ON FEATURES OF INTRODUCING A SMALL PARAMETER IN THE INVESTIGATION OF NONLINEAR OSCILLATIONS IN AUTOMATIC SYSTEMS

(OB OSOBENNOSTIAKH VVEDENIIA MALOGO PARAMETRA PRI ISSLEDOVANII NELINEINYKH KOLEBANII V AVTOMATICHESKIKH SISTEMAKH)

PMM Vol.26, No.1, 1962, pp. 62-69

E. P. POPOV
(Leningrad)

(Received September 29, 1961)

The objective of the present article is to call the attention of mathematicians, interested in applied problems, to an unusual formulation of the problem of a small parameter in the equations of the dynamics of nonlinear automatic systems, which, seemingly, can be fruitfully utilized for the development of a mathematically rigorous basis (to which the author makes no pretense) for widely used approximate methods in the study of nonlinear automatic systems.

The dynamics of automatic systems is frequently described by a system of ordinary differential equations, which in general are nonlinear.

We shall assume, for the sake of simplifying the presentation, that the system involves only one nonlinearity

$$y = F\ (x,\ px) \qquad (p = d/dt)$$

even though the discussions can be extended to the case of several nonlinearities. In the case of one nonlinearity, all remaining equations of the system whatever its complexity may be, are often reduced to one linear equation of a high order

$$Q(p)x + R(p)y = S(p)f$$

Here $Q(p)$, $R(p)$, $S(p)$ are operator polynomials with constant coefficients, f is the given external reaction.

In automatic systems, it is necessary to take account of nonlinearities which differ considerably from the corresponding linear function $y = k_1 + k_2 px$ (for example, relay hysteresis, some of the saturation type, slack, free play, and many others). For the variable x, which occurs in the nonlinear term, one frequently observes in practice processes which are nearly sinusoidal (periodic or decaying) even though for the variable y, and also for the other variables of the system, the process may be far from a sinusoidal nature (for example, rectangular in a relay system).

This important peculiarity will be utilized in what follows. We will consider automatic systems which are known as "coarse" systems (in the terminology of Andronov).

1. On the periodic solution of a homogeneous equations. Let us first consider the homogeneous equation

$$Q(p)\,x + R(p)\,y = 0, \qquad y = F(x, px) \tag{1.1}$$

In looking for a periodic solution for the variable x, we assume that

$$x = x_1 + \varepsilon z(t) \qquad (x_1 = A_1 \sin \Omega_1 t) \tag{1.2}$$

Here $z(t)$ is an arbitrary bounded function of time, ϵ is a small parameter.

In the application of quasilinear analysis, one selects for the nonlinear function $y = F(x, px)$ frequently the form

$$y = k_1 x + k_2 px + \mu f(x, px) \tag{1.3}$$

where μ is a small parameter. In other words, one assumes that the solutions for all variables of the given system are close to the solutions of some equivalent linear system. As was already mentioned, this is not justified for automatic systems with strong nonlinearities. Therefore, we should not use the expression (1.3) which is frequently used in other quasilinear problems, or else we must not treat μ as a small parameter.

We shall make use of the fact that only the solution for x (1.2) is close to the linear one, and we will write for the solution y the expression

$$y = F(x_1, px_1) + \varepsilon\Phi(t) \tag{1.4}$$

i.e. we assume that it is close to the true nonlinear function taken from the found approximate solution x_1 for the variable x.

Furthermore, in practical computations of automatic systems, one usually does not introduce explicitly the Expression (1.4), but one makes use of the approximate expression of the form $y = k_1 x + k_2 px$, dropping $\mu f(x, px)$ even though it is obvious that it is not small. One thus obtains a satisfactory result as a first approximation which, in the majority of cases, turns out to be good enough in practice, both in the qualitative and quantitative aspects. It is known (see, for example, [1]) that such an approach to the approximate solution coincides in automatics with the introduction of a postulate on the presence within the system of the quality of a filter of the linear part (1.1) which can be written in the form

$$\left|\frac{R(ik\Omega_1)}{Q(ik\Omega_1)}\right| \ll \left|\frac{R(i\Omega_1)}{Q(i\Omega_1)}\right| \qquad (i=\sqrt{-1}) \tag{1.5}$$

Here k stands for the orders of the higher harmonics, $k = 2, 3, \ldots$, or $k = 3, 5, \ldots$.

