OSCILLATIONS OF A SYSTEM WITH A RELAY OF ADVANCING CHARACTERISTICS

(KOLEBANIIA SISTEMY, SODERZHASHCHEI RELE S OPEREZHAIUSHCHEI KHARAKTERISTIKOI)

PMM Vol.22, No.1, 1958, pp.50-66.

I.B. CHELPANOV
(Leningrad)

(Received 15 June 1957)

The subject of the paper is an investigation of the dynamic properties of a system with one degree of freedom, whose behavior is described by the differential equation

$$\frac{d^2u}{dt^2} + L\frac{du}{dt} + Mu = y \qquad (y = -F(u)) \tag{0.1}$$

The function $F(u)$ is the characteristics of a relay with a zone of non-sensitivity; it is double-valued at some values of the argument u. Fig.1 shows this function in idealized form. When u increases starting from zero, the relay will be connected at $u = a$. When u begins to decrease after having reached its maximum value u_{max}, it will be disconnected at $u = a$, if $u_{max} < b$, and at $u = b$, if $u_{max} > b$. In the latter case the advancing effect takes place. In the following we shall consider the real characteristics (Fig.2) instead of the

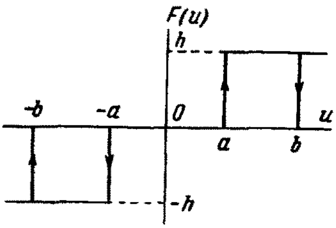

Fig. 1.

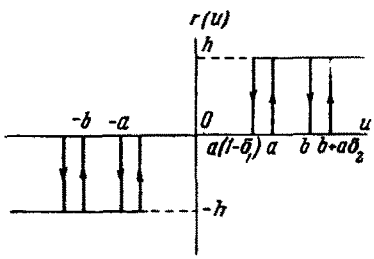

Fig. 2.

idealized one. Here the fact is taken into account, that the disconnection of the relay will actually occur at the dropping of the argument below $u_{max} < b$, not at $u = a$, but at a somewhat smaller value $U = a(1 - \delta_1)$. Analogously, if u starts increasing after having reached the minimum value $u_{min} > a$, the relay becomes connected at $u = b + a\delta_2$, and not at $u = b$.

Reprint Order No. PMM 5.

Depending upon the properties of the system without a relay, the latter may play different roles: (1) If the system by itself is unstable ($L < 0$ or $M < 0$), then the relay may secure the stability of motion at initial deviations, not exceeding definite limits; (2) If the system by itself is stable, the relay may considerably improve its dynamic properties.

The first problem is discussed in the Sections 2 to 6, the second problem in Section 7.

1. Transformation formulas

In the following we consider, instead of the original equation (0.1), the more general non-linear equation

$$\frac{d^2u}{dt^2} + g\left(u, \frac{du}{dt}\right) + f(u) + F(u) = 0 \tag{1.1}$$

We assume that the functions g and f are subjected to restricting conditions of a very general character only, namely:

1. The functions g and f are bounded in any bounded domain;
2. The function f is odd, i.e. $f(-u) = -f(u)$;
3. The function g is even with respect to u and odd with respect to du/dt.

The assumptions 2 and 3 are not fundamental; they are being used only for the purpose of simplifying the subsequent derivations.

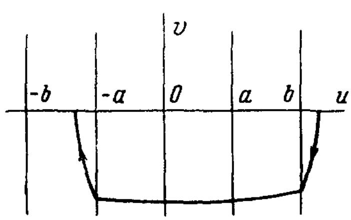

Fig. 3.

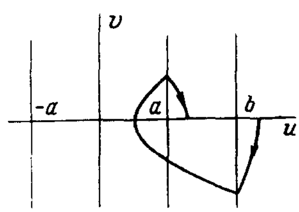

Fig. 4.

We shall consider the passage of a representive point in the phase plane from a starting position on the u-axis to its final position on the same axis. Depending on the functions g and f and on the initial conditions, the consecutive points of intersection of the phase curve with the u-axis can be separated either by a single relay reaction (connection only or disconnection only), or by a double reaction(disconnection followed by connection). Changes of position, which correspond to single relay reactions, will be considered in pairs, in order to obtain combinations consisting of disconnection and connection. Since the functions g and f are assumed to be even, it is permissible to disregard the question, whether the transformation starts from the positive or the negative semi-axis. There are altogether five transformation

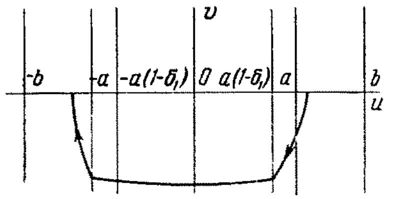

Fig. 5.

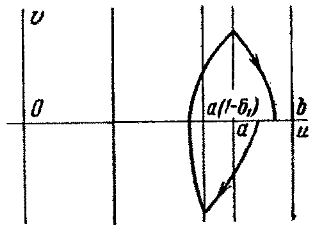

Fig. 6.

possibilities:

1. Transformation ba from one semi-axis to the other. The relay is disconnected at $u = b$ and connected at $u = -a$ (Fig. 3);
2. Transformation $b_1 a_1$ from a semi-axis to itself. The relay is disconnected at $u = b$ and connected at $u = a$ (Fig. 4);
3. Transformation aa from one semi-axis to the other. The relay is disconnected at $u = a(1 - \delta_1)$ and connected at $u = -a$ (Fig. 5);
4. Transformation $a_1 a_1$ to the same semi-axis. The relay is disconnected at $u = a(1 - \delta_1)$ and connected at $u = a$ (Fig. 6);
5. Transformation bb to the same semi-axis. The relay is disconnected at $u = b$ and connected at $u = b + a\delta_2$ (Fig. 7).

The absolute values of the coordinates of consecutive points of the intersection of the phase trajectory with the u-axis will be denoted by $A_1, \ldots, A_n$, respectively.

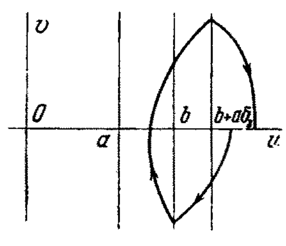

Fig. 7.

