Федеральный исследовательский центр Информатика и Управление Российской академии наук

RUDN university

РЭУ.РФ Российский экономический университет имени Г.В. Плеханова

КОМПЬЮТЕРНАЯ АЛГЕБРА

Материалы 6-й международной конференции

Москва, 23–25 июня 2025 г.

COMPUTER ALGEBRA

6th International Conference Materials

Moscow, June 23–25, 2025

Moscow
2025

UDC 519.6(063)
BBC 22.19;431

Editors: PhD Anna A. Ryabenko, DSc Dmitry S. Kulyabov

Reviewers: PhD A. V. Nesterov, PhD K. P.Lovetskiy

Printed in author's edition

Computer algebra: 6th International Conference Materials. Moscow, 23–25 June, 2025 / ed. A. A. Ryabenko, D. S. Kulyabov. Moscow: RUDN University.

DOI: 10.22363/12585-2025-6-000.
EDN: DZBCNS.

The international conference is organized by Federal Research Center "Computer Science and Control" of RAS, Peoples' Friendship University of Russia and Plekhanov Russian University of Economics. The talks presented at the conference discuss actual problems of computer algebra — the discipline whose algorithms are focused on the exact solution of mathematical and applied problems using a computer.

For scientists, graduate and undergraduate students in mathematics, physics and computer science.

ISBN 978-5-209-12585-3

Scientific publication

Program Committee

Conference Co-Chairs

I. A. Sokolov Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
A. V. Zorin RUDN University, Russia
V. A. Titov Plekhanov Russian University of Economics, Russia

Program Committee General Co-Chairs

Yu. G. Evtushenko Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
O. V. Kitova Plekhanov Russian University of Economics, Russia
K. E. Samuylov Applied Mathematics and Communications Technology Institute, RUDN University, Russia

Program Committee Vice Chairs

L. A. Sevastianov RUDN University, and Joint Institute for Nuclear Research, Russia
T. M. Sadykov Plekhanov Russian University of Economics, Russia
S. A. Abramov Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia

Program Committee

Yu. A. Flerov Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
N. P. Tuchkova Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
M. Barkatou Universite de Limoges, France
T. Cluzeau Universite de Limoges, France
O. V. Druzhinina Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
A. B. Batkhin Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Russia
A. V. Bernstein Skolkovo Institute of Science and Technology, Russia
Yu. A. Blinkov Saratov State University, Russia
M. Wu East China Normal University, Shanghai, P.R.China
E. V. Zima Wilfrid Laurier University, Waterloo, Canada
V. V. Kornyak Joint Institute for Nuclear Research, Russia
D. S. Kulyabov RUDN University, and Joint Institute for Nuclear Research, Russia
M. D. Malykh RUDN University, Russia
A. A. Mikhalev Moscow State University, Russia
A. N. Prokopenya Warsaw University of Life Sciences, Poland
V. A. Kitov Plekhanov Russian University of Economics, Russia

A. A. Bogolubskaya Joint Institute for Nuclear Research, Russia
R. R. Gontsov Institute for Information Transmission Problems of Russian Academy of Sciences, Russia
A. V. Korolkova RUDN University, Russia
A. P. Kryukov Moscow State University, Russia

Organising Committee

Organising Committee Co-Chairs

A. A. Gurin Plekhanov Russian University of Economics, Russia
D. V. Divakov RUDN University, Russia
A. A. Ryabenko Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia

Organising Committee Vice-Chairs

A. I. Medvedeva Plekhanov Russian University of Economics, Russia
Y. A. Zonn Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia
A. A. Tiutiunnik RUDN University, Russia

Organising Committee

G. M. Mikhailov Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia
T. R. Velieva RUDN University, Russia
A. V. Demidova RUDN University, Russia
V. V. Dorodnicyna Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia
Yu. O. Trusova Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia
I. V. Levichev RUDN University, Russia
A. A. Mamonov RUDN University, Russia
S. V. Vladimirova Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia
K. B. Teimurazov Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia

Contents

Invited talks

Contributed talks

Invited talks

Deciding Summability via Residues in Theory and in Practice

Carlos E. Arreche

The University of Texas at Dallas, Department of Mathematical Sciences, 800 W Campbell Road, Richardson, TX 75080, USA

Abstract

In difference algebra, summability arises as a basic problem upon which rests the effective solution of other more elaborate problems, such as creative telescoping problems and the computation of Galois groups of difference equations. In 2012 Chen and Singer introduced discrete residues as a theoretical obstruction to summability for rational functions with respect to the shift and q-dilation difference operators. Since then analogous notions of discrete residues have been defined in other difference settings relevant for applications, such as for Mahler and elliptic shift difference operators. Very recently there have been some advances in making these theoretical obstructions computable in practice.

Keywords

difference equation, difference field, discrete residues, summability problem

1. Difference fields, difference equations, and summability

1.1. Basic notation and conventions

We suppose throughout that $\mathcal{M}$ is a field of characteristic zero equipped with an endomorphism $\sigma : \mathcal{M} \hookrightarrow \mathcal{M}$ of infinite order, and denote by $\mathbb{K}$ the subfield of elements $c \in \mathcal{M}$ such that $\sigma(c) = c$. Thus $(\mathcal{M}, \sigma)$ is a *difference field* and its *subfield of invariants* is $\mathbb{K}$. We assume that $\mathbb{K}$ is relatively algebraically closed in $\mathcal{M}$, and we fix an algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$.

1.2. Linear difference equations

A *linear difference equation* over $\mathcal{M}$ of order r in a formal indeterminate y is one of the form $\sum_{j=0}^{r} a_j \sigma^j(y) = b$, $a_r, \dots, a_0, b \in \mathcal{M}$ such that $a_r a_0 \neq 0$. Such equations help us model many kinds of interesting sequences, such as the Fibonacci numbers or Catalan numbers, and many special functions such as the Euler Gamma function and combinatorial generating functions. Placing the study of such sequences and functions within the abstract setting of difference algebra described above is helpful for designing theoretical and practical algorithms that can further elucidate their properties based on the difference equation(s) that they satisfy.

1.3. The general summability problem

We say that $f \in \mathcal{M}$ is *summable* (in $\mathcal{M}$) if there exists $g \in \mathcal{M}$ such that $f = \sigma(g) - g$. The terminology is justified by the following discrete analogue of the Fundamental Theorem of Calculus: setting $F(n) = \sum_{k=0}^{n} \sigma^k(f)$, we can eliminate the summation symbol and write $F(n) = \sigma^{n+1}(g) - g$ for any g such that $f = \sigma(g) - g$. The study of summability was initiated in [1]. The *summability problem* asks to decide, for a given $f \in \mathcal{M}$, whether f is summable in $\mathcal{M}$. We insist on systematically ignoring the more difficult *summation problem* of computing a *certificate* $g \in \mathcal{M}$ and *reduced form* $h \in \mathcal{M}$ such that $f = \sigma(g) - g + h$ and h is somehow "minimal".

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*

10.22363/12585-2025-6-001

EDN: AAAQBF

arreche@utdallas.edu (C. E. Arreche)

https://personal.utdallas.edu/~arreche/ (C. E. Arreche)

0000-0001-8152-273X (C. E. Arreche)

For many (but not all!) purposes, the explicit computation of such additional data is both very onerous and also not needed beyond the answer to the simple question: is f summable or not?

1.4. Linear obstructions to summability

Let us denote the *forward difference operator* $\Delta := \sigma - \mathrm{id}_{\mathcal{M}}$, so that $\Delta(g) = \sigma(g) - g$ for $g \in \mathcal{M}$. Observing that Δ is a $\mathbb{K}$-vector space endomorphism of $\mathcal{M}$, we may rephrase the summability problem as asking, for a given $f \in \mathcal{M}$, whether it belongs to the $\mathbb{K}$-linear subspace $\mathrm{im}(\Delta)$. It is obvious from a theoretical point of view that there exists another $\mathbb{K}$-linear map on $\mathcal{M}$ whose kernel is precisely $\mathrm{im}(\Delta)$ — namely, the canonical projection $\mathcal{M} \twoheadrightarrow \mathcal{M}/\mathrm{im}(\Delta)$. It is also clear that such a map cannot possibly be unique, since for example post-composing with arbitrary injective $\mathbb{K}$-linear maps into other $\mathbb{K}$-vector spaces will not change the kernel. A *$\mathbb{K}$-linear obstruction to summability* is any $\mathbb{K}$-linear map ρ on $\mathcal{M}$ (to any target $\mathbb{K}$-vector space) such that $\ker(\rho) = \mathrm{im}(\Delta)$. In the next section we describe, for specific examples of difference fields $(\mathcal{M}, \sigma)$ of practical and theoretical interest, explicit $\mathbb{K}$-linear obstructions to summability, which are called (depending on the context) *discrete/orbital residues*. In the first few of these cases we also describe recent and ongoing efforts to design algorithms to efficiently compute ($\mathbb{K}$-rational representations of) these residues. Although the definitions become progressively more technical and complicated, we are optimistic that we shall eventually possess practical algorithms to compute all of them.

2. Discrete residues as obstructions to summability: case studies

2.1. The shift case

In the *shift case* (S), we consider $\mathcal{M} = \mathbb{K}(x)$ and $\sigma(x) = x + 1$. Given $f \in \mathbb{K}(x)$, there exists a unique *complete partial fraction decomposition*

$$f = p + \sum_{k \geq 1} \sum_{\alpha \in \overline{\mathbb{K}}} \frac{c_k(\alpha)}{(x - \alpha)^k}, \tag{1}$$

where $p \in \mathbb{K}[x]$ is a polynomial and all but finitely many of the $c_k(\alpha) \in \overline{\mathbb{K}}$ are 0. The *discrete residue* of f of order k at the *orbit* $\omega \in \overline{\mathbb{K}}/\mathbb{Z}$ is defined in [7] by the finite sum

$$\mathrm{dres}(f, \omega, k) := \sum_{\alpha \in \omega} c_k(\alpha). \tag{2}$$

It is proved in [7] that f is summable if and only if $\mathrm{dres}(f, \omega, k) = 0$ for every $\omega \in \overline{\mathbb{K}}$ and $k \in \mathbb{N}$. In [11, 2, 3] it is shown how to compute efficiently pairs of $\mathbb{K}$-polynomials (B_k, D_k) with the following properties: for each orbit ω such that $\mathrm{dres}(f, \omega, k) \neq 0$, there exists a unique $\alpha \in \omega$ such that $B_k(\alpha) = 0$, and for this α we have $D_k(\alpha) = \mathrm{dres}(f, \omega, k)$. This is desirable in applications where one wishes to compute with discrete residues but the computation of the complete partial fraction decomposition (1) is too expensive or impossible. An alternative computationally feasible $\mathbb{K}$-rational representation of discrete residues is described in [10].

2.2. The q-dilation case

In the *q-dilation case* (Q), we again consider $\mathcal{M} = \mathbb{K}(x)$, but this time we set $\sigma(x) = qx$ for some $q \in \mathbb{K}^\times = \mathbb{K} - \{0\}$ such that q is not a root of unity (so that σ is of infinite order). This time the *orbits* are the cosets $\omega \in \overline{\mathbb{K}}^\times / q^{\mathbb{Z}}$, and we must choose a distinguished representative α_ω in each

orbit ω. Relative to the complete partial fraction decomposition (1) of $f \in \mathbb{K}(x)$, the q-discrete residue of f of order k at the orbit $\omega \in \overline{\mathbb{K}}^\times/q^{\mathbb{Z}}$ is defined in [7] by the finite sum

$$\operatorname{dres}(f,\omega,k) := \sum_{n\in\mathbb{Z}} q^{-nk} c_k(q^n\alpha_\omega); \qquad \text{and} \qquad \operatorname{dres}(f,\infty) := p(0).$$

Note that making a different choice of distinguished representative $\alpha'_\omega = q^\ell \alpha_\omega$ of the orbit ω has the effect of multiplying the corresponding q-discrete residue of order k by $q^{\ell k}$. It is proved in [7] that f is summable if and only if $\operatorname{dres}(f,\infty) = 0$ and $\operatorname{dres}(f,\omega,k) = 0$ for every $\omega \in \overline{\mathbb{K}}^\times/q^{\mathbb{Z}}$ and $k \in \mathbb{N}$. It would be desirable to have in this case also efficient algorithms analogous to those of [11, 2, 2] in the shift case (S) that produce $\mathbb{K}$-rational representations of the q-discrete residues of f whilst bypassing the expensive or impossible computation of the complete partial raction decomposition (1). No such algorithm exists (yet).

2.3. The Mahler case

In the *Mahler case* (M), we again consider $\mathcal{M} = \mathbb{K}(x)$, but this time we set $\sigma(x) = x^m$ for some integer $m \geq 2$. Note that in this case σ is only an endomorphism of $\mathcal{M}$, but not an automorphism.[1] In this case it is helpful to decompose $f \in \mathbb{K}(x)$, relative to the complete partial fraction decomposition (1), as a Laurent polynomial component $f_L := p + \sum_{k\geq 1} c_k(0)x^{-k}$ and a complementary component $f_T := f - f_L$. It is not difficult to see that f is summable if and only if both f_L and f_T are summable. For any Laurent polynomial $L = \sum_{j\in\mathbb{Z}} \ell_j x^j \in \mathbb{K}[x,x^{-1}]$, its *Mahler discrete residue* is a vector indexed by equivalence classes θ under the equivalence relation on integer exponents $i \sim j$ if $i/j \in m^{\mathbb{Z}}$, with components defined by $\operatorname{dres}(L,\infty)_\theta := \sum_{j\in\theta} \ell_j$. In this setting we must similarly decompose $\overline{\mathbb{K}}^\times$, not into orbits but rather into *Mahler trees* τ: these are the equivalence classes in $\overline{\mathbb{K}}$ under the equivalence relation $\alpha \sim \beta$ if $\alpha^{m^r} = \beta^{m^s}$ for some $r,s \in \mathbb{N}$. There is a qualitative dichotomy between *torsion trees* that consist entirely of roots of unity, and *non-torsion trees* that contain no roots of unity. For a non-torsion tree τ and $k \in \mathbb{N}$, the *Mahler discrete residue* $\operatorname{dres}(f,\tau,k)$ is defined in [5] as a vector in $\overline{\mathbb{K}}^\tau$, all of whose components are zero except for those indexed by a certain finite set of $\alpha \in \tau$ of "maximal height" (relative to f), for which the corresponding components are defined by the finite sum

$$\operatorname{dres}(f,\tau,k)_\alpha = \sum_{s\geq k}\sum_{n\geq 0} V_{s,k}^{(m,n)} \alpha^{k-sm^n} c_s(\alpha^{m^n}); \qquad \text{where the} \qquad V_{s,k}^{(m,n)} \in \mathbb{Q}$$

are certain auxiliary structural constants computed explicitly in [4, 5] for $1 \leq s \leq r$ and $n \geq 0$. The Mahler discrete residues at torsion trees are defined similarly, but mediated by additional technical ingredients necessary to retain control over the pre-periodic behavior of roots of unity under the Mahler endomorphism $\zeta \mapsto \zeta^m$. It is proved in [4, 5] that f is Mahler summable if and only if all its Mahler discrete residues vanish. In [5] a generalization of Mahler discrete residues is developed for the "twisted" Mahler summability problem of deciding, for a given $f \in \mathbb{K}(x)$ and $\lambda \in \mathbb{Z}$, whether $f = m^\lambda \sigma(g) - g$ for some $g \in \mathbb{K}(x)$. There seem to be several technical difficulties to overcome in order to develop practical algorithms to compute (twisted) Mahler discrete residues.

2.4. The elliptic case

In the *elliptic shift case* (E), we fix an *elliptic curve* $\mathcal{E} : y^2 = x^3 + Ax + B$ for some $A, B \in \mathbb{K}$ such that $4A^3 + 27B^2 \neq 0$, and we denote by $\mathcal{M}_\mathcal{E} = \mathbb{K}(x,y)$, the field of rational functions on $\mathcal{E}$. The elliptic curve $\mathcal{E}$ can sometimes be modeled in other ways. In case $\mathbb{K}$ is the field $\mathbb{C}$ of complex

[1] In certain theoretical contexts it can be useful to replace $\mathbb{K}(x)$ with $\bigcup_{n\in\mathbb{N}} \mathbb{K}(x^{1/n})$, for which the natural extension of the Mahler endomorphism σ becomes an automorphism.

numbers, there exists a lattice $\Lambda \subset \mathbb{C}$ such that $\mathcal{E} \simeq \mathbb{C}/\Lambda$, and $\mathcal{M}_\mathcal{E}$ is identified with the field of meromorphic functions $f(z)$ on $\mathbb{C}$ such that $f(z+\lambda) = f(z)$ for every $\lambda \in \Lambda$. In case $\mathbb{K}$ is $\mathbb{C}$ or $\mathbb{R}$ (or some other complete valued field, such as a p-adic field) there exists[2] $q \in \mathbb{K}^\times$ with $|q| < 1$ such that $\mathcal{E}$ is isomoprhic to the *Tate curve* $\mathbb{K}^\times/q^\mathbb{Z}$, and $\mathcal{M}_\mathcal{E}$ is identified with the field of meromorphic functions $f(z)$ on $\mathbb{K}^\times$ such that $f(qz) = f(z)$.

