УДК 621.372

BUILDING AND ANALYSIS OF STABILITY THE COMBINED OF NUMERICAL METHODS WITH MINIMAL ERROR OF DISCRETIZATION

V.M. Zaiats

National University of water management and environmental engineering, Ukraine, Rivne;
University of Science and Technology in Bydgoszcz, Poland

An iteration approach to the construction of second-order numerical methods based on the Linigura-Willabi method with minimal discretization error is suggested. The essence of the approach is to identify corrections to the Euler's explicit and implicit method at a time when their contributions are equivalent. Based on an iteration approach, we have proposed a new class of difference equations, which improves the accuracy of the calculation of dynamic (oscillatory) systems compared to the trapezoidal rule. A new formula is called a modified trapezoidal rule (MTR) and there is no error of discretization with the precision of second-order terms. This significantly broadens the area of convergence of the calculation of oscillatory systems. The improving time and accuracy calculation of the characteristics in the calculation process in the quartz oscillators 9th order and the oscillatory system 18th order with high Q level and long transients have been confirmed.

The conditions A - stability of the constructed class of methods on the example of a conservative system without losses are established.

Introduction. In work a new numerical method is offered for forming of the discrete equations during conducting of analysis of systems oscillation both autonomous, and with external forces. A discrete method is built such, that for obtain the next discrete point of state variables on integral curves, that answers the continuous model of process oscillation, need enough limited size of integration step at condition that a previous discrete point is found exactly.

Expedience of application of the developed method is shown to the analysis of quartz generators with high-Q quality [1, 2] and for generators with the long transitional processes [4-8, 13], that work in the regime of discrete time and for objects with difficult dynamic nature of such, as systems of recognition of computer users that based on the probabilistic approach [14, 18-19].

Building the main formula discretization. Among a plenty of implicit numerical methods, which found wide application at the analysis of the hard systems (implicit methods Eylera, Rounge-Coutta, many-stepping methods Adamsa-Moultona, Gira, formulas of differentiation backwards, that based on application of Legandra polynomials [3, 9-14, 17-22]) there are methods with the variable step of calculations, that enable to realize by main advantage of implicit numerical methods in comparative with explicit numerical methods – changing of step in wide scopes at finding of point of phase space in areas with a different rate of movement [23-25]. Methods higher than the second order of complication, practically, are not used. This is due with growth of calculable complication of algorithms, and from the second side that the methods of the second order allow to study dynamics of the system of any nature and difficulty of conduct.

The analysis of literary sources confirms, in considerable part of the applied researches of dynamics of the hard systems is used the method of trapezoids, discrete formula of which has a kind [15]:

$$x[n+1] = x[n] + \frac{h}{2}\cdot(f[x[n],t_n] + f[x[n+1],t_{n+1}]) \quad (1)$$

where x[n], x[n+1] – value of state variables in the n and n+1- points of discretization; $f[x[n],t_n]$ $f[x[n+1],t_{n+1}]$ – value of right parts of the differential equation of the explore system, which written in normalized form Koshy.

A formula (1) is averaging amendments for implicit and explicit of methods Eyler for calculate the next point of discretization, that determined in the moment of time on half size step integration. She coincides with the method of Adamsa-Bashforta first order which near to the implicit method Heyna, in which the exact meaning of derivative at the end of interval of calculations is transferable on the close value of derivative, calculated in the middle of step of integration.

