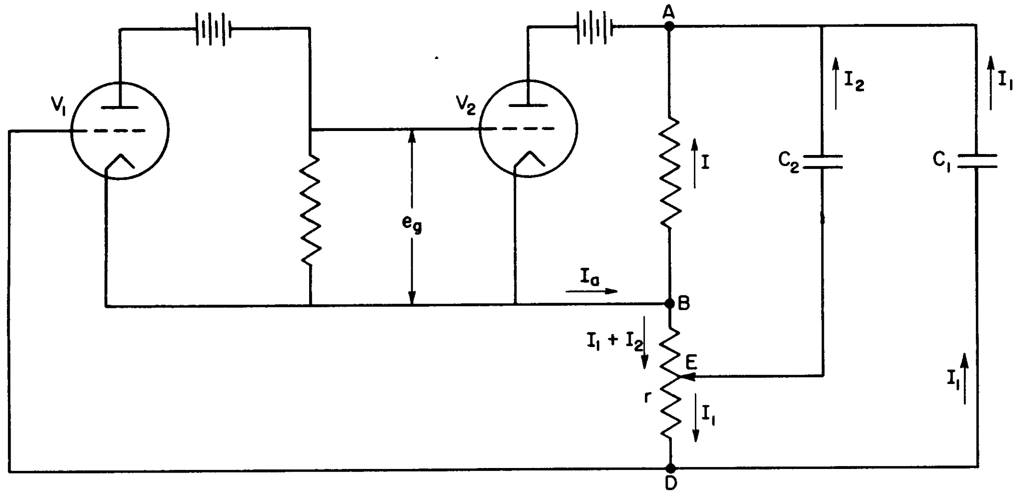


Figure 139.1

It is apparent that for $\beta = 1$, that is, $r_1 = r$, the circuit reduces to that shown in Figure 135.1 where a discontinuous performance occurs. For $\beta = 0$ and $r_1 = 0$, we have Heegner's circuit, which produces continuous oscillations. For some intermediate value of β, the co-factor of $(1 - \beta)$ in the denominator of the second equation [139.1] may vanish, which means that the otherwise continuous oscillation will undergo a discontinuous jump parallel to the I_2-axis as shown by the broken lines in Figure 139.2. This generally occurs when the system [139.1] is characterized by a saddle point, and the transition range is at the point where an unstable focal point degenerates into a saddle point, see Figure 18.1.

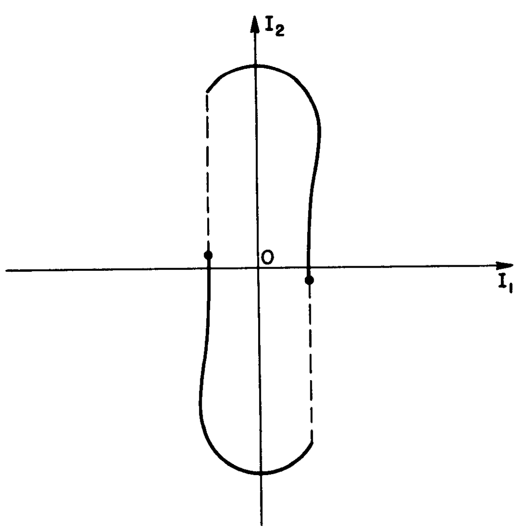


Figure 139.2

We shall not enter into a further study of these complicated and relatively unexplored phenomena. It is sufficient to emphasize once more that their apparent complexity is related to the appearance, or disappearance, of critical points in the differential equations.

CHAPTER XXIII

MECHANICAL RELAXATION OSCILLATIONS

140. INTRODUCTORY REMARKS

To date, mechanical relaxation oscillations have been studied less than electrical ones. As was mentioned in the introduction to this report, which appears in Part I, two principal reasons account for this. First, the determination of the parameters of a mechanical system is generally more complicated than the determination of the parameters of an electric circuit. Secondly, mechanical relaxation oscillations always appear as undesirable parasitic phenomena, and the endeavor of engineers is directed toward their elimination rather than their study. These phenomena generally take the form of "jarring" motions resulting from dry friction, misalignment of machinery, and similar factors. Whether any useful applications of these effects can be found is difficult to say. Very likely this state of affairs will continue until the laws of friction, which usually account for the appearance of such effects, are better understood and can be controlled with a prescribed degree of reliability. For the time being, the whole subject of mechanical relaxation oscillations is purely academic.

In addition to mechanical relaxation oscillations proper, of which we will indicate examples in Sections 141 and 142, there exists a class of oscillations maintained by periodically timed impacts. These oscillations also exhibit quasi-discontinuous behavior describable by piecewise analytic trajectories. These impact-maintained oscillations can hardly be considered to be of a *relaxation* type, although very frequently, and perhaps improperly, they are classified as such. Typical examples of these systems are clocks and quenched spark oscillators, in which the energy of a linear dissipative dynamical system is increased periodically in a quasi-discontinuous manner by special timing. We shall leave the investigation of impulse-excited systems to a later chapter and will investigate here a simple mechanical system of a pure relaxation type.

Let us consider a non-linear differential equation of the form

$$m\ddot{x} + kx = -F(\dot{x}) = -\mu f(\dot{x}) \qquad [140.1]$$

where $F(\dot{x})$ is a certain non-linear function of velocity which we shall identify with friction. If μ is small, Equation [140.1] is of a quasi-linear type and can be discussed by the standard methods of Part II. More specifically, for particular forms of the function $f(x)$ self-excitation may occur, and the oscillation may approach a stable limit cycle, as was studied in Part I in connection with Froude's pendulum; see Section 8.

When μ is large, Equation [140.1] cannot be solved by existing analytical methods, and at present the only guide to its solution is the qualitative analysis of Liénard;* see Section 36. We know from his analysis that the periodic trajectory of the differential equation when μ is large generally consists of two pairs of branches. On one pair of branches the motion of the representative point is slow and the displacements are large, so the system remains for a relatively large fraction of its relaxation period in that region. On the other pair of branches, however, the motion is very rapid. This phase of the motion is of very short duration so that, in spite of the large values of velocity and acceleration prevailing during these short fractions of the cycle, the coordinate has no time to change appreciably. Liénard's analysis fails, however, to determine the "corners" connecting these branches in a closed analytic curve; thus a rigorous analytic solution is still lacking.

The mechanical relaxation oscillations caused by friction exhibit the familiar picture previously studied. In fact, there are time intervals when the slow motion becomes a state of rest, followed by an interval of exceedingly rapid motion, followed by another period of rest, and so on. We are thus confronted with a special type of relaxation oscillation appearing as a kind of "jarring" motion.

141. QUALITATIVE ASPECTS OF A MECHANICAL RELAXATION OSCILLATION

Since a condition of degeneration, as previously shown, is essential for the appearance of quasi-discontinuous solutions of a differential equation of the second order and since we are trying to approximate the relaxation process by a motion in conformity with the Liénard analysis, we must consider a mechanical system with a very small mass.

Under this assumption, the motion of the system is determined mainly by the balance between the restoring force kx and the friction force $\mu f(\dot{x})$. When this balance is momentarily lost at a certain instant of the cycle, the acceleration and the velocity of the system may suddenly reach very large values since, by our assumption, the mass is very small.

When $m = 0$, Equation [140.1] degenerates into the equation

$$F(\dot{x}) = -kx \qquad [141.1]$$

which describes the pair of branches of the Liénard cycle on which the system remains a relatively long time and the coordinate changes appreciably. The balance between the restoring force and the friction force takes place on these branches. Differentiating this equation, one has

---

* Recent research (12) of J.A. Shohat gave hope of bridging this gap, but these efforts have been interrupted by his untimely death.

$$F'(\dot{x})\ddot{x} = -k\dot{x} \qquad [141.2]$$

If the function $F(\dot{x})$ is such that $F'(\dot{x}) < 0$ in a certain region, Equation [141.2] shows that this region is unstable. If at a certain point in this region $F'(\dot{x}) \approx 0$, the acceleration $\ddot{x}$ may acquire a very large value so that a new balance of forces will appear in which the term $m\ddot{x}$ will play a role in spite of the assumed smallness of the mass m. We thus reach the conclusion that this condition characterizes the second pair of branches of the Liénard cycle which are traversed in a very short time. During this phase of the cycle the acceleration is very large and the velocity varies in a quasi-discontinuous manner by a finite quantity, but the coordinate does not change appreciably because of the short duration of this phase.

The qualitative aspect of the motion in this case is similar in all respects to that which occurs when a ball strikes a wall. However, the underlying dynamical facts are different. For a mechanical impact, the quasi-discontinuities in velocity and acceleration described above are due to an external force, the reaction of constraint when the ball strikes the wall. In Equations [141.1] and [141.2] these discontinuities appear as a result of the disappearance of the balance of forces expressed by Equation [141.1] and are due to the assumed peculiarities of the function $F(\dot{x})$.

The preceding argument can be condensed somewhat by putting $\dot{x} = y$ in Equation [141.2], which gives

$$\frac{dy}{dt} = -\frac{ky}{F'(y)} \qquad [141.3]$$

It is apparent that this equation has a critical point when $F'(y) = 0$ and, hence, ceases to describe the phenomenon at that point. In order to determine the discontinuity by which we idealize the quasi-discontinuity of the physical problem, we must again apply the condition of Mandelstam. In the idealized case, $m = 0$; hence, the kinetic energy is zero so that the total energy is the potential energy, which is a function of the coordinate x. From the condition of Mandelstam we infer, therefore, that the function $x(t)$ is continuous although it has a discontinuous second derivative, as is to be expected. The situation is thus similar to that which we have studied in Chapter XXI in connection with electrical relaxation oscillations. It also resembles the results obtained in the theory of mechanical impacts, with the difference, however, that the quasi-discontinuity here arises from the existence of a critical point in the differential equation and is not due to external impulsive excitation as it is for a mechanical impact where the energy also changes abruptly.

142. MECHANICAL RELAXATION OSCILLATIONS CAUSED BY NON-LINEAR FRICTION

Since Equation [141.2] has a critical point for $F'(\dot{x}) = 0$, we conclude that relaxation oscillations are possible in a system of this kind whenever the friction force considered as a function of the velocity $\dot{x}$ possesses extremum values. Moreover, the general condition specified in Section 133, that is, that $F(\dot{x})$ should be a two-valued function of x, must also be fulfilled.

In practice, these theoretical conditions are frequently encountered. Thus, when a shaft rotates in a bearing, Sommerfeld (17) has indicated that the friction force goes through a minimum for a speed v_1 given by the equation

$$v_1 = \frac{1}{15.1}\,\frac{\delta^2 P}{\lambda r^2} \qquad [142.1]$$

where P is the pressure,

λ is the coefficient of viscosity of the lubricant,

r is the radius of the shaft, and

δ is the thickness of the oil film.