For a given $\mathbb{K}$-rational non-torsion point $t \in \mathcal{E}$, we consider the corresponding automorphism σ on $\mathcal{M}_\mathcal{E}$ obtained by pre-composing rational functions f on $\mathcal{E}$ with the addition-by-t map σ_* on $\mathcal{E}$ under the elliptic group law. Concretely, if the coordinates $x(t) = t_x$ and $y(t) = t_y$, then setting $s_t := \frac{y-t_y}{x-t_x}$ we have $\sigma(x) = s_t^2 - x - t_x$ and $\sigma(y) = s_t(x - \sigma(x)) - y$. In the conceptually simpler alternative descriptions of $\mathcal{E}$, the corresponding description of σ is more straightforward: when $\mathcal{E} = \mathbb{C}/\Lambda$, we have $\sigma(f(z)) = f(z+t)$ for some $t \in \mathbb{C}$ such that $nt \notin \Lambda$ for any $n \in \mathbb{N}$; and in the Tate curve setting $\mathcal{E} = \mathbb{K}^\times/q^\mathbb{Z}$, we have $\sigma(f(z)) = f(tz)$ for some $t \in \mathbb{K}^\times$ such that t and q are multiplicatively independent. Working in the algebraic setting, a *compatible system of parameters* $\mathcal{U} = \{u_\alpha \in \mathcal{M}_\mathcal{E} \mid \alpha \in \mathcal{E}(\overline{\mathbb{K}})\}$ is defined in [8, 9] by the conditions that $\mathrm{ord}_\alpha(u_\alpha) = 1$ and $\sigma(u_{\sigma_*(\alpha)}) = u_\alpha$. For $f \in \mathcal{M}_\mathcal{E}$ and $\alpha \in \mathcal{E}(\overline{\mathbb{K}})$, there exist unique $c_k(\alpha) \in \overline{\mathbb{K}}$ for $k \in \mathbb{N}$, almost all 0, such that $f - \sum_{k\geq 1} c_k(\alpha) u_\alpha^{-k}$ is non-singular at α. Relative to these ancillary definitions, the *orbital residue* of $f \in \mathcal{M}_\mathcal{E}$ at the orbit $\omega \in \mathcal{E}(\overline{\mathbb{K}})/\mathbb{Z}.t$ of order $k \in \mathbb{N}$ is defined in [8, 9] by the same formula (2), where it is also proved that if f is summable then all its orbital residues vanish. However, the converse is not true. An additional set of two obstructions, called *panorbital residues*, were introduced in [6] in both the lattice and algebraic settings, where it was proved f is summable if and only if its orbital and panorbital residues all vanish. The definition of panorbital and orbital residues in the setting of Tate curves will appear in a forthcoming publication. It would be desirable also in this case to have algorithms that can compute orbital and panorbital residues of elliptic functions, at least in the algebraic setting.

Remark 2.1. Another interesting operator on $\mathcal{M}_\mathcal{E}$ is obtained by pre-composing $f \in \mathcal{M}_\mathcal{E}$ with the multiplication-by-m map for an integer $m \geq 2$ under the elliptic group law. As far as we know, no one has yet defined a $\mathbb{K}$-linear obstruction to summability in this *elliptic Mahler case*.

References

[1] S. A. Abramov, On the summation of rational functions, USSR Computational Mathematics and Mathematical Physics 11(4) (1971) 324–330. doi:https://doi.org/10.1016/0041-5553(71)90028-0.

[2] C. E. Arreche, H. Sitaula, Computing discrete residues of rational functions, in: Proceedings of the 2024 International Symposium on Symbolic and Algebraic Computation, ISSAC '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 65–73. doi:10.1145/3666000.3669676.

[3] C. E. Arreche, H. P. Sitaula, A computational approach to rational summability and its applications via discrete residues, 2025. URL: https://arxiv.org/abs/2503.15636. arXiv:2503.15636.

[4] C. E. Arreche, Y. Zhang, Mahler discrete residues and summability for rational functions, in: Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation, ISSAC '22, Association for Computing Machinery, New York, NY, USA, 2022, pp. 525–533. doi:10.1145/3476446.3536186.

[5] C. E. Arreche, Y. Zhang, Twisted Mahler discrete residues, International Mathematics Research Notices 2024(23) (2024) 14259–14288. doi:10.1093/imrn/rnae238.

[6] M. Babbitt, Summability of Elliptic Functions via Residues, Ph.D. thesis, The University of Texas at Dallas, 2025. doi:10.48550/arXiv.2503.22770.

[2] After possibly replacing $\mathbb{K}$ with a finite algebraic extension.

[7] S. Chen, M. F. Singer, Residues and telescopers for bivariate rational functions, Advances in Applied Mathematics 49(2) (2012) 111–133. doi:https://doi.org/10.1016/j.aam.2012.04.003.

[8] T. Dreyfus, C. Hardouin, J. Roques, M. F. Singer, On the nature of the generating series of walks in the quarter plane, Inventiones mathematicae 213(1) (2018) 139–203. doi:10.1007/s00222-018-0787-z.

[9] C. Hardouin, M. F. Singer, On differentially algebraic generating series for walks in the quarter plane, Selecta Mathematica 27(5) (2021) 89. doi:10.1007/s00029-021-00703-9.

[10] Q.-H. Hou, R.-H. Wang, An algorithm for deciding the summability of bivariate rational functions, Advances in Applied Mathematics 64 (2015) 31–49. doi:10.1016/j.aam.2014.11.002.

[11] H. Sitaula, Algorithms to Compute Discrete Residues of a Rational Function, Ph.D. thesis, The University of Texas at Dallas, 2023. URL: https://hdl.handle.net/10735.1/9952.

Gröbner Bases for Investigating and Solving Polynomial and Differential Systems

Yury A. Blinkov

Saratov State University, 83 Astrakhanskaya St, Saratov, 410012, Russian Federation

Abstract

This paper explores the theory and applications of Gröbner bases in polynomial and differential systems, emphasizing their role in solving ideal membership problems, analyzing differential equations, and constructing numerical schemes. The computational framework leverages the Python library PyGInv (a lightweight variant of GInv), integrated with SymPy, to demonstrate efficiency in symbolic computations.

Keywords

Gröbner bases, Hilbert polynomial, roots of polynomial equations, symmetries of differential equations, difference schemes, first differential approximation

Historically, works in the direction of creating the theory of Gröbner bases were focused on constructively solving the ideal membership problem. One of the first works was done in 1900 by [12]. An admissible monomial ordering was introduced by Macaulay in 1927 [15]. In 1939 [13], Gröbner used admissible monomial ordering to find a basis of the quotient ring for a zero-dimensional ideal. Gröbner bases were introduced by Buchberger in his dissertation in 1965 [7]. In his subsequent works, Buchberger applied the algorithm for constructing Gröbner bases to the study of systems of polynomial equations.

An alternative approach to the Gröbner basis method emerged during the investigation of differential equations. It was focused on reducing the system to a form that allows determining the dimension of the solution space. The concept of involutiveness was introduced by Cartan [8] while studying Pfaffian-type equations in total differentials. An *involutive* system of differential equations contains all integrability conditions, and prolongations of the system do not yield new compatibility conditions. The completion of a system with its integrability conditions is called *closure*.

Riquier [17], while studying solutions of differential equations in the form of formal series, proposed a complete ordering for partial derivatives. Using this ordering, he identified a subset of derivatives called *principal*, with respect to which the system of PDEs can be resolved. The remaining derivatives, called *parametric*, determine the arbitrariness in the solution and affect the formulation of initial conditions. As a result, Riquier developed a theory containing the Cauchy-Kovalevskaya theorem as a special case.

Janet [14] further developed Riquier's approach. For principal derivatives, he introduced a partition of independent variables into *multiplicative* and *non-multiplicative*. As a result, all prolongations of the system were divided into multiplicative and non-multiplicative, and if prolongations by non-multiplicative variables yielded the same system as prolongations by multiplicative variables, such a system was called *passive*.

The difference between Gröbner bases and their extension by involutive bases is best illustrated in Figure 1. On the left figure, cones of divisible monomials are shown in different colors. As a result of fixing the monomial ordering in the general polynomial case, at the intersection points of the cones, it is necessary to compute S-polynomials with respect to the current basis and add them to the basis. Upon completion of this process, a Gröbner basis in the chosen

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-002
EDN: ACVBOF
blinkovua@sgu.ru (Y. A. Blinkov)
0000-0001-7340-0919 (Y. A. Blinkov)

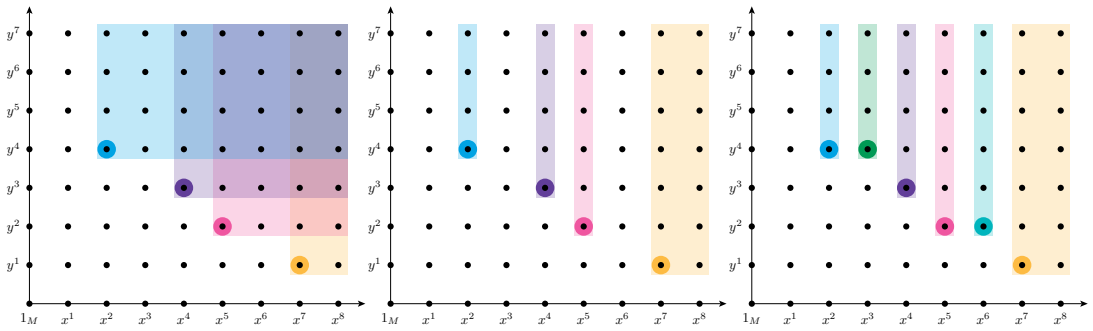


Figure 1: Left: standard division; center and right: Janet division.

ordering will be obtained. Now, to determine membership, it is sufficient to compute the polynomial's value with respect to the current basis. The sequence of reductions is unimportant since all S-polynomials already belong to the basis.

Janet's partition of variables prohibits division by non-multiplicative variables. As a result, the cones in the center and right figures do not intersect, and the dimension of the cones without intersections can be easily calculated. By augmenting the current basis with prolongations by non-multiplicative variables, all S-polynomials will be computed, and the involutive basis will contain the Gröbner basis, as shown in the right figure. For it, the Hilbert polynomial can be easily calculated: $HP_8 = 3s + 10$, where $s \geq 8$ denotes the number of monomials up to total degree s that have no divisors among the leading monomials of the Gröbner basis. The involutive approach to constructing Gröbner bases was developed in the works [19, 9, 10].

The Gröbner basis in the zero-dimensional case allows constructing a matrix representation for each variable. For example, for the well-known cyclic roots example cyclic6, the Hilbert polynomial equals 156, meaning there are exactly 156 monomials w_l that have no divisors among the leading monomials of the Gröbner basis. The transition to matrix representation can be performed using the formula: $x_i w_l = \sum_{k=1}^{156} X^i_{l,k} w_k$. Here, the matrices X^i are integer coefficients obtained by computing the corresponding normal form. The number of non-zero terms in these matrices is $2789, 2728, 2548, 2406, 1952, 799$, and their eigenvalues represent all roots of the system, accounting for their multiplicity.

Using Rabinowitsch's trick [16] to represent inequalities by introducing additional variables, Gröbner bases can be used to solve the graph coloring problem with n colors. For example, in the case of three colors, the colors for adjacent vertices can be defined by the following polynomials: $x(x+1)(x+2), y(y+1)(y+2), (x-y)k-1$. In the last polynomial, the additional variable k cannot vanish, meaning the colors x and y are distinct. The Gröbner basis after eliminating the polynomial with k has the following form: $x^3+3x^2+2x, y^2+yx+3y+x^2+3x+2$. Using these relations, all adjacent vertices can be described, and by constructing the Gröbner basis, all possible colorings of the given graph can be found.

Another interesting application of polynomial Gröbner bases is presented in the work [2].

The condition for higher symmetries of the Korteweg-de Vries equation (abbreviated notation, e.g., $u_{xxx} = u_3$):

$$\mathbf{u_t} + 6u_1u + u_3 = 0$$
$$\mathbf{u}_\varepsilon - F(t, x, u, u_1, u_2, \ldots, u_5) = 0$$

For the second equation of the system, we introduce total derivatives of F.

$$\frac{dF}{dt} = F_t + F_u(6u_1u + u_3) + \sum F_{u_i}(6u_1u + u_3)_i$$
$$\frac{dF}{dx} = F_x + F_u u_1 + \sum F_{u_i} u_{i+1}$$

This form of notation allows introducing a constant ε into the exact solution of the Korteweg-de

Vries equation by integrating the second equation. The integrability (compatibility) condition of this system can be written using the S-polynomial (lexicographic ordering $t \succ x$):

$$\left(\mathbf{u_t} + 6u_1u + u_3\right)_\varepsilon - \mathbf{u}_{\varepsilon\mathbf{t}} + \frac{d}{dt}F = 0$$

In the last equation, due to the Leibniz differentiation rule, the derivative with respect to ε enters the integrability condition linearly.

$$6uu_{\varepsilon 1} + 6u_1u_\varepsilon + u_{\varepsilon 3}\frac{d}{dt}F = 0$$

Here, we can replace u_ε with F using total derivatives of F.

$$6u\frac{d}{dx}F + 6u_1F + \frac{d^3}{dx^3}F + \frac{d}{dt}F = 0 \tag{1}$$

Since F itself does not depend on derivatives higher than fifth order, we can obtain a system of linear equations for F. By constructing a Gröbner basis for the resulting system and refining the form of the solution, the explicit form of higher symmetries can be found after several steps.

Gröbner bases can also be used to construct difference schemes [11, 5, 1, 5]. As an example, consider constructing a compact fourth-order difference scheme for the Laplace equation. A compact scheme for second derivatives is known: $\frac{1}{10}f''_{i-1} + f''_i + \frac{1}{10}f''_{i+1} = \frac{6}{5}\frac{f_{i+1}-2f_i+f_{i-1}}{h^2}$. By introducing three functions with the elimination ordering $uxx \succ uyy \succ u$, the scheme for the Laplace equation can be written as:

$$\begin{array}{l}
\frac{1}{10}uxx_{i-1,\,k} + uxx_{i,\,k} + \frac{1}{10}uxx_{i+1,\,k} - \frac{6}{5}\frac{u_{i+1,\,k}-2u_{i,\,k}+u_{i-1,\,k}}{h^2},\\
\frac{1}{10}uyy_{i,\,k-1} + uyy_{i,\,k} + \frac{1}{10}uyy_{i,\,k+1} - \frac{6}{5}\frac{u_{i,\,k+1}-2u_{i,\,k}+u_{i,\,k-1}}{h^2},\\
uxx_{i,\,k} - uyy_{i,\,k}.
\end{array}$$

This yields the well-known compact fourth-order scheme: $\begin{array}{ccc} 1 & 4 & 1 \\ 4 & -20 & 4. \\ 1 & 4 & 1 \end{array}$

In the 1960s, Yanenko and Shokin [18] formulated the method of differential approximations of difference schemes. The main idea of this method is to replace the study of properties of difference schemes with the study of a certain problem with differential equations that occupy an intermediate position between the original differential problem and the difference scheme approximating it.

The algorithm for constructing the first differential approximation (FDA) for such systems represents a set of simple operations with formal power series, for which the construction of a Gröbner basis allows building FDA not only for evolutionary-type equations [4].

For ordinary differential equations (ODEs), constructing FDA allows estimating the stiffness of the ODE system [6].

Constructing FDA for the Van der Pol oscillator [6] enabled correctly determining the order of the numerical method and its residual term. FDA was constructed for various explicit and implicit Runge-Kutta methods and multistep Adams-Bashforth and Adams-Moulton methods. All results showed that FDA has the form $h^p(C\mu^{p+1}u_1(u^2-1)^{p+1}) + \ldots)$, where p is the order of the method and C is some constant. The conducted computations demonstrate that using FDA, one can estimate the residual of the used numerical method depending on the problem parameters, detect and estimate the stiffness of the ODE system. For computations with variable step size, given a tolerance tol, computations can be performed using the following formula: $\min(h, (tol/(1 + abs(\text{FDA}))^{1/p})$.

The presented methods allow conducting efficient computations using computer algebra systems. All given examples were computed in PyGInv, which is a "lightweight" version of GInv [3] (short for Gröbner INVolutive). It is written in "pure" Python and uses the SymPy computer algebra system for working with parameters.

References

[1] P. Amodio, Y. Blinkov, V. Gerdt, R. La Scala, Algebraic construction and numerical behavior of a new s-consistent difference scheme for the 2D Navier–Stokes equations, Programming and Computer Software 314 (2017) 408–421. doi:10.1016/j.amc.2017.06.037.

[2] R. E. Bayramov, Y. A. Blinkov, I. V. Levichev, M. D. Malykh, V. S. Melezhik, Analytical study of cubature formulas on a sphere in computer algebra systems, Comput. Math. and Math. Phys. 63 (2023) 77–85. doi:10.1134/S0965542523010050.

[3] Y. A. Blinkov, V. P. Gerdt, Specialized computer algebra system GINV, Programming and Computer Software 34 (2008) 112–123. doi:10.1134/S0361768808020096.

[4] Y. A. Blinkov, A. Y. Rebrina, Investigation of difference schemes for two-dimensional Navier–Stokes equations by using computer algebra algorithms, Programming and Computer Software 49 (2024) 26–31. doi:10.1134/S0361768823010024.

[5] Y. A. Blinkov, V. P. Gerdt, K. B. Marinov, Discretization of quasilinear evolution equations by computer algebra methods, Programming and Computer Software 43 (2017) 84–89. doi:10.1134/S0361768817020049.

[6] Y. A. Blinkov, Computer-algebraic approach to first differential approximations: Van der Pol oscillator, Programming and Computer Software 50 (2024) 115–120. doi:10.1134/S0361768824020026.

[7] B. Buchberger, Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal, Journal of Symbolic Computation 41 (2006) 475–511. doi:10.1016/j.jsc.2005.09.007.

[8] E. Cartan, Sur certaines expressions différentielles et le problème de Pfaff, Annales scientifiques de l'École Normale Supérieure Sér. 3 (1899) 239–332. URL: http://archive.numdam.org/article/ASENS_1899_3_16__239_0.djvu.

[9] V. P. Gerdt, Y. A. Blinkov, Involutive bases of polynomial ideals, Mathematics and Computers in Simulation 45 (1998) 519–542. doi:10.1016/S0378-4754(97)00127-4.

[10] V. P. Gerdt, Y. A. Blinkov, Minimal involutive bases, Mathematics and Computers in Simulation 45 (1998) 543–560. doi:10.1016/S0378-4754(97)00128-6.