In order to clarify the nature of the relationship between any two consecutive values A_i and A_{i+1}, we formally integrate the equation (1.1), which leads to the desired formula

$$\Phi(A_{i+1}) - \Phi(A_i) = -\int\limits_{i}^{(i+1)} \left[g\left(u, \frac{du}{dt}\right) + F(u) \right] du$$

$$\left(\Phi(A) = \int\limits_0^A f(A)\, dA \right) \tag{1.2}$$

To fix the ideas we assume that A_i corresponds to a point on the positive portion of the u-axis. Then, if the transformation ba takes place, the function F assumes the following values:

$$F(u) = \begin{cases} +h \text{ for } b < u < A_i \\ 0 \text{ for } -a < u < b \\ -h \text{ for } -A_{i+1} < u < a \end{cases} \tag{1.3}$$

Therefore we will have

$$\int\limits_{(i)}^{(i+1)} F(u)\, du = h(b-a) + h(A_{i+1} - A_i)$$

For the transformation b_1a_1 the integral has the same value.

For the transformation bb

$$\int\limits_{(i)}^{(i+1)} F(u)\, du = -ha\delta_2 + h(A_{i+1} - A_i)$$

For the transformations aa and a_1a_1

$$\int\limits_{(i)}^{(i+1)} F(u)\, du = -ha\delta_1 + h(A_{i+1} - A_i)$$

Thus we obtain three formulas.

For the transformation ba or b_1a_1

$$[hA_{i+1} + \Phi(A_{i+1})] - [hA_i + \Phi(A_i)] = -h(b-a) - \int\limits_{(i)}^{(i+1)} g\left(u, \frac{du}{dt}\right) du \tag{1.4}$$

For the transformation aa or a_1a_1

$$[hA_{i+1} + \Phi(A_{i+1})] - [hA_i + \Phi(A_i)] = ha\delta_1 - \int\limits_{(i)}^{(i+1)} g\left(u, \frac{du}{dt}\right) du \tag{1.5}$$

For the transformation *bb*

$$[hA_{i+1} + \Phi(A_{i+1})] - [hA_i + \Phi(A_i)] = ha\delta_2 - \int_{(i)}^{(i+1)} g\left(u, \frac{du}{dt}\right) du \quad (1.6)$$

Analogously it is possible to link the values of A_i and A_{i+k}, separated by k transformations of any kind in any order, with each other. The resulting relation has the form

$$[hA_{i+k} + \Phi(A_{i+k})] - [hA_i + \Phi(A_i)] = R + \Psi \quad (1.7)$$

where R is a constant, depending on the number of transformations of each type, while

$$\Psi = -\left\{ \int_{(i)}^{(i+1)} g\left(u, \frac{du}{dt}\right) du + \int_{(i+1)}^{(i+2)} g\left(u, \frac{du}{dt}\right) du + \ldots + \int_{(i+k-1)}^{(i+k)} g\left(u, \frac{du}{dt}\right) du \right\} \quad (1.8)$$

All formulas give only formal relations between the A-values, since the path of integration in the phase plane is not known.

2. Stability of motion.

We will show that upon the choice of a sufficiently large value for the parameter h of the relay characteristics, we can obtain the following general picture for the motion of a point in the phase plane (Fig. 8):

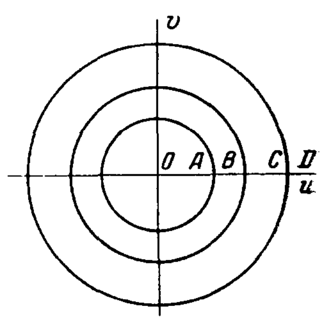

Fig. 8.

(1) If the representative point was originally in the domain C, it will be located in the domain A after further motion;

(2) If the representative point was originally in the domain A, it will not move beyond the limits of the domain B.

A domain D of divergent motions may or may not exist depending on the properties of the functions g and f. By suitable choice of the parameters of the relay characteristics, the boundary of the domains C and D

(if the domain D exists) can be removed from the origin of the system of coordinates as far as desired, and the boundary of the domains B and C can be moved toward the origin of the coordinate system.

Let us prescribe the following conditions for h:

1. The function $[hA + \Phi(A)]$ shall be positive and non-decreasing when $d < |u| < N$, where N corresponds to the boundary of the domains D and C (it can be prescribed arbitrarily), while d is the minimum value of A, corresponding to the transformation ba. The meaning of the condition is that the relay shall compensate for the static instability of the system in the given domain, if such an instability exists.

2. Denote by g_0 the maximum value of the function g in the domain, beyond whose limits the point is known not to move, when the value $A < N$ undergoes a transformation. We require that $h(b - a) > 2Ng_0$. The meaning of this inequality is that the relay shall produce a sufficient "advancing" effect in order to compensate for the tendency of the non-conservative force (if it is non-dissipative) to intensify the oscillations.

If the system is by itself statically stable, then the first condition is satisfied by any positive h. The second condition is automatically fulfilled, if the force g is dissipative, i.e. when $g < 0$ (along the part $du/dt < 0$ of the phase trajectory considered.) In the general case both conditions can be fulfilled by assuming a sufficiently large value for h.

Suppose an initial value A_1 satisfies the condition $d < A_1 < N$. Then A_2 is obtained by transformation ba. The magnitude of A_2 is determined by formula (1.4). Since

$$\int_{(1)}^{(2)} g\left(u, \frac{du}{dt}\right) du > 2Ng_0$$

we shall have

$$[hA_2 + \Phi(A_2)] - [hA_1 + \Phi(A_1)] < 0$$

In consequence of the fact that $[hA + \Phi(A)]$ is a non-decreasing function, we immediately conclude that $A_2 < A_1$. If A_2 is situated within the same limits as A_1, i.e. if $d < A_2 < N$, then we analogously can derive that $A_3 < A_2$, and so forth. As a result we shall have a sequence of values $A_1, \ldots, A_l$ such that $A_1 > A_2 > \ldots > A_l$. The amplitudes will decrease until the condition for the transformation ba becomes violated. This proves the first statement: the point moves from the domain C into the domain A. Note that stability "in the large" may not exist [occur].

The series of transformations ba is followed by one of the transformations b_1a_1, aa_1, a_1a_1, or bb. These transformations may lead, depending on the properties of the functions g and f, to $A_{l+1} < A_l$, as well as to $A_{l+1} > A_l$. In the latter case, after one or several transformations of such type, the point enters the zone B, which unifies states, obtained by the transformations aa, a_1a_1 and bb, to be followed by the transformation ba. The latter returns the point, with the next oscillation, into the zone A. In this manner a point, originally located in the zone C, undergoes several times the transformation ba before arriving at the zone A; then it goes through combinations of various transformations, which do not lead it beyond the limits of the zone B.