With the purpose of increase of exactness of method of trapezoids expediently to modify him thus, to provide the exact hit in the n+1 point of discrete during conducting of integration on condition that the values of state variables in a n point are calculated exactly. This it is possible to attain, if in (1) the step of integration is variable, and derivative (right parts of the differential equation) calculated in such moment on interval of h, that amendments of implicit method Eylera (second element of right part of formula (1)) and explicit method Eylera (third element of right part of formula (1)) are equal. Thus, we build a discrete method in a kind [15]:

$$x[n+1] = x[n] + h\cdot(\mu\cdot f[x[n],t_n] + (1-\mu)\cdot f[x[n+1],t_{n+1}]) \quad (1)$$

where

$$0 < \mu < 1.$$

On the basis of equation of two tangents, conducted on one curve in different moments of time, we can to define the value of parameter in that moment when amendments to implicit and explicit methods Eylera in a formula (2) are equivalent. Thus, from condition

$$\mu\cdot f[x[n],t_n] = (1-\mu)\cdot f[x[n+1],t_{n+1}]$$

we get that

$$\mu = \frac{f[x[n+1], t_{n+1}]}{(f[x[n+1], t_{n+1}] + f[x[n], t_n])}.$$

Consequently, we get a numerical discrete formula for construction of discrete models of the systems with different areas of slow and rapid of motions:

$$x[n+1] = x[n] + h \cdot \frac{2 \cdot f[x[n+1], t_{n+1}] \cdot f[x[n], t_n]}{(f[x[n+1], t_{n+1}] + f[x[n, t_n])} \qquad (3)$$

The results of computer design confirm a right to existence of such method what gives adequate recreation of descriptions of the set mode of oscillatory circuits without the losses, of generators with high quality, quartz generators with protracted transitional processes [2, 14].

The negative in the resulted method, is comparatively with the method of trapezoid, in growth of algorithmic and computational complexity (almost in two times). Advantage in that at the comparable accuracy with the method of trapezoids on one step of integration, the developed method substantially extends the region of convergence of calculable procedures. In result at the calculation of generator with the protracted transitional processes, which are described by the differential equation of high orders, is necessity of application of methods of acceleration of search of the steady modes [4-7]. For providing of convergence of process of deductions, at the use of the discrete method (3), is possible to do without such procedure. Application of method (3) provides winning in time of calculation in 3 – 10 times in comparatively with the method of trapezoid and depends from dimension of equations that describe model generator and it's of quality.

As for the construction of the formula (3) the contribution of each of the Euler method does not exceed half the distance between x_n and x_{n+1}, thus the method (3) provides guaranteed limits on the error of discretization at each step and ensures its with positive sign.

Building the combined of numerical methods. The analysis of error calculation on the example of the conservative model of second order (on example oscillation circuit without losses) confirmed that the error of calculation method (3) is proportional to $h^2/24$, as in the method of trapezoids, but has the opposite sign and twice more the absolute value [16]. After used for the first half step of formula (3), and then formula (1), obtain discrete formula:

$$x_{n+1} = x_n + \frac{h \cdot f_n \cdot f_{n+1}}{(f_n + f_{n+1})} + \frac{h}{4}(f_n + f_{n+1}) \qquad (4)$$

named as a difference combination of the first kind (C1K). The error of calculation when using (4) to the conservative system second order was two times lower compared with the method of trapezoids and opposite in sign. Now, after averaging (1) and (4) we obtain difference combination of the second kind (C2K):

$$x_{n+1} = x_n + \frac{h \cdot f_n \cdot f_{n+1}}{2 \cdot (f_n + f_{n+1})} + \frac{3 \cdot h}{8}(f_n + f_{n+1}). \tag{5}$$

5As shown the results of the analysis error calculation method (5) on the model without loss, then it was 4 times less than the error of the method of trapezoids and two times less than the error of method (4). This error in the C2K coincides with the sign error in the method of trapezoids and opposite with the sign to the error C1K. Thus, we can expect further decrease of the error calculation of methods (5) and (6), which leads to a difference combination of third kind (C3K):

$$x_{n+1} = x_n + \frac{3 \cdot h \cdot f_n \cdot f_{n+1}}{4 \cdot (f_n + f_{n+1})} + \frac{5 \cdot h}{16}(f_n + f_{n+1}). \tag{6}$$