[11] V. P. Gerdt, Y. A. Blinkov, V. V. Mozzhilkin, Gröbner bases and generation of difference schemes for partial differential equations, Symmetry, Integrability and Geometry: Methods and Applications. 2 (2006) 26. doi:10.3842/SIGMA.2006.051.

[12] P. Gordan, Les invariants des formes binaires, Journal de Mathématiques Puves et Appliqués 6 (1900) 141–156.

[13] W. Gröbner, Über die Eliminationstheorie, Monatsh. der Math. 54 (1939) 71–78.

[14] M. Janet, Systèmes d'équations aux dérivées partielles, Journals de mathématiques, 8e série 3 (1920) 65–151. URL: http://eudml.org/doc/192952.

[15] F. S. Macaulay, Some properties of enumeration in the theory of modular systems, Proc. London Math. Soc. 26 (1927) 531–555. doi:10.1112/plms/s2-26.1.531.

[16] J. L. Rabinowitsch, Zum Hilbertschen Nullstellensatz, Math. Ann. 102 (1930) 520. doi:10.1007/BF01782361.

[17] C. Riquier, Les Systèmes d'Equations aux Dérivées Partielles, Gauthier-Villars, Paris, 1910.

[18] N. N. Yanenko, Y. I. Shokin, The first differential approximation to finite-difference schemes for hyperbolic systems of equations, Siberian Mathematical Journal 10 (1969) 868–880. doi:10.1007/BF00971662.

[19] A. Y. Zharkov, Y. A. Blinkov, Involution approach to investigating polynomial systems, Mathematics and Computers in Simulation 42 (1996) 323–332. doi:10.1016/S0378-4754(96)00006-7.

Integrable Cases of the Euler–Poisson Equations

Alexander D. Bruno[1,*], Alexander B. Batkhin[1,2]

[1]Keldysh Institute of Applied Mathematics, Miusskaya Sq. 4, Moscow, 125047, Russian Federation
[2]Technion – Israel Institute of Technology, Technion City, 3200003 Haifa, Israel

Abstract
In the classical problem of the motion of a rigid body around a fixed point, described by the Euler-Poisson equations, we make an attempt to connect integrability of the system with its properties near its stationary points. The method gives all known cases of integrability and several new.

Keywords
rigid top, Euler-Poisson equations, first integral, normal form

1. Introduction

The system of *Euler-Poisson equations* (1750) (or shortly EP-equations) is a real autonomous system of six ordinary differential equations (ODEs).

$$\begin{aligned} &Ap' + (C-B)qr = Mg\,(y_0\gamma_3 - z_0\gamma_2)\,,\\ &Bq' + (A-C)pr = Mg\,(z_0\gamma_1 - x_0\gamma_3)\,,\\ &Cr' + (B-A)pq = Mg\,(x_0\gamma_2 - y_0\gamma_1)\,,\\ &\gamma_1' = r\gamma_2 - q\gamma_3, \quad \gamma_2' = p\gamma_3 - r\gamma_1, \quad \gamma_3' = q\gamma_1 - p\gamma_2, \end{aligned} \tag{1}$$

with dependent variables $p, q, r, \gamma_1, \gamma_2, \gamma_3$ and parameters A, B, C, x_0, y_0, z_0, satisfying the triangle inequalities

$$0 < A \leqslant B + C, \quad 0 < B \leqslant A + C, \quad 0 < C \leqslant A + B. \tag{2}$$

Here, the prime indicates differentiation over the independent variable time t, Mg is the weight of the body, A, B, C are the principal moments of inertia of the rigid body, x_0, y_0, z_0 are the coordinates of the center of mass of the rigid body, $\gamma_1, \gamma_2, \gamma_3$ are the vertical directional cosines.

EP-equations describe the motion of a rigid top around a fixed point [4] and have the following three first integrals:

$$\begin{aligned} \textit{energy}\colon\ & I_1 \stackrel{\text{def}}{=} Ap^2 + Bq^2 + Cr^2 - 2Mg\,(x_0\gamma_1 + y_0\gamma_2 + z_0\gamma_3) = h = \text{const},\\ \textit{momentum}\colon\ & I_2 \stackrel{\text{def}}{=} Ap\gamma_1 + Bq\gamma_2 + Cr\gamma_3 = l = \text{const},\\ \textit{geometric}\colon\ & I_3 \stackrel{\text{def}}{=} \gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1. \end{aligned}$$

EP-equations are integrable if there is a fourth general integral I_4. So far, *five* cases of integrability are known:

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-003
EDN: ADQSAX
*Corresponding author.
abruno@keldysh.ru (A. D. Bruno); batkhin@technion.ac.il (A. B. Batkhin)
http://brunoa.name/ (A. D. Bruno)
0000-0002-7465-1258 (A. D. Bruno); 0000-0001-8871-4697 (A. B. Batkhin)

Case 1. *Euler-Poinsot*: $x_0 = y_0 = z_0 = 0$ and the fourth integral is

$$I_4 \stackrel{\text{def}}{=} A^2p^2 + B^2q^2 + C^2r^2 = \text{const}.$$

Case 2. *Lagrange-Poisson*: $B = C$, $x_0 \neq 0$, $y_0 = z_0 = 0$, and the fourth integral is

$$I_4 \stackrel{\text{def}}{=} p = \text{const}.$$

Case 3. *Kovalevskaya* (1890): $A = B = 2C$, $x_0 \neq 0$, $y_0 = z_0 = 0$, and

$$I_4 \stackrel{\text{def}}{=} \left(p^2 - q^2 + c\gamma_1\right)^2 + (2pq + c\gamma_2)^2 = \text{const},$$

where $c = Mgx_0/C$.

Case 4. *Kinematic symmetry*: $A = B = C$ and $I_4 \stackrel{\text{def}}{=} x_0p + y_0q + z_0r = \text{const}$. It is derived from Case 2.

Case 5. *Bruno–Batkhin* (2024) [1]: $A = B = 2C$, $x_0 \neq 0$, $y_0 \neq 0$, $z_0 = 0$, the fourth integral is

$$I_4 \stackrel{\text{def}}{=} \left(p^2 - q^2 + c\gamma_1 - d\gamma_2\right)^2 + (2pq + d\gamma_1 + c\gamma_2)^2 = \text{const},$$

where $c = Mgx_0/C$, $d = Mgy_0/C$.

Let ℓ be a number of nonzero values among parameters x_0, y_0, z_0. Below we consider theory for all $\ell = 0, 1, 2, 3$. We apply it here for the case $\ell = 1$ only because of economy of space, but in our lecture we will consider the cases $\ell = 2, 3$ as well.

2. Theory

2.1. Local and global integrability

In [2] the notion *local integrability* was introduced: *An ODE system is locally integrable near a stationary point (SP) of the system if it has enough analytic integrals in a vicinity of the SP.* It is evident that an integrable system is locally integrable at each of its stationary points.

Hypothesis 1 (Edneral [3]). *If an autonomous polynomial ODE system is locally integrable in the neighborhood of all its stationary points, then it is globally integrable.*

Therefore, to find global integrability, we must first find all stationary points of the ODE system and then find out whether the system is locally integrable in their neighborhoods.

Let $X = (p, q, r, \gamma_1, \gamma_2, \gamma_3)$, the point X be a stationary point in the system (1) and

$$M = \begin{pmatrix} 0 & \frac{B-C}{A}r & \frac{B-C}{A}q & 0 & -\frac{z_0}{A} & \frac{y_0}{A} \\ \frac{C-A}{B}r & 0 & \frac{C-A}{B}p & \frac{z_0}{B} & 0 & -\frac{x_0}{B} \\ \frac{A-B}{C}q & \frac{A-B}{C}p & 0 & -\frac{y_0}{C} & \frac{x_0}{C} & 0 \\ 0 & -\gamma_3 & \gamma_2 & 0 & r & -q \\ \gamma_3 & 0 & -\gamma_1 & -r & 0 & p \\ -\gamma_2 & \gamma_1 & 0 & q & -p & 0 \end{pmatrix}$$

is a matrix of the linear part of the system (1) near the point X. The characteristic polynomial $\chi(\lambda)$ of the matrix M is $\chi(\lambda) = \lambda^6 + a_4\lambda^4 + a_2\lambda^2$. Canceling it by λ^2 we get a biquadratic form, which discriminant on λ^2 is the following

$$D_\lambda(\chi) = a_4^2 - 4a_2. \tag{3}$$

It is a rational function $D = G/H$, where G and H are polynomials in system parameters.

A stationary point is *locally integrable* [2] if $a_2 < 0$ or $D_\lambda(\chi) < 0$. But this property is not satisfied for definite values of system parameters (1).

The stationary points of the EP system form one-dimensional and two-dimensional families $\mathcal{F}_j^\ell$ in $\mathbb{R}^6$. Here ℓ is the number of nonzero values among x_0, y_0, z_0 and j is the number of the family for given value of ℓ.

The numerator G of the first discriminant $D_\lambda(\chi)$ (3) depends on the set Ξ of parameters

$$\Xi \stackrel{\text{def}}{=} \{s,\ A,\ B,\ C\} \tag{4}$$

where s is the parameter along the family $\mathcal{F}_j^\ell$ and others are parameters of the system (1). Let ξ be one of the parameters (4). By $\Delta_\xi\left(\mathcal{F}_j^\ell\right)$ we denote the *secondary discriminant* of the numerator G of the first discriminant (3) on the parameter ξ.

Hypothesis 2. *If near a stationary point X of the family $\mathcal{F}_j^\ell$ with certain parameters values* (4) *the EP equations are locally integrable, then at these parameter values there is at least one secondary discriminant $\Delta_\xi\left(\mathcal{F}_j^\ell\right) = 0$.*

We have found some general properties of the integrable cases 1–5, which were formulated as Hypothesis 2. So we have to compute all the values of the parameters A, B, C, x_0, y_0, z_0 for which this property is satisfied.

2.2. Zero eigenvalues

Usually, SP of the EP-equations has two zero eigenvalues. The case where the characteristic polynomial $\chi(\lambda)$ has zero roots can be studied using the approach from [5]. But here we assume

Hypothesis 3. *For EP-equations, the stationary point with four zero eigenvalues is locally integrable.*

2.3. Checking for integrability

There are two ways to check the existence of the fourth integral:

- finding this integral in the explicit form;
- using the normal form (NF) of the system near the SP.

In the last case, according to [2, Section 5.3] the coefficients of the resonant terms of the NF at a resonance of order 3, i.e., when there exists a pair of eigenvalues with ratio $2:1$, should be zero in integrable cases. They are zero in some subcases, and in other subcases they are nonzero.

So, for a check of integrability, it is sufficient to provide one step of the normalization transformation $X = QY$, where Q is a matrix composed of the eigenvectors of M and reduces the linear system of EP-equations into a diagonal form. After the reduction, it is only necessary to check the coefficients of the resonant monomials $Y_j^p Y_k^q$.

3. Case $\ell = 1$

3.1. Families of SP for $\ell = 1$

Theorem 1. *For $\ell = 1$ the system* (1) *has four families of SP:*

$$\mathcal{F}_1^1 : \{p = s,\ q = r = 0,\ \gamma_1 = p/k = \pm 1,\ \gamma_2 = \gamma_3 = 0\};$$
$$\mathcal{F}_2^1 : \left\{p = \frac{x_0}{k(C-A)}, q = 0, r = \frac{s}{k},\ \gamma_1 = \frac{p}{k},\ \gamma_2 = 0, \gamma_3 = \frac{r}{k}, A \neq C,\ x_0 \neq 0\right\};$$

$$\mathcal{F}_3^1 : \left\{ p = \frac{x_0}{k(B-A)}, r = 0, q = \frac{s}{k},\ \gamma_1 = \frac{p}{k},\ \gamma_2 = \frac{q}{k}, \gamma_3 = 0, A \neq B,\ x_0 \neq 0 \right\};$$
$$\mathcal{F}_4^1 : \left\{ p = \frac{x_0}{k(B-A)},\ \gamma_1 = \frac{p}{k},\ \gamma_2 = \frac{q}{k}, \gamma_3 = \frac{r}{k}, A \neq B = C,\ x_0 \neq 0 \right\},$$

where s, q, r are parameters. Under the permutation

$$q \leftrightarrow r, \quad \gamma_2 \leftrightarrow \gamma_3, \quad B \leftrightarrow C, \quad y_0 \leftrightarrow z_0, \quad t \to -t. \tag{5}$$

families $\mathcal{F}_3^1 \leftrightarrow \mathcal{F}_2^1$, the families $\mathcal{F}_1^1$ and $\mathcal{F}_4^1$ are invariant under that automorphism.

Let's apply our approach to the case $\ell = 1$. In this case $x_0 \neq 0, y_0 = z_0 = 0$. Now we study local integrability for the families $\mathcal{F}_1^1$, $\mathcal{F}_2^1$ and $\mathcal{F}_4^1$, because family $\mathcal{F}_3^1$ becomes family $\mathcal{F}_2^1$ under the permutation (5).

3.2. Conditions of local integrability for family $\mathcal{F}_1^1$

So, for the family $\mathcal{F}_1^1$ we have the following:
coefficients a_4 and a_2 of the characteristic polynomial $\chi(\mathcal{F}_1^1)$ are

$$a_4 = \frac{(B+C)\,x_0 + s^2\left(A^2 - AB - AC + 2CB\right)}{BC},\ a_2 = \frac{\left((A-C)s^2 + x_0\right)\left((A-B)s^2 + x_0\right)}{BC};$$

numerator G of the discriminant $D_\lambda(\chi)$ is

$$G(\mathcal{F}_1^1) = A^2\,(A-C-B)^2\,s^4 + 2\,(BA + AC - 4CB)\,(A-C-B)\,x_0\,s^2 + (B-C)^2\,x_0^2,$$

the secondary discriminants Δ_ξ are the following

$$\Delta_{s^2}\left(\mathcal{F}_1^1\right) \cong (A-2C)^2\,(A-2B)^2\,(A-B-C)^6\,(B-C)^2\,A^2B^2C^2x_0^6,$$
$$\Delta_A\left(\mathcal{F}_1^1\right) \cong g_{1A}\,(B-C)^2\,x_0^3B^2C^2s^{12},$$
$$\Delta_B\left(\mathcal{F}_1^1\right) \cong (A-2C)^2\,g_{1B}x_0^2C\,s^2.$$

According to permutation $B \leftrightarrow C$ we obtain

$$\Delta_C\left(\mathcal{F}_1^1\right) \cong (A-2B)^2\,g_{1C}x_0^2Bs^2,$$

where

$$g_{1A} = 2\,(B+C)^3\,s^6 + 3\,(C-5B)\,(B-5C)\,s^4x_0 + (24B+24C)\,s^2x_0^2 + 16x_0^3,$$
$$g_{1B} = (A-C)s^2 + x_0, \quad g_{1C} = (A-B)s^2 + x_0.$$

3.3. Conditions of local integrability for families $\mathcal{F}_2^1$, $\mathcal{F}_3^1$

For the family $\mathcal{F}_2^1$ we have the following.

Coefficients of the characteristic polynomial $\chi(\mathcal{F}_2^1)$ are

$$a_4 = \frac{\left(2AB - AC - CB + C^2\right)s^2}{A\,k^2B} + \frac{\left(A^2 - 2AB - 2AC + 3CB + C^2\right)x_0^2}{B\,k^2\,(A-C)^2\,C},$$
$$a_2 = \frac{\left(C(A-C)^2s^2 + (4C-3A)x_0^2\right)(B-C)\,s^2}{C\,(A-C)\,BA\,k^4}.$$

Connection between parameter k and other parameters from the set Ξ is

$$(A-C)^2\,k^4 = x_0^2 + s^2\,(A-C)^2\,.$$

Numerator G of the discriminant $D_\lambda(\mathcal{F}_2^1)$ is the following

$$\begin{aligned}G(\mathcal{F}_2^1) = C^4\,(A-C)^4\,(A+B-C)^2\,s^4+\\ 2AC\,(A-C)^2\,(A+B-C)\left(2A^2B-A^2C-6ABC+2A\,C^2+5B\,C^2-C^3\right)s^2x_0^2\\ +A^2\left(A^2-2AB-2CA+3CB+C^2\right)^2x_0^4\end{aligned}$$

and the secondary discriminants Δ_ξ are

$$\begin{aligned}\Delta_s\left(\mathcal{F}_2^1\right) \cong&(B-C)^2\,(A-2C)^4\,(A-C)^{16}\,(A+B-C)^6\left(A^2-2AB-2AC+3BC+C^2\right)^2\times\\ &\times A^6B^2C^8x_0^{12},\\ \Delta_A\left(\mathcal{F}_2^1\right) \cong&(B-C)^5h_{2A}g_{2A}^2B^8C^{19}x_0^{20}s^8,\\ \Delta_B\left(\mathcal{F}_2^1\right) \cong&(A-2C)^2\,(A-C)^5\,g_{2B}A^3C\,x_0^4s^2,\\ \Delta_C\left(\mathcal{F}_2^1\right) \cong&f_{13}(s,x_0,A,B)g_{2C}^2A^{25}B^{10}x^{28}s^{26},\end{aligned}$$

where

$$\begin{aligned}h_{2A} =&16B^5C^7s^8-8C^5B^3\left(5B^2-10BC-27C^2\right)s^6x_0^2\\ &+3BC^3\left(291B^4+4C\,B^3-638B^2C^2+612BC^3+243C^4\right)s^4x_0^4\\ &-8BC\,(3B+C)\left(9B^3+249B^2C-557BC^2+171C^3\right)s^2x^6+16\,(3B+C)^4\,x_0^8,\\ g_{2A} =&2(B-C)x_0^2+C^2(B+C)s^2,\\ g_{2B} =&(4C-3A)\,x_0^2+C\,(A-C)^2\,s^2,\\ g_{2C} =&8\,(2B-A)\,x_0^2+A^2\,(A+2B)\,s^2,\end{aligned}$$

and $f_{13}(s,x_0,A,B)$ is a very cumbersome expression.

