UDC 618.318

V. M. Zayats

TWO APPROACHES TO BUILDING OF NUMERICAL METHODS SECOND ORDER FOR ANALYSIS DYNAMICAL SYSTEMS

The two approaches to minimize of error discretization for numerical methods second order, which is based on a modification of the method of trapezoids and setting the time when deposits explicit and implicit Euler methods have the same contribution to the amendment to the next discretization point of dynamical system. The expediency of its application to the analysis of nonlinear dynamical systems of oscillatory nature with a high quality factor and long transient's processes.

Keywords: *error discretization, numerical methods, stability, dynamical systems.*

Запропоновано два підходи до мінімізації похибки дискретизації для чисельних методів другого порядку, що ґрунтуються на модифікації методу трапецій і визначенні моменту часу, коли вклади явного і неявного методів Ейлера мають однаковий внесок у поправку до наступної точки дискретизації динамічної системи. Розглянуто доцільність їх застосування для аналізу нелінійних динамічних систем коливного характеру з високою добротністю і тривалими перехідними процесами.

Ключові слова: *похибка дискретизації, числові методи, стійкість, динамічні системи.*

In the analysis of complex dynamic processes and phenomena that can be described by a system of continuous differential equations presented in the normal Koshy form:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}[\mathbf{x}(t).\ t]\ ,$$

where $\mathbf{x}$ – N-dimensional vector of state variables; $\mathbf{f}$ – N-dimensional vector function describing the dynamics of the phase trajectories of the system, using numerical (difference) methods for discretization. Such methods should be convergent and have a small sampling error for the preservation of qualitative and quantitative correspondence between the investigational process or phenomenon and its discrete model [1–5]. The second requirement for difference methods – a property **A** – stability. Otherwise, the presence of a small truncation error deduction admitted one step can lead to the accumulation of errors during movement depicting points along the phase trajectory and complete unfitness for applied uses the final result deductions [6, 7].

The programs for computer analysis of electronic circuits [7], the analysis of the behaviour of systems with complex dynamics [3, 4], the analysis of oscillatory systems with high Q-values for which the transition is long [6], there is a problem between the complexity difference algorithm and its accuracy. Typically, using methods not higher than second order of complexity or these combinations. In particular, the commonly used method of trapezoids [6, 7]. Difference formula for this method is:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{h}{2}(\mathbf{f}_n + \mathbf{f}_{n+1})\,. \tag{1}$$

This formula is a combination of two methods: the first half of the sampling step using an explicit Euler method and in the second half of the implicit Euler method [8]. As a result of the construction of such a combination, as indicated by numerous publications accuracy increases more than an order of magnitude compared to the Euler method. In addition, this method is characterized by the property of **A** – stability, which is confirmed by calculation of generator circuit with high Q-factor [6] and long transients [9].

A method of minimizing error rate. In [7] proposed to consider amendments to the next sampling point is the middle step, but at the time when deposits explicit and implicit Euler's methods are equivalent. For this purpose the difference Eq. (1) are presented in the form proposed by Liniher–Wiulaby [10]:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h(1-\mu)\mathbf{f}_n + h \cdot \mu \cdot \mathbf{f}_{n+1}, \tag{2}$$

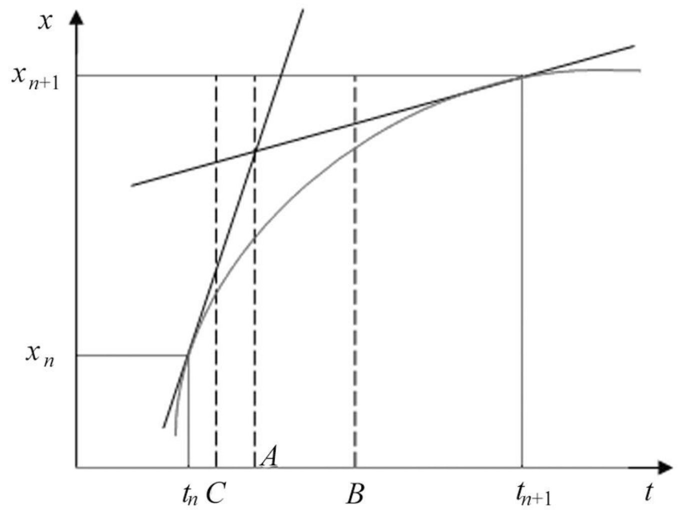
that when $\mu = 0$ corresponds to the explicit Euler method; $\mu = 0.5$ – method of trapezoids; $\mu = 1$ – implicit Euler method. Equating the second and third terms on the right side of Eq. (2) obtained the value of μ at which the explicit and implicit Euler methods for making the same contribution to the amendment to the value x_n:

$$\mu = \frac{\mathbf{f}_n}{\mathbf{f}_n + \mathbf{f}_{n+1}} \ . \tag{3}$$

After substituting (3) in (2) received a new difference formula:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{2 \cdot h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{(\mathbf{f}_n + \mathbf{f}_{n+1})}. \tag{4}$$

Because of the construction of the formula (4) the contribution of each of the methods of Euler's less than half the distance between h_n and h_{n+1}, then the method (4) gives a guaranteed limit on the amount of sampling error at each step and provides its positive.



