https://doi.org/10.31489/2024M1/21-27 *Research article*

On some estimations of deviations between real solution and numerical solution of dynamical equations with regard for Baumgarte constraint stabilization

M.I. Akylbayev[1,*], I.E. Kaspirovich[2]

[1] *Peoples Friendship University named after academician A.Kuatbekov, Shymkent, Kazakhstan;*
[2] *Patrice Lumumba Peoples' Friendship University of Russia, Moscow, Russian Federation*
(E-mail: musabek_kz@mail.ru, kaspirovich_ie@rudn.ru)

The numerical solution of a system of differential equations with constraints can be unstable due to the accumulation of rounding errors during the implementation of the difference scheme of numerical integration. To limit the amount of accumulation, the Baumgarte constraint stabilization method is used. In order to estimate the deviation of real solution from the numerical one the method of constraint stabilization can be used to derive required formulas. The well-known technique of expansion the deviation function to Taylor series is being used. The paper considers the estimation of the error of the numerical solution obtained by the first-order Euler method.

Keywords: constraint stabilization, numerical integration, stability, dynamics, system of differential levels, numerical methods, numerical solution, difference scheme, rounding.

2020 Mathematics Subject Classification: 65D30.

Introduction

The description of the dynamics of the system using Hamilton or Lagrange formalisms assumes the solution of differential equations or a qualitative study of their properties [1]. It is not always possible to obtain analytically the solution of systems of differential equations. Therefore, it is necessary to resort to numerical integration methods [2] or to methods of investigating the properties of solutions using methods of the qualitative theory of differential equations [3].

The use of numerical methods for solving differential equations is associated with the inevitable accumulation of numerical integration errors. Therefore, the result of the numerical solution reflects the real picture only with some degree of accuracy. The fact is that the implementation of one or another difference scheme of numerical integration is accompanied by the accumulation of numerous errors, in particular rounding errors.

Baumgarte showed [4] that the classical method of determining the reactions of contact constraints used in mechanics leads to an inevitable accumulation of numerical integration errors associated with an increase in the values of deviations from the constraints equations caused by errors in setting the initial conditions. To reduce these deviations, Baumgarte proposed using linear combinations of constraints equations together with their derivatives. The equations that establish the relationship between linear combinations of constraints and their derivatives are called the equations of perturbations of constraints. In essence, the Baumgarte method boils down to replacing the constraints equations with servo constraints equations. The method of bond stabilization proposed by Baumgarte proved popular and

*Corresponding author. *E-mail: musabek_kz@mail.ru*

The research was funded by the Russian Science Foundation and Moscow city project No. 23-21-10065, https://rscf.ru/project/23-21-10065/.

Received: 22 November 2023; *Accepted:* 13 December 2023.

caused the emergence of various modifications. Thus, Yu. Ascher proposed a method for stabilizing systems of higher-order differential algebraic equations with constraints [5].

The conditions imposed on the behavior of solving a system of dynamic equations with deviations from the constraints equations leads to additional requirements for determining the constraints reactions. For these purposes, the concept of program constraints was introduced.

The first-order Euler difference scheme is the simplest scheme for numerical integration of systems of first-order differential equations. When integrating, the area under the curve is searched for as an area collections of rectangles. Any set of rectangles of finite length will not be able to completely cover the area of a curved trapezoid, so the numerical solution of the integrable equation does not coincide with the real one. The estimate of the maximum value of this error can be calculated by considering the deviation of the numerical solution from the real one.

In theoretical mechanics there is a specific set of problems that define their goal as constructing a system of ordinary differential equations based on the given properties. Such problems are being called inverse ones. In some cases we need to find the specific constraints equations that provide the system with its requested properties. These constraints are defined as program constraints. Methods for solving systems of differential algebraic equations were investigated in work [6]. If a system contains some ambiguity by its internal random parameters than it can be considered as stochastic. Some inverse dynamical problems for the system with stochastic parameters are considered in papers [7–9]. In some cases the system of motion equations is required to be constructed with regard for Baumgarte constraint stabilization method implemented in it. In papers [10, 11] it was shown that perturbation parameters are connected with dissipative function that pumps energy out of the system. New advanced numerical methods were investigated for inverse-like problems in works [12, 13].