According to permutation $B\leftrightarrow C$ all secondary discriminants Δ_ξ for family $\mathcal{F}_3^1$ are symmetric to the corresponding discriminants of the family $\mathcal{F}_2^1$ (see (7) and (8) below).

3.4. Local integrability for the family $\mathcal{F}_4^1$

For family $\mathcal{F}_4^1$ of SP the coefficient a_2 of the characteristic polynomial $\chi(\mathcal{F}_4^1)$ equals to zero. So, we have here four zero eigenvalues. Matrix $M(\mathcal{F}_4^1)$ has only one 2×2 Jordan block, and Hypothesis 3 is applicable, and gives local integrability for the family $\mathcal{F}_4^1$.

3.5. Integrable cases for $\ell = 1$

According to Subsections 3.2, 3.3, 3.4 local integrabilities for families $\mathcal{F}_j^1$, $j=1,2,3$, are

$$\begin{aligned}\mathcal{F}_1^1:&L_{11}=\{B=C\},L_{12}=\{A=2C\},L_{13}=\{A=2B\},L_{14}=\{A=B+C\};\\ \mathcal{F}_2^1:&L_{21}=\{B=C\};L_{22}=\{A=2C\},L_{23}=\{A=C\},L_{24}=\{C=A+B\};\\ &L_{25}=\left\{A^2-2AB-2AC+3BC+C^2=0\right\}.\\ \mathcal{F}_3^1:&L_{31}=\{B=C\},L_{32}=\{A=2B\},L_{33}=\{A=B\},L_{34}=\{B=A+C\},\\ &L_{35}=\left\{A^2-2AB-2AC+3BC+B^2=0\right\},\end{aligned}\tag{6}$$

For all families $\mathcal{F}_j^1$, $j=1,2,3$, there is a case $L_{11}\equiv L_{21}\equiv L_{31}$, i.e. $B=C$. So, it is an integrable case. For families $\mathcal{F}_1^1$ and $\mathcal{F}_2^1$, there are cases $L_{12}\equiv L_{22}$, i.e. $A=2C$, and for family $\mathcal{F}_3^1$ there is a case L_{33}, i.e. $A=B$. So, case $A=B=2C$ is integrable, similarly the case $A=C=2B$ is also integrable. Any other combinations of cases of local integrability either contradict to the triangle inequalities (2), or give the case $B=C$. These cases of integrability are known.

In (6) we considered only factors, which does not depend on s and x_0. Let us take into account factors, which do depend on mentioned above parameters:

$$\begin{aligned}
&\mathcal{F}_1^1\colon g_{1A},\ g_{1B},\ g_{1C};\\
&\mathcal{F}_2^1\colon g_{2A},\ g_{2B},\ g_{2C},\ h_{2A},\ f_{13}(s,x_0,A,B);\\
&\mathcal{F}_3^1\colon g_{3A},\ g_{3B},\ g_{3C},\ h_{3A},\ f_{13}(s,x_0,A,C),
\end{aligned} \tag{7}$$

where

$$g_{3A}(B,C)=g_{2A}(C,B),\quad g_{3B}(B,C)=g_{2C}(C,B),\quad g_{3C}(B,C)=g_{2B}(C,B),\\ h_{3A}(B,C)=h_{2A}(C,B). \tag{8}$$

Using factor $A=B+C$ for family $\mathcal{F}_1^1$, factors $g_{2B}=0$ and $g_{3B}=0$ for $\mathcal{F}_2^1$ and $\mathcal{F}_3^1$, we obtain a new integrable case $C/B\approx 2.837994222$. Similarly we found other 8 integrable cases with

$$C/B\in\{0.9805246687, 1.003244577, 0.06376380978, 0.9970470612, 1.000408184,\\ 0.3901207502, 0.9660040562, 1.006388720\}.$$

But they are only particular cases of integrability. Searching for all of them is a big computational work.

Cases $\ell=2$ and $\ell=3$ will be considered in lecture.

Author Contributions: Conceptualization, writing—original draft preparation, Alexander D. Bruno; computations, writing—review and editing, Alexander B. Batkhin. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The author declare no conflict of interest.

References

[1] A. D. Bruno, A. B. Batkhin, New cases of integrability of the Euler-Poisson system, in: N. N. Vassiliev (Ed.), International Conference Polynomial Computer Algebra '2024. St. Petersburg, April 15-20, 2024. Euler International Mathematical Institute, VVM Pubishing, St. Petersburg, Russia, 2024, pp. 28–32.

[2] A. D. Bruno, Analysis of the Euler–Poisson equations by methods of power geometry and normal form, Journal of Applied Mathematics and Mechanics 71 (2007) 168–199. doi:10.1016/j.jappmathmech.2007.06.002.

[3] V. F. Edneral, Integrable cases of the polynomial Liénard-type equation with resonance in the linear part, Mathematics in Computer Science 17 (2023). doi:10.1007/s11786-023-00567-6.

[4] V. V. Golubev, Lectures on Integration of the Equations of Motion of a Rigid Body about a Fixed Point, Moscow State Publishing House Of Theoretical ..., Moscow, 1960. (in Russian).

[5] L. Stolovitch, F. Verstringe, Holomorphic normal form of nonlinear perturbations of nilpotent vector fields, Regular and Chaotic Dynamics 21 (2016) 410–436. doi:10.1134/s1560354716040031.

Contributed talks

Matrices of Infinite Series: Checking Non-Singularity Based on Truncations

Sergei A. Abramov, Anna A. Ryabenko

FRC CSC RAS, ul. Vavilova 40, Moscow, 119333 Russia

Abstract

Square matrices over a ring of formal power series are considered. The matrix entries (series) are given "approximately": for each of them, only a finite number of initial terms are known, and these numbers may not coincide for different matrix entries. Issues related to the possibility of guaranteeing the non-singularity of a matrix for which only this kind of "approximate" representation is known are discussed.

Keywords

strongly non-singular matrices, truncated power series, computer algebra

1. Introduction

In [1], an algorithm is proposed that is applicable to an arbitrary non-singular square matrix $P = (p_{ij})$, in which all entries p_{ij} are polynomials in x over a field K; the algorithm allows one to check whether the matrix $P + Q$ is non-singular for any matrix Q with entries from the ring $K[[x]]$ of formal power series that, first, is of the same size as P and, second, has the property

$$\operatorname{val} Q \geqslant 1 + \deg P. \tag{1}$$

Recall that the valuation $\operatorname{val} f(x)$ of a power series or polynomial is the lowest power of x that has a nonzero coefficient in $f(x)$ (for example, for $f(x) = -3x^2 + 5x^3 + \dots$ we have $\operatorname{val} f(x) = 2$; by definition $\operatorname{val} 0 = \infty$). The valuation of a matrix composed of power series or polynomials is the smallest of the valuations of all entries of this matrix. The degree of a matrix composed of polynomials is considered to be the largest of all the degrees of the entries, wherein $\deg 0 = -\infty$.

A matrix P for which the matrix $P + Q$ is non-singular whenever for a matrix Q of the same size as P, the inequality (1) is satisfied, is called in [1] *strongly non-singular*. A strongly non-singular matrix remains non-singular when we add "tails" to its entries, turning the polynomials into series with coefficients in K. When adding "tails", it is necessary that the lowest power (valuation) of each such "tail" exceed $\deg P$.

Let a matrix P be non-singular and be obtained by truncating some matrix

$$M = (m_{ij}), \tag{2}$$

whose entries are formal power series, and the degree of truncation for all entries is equal to a fixed non-negative integer d: all terms of higher degrees than d are removed.

Then, if $d = \deg P$, using the algorithm proposed in [1], we can check whether the matrix M (we do not know this matrix) cannot be singular. For this purpose, the strong non-singularity of the matrix P is checked.

Remark 1. If d is greater than a power of some p_{ij}, then it is implied that the coefficients at the powers $\deg p_{ij} + 1, \dots, d$ in m_{ij} are equal to zero.

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*

10.22363/12585-2025-6-004

EDN: AMDTKY

sergeyabramov@mail.ru (S. A. Abramov); anna.ryabenko@gmail.com (A. A. Ryabenko)

We further consider the problem of checking the strong non-singularity in a more general version, in comparison with [1], by expanding the very concept of the strong non-singularity, omitting the assumption that the degrees of truncations of all entries of the original matrix are the same and, in particular, assuming that the degree of any (and even each) entry of the matrix P can be less than d — see Remark 1.

Below, P is a non-singular $n \times n$-matrix whose entries are polynomials in x over a field K. (The degrees of these polynomials may differ from each other.)

2. Preliminaries

In [1] the criterion was proved, i.e. a necessary and sufficient condition for the strong non-singularity of the matrix P:

Proposition 1. *Let $P = (p_{ij})$, $p_{ij} \in K[x]$, be a non-singular polynomial $n \times n$-matrix and let d be an integer such that $d \geqslant \deg p_{ij}$, $i, j = 1, \ldots, n$. Then the matrix $P + Q$ is non-singular for any $n \times n$-matrix Q with entries from the ring $K[[x]]$, possessing property* $\operatorname{val} Q \geqslant d + 1$ *if and only if*

$$d + \operatorname{val} P^* \geqslant \operatorname{val} \det P, \tag{3}$$

where P^* is the cofactor matrix of P.

This criterion will be used substantially further in the algorithm for solving the problem under consideration.

For a field K, we assume that there exists an algorithm that checks the consistency of an arbitrary given polynomial system of equations over K, provided that the coefficients of the system belong to K and the system has a solution with components in K. In the case of an algebraically closed field, such an algorithm is based on Gröbner bases [2].

Together with the matrix P, the integers $d_{ij} \geqslant -1$, $i, j = 1, \ldots, n$ are considered given. Let

$$P = (p_{ij}), \quad p_{ij} \in K[x], \quad \deg p_{ij} \leqslant d_{ij}, \quad d = \max_{i,j} d_{ij}. \tag{4}$$

With respect to the entry m_{ij} of the matrix (2), it is assumed that $m_{ij} = p_{ij} + q_{ij}$, where $q_{ij} \in K[[x]]$ is some series such that $\operatorname{val} q_{ij} \geqslant d_{ij} + 1$. The equality $d_{ij} = -1$ means that we have no information about the entry m_{ij} of the matrix M, except for its membership in $K[[x]]$.

Example 1. Let the matrix $\begin{pmatrix} 2x + x^2 & 0 \\ 0 & x \end{pmatrix}$ and the truncation degrees $d_{11} = 2$, $d_{12} = 0$, $d_{21} = 2$, $d_{22} = 1$ be given. These data define the truncated matrix $\begin{pmatrix} 2x + x^2 + O(x^3) & O(x^0) \\ O(x^3) & x + O(x^2) \end{pmatrix}$. The notation $O(x^k)$ is used for some (unspecified) formal series, whose valuation is greater than or equal to k.

This problem of checking the non-singularity of a matrix M is more difficult than that considered in [1], where none of the added terms can have a degree less than $d + 1$. Thus, the added term does not affect the initial terms of the determinant.

The solution of this new problem is reduced below to a series of consistency checks (in other words, — checks for the presence of solutions, solvability) of systems of polynomial equations.

We assume that among the integers d_{ij} there are non-negative ones. Thus, $d \geqslant 0$ in (4).

If all d_{ij} are equal to each other, then the matrix P will be called *flat*.

3. Clarification of the concept of the strong non-singularity of a matrix

Definition 1. A matrix P satisfying conditions (4) will be called *strongly non-singular* if for any $n \times n$-matrix $Q = (q_{ij})$ whose entries belong to $K[[x]]$, $\operatorname{val} q_{ij} \geqslant d_{ij} + 1$, the matrix $P + Q$ is non-singular. (If $d_{ij} = -1$, then q_{ij} can be any element of $K[[x]]$, and for any q_{ij} the matrix $P + Q$ is non-singular.)

It will be shown that for a given matrix P, it is possible to find out algorithmically whether this matrix is strongly non-singular in the sense of Definition 1.

Definition 2. Let (4) hold. For a matrix P, its *K-completion* is the matrix obtained by adding to each p_{ij} for $d_{ij} < d$ some monomials of degrees $d_{ij} + 1, \ldots, d$ with coefficients from K. *Formal completion* is a completion in which all added monomials have different indefinite (symbolic) coefficients.

It is easy to see that the following proposition is true:

Proposition 2. *A matrix P is strongly non-singular if and only if any of its K-completions (each of which is a flat matrix) is strongly non-singular.*

Example 2. For the matrix from Example 1, we get $d = 2$ and $\tilde{P} = \begin{pmatrix} 2x + x^2 & t_1 + t_2 x + t_3 x^2 \\ 0 & x + t_4 x^2 \end{pmatrix}$ is its formal completion where t_1, t_2, t_3, t_4 are different unknowns. We get $\det \tilde{P} = 2x^2 + (2t_4 + 1)x^3 + t_4 x^4$. In this case, the valuation of $\det \tilde{P}$ doesn't depend on the unknowns. Then this valuation is equal to 2 for any K-completions and, since $d = 2$, condition (3) holds for any K-completions. Then the given matrix is strongly non-singular.

Example 3. Now consider a more complex case.

For the truncated matrix $P = \begin{pmatrix} x + O(x^3) & O(x^2) \\ 1 + O(x^3) & x + x^2 + O(x^3) \end{pmatrix}$, we get $d = 2$.

Then $\tilde{P}(t) = \begin{pmatrix} x & tx^2 \\ 1 & x + x^2 \end{pmatrix}$ (t is an unknown), and $\det \tilde{P}(t) = x^3 + (-t + 1)x^2$. In this case, the valuation of $\det \tilde{P}(t)$ depends on t. Then this valuation is equal to 2 or 3 for different K-completions and condition (3) holds or does not hold. To check the strong non-singularity of P, we have to consider different K-completions.

In what follows, considering statements that are true for all solutions of some polynomial system of equations, we assume that if the system is empty (does not contain equations), then any set of values of the unknowns is its solution.

Theorem 1. *Let $p(x; t_1, \ldots, t_m)$ be a polynomial in x whose coefficients are polynomials over K in unknowns $t_1, \ldots, t_m$. Let S be a polynomial over K consistent system of equations (possibly empty) with respect to unknowns $t_1, \ldots, t_m$. There exists an algorithm (let us call it $\mathcal{A}$) that allows one to find all values of* $\operatorname{val} p$ *that are realized (arise) for some values $t_1, \ldots, t_m \in K$ that satisfy the system S. Together with each found value ν of the valuation, this algorithm finds a consistent system S_ν of polynomial over K equations, the solutions of which are all those solutions of the system S for which* $\operatorname{val} p = \nu$. *The result of the algorithm is a list of* conditional valuations *of the form $(\nu,\ S_\nu)$.*

4. Checking the strong non-singularity of a matrix

Algorithm $\mathcal{A}$ and criterion (3) lead to an algorithm for solving the main problem. The idea is as follows: using $\mathcal{A}$, for the formal completion $\tilde{P}$ of the matrix P find all pairs

$$\operatorname{val} \det \tilde{P},\ \operatorname{val} (\tilde{P})^*, \tag{5}$$

that are realized for certain specific K-completions of the original matrix; for each of these pairs, check the satisfying of inequality (3). This may indicate the existence of a K-completion of P that is a singular matrix, or a matrix for which (3) does not hold. Then P is not a strongly non-singular matrix. If this does not happen, then any K-completion yields a strongly non-singular matrix, and by Proposition 2, P is strongly non-singular.

Example 4. For P from Example 3, we get $\operatorname{val}\det\tilde{P}(t) = 2$ if $t \neq 1$ (since $d = 2$, condition (3) holds). Otherwise, we get K-completion $\tilde{P}(1) = \begin{pmatrix} x & x^2 \\ 1 & x+x^2 \end{pmatrix}$ with $\det\tilde{P}(1) = x^3$, $\operatorname{val}\det\tilde{P}(1) = 3$. The cofactor matrix for $\tilde{P}(1)$ is $\begin{pmatrix} x+x^2 & -1 \\ -x^2 & x \end{pmatrix}$, its valuation is equal to 0. Condition (3) doesn't hold for $\tilde{P}(1)$, and therefore P is not strongly non-singular.

Above, in Example 2, the case of a strongly non-singular matrix was considered.

The algorithm for checking the strong non-singularity of a matrix with elements in the form of truncated series has been implemented by us in the environment of the computer algebra Maple 2025 (see [4]) as the *StronglyNonSingular* procedure. The Maple library containing this procedure, as well as the Maple session with examples of using the procedure, is available at [5].

To determine the consistency of polynomial over $\mathbb{Q}$ systems that arise during the operation of the procedure, the *Solve* procedure of the *Groebner* package built into Maple is used (see [3]).

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to M. Kauers (University of Linz) for useful advice. The authors also thank Maplesoft (Waterloo, Canada) for consultations and discussions.

References

[1] S. A. Abramov, M. A. Barkatou, On strongly non-singular polynomial matrices, in: C. Schneider, E. Zima (Eds.), Advances in Computer Algebra, Springer International Publishing, Cham, 2018, pp. 1–17.

[2] B. Buchberger, Gröbner bases: An algorithmic method in polynomial ideal theory, in: N. K. Bose (Ed.), Multidimensional systems theory: Progress, directions and open problems in multidimensional systems, Reidel Publishing Company, 1985, pp. 184–232.

[3] Maple, Groebner, Solve, 2018. URL: https://www.maplesoft.com/support/help/Maple/view.aspx?path=Groebner%2fSolve.

[4] Maple, online help, 2025. URL: http://www.maplesoft.com/support/help/.

[5] A. Ryabenko, StronglyNonSingular procedure, 2025. URL: http://www.ccas.ru/ca/_media/StronglyNonSingularSeriesMatrix.zip.