We shall show by means of the simplest mathematical arguments how one can connect the statements (1.2) and (1.4) with each other, and also with the physical postulate on the property of the filter (1.5). Hereby one, seemingly, obtains also a good explanation of the fact that the direct utilization of an expression of the type $y = k_1x + k_2px$ (harmonic balance, harmonic linearization, describing function, etc.) for an obviously not small μ in (1.3), yields a satisfactory result in the first approximation, in spite of the seemingly unjustifiability of such a solution.

The first incomplete attempt in this direction was made in [2].

2. Construction of the periodic solution. Thus we assume that for the variable x in the given system (1.1) there exists a periodic solution of the type (1.2) with still unknown A_1, Ω_1 and $\epsilon z(t)$.

Let us express the given nonlinear function $y = F(x, px)$ in the form

$$y = F(x_1, px_1) + [F(x_1 + \varepsilon z, px_1 + \varepsilon pz) - F(x_1, px_1)] \tag{2.1}$$

The first term can be expanded into a Fourier series

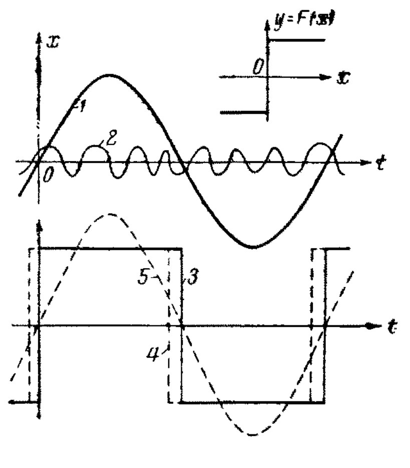
$$F(x_1, px_1) = F_0 + C\sin\Omega_1 t + B\cos\Omega_1 t + \sum_{k=2}^{\infty} F_k(t) \tag{2.2}$$

in which the higher harmonics F_k are small.

The expression which is enclosed in the square brackets in the Formula (2.1) can be expressed by the use of Taylor's series in the form

$$[F(x_1 + \varepsilon z, px_1 + \varepsilon pz) - F(x_1, px_1)] =$$
$$= \varepsilon\left[\frac{\partial}{\partial x}F(x_1, px_1)\,z + \frac{\partial}{\partial px}F(x_1, px_1)\,pz\right] + \varepsilon^2[\ldots] + \varepsilon^3[\ldots] + \ldots \tag{2.3}$$

This expression will be small if the partial derivatives with respect to x and px are bounded. This will be true for continuous nonlinear functions $F(x, px)$ that occur in real automatic systems. One can depend on the smallness of the Expression (2.3) also for discontinuous nonlinear functions met in practical applications, for example, in relay systems, which the indicated derivatives will be delta functions. This is easily seen in the example given in the figure. The increment ϵz (the curve 2) of the basic term x_1 (curve 1) causes only a small displacement 4 of the graph 3 for $F(x_1, px_1)$.



When these weak conditions are imposed on the nonlinear function, one obtains from (2.1) and (2.3) the fundamental expression (1.4), namely,

$$y = F(x, px) = F(x_1, px_1) + \varepsilon\Phi(t) \tag{2.4}$$

where

$$\Phi(t) = \Phi_1(t) + \varepsilon\Phi_2(t) + \varepsilon^2\Phi_3(t) + \dots \tag{2.5}$$

while $\Phi_1(t)$, $\Phi_2(t)$, ... are determined by the corresponding expressions in the square brackets of the right-hand side of (2.3). As a result one obtains

$$y = F_0 + C\sin\Omega_1 t + B\cos\Omega_1 t + \sum_{k=2}^{\infty} F_k(t) + \varepsilon\Phi(t) \tag{2.6}$$

where the $F_k(t)$ are finite higher harmonics, $\epsilon\Phi$ stands for the small terms of all frequencies. Hence, one does not assume here that the difference between the nonlinear function and the linear function is small.