3. Stability of limit cycles

We restrict the class of functions g under consideration, by imposing upon them the following conditions:

(1) The product $g du/dt$ does not change its sign anywhere in the entire zone A; in other words, the force g is either dissipative or non-dissipative;

(2) The absolute magnitude of g is a non-decreasing function of du/dt at any u, in other words

$$\left| g\left[u, \left(\frac{du}{dt}\right)_1\right]\right| \leqslant \left| g\left[u, \left(\frac{du}{dt}\right)_2\right]\right| \qquad \text{when } 0 < \left(\frac{du}{dt}\right)_1 < \left(\frac{du}{dt}\right)_2 \tag{3.1}$$

We shall prove now the fundamental theorem: if there are limit cycles in the system, all of them are stable if $g(u, du/dt)(du/dt) > 0$ and they are all unstable if $g(u, du/dt)(du/dt) < 0$.

For the special case of a linear equation and of basically compatible relay characteristics, the particular feature of such systems - - the possibility of the existence of several stable limiting cycles in the absence of any unstable ones - - was stated in a paper by A.S. Alekseev [1]. He has also shown that depending on the relationship between the parameters, the system may contain complex limit cycles. It is natural to expect the existence of such limit cycles in the general case as well. We are going to examine their stability independently of their complexity.

We refer to our formula (1.7), expressing the relationship between the values A_i and A_{i+k}, separated from each other by any sequence of k transformations. The total differential of (1.7) is

$$[h + f(A_{i+k})]\, dA_{i+k} - [h + f(A_i)]\, dA_i = \frac{d\Psi}{dA_i} dA_i$$

from which

$$\frac{dA_{i+k}}{dA_i} = \frac{[h + f(A_i)] + d\Psi / dA_1}{h + f(A_{i+k})} \tag{3.2}$$

Assume that the sequence $A_i, \ldots, A_{i+k}$ belongs to a limit cycle, which closes at the kth, oscillation, so that $A_i = A_{i+k}$. It is known [2] that the stability of the cycle is determined by the derivative (3.2) with $A_{i+k} = A_i$ substituted into its right-hand side. This substitution leads to

$$\frac{dA_{i+k}}{dA_i} = \frac{[h + f(A_i)] + d\Psi / dA_i}{h + f(A_i)} \tag{3.3}$$

The stability of the cycle depends on the sign of the derivative $d\Psi/dA_i$. Differentiating with respect to A_i the first term in the expression for the function Ψ we find

$$\frac{d}{dA_i} \int\limits_{(i)}^{(i+1)} g\left(u, \frac{du}{dt}\right) du = g(A_i, 0) + g(A_{i+1}, 0) \frac{dA_{i+1}}{dA_i} + \int\limits_{(i)}^{(i+1)} \frac{dg}{dA_i} du \tag{3.4}$$

All of the three terms just obtained have the same sign identical with that of g on the portion of the phase trajectory considered. Indeed, the phase trajectories, starting from two points at infinitesimal distance from each other, cannot intersect each other; therefore $dA_{i+1}/dA_i > 0$. The phase trajectory starting from the point $A_i + dA_i$ goes upward. Therefore, taking into account that g is a non-decreasing function of du/dt, we find that the integral within the same limits along that path is absolutely larger, so that

$$\operatorname{sgn} \int\limits_{(i)}^{(i+1)} \frac{dg}{dA_i} du = \operatorname{sgn} g$$

The same considerations apply to the remaining integrals appearing in the expression for Ψ. Thus we finally have

$$\begin{aligned} \frac{dA_{i+k}}{dA_i} &> 1 \quad \text{if } \frac{d\Psi}{dA_i} > 0, \text{ i.e. if } g\,\frac{du}{dt} < 0 \\ \frac{dA_{i+k}}{dA_i} &< 1 \quad \text{if } \frac{d\Psi}{dA_i} < 0, \text{ i.e. if } g\,\frac{du}{dt} > 0 \end{aligned} \tag{3.5}$$

The result is instability of the limit cycle in the first case and its stability in the second case.

The general picture of motion, derived in Section 2, remains the same in both cases: there is a region C, from which the phase trajectories lead into the region A, nearer to the origin of the system of coordinates. In the second case, when there may be stable limit cycles in the system, it is not difficult to form an idea about the further motion of the point. After having entered the region A, the phase curve travels around one of the limit cycles or approaches the origin of the system of coordinates. The motion becomes much more complicated, if there cannot be any stable limit cycles. After having entered the zone A, the phase trajectory does not cross the limits of the zone B, it does not approach any limit cycle (since there cannot be any stable limit cycles), nor does it travel around the origin of the system of coordinates (since the origin of the system of coordinates is unstable, because $g du/dt < 0$). It is, therefore, natural to imagine that the phase trajectory, if continued indefinitely, fills out some region. To describe such a motion for extended periods of time it appears proper to use some ideas of the theory of probability.

4. Representation of the motion by means of probability density

We introduce into our considerations a function $p(A)$, the probability density, such that the element of probability $dW(A)$ of the intersection of the phase trajectory with the u-axis in the interval $(A, A + da)$ is expressed by the formula

$$dW(A) = p(A)\, dA \tag{4.1}$$

In the following we shall consider the motion only in the zones A and B, the duration of which can be unlimited, as shown above. Therefore the function $p(A)$ just introduced has the property

$$\int_{\alpha}^{\beta} p(A)\, dA = 1 \tag{4.2}$$

where α is the inner boundary of the domain A, while β is the outer boundary of the domain B.

