

# Theoretical Aspects of Nonlinear Oscillations\*

N. MINORSKY†

## INTRODUCTION

NONLINEAR problems acquire a gradually increasing importance in various branches of applied science. It is therefore useful to give a brief outline of these questions.

The theory of oscillations is the best-explored branch; here the nonlinear theory was gradually worked out and applied to numerous phenomena, some of which are outlined below. There are other branches into which nonlinear problems began to penetrate in recent years, such as theory of automatic control systems, econometrics, biology, astronomy, atomic theory, etc., but in all these new developments the fundamentals remain very much the same.

On the theoretical side the situation seems to be gradually codified owing to an intense activity going on steadily for the last three decades or so. It must be said, however, that the whole situation is not definitely crystallized, so to speak. There are some topics where mathematical developments are ahead of the experimental evidence; at other points, on the contrary, there exists no definite theory capable of accounting for the observed facts. Finally, there is a domain in which there are two competing theories, and it is impossible to say at present which is "better" since much depends on one's point of view.

However, there is a domain, *the nearly linear domain*, in which the final codification has been apparently reached; that is, both the theory and the experimental evidence proceed in agreement with one another. It seems likely that when the synthesis (*i.e.*, engineering applications) eventually start, the principal developments are to be expected in this domain where everything seems to be clear and definite

For systems with one degree of freedom (electrical or mechanical) these nearly linear problems are amenable to differential equations (DE) of the form

$$\ddot{x} + x + \mu f(x, \dot{x}) = 0 \quad (1)$$

where  $f(x, \dot{x})$  is a nonlinear function of  $x$  and  $\dot{x}$  and  $\mu$  are a small parameter. In such a case (1) differs but little from the linear DE  $\ddot{x} + x = 0$ . However, the fact that the two DE are "near" or "in the neighborhood," or another such condition, does not mean yet that their solutions are also in a corresponding "neighborhood." Thus, for instance, if  $f(x, \dot{x}) = b\dot{x}$ , the two DE  $\ddot{x} + x = 0$  and  $\ddot{x} + \mu b\dot{x} + x = 0$  are in the neighborhood, but their solutions are not if  $t \rightarrow \infty$ .

The investigation of conditions under which (1) has periodic solutions constitutes the *problem of Poincaré* [1] outlined in section 3.

As the nonlinearity is "gauged," so to speak, by the smallness of  $\mu$ , it is customary also to speak about *the method of small parameter* (or parameters).

If  $\mu$  is not small, nonlinear problems become far more complicated. The physical interpretation of the "large parameter" domain is in the so-called *relaxation oscillations*, which we will not investigate here.

In the range of small-parameter problems the mathematical treatment is somewhat different, according to whether the DE is of "autonomous" or "nonautonomous" type. In the former the independent variable  $t$  does not enter explicitly in the DE, whereas in the latter it does. From the standpoint of analytical treatment, there is not much difference between these two types of DE, but the topological approach, so useful in autonomous problems, ceases to be applicable in nonautonomous ones.

Besides the main body of problems amenable to ordinary nonlinear DE, phenomena also appeared whose treatment is amenable to certain functional equations of the difference-differential type (DDE). The theory of these DDE is different, and we merely mention it in reference to publications [2] on the subject. These phenomena appear whenever there are "retarded actions" in physical systems caused by time lags.

A classification of nonlinear problems can be summarized in Table I.

TABLE I

A. Nearly linear problems ( $\mu$ small)	1) Autonomous (self-sustained oscillations) 2) Nonautonomous (subharmonic resonance, synchronization, parametric excitation etc.)
B. Strongly nonlinear problems ( $\mu$ large) (relaxation oscillations)	1) Discontinuous theory 2) Analytical theory
C. Retarded actions; DDE.	

There are still some special problems of a more complicated functional type but these are hardly explored at present.

In problems B-1, use is made of certain idealizations which result in a discontinuous theory resembling that used in the theory of shocks in theoretical mechanics; topological concepts play an important role in that theory

\* Manuscript received by the PGCT, August 3, 1959.

† Aix-en-Provence, France.

which, for that reason, is essentially qualitative and can be used easily in applied problems [3]. The analytical approach to these problems [4] is at present very complicated and can be regarded as an advance in the theory of nonlinear DE rather than as a practical tool in the hands of physicists or engineers.

As regards problems *C*, these oscillatory phenomena begin to play an important role in the theory of automatic control systems where time lags are inevitable; their theory is beyond the reach of ordinary nonlinear DE, and the reader has to consult special treatises on this subject [2] as even a short outline of this matter is impossible in this review.

It is likely however, that problems *A* will continue to be of a primary importance, particularly when the synthesis stage of these studies is ultimately reached.

As regards the contents of this paper, Section I gives a brief summary of existing topological concepts and methods; it is assumed that the reader has already a certain preliminary knowledge of these questions. Section II gives a similar review of principal analytical methods; the last section in this chapter outlines the recently developed stroboscopic method which is used in the calculation of certain nonlinear phenomena in Section III.

## I. TOPOLOGICAL ASPECTS OF THE THEORY

### A. Singular Points

This subject develops from the important paper of Poincaré [5] which can be found now in any text on the theory of DE [6]. We merely mention here some of the important topics and definitions.

The plane of the variables  $x, \dot{x}$  is called *phase plane*, and the study of integral curves (characteristics) is usually conducted in this plane. An integral curve with the direction of motion of the *representative point R* (the instantaneous state of motion) is called *trajectory*.

For a physical system with one degree of freedom of autonomous type, the DE are of the form

$$\dot{x} = P(x, y); \quad \dot{y} = Q(x, y), \quad (2)$$

where  $P$  and  $Q$  are either analytic function of  $x$  and  $y$  or polynomials. A point  $x, y$  for which  $P$  and  $Q$  do not vanish simultaneously is called an *ordinary point*, while a special point  $x_0, y_0$  for which both  $P$  and  $Q$  vanish is called a *singular point*. In the theory of oscillations singular points are identified with *positions of equilibria*. Singular points are characterized by the form of trajectories in their neighborhood, as well as by the direction in which  $R$  moves: if  $R$  moves *toward* a singular point, the latter is stable; if it moves away from it, it is unstable.

One often writes the functions  $P$  and  $Q$  in (2) in the form  $P(x, y) = ax + by + P_2(x, y)$ ;

$$Q(x, y) = cx + dy + Q_2(x, y), \quad (2a)$$

where  $P_2$  and  $Q_2$  are either the entire functions or polynomials whose degree begins with  $x^2, y^2, xy$  or higher. In a great majority of cases encountered in applications, one

can neglect  $P_2$  and  $Q_2$  for the determination of singular points, in which case the problem is simple, and it is shown that the nature of a singular point is determined by the nature of roots of the *characteristic equation*

$$S^2 - (a + d)S + (ad - cd) = 0. \quad (3)$$

If  $S_1$  and  $S_2$  are real and of the same sign, the singular point is a *node*;

if  $S_1$  and  $S_2$  are real and of opposite signs, the singular point is a *saddle point*;

if  $S_1$  and  $S_2$  are conjugate complex, the singular point is a *focus*;

if  $S_1$  and  $S_2$  are purely imaginary, the singular point is a *center*.

The first three singularities are called *simple*; as to the last one, the center, it is already a *special* singular point. In fact, in some cases it is impossible to distinguish a center from a special form of a focus on the basis of the first approximation (*i.e.*, assuming that  $P_2 = Q_2 = 0$ ). The problem then becomes more complicated and requires the study of terms of higher degrees in  $P_2$  and  $Q_2$ .

The simple singularities have a definite physical meaning: a node characterizes the position of equilibrium toward which an aperiodically damped motion approaches; a focus, the equilibrium toward which an oscillatory damped motion approaches; finally, the saddle point always characterizes an unstable motion, for instance that of a pendulum in its upright unstable position of equilibrium. As the representation of these motions and trajectories in the phase plane is likely to be known, we do not reproduce them here. If the word "approaches" in the above definitions is replaced by "moves away," instead of *stable* singular points (node or focus), one has *unstable* ones, the form of trajectories remaining the same in both cases. Trajectories near a node are stable (unstable) if the roots of (3) are negative (positive). Near a focus they are stable (unstable) if the real part of the roots is negative (positive). The saddle point is always unstable. As to the center, its equilibrium is *indifferent* in this classification. Nodes and foci always characterize nonconservative physical systems. Centers appear only in conservative systems under very special conditions. Saddle points occur both in conservative and nonconservative systems.

### B. Limit Cycles

The most important discovery of Poincaré was the establishment of certain closed trajectories which he calls *limit cycles* ("cycles limités;" in the following we shall use often the term *cycle*) having the following properties:

1) There exists one single (isolated) closed trajectory  $C$ , the "cycle" (at least in a finite region of the phase plane), all other trajectories being nonclosed and having the form of spirals  $C'$  winding (unwinding) themselves on the stable (from the unstable) cycle as shown in Figs. 1 and 2, which are self-explanatory. One consequence of this definition is obvious: since any point of the phase plane

represents some initial condition, and through each such point a spiral  $C'$  passes, it is clear the ultimate motion on the cycle (for  $t \rightarrow +\infty$  if the cycle is stable, or for  $t \rightarrow -\infty$ , if it is unstable) *does not depend on the initial conditions*. In this respect the motion on a cycle differs radically from the motion on closed trajectories around the center where, on the contrary, any change in the initial conditions results in a new trajectory, which has no tendency to return to its former path (prior to this change).