In this manner, the small term $\epsilon\Phi(t)$ in the Formula (2.4) or in (1.4) symbolizes the small effect of only the higher (small) harmonics of the variable x on the form of the nonlinear oscillations y, but not the smallness of the higher harmonics which are generated by the nonlinearity itself when one substitutes into it the sinusoids (1.2), as would have been the case if the Expression (1.3) had been used with a small μ.

In particular, for discontinuous nonlinear characteristics, the quantity $\epsilon\Phi$ at the points of discontinuities corresponds to small displacements of the lines of jumps (for example, 3-4 in the figure).

Let us now substitute (2.6) and (1.2) into the given Equation (1.1). We thus obtain

$$Q(p)x_1 + R(p)F_0 + R(p)(C\sin\Omega_1 t + B\cos\Omega_1 t) + R(p)\sum_{k=2}^{\infty} F_k = \\ = -Q(p)\varepsilon z - R(p)\varepsilon\Phi \qquad (2.7)$$

If $x_1 = A_1 \sin \Omega_1 t$ represents the exact first harmonic of the periodic solution of (1.2), then the functions $\epsilon z(t)$ and $\epsilon\Phi(t)$ can be represented in the form

$$\varepsilon z = \varepsilon \sum_{k=2}^{\infty} A_k \sin(k\Omega_1 t + \varphi_k) \qquad (2.8)$$

$$\varepsilon\Phi = \sum_{j=1}^{\infty} \varepsilon^j \sum_{k=0}^{\infty} \Phi_{jk} = \sum_{j=1}^{\infty} \varepsilon^j \Phi_{j0} + \sum_{j=1}^{\infty} \varepsilon^j \sum_{k=1}^{\infty} G_{jk} \sin(k\Omega_1 t + \vartheta_{jk}) \qquad (2.9)$$

Besides that, we shall write

$$F_k = N_k \sin(k\Omega_1 t + \eta_k) \qquad (2.10)$$

where in the general case $k = 2, 3, \ldots$, while in a particular case $k = 3, 5, \ldots$ when the nonlinearity of (1.1) is an odd function of x and does not depend on px. The quantities N_k are finite, since the nonlinearity is a real one but

$$N_k \to 0 \quad \text{when } k \to \infty \qquad (2.11)$$

From (2.7) we now obtain a number of exact equations for the corresponding harmonics

$$F_0 = -\varepsilon\Phi_{10} - \varepsilon^2\Phi_{20} - \ldots \qquad (2.12)$$

$$Q(p)x_1 + R(p)(C\sin\Omega_1 t + B\cos\Omega_1 t) = -R(p)\varepsilon\Phi_{11} - R(p)\varepsilon^2\Phi_{21} - \ldots \qquad (2.13)$$

$$Q(p)\varepsilon A_k \sin(k\Omega_1 t + \varphi_k) + R(p)N_k\sin(k\Omega_1 t + \eta_k) = \\ = -R(p)\varepsilon\Phi_{1k} - R(p)\varepsilon^2\Phi_{2k} - \ldots \qquad (k = 2, 3, \ldots) \qquad (2.14)$$

The first of these equations (2.12) gives information only about the smallness of the constant component of the Fourier series. This implies the requirement on the symmetry (with an accuracy ϵ) of the nonlinearity

$$\int_0^{2\pi} F(A\sin\psi, A\Omega\cos\psi)\,d\psi = 0 \qquad (2.15)$$

This restriction can be removed (see below, Section 4). The Equations (2.13) and (2.14) are more important.

The Equation (2.13) for the exact first harmonic $x_1 = A_1 \sin \Omega_1 t$ contains on the right-hand side small terms. We shall neglect these small terms in determining the approximate periodic solution which we will denote by

$$x^* = A \sin \Omega t \tag{2.16}$$

We thus obtain a known equation for the approximate determination of the periodic solution:

$$\left[Q(p) + R(p)\left(q + \frac{q'}{\Omega} p\right)\right] x^* = 0 \tag{2.17}$$

where

$$q = \frac{G}{A} = \frac{1}{\pi A} \int_0^{2\pi} F(A \sin \psi, A\Omega \cos \psi) \sin \psi \, d\psi$$

$$q' = \frac{B}{A} = \frac{1}{\pi A} \int_0^{2\pi} F(A \sin \psi, A\Omega \cos \psi) \cos \psi \, d\psi$$