Assume the probability density $p_1(A_1)$ to be given in the interval (α,β); it is required to determine the probability density $p_2(A_2)$ after transformation, if the transformation function $A_2 = \phi(A_1)$ is known. To simplify the problem, we first consider the case when the function just mentioned produces a single-valued transformation of the interval (α,β) of the values of A_1 into the same interval of the values of A_2 and vice versa, so that only one value of A_2 corresponds to one value of A_1, and vice versa. Besides, we assume that $\phi' \neq 0$ and $\phi' \neq \infty$ in the entire interval. After transformation the value A_1 is changed to A_2 and the value $A_1 + dA_1$ to $A_2 + dA_2$. Since the transformation is single valued in both directions, the elements of probability are equal in the intervals considered, so that

$$p_1(A_1)\, dA_1 = p_2(A_2)\, dA_2$$

Thus we have for this simplest case the following formula for the transformation of probability density

$$p_2(A_2) = p_1(A_1) \frac{dA_1}{dA_2} = p_1(A_1) \frac{1}{\varphi'(A_1)}, \qquad \varphi(A_1) = A_2 \tag{4.3}$$

This result can be generalized to include the case of transformations, which are not single valued in both directions. Assume that for various parts of the basic interval we have the following transformation formulas:

$$\begin{aligned} A_2 &= \varphi_1(A_1) \text{ when } \alpha < A_1 < \alpha_1 \\ A_2 &= \varphi_2(A_1) \text{ when } \alpha_1 < A_1 < \alpha_2 \\ &\cdots\cdots\cdots\cdots \\ A_2 &= \varphi_n(A_1) \text{ when } \alpha_{n-1} < A_1 < \beta \end{aligned} \tag{4.4}$$

We furthermore assume that the basic interval is subdivided in such a manner that for each separate transformation the mutual correspondence remains single valued. The superposition of the transformations violates, however, in the general case, the single-valued correspondence in the reversed direction (Fig. 9). On some parts (e.g. $\alpha'\beta'$ and $\alpha''_2\alpha'_2$) the transformation takes place several times from various portions of the original interval. On other parts (e.g. $\alpha'_1\alpha''_1$) there will be. on the contrary, no transformations from any part of the original interval. It is obvious that on all portions of the original interval of the values of A_2, the probability densities from separate elemental transformations will add up.

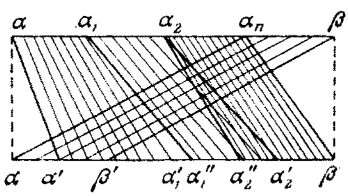

Fig. 9.

This leads to the following generalization of the formula (4.3):

$$p_2(A_2) = \sum_i{}' p_1(A_1)\frac{1}{\varphi'_i(A_1)}, \quad \varphi_i(A_1) = A_2 \tag{4.5}$$

The summation extends, on each part of the interval (α,β), over such values of i, which correspond to transformation functions on the part considered.

As stated already above, there are five functions for the equation of the second order, each of which carries out a single valued transformation $(ba, b_1a_1, aa, a_1a_1, bb)$ in both directions. Therefore, the transformation problem can become quite complicated in the general case.

5. Stationary probability density

In the general case, when the function p_1 is given arbitrarily, the function p_2 arising after transformation, will differ from the original function. In the following we consider the question, how to find such a probability density, that it remains unchanged after transformation by formulas (4.5), that is $p_1 \equiv p_2 = p_0$. It is obvious that such a function will remain unchanged after any number of transformations, since the transformation formulas are always the same. Therefore we shall call p_0 stationary probability density. We assume that the stationary probability density is finite at every point of the interval, that it exists, and that it is unique.

For a practical determination of the function p_0, the method of successive approximations can be used. To this end we start from an arbitrary function p_1, satisfying the integral condition (4.2), and we subject it to multiple transformation by formulas (4.5), continuing the process until the approximations start to coincide with each other within a prescribed degree of accuracy. The question of the convergence of the process is not to be examined here.

Let us study some properties of the function p_0 which become evident in the course of application of the method of successive approximations. For the purpose of simplifying the subsequent discussion we exclude from consideration statically unstable systems. Then the five elementary functions of transformation reduce to only two: ϕ_1, corresponding to the transformation *ba*, and ϕ_2, corresponding to the transformation *aa*.

Depending upon the properties of the transformation functions, we can obtain three fundamentally different types of transformation:

1. The transformation of the first type produces single-valued correspondence in both directions in the entire interval (Fig. 10);
2. The transformation of the second type produces single-valued correspondence in both directions in the entire interval $\alpha < A_1 < \beta$ and in a part of the interval $\alpha < A_2 < \beta$ (for $\alpha < A_2 < \beta'$ and for $\alpha' < A_2 < \beta$). There are no transformations on the part $\beta' < A_2 < \alpha'$ (Fig. 11);
3. The transformation of the third type is not single-valued: on parts $(\alpha\alpha')$ and $(\beta'\beta)$ a single-valued transformation in both directions takes place, but on part $(\alpha'\beta')$ the transformations are superposed, (Fig. 12).

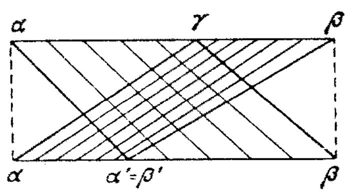

Fig. 10.

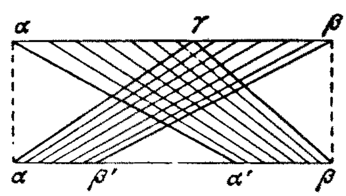

Fig. 11.

The transformation of the first type occupies a limiting position between the transformations of the second and the third types.

It is possible to establish a relation between the type of transformation and the stability conditions of the limit cycles, obtained in Section 3. We are going to show that $g du/dt > 0$ necessitates transformations of the second type, while $g du/dt < 0$ leads to those of the third type, and $g \equiv 0$ to those of the first type.

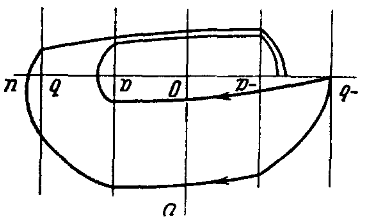

Fig. 12.

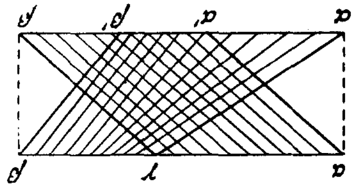

Fig. 13.

Let us apply formula (1.7) to a sequence of two transformations of the values $A_1 = \gamma - 0$ and $A_1 = \gamma + 0$ (the point $A_1 = \gamma$ separates portions of value A_1, which are to undergo the transformations aa and ba, respectively). For the point $\gamma - 0$ we will have a sequence of transformations $(aa - ba)$, for the point $\gamma + 0$ a sequence of transformations $(ba - aa)$. The phase trajectory of the point $\gamma - 0$ embraces the trajectory of the point $\gamma + 0$, (Fig.13). By virtue of the same considerations, which we have used in the proof of the stability of the limit cycles in Section 3, we find

$$|\Psi_-| > |\Psi_+|$$

where Ψ_- and Ψ_+ correspond to the transformations of the points $\gamma - 0$ and $\gamma + 0$, respectively.