Geometric interpretation of approach iterative to minimize the error rate.

Geometric illustration of the proposed method of reducing sampling error is illustrated in Figure. If contribution for explicit and implicit Euler method to the next sampling point to consider at the moment that corresponds to point *C* (see Figure), we get the trapezoid method, in point *B* – we have proposed a method that made equivalent contribution Euler's methods, at point *A* – the optimum combination is obtained, which corresponds to the point of intersection of tangents to x_n and $x_n + 1$ points rate.

To estimate the error of the method (4) an analysis of sampling error on the example of the conservative model of second order

$$\frac{d^2\mathbf{x}}{dt^2} = -\omega_0^2 \mathbf{x} .$$

confirmed that sampling error method (4) is proportional to h2/24, as in the method of trapezoids, but has the opposite sign and twice the smaller absolute value. Studies have shown that the method (4) as a method of trapezoids has the property of *A*-stability.

Iterative approach to minimize the error rate. Given that the error of the method (4) and the method of trapezoids (1) have opposite signs, can hold their arithmetic averaging, thus reducing the value of error. Applying the first half step (4) and the second (1), we obtain the difference formula

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{(\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{h}{4}(\mathbf{f}_n + \mathbf{f}_{n+1}), \tag{5}$$

what is the name of difference combination of the first kind (C1K). Sampling error when using (5) to the conservative system was two times less than the method (4) and opposite in sign with respect to the method of trapezoids. Now, after averaging (1) and (5) we obtain the difference combination of the second kind (C2K):

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{2 \cdot (\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{3 \cdot h}{8}(\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{6}$$

As shown sampling error analysis method (6) when considering the model without loss, it was 4 times less than the error of the method of trapezoids and two times smaller than the error of the method (5). The sign error in C2K coincides with the sign of the error in the method of trapezoids and opposite to the error that gives C1K. Thus, we can expect further reduction of the sampling error of a combination of methods (5) and (6), which leads to difference combinations of the third kind (C3K):

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{3 \cdot h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{4 \cdot (\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{5 \cdot h}{16}(\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{7}$$

Note that the considered combination of (6) with (4) inappropriate (although she has a right to exist), since (5) is 4 times smaller error rate, compared to (4). In addition, signs of error in (4) and (6) coincide.

The winning combination for minimizing error rate. The proposed combination of difference schemes constructed in such a way that the combination of odd genus (C1K, C3K) is more significant is the contribution of the second term in the formulas obtained, compared with the third, and combinations of even genus (C2K) these deposits are virtually aligned. This structure provides a change of sign of error in obtaining new combination. Thus, we can construct a second-order method, which will provide up to the members of the second order of smallness arbitrarily small error rate. After arithmetic averaging (6) and (7) we arrive at the difference scheme fourth generation (C4K):

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{5 \cdot h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{8 \cdot (\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{11 \cdot h}{32}(\mathbf{f}_n + \mathbf{f}_{n+1}) \tag{8}$$

Analyzing the Eqs. (5)–(8), the *k*-th step, half step using a combination of steam and the odd half-step, we obtain the difference scheme for a combination of *k*-th kind (CKK):

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{a_k \cdot h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{(\mathbf{f}_n + \mathbf{f}_{n+1})} + a_{k+1} \cdot h \cdot (\mathbf{f}_n + \mathbf{f}_{n+1}), \tag{9}$$

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* values of the coefficients a_k and a_{k+1} are reduced, leading to a reduction in the error rate.

In this case, the error discretization any k-th combination can be calculated by the equation

$$\delta = \frac{(-1)^k}{2^{k+1}}, \tag{10}$$

confirming the analysis conservative second order systems and systems with a high figure of merit of high order.

In order to minimize sampling error in (9) feasible limit switch, sending k to infinity. Get the difference scheme (11) for which to terms of second order smallness sampling error missing:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{2 \cdot h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{3 \cdot (\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{1}{3} \cdot h \cdot (\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{11}$$

The conclusion that there is no sampling error of the difference scheme (11) follows from the equation (9), if it target k to infinity.