Next, since the quantities N_k in the equations for the higher harmonics (2.14) are finite (at least for some values of k), the quantities

$$\left| \frac{R(ik\Omega_1)}{Q(ik\Omega_1)} \right| \qquad (k = 2, 3, \ldots)$$

must be of order ϵ, at least, for it follows from (2.14) that

$$\varepsilon A_k \sin(k\Omega_1 t + \varphi_k) = -\left| \frac{R(ik\Omega_1)}{Q(ik\Omega_1)} \right| N_k \sin(k\Omega_1 t + \eta_k + \beta_k) - \cdots$$

where

$$\beta_k = \arg \frac{R(ik\Omega_1)}{Q(ik\Omega_1)}$$

At the same time, the quantity $\left| \frac{R(i\Omega)}{Q(i\Omega)} \right|$ is finite since it follows from (2.17) that

$$\sin \Omega t = -\left| \frac{R(i\Omega)}{Q(i\Omega)} \right| \sqrt{q^2 + (q')^2} \sin(\Omega t + \gamma + \beta)$$

where

$$\gamma = \tan^{-1} \frac{q'}{q}, \qquad \beta = \arg \frac{R(i\Omega)}{Q(i\Omega)}$$

Comparing these two results (and replacing in the first one of them the exact Ω_1 by the approximate Ω, which is inessential in the given case) we come to the conclusion that the equation holds

$$\left|\frac{R(ik\Omega)}{Q(ik\Omega)}\right| = \varepsilon^m c_k \left|\frac{R(i\Omega)}{Q(i\Omega)}\right| \qquad (m \geqslant 1) \tag{2.18}$$

where the c_k are arbitrary finite positive numbers. This is an important condition which must be satisfied by the initial Equation (1.1), if one is to be able to find a solution of the form (1.2) in the presence of a strong nonlinearity (1.1). It corresponds to the above mentioned property of a filter (1.5) of the linear part of the system (1.1). One can verify whether this condition is satisfied after one finds the frequency Ω of the approximate solution (2.16).

From what has been said there follow all published methods for the computation of auto-oscillations in nonlinear automatic systems on the basis of the first harmonic by means of the harmonic linearization, the harmonic balance or describing functions. All of them correspond to the equation of the first approximation (2.17) (see for example [3]).

The following observations are here of basic importance. The expression $qx^* + (q'/\Omega)px^*$ in Equation (2.17) corresponds to the linear terms $k_1x + k_2px$ in Formula (1.3). But from what has been said, it is clear that we restrict ourselves to these terms in the first approximation not because $\mu f(x, px)$ is small, but because finite higher harmonics do not enter into the approximate Equation (2.13). These higher harmonics are generated by the strong nonlinearity and are contained, therefore, in the term $\mu f(x, px)$ of the Formula (1.3) or in the Expression $F(x_1, px_1)$ of the Formula (1.4). These higher harmonics do not enter into the approximate Equation (2.17) because this equation is obtained from the exact Equation (2.13) of the first harmonic. The approximate Equation (2.17) must in this connection be considered as the first harmonic of the sought periodic solution, which is only for one variable x close to the true periodic solution. For the other variables of such a system, the first harmonic (for example the curve 5 in the figure) is far from the periodic solution.

All derivations were made assuming the existence of a periodic solution (1.2). The approximate Equation (2.17) can, however, be justified also in the case when the function $z(t)$ is an arbitrary (bounded) function, i.e. when the exact solution is not periodic but nearly so.

3. Further realization of the result obtained. Making use of the equations for the higher harmonics (2.4), one can develop the obtained approximate solution. Taking into account the condition (2.18), one can drop, on the exact Equations (2.14), the right-hand sides (infinitesimals of higher order). Further, we replace the exact Ω_1 by the approximate value Ω of the frequency, and express the amplitudes of the higher harmonics ϵA_k in fractions of the already found approximate value of the amplitude A of the first harmonic, i.e. $\epsilon A_k = \delta_k A$ where δ_k is a small quantity that has to be evaluated. Each of the higher harmonics of the variable x can then, in accordance with (2.14), be determined separately and approximately in the form

$$x_k = \delta_k A \sin(k\Omega t + \varphi_k) \qquad (k = 2, 3, \ldots) \tag{3.1}$$