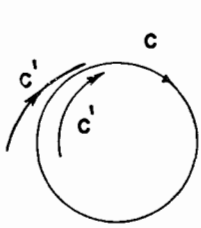


Fig. 1.

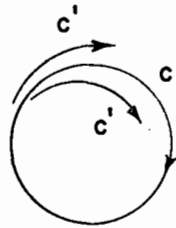


Fig. 2.

Contrary to singular points, which very frequently are determined by the linear terms of functions  $P$  and  $Q$ , the cycles are manifestations of the presence of nonlinear terms in these functions. Unfortunately the matter here is more complicated, for the *form* alone of these functions is not sufficient to determine the presence of cycles, as we shall see later.

C. Topological Configurations

Poincaré has shown that limit cycles and singular points form certain *topological configurations*, that is, that they coexist together. The theorem of Poincaré is sufficiently general and is based on his *theory of indices*, but for the sake of simplicity we use here a simplified statement: *Every limit cycle contains at least one singular point in its interior of stability opposite to that of the cycle.*

Thus an unstable singular point (a focus or a node) is surrounded by a stable cycle (Fig. 3) and vice versa (Fig. 4). We can use the notation IS for the configuration of Fig. 3 and SI for that of Fig. 4, the first letter relating always to the stability of the singular point and the second to the stability of the cycle. From the physical point of view the configuration IS usually designates a process of self-excitation (*e.g.*, of an electron tube circuit) when a trajectory unwinds itself from the unstable singular point (state of rest) and winds itself onto the stable cycle from inside. The configuration SI, on the contrary, designates the disappearance of an oscillatory process.

It is convenient to consider trajectories as "lines of flow" (of a certain "fluid of trajectories") which facilitates the grasp of these phenomena. From that viewpoint the unstable elements of configuration (singular points or cycles) appear always as "sources," and the stable ones, as "sinks." The interpretation of Figs. 3 and 4 is then obvious.

Occasionally one encounters more complicated configurations involving several cycles (of a "concentric"

type), *e.g.*, ISIS (Fig. 5) and SIS (Fig. 6). In the latter case, for instance, as the singular point is stable, the system is not self-excited. If, however, one communicates an impulse SA transferring the representative point beyond the unstable cycle, the system is able to settle on the external stable cycle. Such phenomena are called "hard self-excitation."

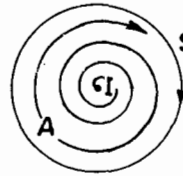


Fig. 3.

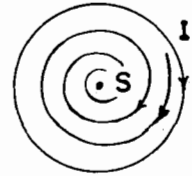


Fig. 4.

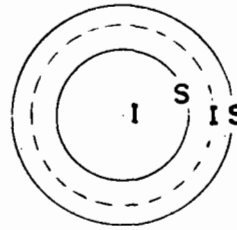


Fig. 5.

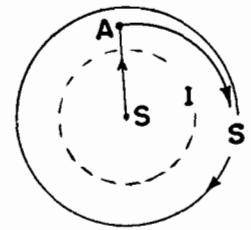


Fig. 6.

D. Bifurcations

In his work on a cosmogonic problem [7] Poincaré investigated the behavior of solutions (trajectories) of a DE containing a parameter  $\lambda$  (not to be confused with the parameter  $\mu$ ). If for a small variation  $\lambda$  around some value the topological configuration remains the same *qualitatively*, such a value is called the *ordinary value* of  $\lambda$ . If, however, for a small change  $\Delta\lambda$  around some *special value*  $\lambda = \lambda_0$ , the qualitative aspect of the configuration changes, such a value is called the *bifurcation value*. This theory found considerable application in oscillation theory. There are two principal types of these phenomena.

1) *Bifurcation of the first kind* occur when a stable (unstable) singular point becomes unstable (stable) with the appearance of a stable (unstable) cycle. In our notations this can be indicated by the schemes

$$S \rightleftharpoons (IS) \rightleftharpoons IS; \quad I \rightleftharpoons (SI) \rightleftharpoons (SI).$$

The letters in brackets indicate the elements in a state of coalescence (*i.e.*, a semi-stable cycle) just before they split into a singular point and a cycle; the double arrows merely show that the phenomenon is reversible.

The first of these schemes is well known in radio technique; that is, a regenerative amplifier circuit in absence of any signal remains at rest, which we indicate here by the symbol  $S$ .

If, however, the parameter  $\lambda$  (the coefficient of coupling between the anode and the grid circuits) begins to increase, the bifurcation point will be reached when the circuit is just on the threshold between the amplification and the oscillation ranges. If this threshold is crossed,

the circuit begins to operate as oscillator on a limit cycle; the configuration is then IS in these notations.

2) *Bifurcations of the second kind* occur in polycyclic configurations when, as the result of parameter variation, two adjoining cycles (one stable and the other unstable), approach each other indefinitely and coalesce for  $\lambda = \lambda_0$ . Beyond this value of the parameter, they destroy one another. The phenomenon is reversible if the parameter varies in the opposite direction. This effect can be represented by the following scheme:

$$\text{ISI} \rightleftharpoons \text{I(SI)} \rightleftharpoons \text{I}; \quad \text{or} \quad \text{SIS} \rightleftharpoons \text{S(IS)} \rightleftharpoons \text{S}.$$

Summing up, a bifurcation of the first kind changes the number of cycles by one unit, while that of the second kind changes it by two units.

### E. Poincaré-Bendixson Theorem

The determination of limit cycles, in general, is a very difficult problem and could be solved only in a few simple cases.<sup>1</sup> The difficulty is due to the fact that it is impossible to ascertain the presence of a cycle from the form of the DE. What permits ascertaining the existence of a cycle is the knowledge of the *topology of the solutions of a DE*, but the latter is generally unknown; one finds oneself in a vicious circle from which there is no easy issue.

Poincaré indicated certain necessary criteria (theory of indices, curves of contact) which were generalized into a theorem and later improved by Bendixson. This criterion is known now as the Poincaré-Bendixson theorem. It gives the necessary and sufficient conditions for the existence of a cycle but its application is not without a difficulty. The theorem states:

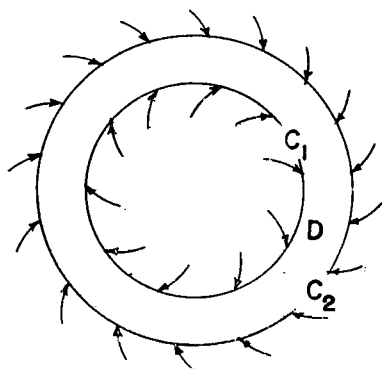


Fig. 7.

If it is possible to determine an annular  $D$  (Fig. 7) limited by two closed curves  $C_1$  and  $C_2$  such that: 1) there are no singular points either in  $D$  or on the boundary curves  $C_1$  and  $C_2$ , and 2) if all trajectories enter (leave)  $D$  through *every point* of  $C_1$  and  $C_2$ , then there exists at least one stable (unstable) limit cycle in  $D$ .

<sup>1</sup> This remark relates to the general theory; it will be shown later that in the theory of approximations, on the contrary, the problem does not present any difficulty.

Although the significance of the theorem is obvious on the basis of the concept of "sources" and "sinks," its application generally is difficult on account of condition 2). In fact, in order to determine correctly the bounding curves, it is necessary to know the behavior of solutions, but these are unknown, as was just mentioned. It is often necessary to apply a great deal of ingenuity in order to construct such a domain. In cases when one strikes a problem in which no such a difficulty exists, the theorem gives an immediate answer.

We shall not go beyond these few remarks, but we shall refer the reader to existing texts [6].

## II. ANALYTICAL ASPECTS OF THE THEORY

### A. Method of Poincaré

The essence of the method is to establish the existence of a periodic solution of a nonlinear DE, which we shall consider here to be in the form (1). We have already mentioned this problem in the Introduction. The solution can be written in the form

$$x = x(t, \mu, K); \quad y = \dot{x} = y(t, \mu, K) \quad (4)$$

where  $K$  is the amplitude. It is useful to introduce two additional parameters

$$\begin{aligned} \beta_1 &= x(0, \mu, K) - x(0, 0, K); \\ \beta_2 &= y(0, \mu, K) - y(0, 0, K), \end{aligned} \quad (5)$$

which determine the difference between the initial conditions of a nonlinear ( $\mu \neq 0$ ) and linear ( $\mu = 0$ ) solutions; it is obvious that  $\beta_1(\mu) \rightarrow 0, \beta_2(\mu) \rightarrow 0$  when  $\mu \rightarrow 0$ .

At this point a difference appears between the autonomous and nonautonomous systems. For the former the period is determined by the DE, whereas for the latter it is fixed by the external periodic excitation (more precisely, by the term containing  $t$  explicitly). There is also another difference: in the autonomous systems one can always replace  $t$  by  $t + t_0$ ,  $t_0$  being a constant without changing the solution. One can therefore select  $t_0$  so as to make one of  $\beta$  equal to zero. Suppose that we make  $\beta_2 = 0$  and set  $\beta_1 = \beta$ . Thus, in an autonomous system the parameters will be  $\mu, \beta$ , and  $\tau$ , the latter being the so-called *correction for period*. In the nonautonomous system (which has no "translation" property), one cannot dispose of arbitrary  $t_0$  and has to keep both  $\beta_1$  and  $\beta_2$ . However,  $\tau$ , the period correction, does not exist, since the period is fixed by the term containing  $t$  explicitly and not by the parameters of the DE. The further treatment of both cases is analogous,<sup>2</sup> so that it will be sufficient to study the autonomous systems which are more important in application, at least at present.