1 Problem Statement

Let the state of a mechanical system be given by the set of generalized coordinates $q = \left(q^1, ..., q^n\right)$. The change in the position of the mechanical system in time implies the dependence of the vector q on time t: $q = q(t)$. The rate of change in the position of the system is determined by the velocity vector: $v(t) = dq(t)/dt = \dot{q}(t) = (\dot{q}^1, ..., \dot{q}^n)$. Let's consider that the system of motion equations is presented in form:

$$\begin{aligned} \dot{q} &= \nu; \\ \dot{\nu} &= a(q, t), \end{aligned} \tag{1}$$

where $a(q, t)$ is a given function. Let's introduce a vector state $x = (q, \nu)$ and rewrite (1) in a matrix form:

$$\dot{x} = F(x, t). \tag{2}$$

Suppose that the motion is restricted and the kinematic state vector $x(t)$ is limited by a set of mechanical constraints described by the equations:

$$h_i(q, t) = 0, \;\; i = 1, ..., m, \;\; m < n. \tag{3}$$

Here and in the future, corresponding to Einstein's notation, repeating indices imply summation by the same indices.

In order to solve system (2) with constraints (3) the method of Lagrange multipliers is used. But during the numerical integration with Euler first order scheme we will inevitably face with solution's instability. To solve this problem J. Baumgarte [1] suggested to consider an arbitrary linear of constraints and its full time derivatives while solving the system of differential algebraic equations with constraints. According to this stabilization method our system will take form:

$$\begin{cases} \dot{x} = F(x, t), \\ \ddot{h} + A\dot{h} + Bh = 0, \end{cases}$$

where A and B are matrices with arbitrary components that are called perturbation parameters. By manipulating the values of these components, we can achieve a stable numerical solution. We can algebraically solve obtained system and derive $\dot{x}$:

$$\dot{x} = X(x + \aleph, t). \tag{4}$$

Symbol $\aleph$ here stands for the terms with perturbation parameters and provides numerical stability.

2 Numerical Integration

In numerical integration, it is assumed that the differentials of the function and the independent argument are represented in finite differences $dx(t) \approx \Delta x(t)$, $dt \approx \Delta t$. Thus, the functions become functions of a discrete argument.

Let equation (4) be determined on a set $[t_1, t_2]$. In the simplest difference schemes, this set can be divided by points $t_1 = t_{(1)}, t_{(2)}, ..., t_{(2)} = t_{(N)}$ and $(N-1)$ equal length segments $\tau = t_{(\alpha+1)} - t_{(\alpha)}$, corresponding to the integration step. Finite increment of a state vector $x(t)$ can be represented as a difference:

$$\Delta x\left(t_{(\alpha)}\right) = x\left(t_{(\alpha+1)}\right) - x\left(t_{(\alpha)}\right), \;\; \alpha = 1, ..., N-1.$$

We use Euler first order difference scheme to solve equation (4) numerically:

$$x_{(\alpha+1)} = x_{(\alpha)} + \tau X_{(\alpha)}, \qquad \alpha = 1, ..., N-1. \tag{5}$$

To estimate the deviation error, the real solution will be denoted $\tilde{x}(t)$. It satisfies the system (2) with constraints (3). It does not include stabilization terms. Consider the deviation of a real solution from a numerical one (5) at the moment $t_{(\alpha)} : \tilde{x}(t_{(\alpha)}) - x_{(\alpha)}$. Let's expend $\tilde{x}(t_{(\alpha)})$ at $t_{(\alpha)}$ to Taylor series:

$$\tilde{x}(t_{(\alpha)}) = \tilde{x}(t_{(\alpha-1)}) + \tau\dot{\tilde{x}}(t_{(\alpha-1)}) + \frac{\tau^2}{2}\ddot{\tilde{x}}(\zeta),$$