Chaikin and Kaidanowski (18) have investigated mechanical relaxation oscillations by the device described in the following paragraphs.

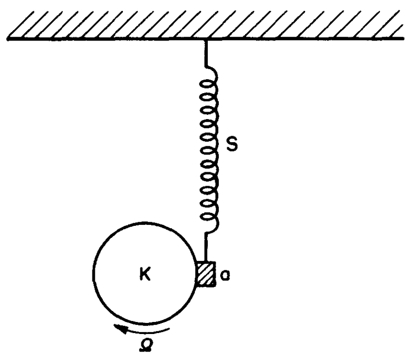


Figure 142.1

A relatively small mass a forming a Prony brake engaged frictionally a rotating shaft K, as shown in Figure 142.1. The mass a was centralized in a definite position by a rather strong spring S. A definite friction force was secured by means of another spring not shown, pressing a against the shaft K.

The differential equation of motion of a, neglecting its mass (which results in the degeneration of the equation to the first order), is

$$r\Big(F\big[(\Omega - \dot{\phi})r\big]\Big) = c\phi \qquad [142.2]$$

where $F(v) = F\big[(\Omega - \dot{\phi})r\big]$ is the friction force, a non-linear function of the relative peripheral velocity $v = (\Omega - \dot{\phi})r$,

Ω is the angular velocity of the shaft K,

$\dot{\phi}$ is the angular velocity of a for small departures ϕ, and

c is a constant depending on the spring's strength.

The approximate form of the function $F(v)$ is shown in Figure 142.2. It is noted that the tangent to this characteristic is horizontal at the point B(v = 0). When $v = 0$, a moves together with K, and $\dot{\phi} = \Omega$. When a is not moving, $\phi = 0$ and $v = v_0 = \Omega r$. To the right of the point $v = v_0$ the mass a and

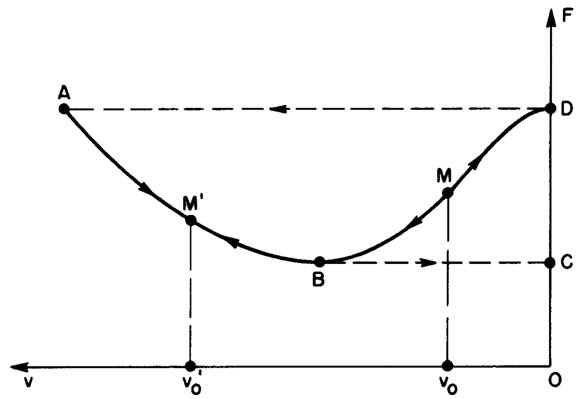


Figure 142.2

the shaft K move in the same direction; to the left of $v = v_0$ they move in opposite directions. Because of the constraining spring S no continuous motion of a is possible; there may, however, be a position of equilibrium for which $\dot{\phi} = 0$. As long as $v = 0$, that is, $\dot{\phi} = \Omega$, the spring S is gradually stretched, and the friction force F is a static force which may have any value whatever provided $F \leqq F_1$, where F_1 is the limit of static friction. When a is moved by the amount ϕ_1, for which $c\phi_1 = F_1 r$, the limit of static friction is reached. During this phase ($v = 0$) of the phenomenon, the representative point moves along the axis of F until it reaches the point D, the limit of static friction.

Differentiating Equation [142.2], we have

$$- r^2 F'[(\Omega - \dot{\phi})r]\ddot{\phi} = c\dot{\phi} \qquad [142.3]$$

where F' designates $\frac{dF}{d\phi}$. It is seen that $F' < 0$ when $\dot{\phi}$ and $\ddot{\phi}$ have the same signs, and $F' > 0$ when they have opposite signs. Moreover, for $v = v_0$ the quantity $\dot{\phi}$ goes through zero and changes its sign. If the speed Ω of the shaft is relatively small, the point M will be on the part BD of the dynamical characteristic of friction which may be considered here as the *phase line* of the process, see Section 134. On the part BD of the phase line, $F' < 0$, so that $\dot{\phi}$ and $\ddot{\phi}$ have the same sign; this part of the phase line is therefore unstable, as shown by the arrows. It follows, therefore, that when the representative point has reached the point D during the "static friction period" CD of the process, it cannot pass onto the dynamic characteristic DBA. We conclude that the point D is a critical point.

One can illustrate this also by putting $\dot{\phi} = y$ in Equation [142.3], which gives

$$\frac{dy}{dt} = - \frac{cy}{r^2 F'[(\Omega - y)r]}$$

From this expression it follows that $\frac{dy}{dt} = \ddot{\phi} = \infty$ at the point D since $F'[(\Omega - \dot{\phi})r] = 0$ at this point. At D, the representative point undergoes a discontinuity, the direction of which is parallel to the v-axis because here the discontinuity occurs at a constant potential energy, as required by the condition of Mandelstam. The jump terminates at the point A where the characteristic is met, and the subsequent motion of the representative point is again continuous along the branch BC until another critical point is encountered at the point B where $F' = 0$. Here another jump DA occurs, after which the phase CD of static friction begins anew; during that phase of the motion the representative point again moves continuously along the branch CD of the cycle DABCD.

One notes an analogy between this example of a mechanical relaxation oscillation and the example of the relaxation oscillation of a neon tube, Section 134. In both examples, relaxation oscillations are possible if the phenomenon is confined to an unstable region of the characteristic. In the neon-tube oscillator this is accomplished by adjusting the resistance so that the straight line $\frac{E - V}{R}$ cuts the non-linear characteristic, the phase line of the process, on its unstable branch. With the brake described in this section, a similar effect is obtained by running the shaft K at a relatively slow speed Ω so as to have the point M in Figure 142.2 in the region where $F' < 0$. There are no relaxation oscillations in a neon-tube circuit if the line $\frac{E - V}{R}$ intersects the characteristic on its upper stable branch, see Figure 134.2; likewise, no mechanical relaxation oscillations are observed, see Figure 142.2, if the angular velocity Ω of the shaft is large enough so that the point M$'$ corresponding to $v_0' = \Omega r$ is on the stable branch CD of the friction characteristic. The piecewise analytic cycles ABCDA in both Figures 134.2 and 142.2 are indicated by the corresponding letters. Thus, for example, the branch CD corresponds to the period when the capacitor is charged (Figure 134.2) and to the period when the static friction force F increases with no slipping between a and K (Figure 142.1). Branch AB in both examples corresponds to continuous trajectories of a non-linear differential equation. In both figures the discontinuous stretches DA and BC are determined by the condition of Mandelstam, and so on. Plotted in the (ϕ, t)-plane, the curves of the mechanical relaxation oscillations observed by Chaikin and Kaidanowski have a typical "saw-tooth" appearance, see Figure 142.3, characterizing electrical relaxation oscillations of a similar nature.

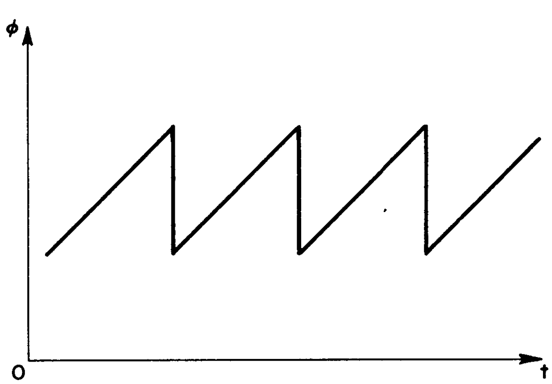


Figure 142.3

This discussion emphasizes the features which all relaxation oscillations have in common.

CHAPTER XXIV

OSCILLATIONS MAINTAINED BY PERIODIC IMPULSES

143. INTRODUCTORY REMARKS

We have defined "relaxation oscillations" as stationary self-excited oscillations exhibiting quasi-discontinuities at some points of their cycle. One of the fundamental properties of these oscillations is that their stationary state does not depend on the initial conditions of the system. In that respect they resemble the continuous oscillations of the limit-cycle type studied in Chapter IV with the difference, however, that because of the existence of discontinuous stretches their properties are different from those of continuous oscillations.

As a somewhat different type of oscillation appear the so-called *impulse-excited oscillations.* In some respects these oscillations resemble relaxation oscillations as defined above; in some other respects they differ from them.

The feature common to both types of oscillations is that they can be represented in a phase plane by piecewise analytic trajectories. For both types, also, the stationary state does not depend on initial conditions; moreover, the stationary trajectories of both are "closed" by discontinuous stretches.

The essential difference between these two types of oscillations lies in the physical process during the quasi-discontinuous rapid changes, idealized as mathematical discontinuities. In the pure relaxation oscillations with which we have been concerned in preceding chapters of Part IV, the energy stored in the system during the quasi-discontinuous interval does not change appreciably. This results in an idealized picture of these phenomena with an assumption that the energy *does not change* in the interval $(t_0 - 0, t_0 + 0)$. Moreover, the existence of these discontinuities, as we saw, was formally reduced to that of the critical points in the differential equations describing the system.

For the impulse-excited oscillations which we are going to investigate in this chapter, a difference exists, in that the discontinuous stretches in the representation of the phenomenon by phase trajectories correspond precisely to the impulsive changes of the energy of the system.

In the investigation of relaxation oscillations proper, we had to rely mostly on the theory of electric circuits. We shall begin the study of impulse-excited oscillations by investigating the behavior of a mechanical device known for centuries, the clock.

144. ELEMENTARY THEORY OF THE CLOCK

A clock is a mechanical device consisting of three principal elements:

1. An oscillatory dissipative system, for example, an ordinary pendulum, a torsional pendulum with a hair spring, etc.

2. A source of energy, for example, a weight, a main spring, etc., which has to be replenished periodically ("winding" the clock).

3. An escapement connecting periodically the first and second elements.

The purpose of the escapement is to release periodically the energy stored in the second element in the form of an impulse and to apply this impulse to the first element at an appropriate instant of the oscillation.

Since the instant at which the impulse is released by the escapement is uniquely determined by the motion of the system, the system is *autonomous* in the sense that all three elements, 1, 2, and 3, of the mechanism are connected so as to form a single unit, just as a thermionic generator forms a single autonomous unit comprising an oscillating circuit, a battery, and an electron tube with its grid control circuit. We shall see later that it is possible to arrange the operation of an electron-tube circuit so that it resembles the performance of a clock, but in general the mode of operation of a clock differs in other respects from that of the thermionic oscillators commonly used.