<sup>2</sup> This statement applies to the existence of the solution but not to its stability; regarding stability, while this problem is simple for autonomous systems, for the nonautonomous ones it is far more complicated.

The conditions of periodicity can be written as

$$\begin{aligned} x(2\pi + \tau, \mu, \beta, K) - x(0, \mu, \beta, K) &= \mu\Phi(\tau, \mu, \beta, K) = 0 \\ y(2\pi + \tau, \mu, \beta, K) - y(0, \mu, \beta, K) &= \mu\Psi(\tau, \mu, \beta, K) = 0. \end{aligned} \tag{6}$$

On the right-hand side  $\mu$  is taken as a factor in order to take into account the fact that for  $\mu = 0$  there is always a periodic solution. Hence, if  $\mu \neq 0$  (*i.e.*, in the nonlinear case), periodicity is possible only if

$$\Phi(\tau, \mu, \beta, K) = 0 \quad \text{and} \quad \Psi(\tau, \mu, \beta, K) = 0. \tag{7}$$

The essential part of the theory is the theorem of Poincaré [9], which states that in the case of a DE containing a parameter  $\mu$ , like (1), the solution is analytic in terms of this parameter (or parameters, if there are several). This means that the solution can be represented by an entire series arranged according to the ascending powers of these parameters.

As we have here three parameters, of which  $\mu$  is independent and the other two are  $\beta(\mu)$  and  $\tau(\mu)$  such that  $\alpha(\mu) \rightarrow 0, \tau(\mu) \rightarrow 0$  when  $\mu \rightarrow 0$ , the functions  $\Phi$  and  $\Psi$  are of the form

$$\begin{aligned} \Phi &= \Phi_0 + a\mu + b\tau + c\beta + \dots; \\ \Psi &= \Psi_0 + a_1\mu + b_1\tau + c_1\beta + \dots \end{aligned} \tag{8}$$

For the first approximation, it is sufficient to take only linear terms in (8). For approximations of higher orders it is necessary to take more terms in these series. We shall limit ourselves to the first approximation. In fact, the method becomes clear from the first approximation without complicating it too much; moreover, if  $\mu$  is small, as we assume, higher order approximations hardly add anything of interest but complicate calculations considerably. It is clear that functions  $\beta(\mu)$  and  $\tau(\mu)$  in the first approximation can be taken as  $\tau(\mu) = \alpha\mu$  and  $\beta(\mu) = \gamma\mu$  where  $\alpha$  and  $\gamma$  are unknowns. We have thus a system of two algebraic equations,

$$\begin{aligned} \Phi &= \Phi_0 + \mu(a + b\alpha + c\gamma) = 0 \\ \Psi &= \Psi_0 + \mu(a_1 + b_1\alpha + c_1\gamma) = 0 \end{aligned} \tag{9}$$

which are valid in the first approximation and which must be solved for any arbitrary  $\mu$  (provided it is sufficiently small). It is clear that, in order to determine  $\alpha$  and  $\gamma$  from these equations, the following conditions must be fulfilled:

$$1) \quad \Phi_0 = \Psi_0 = 0 \tag{10a}$$

and

$$2) \quad \begin{vmatrix} b & c \\ b_1 & c_1 \end{vmatrix} \neq 0. \tag{10b}$$

Since

$$b = \frac{\partial\Phi}{\partial\tau}; \quad c = \frac{\partial\Phi}{\partial\beta}; \quad b_1 = \frac{\partial\Psi}{\partial\tau}; \quad \text{and} \quad c_1 = \frac{\partial\Psi}{\partial\beta},$$

the second condition can be written as

$$J = \begin{vmatrix} \frac{\partial\Phi}{\partial\tau} & \frac{\partial\Phi}{\partial\beta} \\ \frac{\partial\Psi}{\partial\tau} & \frac{\partial\Psi}{\partial\beta} \end{vmatrix} \neq 0. \tag{11}$$

Hence, if  $\Phi_0 = \Psi_0 = 0$  and the Jacobian (11) is different from zero, there exists at least one periodic solution of the nonlinear problem, small  $\mu$ . The rest of the calculation is simple but long. One begins by expressing  $x(t)$  and  $y(t) = \dot{x}(t)$  in terms of the series of Poincaré. Then one develops the function  $f(x, \dot{x})$  in (1) around the values  $x_0(t)$  and  $y_0(t)$ , which are the so-called "generating solutions" for  $\mu = 0$ , and substitutes these series into (1). Then one arranges the various terms according to the like powers of  $\mu$ . In this way one obtains finally a system of linear DE, which can be integrated successively and yield the coefficients of the series. These coefficients are functions of  $t$ .

The method of Poincaré is very general and yields successive approximations; in its original form it was established for astronomical calculations. Later on it was adapted [11] for other applied problems, but it is still too complicated because of its generality. Its importance is in that it opened an entirely new avenue of approach to the solution of nonlinear problems.

### B. Methods of van der Pol and of Krylov-Bogoliubov

These two methods are very close to each other. The idea is to take a simple harmonic solution and to "fit" it into the nearly linear equation. For this purpose the constant parameters of the harmonic solution are considered as function of  $t$  subject to the determination by a method similar to the classical method of the "variation of parameters" in the theory of DE.

The whole difference between the two methods<sup>3</sup> is that van der Pol takes the harmonic solution in the form,  $A \sin t + B \cos t$ , while in the Krylov-Bogoliubov method this solution is taken in the form,  $a \cos(\omega t + \phi)$ .

We outline here the Krylov-Bogoliubov method as a little more convenient to use. The DE in this method is considered in the form

$$\ddot{x} + \omega^2 x + \mu f(x, \dot{x}) = 0. \tag{12}$$

Substituting the solution  $x = a \sin(\omega t + \phi)$ , and imposing an additional condition that  $\dot{x}$  should be of the form

$$\dot{x} = a\omega \cos(\omega t + \phi), \tag{13}$$

<sup>3</sup>Historically, the method of van der Pol is earlier (around 1927) than the method of Krylov-Bogoliubov (1937).

one obtains the following DE expressing this condition:

$$\dot{a} \sin(\omega t + \phi) + a\dot{\phi} \cos(\omega t + \phi) = 0. \tag{14}$$

Replacing  $x$ ,  $\dot{x}$  and  $\ddot{x}$  (in which differentiations are carried out with respect to  $a$  and  $\phi$ ), one gets the second DE,

$$\dot{a}\omega \cos \gamma - a\omega\dot{\phi} \sin \gamma + \mu f(a \sin \gamma, a\omega \cos \gamma) = 0;$$

where

$$\gamma = \omega t + \phi. \tag{15}$$

Solving these equations with respect to  $\dot{a}$  and  $\dot{\phi}$ , one finds that the right-hand side contains the small factor  $\mu$ ; this shows that  $a(t)$  and  $\phi(t)$  are slowly varying functions of  $t$ . Hence during the longest period  $T$  of the trigonometric functions (which appear as the result of the development into a trigonometric series), one can assume that  $a$  and  $\phi$  are constant, thus approximating a trigonometric series by the regular Fourier series. If one integrates between 0 and  $T$ , all trigonometric terms drop out, and only the constant terms  $K_0(a)$  and  $P_0(a)$  remain; if one replaces them by their Fourier expressions, one finds the usual form of equations of the first approximation,

$$\begin{aligned} \frac{da}{dt} &= -\frac{\mu}{\omega} \frac{1}{2\pi} \int_0^{2\pi} f(a \sin \gamma, a\omega \cos \gamma) \cos \gamma \, d\gamma = \Phi(a), \\ \frac{d\psi}{dt} &= \omega + \frac{\mu}{a\omega} \frac{1}{2\pi} \int_0^{2\pi} f(a \sin \gamma, \omega \cos \gamma) \sin \gamma \, d\gamma = \Omega(a). \end{aligned} \tag{16}$$

The first equation gives the stationary amplitude  $a_0$  from the equation  $\Phi(a_0) = 0$ , and the second, the nonlinear frequency correction by the second term on the right-hand side of the second equation.

For the first approximation, these two methods are simpler than Poincaré's method, but initially the van der Pol method was not adapted for approximations of higher orders; later Krylov and Bogoliubov [12] developed a method based on an earlier procedure of Lindstedt used in astronomy for approximations of higher orders. Finally, recently Bogoliubov and Mitropolsky [13] generalized this method still further.

C. Stroboscopic Method

This method is applicable to both nonautonomous and autonomous systems and occupies a somewhat intermediate position between topological and analytical methods; it has been worked out in collaboration with M. Schiffer and applied later to various problems, as will appear from the following section.

The method is based on the transformation theory of DE. Its primary purpose is to replace the difficult problem of the determination of a stable periodic solution of a nonautonomous system by the simpler problem of establishing conditions for the existence of a stable singular

point of an auxiliary "stroboscopic" system. We give a brief outline of this method.

Consider a simple problem with a DE:  $\ddot{x} + x = 0$ . Its trajectories are circles concentric with the origin. The representative point  $R$ , starting from some point  $R_0$  on the circle, comes back to this point after the period  $2\pi$ . This defines a transformation effected by the DE during the period  $2\pi$ . Clearly, from the standpoint of the transformation it is immaterial what happens *during* the time interval  $0, 2\pi$ ; the essential point is in the *result* of the transformation. This result could be obtained if, instead of illuminating motion continuously, one illuminated it by stroboscopic flashes occurring once during the time  $2\pi$ . In such a case one would see only the fixed point  $R_0$ .