Note that consider a combination of (6) to (4) inappropriate (although she has a right to exist), because (5) has in 4 times smaller of error calculation, compared to (4). In addition the signs of errors in (4) and (6) coincide. The proposed combination of difference schemes constructed so, that in combinations of odd genus (C1K, C3K) is more significant contribution to the second member in the derived formulas, compared with the third, but in even combinations of type (C2K) these deposits are almost equivalents. This construction provides a change of sign of error in obtaining the new combination. Thus, we can construct a метод of any higher order, which will provide up to the members of the second order smallness arbitrarily small error. After arithmetic averaging (5) and (6) we arrive at the difference scheme fourth generation (C4K):

$$x_{n+1} = x_n + \frac{5 \cdot h \cdot f_n \cdot f_{n+1}}{8 \cdot (f_n + f_{n+1})} + \frac{11 \cdot h}{32}(f_n + f_{n+1}) \tag{7}$$

Analyzing formula (4) - (7), for k-th step at using on half-step steam combination and half-step odd, we obtain the difference scheme for the combination of *k*-th family (CKK):

$$x_{n+1} = x_n + \frac{a_k \cdot h \cdot f_n \cdot f_{n+1}}{(f_n + f_{n+1})} + a_{k+1} \cdot h \cdot (f_n + f_{n+1}), \tag{8}$$

where

$$a_k - \frac{2^k - (-1)^k}{3 \cdot 2^{k-1}}; \qquad a_{k+1} = \frac{2^{k+1} + (-1)^k}{3 \cdot 2^{k+1}}.$$

Obviously, with increasing k the coefficients a_k and a_{k+1} in decrease, leading to a decrease of error calculation.

Thus, the error of calculation any *k*-th combinations can be estimated by the formula:

$$\delta = \frac{(-1)^k}{2^{k+1}}, \tag{9}$$

that confirming the analysis of conservative second-order systems, systems of high order and with high quality factor.

Optimization method with minimal error of discretization. In order to minimize errors in (8) we will make the passage to the limit, directing k to infinity. Thus, we get the difference scheme (10). In her with accuracy to members of the second-order of smallness, the errors are absent:

$$x_{n+1} = x_n + \frac{2 \cdot h \cdot f_n \cdot f_{n+1}}{3 \cdot (f_n + f_{n+1})} + \frac{1}{3} \cdot h \cdot (f_n + f_{n+1}). \tag{10}$$

The conclusion that there is no error of difference scheme (10) that is the modification of method trapezoid (MMT) follows from formula (9), if in its parameter k goes to infinity.

The obtained results of error evaluation for difference schemes (5 - (8) and (10) confirmed at the modeling steady-state regimes of the generator Van der Pol [1, 4, 7]. In that model is introduced cubic nonlinearity to delay the transition process. The above analytical error of calculations (9) fully confirmed our results of computer simulation for different values of the parameters of the generator Van der Pol.

The analysis high quality generator that described by a system of differential equations of order 18 at the choice 200 points in every steep, when we used method trapezoids (1) or method (10), that provide normal mode search steady state. At that the use of (1) requires twice more time consuming than (10). With increasing step twice, trapezoid method does not provide the convergence process of deduction, while MMT provides a deduction of the exact value of state variables and frequency fluctuations. A good correspondence between auto vibration system and its discrete model is stored and at the choice of 45 points using the difference scheme MMT.

Analysis of stability the combined of numerical methods. To perform the stability analysis of the obtained difference schemes, we use a conservative second-order system of the form (11):

$$\frac{d^2\{}{dt^2} = -\omega_0^2 \{ \tag{11}$$

For analysis, it is more convenient to bring it to the normal Cauchy form in the form of two first-order equations:

$$\frac{dx}{dt} = x^1; \qquad \frac{dx^1}{dt} = -\omega^2 \cdot x. \tag{11a}$$

After applying (3) to equations (11), we proceed to difference equations

$$x_{n+1} = x_n + \frac{2 \cdot h \cdot x_{n+1}^1 \cdot x_n^1}{x_{n+1}^1 + x_n^1}$$