The escapement communicates to the oscillatory system periodically timed impulses which we shall first idealize as instantaneous and of a constant value a. Let v_1 be the velocity of the system immediately after the first impulse, from which instant we wish to study the motion. If we assume that the system is a linear dissipative one with a decrement d per cycle, the velocity v_2' immediately before the second impulse will be $v_2' = v_1 e^{-d}$. Immediately after the second impulse, the velocity will be $v_2 = v_2' + a = v_1 e^{-d} + a$, where a is the impulsive increment of the velocity, and so on for subsequent impulses. A stationary state will be reached when, beginning with a certain number n of impulses, $v_n \approx v_{n+1} \approx v_{n+2} \approx \cdots \approx v_0$, which gives

$$v_0 = \frac{a}{1 - e^{-d}} \qquad [144.1]$$

The representation of the process in the phase plane is shown in Figure 144.1, where one spiral trajectory S of the system is indicated. As was shown in Section 5, these spirals form a continuous family so that through every ordinary point of the phase plane passes one and only one such spiral. Let us designate as u the segment AB of the $\dot{x}$-axis. This intercept represents the change of the initial radius vector OA after one revolution 2π. From the properties of the logarithmic spiral it is apparent that the intercept u, considered as a function of r_0 = OA, is a monotonic function $u(r_0)$. Since the spirals S form a continuous family, it is also obvious that the function $u(r_0)$ is a continuous monotonically increasing function of the radius vector r. On the other hand, by our hypothesis, the impulsive changes of the velocity v due to the operation of the escapement are constant and hence are represented by constant jumps a along the $\dot{x}$-axis. It follows, therefore, that a steady-state condition will be attained when a radius vector r_0 = OA is reached for which $u(r_0) = a$. This stationary condition is represented by a piecewise analytic cycle BCAB shown in Figure 144.1 by a heavy line; this cycle consists of a convergent spiral BCA (since, by our assumption, the system is dissipative) "closed" by a discontinuous jump AB representing the impulsive increase of the velocity v caused by the impact. It is also clear that such a stationary piecewise analytic cycle BCAB is of a *stable limit-cycle* type in the sense that, if one starts from the relatively small values of the initial radius vector r_0 = OA, the value of a is initially greater than the corresponding value of the intercept $u(r_0)$ in this region of the phase plane. Physically this means that the energy communicated by the impulse is greater than the amount of energy which the system is able to dissipate during one cycle 2π, which will result in the fact that the subsequent radius vectors r_0 = OA will grow initially. If, on the contrary, the initial velocity of the system is sufficiently great, the dissipation of energy per cycle is greater than the impulsive increments communicated by the escapement so that the spirals will gradually shrink. The stable condition is reached when the impulsive increments a are just equal to the intercepts $u(r_0)$ in a particular region. In view of the continuity and the monotonic character of the function $u(r_0)$, it is clear that there exists one and only one piecewise analytic limit cycle BCAB of the clock and it is stable. This elementary discussion accounts for the fact that the clock's performance is of a *limit-cycle* type and as such does not depend on initial conditions, see Chapter IV; in other words, the ultimate performance of a clock does not depend on how it has been started. Once it is started the operation of the clock depends

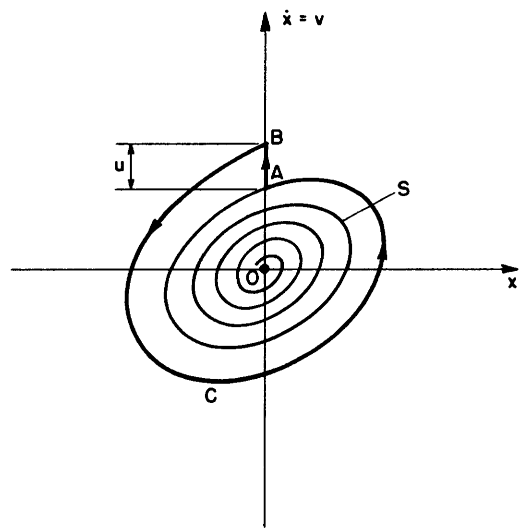


Figure 144.1

entirely on the parameters of the system and has nothing to do with the initial conditions. The converse is true for a mathematical pendulum.

In one respect the elementary theory discussed above fails to account completely for the observed facts. For, if a clock is wound, this simplified theory indicates that the clock should start by itself even if the initial disturbance is infinitely small. In other words, it predicts a *soft* self-excitation of the clock. In reality, unless a clock is given a certain minimum disturbance, for example, shaking, it will not start. This threshold is partially due to the existence of Coulomb friction not taken into consideration here. There is, however, another reason why the elementary theory is not complete, namely, our assumption that the jump a in velocity caused by the impact remains the same whatever the velocity of the system. Andronow and Chaikin (13) have shown that by a slight refinement of the preceding theory it is possible to explain better the observed facts.

Let us assume that the change of kinetic energy during the impact remains constant, that is,

$$\frac{mv_1^2}{2} - \frac{mv_0^2}{2} = \text{constant}$$

In the phase-plane representation this is equivalent to the condition

$$y_1^2 - y_0^2 = h^2 \qquad [144.2]$$

where h is a constant determined by the properties of the escapement and y_0 and y_1 are the values of the radius vector immediately before and immediately after the impact. It follows, therefore, that the jump a in the phase plane, instead of being constant as was originally assumed, is now given by the equation

$$a = \sqrt{y_0^2 + h^2} - y_0 \qquad [144.3]$$

that is, it decreases with increasing velocity y_0 according to a hyperbolic law, as shown in Figure 144.2.

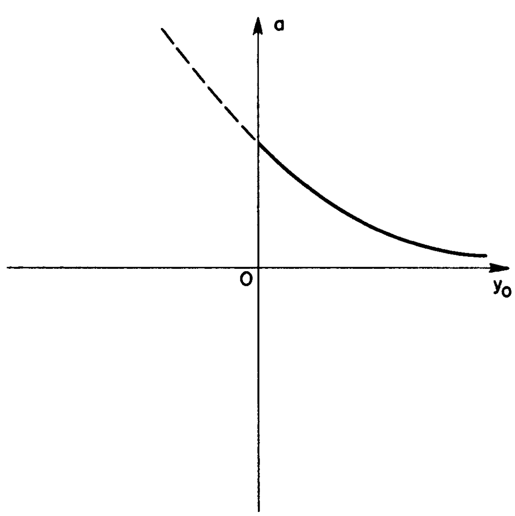


Figure 144.2

This refinement of the theory, although it leads to conclusions concerning the existence of a stable piecewise analytic limit cycle which are similar to those obtained by the original simplified theory, is not yet sufficient to explain the fact that a clock is a system with a *hard* self-excitation.

In order to extend the theory still further it is necessary to investigate how Coulomb friction manifests itself in the representation of motion by phase trajectories.

145. PHASE TRAJECTORIES IN THE PRESENCE OF COULOMB FRICTION

The customary idealization of Coulomb, or dry, friction consists in assuming that the friction force f_0 remains constant during the motion and changes its sign with a change of the direction of motion, as shown in Figure 145.1.

The motion of a system with one degree of freedom in the presence of Coulomb friction cannot be described by one differential equation but requires two equations, one for $\dot{x} < 0$ and the other for $\dot{x} > 0$, namely,

$$m\ddot{x} + kx = +f_0; \qquad m\ddot{x} + kx = -f_0 \qquad [145.1]$$

Putting $\frac{k}{m} = \omega_0^2$ and $|f_0| = a\omega_0^2$, these two equations are

$$\begin{aligned} \ddot{x} + \omega_0^2 x &= +a\omega_0^2 \quad \text{when } \dot{x} < 0 \\ \ddot{x} + \omega_0^2 x &= -a\omega_0^2 \quad \text{when } \dot{x} > 0 \end{aligned} \qquad [145.2]$$

Introducing the variables $x_1 = x - a$ for the first equation [145.2] and $x_2 = x + a$ for the second, one obtains two equations, namely,

$$\begin{aligned} \ddot{x}_1 + \omega_0^2 x_1 &= 0 \quad \text{when } \dot{x} < 0 \\ \ddot{x}_2 + \omega_0^2 x_2 &= 0 \quad \text{when } \dot{x} > 0 \end{aligned} \qquad [145.3]$$

These equations are obviously identical, with the difference, however, that the center of oscillation is displaced from $+a$ to $-a$, and vice versa, each time $\dot{x}$ changes its sign. It is to be noted that $\dot{x}$ itself does not enter into the equations of motion. The "change of equations" occurs at the instant when $\dot{x} = 0$. In the (x,t)-plane such a motion can be represented by the curve shown in Figure 145.2. Assume that initially the system has been deviated to the right ($x_{01} > 0$) and then released with an initial velocity $\dot{x}_{01} = 0$. The system will then move to the left ($\dot{x} < 0$). As was just mentioned, this is a sinusoidal motion with respect to the axis t_1 displaced a distance $+a$ from

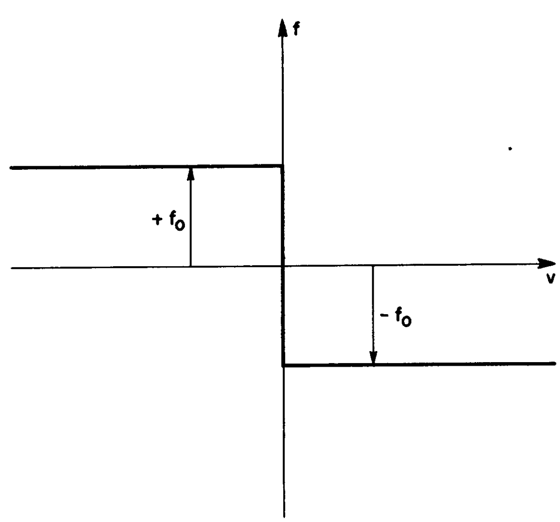


Figure 145.1

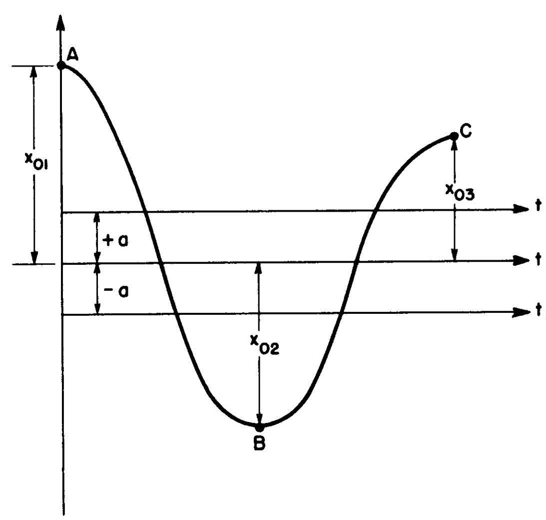


Figure 145.2

the t-axis. At the point B the velocity $\dot{x}$ is equal to zero, and the amplitude is $|x_{02}| = |x_{01} - 2a|$ *with respect to the t-axis.* At the point B the change of equations takes place, and the "acting abscissa axis" is now t_2 displaced a distance $-a$ from the t-axis. With respect to the t_2-axis, the motion is again sinusoidal, so that the point C (at which $\dot{x} = 0$) is at the same distance from the t_2-axis as the point B was originally. With respect to the t-axis, however, the amplitude of the point C is $x_{03} = x_{02} - 2a$. It is seen that with respect to the t-axis the amplitudes decrease in an *arithmetic progression* with the constant difference $2a$. The motion stops after a certain number of swings. The time interval between two consecutive maximums (A and C in Figure 145.2) is the same as for a harmonic oscillator, as follows from the preceding discussion. It is to be noted again that the curve ABC··· is a piecewise analytic curve which loses its analyticity at the points A, B, C, ···, at which there is a discontinuity in the second derivative.