It is useful to define two planes: 1) the plane ( $\psi$ ) in which one sees the continuous motion and 2) the *stroboscopic plane* ( $\phi$ ), in which one sees only the *stroboscopic image*, a fixed point in this case. One can designate the *identical* transformation just mentioned by a symbol,  $T^n(R_0) \rightarrow_{n \rightarrow \infty} R_0$ , which means that the transformation  $T$  applied  $n$  times to the point  $R_0$  of the phase plane always results in the same point  $R_0$ .

In a slightly more complicated case, when the van der Pol equation acts as an operator of the transformation, one has the transformation  $T^n(r) \rightarrow_{n \rightarrow \infty} r_0 = 2$ , which means that the radius  $r$  approaches the value  $r_0 = 2$  after an infinite number of transformations. In the stroboscopic plane, there will be a discrete sequence of "stroboscopic points,"  $A_0, A_1, A_2, \dots$ , converging to a limit point  $A$  (Fig. 8) both from the inside and the outside of the circle along the radius. This is valid in the first approximation, since, as is known, the period then is  $2\pi$ , so that for the phase the transformation is identical. If one had to take into account the correction for the period, this would cause a slight departure from radial motion. One can consider the limit cycle as a locus of limit points  $A$  if the phase of the flashes changes, since everything is symmetrical with respect to the circle  $r_0 = 2$ . So far this amounts to presenting the known facts in a different manner, but it will be shown that this argument leads to important results.

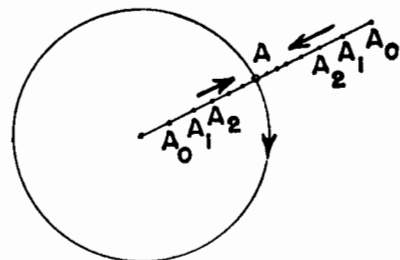


Fig. 8.

It is useful (although not necessary) to introduce two new variables  $\rho$  and  $\psi$ , defined by the relations:  $\rho = r^2 = x^2 + \dot{x}^2$ ;  $\psi = \tan^{-1}(\dot{x}/x)$ , where  $\rho = r^2$ ,  $\dot{x} = y$ , and  $x = r \cos \psi$  and  $y = r \sin \psi$ . Clearly, the variable  $\rho$  is a measure of *total energy* (stored in the oscillation) up to a constant

factor. One writes (1) as an equivalent system of two first-order DE (setting  $y = \dot{x}$ ), and introduces these variables, noting that  $x\dot{x} + y\dot{y} = 1/2 d\rho/dt$  and  $x\dot{y} - y\dot{x} = \rho d\psi/dt$ . It is sufficient to replace  $x$  and  $y$  by their values in order to have equations in the new variables. If one applies this to the DE of harmonic oscillator,  $\ddot{x} + x = 0$ , one finds

$$\frac{d\rho}{dt} = 0; \quad \frac{d\psi}{dt} = -1. \tag{17}$$

The first expression shows that the energy is conserved, and the second indicates that the period is  $2\pi$ ; these are obviously equations of a harmonic oscillator in our new variables.

One ascertains easily that any nearly linear DE of the type (1) can be written in the form

$$\frac{d\rho}{dt} = \mu f(\rho, \psi, t); \quad \frac{d\psi}{dt} = -1 + \mu g(\rho, \psi, t). \tag{18}$$

If one replaces in these DE the series solutions of the form

$$\begin{aligned} \rho(t) &= \rho_0(t) + \mu\rho_1(t) + \dots; \\ \psi(t) &= \psi_0(t) + \mu\psi_1(t) + \dots \end{aligned} \tag{19}$$

one finds that for the zero-order terms (those which do not contain  $\mu$ ) one has

$$\rho_0(t) = \rho_0; \quad \psi_0(t) = \phi_0 - t, \tag{20}$$

where  $\rho_0$  and  $\phi_0$  are the initial conditions. The first-order corrective terms are

$$\begin{aligned} \rho_1(t) &= \int_0^t f(\rho_0, \phi_0 - \sigma, \sigma) d\sigma = K(\rho_0, \phi_0), \\ \psi_1(t) &= \int_0^t g(\rho_0, \phi_0 - \sigma, \sigma) d\sigma = L(\rho_0, \phi_0). \end{aligned} \tag{21}$$

The first-order approximation is then

$$\rho(t) = \rho_0 + \mu\rho_1(t); \quad \psi(t) = \phi_0 - t + \mu\psi_1(t). \tag{22}$$

It is clear that these expressions cannot be used for  $t \rightarrow \infty$  since we have neglected the higher order terms (with  $\mu^2, \mu^3, \dots$ ) and these may give rise to the so-called *secular terms*, which would impair the accuracy of approximation in the long run. Instead of this we apply the following procedure: we let time run from 0 to  $2\pi$  and determine  $\rho(2\pi)$  and  $\psi(2\pi)$  from (22). The values so obtained will be used as the initial conditions for the interval ( $2\pi$  to  $4\pi$ ), and again the time will run for the interval  $2\pi$  and so on. In this way the error owing to neglecting higher order terms will be kept under a certain value and will not accumulate. It is seen, however, that the initial conditions *in each* interval will play the role of a *discrete variable* resulting from the fact that we have replaced the original DE (18) by difference equations, as we are going to show now.

Clearly (22) may be regarded as a transformation,

$$\rho' = \rho + \mu\rho_1; \quad \phi' = \phi + \mu\psi_1, \tag{23}$$

where  $\rho'$  and  $\phi'$  are the terminal values in some interval and  $\rho$  and  $\phi$  the initial values. Setting  $\rho' - \rho = \Delta\rho$ ,  $\phi' - \phi = \Delta\phi$ , and taking into account (21), (23) can be written as

$$\Delta\rho = 2\pi\mu K(\rho, \phi); \quad \Delta\phi = 2\pi\mu L(\rho, \phi). \tag{24}$$

We have taken  $2\pi$  as a factor, as it usually appears in integrations. It is noted that the variable  $t$  has disappeared in integrations (21) and  $K$  and  $L$  are merely some numbers. It is convenient, however, to re-introduce the temporal element by defining

$$2\pi\mu = \Delta\tau, \tag{25}$$

where  $\tau$  is the new independent variable, *the stroboscopic time*. The difference equations can now be written as

$$\frac{\Delta\rho}{\Delta\tau} = K(\rho, \phi); \quad \frac{\Delta\phi}{\Delta\tau} = L(\rho, \phi). \tag{26}$$

These equations permit calculating successively the "stroboscopic points" in the  $\phi$  plane. If the phenomenon lasts long enough in terms of one time interval  $2\pi$ , one can introduce a continuous variable (since  $\Delta\tau \rightarrow d\tau$ ;  $\Delta\rho \rightarrow d\rho$  and  $\Delta\phi \rightarrow d\phi$ ).

At the limit we obtain *the stroboscopic DE*,

$$\frac{d\rho}{d\tau} = K(\rho, \phi); \quad \frac{d\phi}{d\tau} = L(\rho, \phi). \tag{27}$$

It is noted that the stroboscopic system is *autonomous* although the original system (18) is nonautonomous. Hence, the topological argument can be used in connection with (27), whereas it was impossible for (18).

One thus arrives at the theorem:

The existence of a stable singular point of (27) is the criterion for the existence of a stable periodic solution (limit cycle) of (18).

In fact for a singular point one must have

$$K(\rho_0, \phi_0) = L(\rho_0, \phi_0) = 0, \tag{28}$$

which means  $d\rho/d\tau = 0$  and  $d\phi/d\tau = 0$ . In other words, there exists a stable fixed point in the  $\phi$  plane, and this in turn means that the trajectory in the  $\psi$  plane passes always through the same point  $(\rho_0, \psi_0)$ . In other words, the trajectory is fixed in all its points and, since it is periodic with period  $2\pi$ , this is a closed trajectory (the limit cycle).

We omit some further remarks regarding the properties of the stroboscopic system and pass directly to the survey of applied problems.

### III. PRINCIPAL NONLINEAR PHENOMENA

#### A. Parametric Excitation

It has been known for a long time [14] that if a parameter of an oscillating system is varied with double frequency (compared to the free frequency of the system), an oscillation with the free frequency builds itself up in that system. In recent times such experiments were produced

by Mandelstam and Papalexli with an oscillating circuit, and it was ascertained that if the system is linear, the oscillation builds up indefinitely until the circuit is punctured by an excessive voltage. If, however, a nonlinear conductor is inserted, the amplitude of oscillation reaches a finite value and is stable [15].