$$and \qquad x_{n+1}^1 = x_n^1 - \frac{2 \cdot h \cdot \omega^2 \cdot x_{n+1} \cdot x_n}{x_{n+1} + x_n}, \tag{11b}$$

the solution of which $x_{p,}$ approximated to the solution (3a) has the form:

$$\begin{bmatrix} x_{n+1} \\ x^1_{n+1} \end{bmatrix} = \mathbf{C} \cdot \boldsymbol{\rho} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}^n ,$$

where ρ - the vector column of the multipliers of the difference system (11b); ***C*** is a matrix of eigenvalues of eigenvectors of system (11b). It is easy to see that the multipliers of the discrete model (11b) are determined by the eigenvalues of the continuous system (11). Instead of the second-order system, it is appropriate to analyze immediately two equations of the first order written with respect to eigenvalues (11):

$$\frac{dx}{dt} = j \cdot \omega \cdot x \quad i \quad \frac{dx}{dt} = -j \cdot \omega \cdot x , \tag{11c}$$

to which the multipliers of the discrete system correspond (11b)

$$\rho_1 = j \cdot \omega \cdot h + \cdot\sqrt{(1-\omega^2 \cdot h^2)}$$
$$i\, \rho_2 = -j \cdot \omega \cdot h + \cdot\sqrt{(1-\omega^2 \cdot h^2)}$$

Since the elements of the matrix ***C*** are determined by the initial conditions that significantly affect the area of convergence of the calculation process, and the sampling error of the numerical method does not depend on them, it is sufficient to evaluate the modules and arguments of the multipliers:

$$|\rho_1| = |\rho_2| = 1 + h^4 \cdot \omega^4 \quad i \quad \varphi_1 = -\varphi_2 = \operatorname{arctg} \frac{\omega \cdot h}{(1 - h^2 \cdot \omega^2 / 2)} .$$

Thus, there is no error in determining the amplitude of oscillations up to the members of the second order of smallness when using method (4). Given that the approximate magnitude of the oscillation period can be determined from equality

$$T_p = \frac{2\pi \cdot h}{\varphi} ,$$

the relative magnitude of the error in determining the period of oscillation is:

$$\delta T = \frac{T_0 - T_p}{T_0} = \frac{h^2 \omega^2}{6} .$$

Compared to the trapezoid method (2) from [3], this value is twice as large in absolute value, but has the opposite sign. For this combination, the multipliers determined by sampling equations (11c) using C1K (6) have the form:

$$\rho_1 = \frac{6 \cdot j \cdot \omega \cdot h + 8 \cdot \sqrt{(1-\omega^2 \cdot h^2/2)}}{8 \cdot (1 - j \cdot \omega \cdot h/4)} \quad i \quad \rho_2 = \frac{-6 \cdot j \cdot \omega \cdot h + 8 \cdot \sqrt{(1-\omega^2 \cdot h^2/2)}}{8 \cdot (1 + j \cdot \omega \cdot h/4)}$$

Such,

$$|\rho_1| = |\rho_2| = 1 \qquad i \quad \varphi_1 \quad = -\varphi_2 = \text{arctg}\frac{3\cdot\omega\cdot h\cdot\sqrt{(1+h^2\cdot\omega^2/2)}}{4} + \text{arctg}\frac{\omega\cdot h}{4}.$$

Since the modules of these multipliers are equal to one, there is no error in determining the amplitude of oscillations. The sampling error in using (6) for the conservative system (3a) is twice smaller and opposite in sign than in the trapezoid method, that is, h2ω2 / 24. Note that this combination has the property A - stability, since the multiplier modules do not go beyond the circle of a single radius.