The representation of such a motion by phase trajectories is obtained if we observe that each of the equations [145.3] is represented by an elliptic trajectory around the vortex point, see Section 1.

If we let $\frac{dx}{dt} = y$, Equations [145.3] become

$$\frac{dy}{dx} = -\frac{\omega_0^2(x-a)}{y} \quad \text{when } y < 0$$

$$\frac{dy}{dx} = -\frac{\omega_0^2(x+a)}{y} \quad \text{when } y > 0 \qquad [145.4]$$

Integrating these equations, one gets

$$\frac{(x-a)^2}{R_1^2} + \frac{y^2}{R_1^2\omega_0^2} = 1 \quad \text{when } y < 0$$

$$\frac{(x+a)^2}{R_2^2} + \frac{y^2}{R_2^2\omega_0^2} = 1 \quad \text{when } y > 0 \qquad [145.5]$$

where R_1 and R_2 are the constants of integration determined at the end of each preceding interval, that is, when $\dot{x} = 0$. Equations [145.5] represent two families of ellipses, Γ_1 and Γ_2, whose centers are displaced on both sides of the origin by a constant quantity $\pm a$, as shown in Figure 145.3. Let us consider the motion beginning with a point A representing the initial conditions on one of the trajectories Γ_2 which is drawn in a heavy line. At the point B, where $y = 0$, the change of equations takes place, and the representative point goes onto the trajectory Γ_1 passing through B. The motion on that trajectory will continue up to the point C at which the representative point will go onto a trajectory Γ_2 which passes through C. On that trajectory the change of equations should occur at the point D. It is apparent, however, that such a trajectory would be confined within the zone limited by the broken lines $x = +a$ and $x = -a$, which characterize the zone of static friction. Hence once a trajectory is reached which does not emerge from that zone, the motion ceases.

It is thus seen that idealized Coulomb friction is capable of being represented by piecewise analytic trajectories formed by elliptic arcs of the two families Γ_1 and Γ_2 with centers symmetrically displaced by a constant quantity a on both sides of the origin O. The junction points of the arcs are situated on the x-axis. At these points the curve has a continuous first, but a discontinuous second, derivative as may be expected from the fact that the friction force changes discontinuously at these points. Instead of having a point of equilibrium, systems possessing Coulomb friction have a *line of equilibrium* O_1O_2, which means that any point on the segment O_1O_2 is a point of equilibrium and that the segment O_1O_2 itself is the zone of *static friction*.

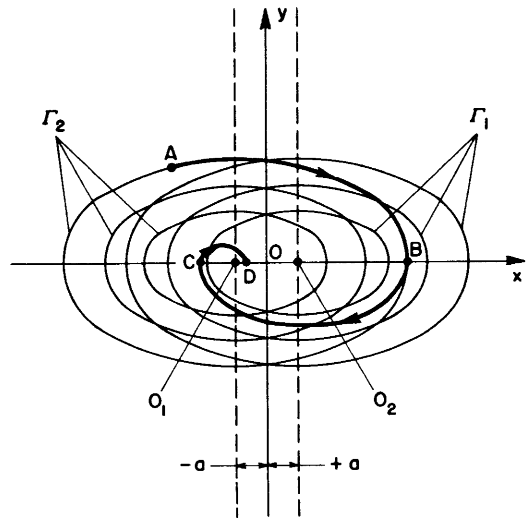


Figure 145.3

Making use of these conclusions regarding the form of trajectories in the presence of Coulomb friction, Andronow and Chaikin have further elaborated the theory of the clock. We shall omit the details of their theory and will merely indicate their argument as well as their final conclusions.

By introducing "angular time" $\tau = \omega_0 t$, as we did on a number of occasions, Equations [145.3] can be reduced to the form

$$\frac{d^2x_1}{d\tau^2} + x_1 = 0 \qquad \text{when } \dot{x} < 0$$
$$\frac{d^2x_2}{d\tau^2} + x_2 = 0 \qquad \text{when } \dot{x} > 0 \qquad [145.6]$$

The trajectories of these equations are two families C_1 and C_2 of concentric circles whose centers are displaced on the x-axis on both sides of the origin by a fixed quantity a. By reproducing the preceding argument and by expressing the condition that the ultimate trajectory always emerges from the "dead zone" of static friction, these authors show that, if the trajectory is to emerge ultimately from the dead zone, the impulsive change of kinetic energy h^2 must be equal to or greater than $16f_0^2$, where f_0 is the Coulomb force. This means that a clock is a mechanism with a *hard* self-excitation. In other words, unless the condition stated above concerning the relation between h^2 and $16f_0^2$ is satisfied, a clock will not start. If, however, this condition is satisfied, the clock will start and will approach its piecewise analytic limit cycle irrespective of any other aspects of the initial conditions.

146. ELECTRON-TUBE OSCILLATOR WITH QUASI-DISCONTINUOUS GRID CONTROL

We will investigate in this section the behavior of an electron-tube circuit exhibiting features resembling those of a mechanical system possessing Coulomb friction and acted on by impulses. The analogy, as we shall see, is purely formal, but it will be helpful in approaching from a somewhat new viewpoint the difficult subject of self-excited oscillations. The interesting feature of the analysis made by Andronow and Chaikin is that it permits investigating questions concerning the establishment and stability of self-excited oscillations by a method similar to that used in the preceding section, except that here one may have *negative Coulomb friction* maintaining the oscillations.

Consider the circuit shown in Figure 146.1 with positive directions and customary notations indicated. The symbol x designates the current in the inductance L. The differential equation of the circuit is

$$L\frac{dx}{dt} + Rx + \frac{1}{C}\int(x - I_a)\,dt = 0 \qquad [146.1]$$

Differentiating and rearranging, we obtain

$$L\ddot{x} + R\dot{x} + \frac{1}{C}x = \frac{1}{C}f(e_g) \qquad [146.2]$$

where $I_a = f(e_g)$ is the non-linear characteristic of the electron tube. We will consider a peculiar performance of this circuit, namely, the one which appears when the coefficient of mutual inductance M is made very large. As a result of this, the amplitude $|e_g|$ of the grid-voltage variation will also be large; this will cause the electron tube to act more or less as a switch operating between the points A, when $I_a = 0$, and B, when $I_a = I_s$, where I_s is the saturation plate current. It is apparent that when the grid potential varies between the limits $\pm e_g$, as shown in Figure 146.2, the interval A_1B_1 corresponding to the variations $\pm e_{g1}$, which are instrumental in causing the plate current to vary, is small in comparison with the total interval AB, and it is possible to idealize the performance as a discontinuous characteristic such as

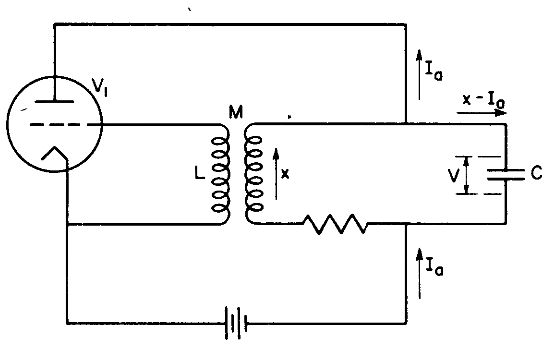


Figure 146.1

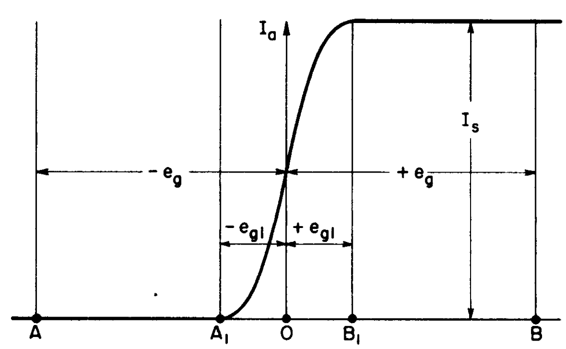


Figure 146.2

that shown in Figure 145.1. Instead of being described by one non-linear differential equation [146.2], the phenomenon will be described alternately by two linear differential equations

$$\ddot{x} + 2h\dot{x} + \omega_0^2 x = 0; \qquad \ddot{x} + 2h\dot{x} + \omega_0^2 x = \omega_0^2 I_s \qquad [146.3]$$

where $2h = \frac{R}{L}$, $\omega_0^2 = \frac{1}{LC}$, and $\omega_0^2 I_s = \frac{1}{LC} f(+e_g)$. It is obvious that the change from one equation to the other occurs at the point where $\dot{x}$ changes its sign.

We shall assume that the connections are made so that $M > 0$, $e_g > 0$ when $\dot{x} < 0$, and $e_g < 0$ when $\dot{x} > 0$. We shall investigate the second alternative later.

In view of our idealization, the anode current I_a will undergo discontinuities in the neighborhood of $\dot{x} = 0$. These discontinuities have to pass through the capacitor branch of the oscillating circuit since in the inductive branch no discontinuities of current are possible. However, the discontinuities $L\frac{dx}{dt}$ in the voltage across the inductance are possible. We conclude, therefore, that the quantities which are capable of varying discontinuously are I_a in the capacitor branch and $L\frac{dx}{dt}$ across the inductance. The quantity x remains continuous. Another continuous quantity is the voltage V across the capacitor, as was mentioned in Section 131.