These phenomena are amenable to the DE of Mathieu [16], which for a linear case has the form

$$\ddot{x} + (1 + a \cos 2t)x = 0, \quad (29)$$

and for the nonlinear case (we refer to the above-mentioned tests) it is

$$\ddot{x} + b\dot{x} + (1 + a \cos 2t)x + cx^3 = 0. \quad (30)$$

We shall treat both cases by the stroboscopic method [17] in the first approximation and will assume that all coefficients are small numbers in order to be in the nearly linear domain. Consider first (29). Its equivalent system is  $\dot{x} = y; \dot{y} = -x - a[\text{Cos } 2t]x$ . Introducing the variables  $\rho$  and  $\psi$  (Section II-C) one obtains the differences,

$$\Delta\rho = -\frac{1}{2}2\pi a\rho_0 \sin 2\phi_0; \quad \Delta\phi = -\frac{1}{4}2\pi a \cos 2\phi_0, \quad (31)$$

of  $\rho$  and  $\psi$  for one period  $2\pi$ . Here  $a$  plays the role of  $\mu$ , so that the stroboscopic time element is  $\Delta\tau = 2\pi a$ . The stroboscopic system is then

$$\frac{d\rho}{d\tau} = -\frac{1}{2}[\sin 2\phi]\rho; \quad \frac{d\phi}{d\tau} = -\frac{1}{4} \cos 2\phi. \quad (32)$$

It is seen that (32) has no singular point, as  $\sin 2\phi_0$  and  $\cos 2\phi_0$  cannot vanish together, but the second equation shows that the phase has an equilibrium point when  $\cos 2\phi_0 = 0$ , that is, when  $\sin 2\phi_0 = \pm 1$ . The choice between these two values is dictated by the stability of  $\phi_0$ . Forming the *variational equation* for the second equation in (32) (*i.e.*, replacing  $\phi$  by  $\phi_0 + \delta\phi$ , where  $\delta\phi$  is a small perturbation), one finds that the phase is stable when  $\phi_0 = 3\pi/4$ , that is, when  $\sin 2\phi_0 = -1$ , and for this value the first equation shows that  $\rho$  increases beyond any limit.

The procedure for the DE (30) remains the same; here it is impossible to eliminate the term  $b\dot{x}$  by the classical transformation of the dependent variable as in the linear case. The stroboscopic system in this case is

$$\frac{d\rho}{d\tau} = -\frac{1}{2}\rho[2B + A \sin 2\phi] = K(\rho, \phi), \quad (33)$$

$$\frac{d\phi}{d\tau} = -\frac{1}{4}[A \cos 2\phi + \frac{3}{2}C\rho] = L(\rho, \phi),$$

where  $A = a/\mu$ ,  $B = b/\mu$ , and  $C = c/\mu$  are the parameter of the series solution. The singular point of (33) exists if

$$\sin 2\phi_0 = -2B/A; \quad \cos 2\phi_0 = -3C\rho/2A. \quad (34)$$

Since  $|\sin 2\phi_0| \leq 1$ , we have the first condition

$$A \geq 2B. \quad (35)$$

Forming the expression  $\sin^2 2\phi_0 + \cos^2 2\phi_0 = 1$ , we obtain

$$\rho_0 = \frac{2}{3C} \sqrt{A^2 - 4B^2}. \quad (36)$$

The quantity  $\rho_0$  is real if the condition (35) is fulfilled. In order to see whether the singular point is stable, it is necessary to form the characteristic (3). From the general theory it follows that  $a = K_\rho$ ,  $d = L_\phi$ ,  $b = K_\phi$ , and  $c = L_\rho$ , where  $K_\rho$ ,  $K_\phi$ ,  $L_\rho$  and  $L_\phi$  are partial derivatives of the functions  $K$  and  $L$  in (33) with respect to  $\rho$  and  $\phi$  at the point  $\rho_0$  and  $\phi_0$ . If one carries out this calculation, (3) in this case, one gets

$$S^2 + BS + \frac{1}{4}(A^2 - 4B^2) = 0, \quad (37)$$

and for a *stable* singular point, it is necessary to have to fulfill again the condition (35), which is thus the necessary and sufficient condition for the existence of amplitude and phase as well as for their stability. This result is again in agreement with the above-mentioned experiments.

In recent years the theory of parametric excitation was further generalized [18], which permits studying a number of phenomena which are similar but of a more complicated nature.

In all problems of parametric excitation the essential feature is that the phenomenon always starts from rest. In other words, the degree of freedom of the parameter variation always *absorbs* the energy which is transferred into the principal degree of freedom where the oscillation builds up. This is because of the fact that the stable phase occurs for  $\phi_0 = 3\pi/4$  for which the amplitude is unstable as we mentioned previously. One can question whether the "inverse parametric effect" is possible. It is clear that it cannot occur *spontaneously* as it is inherent in the unstable phase  $\phi_0 = \pi/4$ . However, if such a phase could be *imposed* artificially by an additional circuit, the inverse phenomenon undoubtedly could be produced. In such a case a bar, or string, oscillating laterally (owing to an external periodic excitation of some kind) could be brought to rest, or its oscillation reduced, by an axial periodic force with the phase  $\phi_0 = \pi/4$ . In such a case, instead of *excitation*, one would have a *parametric damping*.

### B. Subharmonic Resonance

It has been known for a long time that if a nonlinear system is acted on by two periodic forces of frequencies  $\omega_1$  and  $\omega_2$ , the response of the system occurs not only with these frequencies and their harmonics but also with the so called *combination tones* [19]. Thus, for instance, if one impresses on the grid of an electron-tube oscillator with the nonlinear characteristic of the form  $i_a = a_1 v + a_2 v^2 + a_3 v^3$  (where  $i_a$  is the anode current and  $v$  the grid voltage), a voltage of the form  $v = v_0 x (\sin \omega_1 t + \sin \omega_2 t)$ , it is easy to show that the frequencies  $\omega_1$ ,  $\omega_2$ ,  $2\omega_1$ ,  $2\omega_2$ ,  $3\omega_1$ ,  $3\omega_2$ ,  $\omega_1 + \omega_2$ ,  $\omega_1 - \omega_2$ ,  $2\omega_1 + \omega_2$ ,  $2\omega_1 - \omega_2$ ,  $2\omega_2 + \omega_1$ ,  $2\omega_2 - \omega_1$  will appear in the plate current. The first frequencies are the original frequencies and their harmonics, but the other ones are the *combination tones*. Those of the latter whose frequencies are lower than the lowest of the two frequencies  $\omega_1$  and  $\omega_2$  are called *subharmonics*. The combination tones are given by the formula  $\omega = m\omega_1 + n\omega_2$ , where  $m$  and  $n$  are (positive or negative) integers. One readily sees that, if  $\omega_0$  is the free frequency



of the system, and if the condition  $\omega_0 = m\omega_1 + n\omega_2$  is fulfilled, a resonance effect is to be expected in such a system.

On this somewhat intuitive basis, the phenomena of the subharmonic resonance are obvious, once the existence of subharmonic is recognized. If, however, one tries to proceed analytically, one encounters considerable difficulty because, given a differential equation, the determination of a subharmonic as well as its stability generally involves very long calculations. These difficulties very likely account more than anything else for our present relatively limited knowledge of these phenomena, although their physical nature is simple enough.

The first complete theory of the subharmonic resonance was given by Mandelstam and Papalexi [20]. These authors dealt directly with the nonautonomous DE of the resonance. The calculations of stability are rather long, as this involves the calculation of the characteristic exponents, as the variational equations here appear with periodic coefficients.

We prefer to treat this problem by the stroboscopic method which simplifies the problem to some extent.

We consider the DE,

$$\ddot{x} + x + \mu f(x, \dot{x}) = \sin nt. \quad (38)$$

It is noted that the amplitude of the external periodic excitation here is *one*, as this can always be done by a change of the independent variable.

We wish to show that there may still be a periodic solution with frequency *one* in spite of the fact that the frequency of the external excitation is  $n$  (*i.e.*, the subharmonic resonance of the order  $n$ ). If  $\mu = 0$ , the generating solution  $x_0(t)$  is

$$x_0(t) = A \sin t + B \cos t + \frac{1}{1-n^2} \sin nt, \quad (39)$$

and likewise for  $y(t) = \dot{x}(t)$ . If one forms now the conditions of periodicity (6) one finds

$$x_1(t) = -\int_0^t \sin(t-\sigma)f(x_0, y_0) d\sigma; \quad (40)$$

$$y_1(t) = -\int_0^t \cos(t-\sigma)f(x_0, y_0) d\sigma,$$

which results in the first approximation,

$$x(t) = x_0(t) + \mu x_1(t); \quad y(t) = y_0(t) + \mu y_1(t). \quad (41)$$

At this point there is a ramification of the problem into two cases: 1) when the resonance is *exact*, and 2) when the system is in the *neighborhood* of the exact resonance. In the latter case there appears also the phenomenon of synchronization, which we shall investigate later. As the first problem is simpler, we investigate it here. In fact, the difference between the two cases is in the length of calculations only. For the exact resonance, all integrations are between 0 and  $2\pi$ , so that the integrated parts vanish, whereas in the second case the integrations are between 0 and  $2\pi + \epsilon$  where  $\epsilon$  is "the detuning" from the exact

resonance. In order to simplify the problem still further, we investigate the subharmonic resonance of the order 1/2 for the DE,

$$\ddot{\xi} - \mu(\alpha - \beta\xi^2)\xi + \xi = e \sin 2t,$$

which by the change of the dependent variable  $\xi = ex$  reduces to

$$\ddot{x} - \mu(\alpha - \gamma x^2)\dot{x} + x = \sin 2t; \quad \gamma = \beta e^2. \quad (42)$$