where $\zeta : \zeta \in \left[t_{(\alpha-1)}, t_{(\alpha)}\right]$. Taking into account (2) deviation $\tilde{x}(t_{(\alpha)}) - x_{(\alpha)}$ will be written in the form:

$$\tilde{x}(t_{(\alpha)}) - x_{(\alpha)} = \tilde{x}(t_{(\alpha-1)}) - x_{(\alpha-1)} + \tau\left(\tilde{X}\left(t_{(\alpha-1)}\right) - X_{(\alpha-1)}\right) + \frac{\tau^2}{2}\ddot{\tilde{x}}(\zeta). \tag{6}$$

If we apply mean value theorem to the term $\tilde{X}\left(t_{(\alpha-1)}\right) - X_{(\alpha-1)}$ we will obtain the following relation:

$$\tilde{X}\left(\tilde{x}(t_{(\alpha-1)}), t_{(\alpha-1)}\right) - X\left(x_{(\alpha-1)} + \aleph, t_{(\alpha-1)}\right) = \frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\left(\tilde{x}(t_{(\alpha-1)}) - x_{(\alpha-1)} - \aleph\right),$$

where $x_\zeta \in \left[\tilde{x}(t_{(\alpha-1)}), x_{(\alpha-1)}\right]$ or $x_\zeta \in \left[x_{(\alpha-1)}, \tilde{x}(t_{(\alpha-1)})\right]$ depending on which value is greater. $\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right) -$ matrix $[2n \times 2n]$.

Let's denote the deviation $\tilde{x}(t_{(\alpha)}) - x_{(\alpha)} = \Delta_{(\alpha)}$, then the ratio (6) can be rewritten as:

$$\Delta_{(\alpha)} = \left(I_{2n} + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right)\Delta_{(\alpha-1)} - \tau\aleph\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right) + \frac{\tau^2}{2}\ddot{\tilde{x}}(\zeta), \tag{7}$$

where I_{2n} is a unit matrix.

Denote $\Im = \max\limits_{t \in [t_0, t_k]}\left(\frac{\tau}{2}\ddot{\tilde{x}}(\zeta) - \aleph\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right)$. Taking into account the triangle inequality, the ratio (7) will take form:

$$\left|\Delta_{(\alpha)}\right| \le \left|\Delta_{(\alpha-1)}\right|\left|I_{2n} + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right| + \tau\Im.$$

The norm of a vector or matrix by components is understood as the maximum value of the modulus of its components: $\left|\Delta_{(\alpha)}\right| = \max\limits_{k=1,\ldots,2n}\left|\Delta_{k(\alpha)}\right|$. Solving this recursive inequality with respect to the first element, we obtain:

$$\left|\Delta_{(\alpha)}\right| \leq \left|\Delta_{(1)}\right| \left|I_{2n} + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right|^{\alpha-1} + \tau\Im\sum\nolimits_{l=1}^{\alpha-2}\left|I_{2n} + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right|^{l}.$$

Let's assume that the integration step is small enough, so the expression $1 + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)$ is positive $\forall t \in [t_0, t_k]$ even, if the derivative $\frac{\partial X}{\partial x}$ is negative. Also apply to the second term the formula of the sum of a finite number of elements of geometric series, we get:

$$\left|\Delta_{(\alpha)}\right| \leq \left|\Delta_{(1)}\right| \left|I_{2n} + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right|^{\alpha-1} + \Im\frac{\left|I_{2n} + \tau\frac{\partial X}{\partial x}(x_\zeta, t_{(\alpha-1)})\right|^{\alpha-1} - 1}{\left|\frac{\partial X}{\partial x}(x_\zeta, t_{(\alpha-1)})\right|}. \tag{8}$$