Formal Stability Investigation in Hamiltonian Systems with Resonances

Alexander B. Batkhin[1,2], Zafar Kh. Khaydarov[3,*]

[1]Keldysh Institute of Applied Mathematics, Miusskaya Sq. 4, Moscow, 125047, Russian Federation
[2]Technion – Israel Institute of Technology, Technion City, 3200003 Haifa, Israel
[3]Samarkand State University named after Sh.Rashidov, 140103 Samarkand, Uzbekistan

Abstract
A method for studying formal stability of equilibrium points in multiparameter Hamiltonian systems with three degrees of freedom is proposed in the case of resonances present. We detail a technique for the symbolic computation of resonance conditions, employing computer algebra and power geometry to represent the resonant variety as a rational algebraic curve and provide its polynomial parametrization.

Keywords
Hamiltonian system, equilibrium state, normal form, formal stability, resonance condition, elimination ideal

1. Introduction

Resonances are fundamental to the behavior of oscillatory systems, significantly impacting their dynamics and stability. This study investigates the formal stability analysis within a multiparameter Hamiltonian framework possessing three degrees of freedom, with particular emphasis on the calculation of resonance conditions. The employment of computer algebra alongside power geometry facilitates the symbolic representation of these conditions in the form of rational algebraic curves.

Here we focus on studying stability in the case of resonances of the third and fourth orders of multiplicity 1 (see Definition 1). Earlier in [4] a schematic description of a method was proposed to study the formal stability of the stationary point of a Hamiltonian system. The drawback of this approach is that it does not take into account multi-frequency resonances of order three or more, which appear in systems with more than two degrees of freedom. This drawback was corrected in papers [1, 2], where a method for computing the parametric representation of a resonant variety of an arbitrary 3-frequency resonance of multiplicity 1 was proposed. However, in those papers, the problem of stability in the presence of such a resonance was left aside. Here we correct for this issue.

The principal characteristics can be outlined as follows.

1. Definition of a resonant variety for a specific resonant vector using the characteristic polynomial coefficients of the linear Hamiltonian system.
2. Utilization of computer algebra and power geometry techniques to obtain polynomial parametrization of the resonant variety.
3. Application of these results to investigate the formal stability regions of equilibrium in a Hamiltonian multiparameter system, including the presentation of a non-trivial example.

6th International Conference "Computer Algebra", Moscow, June 23–25, 2025
10.22363/12585-2025-6-005
EDN: AMTNNK
*Corresponding author.
batkhin@technion.ac.il (A. B. Batkhin); zafarxx@gmail.com (Z. Kh. Khaydarov)
0000-0001-8871-4697 (A. B. Batkhin); 0000-0002-2580-4887 (Z. Kh. Khaydarov)

2. Normal form and scheme of formal stability investigation

We consider an autonomous Hamiltonian system with an analytic function $\gamma(\boldsymbol{\xi}, \boldsymbol{\eta})$ and with n degrees of freedom (DOF)

$$\dot{\boldsymbol{\xi}} = \frac{\partial \gamma}{\partial \boldsymbol{\eta}}, \quad \dot{\boldsymbol{\eta}} = -\frac{\partial \gamma}{\partial \boldsymbol{\xi}}, \tag{1}$$

whose *stationary point* (SP) coincides with the origin $\boldsymbol{\zeta} \stackrel{\text{def}}{=} (\boldsymbol{\xi}, \boldsymbol{\eta}) = 0$.

If the Hamilton function $\gamma(\boldsymbol{\zeta})$ is analytic at this point, then it expands into a convergent power series

$$\gamma(\boldsymbol{\zeta}; \mathbf{P}) = \sum \gamma_{\mathbf{pq}}(\mathbf{P}) \boldsymbol{\xi}^{\mathbf{p}} \boldsymbol{\eta}^{\mathbf{q}}, \tag{2}$$

where $\mathbf{p},\ \mathbf{q} \in \mathbb{Z}^n$, $\mathbf{p}, \mathbf{q} \geqslant 0$, $\gamma_{\mathbf{pq}}(\mathbf{P})$ are constant coefficients that can depend smoothly on a parameter vector $\mathbf{P} \in \Pi$.

Definition 1. *For each resonance, the following are defined:*

- ***multiplicity*** $\mathfrak{k}$*: the number of linearly independent solutions* $\mathbf{p} \in \mathbb{Z}^n$ *to the resonant equation*

$$\langle \mathbf{p}, \boldsymbol{\lambda} \rangle = 0; \tag{3}$$

- ***order*** $\mathfrak{q}$*:* $\mathfrak{q} = \min |\mathbf{p}|$ *over* $\mathbf{p} \in \mathbb{Z}^n \backslash \{0\}$*, satisfying Eq. (3);*
- k***-frequency resonance****: if exactly* k *nonzero eigenvalues* λ_j *are included in the nontrivial solution of the resonance equation;*
- ***strong resonances*** *are called the resonances of orders 2, 3 or 4.*

Definition 2. *Condition* A_k^n *for a system with* n *DOF takes place if the resonant equation (3) has no integer solutions* $\mathbf{p} \in \mathbb{Z}^n$ *with* $|\mathbf{p}| \leqslant k$.

This condition means that there are no resonances up to and including the order k.

In particular, under the condition A_2^n we have $g = \langle \boldsymbol{\rho}, \boldsymbol{\lambda} \rangle + \tilde{g}^{(3)}(\mathbf{z}, \bar{\mathbf{z}})$, and under the condition A_4^n we have

$$g = \langle \boldsymbol{\rho}, \boldsymbol{\lambda} \rangle + \langle C\boldsymbol{\rho}, \boldsymbol{\rho} \rangle + \tilde{g}^{(5)}(\mathbf{z}, \bar{\mathbf{z}}), \tag{4}$$

where C is $n \times n$ matrix.

Theorem 1 (Dirichlet). *If Condition* A_2^n *in Definition 2 is satisfied and the values* $\Re\lambda_j$, $j = 1, \ldots, n$, *are of the same sign, then the stationary point* $\boldsymbol{\zeta} = 0$ *is stable according to Lyapunov.*

In the cases of resonances of multiplicity 1 and order 3 or 4 so called *Birkhoff stability* or instability are considered as Lyapunov stability of the truncated up to the 4th order normal form.

Theorem 2 ([7]). *For resonance of order* 3 *with resonant vector* $\mathbf{p}$*, the truncated up to the 3rd order NF* $\tilde{g}$ *takes the form* $\tilde{g} = \langle \boldsymbol{\rho}, \boldsymbol{\lambda} \rangle + 2A\sqrt{\boldsymbol{\rho}^{\mathbf{p}}} \cos \psi$, $\psi = \langle \mathbf{p}, \boldsymbol{\varphi} \rangle$*. If* $A \not\equiv 0$ *then the SP is Lyapunov unstable. For resonance of order* 4 *with resonant vector* $\mathbf{p}$*, the truncated up to the 4th order NF* $\tilde{g}$ *takes the form* $\tilde{g} = \langle \boldsymbol{\rho}, \boldsymbol{\lambda} \rangle + 2A\sqrt{\boldsymbol{\rho}^{\mathbf{p}}} \cos \psi + \langle C\boldsymbol{\rho}, \boldsymbol{\rho} \rangle$, $\psi = \langle \mathbf{p}, \boldsymbol{\varphi} \rangle$*. If* $A < S \stackrel{\text{def}}{=} |\langle C\mathbf{p}, \mathbf{p} \rangle| / (2\sqrt{\mathbf{p}^{\mathbf{p}}})$ *then this condition is sufficient for Lyapunov stability of the SP, otherwise it is unstable.*

Definition 3 ([8]). *A stationary point* $\boldsymbol{\zeta} = 0$ *of a real Hamiltonian system (1) is formally stable if there exists a (possibly diverging) power series* $f(\boldsymbol{\zeta})$ *that is a formal positive definite first integral* $\{f, \gamma\} = 0$.

Formal stability means that the divergence of solutions from the vicinity of SP is very slow: slower than any finite degree of t.

Let $\mathcal{L} \subset \mathbb{R}^n$ be a linear shell of integers $\mathbf{q}$ satisfying equation $\langle \boldsymbol{\alpha}, \mathbf{q} \rangle = 0$, and $\mathcal{Q} = \{\mathbf{q} \geqslant 0, \mathbf{q} \neq 0\} \subset \mathbb{R}^n$ is a nonnegative orthant without the origin.

Theorem 3 (Formal Stability Theorem [5]). *If Condition A_4^n is satisfied and in* (4)

$$\langle C\mathbf{q}, \mathbf{q}\rangle \neq 0 \text{ for } \mathbf{q} \in \mathcal{L} \cap \mathcal{Q},$$

then the point $\boldsymbol{\zeta} = 0$ is formally stable in the sense of Definition 3.

We finally discuss the situation when any resonance of multiplicity 1 takes place, i.e. there exists the only integral vector $k\mathbf{p}$, $k \in \mathbb{Z}\backslash\{0\}$, $\mathbf{p} \in \mathbb{Z}^n$, satisfying the resonant equation (3). Let $\boldsymbol{\omega}_j$, $j = 1, \ldots, n-1$, be the basis of the orthogonal complement to the one-dimensional solution space, then $\langle \boldsymbol{\omega}_j, \boldsymbol{\rho}\rangle$ is the first integral of the normalized system with Hamiltonian $g(\mathbf{z}, \bar{\mathbf{z}})$ [6, Ch.I, Sect. 3]. Note that in the case of resonance of multiplicity 1, the normal form g is always integrable since it has n independent first integrals: $\langle \boldsymbol{\omega}_j, \boldsymbol{\rho}\rangle$, $j = 1, \ldots, n-1$, and G from equation $g(\mathbf{z}, \bar{\mathbf{z}}) = g_2(\mathbf{z}, \bar{\mathbf{z}}) + G(\mathbf{z}, \bar{\mathbf{z}})$.

Theorem 4 ([3]). *If there exists only one resonant vector $k\mathbf{p}$, $k \in \mathbb{Z}$, which does not belong to the positive orthant $\mathcal{Q}$, then SP $\boldsymbol{\zeta} = 0$ is formally stable.*

Remark 1. *If the coordinate structure of the resonant vector $\mathbf{p}$ does not meet the criteria outlined in Theorem 4, it is necessary to examine the definiteness of the normal form portion $g_3 + g_4$ along the ray where the integral $\mathcal{I} \equiv 0$ is defined, as referenced in [7]. Should this examination reveal positive definiteness, a formally real positive definite integral $\mathcal{I} + G^2$ is achieved.*

Theorem 4 has an application area essentially wider than Theorem 2.

3. Description of the method

Here we describe a method for investigating formal stability of the SP for a multiparameter Hamiltonian system with 3 degrees of freedom. Consider a Hamiltonian system in the vicinity of the SP for which the following conditions are satisfied:

- the number of degrees of freedom of the system is greater than two;
- the quadratic form γ_2 in expansion (2) is nondegenerate and is not definite;
- the Hamiltonian function γ smoothly depends of the vector of parameters $\mathbf{P}$.

Corollary 1 (of Formal Stability Theorem 3). *If under the condition of Theorem 3 in $\mathbb{R}^3$ the intersection of the plane $\langle \boldsymbol{\lambda}, \mathbf{q}\rangle = 0$ and the cone $\langle C\mathbf{q}, \mathbf{q}\rangle$ either does not belong to $\mathcal{Q}$, or belongs to $\mathcal{Q} = \mathbb{R}^3_+$, but does not contain the integral vector $\mathbf{q}$, then the stationary point is formally stable.*

To examine the formal stability of a stationary point of a Hamiltonian system (1), we should:

- find in the space of parameters Π the stability set Σ of the linear part of system (1)
- find such domains, in which the quadratic form $\gamma_2(\mathbf{z})$ is not sign definite
- find parts S_k in these domains that do not contain strong resonances
- normalize the Hamiltonian in each of these parts S_k up to order four, and then
- apply Formal Stability Theorem 3

The border of the stability set Σ in generic case is the resonant variety $\mathcal{R}_3(1, 1, 0)$. For $n = 3$ the borders between the parts S_k are defined by the following resonant varieties:

- $\mathcal{R}_3^{(2,1,0)}$, $\mathcal{R}_3^{(3,1,0)}$ corresponding two-frequency resonances and
- $\mathcal{R}_3^{(1,1,1)}$, $\mathcal{R}_3^{(2,1,1)}$ corresponding three-frequency resonances

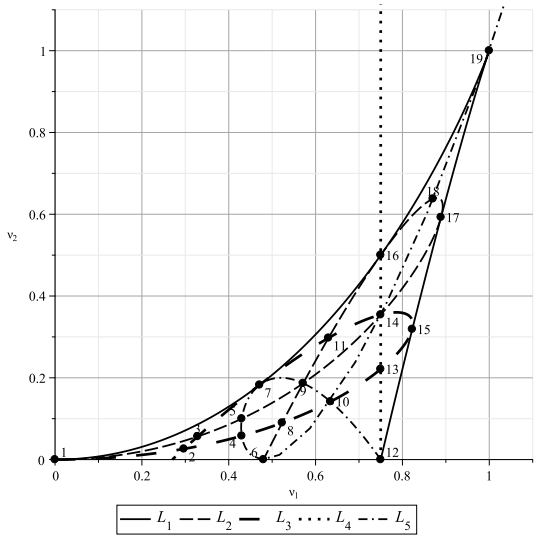


Figure 1: Resonance manifolds in variables ν_1, ν_2.

The implicit and parametric representation of all resonant varieties is given in [3]. The projective representation is shown in Fig. 1.

The application of the method would be demonstrated on an nontrivial example in the lecture.

Author Contributions: Conceptualization, writing—review and editing, Alexander B. Batkhin; computation, writing—original draft preparation, Zafar Kh. Khaydarov. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] A. B. Batkhin, Z. K. Khaydarov, Calculation of a strong resonance condition in a Hamiltonian system, Computational Mathematics and Mathematical Physics 63 (2023) 687–703. doi:10.1134/S0965542523050068.

[2] A. B. Batkhin, Z. K. Khaydarov, Symbolic computation of an arbitrary-order resonance condition in a Hamiltonian system, Programming and Computer Software 49 (2023) 842–853. doi:10.1134/S0361768823080030.

[3] A. Batkhin, Z. Khaydarov, Resonances and formal stability investigation in Hamiltonian systems, Mathematics in Computer Science 18 (2024) 22. doi:10.1007/s11786-024-00596-9.

[4] A. D. Bruno, A. B. Batkhin, Survey of eight modern methods of Hamiltonian mechanics, Axioms 10 (2021) 1–32. doi:10.3390/axioms10040293.

[5] A. D. Bruno, Formal stability of Hamiltonian systems, Math. Notes 1 (1967) 216–219.

[6] A. D. Bruno, The Restricted 3-body Problem: Plane Periodic Orbits, Walter de Gruyter, Berlin, 1994. = Nauka, Moscow, 1990. 296 p. (in Russian).

[7] L. G. Khazin, On the stability of Hamiltonian systems in the presence of resonances, J. Appl. Math. Mech. 35 (1971) 384–391.

[8] J. K. Moser, New aspects in the theory of stability in Hamiltonian systems, Comm. Pure Appl. Math. 11 (1958) 81–114.

Parametric Expansions of an Algebraic Variety Near its Singularities

Alexander D. Bruno[1], Alijon A. Azimov[2,*]

[1]*Keldysh Institute of Applied Mathematics, Miusskaya Sq. 4, Moscow, 125047, Russian Federation*
[2]*Department of Algebra and Geometry, Samarkand State University named after Sh. Rashidov,Samarkand 140104, Uzbekistan*

Abstract
In works [2, 3], parametric expansions near 5 singular points and 3 curves consisting of singular points were computed for a two-dimensional algebraic variety Ω. This paper presents general methods for computing expansions of the variety near its singularities.

Keywords
algebraic variety, singular point, local parametrization, power geometry

1. Introduction

While investigating homogeneous spaces in [1], a set Ω of parameter values a_1, a_2, a_3 was encountered, described by the equation:

$$\begin{aligned} Q(s_1,s_2,s_3) \equiv &(2s_1+4s_3-1)\left(64s_1^5-64s_1^4+8s_1^3+240s_1^2s_3-1536s_1s_3^2-\right.\\ &\left.-4096s_3^3+12s_1^2-240s_1s_3+768s_3^2-6s_1+60s_3+1\right)-\\ &-8s_1s_2(2s_1+4s_3-1)(2s_1-32s_3-1)(10s_1+32s_3-5)-\\ &-16s_1^2s_2^2\left(52s_1^2+640s_1s_3+1024s_3^2-52s_1-320s_3+13\right)+\\ &+64(2s_1-1)s_2^3(2s_1-32s_3-1)+2048s_1(2s_1-1)s_2^4=0, \end{aligned}$$

where s_1, s_2, s_3 are elementary symmetric polynomials, given by: $s_1 = a_1 + a_2 + a_3$, $s_2 = a_1a_2 + a_1a_3 + a_2a_3$, $s_3 = a_1a_2a_3$. In [4], for symmetry reasons, a transition was made from coordinates $\mathbf{a} = (a_1, a_2, a_3)$ to coordinates $\mathbf{A} = (A_1, A_2, A_3)$ via a linear transformation $\mathbf{a}^T = M\mathbf{A}^T$, where M is a specific matrix.

Let $\varphi(X)$ be a polynomial, $X = (x_1, \dots, x_n)$. A point $X = X^0$ in the set $\varphi(X) = 0$ is called a **singular point of order k** if at this point all partial derivatives of the polynomial $\varphi(X)$ with respect to $x_1, \dots, x_n$ up to order k vanish and at least one partial derivative of order $k+1$ does not vanish.