The generating solution in this case is

$$x_0(t) = A \sin t + B \cos t - \frac{1}{3} \sin 2t; \quad (43)$$

$$y_0(t) = A \cos t - B \sin t - \frac{2}{3} \cos 2t$$

and the conditions of periodicity (6) are

$$x(2\pi) - x(0) = x_0(2\pi) - x_0(0) + \mu x_1(2\pi) = 0, \quad (44)$$

$$y(2\pi) - y(0) = y_0(2\pi) - y_0(0) + \mu y_1(2\pi) = 0,$$

since  $x_1(0) = y_1(0) = 0$ , as follows from (40). Eqs. (44) are adequate for the determination of the *existence* of the subharmonic but not of its stability. For the latter it is preferable to investigate the *approach* to the subharmonic solution. For this it is useful to consider the case when the right-hand terms in (6), instead of being zero, are certain quantities, say  $\Delta x$  and  $\Delta y$ . It is noted also that  $x(0) = B$  and  $y(0) = A - 2/3$ . Since in the stroboscopic method the ultimate variables are the initial conditions in each interval  $2\pi$ , it is useful to introduce  $A$  and  $B$  as new variables. Calculations at this point are simple but long, as expressions (39) are to be introduced into  $f(x_0, \dot{x}_0)$  and the results so obtained have to be integrated between 0 and  $2\pi$ . The stroboscopic system in this case becomes

$$\begin{aligned} \frac{dB}{d\tau} &= \frac{1}{2}B \left\{ \alpha - \gamma \left[ \frac{1}{4}(A^2 + B^2) + \frac{1}{18} \right] \right\}; \\ \frac{dA}{d\tau} &= \frac{1}{2}A \left\{ \alpha - \gamma \left[ \frac{1}{4}(A^2 + B^2) + \frac{1}{18} \right] \right\}. \end{aligned} \quad (45)$$

The introduction of polar coordinates  $\rho = A^2 + B^2$  and  $\phi = \tan^{-1}(B/A)$  simplifies matters. Forming the combination  $B\dot{B} + A\dot{A} = \frac{1}{2}d\rho/d\tau$ , one gets

$$\frac{d\rho}{d\tau} = \rho \left[ (\alpha - \frac{1}{8}\gamma) - \frac{1}{4}\gamma\rho \right] = R_0(\rho). \quad (46)$$

The second combination is of no interest, as one has simply  $d\phi/d\tau = 0$ . From (46) one obtains an expression for the stationary amplitude

$$\rho_0 = \frac{A}{\beta e^2} \left( \alpha - \frac{1}{18} \beta e^2 \right), \quad (47)$$

since  $\gamma = \beta e^2$ .

Hence the stationary amplitude  $\rho_0 = r_0^2$  can exist only if

$$e < \sqrt{\frac{1}{18} \frac{\alpha}{\beta}}. \quad (48)$$

This is the essential feature of the nonlinear resonance. In ordinary (linear) resonance, the amplitude of oscillation exists for any value of the external periodic excitation;

here it exists only as long as the external excitation is not too large.

Since the stroboscopic method reduces the problem of stability of the periodic solution of the original system to that of the singular point of the stroboscopic system, the condition of stability is

$$R_\rho(\rho_0) < 0, \tag{49}$$

and one finds easily that the condition of stability is always fulfilled in this case.

The problem of determining a stable subharmonic solution in this case was simple because it was possible to introduce the variable  $\rho$ . However, this is not always possible.

Thus, for instance, if one tries to ascertain the existence of the subharmonic resonance of order 1/3 in the case of a Duffing equation,

$$\ddot{\xi} + \alpha\dot{\xi} + c\xi^3 + \xi = e \sin 3t, \tag{50}$$

and proceeds as previously shown, the coordinates of the singular point in the stroboscopic plane are given as real roots common to two algebraic equations,

$$\begin{aligned} \alpha A + \frac{3}{4}\gamma e^2 \left[ B(A^2 + B^2) - \frac{1}{4}AB + \frac{1}{64}B \right] &= 0, \\ -\alpha B + \frac{3}{4}\gamma e^2 \left[ (A - \frac{1}{8})(A^2 + B^2) + \frac{1}{64}B \right] &= 0, \end{aligned} \tag{51}$$

where  $a = \mu\alpha$ ,  $c = \mu\gamma$ . The difficulty is thus in the algebraic part of the problem. It is likely that these long calculations account more than anything else for a relative scarcity of information concerning the quantitative data regarding subharmonic resonances of higher orders.

### C. Synchronization

The discovery of the phenomenon of synchronization (or "entrainment" of frequency) goes back to Huyghens (1629-1695). Being inventor of the mechanism of escapement in clocks, he observed the following fact: two clocks were slightly unsynchronized with one another when hung on a wall but became synchronized when suspended on a thin wooden board. More than two centuries elapsed before similar effects were observed in electric circuits.

The simplest way of observing this phenomenon is to impress on the grid of an electron tube oscillator (in addition to the normal feedback voltage) an extraneous voltage of a different frequency. If  $\omega_0$  is the frequency of the oscillator and  $\omega$  that of the applied voltage, one observes the following effect: if  $\omega$  is sufficiently far away from  $\omega_0$ , there is the usual phenomenon of interference or "beats" of the two frequencies as this happens also in linear systems. If, however,  $\omega$  approaches sufficiently near to  $\omega_0$ , the beats disappear suddenly and there remains only one frequency  $\omega$ . Everything happens as if the free frequency of the oscillator were "entrained" by the external frequency  $\omega$ . This occurs within a certain band ( $-\omega^*$ ,  $+\omega^*$ ) of frequencies around  $\omega_0$ ; if  $\omega$  leaves this interval, the two frequencies separate.

Van der Pol gave a theory for this phenomenon [21] by assuming a solution of the form

$$x(t) = b_1 \sin \omega t + b_2 \cos \omega t. \tag{52}$$

If one inserts this solution into the DE (which is not written here), it is necessary to assume that  $b_1$  and  $b_2$  are slowly varying functions of time in order to satisfy the DE. Then it is possible to reduce the original DE to a system of the form

$$\dot{b}_1 = f(b_1, b_2); \quad \dot{b}_2 = g(b_1, b_2). \tag{53}$$

This is now an autonomous system whose singular point occurs for  $f = g = 0$ ; in such a case  $b_1$  and  $b_2$  become constant and the solution represents the phenomenon of synchronization. From this point Andronov and Witt [22] developed a purely topological theory of synchronization which can be found also in reference [23].

One can also apply the stroboscopic method. It permits approaching this problem from a quantitative point of view, but it does not possess a simple intuitive approach to this question as does the topological method.

We consider again the DE

$$\ddot{x} + (cx^2 - a)\dot{x} + x = e \sin \omega t$$

where  $a$ ,  $c$ , and  $e$  are small numbers (the assumption that  $e$  is small is not essential; it merely simplifies the argument) and  $\omega = 1 + \epsilon$ ,  $\epsilon$  being generally small. In other words, we wish to consider the problem of nonlinear resonance of the first order but not the *exact resonance*, since we wish to consider the synchronization feature of the phenomenon which occurs in the neighborhood of the exact resonance.

If one reproduces the calculations inherent in the stroboscopic transformation with which we are now familiar, one easily reaches the stroboscopic equations which are of the form

$$\begin{aligned} \frac{dr}{d\tau} &= -\frac{1}{\lambda} (a_0 r^3 + a_2 r + a_3); \\ \frac{d\phi}{d\tau} &= -\frac{1}{\lambda_r} (b_0 r^3 + b_2 r + b_3). \end{aligned} \tag{54}$$

It is more convenient in this case to use the variable  $r$  (and not  $\rho$ ); in addition to this, the integrations are to be carried between 0 and  $\lambda = 2\pi/\omega$ , since in the synchronization the frequency is *imposed* by the external periodic excitation.

The coefficients in (54) are certain trigonometric functions of  $\sin(\phi_0 - \lambda/2)$  and  $\cos(\phi_0 - \lambda/2)$  and their multiple arguments, where  $\phi_0$  is the phase angle at the exact resonance. It can easily be calculated (we omit this calculation), and one finds that  $\phi_0 = \pi$ . With this value one can calculate the coefficients for the various values of "detuning." The problem reduces then to one of ascertaining for which values of coefficients in (54) the cubic polynomials have a real common root. Synchronization exists if such a root exists and disappears together with this root.

Summing up, in this problem as in the subharmonic

resonance problem, the stroboscopic method clears up analytical difficulties, but the algebraic difficulties appear at the end in the computational part of the problem.

Unfortunately, this seems to be a common feature of all nonlinear problems when one tries to obtain quantitative results. If these problems are to be cleared up quantitatively, perhaps the use of computing machines would be a proper way of proceeding without too much loss of time.

#### D. Asynchronous Actions

When the frequency of external periodic excitation is outside the zones of synchronization and is in no rational ratio with the frequency of the self-excited oscillation, it is customary to speak about *asynchronous actions*. The term "asynchronous" specifies the lack of any rational ratio between these two frequencies. If the "autoperiodic" (self-excited) frequency is present, there are usual "beats" between the two frequencies; if it is absent, there remains a small "hetero-periodic" (forced) frequency (sometimes it does not exist).

These phenomena are as yet little explored except in two cases: 1) If the heteroperiodic frequency  $\omega$  is very high in comparison with the autoperiodic frequency  $\omega_0$ , the existing self-excited oscillation (with frequency  $\omega_0$ ) is destroyed or "quenched" by the external frequency  $\omega$ . This phenomenon is usually called the *asynchronous quenching*. 2) If the nonlinear characteristic is represented by a polynomial of at least fifth degree (that is, with an additional inflexion point), another nonlinear phenomenon, the *asynchronous excitation*, is occasionally possible. It can be described as follows: in the absence of the external frequency  $\omega$ , the circuit is at rest, but as soon as  $\omega$  is applied, it begins to oscillate *with its own frequency*. The frequency  $\omega$  appears thus as a kind of a trigger action releasing oscillation with frequency  $\omega_0$ . Here the magnitude of  $\omega$  is not of importance (as it is in quenching); what is important here is *the form* of the characteristic without which the effect does not appear.