As

$$I_{2n} + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right) \leq \exp\left(\tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right),$$

then:

$$\left|I_{2n} + \tau\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right|^{\alpha-1} \leq \left|\exp\left(\tau(\alpha-1)\frac{\partial X}{\partial x}\left(x_\zeta, t_{(\alpha-1)}\right)\right)\right|$$

and $t_k = t_0 + (N-1)\tau,\ \alpha \leq N,\ \tau(\alpha-1) \leq t_k - t_0$. Then the ratio (8) will be written in the form:

$$\left|\Delta_{(\alpha)}\right| \leq \left|\Delta_{(1)}\right| \left|\exp\left(t_k - t_0\right)\frac{\partial X}{\partial x}\left(x_\zeta, t_\zeta\right)\right| + \Im\frac{\left|\exp\left(t_k - t_0\right)\frac{\partial X}{\partial x}\left(x_\zeta, t_\zeta\right)\right| - 1}{\left|\frac{\partial X}{\partial x}\left(x_\zeta, t_\zeta\right)\right|}. \tag{9}$$

The right side of this ratio does not include the node number, so you can also enter the norm for nodes: $\alpha\,||\Delta|| = \max\limits_{\alpha=1,\ldots,N-1}\left|\Delta_{(\alpha)}\right|$. Then the relation (9) allows you to set the ratio for the maximum possible error in numerical integration using the Euler difference scheme:

$$||\Delta|| \leq \left|\Delta_{(1)}\right| \left|\exp\left(t_k - t_0\right)\frac{\partial X}{\partial x}\left(x_\zeta, t_\zeta\right)\right| + \Im\frac{\left|\exp\left(t_k - t_0\right)\frac{\partial X}{\partial x}\left(x_\zeta, t_\zeta\right)\right| - 1}{\left|\frac{\partial X}{\partial x}\left(x_\zeta, t_\zeta\right)\right|}.$$

Conclusion

It follows from this relation that the maximum possible error exponentially depends on the length of the segment on which the integration takes place. Also, the second term of this relation contains the stabilization term $\aleph$ associated with the equations of perturbed constraints. Therefore, a change in the values of the perturbation parameters affects the maximum deviation error during numerical integration. However, due to the arbitrariness of the type of functions $X(x,t)$, it is extremely difficult to draw a conclusion about the direct relationship between the perturbation parameters and the maximum deviation value. Only in some cases, discussed below, the estimates of the perturbation parameters can be determined by the formula, while ensuring the stability of the numerical solution.

Acknowledgments

The research was funded by the Russian Science Foundation and Moscow city project No. 23-21-10065, https://rscf.ru/ project/23-21-10065/.

Author Contributions

All authors contributed equally to this work.

Conflict of Interest

The authors declare no conflict of interest.

References

1 Степанов В.В. Курс дифференциальных уравнений / В.В. Степанов // URSS. — 2016. — 512 с.

2 Atkinson K. Numerical solutions of ordinary differential equations / K. Atkinson. — John Wiley & Sons, Inc., Hoboken, New Jersey, 2009. — 261 p.

3 Немыцкий В.В. Качественная теория дифференциальных уравнений / В.В. Немыцкий, В. Степанов. — 4-е изд. — Ленанд, 2017. — 552 с.

4 Baumgarte J. Stabilization of constraints and integrals of motion in dynamical systems / J. Baumgartee // Computer Methods in Applied Mechanics and Engineering. — 1972. — Vol. 1. — No. 1. — P. 1–16.

5 Ascher U.M. Stabilization of DAEs and invariant manifolds / U.M. Ascher, H. Chin, S. Reich // Numerische Mathematik. — 1994. — Vol. 67. — No. 2. — P. 131–149.