The constant R in formula (1.7) does not depend on the sequence of the elemental transformations and has, therefore, the same value in both cases. Taking into account that $[hA + f(A)]$ is a non-decreasing function of A, we find

$$\alpha' > \beta' \text{ when } g\frac{du}{dt} > 0, \qquad \alpha' = \beta' \text{ when } g \equiv 0, \quad \alpha' < \beta' \text{ when } g\frac{du}{dt} < 0$$

In other words, if only stable limit cycles are possible, then the transformation of the second type takes place, if only unstable ones are possible, we will have transformations of the third type; in the limiting case we have transformations of the first type. Let us consider the course of the process of successive approximations for each type of transformations.

Starting with the transformation of the second type, we assume that in the interval (α, β) the function p_1 is arbitrarily given to be finite and never vanishing. After the first transformation there will appear a portion $(\beta'\alpha')$, along which the function p_2 of the second approximation will be zero. With each new transformation the specific weight of the portions, where the function of the ith approximation equals zero, will rise. It is possible to imagine that in the limit, after an unrestrictedly large number of transformations, we will obtain a system of δ-functions; this corresponds to one or several stable limit cycles. Of course, the procedures indicated are not applicable for practical determination of stable limit cycles; they are given here in order to establish a general point of view for all types of transformation.

Apparently the process of successive approximations for transformations of the third type will not lead to a system of δ-functions; this corresponds to the fact of absence of stable limit cycles in the system.

In the case of transformations of the first type there is no need for the method of successive approximations for determination of the stationary probability density. It can be obtained from the following considerations.

For the transformation of the first type we have $\Psi \equiv 0$. Therefore equation (3.2) assumes the form

$$\frac{dA_2}{dA_1} = \frac{h + f(A_1)}{h + f(A_2)}$$

This formula is equally valid for the transformation ba and for the transformation aa. Thus we have the following single formula for the transformation of the probability density

$$p_2(A_2) = p_1(A_1)\frac{h + f(A_1)}{h + f(A_2)} \tag{5.1}$$

This relationship determines the transformation, single-valued in both directions, of the function p_1 from the interval (α, β) to the same interval. Consequently, having transformed by means of this formula the function

$$p^*(A) = \frac{1}{T}[h + f(A)] \tag{5.2}$$

where T is a constant, we obtain as a result the same function $p^*(A)$. This means that $p^*(A)$ represents the desired function $p_0(A)$ multiplied by some constant factor. This constant factor is to be chosen in such a way as to fulfil the fundamental integral condition (4.2).

Ultimately we obtain the following formula for the stationary probability density

$$p_0(A) = [h + f(A)]\left(\int_\alpha^\beta [h + f(A)]\, dA\right)^{-1} \tag{5.3}$$

The results of examination of the first type of transformation, corresponding to complete absence of a non-conservative force, may be, apparently used for an approximate description of the motion in the case of a transformation of the third type, provided the absolute magnitude of the function g is small enough.

Thus, in the case of the transformation of the second type, the motion is described by means of limit cycles, and in the case of transformations of the first and third types by means of stationary probability density. A well-known analogy can be used for the behavior of the system when it has stable limit cycles and when it has none. In the first case the motion, having started under arbitrary initial conditions, tends in the course of time to assume a stationary state independent of the initial conditions. In the second case the motion approaches again a stationary state, independent of the initial conditions (or of the distribution of probability over the initial conditions). In the first case, however, the stationary state represents a fully defined periodical motion, fixed precisely in the phase plane, and in the second the motion is non-periodical, with varying non-repeating amplitudes; the latter state is described by stationary probability density.

Apparently, if it is sufficient to have averaged characteristics of motion for large values of time, the description of the behavior of the system under consideration by means of probability density is by no means inferior to that by means of cycle parameters for systems with stable limit cycles. If we know the stationary probability density, we have the means of calculating the average "period", the average value of the amplitude, etc. We note that limiting values of some quantities (minimum period, maximum amplitude) can be obtained simply from inspection of motion in the phase plane without searching for the stationary probability density.

6. Linear equation (absence of stable cycles)

Returning to the original linear equation (0.1), we change the variables as follows:

$$\tau = t\sqrt{h}, \quad u^\circ = \frac{u}{a}$$

This gives

$$\frac{d^2u^\circ}{d\tau^2} + \varepsilon\frac{du^\circ}{d\tau} + ru^\circ + F_0(u^\circ) = 0 \tag{6.1}$$

where the function $F_0(u^0)$ has the shape shown in Fig.14. There are five independent parameters in the equation, namely $\epsilon, r, k, \delta_1, \delta_2$.

In the present section we will consider the cases $\epsilon \leqslant 0$, which correspond to absence of stable limit cycles.

Let us start with the simplest case $\epsilon = 0$. Under this condition, the boundary of the domains C and D moves toward infinity and the system is stable "in the large". According to Section 5 the condition $\epsilon = 0$ leads

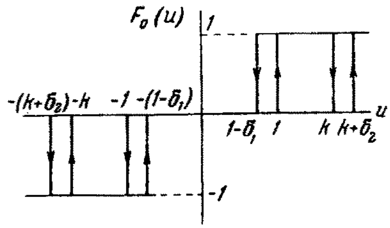

Fig. 14.