It is noted that the change from one differential equation to the other, mentioned above, resembles the situation which we have already encountered in connection with Equations [145.2] describing the behavior of a system possessing Coulomb friction, and may follow a similar argument. We can write the second equation [146.3] as

$$\ddot{\overline{x}} + 2h\dot{\overline{x}} + \omega_0^2 \overline{x} = 0 \qquad [146.4]$$

where $\overline{x} = x - I_s$. In this form the first equation [146.3] and Equation [146.4] are two linear differential equations possessing focal points on the t-axis at a distance I_s from each other. We can, therefore, repeat the argument used in connection with Figure 145.2. Let us consider the phenomenon from the instant when the amplitude is x_1 and the electron-tube switch is off, which corresponds to the first equation [146.3]. During the first half-cycle the original amplitude x_1, see Figure 146.3, will be reduced by damping and will become x_2. It is noted that $x_2 < x_1$ if referred to the t-axis. At the point B the electron-tube current jumps suddenly to the value I_s and, as we saw, the new abscissa axis is now t_1, displaced above the t-axis by the quantity I_s. With respect to this t_1-axis, the initial amplitude is now $\overline{x}_2 = x_2 + I_s$ and although during the following half-cycle the amplitude is again reduced with respect to this axis because of the dissipation of energy, the position of the point C with respect to the t-axis may be at a higher level than that of A if

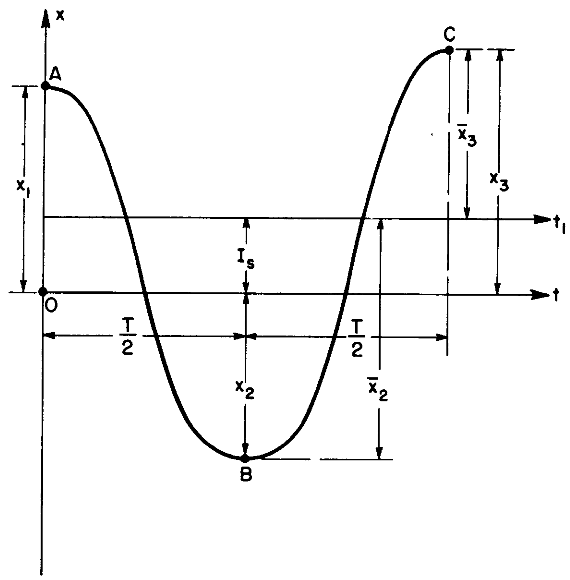


Figure 146.3

I_s is large enough. Beginning with the point C, we can resume the same argument which we used at the point A, and so on.

It is seen that, because of this quasi-discontinuous timing of the electronic switch, the amplitudes in the oscillating circuit may grow, thus outweighing the effect of the dissipation of energy. It is easy to express this condition mathematically. If one designates by x the amplitudes referred to the t-axis and by $\bar{x}$ those referred to the t_1-axis, one has

$$x_2 = x_1 e^{-\frac{hT}{2}}; \qquad \bar{x}_2 = x_1 e^{-\frac{hT}{2}} + I_s$$

$$\bar{x}_3 = \bar{x}_2 e^{-\frac{hT}{2}} = x_1 e^{-\frac{2hT}{2}} + I_s e^{-\frac{hT}{2}}; \qquad x_3 = \bar{x}_3 + I_s = x_1 e^{-\frac{2hT}{2}} + I_s e^{-\frac{hT}{2}} + I_s$$

. . . .

If the amplitude reaches a stationary value, $x_3 = x_1 = x_0$, which gives

$$x_0 = \frac{I_s(1 + e^{-\frac{hT}{2}})}{1 - e^{-hT}} = \frac{I_s}{1 - e^{-\frac{hT}{2}}} \qquad [146.5]$$

This shows that a stationary amplitude x_0 is determined solely by the properties of the circuit and is independent of the initial conditions.

The question of the stability of these oscillations can be investigated graphically by comparing their subsequent amplitudes. For instance, take the relation between x_1 and x_3, namely,

$$x_3 = x_1 e^{-hT} + I_s(1 + e^{-\frac{hT}{2}}) \qquad [146.6]$$

In the (x_1, x_3)-plane, see Figure 146.4, this relation is a straight line with a slope $\tan \alpha = e^{-hT}$, and an intercept $I_s(1 + e^{-\frac{hT}{2}}) = A$. On the other hand, if the oscillation is stationary, $x_1 = x_3$, which represents the bisector of the angle between the x_1- and x_3-axes. The point M of intersection of these lines determines the stationary amplitude x_0. If the initial amplitude x_1' is not stationary, the subsequent amplitude will be x_3' and the corrected amplitude will be x_3' on the x_1-axis to which the following amplitude x_5' will correspond on the x_3-axis, and its corrected value on the x_1-axis will be x_5'.

By the interplay of the subsequent corrections it is seen that any non-stationary amplitude tends to approach the stationary value x_0. This argument is also valid when the initial non-stationary amplitude is larger than the stationary amplitude x_0. Hence the stationary amplitude is *stable*. It should be noted that if we reverse the connection of the coils, the displacement of the x_1-axis with respect to the x-axis by the quantity I_s will occur in a direction opposite to that shown in Figure 146.3. This will result in a *reduction* of the subsequent amplitudes similar to that found in connection with Figure 145.2 illustrating the action of Coulomb friction. Viewed from this standpoint, the circuit behaves as if it had a kind of *negative Coulomb damping* when $M > 0$ and a *positive Coulomb damping* when $M < 0$. The self-excitation of the circuit occurs when the circuit is characterized by *negative Coulomb damping.*

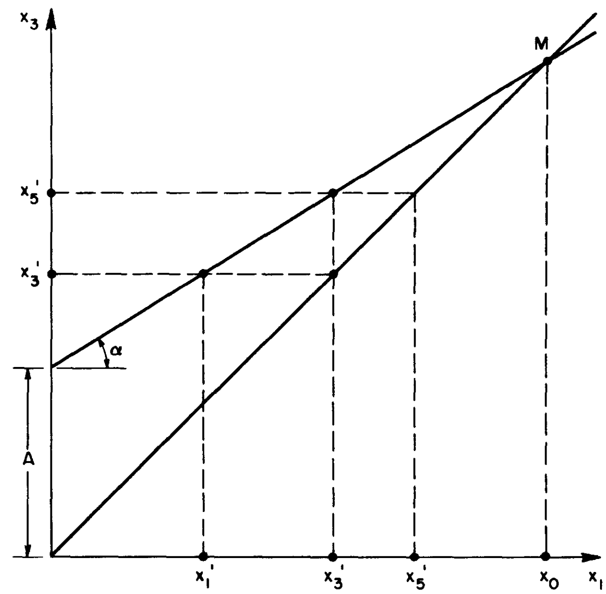


Figure 146.4

147. PHASE TRAJECTORIES OF AN IMPULSE-EXCITED OSCILLATOR

The property of negative Coulomb damping of the circuit investigated in the preceding section becomes still more striking if we analyze the phase trajectories of the system [146.3] which we write in the form

$$\begin{aligned} \ddot{x} + 2h\dot{x} + \omega_0^2 x &= 0 \\ \ddot{x}_1 + 2h\dot{x}_1 + \omega_0^2 x_1 &= 0 \end{aligned} \qquad [147.1]$$

where $x_1 = x - I_s$. Since each of these equations describes a damped oscillator, each is represented by a spiral trajectory approaching a stable focal point; see Section 5. The focal points of these equations are separated by a constant distance I_s along the x-axis in the $(x,\dot{x})$ phase plane. Figure 147.1 shows the phase diagram of the motion corresponding to Equations [147.1]. Point O is the focal point of the first equation [147.1], and $\overline{O}$ that of the second equation. Let us start from a certain arbitrary initial condition, represented, for instance, by a point 1 on the negative abscissa axis. Through Point 1 passes the spiral trajectory 1m2 having O as its focal point and representing the first equation [147.1]. At the point 2 the "change of equations"

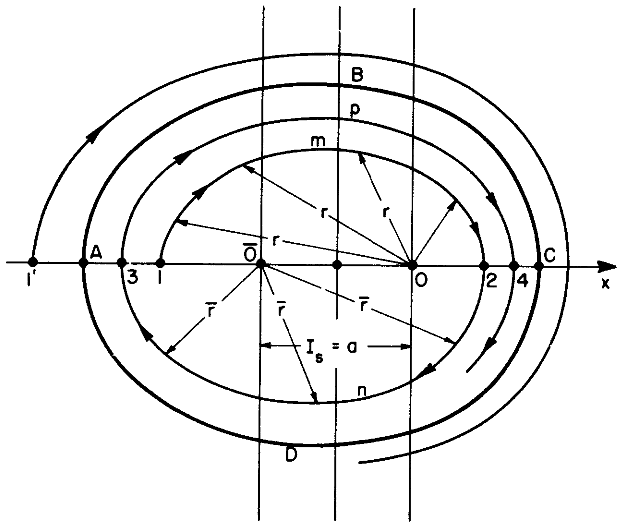


Figure 147.1

occurs, and the phenomenon is now governed by the second equation [147.1] having its focal point at $\overline{O}$. The second stretch of the spiral trajectory $2n3$ ends at the point 3, where the first equation begins to describe the phenomenon again; this arc of the spiral trajectory $3p4$ ends at the point 4, and so on. It can be shown that the piecewise analytic trajectory $1m2n3p4\cdots$ ultimately approaches a closed curve ABCDA formed by two spiral arcs, ABC with O as its focal point and CDA with $\overline{O}$ as its focal point. One obtains a similar conclusion if, instead of starting with a point 1, one starts with a point 1′ exterior to the closed curve ABCDA and repeats the argument. We shall designate by r the radius vectors with respect to the focal point O when the representative point moves on the upper spiral arcs, and by $\overline{r}$ the radius vectors with a focal point at $\overline{O}$ when the representative point moves on the lower arcs. For the subsequent radius vectors along the x-axis, we have the following obvious relations:

$$r_2 = r_1 e^{-\frac{\pi h}{\omega}}; \quad \overline{r}_3 = \overline{r}_2 e^{-\frac{\pi h}{\omega}} = (r_2 + a) e^{-\frac{\pi h}{\omega}} = (r_1 e^{-\frac{\pi h}{\omega}} + a) e^{-\frac{\pi h}{\omega}}; \ \cdots$$

One obtains easily the following general expressions:

$$r_k = a\left[1 + e^{-\frac{\pi h}{\omega}} + e^{-\frac{2\pi h}{\omega}} + \cdots + e^{-(k-2)\frac{\pi h}{\omega}}\right] + r_1 e^{-(k-1)\frac{\pi h}{\omega}}; \qquad \overline{r}_k = r_k - a$$

$$\overline{r}_s = a\left[1 + e^{-\frac{\pi h}{\omega}} + \cdots + e^{-(s-2)\frac{\pi h}{\omega}}\right] + r_1 e^{-(s-1)\frac{\pi h}{\omega}}; \qquad r_s = \overline{r}_s - a \qquad [147.2]$$

where k is an odd integer and s is an even integer. If k and s increase indefinitely, that is, if the number of turns of the piecewise analytic spiral $1m2n3p4\cdots$ increases indefinitely, the spiral approaches the closed curve ABCDA, previously mentioned, in the manner specified in Section 22. One sees this immediately from the expressions given above for r_k and $\overline{r}_s$ when $k \to \infty$ and $s \to \infty$, namely,

$$\lim_{k\to\infty}(r_k) = \frac{a}{1 - e^{-\frac{\pi h}{\omega}}} = \lim_{s\to\infty}(\overline{r}_s) \qquad [147.3]$$

The conclusions remain exactly the same if one repeats the argument for a piecewise analytic spiral $1'm'2'n'3'p'\cdots$ starting at a relatively distant point $1'$.