We shall not enter into the theory of these effects; it is sufficient to say that the quenching phenomenon is easily treated by the stroboscopic method [24] on condition that we take into account that  $\omega$  is large in the final result. In the excitation phenomenon the matter reduces to a simple topological representation if one recalls that for the characteristic of the fifth degree there are two limit cycles, and since normally (without  $\omega$ ) the system is at rest, the configuration is SIS in the notations of Section I. D. If, however, the hetero-periodic frequency  $\omega$  is applied, the nonlinear coupling (since the principle of superposition does not hold here) acts so as to modify the configuration according to the bifurcation of the first kind; SIS  $\rightarrow$  (SI)S  $\rightarrow$  IS which releases the autoperiodic oscillation.

The two phenomena (especially the second one) are relatively little known but seem to be of a considerable applied interest in view of the possibility of influencing one oscillation by another one of an entirely different frequency.

#### E. Piece-Wise Linear Phenomena

In the preceding sections we have followed the *classical* nonlinear theory which requires that the solutions of a DE should not only be continuous but should also have a certain number of continuous derivatives. The question of nonlinearity was thus limited to an implicit requirement of the *analyticity* of the solution; the latter was supposed to be an analytic curve (not only continuous but possessing continuous tangents, etc.). Under these assumptions, the existence of self-excited oscillations becomes possible only in nonlinear systems as defined by the DE (2a), that is, only when  $P_2$  and  $Q_2$  in (2a) are different from zero; these oscillations are thus impossible in linear systems.

The above requirements are realized in the classical theory, but a number of experimental facts have shown that weaker conditions are sufficient for the existence of self-excited oscillations. In fact, it has been known for some time that such oscillations appear also in perfectly linear systems whose solutions exhibit certain points at which the analyticity is lost (but, generally, not the continuity); this usually indicates the presence of impulsive actions. Such phenomena are of a frequent occurrence in modern control systems where a number of secondary effects such as dry friction, backlash, hysteresis, relay actions, oversaturation, etc. may occasionally prepare a favorable ground for their appearance in otherwise perfect linear systems.

As it is impossible to go into this matter here we shall indicate an example [3] which will clarify the physical nature of these phenomena.

Consider a discontinuous characteristic  $B'A'OAB$  (Fig. 9) which may be regarded as a limit of functions 1, 2, 3,  $\dots$ , which approach the broken line  $B'A'OAB$ . Physically this may be the electron tube characteristic for a gradually increasing grid voltage (with a corresponding contraction of the  $x$  scale), or a relay switching "on" and "off" a certain circuit, etc.

Whatever the nature of this discontinuous action may be, what is essential is that a certain "step function" is so introduced.

We consider a simple linear DE with a positive damping,

$$\ddot{x} + 2h\dot{x} + \omega_0^2 x = 0, \quad (55)$$

and require that during the time when  $\dot{x} < 0$ , this DE does not have any external action, but if  $\dot{x} > 0$  this DE acquires a constant right-hand term, say  $\omega_0^2$  representing a kind of a "step function," due to "on" and "off" conditions.

We have thus a phenomenon governed by an alternate sequence of two DE which have exactly the same left-hand sides but differ only by the right-hand sides, and this replacement of one DE by the other takes place when  $\dot{x} = 0$ . We have thus

$$\ddot{x} + 2h\dot{x} + \omega_0^2 x = 0 \quad \text{for } \dot{x} < 0$$

and

$$\ddot{x} + 2h\dot{x} + \omega_0^2 x = \omega_0^2 \quad \text{for } \dot{x} > 0. \quad (56a)$$

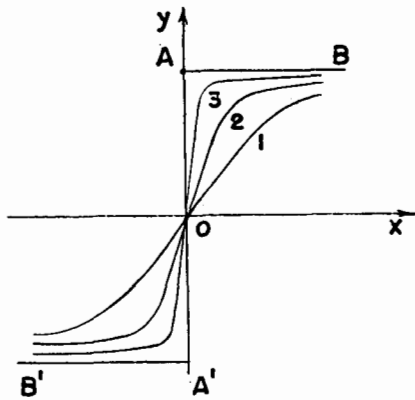


Fig. 9.

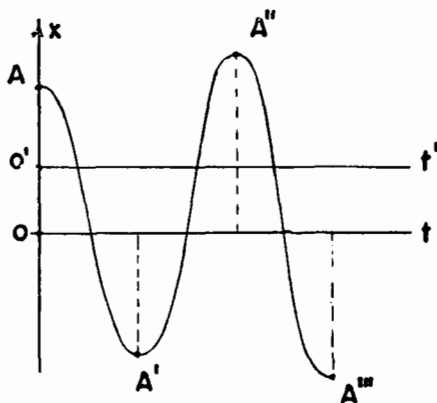


Fig. 10.

Clearly, the second equation can be written as  $\ddot{x}' + 2h\dot{x}' + \omega_0^2 x' = 0$ , if we set  $x' = x - 1$ . This means that in the  $(t, x)$  diagram (Fig. 10) we have to use the regular  $Ot$  axis when  $\dot{x} < 0$  and use the "displaced" axis  $O't'$  when  $\dot{x} > 0$ .

Suppose we start ( $t = 0$ ) at the point  $A$ . As  $\dot{x} < 0$ , from  $A$  to  $A'$  we use the  $t$  axis; at the point  $A'$  we change from  $t$  axis to  $t'$  axis. It is noted that, although with respect to the  $t$  axis the point  $A'$  is nearer to that axis (because of the dissipation of energy),  $A''$  may be more distant from the  $t$  axis than  $A'$  because from  $A'$  to  $A''$  intervenes the displaced axis  $t'$ . In other words, if the constants are properly chosen, what is lost in the dissipation can be made smaller than what is gained, owing to the displacement of the abscissa axis when  $\dot{x} > 0$ . A scheme of this kind may exhibit the feature of self-excitation (its amplitudes grow) in spite of the presence of the positive damping.

It is necessary to ascertain whether a stationary state is possible [3]. Let the ordinates corresponding to the points  $A, A', A'', \dots$  be  $x_1, x'_1, x''_1, \dots$ . During the motion from  $A$  to  $A'$  we have the relation

$$x'_1 = x_1 \exp(-d/2), \tag{56}$$

where  $d = hT = 2\pi h \sqrt{\omega_0^2 - h^2}$  is the logarithmic decrement. For motion between  $A'$  and  $A''$  (when the

"displaced axis" operates), we have

$$x''_1 - 1 = (x'_1 + 1) \exp(-d/2). \tag{57}$$

Eliminating  $x'_1$  between these two equations, we have

$$x''_1 = 1 + \exp(-d/2) + x_1 \exp(-d). \tag{58}$$

This equation determines the transformation we are interested in; if one knows the amplitude  $x_1$ , it is possible to obtain the following amplitude  $x''_1$ .

Clearly, if the trajectory is closed,  $x_1 = x''_1 = \bar{x}$ , and we have thus the following condition for the existence of a limit cycle:

$$\bar{x} = \frac{1}{1 - \exp(-d/2)} > 1. \tag{59}$$

Instead of this argument one can use a more direct topological argument. In fact, each of the two DE's (56a) and (56b) can be represented in the phase plane by spiral trajectories but referred to two different local points displaced along the  $x$  axis. In this manner, the ultimate closed curve (which represents the periodic motion) consists of two "pieces" (arcs) of logarithmic spirals which are "fitted together" on the  $x$  axis.

Along the spiral arcs (the analytic region) the system is dissipative as  $h > 0$  but, since the stationary motion is periodic, energy (on the spiral arcs) is compensated for by the impulsive inputs of energy occurring of the non-analytic points (the spirals of which are "fitted together").

A system of this kind, from the standpoint of energy relations, behaves as a clock where a similar situation exists owing to the impulsive replenishing of energy on the part of the escapment mechanism.

It is to be noted that we have here a closed trajectory, a limit cycle, but that this cycle is composed of two analytic arcs joined to each other at the points  $A, A', A'', \dots$ . At these junction points, both the continuity of the solution and of its first derivative are preserved, but the discontinuity takes place in the second derivative (radius of the curvature). This is sufficient already to render the conclusions of the classical theory inapplicable. In fact we have here self-excited oscillations in a purely linear system, which at a first glance contradicts all that has been said before. In reality this is not so, as in the analytic theory we have assumed that the integral curves are also analytic. Here we have come across phenomena which show that this does not occur here, and one should not wonder that the conclusions are totally different.

These physical considerations (of Andronov) were later generalized in the form of the *point transformation method* [25], which we will outline briefly.

Assume that there are in the phase plane a certain number of arcs of analytic trajectories,  $A_1, A_2, \dots, A_n$ , traversed by the representative point  $R$ . Each of these "pieces" (arcs) is connected to the others, but at the points of junctions the analyticity is lost (e.g., points with two distinct tangents). The point  $R$  traverses this polygone of arcs passing from the beginning, say  $a'_1$  of arc  $A_1$  to

its end,  $a_1''$ ; then passing to the next arc  $A_2$  (obviously:  $a_1'' = a_2'$ ) etc.

Clearly, a periodic motion is possible if this polygone is closed which requires that  $a_1'$  should coincide with  $a_n''$ . This amounts to the existence of the fixed point of the transformation:  $T = T_{A_1} T_{A_2} \cdots T_{A_n}$  where  $T_{A_i}$  are partial transformations, transferring  $R$  from the beginning  $a_i'$  of  $A_i$  to its end  $a_i''$ . Each such partial transformation can be easily calculated as the motion of  $R$  on each arc is governed by a simple linear DE.