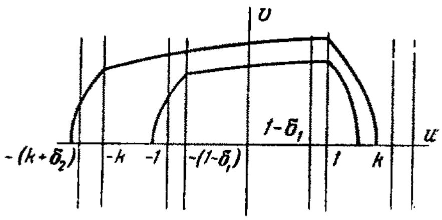

Fig. 15.

to transformations of the first type. Formula (5.3) for the stationary probability density becomes

$$p_0(A) = (1+rA)\left(\int_\alpha^\beta (1-rA)\,dA\right)^{-1} = \frac{1+rA}{(\beta-\alpha)+{}^1/{}_2 r(\beta^2-\alpha^2)} \qquad (6.2)$$

The integration limits are to be established by inspection of the motion in the phase plane (Fig.15). The equations of the portions of the phase trajectories are of the form

$$\tfrac{1}{2}v^2 + \tfrac{1}{2}ru^{\circ 2} = C_1 \qquad \text{on parts where } F_0(u^\circ) = 0$$

$$\tfrac{1}{2}v^2 + \tfrac{1}{2}ru^{\circ 2} + u^\circ \operatorname{sgn} u^\circ = C_2 \qquad \text{on parts where } F_0(u^\circ) = 1 \operatorname{sgn} u^\circ$$

After "sewing up" the portions of the phase trajectories along the straight lines $u^0 = -(1-\delta_1)$ and $u^0 = 1$ for the transformation aa and along the straight lines $u^0 = -k$ and $u^0 = 1$ for the transformation ba we obtain

$$A_2 = \sqrt{\left(A_1 + \frac{1}{r}\right)^2 - \frac{2}{r}(k-1)} - \frac{1}{r} \qquad \text{(transformation } ba\text{)}$$

$$A_2 = \sqrt{\left(A_1 + \frac{1}{r}\right)^2 + \frac{2}{r}\delta_1} - \frac{1}{r} \qquad \text{(transformation } aa\text{)} \qquad (6.3)$$

Using these formulas for the transformation of the quantity $A_1 = k + \delta_2$ we obtain the values of α and β:

$$\alpha - \sqrt{\left(k + \delta_2 + \frac{1}{r}\right)^2 - \frac{2}{r}(k-1)} - \frac{1}{r}$$

$$\beta = \sqrt{\left(k + \delta_2 + \frac{1}{r}\right)^2 + \frac{2}{r}\delta_1} - \frac{1}{r} \qquad (6.4)$$

With the values of α and β thus calculated, the mean amplitude becomes

$$A_0 = \int_{\alpha}^{\beta} A p_0(A)\,dA = \frac{{}^1/_2(\beta^2-\alpha^2) + {}^1/_3 r(\beta^3-\alpha^3)}{(\beta-\alpha) + {}^1/_2 r(\beta-\alpha^2)} \tag{6.5}$$

Figs. 16 and 17 show the mean amplitude A_0 as a function of k for $\delta_1 = 1.0$ and $\delta_1 = 0.5$ and for various values of r. The same figures give the maximum and minimum values of A ($A = \alpha$ and $A = \beta$).

When $r = 0$, the values of A in the region (α, β) are of equal probability and formula (6.5) becomes

$$A_0 = {}^1/_2(\alpha + \beta)$$

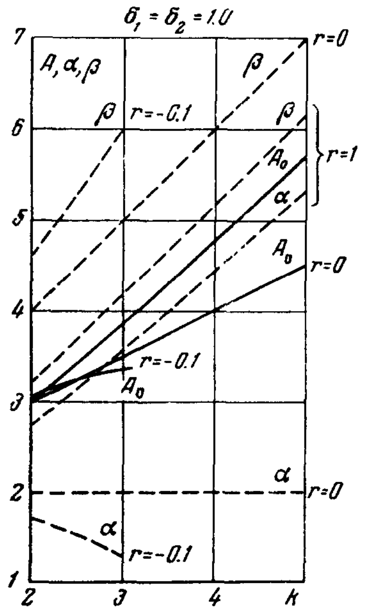

Fig. 16.

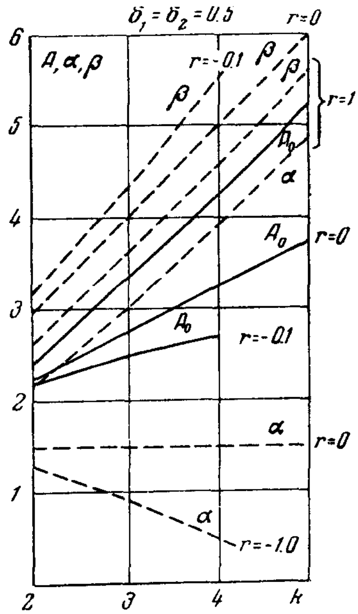

Fig. 17.

The magnitude of the "half-period" as a function of the magnitude of the initial amplitude A_1 for $r = 0$ is expressed by the following formulas:

for the transformation ba

$$T = 2\sqrt{2(A_1 - k)} + \frac{k+1}{\sqrt{2(A_1 - k)}} \tag{6.6}$$

for the transformation aa

$$T = 2\sqrt{2(A_1 - 1 + \delta_1)} + \frac{2-\delta_1}{\sqrt{2(A_1 - 1 + \delta_1)}} \tag{6.7}$$

The mean half-period T_0 can be calculated from the formula

$$T_0 = \frac{1}{\beta - \alpha} \int_\alpha^\beta T(A)\, dA \tag{6.8}$$

T_0 as a function of k for $\delta_1 = 1.0$ and $\delta_1 = 0.5$ is shown in Fig. 18.

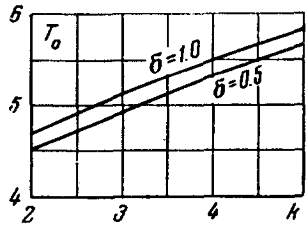

Fig. 18.

We turn now to a study of the characteristics of the motion when $\epsilon < 0$. The relation between A_1 and A_2, to be obtained, as in (6.3), by means of the method of "sewing up" the solutions, is not derived in explicit form. We confine ourselves to the case $r = 0$. In this case the equations, which determine A_2 in terms of A_1, have the following form:

for the transformation *aa*

$$\begin{aligned} -\varepsilon^2 [A_1 - (1 - \delta_1)] &= \varepsilon v_1 + \ln(1 - \varepsilon v_1) \\ \varepsilon^2 [A_2 - 1] &= \varepsilon v_2 - \ln(1 + \varepsilon v_2) \\ v_1 - v_2 &= \varepsilon(2 - \delta_1) \end{aligned} \tag{6.9}$$

for the transformation *ba*

$$\begin{aligned} -\varepsilon^2 [A_1 - k] &= \varepsilon v_1 + \ln(1 - \varepsilon v_1) \\ \varepsilon^2 [A_2 - 1] &= \varepsilon v_2 - \ln(1 - \varepsilon v_2) \\ v_1 - v_2 &= \varepsilon(k + 1) \end{aligned} \tag{6.10}$$

The expression for the derivative for both transformations can be written in the form

$$\frac{dA_2}{dA_1} = \frac{v_2 (1 - \varepsilon v_1)}{v_1 (1 + \varepsilon v_2)}$$

Practically, the computation of A_2 in terms of A_1 is carried out in the following order: starting from a chosen value of v_1, we determine v_2, and then we find A_1 and A_2.