6 Mukharlyamov R.G. Differential-algebraic equations of programmed motions of Lagrangian dynamical systems / R.G. Mukharlyamov // Mechanics of Solids. — 2011. — Vol. 46. — No. 4. — P. 534–543. https://doi.org/10.3103/S0025654411040042

7 Tleubergenov M.I. Construction of the differential equations system of the program motion in Lagrangian variables in the presence of random perturbations / M.I. Tleubergenov, G.K. Vassilina, D.T. Azhymbaev // Bulletin of the Karaganda University. Mathematics series. — 2022. — No. 1(105). — P. 118–126. https://doi.org/10.31489/2022M1/118-126

8 Tleubergenov M.I. Construction of stochastic differential equations of motion in canonical variables / M.I. Tleubergenov, G.K. Vassilina, S.R. Seisenbayeva // Bulletin of the Karaganda University. Mathematics series. — 2022. — No. 3(107). — P. 152–162. https://doi.org/10.31489/2022M3/152-162

9 Tleubergenov M.I. On construction of a field of forces along given trajectories in the presence of random perturbations // M.I. Tleubergenov, G.K. Vassilina, G.A. Tuzelbaeva // Bulletin of the Karaganda University. Mathematics series. — 2021. — No. 1(101). — P. 98–103. https://doi.org/10.31489/2021m1/98-103

10 Kaspirovich I.E. On Constructing Dynamic Equations Methods with Allowance for Atabilization of Constraints / I.E. Kaspirovich, R.G. Mukharlyamov // Mechanics of Solids. — 2019. — Vol. 54. — No. 4. — P. 589–597.

11 Kaspirovich I.E. Possible solutions of inverse dynamical problems with regards for nonlinear constraint stabilization function / I.E. Kaspirovich, R.G. Mukharlyamov // Journal of Physics: Conference Series. — 2020. — Vol. 1705. — P. 012013.

12 Omkar R. Numerical solution of differential-difference equations having an interior layer using nonstandard finite differences / R. Omkar, M. Lalu, K. Phaneendra // Bulletin of the Karaganda University. Mathematics series. — 2023. — No. 2(110). — P. 104–115. https://doi.org/10.31489/2023M2/104-115

13 Sinsoysal B. Numerical method to solution of generalized model Buckley-Leverett in a class of discontinuous functions / B. Sinsoysal, M. Rasulov, R. Iskenderova // Bulletin of the Karaganda University. Mathematics series. — 2023. — No. 1(109). — P. 131–140. https://doi.org/10.31489/2023M1/131-140

Баумгарт байланысының тұрақтануын ескере отырып, динамикалық теңдеулердің нақты және сандық шешімі арасындағы ауытқулардың кейбір бағалаулары туралы

М.И. Ақылбаев[1], И.Е. Каспирович[2]

[1] *Академик Ә. Қуатбеков атындағы Халықтар достығы университеті, Шымкент, Қазақстан;*
[2] *П. Лумумба атындағы Ресей халықтар достығы университеті, Мәскеу, Ресей*

Сандық интегралдауда айырымдық схемасын жүзеге асыру кезінде дөңгелектеу қателерінің жиналуына байланысты байланысы бар дифференциалдық теңдеулер жүйесінің сандық шешімі тұрақсыз болуы мүмкін. Жиналу мөлшерін шектеу үшін Баумгарт байланысын тұрақтандыру әдісі қолданылады. Нақты шешімнің сандық шешімнен ауытқуын бағалауда қажетті формулаларды алу үшін тұрақтандыру әдісін пайдалануға болады. Ауытқу функциясын Тейлор қатарына жіктеудің белгілі әдісі қолданылған. Мақалада бірінші ретті Эйлер әдісімен алынған сандық шешімнің қателігін бағалау қарастырылды.

Кілт сөздер: байланыстарды тұрақтандыру, сандық интегралдау, тұрақтылық, динамика, дифференциалдық теңдеулер жүйесі, сандық әдістер, сандық шешім, айырымдық схемасы, дөңгелектеу.

О некоторых оценках отклонений между реальным и численным решениями динамических уравнений с учетом стабилизации связи Баумгарта

М.И. Акылбаев[1], И.Е. Каспирович[2]

[1] *Университет дружбы народов имени академика А. Куатбекова, Шымкент, Казахстан;*
[2] *Российский университет дружбы народов имени П. Лумумбы, Москва, Россия*

Численное решение систем дифференциальных уравнений со связями может быть нестабильным из-за накопления ошибок округления при реализации разностной схемы численного интегрирования. Для ограничения величины накопления использован метод стабилизации связей Баумгарта. Для оценки отклонения реального решения от численного может быть применен метод стабилизации для получения требуемых формул. Использован хорошо известный метод разложения функции отклонения в ряд Тейлора. В статье рассмотрена оценка погрешности численного решения, полученного методом Эйлера первого порядка.