Note that all the obtained difference formulas (4)–(9), (11) for continuous sampling systems have the property **A** – resistance that prevents the accumulation of sampling error during long transients that are characteristic of dynamic systems with high quality factor. This result is confirmed by calculation quartz generator devices and high good-quality generator circuits with long transients [6].

Direct search for optimal combinations. In order to obtain direct analytical expression for which the error rate in the first approximate missing, find the coordinates of point A Fig. 1, corresponding to the intersection of tangents

$$\mathbf{x} = \mathbf{x}_0^n + \mathbf{f}_n t \quad \text{and} \quad \mathbf{x} = \mathbf{x}_0^{n+1} + \mathbf{f}_{n+1} t$$

conducted in two adjacent n and $n+1$ points rate. Equaling right parts in the last two equations, we obtain:

$$t = \frac{\mathbf{x}_0^{n+1} - \mathbf{x}_0^n}{\mathbf{f}_n - \mathbf{f}_{n+1}}. \tag{12}$$

where

$$\mathbf{x}_0^n = \mathbf{x}_n - nh\mathbf{f}_n \quad з \quad \mathbf{x}_0^{n+1} = \mathbf{x}_{n+1} - (n+1)h\mathbf{f}_{n+1},$$

as shown in Figure. On the other hand, since the intersection of tangents corresponding values

$$t = h(n + \mu). \tag{13}$$

Equaling right sides of Eqs. (12) and (13) we find the value of μ, the choice of which deposits explicit and implicit Euler method provides a point of contact with phase n to $n+1$ point sampling:

$$\mu = \frac{\mathbf{x}_{n+1} - \mathbf{x}_n}{(\mathbf{f}_n - \mathbf{f}_{n+1})} + \frac{h\mathbf{f}_{n+1}}{\mathbf{f}_n - \mathbf{f}_{n+1}}. \tag{14}$$

Substituting the value of μ from (14) in Eq. (2) we obtain an optimal combination of second-order numerical method for which no sampling error:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h/2(\mathbf{f}_n + h\mathbf{f}_{n+1}). \tag{15}$$

For algorithmic complexity (15) simpler method (4) and slightly inferior method (1), while ensuring minimum error deductions associated only with precision representation of numbers in the environment deductions.

Note that all the resulting difference formulas (5)–(9), (11), (15) for continuous sampling systems have property A – stability, which prevents the accumulation of error rate during long transients that are characteristic of dynamic systems with a high quality factor. This result confirmed the expectation quartz generator devices and high good-quality generator circuits with long transients [6].

1. *Анализ* точности и качественного соответствия дискретных моделей электрических цепей / В. М. Бондаренко, И. И. Герасымив, Б. А. Мандзий, А. В. Маранов. – К., 1983. – 44 с. (Препр. НАН Украины. Ин-т электродинамики, № 307).
2. *Бутенин Н. В., Неймарк Ю. И., Фуфаев Н. А.* Введение в теорию нелинейных колебаний. – М.: Наука, 1987. – 384 с.
3. *Васильев В. И., Шевченко А. И.* Комбинированный алгоритм оптимальной сложности // Пр. Міжнар. конф. "Штучний інтелект". – Крим, 2002. – Т. 1. – С. 308–310.
4. *Заяць В. М.* Построение и анализ дискретной модели дискретной колебательной системы // Кибернетика и системный анализ. – 2000. – № 4. – С. 161–165.
5. *Заяць В. М.* Аналіз динаміки та умов стійкості дискретних моделей коливних систем // Вісник НУ "Львівська політехніка" "Інформаційні системи та мережі". – 2004. – № 519. – С. 132–142.
6. *Заяць В. М.* Ускоренный поиск установившихся режимов в высокочастотных автогенераторах с длительными переходными процессами // Известия вузов. Радиоэлектроника. – 1993. – № 3. – С. 26–32.
7. *Чуа Л. О., Пен-Мин Лин.* Машинный анализ электронных схем (алгоритмы и вычислительные методы). – М.: Энергия, 1980. – 640 с.
8. *Заяць В. М.* Побудова комбінованих різницевих методів другого порядку // Зб. праць наук.-техн. конф. "Обчислювальні методи і системи перетворення інформації", Львів, 7–8 жовтня 2011. – Львів: Фіз.-мех. ін-т ім. Г. В. Карпенка НАН України. – 2011. – С. 34–36.
9. *Aprille T., Trick T.* A computer algorithm to determine the steady-state response of nonlinear oscillators // IEEE Trans. on Circuit Theory. – 1972. – **CT-19**, № 4. – P. 354–360.
10. *Liniger W., Willouby B.* Efficient integrative methods for still systems off ordinary differentional equations // Num. Anal. – 1972. – № 7. – P. 47–66.

National University "Lviv polytechnic", Ukraine

Received 02.07.2013