As the results of estimating, the sampling error of method (5) showed when considering the lossless model (11), it proved to be 4 times smaller than the error of the trapezoid method and twice less than the error of method (4). Thus, for method (5) δT = –h2ω2 / 48..In this case, the error sign in C2K coincides with the error sign of the trapezoid method and has the opposite sign to the error given by C1K. As in the previous combination (5), it has the property **A** - stability.

We note that all obtained difference formulas (5) - (8), (10) for the discretization of continuous systems have the property A - stability, which makes it impossible to accumulate the sampling error for long transient processes, which are characteristic of dynamic systems with high quality. This result is confirmed by the calculation of quartz generator schemes and high-quality generator circuits with long transient processes [7, 16].

Conclusions. In the analysis of conservative systems with no energy loss, the calculation of wireless devices with high quality factor, analysis of systems with long transient finding periodic modes in systems with complex dynamics it is advisable to use the discrete formula (10), which in the second order of smallness provides no error in determining both the amplitude and period of oscillations.

Using the difference scheme (10) in the construction of discrete models in applications of computer analysis of electronic systems will increase their effectiveness and will ensure the reliability of their work.

In the future, we intend to rigorously prove the property of A-stability MMT (10) and show the feasibility of its application to the analysis of oscillating processes dynamic systems of high dimensionality.

REFERENCES

1. V.M. Zaiats, "Constraction and analises of model descrete oscillation systems", Cybernetics and Systems Analysis, vol. 2,, Kiev, NAS of Ukraine, 2000, pp. 161-165.
2. V.M. Zaiats, "The Models Descrete Oscillation Systems", Computer Tegnology of Print, Lviv,UAP, 2011, pp.37-39.
3. J. Vlach and K. Singhall, Computer Methods for Circuit Analysis and Design", New York, VNRC, 1980, 560 p.
4. V.M. Zaiats, "Amplitude and frequency errors of autooscillation of Linigur-Whillaby numerical integration formula", Theoretical Electrical Engineering ,vol. 37, Lviv, LNU by Franko, 1984, pp. 83-88.
5. L.A. Synitsky and U.M. Shumkow, "Analysis periodical ragimesin nonlinear cirquits by numerical methods", Theoretical Electrical Engineering vol. 9, 1970, pp. 110-115.