In [4], all singular points of the variety Ω were found in the coordinates $\mathbf{A} = (A_1, A_2, A_3)$: five third-order points $P_1^{(3)} = (0, 0, 3/4)$, $P_2^{(3)} = (0, 0, -3/2)$, …, $P_5^{(3)} = (1, 1, 1/2)$; three second-order points $P_1^{(2)}, P_2^{(2)}, P_3^{(2)}$; and three algebraic curves of first-order singular points $\mathcal{F}, \mathfrak{I}, \mathcal{K}$. The points $P_3^{(3)}, P_4^{(3)}, P_5^{(3)}$ are similar, as they transform into each other under a rotation of the plane A_1, A_2 by $2\pi/3$ around the origin, as do all points $P_1^{(2)}, P_2^{(2)}, P_3^{(2)}$. The curves $\mathcal{F}, \mathfrak{I}, \mathcal{K}$ each have two similar counterparts. Therefore, it suffices to study the variety Ω in the neighborhoods

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-006
EDN: AOCHHS
*Corresponding author.
abruno@keldysh.ru (A. D. Bruno); azimov_alijon_akhmadovich@mail.ru (A. A. Azimov)
http://brunoa.name/ (A. D. Bruno)
0000-0002-7465-1258 (A. D. Bruno); 0000-0002-7799-2525 (A. A. Azimov)

of the points $P_1^{(3)}, P_2^{(3)}, P_5^{(3)}, P_3^{(2)}$ and the curves $\mathcal{F}, \mathfrak{I}, \mathcal{K}$. Here, we consider the structure of Ω near the singular point $P_1^{(3)}$.

2. Structure of the variety near the singular point $\mathbf{P_1^{(3)}}$

Near the point $P_1^{(3)} = (0, 0, 3/4)$, we introduce local coordinates x_1, x_2, x_3:

$$A_1 = x_1, \quad A_2 = x_2, \quad A_3 = \frac{3}{4} + x_3.$$

From the polynomial $R(\mathbf{A}) \stackrel{\text{def}}{=} Q(\mathbf{s}) = 0$, we obtain a 12th-degree polynomial:

$$S(x_1, x_2, x_3) = R(\mathbf{A}) = Q(s_1, s_2, s_3).$$

We compute its support, the Newton polyhedron Γ_1, its faces $\Gamma_j^{(2)}$, and their outer normals using the `PolyhedralSets` package in the `Maple 2021` computer algebra system (CAS). We obtain 5 faces $\Gamma_j^{(2)}$. The graph of the polyhedron Γ_1 is shown in Fig. 1.

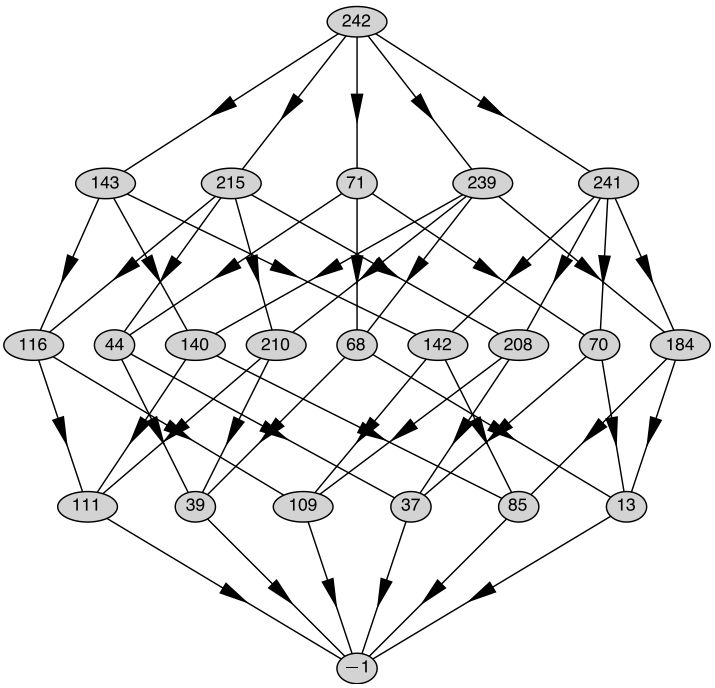


Figure 1: Graph of the polyhedron Γ_1.

The top row shows the entire polyhedron, the next row shows all two-dimensional faces, followed by edges, vertices, and the empty set at the bottom. The outer normals to its two-dimensional faces are:

$$N_{71} = \left(-1, -1, -\frac{1}{2}\right), \; N_{143} = (1, 1, 1), \; N_{215} = (-1, 0, 0), \quad N_{239} = (0, -1, 0), \; N_{241} = (0, 0, -1).$$

Only one normal, N_{71}, has all negative coordinates. Therefore, the neighborhood of the point $x_1 = x_2 = x_3 = 0$ is approximately described by the zeros of the truncated polynomial:

$$\widehat{f}_1 = -\frac{4096}{81}x_3^8 + \frac{3}{4}x_1^4 + \frac{3}{4}x_2^4 + \frac{64}{3}x_1^2x_3^4 - \frac{16}{3}x_1^3x_3^2 + \frac{64}{3}x_2^2x_3^4 - \frac{16}{3}x_2^3x_3^2 + \frac{3}{2}x_1^2x_2^2 + 16x_1^2x_2x_3^2 + 16x_1x_2^2x_3^2,$$

corresponding to face 71 with normal $N_{71} = (-2, -2, -1)$. We find a unimodular matrix:

$$\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & -2 & 1 \end{pmatrix},$$

such that $N_{71}\alpha = (0, 0, -1)$. Consequently, we perform a power transformation: $(\ln y_1, \ln y_2, \ln y_3) = (\ln x_1, \ln x_2, \ln x_3)\,\alpha$, i.e., $(\ln x_1, \ln x_2, \ln x_3) = (\ln y_1, \ln y_2, \ln y_3)\,\alpha^{-1}$.

Since $\alpha^{-1} = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 2 & 1 \end{smallmatrix}\right)$, we have:

$$x_1 = y_1 y_3^2, \quad x_2 = y_2 y_3^2, \quad x_3 = y_3. \tag{1}$$

At the same time, $\widehat{f}_1(x_1, x_2, x_3) = y_3^8 \cdot F_1(y_1, y_2)$, where

$$F_1(y_1, y_2) = -\frac{4096}{81} + \frac{3}{4}y_1^4 + \frac{3}{4}y_2^4 + \frac{64}{3}y_1^2 - \frac{16}{3}y_1^3 + \frac{64}{3}y_2^2 - \frac{16}{3}y_2^3 + \frac{3}{2}y_1^2y_2^2 + 16y_1^2y_2 + 16y_1y_2^2. \tag{2}$$

According to the `algcurves` program in the CAS `Maple`, the curve $F_1(y_1, y_2) = 0$ has genus 0 and the parametrization:

$$\begin{aligned} y_1 = b_1(t) = &- 8(21434756829626557083983t^4 + 14170747278915941772025 60t^3 + \\ &+ 31706038193372580461588706t^2 + 335726200061958227448792184t + \\ &+ 8333103427347345384379)/\delta, \\ y_2 = b_2(t) = &- 56(30534309009669314405 69t^4 + 19840750299173693831 6080t^3 + \\ &+ 38835332085532533132581 58t^2 + 91935591048204912797158 48t - \\ &- 26226282218333750665865 0323)/\delta, \\ \delta = &9\left(85576987369t^2 + 3099727166140t + 37630556816821\right)^2, \end{aligned} \tag{3}$$

The graph of this curve is shown in Fig. 2.

This is a curvilinear triangle with vertices:

$$(y_1, y_2) = -\frac{8}{3}(1, 1), \quad -\frac{8}{3}\left(-\frac{1+\sqrt{3}}{2}, \frac{\sqrt{3}-1}{2}\right), \quad -\frac{8}{3}\left(\frac{\sqrt{3}-1}{2}, -\frac{\sqrt{3}+1}{2}\right).$$

Now, to describe the structure of the variety Ω near the point $P_1^{(3)}$, we perform the power transformation (1) in the polynomial S and obtain the polynomial $T(y_1, y_2, y_3)$. It decomposes into the sum:

$$T(y_1, y_2, y_3) = y_3^8 \sum_{k=0}^{m} T_k(y_1, y_2) y_3^k,$$

where $T_0(y_1, y_2) = F_1(y_1, y_2)$. Using the Maple command `coeff(f,y[k],m)`, we select monomials containing y_k^m. For $k = 3$ and $m = 1$, we obtain:

$$\begin{aligned} T_1 \overset{\text{def}}{=} G(y_1, y_2) = 8y_1^4 + 16y_1^2y_2^2 + 8y_2^4 - \frac{1216}{27}y_1^3 + \frac{1216}{9}y_1^2y_2 + \frac{1216}{9}y_1y_2^2 - \frac{1216}{27}y_2^3 + \\ + \frac{3584}{27}y_1^2 + \frac{3584}{27}y_2^2 - \frac{65536}{729}. \end{aligned} \tag{4}$$

In the polynomials $T_k(y_1, y_2)$, we substitute:

$$y_1 = b_1(t) + \varepsilon, \quad y_2 = b_2(t) + \varepsilon. \tag{5}$$

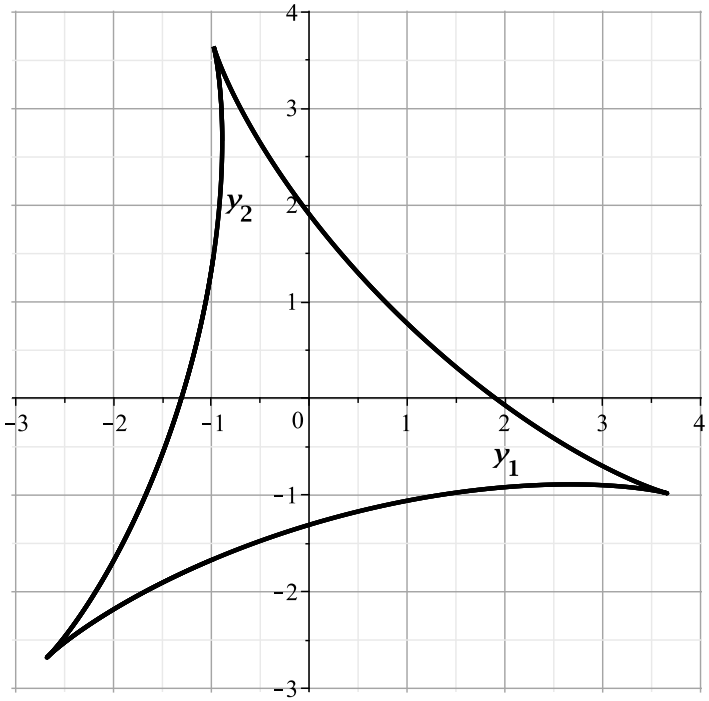


Figure 2: Graph of the curve $F_1(y_1, y_2) = 0$.

We obtain the polynomial $u(\varepsilon, y_3) = T(y_1, y_2, y_3)/y_3^8$ with coefficients depending on t through $b_1(t)$ and $b_2(t)$. In this polynomial:

$$u(\varepsilon, y_3) = \sum_{k=0}^{m} T_k(b_1 + \varepsilon, b_2 + \varepsilon) y_3^k = \sum_{p,q \geq 0} u_{pq} \varepsilon^p y_3^q,$$

where $u_{00} = F_1(b_1(t), b_2(t))$ from (2), hence $u_{00} = 0$.

By the implicit function theorem (see [2]), the equation $u(\varepsilon, y_3) = 0$ has a solution in the form of a power series in y_3:

$$\varepsilon = \sum_{k=1}^{\infty} c_k(t) \cdot y_3^k, \tag{6}$$

where $c_k(t)$ are rational functions expressed in terms of the coefficients $u_{pq}(t)$, which in turn are expressed in terms of $b_1(t)$ and $b_2(t)$.

Author Contributions: Conceptualization, writing—review and editing, Alexander D. Bruno; computation, writing—original draft preparation, Alijon A. Azimov. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] N. A. Abiev, Y. G. Nikonorov, The evolution of positively curved invariant Riemannian metrics on the Wallach spaces under the Ricci flow, Annals of Global Analysis and Geometry 50 (2016) 65–84. doi:10.1007/s10455-016-9502-8.

[2] A. D. Bruno, A. A. Azimov, Parametric expansions of an algebraic variety near its singularities, Axioms 12 (2023) 469. doi:10.3390/axioms12050469.
[3] A. D. Bruno, A. A. Azimov, Parametric expansions of an algebraic variety near its singularities II, Axioms 13 (2023) 106. doi:10.3390/axioms13020106.
[4] A. D. Bruno, A. B. Batkhin, Investigation of a real algebraic surface, Programming and Computer Software 41 (2015) 74–82. doi:10.1134/S0361768815020036.

Block Fermat Numbers in Modular Arithmetic

Benjamin Chen, Eugene Zima*

University of Waterloo, Waterloo, Canada

Abstract

We show how to select balanced in size moduli of the form $2^n + 1$ to achieve accelerated modular reduction and reconstruction. Experimental implementation demonstrates that this choice of moduli provides better performance of modular arithmetic compared to the popular moduli of the form $2^n - 1$.

Keywords

accelerated modular arithmetic, fast modular reduction, fast reconstruction, scalable moduli, Fermat type numbers and polynomials

1. Introduction

Conversion to a modular representation is a popular technique to accelerate the arithmetic of computer algebra systems. The choice of specific values of moduli can significantly influence the time complexity of both the calculations modulo m_i and the reconstruction of the result. One of the oldest examples of such an approach is described in [7, 6]: several relatively prime moduli of the form $2^n - 1$ (sometimes referred to as Mersenne type moduli) are selected (using $\gcd(2^n - 1, 2^m - 1) = 1$ if and only if $\gcd(n, m) = 1$). This replaces division with remainder in the residue computation by shift and addition operations that are much simpler. In [8] this was expanded to moduli of the forms $2^n + 1$ (sometimes referred to as Fermat type moduli). Two schemes for generating the moduli set: "shift" scheme and "block" scheme were considered, both based on the fact that $\gcd(2^n + 1, 2^m + 1) = 1$ if and only if $\nu_2(n) \neq \nu_2(m)$ (here $\nu_2(x)$ is binary valuation of a natural number x). One important feature of Fermat type moduli is that (unlike in the case of Mersenne type moduli) the relative primality of two moduli is preserved under scaling (i.e. when 2 is replaced by 2^c for a natural c). Another feature that was explored in [8] for shift scheme is that inverses that appear during reconstruction from modular representation [2, 6, 3] preserve bit pattern under scaling: for moduli of the form $m_i = 2^{a2^i} + 1, \ \ i = 0, 2, \ldots, k$, ($a$ is an arbitrary positive integer) it was shown that

$$M_i^{-1} \bmod m_i = 2^{a2^i - 1} - 2^{a-1} + 1, \ \ i = 1, 2, ..., k, \tag{1}$$

where $M_i = \prod_{j=0}^{i-1} m_j, \ \ i = 0, 1, \ldots, k-1$. With this choice of moduli, there is no need to (pre-)compute and to store inverses. Inverse is defined by the value of a (which is the same for all moduli) and the index i. This allows reconstruction to become essentially multiplication-free. Shift-based moduli are also scalable (with scaling factor a) which means informally that setting a to an arbitrary value will not change the relative primality of the moduli m_i and will not change the sparsity (number of set bits) of the inverses in sparse balanced binary representation [4, 5]. However, the moduli m_i in the shift scheme are imbalanced in size: the bit length of m_i is larger than the bit-length of the product $m_0 m_1 \ldots m_{i-1}$, which makes the practicality of such a choice questionable.

The "block" schemes proposed in [8] generates more balanced in size moduli with exponents of the same bit length, however scalability of inverses was not explored in [8] and performance

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*

10.22363/12585-2025-6-007

EDN: AQXBEC

*Corresponding author.

benjamin.chen@uwaterloo.ca (B. Chen); ezima@uwaterloo.ca (E. Zima)

of block scheme Fermat type moduli reported by authors was similar to the performance of Mersenne type moduli.

We present new results related to the scalability of inverses for Fermat type moduli. In fact we established that any pair or relatively prime moduli $2^n + 1$,$2^m + 1$ has scalable inverses. This makes block scheme more competitive compared to similar in size set of Mersenne type moduli, which is also confirmed by our GMP experimental implementation.

2. Fermat polynomials and their properties

It is more convenient to represent a modulus as the result of the evaluation of a polynomial. For example, $m = 2^n + 1$ can be viewed as $f(2)$ for $f(x) = x^n + 1$, scaled modulus $m = (2^c)^n + 1$ can be viewed as $f(2^c)$. The following properties of Fermat polynomials are useful to select a block of relatively prime Fermat type moduli with scalable inverses: consider two Fermat type polynomials $f(x) = x^n + 1$ and $g(x) = x^m + 1$ in $\mathbb{Q}[x]$.

1. Then
$$\gcd(x^n + 1, x^m + 1) = \begin{cases} 1, & \nu_2(n) \neq \nu_2(m) \\ x^{\gcd(n,m)} + 1, & \text{otherwise} \end{cases}$$
2. If $gcd(f(x), g(x)) = 1$ then for any integer $\ell \geq 1$ numbers $f(2^\ell)$ and $g(2^\ell)$ are relatively prime.
3. If $gcd(f(x), g(x)) = 1$ then the coefficients of Bézout cofactors $s(x)$ and $t(x)$ in equation
$$s(x)f(x) + t(x)g(x) = 1$$
are dyadic with the numerators ± 1 and denominators equal to 2. Thus $2^{\ell n} + 1$ and $2^{\ell m} + 1$ is scalable pair of moduli with scalable inverses for any $\ell \geq 2$.
4. If the cofactors $s(x), t(x)$ have more than 1 term, the difference in adjacent degrees of the terms in the cofactors ordered by the degree is $h = \gcd(n, m)$. Specifically, $s(x) = \pm\frac{1}{2}x^{m-h} \pm \frac{1}{2}x^{m-2h} \pm \cdots \pm \frac{1}{2}$, $t(x) = \pm\frac{1}{2}x^{n-h} \pm \frac{1}{2}x^{n-2h} \pm \cdots \pm \frac{1}{2}$.
5. If $gcd(f(x), g(x)) = 1$ then there exist polynomials $p(x)$ with $\deg p(x) \leq m$ and $q(x)$ with $\deg g(x) \leq n$ having coefficients from the set $\{0, \pm 1/2, \pm 1\}$ such that $f(2^\ell)^{-1} \bmod g(2^\ell) = p(2^\ell)$, $g(2^\ell)^{-1} \bmod f(2^\ell) = q(2^\ell)$.