where

$$\delta_k = \left|\frac{R(ik\Omega)}{Q(ik\Omega)}\right| \sqrt{r_k^2 + s_k^2}, \qquad \varphi_k = \pi + \tan^{-1}\frac{s_k}{r_k} + \arg\frac{R(ik\Omega)}{Q(ik\Omega)} \tag{3.2}$$

whereby

$$\begin{aligned} r_k &= \frac{1}{\pi A}\int_0^{2\pi} F(A\sin\psi,\ A\Omega\cos\psi)\sin k\psi\, d\psi \\ s_k &= \frac{1}{\pi A}\int_0^{2\pi} F(A\sin\psi,\ A\Omega\cos\psi)\cos k\psi\, d\psi \end{aligned} \tag{3.3}$$

After this, one can refine the first harmonic by taking into account the term with ϵ to the first degree in (2.13). This gives the increments

$$\begin{aligned} \Delta q &= \frac{1}{\pi A}\int_0^{2\pi}\left[\frac{\partial}{\partial x}F(x^*, px^*)\sum_{k=2}^{n} x_k + \frac{\partial}{\partial px}F(x^*, px^*)\sum_{k=2}^{n} px_k\right]\sin\psi d\psi \\ \Delta q' &= \frac{1}{\pi A}\int_0^{2\pi}\left[\frac{\partial}{\partial x}F(x^*, px^*)\sum_{k=2}^{n} x_k + \frac{\partial}{\partial px}F(x^*, px^*)\sum_{k=2}^{n} px_k\right]\cos\psi d\psi \end{aligned} \tag{3.4}$$

By substituting these into the previous equation of the first approximation (2.17), one can obtain an improved value of the amplitude A_1 and of the frequency Ω_1 of the first harmonic (for the computational procedure see [3]). For practical computations with the Formulas (3.4), we restrict ourselves to some finite number of harmonics n, which is entirely permissible because of the property (2.11).

As a final result of such an approximation process, one may write for the nonlinear function (1.1) the following approximate expression:

$$y = A_1\left[(q + \Delta q)\sin\Omega_1 t + (q' + \Delta q')\cos\Omega_1 t + \sum_{k=2}^{\infty}(r_k \sin k\Omega_1 t + s_k\cos k\Omega_1 t)\right] \tag{3.5}$$

in which the finite higher harmonics really occur, in spite of their smallness in the variable x

$$x = A_1\left[\sin \Omega_1 t + \sum_{k=2}^{\infty} \delta_k \sin (k\Omega_1 t + \varphi_k)\right] \tag{3.6}$$

This is a computational realization of the scheme described in Section 2 for the construction of a periodic solution.

4. Some generalizations. An analogous approach can be used for the construction of a solution in a number of other important practical problems. The simplest ones of them are: single frequency forced oscillations under a periodic external force, oscillatory transient processes with gradually changing frequency and a decay index, and also nonsymmetric auto-oscillations when the condition (2.15) is not fulfilled.

The approximate investigation of processes with two frequencies in automatic systems on the basis of the harmonic linearization of the nonlinearity, is of special importance. The point is that the most typical cases of nonlinear automatic systems are those when the auto-oscillations or the forced oscillations are vibrational processes (of relatively high frequency) that are superimposed upon the basic slowly changing control processes (of relatively low frequency). Due to the nonlinearity of the system, both types of processes affect each other considerably.

Here one considers a nonlinear equation of the form

$$Q(p)x + R(p)F(x, px) = S(p)f \tag{4.1}$$

and looks for an approximate solution* of the form

$$x = x^\circ + x^*, \qquad x^* = A \sin \Omega t' \tag{4.2}$$

Here x^* is a vibrational (auto-oscillatory) component, x° is a slowly changing component, i.e. x° changes little during the oscillation period of x^*; the external force $f(t)$ is assumed to be a slowly changing one. The time t' is measured within each period separately.