6. T. Aprille and T. Trick.”A computer algorithm to determine the steady-state response of nonlinear oscillators”, IEEE Trans. on Circuit Theory, 1972, V. CT-19,N 4, pp. 354-360.
7. V.M. Zaiats, “Fast search of established modes of high-frequency oscillators with long transients”, Proceedings Universities. Radioelektronika, Kiev, vol. 3, 1993, pp. 26-32.
8. L.A. Synitsky and O. Yo. Felyshtyn “Accelerated Search of Periodical Regimes in Autonomous Systems”, Theoretical Electrical Engineering,vol. 39, Lviv, LNU by I. Franko,1985, pp. 93-103
9. L. Feldman, A. Petrenko and O. Dmitrieva, Numerical Methods in informatic, Kiev, BHV, 2006, 480 p.
10. G. Holl and J.M. Watt, Modern Numerical Methods for Ordinary Differential Equations,. Clarendon press, Oxford, 1979, 312 p.
11. D. Forsait, H.Malkom, K. Mouler, Computer Methods of Mathematics Calculations, Moscow, Peace,1980, 282.p.
12. X. Shtetter, Analises Descretization for Original Differencial Equations, Moskow, Peace,1978, 462 p.
13. V.M. Zaiats, Discrete Models of Oscillatory Systems for the Analysis of their Dynamics, Lviv,UAP, 2011, 284 p.
14. V.M. Zaiats, Methods, Algorithms and Software for Simulation and Analysis of the Dynamics of Complex Objects and Systems Based on Discrete Models, Lviv, New World, 2009, 400 p
15. Liniger W. Efficient integration methods for stiff systems of ordinary differential equation / W. Liniger, R. Willoughby // SIAM J. Numer. Anal. 1970, V. 7, №1. P. 47-66.
16. Zaiats V.M. Iterative approach to minimize errors of the second order numerical methods and their application to the analysis of nonlinear dynamical systems / V.M. Zaijats // Journal “ Reports National Academy Sciences Ukraine”, 2013, № 8. P. 33-37.
17. L..A, Sinitsky, J.A. Shmyhelskyy. “Synthesis periodic self-excited oscillator with multiple modes with different frequencies and changes in the form of vibrations”, Theoretical Electrical Engineering. – 2004, vol . 57, pp.153-158.
18. Zaiats V. “Dynamical regimes classification of discrete oscillatory system“. Pros. International Conf. “The experience of designing and application of CAD systems in microelectronics”, Lviv, Slavske, 2003, pp. 157-158.
19. V. Zaiats “Comparative characteristic objects and regimes of discrete nature models”. Proc. International conf. TCSET’2004, Lviv, 2004, pp. 84-85.
20. J. Vlach and K. Singhall, Computer Methods for Circuit Analysis and Design”, New York, VNRC, 1980, 560 p.
21. V.M. Zaiats, “Amplitude and frequency errors of autooscillation of Liniger-Willouhby numerical integration formula”, Theoretical Electrical Engineering ,vol. 37, Lviv, LNU by Franko, 1984, pp. 83-88.
22. V. Zaiats. Numerical methods for analysis of linear and nonlinear systems. Textbook / V. Zaiats, B. Durnyak, I. Jaworskiy // Lviv: Publishing UAH, 2009.- 302 p.
23. Zaiats V.M. Combinational numerical methods with a minimum discretization error /V. Zaiats // Cybernetics and Systems Analysis, 2013, No. 2. P. 115-120.v
24. Zayats Vasyl M. Approach to Working Features Dictionary Construction Based on Defined Priorities of the Primary Features / Vasyl M. Zayats, Galyna Ya. Shokyra // Telecommunications and Electronics. N 17 (262), 2013, pp. 27-36.
25. V. Zaiats. Methods and means of computer information technologies in applied applications. Textbook / V. utp.edu.plZaiats // Lviv: Publishing “Ukrainian technologies”, 2017. 262 p.

DOI: 10.32403/2411-9210-2019-2-42-48-56

ПОБУДОВА І АНАЛІЗ СТІЙКОСТІ КОМБІНОВАНИХ ЧИСЛОВИХ МЕТОДІВ ДРУГОГО ПОРЯДКУ З МІНІМАЛЬНОЮ ПОХИБКОЮ ДИСКРЕТИЗАЦІЇ

В.М. Заяць

Університет Технологічно-природничий, Бидгощ, Польща
Національний університет водного господарства та природокористування, Рівне, Україна
vasyl.zaiats@utp.edu.pl, vasyl.zayats@nuwm.edu.ua

Запропоновано ітераційний підхід до побудови числових методів другого порядку на основі методу Лінігера-Віллабі з мінімальною помилкою дискретизації. Суть підходу полягає у визначенні поправок до явного та неявного методів Ейлера в той момент, коли їх внески є рівнозначними. На основі ітераційного підходу ми запропонували новий клас різницевих рівнянь, що підвищує точність обчислення динамічних (коливальних) систем порівняно з правилом трапеції. Нова формула називається модифікованим трапецієподібним правилом (МТП) і вона немає помилок дискретизації з точністю до членів другого порядку малості. Це значно розширює область збіжності обчислення характеристик коливальних систем. Підтверджено поліпшення часу та точності розрахунку характеристик у процесі аналізу кварцових генераторах 9-го порядку та автоколивальної системі 18-бротностіго порядку з високим рівнем Q (добротності) та тривалими перехідними процесами.

Встановлено умови A - стійкість побудованого класу методів на прикладі консервативної системи без втрат.

Стаття надійшла до редакції 25.04.2019.
Received 25. 04.2019.