It is thus apparent that the closed curve ABCDA formed by the two arcs of the logarithmic spirals is a *piecewise analytic limit cycle* which the oscillation approaches as $t \to \infty$. Moreover, this limit cycle is *stable*. It is useful to emphasize once more that this limit cycle characterizing stationary periodic motion is determined exclusively by the constant parameters $a = I_s$, h, and ω of the oscillatory system and is entirely independent of initial conditions. This is an essential property of self-excited oscillations, as was pointed out in Section 22.

It is also worth mentioning that we have obtained these results of the classical theory by replacing the actual non-linear equation [146.2] by a system [147.1] of two alternating linear differential equations. Such a procedure becomes possible only after we have idealized the relatively complicated non-linear characteristic, shown in Figure 146.2, as a discontinuous step-function like that shown in Figure 145.1. In connection with this it may be useful to recall the statement made in Section 3 that the really important difference between linear and non-linear systems is their behavior *in the large* and that *local properties* are of relatively minor importance. This argument is a typical example of a situation of this kind where a certain idealization of the local properties of trajectories does not change their properties in the large and merely simplifies a problem which would otherwise be extremely complex.

CHAPTER XXV

EFFECT OF PARASITIC PARAMETERS ON STATIONARY STATES OF DYNAMICAL SYSTEMS

148. PARASITIC PARAMETERS

The study of relaxation oscillations in the preceding sections was made under the assumption of certain simplifying idealizations which resulted in degenerate equations instead of complete ones containing both *oscillatory* parameters (L and C in electrical problems and m and k in mechanical ones).

Although the advantages of such simplifications, as we saw, are numerous and the results obtained in this manner are generally found to be in agreement with experimental facts, the introduction of these simplifying assumptions is not without theoretical difficulties and, in some cases, which we will analyze here, may lead to certain complications. In spite of the fact that such cases appear generally as rare exceptions, it is important to analyze this matter in greater detail now that we are acquainted with the general method, at least in its present scope.

One of the principal difficulties, noted on several occasions, is the inconsistency between the number of integration constants appearing in a degenerate problem and that in the corresponding non-degenerate one. Thus, for example, if the non-degenerate problem involves a differential equation of the second order, two constants of integration are necessary to determine the solution; these constants appear as certain definite physical "initial conditions." In the corresponding degenerate problem, where the differential equation is of the first order, one constant of integration is sufficient to determine a solution. If, however, there still exist two physical factors to which some initial values can be assigned, these factors cannot have entirely arbitrary values but must readjust themselves eventually so that a definite relation exists between them, as was explained in Section 128. The discontinuous theory of relaxation oscillations *ignores* this rather delicate passage from the solutions of one form to those of the other form, just as the classical theory of mechanical impacts ignores the unknown dynamics of the collision process. Both theories follow a somewhat similar argument, namely, the rapidly changing motion is idealized during a very short time interval as a mathematical discontinuity occurring in the *infinitely short* time interval ($t - 0$, $t + 0$), and the loss of information resulting from intentionally overlooking the dynamics of the process is supplemented by *additional physical information* which permits solution of the problem *in the large* although certain *local* details are inevitably lost in such a procedure. In the theory of mechanical impacts this additional physical information is provided by the theorems of momentum and kinetic energy as well as by defining a certain empirical *coefficient of restitution* characterizing the instantaneous dissipation of energy during the impact; likewise, in the discontinuous theory of relaxation oscillations this additional information appears in the form of the conditions of Mandelstam.

The use of the concepts "infinitely short time interval," "infinitely large acceleration," and so on, instead of the more correct terms "very short time interval," "very large acceleration," and so on, while convenient for the description of a phenomenon *in the large*, is sometimes capable of introducing serious errors in a theoretical argument and of leading to conclusions at variance with experimental facts.

Thus, for instance, in dealing with an idealized (L,R)-circuit, we describe its behavior by the equation

$$L\frac{di}{dt} + Ri = E(t) \qquad [148.1]$$

This equation describes the phenomenon with adequate accuracy in a range sufficient for practical purposes (on the scale, say, of milliseconds). If, however, the behavior of the same circuit is studied in a different range (say, on the scale of microseconds), this equation may not give a correct answer. In fact, any resistor or inductance coil inevitably has a small *parasitic capacity* C_p, and if we take this capacity into account a more correct equation will be

$$RLC_p\frac{d^2 i_1}{dt^2} + L\frac{di_1}{dt} + Ri_1 = E(t) \qquad [148.2]$$

It was shown in Section 128 that under certain conditions the solutions $i(t)$ and $i_1(t)$ of Equations [148.1] and [148.2] may have entirely different features locally, although in the large these solutions are practically indistinguishable. For that reason an electrical engineer would prefer to use the simpler equation [148.1] rather than the more complicated one [148.2]. These facts are too well known to need further emphasis here. However, they are frequently very troublesome. Thus, for instance, a student attempting to investigate the behavior of a simple circuit by means of a cathode-ray oscillograph, instead of observing the pattern of curves in accordance with his differential equation, usually observes what is commonly called "hash," that is, a far more complicated pattern of numerous harmonics with cross modulation, and so on. The reason for this is that too many parasitic parameters have been neglected in forming the differential equation, and consequently the true differential equation is far more complicated than the assumed one.

In most practical cases such discrepancies are not very serious. However, in some special instances, analyzed in what follows, difficulties may arise which lead to paradoxes already noted by certain authors (19) (20).

A further remark may be useful. Consider the differential equation of an (L,R,C)-circuit,

$$L\ddot{q} + R\dot{q} + \frac{1}{C}q = 0 \qquad [148.3]$$

and assume that L is very small. If L is neglected, one has a differential equation of the first order. The expression "L is very small" is, however, rather indefinite in the sense that it is not important to know that L is small but rather that the consideration of the problem is restricted to the range in which the term $L\ddot{q}$ is negligible compared with the other two terms. Still more confusing is the effect of the degeneration when $C \to 0$, which leads to a meaningless result. If, however, we make $C \to \infty$, at the limit we obtain

$$L\frac{di}{dt} + Ri = 0 \qquad [148.4]$$

which is an absolutely degenerate equation of an idealized (L,R)-circuit. Here the degeneration occurs not because a parameter is small but because it is *large*. This confusion disappears if, instead of the parameter C, the capacity, we use the *elastance* $k = \frac{1}{C}$ as a parameter. It is convenient, for this reason, to consider parameters as parasitic if they are small or if the corresponding terms in the differential equation are small. The first condition is merely a definition, whereas the second condition specifies the range in which a small parameter can be neglected.

149. INFLUENCE OF PARASITIC PARAMETERS ON THE STATE OF EQUILIBRIUM OF A DYNAMICAL SYSTEM

Consider a system with several degrees of freedom expressible by the differential equation

$$a_0 x^{(n)} + a_1 x^{(n-1)} + \cdots + a_{n-2}\ddot{x} + a_{n-1}\dot{x} + a_n x = 0 \qquad [149.1]$$

where $x^{(n)} = \frac{d^n x}{dt^n}$, and so on. The corresponding characteristic equation is

$$a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-2}\lambda^2 + a_{n-1}\lambda + a_n = 0 \qquad [149.2]$$

The equilibrium is stable if the real parts of the roots of [149.2] are negative, which can be ascertained by the well-known criteria of Routh-Hurwitz. As was previously stated, we never know the exact form of the differential equations in a physical problem because of a number of parasitic parameters either entirely unknown or known only approximately.

Let us assume that Equation [149.1] is a degenerate one derived from some other equation describing more correctly the behavior of the system by taking into account a parasitic parameter α. There are two ways in which this parasitic parameter may appear, depending on whether it is of the "inductance" type or of the "elastance" type.

If it is of the inductance type, the non-degenerate equation will be

$$\alpha x^{(n+1)} + a_0 x^{(n)} + a_1 x^{(n-1)} + \cdots + a_{n-1}\dot{x} + a_n x = 0 \qquad [149.3]$$

If it is of the elastance type, it will be

$$a_0 x^{(n)} + a_1 x^{(n-1)} + \cdots + a_{n-1}\dot{x} + a_n x + \alpha\int x\,dt = 0 \qquad [149.4]$$

Differentiating this equation, we get

$$a_0 x^{(n+1)} + a_1 x^{(n)} + \cdots + a_{n-1}\ddot{x} + a_n\dot{x} + \alpha x = 0 \qquad [149.5]$$

For both types of parasitic parameter the order of the original differential equation has been raised by one unit. The corresponding characteristic equations are

$$\alpha\lambda^{n+1} + a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0 \qquad [149.6]$$

$$a_0\lambda^{n+1} + a_1\lambda^n + a_2\lambda^{n-1} + \cdots + a_n\lambda + \alpha = 0 \qquad [149.7]$$

They now have $(n + 1)$ roots, whereas the original characteristic equation [149.2] had only n roots. Since α is small, by our assumption, the old n roots change but little owing to the appearance of the new root, and we may neglect this change. The matter hinges, therefore, on the new root λ_{n+1}.

We can write Equation [149.2], in terms of its roots, as

$$a_0(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) = 0 \qquad [149.8]$$

The new characteristic equation [149.6] of the "parasitic inductance" type can accordingly be written as

$$\alpha(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)(\lambda - \lambda_{n+1}) = \frac{\alpha}{a_0}\Big[a_0\lambda^n + $$

$$+ a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n\Big](\lambda - \lambda_{n+1}) \qquad [149.9]$$

Expanding the rignt-hand side of this equation and identifying it with Equation [149.6], we get

$$\lambda_{n+1} = -\frac{a_0}{\alpha} \qquad [149.10]$$

Applying a similar argument to Equation [149.7], we obtain

$$\lambda_{n+1} = -\frac{\alpha}{a_n} \qquad [149.11]$$

The sign of the new root λ_{n+1} indicates its effect on the stability of equilibrium. The general integral of the new system is

$$x = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + \cdots + C_n e^{\lambda_n t} + C_{n+1} e^{\lambda_{n+1} t} \qquad [149.12]$$

If the conditions of stability for the original degenerate equation [149.1] are fulfilled, the stability of the new equation [149.3] or [149.5] will depend solely on the new root λ_{n+1}. If this root is *stable*, that is, if its real part is negative, the equilibrium will still be stable. If, however, the root λ_{n+1} is unstable, the non-degenerate system will be unstable although the corresponding degenerate system is stable. For a complete discussion, one must apply the Routh-Hurwitz criteria, as usual.