The last property means that the number of set bits in the scaled inverses is the same as the support of polynomials $p(x)$ and $q(x)$ and will be preserved under scaling. For example, if $f(x) = x^{15} + 1$ and $g(x) = 2^{20} + 1$ then $q(x) = \frac{1}{2}x^{15} + \frac{1}{2}x^{10} + \frac{1}{2}x^5 + 1$ and for any integer $\ell \geq 1$

$$g(2^\ell)^{-1} \mod f(2^\ell) = q(2^\ell) = 2^{15\ell-1} + 2^{10\ell-1} + 2^{5\ell-1} + 1.$$

3. Selecting better block of Fermat type moduli

If we need to find a moduli set of size b, there exist many possible "block" style moduli sets. Two greedy schemes to generate the exponents of a moduli set of size b are presented in [8]:

1. $e_k = 2^b - 2^{k-1}$ for $k = 1, 2, \ldots, b$.
2. $e_k = 2^{b-1} + 2^{b-k-1}$ for $k = 1, 2, \ldots, b-1$ and $e_b = 2^{b-1}$.

The key to generating a moduli set is to ensure that every exponent has a different binary valuation. For a moduli set of size b, as long as we keep the binary valuation of the exponents different, we have the freedom to choose the bits prior to the first 1-bit counting from the least significant digit in the binary form of the exponents. For example, let $b = 4$, the generated moduli exponents using the first scheme are $15, 14, 12, 8$. The moduli set is $\{2^{15}+1, 2^{14}+1, 2^{12}+$

$1, 2^8+1\}$. The generated moduli exponents using the second scheme are $12, 10, 9, 8$. The moduli set is $\{2^{12}+1, 2^{10}+1, 2^9+1, 2^8+1\}$. However, the exponent set $12, 10, 11, 8$ is also a valid set.

It is worth noting that the two "greedy" block schemes are not optimal in terms of the total amount of set bits in the inverses. This number is the same as the total support (total number of non-zero terms) of Bézout cofactors for every pair of relatively prime Fermat polynomials in the set. For example, if we want to generate a set of moduli of size $b = 6$, the first scheme generates the exponents $\{63, 62, 60, 56, 48, 32\}$. The total number of terms in the inverses is 289. The second scheme generates the exponents $\{48, 40, 36, 34, 33, 32\}$. The total number of terms in the inverses is 233. However, if we consider an exponent set that follows neither of the two schemes, $\{63, 56, 48, 42, 36, 32\}$, the total number of terms in the inverses is 141. Which means that the overhead related to multiplication by these inverses is almost twice as small compared to greedy choice of exponents.

Note, that due to scalability the selection of moduli set of a given size b is conducted only once, and we can use exhaustive search with back-tracking to find a set of b exponents of different binary valuations with the least total support of Bézout cofactors.

The bit-size of the moduli generated by "block" schemes grows exponentially with the number of needed moduli b. For example, to obtain 9 moduli, the moduli have exponents of the magnitude 2^9. To obtain 10 moduli, the moduli have exponents of the magnitude 2^{10}. This makes Fermat type moduli practical only for moderate values of b. Nevertheless, small sets of relatively large moduli with scalable inverses can be used in upper layer of layered modular arithmetic and provide improvement in performance, as shown in [1].

4. Implementation and Benchmark

We compared performance of Mersenne type moduli and Fermat type moduli. The results of the running time on the standard benchmark of integer matrix multiplication are presented in Table 1.

Dim	Bitsize	M Mult	M Overh	M Total	F Mult	F Overh	F Total
8	2^{18}	0.124	0.282	0.407	0.114	0.093	0.207
-	2^{19}	0.287	0.693	0.980	0.297	0.188	0.485
-	2^{20}	0.768	1.583	2.351	0.783	0.376	1.167
16	2^{18}	0.931	1.056	1.988	0.887	0.387	1.274
-	2^{19}	2.229	2.742	4.972	2.271	0.786	3.058
-	2^{20}	5.970	6.338	12.308	6.004	1.571	7.613
32	2^{18}	7.376	4.305	11.683	7.165	1.542	8.709
-	2^{19}	14.882	8.980	23.864	16.124	2.781	18.908
-	2^{20}	39.726	20.806	60.536	39.887	5.652	45.675
64	2^{18}	49.410	14.472	63.890	47.457	5.460	52.925
-	2^{19}	119.274	36.020	155.304	120.622	11.566	132.199
-	2^{20}	318.661	83.032	401.706	318.602	22.857	341.985

Table 1
Timing (in seconds) of matrix multiplication benchmark: Dim - matrix dimension, Bitsize - bitsize of elements of matrices, M - 7 moduli of the form $2^n - 1$, F - 7 moduli of the form $2^n + 1$, Mult - the time spent for multiplication in RNS, Overh - the time spent for conversion to and from RNS (note, that column M Overh includes the time to compute inverses, while column F Overh inverses are not computed and obtained from scaling), Total - the total time spent performing computations (measured separately), Hardware used is AMD EPYC 7502P 32C @ 2.5 GHz with 503GB of RAM.

From the table, we can see that Fermat-type moduli outperform Mersenne-type moduli in terms of overhead (conversion to RNS and reconstruction from RNS) due to the scalability of the inverses.

5. Conclusion and open question

Block Fermat type moduli provide the ability to scale up a given moduli set without recomputation of the inverses at the cost of the moduli being only slightly unbalanced. Implementation and standard benchmarks shows significant reduction in the overhead when using Fermat type moduli compared to Mersenne type moduli, aligning with the theoretical performance gain.

Our current approach to finding the optimal block of b exponents is exhaustive search with pruning. The size of search space is proportional to $2^{2^{b-2}}$. Recall that for b moduli of the form $2^n + 1$ to be pairwise relatively prime the exponents must have at least b bits. Greedy schemes of constructing exponents provide sets with reasonably good characteristics, but it would be interesting to know if there is better algorithm to find close to optimal block of balanced in size exponents that guaranties pairwise relative primality of the moduli.

References

[1] B. Chen, Y. Li, E. Zima, On a two-layer modular arithmetic, ACM Commun. Comput. Algebra 57 (2023) 133–136. doi:10.1145/3637529.3637534.

[2] H. L. Garner, The residue number system, in: Papers presented at the the March 3–5, 1959, western joint computer conference on XX - IRE-AIEE-ACM '59 (Western), ACM Press, San Francisco, California, 1959, pp. 146–153. doi:10.1145/1457838.1457864.

[3] K. O. Geddes, S. R. Czapor, G. Labahn, Algorithms for Computer Algebra, Springer US, Boston, MA, 1992. doi:10.1007/b102438.

[4] D. Hankerson, S. Vanstone, A. Menezes, Guide to Elliptic Curve Cryptography, Springer-Verlag, New York, 2004. OCLC: 1058873580.

[5] R. Hewlitt, E. Swartzlantler, Canonical signed digit representation for FIR digital filters, in: 2000 IEEE Workshop on SiGNAL PROCESSING SYSTEMS. SiPS 2000. Design and Implementation (Cat. No.00TH8528), IEEE, Lafayette, LA, USA, 2000, pp. 416–426. doi:10.1109/SIPS.2000.886740.

[6] D. E. Knuth, The art of computer programming. 2: Seminumerical algorithms, Addison-Wesley, Reading, Mass, 1969.

[7] A. Schönhage, Multiplikation großer Zahlen, Computing 1 (1966) 182–196. doi:10.1007/BF02234362.

[8] E. V. Zima, A. M. Stewart, Cunningham numbers in modular arithmetic, Programming and Computer Software 33 (2007) 80–86. doi:10.1134/S0361768807020053.

Non-Minimality of Minimal Telescopers Explained by Residues*

Shaoshi Chen[1,2,*], Manuel Kauers[3], Christoph Koutschan[4], Xiuyun Li[1,2], Rong-hua Wang[5] and Yisen Wang[1,2]

[1]KLMM, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 6 Beijing, 100190, China
[2]School of Mathematical Sciences, University of Chinese Academy of Sciences, 6 Beijing 100049, China
[3]Institute for Algebra, Johannes Kepler University, 6 Linz A-4040, Austria
[4]RICAM, Austrian Academy of Sciences, 6 Linz A-4040, Austria
[5]School of Mathematical Sciences, Tiangong University, 6 Tianjin, 300387, China

Abstract

Elaborating on an approach recently proposed by Mark van Hoeij, we continue to investigate why creative telescoping occasionally fails to find the minimal-order annihilating operator of a given definite sum or integral. We offer an explanation based on the consideration of residues.

Keywords

creative telescoping, residues, symbolic integration, symbolic summation, telescoper

Creative telescoping is the standard approach to definite summation and integration in computer algebra. Its purpose is to find an annihilating operator for a given definite sum $\sum_k f(n,k)$ or a given definite integral $\int_\Omega f(x,y)dy$.

Such operators are obtained from annihilating operators of the summand or integrand that have a particular form. In the case of summation, suppose that we have

$$(L - (S_k - 1)Q) \cdot f(n,k) = 0 \tag{1}$$

for some operator L that only involves n and the shift operator S_n but neither k nor the shift operator S_k, and another operator Q that may involve any of n, k, S_n, S_k. Summing the equation over all k yields

$$L \cdot \sum_k f(n,k) = \big[Q \cdot f(n,k)\big]_{k=-\infty}^{\infty}.$$

If the right-hand side happens to be zero, we find that L is an annihilating operator for the sum.

In the case of integration, having

$$(L - D_y Q) \cdot f(x,y) = 0 \tag{2}$$

for some operator L that only involves x and the derivation D_x but neither y nor the derivation D_y, and some other operator Q that may involve any of x, y, D_x, D_y, implies the equation

$$L \cdot \int_\Omega f(x,y)\,dy = \big[Q \cdot f(x,y)\big]_\Omega.$$

If the right-hand side happens to be zero, we find that L is an annihilating operator for the integral.

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*

*This is the extended abstract of a submission to ISSAC2025.

10.22363/12585-2025-6-008
EDN: AXOIAW

*Corresponding author.
schen@amss.ac.cn (S. Chen); manuel.kauers@jku.at (M. Kauers); christoph.koutschan@oeaw.ac.at (C. Koutschan); lixiuyun@amss.ac.cn (X. Li); wangronghua@tiangong.edu.cn (R. Wang); wangyisen@amss.ac.cn (Y. Wang)

An operator L as in equations (1) and (2) is called a *telescoper* for f, and Q is called a *certificate* for L. The degree of S_n or D_x in L is called the order of L. If L is such that there is no telescoper of lower order, then L is called a *minimal telescoper*. The minimal telescoper is unique up to multiplication by rational functions (from the left).

Algorithms for testing the existence of telescopers and computing them if they exist meanwhile have a long history in computer algebra, see [9, 10, 7, 1, 2, 6] for classical results and recent developments on the matter. In his recent paper [8], van Hoeij proposed a fresh view on creative telescoping. He explains why a telescoper can often be written as a least common left multiple of smaller operators, and why the minimal telescoper is sometimes not the minimal-order annihilating operator for the sum or integral under consideration.

Let C be a field of characteristic zero and $C(n,k)$ be the field of rational functions in n,k over C. Let $A_{n,k} = C(n,k)\langle S_n, S_k\rangle$ be the ring of all linear recurrence operators in S_n, S_k with rational function coefficients, and $A_n = C(n)\langle S_n\rangle$ be the subalgebra consisting of all operators that do not involve k or S_k. For a given summand $f(n,k)$, consider the A_n-module $\Omega := A_{n,k}\cdot f(n,k)$ and the quotient module $M := \Omega/((S_k-1)\Omega)$. An operator $L \in A_n$ is then a telescoper for $H = f(n,k)$ if and only if it is an annihilating operator of the image $\overline{H}$ of H in M.

In this setting, van Hoeij makes the following observations:

- If M can be written as a direct sum of submodules, say $M = M_1 \oplus M_2$, then the minimal telescoper of H is the least common left multiple of the minimal annihilating operators of the projections $\pi_1(H)$ and $\pi_2(H)$ of H in M_1 and M_2, respectively.
- If, moreover, the definite sum whose summand corresponds to $\pi_1(H)$ happens to be zero identically, then every annihilating operator of $\pi_2(H)$ is already an annihilating operator of the definite sum over H, even though it may not be a telescoper for H.

In order to take advantage of the second observation, it is necessary to understand under which circumstances a definite sum can be zero. Such "vanishing sums" are themselves examples when a minimal telescoper fails to be a minimal annihilator. For example, we have

$$\sum_k (-1)^k \binom{2n+1}{k}^2 = 0,$$

so the minimal annihilator is 1. However, the minimal telescoper of $(-1)^k\binom{2n+1}{k}^2$ is

$$L = (2n+3)S_n + (8n+8).$$

Note that since L is irreducible, the module M, which is isomorphic to A_n/LA_n, has no nontrivial submodules.

We propose an explanation of why certain sums are identically zero which is based on the investigation of residues. Also based on residues, we will explain why telescopers tend to be least common left multiples. We are not the first to use residues in the context of creative telescoping. For rational functions and algebraic functions in the differential case, it was observed by Chen, Kauers, and Singer [5] that telescopers and residues are closely related. Chen and Singer also used residues in the context of summation problems [4]. Residues are also tied to creative telescoping through the equivalence of extracting residues with taking diagonals and positive parts and the computation of Hadamard products [3].

References

[1] S. A. Abramov, Applicability of Zeilberger's algorithm to hypergeometric terms, in: Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation, ISSAC '02, Association for Computing Machinery, New York, NY, USA, 2002, pp. 1–7. doi:10.1145/780506.780507.

[2] S. A. Abramov, When does Zeilberger's algorithm succeed?, Adv. Appl. Math. 30 (2003) 424–441. doi:10.1016/S0196-8858(02)00504-3.
[3] A. Bostan, F. Chyzak, M. van Hoeij, M. Kauers, L. Pech, Hypergeometric expressions for generating functions of walks with small steps in the quarter plane, Eur. J. Comb. 61 (2017) 242–275. doi:https://doi.org/10.1016/j.ejc.2016.10.010.
[4] S. Chen, M. F. Singer, Residues and telescopers for bivariate rational functions, Adv. Appl. Math. 49 (2012) 111–133. doi:10.1016/j.aam.2012.04.003.
[5] S. Chen, M. Kauers, M. F. Singer, Telescopers for rational and algebraic functions via residues, in: Proc. ISSAC'12, ACM, New York, 2012, pp. 130–137. doi:10.1145/2442829.2442851.
[6] M. Kauers, D-Finite Functions, Springer, Cham, 2023.
[7] M. Petkovšek, H. S. Wilf, D. Zeilberger, $A = B$, A. K. Peters Ltd., Wellesley, MA, 1996.
[8] M. van Hoeij, Submodule approach to creative telescoping, J. Symbolic Comput. 126 (2025) 102342. doi:https://doi.org/10.1016/j.jsc.2024.102342.
[9] D. Zeilberger, A fast algorithm for proving terminating hypergeometric identities, Discrete Math. 80 (1990) 207–211. doi:10.1016/0012-365X(90)90120-7.
[10] D. Zeilberger, The method of creative telescoping, J. Symbolic Comput. 11 (1991) 195–204. doi:https://doi.org/10.1016/S0747-7171(08)80044-2.

On Relative Primality and Other Properties of Trinomials

Robert Dougherty-Bliss[1,*], Mits Kobayashi[1], Natalya Ter-Saakov[2] and Eugene Zima[3]

[1]*Dartmouth College, Hanover, NH, USA*
[2]*Rutgers University, New Brunswick, NJ, USA*
[3]*University of Waterloo, Waterloo, Canada*

Abstract
We discuss some properties of trinomial polynomials. These can be used to generate new balanced-size sets of moduli for accelerated modular arithmetic.

Keywords
modular arithmetic, cyclotomic polynomials, resultant, dyadic rational.

1. Introduction

A popular technique to speed up computations with integer arithmetic is to reduce the input modulo several relatively prime numbers, compute with the residues, then reconstruct the result with the Chinese remainder theorem [3, 2, 5]. In [1], it was proposed to use "trinomial" moduli of the form

$$2^n - 2^k + 1, \quad 0 < k < n, \tag{1}$$

which are not only pairwise relatively prime, but also have pairwise scalable inverses. By that, we mean that replacing 2 with 2^c does not change the sparsity or bit-pattern of the modular inverses. For example, the moduli $2^{20} - 2^{12} + 1$ and $2^{20} - 2^4 + 1$ satisfy the following property: for any integer $c \geq 1$,

$$\begin{aligned} ((2^c)^{20} - (2^c)^{12} + 1)^{-1} &\equiv (2^c)^8 + 1 && (\text{mod } (2^c)^{20} - (2^c)^4 + 1) \\ ((2^c)^{20} - (2^c)^4 + 1)^{-1} &\equiv (2^c)^{20} - (2^c)^{12} - (2^c)^8 + 1 && (\text{mod } (2^c)^{20} - (2^c)^{12} + 1). \end{aligned}$$

Using ad-hoc methods, the authors of [1] discovered a set of five pairwise relatively prime moduli of this form and used them in upper-layer on top of [2] to show improvement in a standard integer matrix multiplication benchmark.

Their work left open the following questions:

1. When are two trinomial moduli relatively prime?
2. Are there arbitrarily large sets of trinomial moduli with the same bit length?
3. For a fixed bit length, how can we efficiently find these sets?

Our aim is to answer some of these questions by examining the pure polynomial trinomials

$$x^n - x^k + 1, \quad 0 < k < n. \tag{2}$$

It seems reasonable to establish the relative primality of trinomial pairs of the form (2), then use this to deduce the relative primality of moduli of the form (1). Unfortunately, this does not work. Almost all pairs of trinomials of the form (2) are relatively prime over the rationals, yet

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-009
EDN: BCOMHM
*Corresponding author.
rdbliss@dartmouth.edu (R. Dougherty-Bliss); mits.kobayashi@dartmouth.edu (M. Kobayashi); natalya.tersaakov@rutgers.edu (N. Ter-Saakov); ezima@uwaterloo.ca (E. Zima)

only a small proportion of moduli of the form (1) are relatively prime integers. In other words, the substitution $x = 2$ in (2) does not preserve relative primality.