As above, the nonlinearity is given in the form

$$F(x, px) = F^\circ + \left(q + \frac{q'}{\Omega} p\right)x^* + \sum_{k=2}^{\infty} F_k(t) + \varepsilon\Phi(t) \tag{4.3}$$

but with a finite (nonzero) value of F°. In the general case, we have a definite expression for each nonlinearity

* More precisely, $x = x_0 + x_1 + \epsilon z$.

$$F^{\circ} = F^{\circ}(x^{\circ}, A, \Omega), \qquad q = q(x^{\circ}, A, \Omega), \qquad q' = q'(x^{\circ}, A, \Omega) \tag{4.4}$$

where all the quantities vary slowly with x°.

The finite higher harmonics $F_k(t)$, which enter into (4.3), will again, as above, not play any role in the approximate equation for the variable x. After the substitution of (4.3) into (4.1), the latter equation can be broken up into two interconnected equations of first approximation

$$Q(p)x^{\circ} + R(p)F^{\circ} = S(p)f, \qquad \left[Q(p) + R(p)\left(q + \frac{q'}{\Omega}p\right)\right]x^{*} = 0 \tag{4.5}$$

which correspond to the slowly changing component and to the auto-oscillatory, vibrational component. In this separation of equations, the nonlinear properties are essentially preserved as well as the nonvalidity of the principle of superposition of solutions, because the nonlinear interconnection is preserved between the quantities (4.4) that are contained in these equations.

From the second Equation (4.5) one can determine A and the frequency Ω as a function of x°. After this one can, with the aid of the first Equation (4.4), determine the "displacement function"

$$F^{\circ} = \Phi^{\circ}(x^{\circ}) \tag{4.6}$$

With this substitution, the first Equation (4.5) takes on the form

$$Q(p)\,x^{\circ} + R(p)\,\Phi^{\circ}(x^{\circ}) = S(p)\,f \tag{4.7}$$

Hence, the displacement function $\Phi^{\circ}(x^{\circ})$ serves as a new nonlinearity with respect to the slowly changing signal in the presence of the vibration superimposed in the signal. Ordinarily, the graph of the function $\Phi^{\circ}(x^{\circ})$ is a smooth curve (even for nonsmooth and loop type nonlinearities) that can be linearized in the usual way as

$$F^{\circ} \approx K_{\text{H}}x^{\circ} \quad \left(K = \left(\frac{d\Phi^{\circ}}{dx^{\circ}}\right)_{x_{\circ}=0}\right) \tag{4.8}$$

Then Equation (4.7) can be solved as a linear equation. One must, however, take account of the peculiarity of the coefficient K which essentially reflects the nonlinear property of the system.

If the function $\Phi^{\circ}(x^{\circ})$ cannot be linearized in the usual way, and if it is desirable to investigate the nonlinear oscillations of the slowly changing component on the basis of the nonlinear Equation (4.7), then one can perform the harmonic linearization of the nonlinear function $\Phi^{\circ}(x^{\circ})$. This will be a repeated harmonic linearization over a new, lower frequency. In this manner one actually investigates nonlinear free oscillations (if $f = 0$) of two frequencies, or mixed oscillations (if f

is periodic with a low frequency).

In an analogous manner one can investigate processes in nonlinear systems with forced oscillations, and also random processes. This makes it possible to take into consideration the effect of vibrational and random interferences on the quality of work of automatic control systems

The solution of the listed problems with examples is described in [3].

The practical effectiveness of this type of method holds great promise for the further development of the applied as well as of the mathematical side of problems in automatic control systems with restrictive conditions. The idea of the introduction of a small parameter, presented in Section 2, may, seemingly, be fruitful also in the more complex problems indicated here. However, one may have to take recourse to certain auxiliary considerations which may take account of the peculiarities of these problems.

BIBLIOGRAPHY

1. Aizerman, M.A., Fizicheskie osnovy primeneniia metodov malogo parametra (Physical bases for the application of the methods of a small parameter). *Avtomatika i telemekhanika* No. 5, 1953.

2. Popov, E.P., O primenenii metoda garmonicheskoi linearizatsii v teorii regulirovaniia (On the application of the method of harmonic linearization in control theory). *Dokl. Akad. Nauk SSSR* Vol. 106, No. 2, 1956.

3. Popov, E.P. and Pal'tov, I.P., *Priblizhennye metody issledovaniia nelineinykh avtomaticheskikh sistem (Approximate Methods of the Investigation of Nonlinear Automatic Systems)*. Fizmatgiz, 1960.

Translated by H.P.T.