In practical problems the existence of parasitic parameters is generally unknown; still less known is the magnitude of such parameters. For these reasons in a complicated physical problem one is never certain about the actual conditions of stability. This is why one occasionally observes somewhat puzzling departures from theoretical predictions based on a simplified study of a system expressed in terms of differential equations with known parameters.

Conclusions obtained from the simple discussion given above may appear to some extent paradoxical. Thus, for instance, if the parasitic parameter is of the "inductance" type, see Equation [149.10], the new root λ_{n+1} is larger as the parasitic parameter α is smaller. If, however, it is of the "elastance" type, see Equation [149.11], the magnitude of the root λ_{n+1} is proportional to α.

The rather erratic "floating" of a system around its theoretical position of equilibrium can sometimes be traced to the effects of parasitic parameters, as we shall see from a few typical examples in the following sections.

150. EFFECT OF PARASITIC PARAMETERS ON STABILITY OF AN ELECTRIC ARC

The question of the stability of an electric arc was the subject of considerable controversy in the past (19) (20). In Section 21 this matter was investigated in detail by means of Liapounoff's equations of the first approximation on the basis of a differential equation with finite parameters L, C, and ρ. The behavior of the arc is far more complex if some of these parameters are small, a condition which results in a degenerate form of the differential equation describing the process.

Using the notation of Section 21, we recall that the differential equations are

$$L\frac{di}{dt} = V - \psi(i); \qquad C\frac{dV}{dt} = \frac{E - V - Ri}{R} \qquad [150.1]$$

where i is the arc current, V is the potential difference across the capacitor, and $\psi(i) = V_a$ is the non-linear characteristic of the arc, see Figure 150.1. As was explained in Section 21, there are either three equilibrium

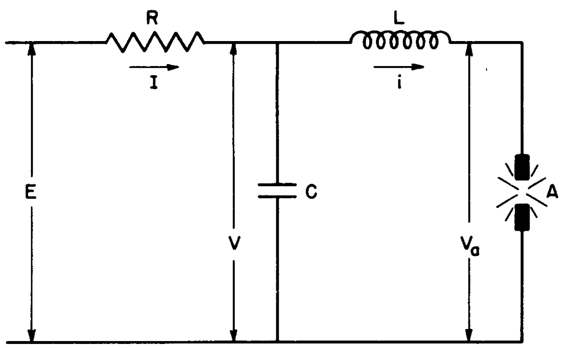


Figure 150.1

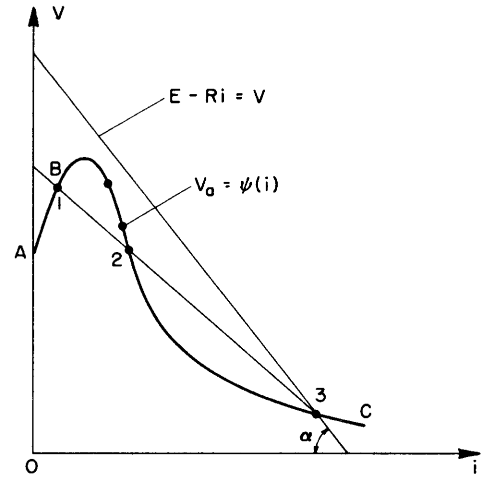


Figure 150.2

points (Points 1, 2, and 3) or one such point (Point 3), as shown in Figure 150.2, depending on the orientation of the straight line $E - Ri$ with respect to the characteristic $\psi(i)$.

It was shown that for small departures, v of voltage and j of arc current, from the corresponding equilibrium values, the equations of the first approximation are

$$\frac{dv}{dt} = -\frac{v}{RC} - \frac{j}{C}; \qquad \frac{dj}{dt} = \frac{v}{L} - j\frac{\rho}{L} \qquad [150.2]$$

where $\rho = \psi'(i_0)$ is the slope of the tangent to the characteristic $\psi(i_0)$ at the point $i = i_0$. The characteristic equation of the system is

$$\lambda^2 + \left(\frac{1}{RC} + \frac{\rho}{L}\right)\lambda + \frac{1}{LC}\left(1 + \frac{\rho}{R}\right) = 0 \qquad [150.3]$$

When there are three points of equilibrium, it is observed that Point 1 corresponds to $\rho = \psi'(i_0) > 0$. This point is always stable because the real parts of the roots of Equation [150.3] are negative. It is, therefore, either a stable nodal, or a stable focal, point according to whether the roots are real or conjugate complex. At Point 2 the equilibrium is unstable because $\rho < 0$ and, moreover, $|\rho| > R$. This point, therefore, is a saddle point. Point 3 is a point of stable equilibrium because, although at this point $\rho < 0$, $|\rho| < R$.

Having recalled these conclusions of Section 21, we propose to investigate now what happens to these various conditions of equilibrium when the system undergoes a degeneration, that is, when one of the two *oscillatory parameters* L and C approaches zero. We may distinguish two cases of degeneration, namely, C-degeneration when $C \to 0$ and L-degeneration when $L \to 0$.

In order to investigate the C-degeneration, it is convenient to write Equation [150.3] as

$$C\lambda^2 + \left(\frac{1}{R} + \frac{C\rho}{L}\right)\lambda + \frac{1}{LR}(R + \rho) = 0 \qquad [150.4]$$

For the investigation of the L-degeneration we will write it as

$$L\lambda^2 + \left(\frac{L}{RC} + \rho\right)\lambda + \frac{1}{CR}(R + \rho) = 0 \qquad [150.5]$$

In this manner we shall be able to use the conclusions of Section 128. Putting $C = 0$ in Equation [150.4], we obtain

$$\lambda = -\frac{R + \rho}{L} \qquad [150.6]$$

Putting $L = 0$ in Equation [150.5], we have

$$\lambda = -\frac{1}{CR\rho}(R + \rho) \qquad [150.7]$$

Let us consider the equilibrium at Point 2 of Figure 150.2. It is observed that in the C-degeneration this point is still unstable. In the L-degeneration, it is stable. A considerable controversy existed on this subject in the technical literature before this question was completely understood. The following explanation was suggested by S. Chaikin (20) (21).

Let us consider the characteristic equation

$$a\lambda^2 + b\lambda + c = 0 \qquad [150.8]$$

with $a > 0$ and $c < 0$, which characterizes a saddle point; see Section 18. We can define saddle points as *positive* when $b > 0$ and as *negative* when $b < 0$. In the case of a-degeneration, a positive saddle point gives rise to a positive root and, hence, to unstable equilibrium. If, however, the saddle point is negative, the degenerate system has stable equilibrium at this point. Thus the a-degeneration reverses the stability at Point 2 of Figure 150.2. Since in both cases Point 2 corresponds to a saddle point of the system, we conclude that the only way to obtain stability under this condition is to have the representative point follow a stable separatrix.* If, however, it follows the stable separatrix, there exists a fixed relation between x and $y = \dot{x}$ in the phase plane. But this is precisely what happens in a degenerate system described by a differential equation of the first order, see Equation [128.3]. The condition of degeneration thus *imposes* this singular trajectory, the only one which is stable in the neighborhood of a saddle point. In this case, therefore, a threshold exists similar to that which exists for the asymptotic motion of a pendulum approaching a point of unstable equilibrium, see Section 4, Case 3; such a case is never obtained in practice since there exist no absolutely degenerate systems, that is, systems in which $a = 0$.

---

* Only in a purely theoretical case when the motion of the representative point takes place along the stable separatrix or asymptote of a saddle point, may the latter be considered as a *stable* singularity. If the actual motion of the representative point occurs in the neighborhood of this theoretical motion, it may appear that the saddle point is a stable singularity, at least for a limited time. In view of this, one could characterize the saddle point as an *almost unstable* singularity.

For practical purposes, if precautions are made to reduce the parasitic parameter to a very low value, the representative point will follow a trajectory very closely to the stable separatrix so that one has the impression that the phenomenon would ultimately "settle" at Point 2 as if it were a point of stable equilibrium. In reality, sooner or later the representative point will depart from this point as a result of accidental disturbances.

By a similar argument one finds that Point 3 of stable equilibrium for a non-degenerate system may become unstable in the case of L-degeneration, as follows from Equation [150.1] in which $\rho < 0$ and $|\rho| < R$.

Point 1 is stable both for the non-degenerate and the degenerate systems. It is thus seen that the assumption of absolute degeneration may lead to conclusions at variance with experimental facts as far as the question of stability is concerned.

151. EFFECT OF PARASITIC PARAMETERS ON STABILITY OF RELAXATION OSCILLATIONS

As another example of the same sort, we shall investigate the behavior of the circuit shown in Figure 135.1, taking into account the presence of the small parasitic inductances L and L' shown in Figure 151.1. Kirchhoff's equations for this circuit are

$$I + i = I_a; \qquad L\frac{dI}{dt} + RI - ri - L_1\frac{di}{dt} - \frac{1}{C}\int i\,dt = 0 \qquad [151.1]$$

Moreover, $I_a = \phi(kri)$ as before. Elimination of I_a and I between these equations results in the equations

$$\frac{di}{dt} = \frac{R\phi(kri) - (R + r)i - V}{L + L' - Lkr\phi'(kri)}; \qquad \frac{dV}{dt} = \frac{i}{C} \qquad [151.2]$$

The conditions of equilibrium are obviously $i = 0$ and $R\phi(0) - V = 0$. On the