Ключевые слова: стабилизация связей, численное интегрирование, устойчивость, динамика, система дифференциальных уровней, численные методы, численное решение, разностная схема, округление.

References

1 Stepanov, V.V. (2016). *Kurs differentsialnykh uravnenii [Course of differential equations]*. URSS [in Russian].

2 Atkinson, K. (2009). *Numerical solutions of ordinary differential equations.* John Wiley & Sons, Inc., Hoboken, New Jersey.

3 Nemytskii, V. (2017). *Kachestvennaia teoriia differentsialnykh uravnenii [Qualitative theory of differential equations].* Lenand [in Russian].

4 Baumgarte, J.(1972). Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering, 1*(1), 1–16.

5 Ascher, U.M., Chin, H., & Reich, S. (1994). Stabilization of DAEs and invariant manifolds. *Numerische Mathematik, 67*(2), 131–149.

6 Mukharlyamov, R. (2011). Differential-algebraic equations of programmed motions of Lagrangian dynamical systems. *Mechanics of Solids, 46*(4), 534–543. https://doi.org/10.3103/S0025654411040042

7 Tleubergenov, M., Vassilina, G.K., & Azhymbaev, D.T. (2022). Construction of the differential equations system of the program motion in Lagrangian variables in the presence of random perturbations. *Bulletin of the Karaganda University. Mathematics series*, 1(105), 118–126. https://doi.org/10.31489/2022M1/118-126

8 Tleubergenov, M. (2022). Construction of stochastic differential equations of motion in canonical variables. *Bulletin of the Karaganda University. Mathematics series*, 3(107), 152–162. https://doi.org/10.31489/2022M3/152-162

9 Tleubergenov, M. (2021). On construction of a field of forces along given trajectories in the presence of random perturbations. *Bulletin of the Karaganda University. Mathematics series*, 1(101), 98–103. https://doi.org/10.31489/2021m1/98-103

10 Kaspirovich, I., & Mukharlyamov, R. (2019). On Constructing Dynamic Equations Methods with Allowance for Atabilization of Constraints. *Mechanics of Solids, 54*(4), 589–597.

11 Kaspirovich, I., & Mukharlyamov, R. (2020). Possible solutions of inverse dynamical problems with regards for nonlinear constraint stabilization function. *Journal of Physics: Conference Series, 1705*, 012013.

12 Omkar, R., Lalu, M., & Phaneendra, K. (2023). Numerical solution of differential-difference equations having an interior layer using nonstandard finite differences. *Bulletin of the Karaganda University. Mathematics series*, 2(110), 104–115. https://doi.org/10.31489/2023M2/104-115

13 Sinsoysal, B. (2023). Numerical method to solution of generalized model Buckley-Leverett in a class of discontinuous functions. *Bulletin of the Karaganda University. Mathematics series*, 1(109), 131–140. https://doi.org/10.31489/2023M1/131-140

*Author Information**

Musabek Islamovich Akylbayev (*corresponding author*) — Candidate (PhD) of Technical Sciences, Associate Professor of Mathematics, Peoples' Friendship University named after Academician A. Kuatbekov, 14 Tereshkova street, Shymkent, 160000, Kazakhstan; e-mail: *musabek_kz@mail.ru*; https://orcid.org/0000-0003-1383-4592

Ivan Evgenyevich Kaspirovich — Doctor of physical and mathematical sciences, Assistant at Institute of Physical Research and Technologies at Patrice Lumumba Peoples' Friendship University of Russia, 6 Miklukho-Maklaya street, Moscow, 117198, Russian Federation; e-mail: *kaspirovich_ie@rudn.ru*; https://orcid.org/0000-0002-6583-6621

*The author's name is presented in the order: First, Middle and Last Names.