In a study of the motion, the question concerning the boundary of the stability domain arises first. To determine that boundary it is necessary to find the periodical solution ($A_1 = A_2$) of the system of equations (6.10) corresponding to the transformation *ba*, in other words to find the amplitude of the largest unstable limiting cycle A_*. The function A_* of ϵ is shown in Fig.19 for various values of k.

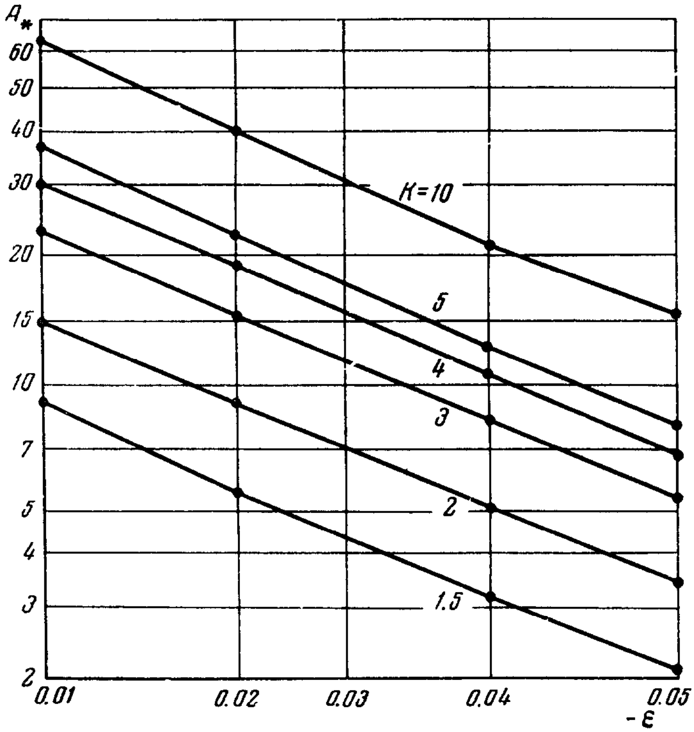

Fig. 19.

The computation of the stationary probability density by means of the method of successive approximations represents in practice a very laborious process. The formulas become too complicated already at the second approximation, if the computations are carried out in general form. It is more convenient to proceed numerically using graphical constructions. A further complication is caused by the fact that it is necessary to deal with discontinuous functions, and the number of discontinuities increases with each further approximation. In some cases, however, when great precision is not required, some ways of simplification can be indicated.

Thus, for instance, when $r = 0$ in the formula (4.5) for recomputation of probability, the coefficients of $p_1(A_1)$ do not depend strongly on A_1 in a wide interval of variability of A_1. Therefore these coefficients can be considered as approximately constant, having different values for the transformations aa and ba. If we take a constant as a starting approximation, then all further transformations reduce to a consecutive superposition of discontinuous functions. The computation consists in the determination of the points, where discontinuities appear, and in the determination of the magnitude of the discontinuities themselves.

As an example we give here the results of computation of the stationary probability density for the following values of the parameters (Figs.20,21):

$$1)\quad k=1.9,\quad \varepsilon=-0.1,\quad \delta_1=0.1,\quad \delta_2=0.1$$

$$2)\quad k=2.9,\quad \varepsilon=-0.01,\quad \delta_1=0.1,\quad \delta_2=0.1$$

The graphs show broken lines corresponding to the zero, the first, the next to the last and the last approximations. As is evident from the examples, the first approximation represents the course of the curve of

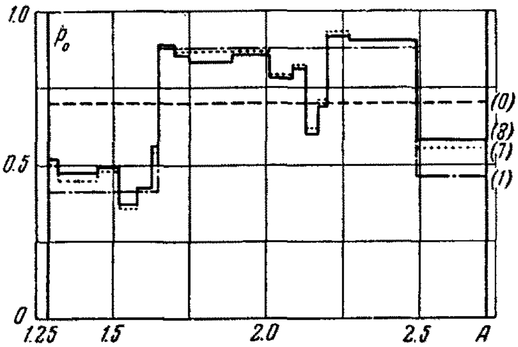

Fig. 20.

stationary probability density, already in fundamental agreement with the general character of the curve, although there are, in certain spots, still considerable differences between the first and the last approximations.

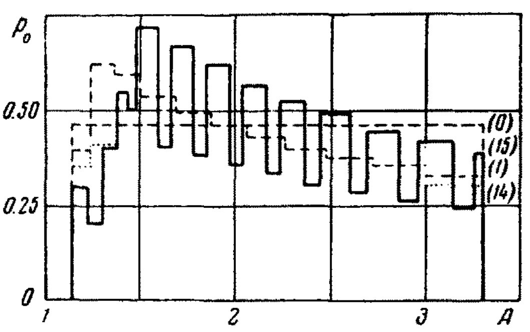

Fig. 21.

In view of scarcity of calculated material, it is impossible to state conclusions concerning the necessary number of approximations. It is merely possible to state that in some cases that number can be quite large. Referring to Fig.20, we see that the seventh and the eighth approximations differ still notably from each other. In actual problems, however, the exact knowledge of the curve of the stationary probability density will not be necessary, since the ultimate result of the computation will consist of average values of quantities, and these values will not be influenced strongly by fine variations of the shape of the curve. Thus, for the two given examples the mean values of the amplitudes A_0, computed from the first approximation, are $A_0^{(1)} = 2.01$, $A_0^{(1)} = 2.14$, respectively, while the last approximation leads to $A_0^{(8)} = 2.03$, $A_0^{(15)} \equiv 2.15$, respectively.

It appears desirable to find ways of obtaining a rough approximation for the curve of stationary density of probability with a minimum amount of labor. This proves to be possible at least for such values of ϵ and δ_1, which are small as compared with k.

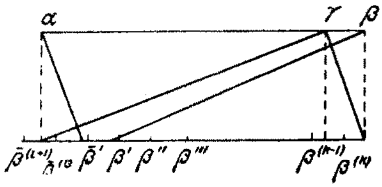

Fig. 22.