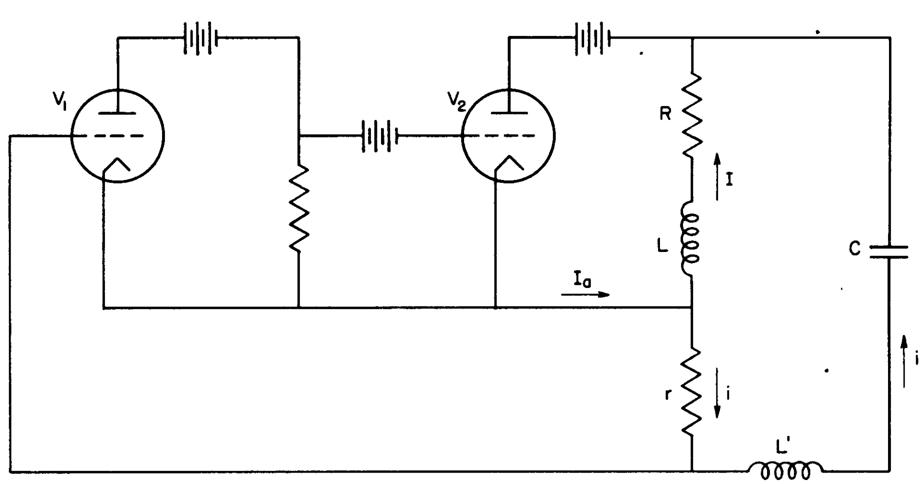


Figure 151.1

other hand, for small departures from equilibrium, we can write

$$\phi(kri) = \phi(0) + kri\phi'(0) + \cdots$$

Limiting the expansion to the first term and noting that $\phi'(0) = S$, the Liapounoff equations of the first approximation are

$$\frac{di}{dt} = -\frac{\delta}{\gamma}i - \frac{V}{\gamma}; \quad \frac{dV}{dt} = \frac{i}{C} \qquad [151.3]$$

where

$$\delta = r + R(1 - Skr) \quad \text{and} \quad \gamma = L' + L(1 - Skr) \qquad [151.4]$$

The characteristic equation of the system [151.3] is

$$\lambda^2 + \frac{\delta}{\gamma}\lambda + \frac{1}{\gamma C} = 0 \qquad [151.5]$$

and its roots are

$$\lambda_{1,2} = \frac{-\delta \pm \sqrt{\delta^2 - \frac{4\gamma}{C}}}{2\gamma} \qquad [151.6]$$

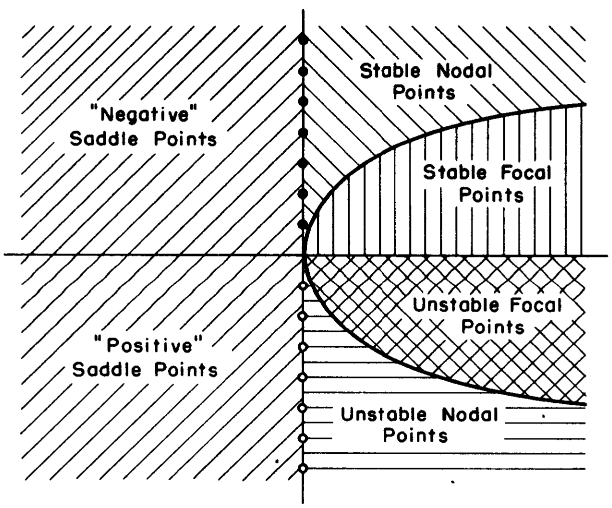


Figure 151.2

For $\gamma < 0$ the roots are real and of opposite sign; hence, the singularity is a saddle point. For $\gamma > 0$, the singularity is either a nodal point or a focal point, depending on whether the roots are real or conjugate complex. If they are conjugate complex, with $\delta > 0$, the singularity is stable; with $\delta < 0$, it is unstable. The distribution of singularities is shown in Figure 151.2, which is the same as Figure 18.1 with the exception that now we make a distinction between "positive" and "negative" saddle points as was explained in Section 150.

For an absolute degeneration, $L = L' = \gamma = 0$. Multiplying the first equation [151.3] by γ, one has $\delta i + V = 0$, that is, $i = -\frac{V}{\delta}$. The second equation gives

$$\frac{dV}{dt} = -\frac{V}{\delta C} \qquad [151.7]$$

The point of equilibrium $V = R\phi(0)$ is the same as in the non-degenerate system. The equilibrium is stable for $\delta > 0$ and unstable for $\delta < 0$. The points of stable equilibrium are indicated by solid points on the positive axis of ordinates and those of unstable equilibrium by circles on the negative axis of ordinates. It is seen that the region of stable equilibrium is localized in the region of "negative" saddle points and that of unstable equilibrium in the region of "positive" ones. It was shown, however, in the preceding section that such stability is of theoretical interest only, because it is impossible in practice to prescribe the initial conditions accurately enough to make the representative point follow the stable separatrix, just as it is impossible to impart to a pendulum a velocity just sufficient to produce an asymptotic motion. Moreover, no physical system is completely degenerate, but is only quasi-degenerate.

The situation is still further complicated by the fact that stability conditions are influenced not only by the existence of parasitic parameters but also by their *relative importance*. In order to show this, assume, for instance, that $L \neq 0$, $L' = 0$, $1 - Skr < 0$, but R is sufficiently small to render $\delta = r + R(1 - Skr)$ positive. In such a case the equilibrium is unstable for any value of L, however small, because the real part of the roots is positive, that is, $\delta > 0$ and $\gamma < 0$. In a degenerate system, where $L = L' = 0$, the equilibrium is stable, however, as follows from Equation [151.7]. Since any physical system is only quasi-degenerate, we conclude that the equilibrium is unstable in this case.

If we consider now the second case, when $L' \neq 0$, the situation is different. Let us assume, as previously, that $1 - Skr < 0$. It is obvious that, by giving L' a suitable value, the quantity $\gamma = L' + L(1 - Skr)$ can be made positive and since $\delta = r + R(1 - Skr)$ is also positive, as previously assumed, it follows from Equation [151.6] that the equilibrium is stable. Hence, the equilibrium which was unstable for $L \neq 0$, $L' = 0$ becomes stable for $L = 0$, $L' \neq 0$. Thus the change from stability to instability depends here on the ratio L/L' of the parasitic parameters.

Since in practice one is seldom sure of the existence of parasitic parameters, much less of their relative values, the predetermination of stability conditions for a circuit of this kind becomes altogether meaningless. It is likely that floating, erratic excitation and similar phenomena still little explored can be explained on this basis.

To sum up the results of this analysis, it can be stated that, in general, the effect of parasitic parameters is negligible in practice unless the system happens to be in the neighborhood of a branch point of equilibrium. If the system is in such a neighborhood, even a small "cause" may exert an appreciable "effect" on the system, and its behavior then becomes entirely unpredictable.

REFERENCES

(1) "On Relaxation Oscillations," by B. Van der Pol, Philosophical Magazine, 7th Series, Vol. 2, November 1926.

(2) "Über Relaxationschwingungen" (On Relaxation Oscillations), by B. Van der Pol, Zeitschrift für Hochfrequenztechnik, Vol. 28, 1926, and Vol. 29, 1927.

(3) "Über Kippschwingungen insbesondere bei Elektronenröhren" (On Sweep Oscillations in Electron Tubes), by E. Friedländer, Archiv für Elektrotechnik, Vol. 17, No. 1, September 1926, and Vol. 18, No. 2, October 1926.

(4) "Reissdiagramme von Senderröhren" (Graphical Diagram for Transmitting Tubes), by H. Rukop, Zeitschrift für Technische Physik, Vol. 5, 1924.

(5) "Eine neue Schaltung zur Erzeugung von Schwingungen mit linearem Spannungsverlauf" (New Wiring Connection for Generating Oscillations with Linear Voltage Curvature), by G. Frühauf, Archiv für Elektrotechnik, Vol. 21, No. 5, February 1929.

(6) "Zur Theorie des rückgekoppelten Röhrensenders" (On the Theory of Regenerative Coupled Vacuum Tube), by F. Kirschstein, Zeitschrift für Hochfrequenztechnik, Vol. 33, No. 6, June 1929.

(7) "Über Schwingungszeugung mittels eines Elektronröhrensystems bei welchem die Kapazität von untergeordneter Bedeutung ist" (On the Excitation of Oscillations by Means of a Vacuum-Tube System in Which Capacity Is of Subordinate Importance), by K. Heegner and I. Watanabe, Zeitschrift für Hochfrequenztechnik, Vol. 34, No. 2, August 1929.

(8) "Über symmetrische Kippschwingungen und ihre Synchronisierung" (On Symmetrical Relaxation Oscillations and Their Synchronization), by H.E. Hollmann, Elektrische Nachrichtentechnik, Vol. 8, No. 10, October 1931.

(9) "On the Theory of the Frühauf Circuit," by L.P. Cholodenko, Journal of Technical Physics, USSR, (Russian), Vol. 5, 1935.

(10) "Les systemes autoentretenues et les oscillations de relaxation" (Self-Excited Systems and Relaxation Oscillations), by Ph. LeCorbeiller, Hermann, Paris, 1931.

(11) "On the Existence of Periodic Solutions of Second Order Differential Equations," by N. Levinson, Journal of Mathematics and Physics, Vol. 22, 1943.

(12) "On Van der Pol's and Related Non-Linear Differential Equations," by J.A. Shohat, Journal of Applied Physics, Vol. 15, 1944.

(13) "Theory of Oscillations," by A. Andronow and S. Chaikin, Moscow, (Russian), 1937.

(14) "Oscillations discontinues" (Discontinuous Oscillations), by S. Chaikin and L. Lochakow, Journal of Technical Physics, USSR, (Russian), Vol. 5, 1935.

(15) "Mesure en valeur absolue des périodes des oscillations électriques de haute fréquence" (Absolute Value of Periods of High-Frequency Electric Oscillations), by H. Abraham and E. Bloch, Annales de Physique, Vol. 12, No. 9, 1919.

(16) "Über Schwingungszeugung mittels eines Elektronröhrensystem welche Selbstinduction nicht erhalten" (On the Excitation of Oscillations by Means of an Electron-Tube System without Self-Induction), by K. Heegner, Zeitschrift für Hochfrequenztechnik, Vol. 29, 1927.

(17) "Zur hydrodynamischen Theorie der Schmiermittelreibung" (On the Hydrodynamic Theory of Lubricating-Oil Friction), by A. Sommerfeld, Zeitschrift für Mathematik und Physik, Vol. 50, 1904.

(18) "Mechanical Relaxation Oscillations," by S. Chaikin and N. Kaidanowsky, Journal of Technical Physics, USSR, (Russian), Vol. 3, 1933.

(19) "Warum kehren sich die für den Lichtbogen gültigen Stabilitätsbedingungen bei Elektronenröhren um?" (Why Do the Conditions of Stability Valid for an Electric Arc Light Reverse in Electron Tubes?), by H. Barkhausen, Physikalische Zeitschrift, Vol. 27, No. 2, February 1926.

(20) "Über Stabilitätsbedingungen und ihre Abhängigkeit von Steuerorganen" (On Stability Conditions as Functions of Control Elements), by E. Friedländer, Physikalische Zeitschrift, Vol. 27, 1926.

(21) "Effect of Small Parameters on the Stationary State of a Dynamical System," by S. Chaikin, Journal of Technical Physics, USSR, (Russian), Vol. 5, 1935.

INDEX

This treatise has been published in four parts. In this index the roman numerals indicate the parts, the arabic numerals indicate pages in the respective parts.