In spite of the apparent simplicity of the method, its application is not simple, as the relation connecting the initial point  $a_i'$  of an arc  $A_i$  with its terminal point  $a_i''$  introduces inevitably transcendental functions (the exponential as well as the trigonometric ones).

Hence, if one writes all these conditions, adding also the condition of re-entrant path, one has to solve a number of transcendental equations which is a complicated problem.

In a simple case of the fourth order (amenable to a system of two DE's of the second order whose solutions are joined nonanalytically so as to obtain a fixed point of the transformation), one can carry out the calculations and discuss the results [25].

This method has been recently applied to the theory of nonlinear servomechanisms possessing nonanalytic features of the piece-wise linear type [26].

It must be noted that this approach is entirely outside the scope of the nonlinear theory of DE with the analytic singularities which has been outlined in the preceding sections, and as was mentioned, there are yet considerable difficulties to be overcome before it can be used in more complicated problems.

As we have already mentioned, these phenomena are of a frequent occurrence in the control theory, and this is the reason why the study of self-excited oscillations in that theory presents itself differently.

#### F. Conclusions

We have attempted to condense in this review the salient points of the modern theory of oscillations. Two methods, or groups of methods, dominate the whole situation: 1) topological (or qualitative) and 2) analytical (or quantitative), but it must be recalled that they apply to a rather restricted class of problems from the point of view of the theory of DE, namely those which appear with a parameter. The well-explored field is still more restricted and lies in the domain of small parameters. In this domain there are again two subdivisions: the autonomous problems and the nonautonomous ones. The former appeared first, and here both major lines of attack, 1) and 2), are available; in the latter only 2) is available. The principal practical difficulty here is the question of stability. The DE's in this case have periodic coefficients. This fact raises the very difficult problem of determining the characteristic exponents. The development of the stroboscopic method was due to a large extent to a desire to obviate this difficulty.

If one tries to go beyond this domain of small parameters, one encounters considerable difficulties which have not been completely overcome. This applies to the domain "of large parameters," that is, the so-called *relaxation oscillations* and also to the piece-wise linear (or, more generally, piece-wise analytic) problems which we mentioned in the preceding section. The two problems are to some extent related to one another if one approaches the relaxation problems on the basis of an idealized discontinuous treatment. The theoretical approach to these problems is still in a very early stage. There is, finally, still another approach on the basis of certain functional equations of the so-called difference-differential type which appears each time when certain *constant* time lags appear in the chain of cause-effect actions.

If, instead of going more or less into the unknown, or at least into something which is as yet little known, we wish to consider the well-explored domain of small parameters, the situation presents itself as follows. In recent years the theory of approximations, the so-called asymptotic methods [13], has reached a stage when the solution of a nearly-linear problem can be determined with any accuracy one wishes at the cost of the length of calculations.

In applied problems one comes across a purely practical question: How far is it necessary to go in successive approximations in view of the fact that in general the accuracy with which the nonlinear data are known is very limited? In most cases the answer to this question is that, usually the first approximation is amply sufficient in engineering problems. In fact, the first approximation generally reveals all essential features of a nonlinear problem such as whether limit cycles are present or not, whether there are any bifurcation points, how the phenomenon behaves near the state of rest, etc. Already the second approximation increases considerably the computational work without adding anything of interest to the *qualitative* side of the problem; it adds merely a very small *quantitative* correction to what has been yielded by the theory of the first approximation. For these reasons the theory of first approximation is very important in applied problems. There are cases when the theory of the first approximation may fail and the higher order approximations become necessary, but such situations are very rare in commonly-encountered nonlinear problems.

If we wish to consider now the physical aspects of the theory, one must admit that our ability to produce the *predetermined nonlinearities* is yet very limited. We know, for instance, that in order to produce a "soft" characteristic, one has to use a cubic law in its polynomial approximation; to produce a "hard" characteristic (with one additional cycle), this polynomial must be of the fifth degree, etc. If one wishes not only to produce the phenomenon but to "gauge" it sufficiently accurately (DE to establish definite relations between the limit cycles, etc.) one has to know much more: For instance, what is the relation between the coefficient of the fifth degree to that of the third degree? One has to admit that such a

synthetic production of predetermined characteristics is still beyond our reach. We are still in the *qualitative* phase of our knowledge of these numerous phenomena and are unable to verify theoretical predictions except in some isolated cases such as was the case of the Mandelstam-Papalexi experiments.

Paraphrasing the words of Leonardo da Vinci, "mechanics is a paradise for a mathematician," one can perhaps say that the nonlinear field also presents a kind of a "paradise" of potential possibilities, once the theoretical knowledge of these numerous phenomena is supplemented by adequate means for their experimental realization.

#### BIBLIOGRAPHY

*Note:* Complete lists of references can be found in the books mentioned in [6]. Other references given here concern topics mentioned in this review.

- [1] H. Poincaré, "Les méthodes nouvelles de la mécanique céleste," Gauthier Villars, Paris, France, vol. 1; 1892.
- [2] R. E. Bellman, *et al.*, "Differential Difference Equations," Project Rand R-133, RAND Corp., Santa Monica, Calif.; 1958.
- [3] E. Pinney, "Ordinary Difference-Differential Equations," University of California Press, Berkeley; 1958.
- [4] A. Andronov and S. Chaikin, "Theory of Oscillations," Moscow, USSR; 1937. English translation by S. Lefschetz, Princeton University Press, Princeton, N. J.; 1949.
- [5] M. L. Cartwright and J. E. Littlewood, "On nonlinear differential equations of second order," *ANN. Math.*, vol. 48; 1947.
- [6] H. Poincaré, "Sur les courbes définies par une equation differentielle," *J. Math.*, vol. 3, p. 7; 1881.
- [7] L. Bieberbach, "Theorie der Differentialgleichungen," Dover Publications, Inc., New York, N. Y.; 1944.
- [8] W. Hurewicz, "Ordinary Differential Equations," Brown University Press, Providence, R. I.; 1943.
- [9] S. Lefschetz, "Differential Equations, Geometric Theory," Interscience Press, New York, N. Y.; 1957.
- [10] N. Levinson and E. A. Coddington, "Theory of Differential Equations," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.
- [11] G. Sansone and R. Conti, "Equazioni differenziali non-lineari," Roma, Italy, 1956.
- [12] Y. H. Ku, "Analysis and control of nonlinear systems," Ronald Press, New York, N. Y.; 1958.
- [13] H. Poincaré, "Figures d'équilibre d'une masse fluide," *Acta Math.*, Paris, France; 1903.
- [14] H. Poincaré, *op. cit.*, [6].
- [15] I. Bendixson, "Sur les courbes définies par des equations differentielles," *Acta Math.*, vol. 24, 1901.
- [16] Poincaré, *op. cit.*, [1], p. 58.
- [17] Poincaré, *op. cit.*, [1]. See also [3].
- [18] I. G. Malkin, "Problems of the theory of nonlinear oscillations," Moscow Press, Moscow, USSR; 1956.
- [19] N. Krylov and N. Bogoliubov, "Introduction to Nonlinear Mechanics," *Acad. Sci.*, Kiev, USSR; 1937.
- [20] N. Bogoliubov and N. Mitropolsky, "Asymptotic methods," Moscow Press, Moscow, USSR; 1958.
- [21] F. Melde, "Ueber Erregung Stehender Wellen," *Pogg. Ann.*, vol. 109; 1860.
- [22] Lord Rayleigh, "On maintained vibrations," *Phil. Mag.*, vol. 15; April, 1883.
- [23] M. Brillouin, "Théorie d'un alternateur auto-exciteur," *Ecl. Electr.*; April, 1897.
- [24] H. Poincaré, "Sur quelques theoremes généraux de l'électrotechnique," *J. Math.*; March, 1907.
- [25] L. Mandelstam and N. Papalexi, "Expose des recherches recentes sur les oscillations non linéaires," *J. Tech. Phys. USSR*; 1934.
- [26] G. Floquet, "Sur les equations differentielles lineaires à coefficients periodiques," *Ann. Ec. Norm.*, Suppl. No. 12; 1883.
- [27] B. van der Pol and M. J. O. Strutt, "On the stability of solutions of Mathieu's equation," *Phil. Mag.*, vol. 5; 1928.
- [28] N. Minorsky, *J. Franklin Inst.*, p. 240; 1945.
- [29] N. Minorsky, "Colloque Intern," Porquerolles, France; 1951.
- [30] H. Helmholtz, "Sensations of tone," Longmans, Green and Co., New York, N. Y.; 1895.
- [31] L. Mandelstam and N. Papalexi, *Zeit. Phys.*, vol. 73; 1932.
- [32] B. van der Pol, "Forced oscillations in a circuit with nonlinear resistance," *Phil. Mag.*, vol. 3; 1927.
- [33] A. Andronov and A. Witt, "Zur theorie der Mitnehmens van der Pol," *Arch. Electrotech.*, vol. 24; 1930.
- [34] N. Minorsky, "Introduction to Nonlinear Mechanics," J. W. Edwards, Ann Arbor, Mich.; 1947.
- [35] J. J. Stoker, "Nonlinear Vibrations," Interscience Press, New York, N. Y.; 1950.
- [36] N. Minorsky, *Comp. rend. Akad. Sci.*, vol. 248, p. 631; February, 1959.
- [37] A. Andronov, A. Witt, and S. Chaikin, "Theory of oscillations," Moscow Press, Moscow, Russia; 1959.
- [38] N. Gorskaya, I. Kroutova and V. Routkovsky, "Dynamics of nonlinear servomechanisms," Moscow Press, Moscow, USSR; 1959.