Let us investigate the procedure of transformation for small values of ϵ and δ_1 somewhat more in detail (Fig.22). The zone (γ,β), whose points are to undergo the transformation ba, represents a small part of the total interval (α,β); it is the smaller, the smaller ϵ and δ_1 are. Therefore, the portion, which at the second transformation becomes the portion $\beta'\gamma$ (this portion does not contain any transformation from portion $(\gamma\beta)$), represents the larger part of the entire interval. Assume that the point β' transforms consecutively into the values β'', β''', ..., $\beta^{(k)}$, so that $\beta^{(k)} > \gamma$, but $\beta^{(k-1)} < \gamma$. The stationary probability density must satisfy the following condition:

$$\int_{\beta'}^{\beta''} p_0(A)\,dA = \int_{\beta''}^{\beta'''} p_0(A)\,dA = \ldots =$$

$$= \int_{\beta^{(k-1)}}^{\beta^{(k)}} p_0(A)\,dA = \Delta W \qquad (6.11)$$

The validity of this chain of equalities follows from the fact of single valued reciprocal correspondence for the transformation of the intervals.

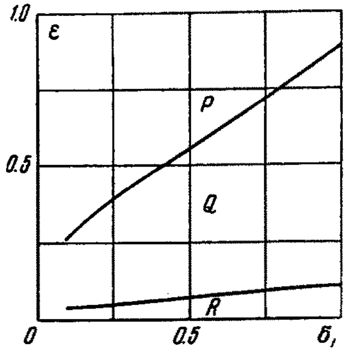

Fig. 23.

The magnitude of the total probability for each interval is unknown; it is, however, possible to estimate it. The fundamental integral property (4.2) of the stationary probability density can be written, in the case considered, in the form

$$(k-1)\Delta W + W_1 + W_2 = 1 \tag{6.12}$$

where W_1 and W_2 are the total probabilities for the intervals (α,β') and $(\beta^{(k)},\beta)$, respectively.

Since a part of the portion $(\beta^{(k-1},\beta^{(k)})$ was transformed on the portion $(\beta^{(k)},\beta)$ at the preceding transformation, we must have

$$0 < W_2 < \Delta W \tag{6.13}$$

It remains to estimate the magnitude of W_1. We find the value of $\beta^{0\prime}$, which after transformation as gives β', then of $\beta^{0\prime\prime}$, which gives $\beta^{0\prime}$, and so forth, until we get $\beta^{0(i)}$, so that $\beta^{0(2)} > \alpha$, but $\beta^{0(i+1)} < \alpha$. We have then obviously

$$\Delta W < W_1 < (i+1)\Delta W \tag{6.14}$$

Substituting (6.13)and (6.14) into (6.12), we obtain an estimated value for ΔW:

$$\frac{1}{k+i+1} < \Delta W < \frac{1}{k}$$

If k is large and i is small, then the value of ΔW can be determined with practically sufficient accuracy. Knowing ΔW we can obtain an approximation for the function p_0, say, in the form of a discontinuous curve, the discontinuities in the points β', β'', ..., $\beta^{(k)}$ being determined from the conditions (6.11), or in the form of a continuous curve. The total probabilities W_1 and W_2 on the limiting portions remain undetermined in this method of calculation.

7. Linear equation (possibility of stable cycles)

We consider again the equation (6.1), in which we now assume $\epsilon > 0$.

In this case the system may have stable limit cycles, as can be concluded from Section 4. A more detailed study shows that the system has either no cycles at all, or it has one stable cycle. Fig.23 shows, in a $\delta_1 \epsilon$ system of coordinates, the region P of no auto-oscillations, the region Q of simple limit cycles and the region R of complex limit cycles for the case $r = 0$ and $k = 2$. It is important to note that the system is always stable "in the large" and that for an arbitrary relation between the parameters, the amplitude of the limit cycle does not exceed the value $k + 2\,\delta_2$.

If the system has no limit cycles, then it is necessary to distinguish between the effects of the relay at small and at large deviations. At small deviations the presence of a non-vanishing δ_1 leads to a decrease of damping. At deviations in excess of k, when the "advancing" effect of the relay is being utilized, the damping rises sharply. Fig.24 shows the

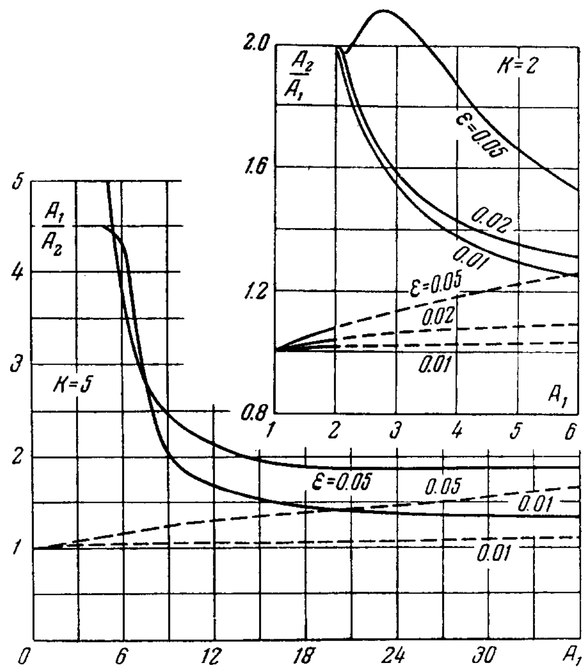

Fig. 24.

relationship of the ratio of two consecutive amplitudes for a system with $r = 0$ for various values of δ in terms of the initial amplitude A_1 at $r = 0$. The broken lines in the domains $A_1 > k$ show the ratio of the amplitudes for the case of a relay of usual characteristics. The curves show that the damping effects of the relay with advancing characteristics become most pronounced at initial deviations somewhat exceeding the value $k + \delta_2$.

BIBLIOGRAPHY

1. Alekseev, A.S., *Dvukhpositsionnyi reguliator temperatury s zonoi operezheniia. Sbornik pamiati A.A. Andronova. (Two-Positional Temperature Regulator with a Zone of Advancing. Collection of Papers in Memory of A.A. Andronov).* Academy of Sciences of the U.S.S.R., Moscow, 1955.

2. Andronov, A.A. and Khaikin, S.E., *Teoriia kolebanii (Theory of Oscillations)* Part 1. Division of Scientific-Technical Publications, 1937.

Translated by I.M.