Our results include a simple condition based on the resultant of two trinomials which ensures the relative primality of moduli of the form (1), a proof that arbitrarily large sets of relatively prime moduli exist, and some discussions on how to find such sets.

2. Dyadically resolving pairs

Definition 1. A pair of monic polynomials $f(x)$ and $g(x)$ in $\mathbb{Z}[x]$ *dyadically resolve* if their resultant is a signed power of 2.

If a pair of monic polynomials dyadically resolve, then they are relatively prime in $\mathbb{Q}[x]$. Their unique Bézout cofactors $a(x)$ and $b(x)$ with $\deg a < \deg g$ and $\deg b < \deg f$ in the equation

$$a(x)f(x) + b(x)g(x) = 1$$

will have dyadic coefficients. This is because the resultant is the determinant of the Sylvester matrix of $f(x)$ and $g(x)$, which is the coefficient matrix used to construct $a(x)$ and $b(x)$. The converse is also true, but it is not obvious.

Theorem 1. *The coefficients of the Bézout cofactors of $f(x)$ and $g(x)$ are all dyadic if and only if $f(x)$ and $g(x)$ dyadically resolve.*

If the trinomials $x^n - x^k + 1$ and $x^n - x^j + 1$ dyadically resolve then there exists an integer-coefficient polynomial $p(x)$ such that

$$\begin{gathered}\gcd(2^{cn} - 2^{ck} + 1, 2^{cn} - 2^{cj} + 1) = 1 \\ (2^{cn} - 2^{ck} + 1)^{-1} \bmod (2^{cn} - 2^{cj} + 1) = p(2^c)\end{gathered} \tag{3}$$

for sufficiently large positive integers c. In other words, the corresponding moduli that arise after setting $x = 2^c$ are relatively prime, and their inverses have a stable bit pattern.

We do not know of any simple, widely applicable condition that implies resolvability. For example, observe the sporadic behavior of powers of 2 in the following matrix.

$$\begin{pmatrix}
0 & 1 & 3 & 1 & 3 & 31 & 9 & 8 & 3 \\
1 & 0 & 1 & 1 & 4 & 1 & 31 & 16 & 8 \\
3 & 1 & 0 & 1 & 3 & 1 & 3 & 31 & 9 \\
1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 31 \\
3 & 4 & 3 & 1 & 0 & 1 & 3 & 4 & 3 \\
31 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\
9 & 31 & 3 & 1 & 3 & 1 & 0 & 1 & 3 \\
8 & 16 & 31 & 1 & 4 & 1 & 1 & 0 & 1 \\
3 & 8 & 9 & 31 & 3 & 1 & 3 & 1 & 0
\end{pmatrix}$$

Figure 1: $\operatorname{res}(x^{10} - x^i + 1, x^{10} - x^j + 1)$ for $1 \le i, j \le 9$.

As a first step to understanding these resultants, we report the following results.

Theorem 2. *As $n \to \infty$, there are arbitrarily large sets of trinomials of the form (2) which dyadically resolve pairwise.*

Our proof of this theorem is constructive, meaning that we can write down arbitrarily large sets at will. Unfortunately, the bit lengths used in our proof grow too quickly to be useful in practice. See Table 1.

size	exponents
1	{1}
2	{1, 2}
3	{2, 3, 4}
4	{12, 15, 16, 18}
5	{720, 760, 765, 768, 780}
6	{48372480, 48434496, 48435465, 48435712, 48436128, 48439664}

Table 1
Exponents constructed in the proof of Theorem 2. For each size, the given set consists of k such that the polynomials $x^n - x^k + 1$ dyadically resolve pairwise. The degree n is any fixed integer larger than the largest element of the set.

Theorem 3. *Let $i, j < n$ be positive integers.*

1. *$x^n - x^k + 1$ and $x^n - x^j + 1$ dyadically resolve if $k - j$ divides k.*
2. *$x^n - x^k + 1$ and $x^n - x^j + 1$ dyadically resolve if and only if $x^n - x^{n-k} + 1$ and $x^n - x^{n-j} + 1$ are.*
3. *If 2^v is the largest power of 2 that divides $k - j$, then $x^n - x^k + 1$ and $x^n - x^j + 1$ do not dyadically resolve if 2^v divides either k or $n - k$.*

There is a closed-form expression for the resultant of two binomials which has been known for many years (see [4]), but this is not true for trinomials. We can report a formula in a very special case.

Theorem 4. *If $k - j$ divides n, then*

$$\operatorname{res}(x^n - x^k + 1, x^n - x^j + 1) = \pm \left(\prod_{m \mid \frac{(k-l)}{\gcd(k,k-l)}} \Phi_m(2) \right)^{\gcd(k,k-l)}$$

where $\Phi_m(x)$ is the mth cyclotomic polynomial.

3. Relative primality

For a fixed n, almost all pairs of trinomials of the form (2) are relatively prime. In fact, if n is odd, then all of them are. The following theorem gives very strict conditions under which two trinomials can share a common factor over the rationals.

Theorem 5. *If $g(x) = \gcd(x^n - x^k + 1, x^n - x^j + 1)$ is not constant, then:*

1. *n is even;*
2. *$k - j$ is divisible by 6; and*
3. *$g(x)$ is a product of cyclotomic polynomials whose orders are multiples of 6.*

The preceding result shows that the only common factors two trinomials can have are cyclotomic polynomials. We can further describe exactly which cyclotomic polynomials divide which trinomials.

Theorem 6. *$\Phi_d(x)$ divides $x^n - x^k + 1$ if and only if d is a multiple of 6, and*

$$(n, k) \equiv \pm(d/3, d/6) \pmod{d}.$$

Using these results and some number theory, it is possible to show that roughly 97% of all pairs of trinomials for large even n are relatively prime.

4. Open questions

Our results answer some of the questions inspired by [1], but there remain open questions.

Minimal pairwise resolvability What is the smallest n such that there exists a set of k trinomials of the form (2) with degree n that dyadically resolve pairwise? The below table shows the first few values of this sequence.

set size k	smallest n
2	3
3	5
4	5
5	10
6	11
7	22
8	41
9	82
10	1668
11	???

Resultant of two trinomials Is there a formula for the resultant of two trinomials of the form (2) in terms of the cyclotomic polynomials evaluated at integers?

Fast construction of trinomial moduli What is the fastest way to construct large sets of pairwise dyadically resolving trinomials? Using maximal clique algorithms we have a method which is faster than a naive brute force approach, but there could be better ways.

Acknowledgments: We thank the Digital Research Alliance of Canada and the Office of Research Computing and Data Services at Dartmouth College for computational resources to run experiments.

References

[1] B. Chen, Y. Li, E. Zima, On a two-layer modular arithmetic, ACM Communications in Computer Algebra 57(3) (2023) 133–136. doi:10.1145/3637529.3637534.

[2] J. Doliskani, P. Giorgi, R. Lebreton, E. Schost, Simultaneous conversions with the residue number system using linear algebra, ACM Transactions on Mathematical Software 44(3) (2018) 1–21. doi:10.1145/3145573.

[3] D. E. Knuth, The Art of Computer Programming, Volume 2: Seminumerical Algorithms, Addison-Wesley Professional, 2014.

[4] R. Swan, Factorization of polynomials over finite fields, Pacific Journal of Mathematics 12(3) (1962) 1099–1106.

[5] E. V. Zima, A. M. Stewart, Cunningham numbers in modular arithmetic, Programming and Computer Software 33(2) (2007) 80–86. doi:10.1134/S0361768807020053.

On the Integrability of an ODE System with an Inhomogeneous Right-Hand Side

Victor F. Edneral

Skobeltsyn Institute of Nuclear Physics of Lomonosov Moscow State University, SINP MSU, 1(2), Leninskie gory, GSP-1, Moscow 119991, Russian Federation

Abstract

The report describes the application of the normal form method to the study of a polynomial modification of the Liénard equation with a 4th-order nonlinearity. A two-dimensional set of parameter values is found for which the corresponding system has a first integral.

Keywords

resonance normal form, integrability, Liénard equation, computer algebra

1. Introduction

In previous works we proposed an algorithmic approach that involves reducing the study of integrability of autonomous ODE systems with a polynomial right-hand side to solving a system of polynomial equations for the parameters of the system [1]. Although this approach works effectively in cases of resonances in the linear part of the ODE, it also sometimes allows us to study the non-resonant case as well [6]. The approach is based on the hypothesis that local integrability of an ODE system in the neighborhoods of its fixed points is necessary for its integrability in a certain region of the phase space.

In the works [7, 8] a method was also proposed for reducing the study of the integrability of an ODE system to solving a system of algebraic equations, based on the study of the algebraic possibility of integrating an ODE system using the Darboux method [10]. However, the current version of this method assumes a 1:1 resonance in the linear part, as well as a homogeneous nonlinearity of the polynomial on the right-hand side. In addition, not every definition of integrability can be said to reduce it to Darboux integrability. In this report, we will consider obtaining integrability conditions for the coefficients of an ODE system with a $1:3$ resonance and a nonhomogeneous polynomial on the right-hand side. We will also briefly dwell on the problems that arise when solving this problem.

2. The problem

To demonstrate the method, we chose an autonomous ODE system constructed based on the Liénard equation [9]

$$\frac{d^2\,x(t)}{d\,x^2} + P(x(t))\frac{d\,x(t)}{(d\,x} + Q(x(t)) = 0. \tag{1}$$

In the original notation, this equation assumes that P is even and Q is odd. We will treat these functions as polynomials without defined parity properties and write equation (1) as a

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-010
EDN: BKBNIP
edneral@theory.sinp.msu.ru (V. F. Edneral)

parameterized polynomial system

$$\begin{aligned} \tfrac{d\,x(t)}{d\,t} &= y(t), \\ \tfrac{d\,y(t)}{d\,t} &= (a_0 + a_1 x(t) + a_2 x^2(t))\, y(t) + b_1 x(t) + b_2 x^2(t)\, y + b_3 x^3(t) + b_4 x^4(t), \end{aligned} \tag{2}$$

here t is independent, $x(t)$ and $y(t)$ are dependent variables, and $a_0, \ldots, b_4$ are parameters. Our task is to determine for what values of the parameters the system (2) has a first integral and, thus, is integrable.

3. Calculation of the local integrability condition at the origin

According to the hypothesis about the necessity of local integrability at stationary points of an ODE system for its integrability [5], we need to construct the corresponding algebraic systems of equations for the parameters, and do this for the case of resonance in its linear part [2, 3]. It turned out that if there are several stationary points in the phase space of the system, it is easier to obtain and then solve this condition for one of these points, and then check its fulfillment for the others.

We will be interested in the system (2) with a resonance $1:3$ in the linear part. The resonance $1:M$ takes place there if the relation

$$b_1 = \frac{a_0^2 M}{(M-1)^2}.$$

Calculating further at resonance with $M = 3$ the normal form up to the 12th order, we obtain the first three algebraic equations of the local integrability condition at the origin.

A remarkable phenomenon observed for autonomous polynomial ODE systems is the ”saturation” of the standard basis of equations of the local integrability condition with an increase in their number. This circumstance was discovered for the Liénard and Bautin systems [5, 6] and is manifested in the fact that, starting from a certain order, when higher equations are added to this condition, the solutions of the system cease to change. In our example, it is apparently sufficient to limit ourselves to three equations.We do not write out this system here due to its bulkiness, we will designate it as System1, it is given in the files [11, 12].

Next, it was necessary to obtain rational solutions of the algebraic system System1 of the 24th order from 6 variables $a_0, a_1, a_2, b_2, b_3, b_4$. It was not possible to do this “head-on”.

4. A special case of integrability

In the work [4] an example of an integrable case of the Liénard system was written out

$$\begin{aligned} \tfrac{d\,\tilde{x}(t)}{d\,t} &= y(t), \\ \tfrac{d\,y(t)}{d\,t} &= \left(\tfrac{25}{12} - 3\tilde{x}^2(t)\right) y(t) - \tfrac{125}{432} + \tfrac{25}{36}\tilde{x}(t) + 5\tilde{x}(t) + 7\tilde{x}^3(t) + 3\tilde{x}^4(t). \end{aligned} \tag{3}$$

By shifting one of the fixed points to the origin, this system can be rewritten as (2)

$$\begin{aligned} \tfrac{d\,x(t)}{d\,t} &= y(t), \\ \tfrac{d\,y(t)}{d\,t} &= \left(2 - x(t) - 3x^2(t)\right) y(t) + 3x(t)(1 + x(t))^3, \\ &\text{where } \ \tilde{x}(t) = x(t) + \tfrac{1}{6}. \end{aligned} \tag{4}$$

Substitution of values of system parameters (2) corresponding to case (4)

$$a_0 = 2, a_1 = -1, a_2 = -3, b_1 = 3, b_2 = 9, b_3 = 9, b_4 = 3,$$

$$b_1 = \tfrac{3}{4} a_0^2$$

into the algebraic system System1, satisfies it identically. Thus, we have at least one zero-dimensional solution and it really corresponds to the integrable case. It is interesting to look for families of solutions that include the one found. To do this, we parametrize the system in a similar (4) form

$$\begin{aligned} \tfrac{d\,x(t)}{d\,t} &= y(t), \\ \tfrac{d\,y(t)}{d\,t} &= (c_0 + c_1 x(t) + c_2 x^2(t))\, y(t) + d_1 x(t)(1 + d_2 x(t))^3. \end{aligned} \tag{5}$$

In this case, the parameters in (2) are related to the current ones as follows

$$\begin{aligned} &a_0 = c_0,\ a_1 = c_1,\ a_2 = c_2,\ b_1 = d_1,\ b_2 = 3 d_1 d_2,\ b_3 = 3 d_1 d_2^2,\ b_4 = d_1 d_2^3, \\ &\text{where} \quad d_1 = \tfrac{3}{4} c_0^2. \end{aligned} \tag{6}$$

Replacing variables in the system System1 using formulas (6), we obtain an algebraic system of three equations with respect to parameters c_0, c_1, c_2, d_2. We denote it as System2, see file [11, 12]. Using the Solve procedure of the MATHEMATICA-11 system, we can obtain its rational solutions

$$\{c_1 = 0, c_2 = 0, d_2 = 0\},\ \{c_1 = -\frac{1}{2} c_0\, d_2,\ c_2 = -\frac{3}{2} c_0\, d_2^2\}. \tag{7}$$

These same solutions also identically satisfy System1 with the appropriate change of variables (6).

5. Results

The first of the solutions (7) reduces the system to a linear integrable form and is of no interest, the second gives a system with two free parameters c_0 and d_2

$$\begin{aligned} \tfrac{d\,x(t)}{d\,t} &= y(t), \\ \tfrac{d\,y(t)}{d\,t} &= \tfrac{1}{4} c_0 (d_2 x(t) + 1)[3 c_0 x(t)(d_2 x(t) + 1)^2 + 2 y(t)(3 d_2 x(t) - 2)]. \end{aligned}$$

This system is integrable because it has a first integral of the form

$$\begin{aligned} I(x(t), y(t)) = \quad & \sqrt[3]{c_0 x(t)(d_2 x(t) + 1)^2 + 2 y(t)} \times \\ & \left(\frac{c_0 (d_2 x(t)+1)^2\, {}_2F_1\left(1,1;\tfrac{5}{3};-\frac{c_0 (d_2 x(t)+1)^3}{2 d_2 y(t) - c_0 (d_2 x(t)+1)^2}\right)}{c_0 (d_2 x(t)+1)^2 - 2 d_2 y(t)} + 2 \right). \end{aligned}$$

Here ${}_2F_1$ is a hypergeometric function.

References

[1] A. D. Bruno, V. F. Edneral, On the integrability of a planar system of ODEs near a degenerate stationary point, Journal of Mathematical Sciences 168(3) (2010) 326–333. doi:10.1007/s10958-010-9983-0.

[2] A. D. Bruno, Analytical form of differential equations (I, II), volume 25, 26, Nauka, M., 1971, 1972. Trans. Moscow Math. Soc., vol. 25, 131–288 (1971), vol. 26, 199–239 (1972).
[3] A. D. Bruno, Local Methods in Nonlinear Differential Equations, Springer Verlag, Berlin Heidelberg NewYork London Paris Tokyo, 1989.
[4] M. V. Demina (2020). Private Communications.
[5] V. F. Edneral, Integrable cases of the polynomial Liénard-type equation with resonance in the linear part, Mathematics in Computer Science 17 (2023) 19. doi:10.1007/s11786-023-00567-6.
[6] V. F. Edneral, Integrable cases of the Bautin system, Mathematics in Computer Science (2025). In printing.
[7] B. Ferčec, J. Giné, V. G. Romanovski, V. F. Edneral, Integrability of complex planar systems with homogeneous nonlinearities, Journal of Mathematical Analysis and Applications 434 (2016) 894–914. doi:10.1016/j.jmaa.2015.09.037.
[8] B. Ferčec, M. Žulj, M. Mencinger, On the integrability of persistent quadratic three-dimensional systems, Mathematics 12(9) (2024) 1338. doi:10.3390/math12091338.
[9] A. Liénard, Etude des oscillations entretenues, Revue générale de l'électricité 23 (1928) 901–912 and 946–954.
[10] V. G. Romanovski, D. S. Shafer, The Center and Cyclicity Problems: A Computational Algebra Approach, Birkhüser, Boston, 2009.
[11] Syst1 and Syst2 algebraic systems, 2025. The file "Systems.pdf" is available at https://disk.yandex.ru/i/UQQHOowSeq6y6w.
[12] Syst1 and Syst2 algebraic systems, 2025. The file "Systems.nb" is available at https://disk.yandex.ru/d/JpR1c52cEzKGHQ.

The Dual Quaternion Algebra and its Implementation in Asymptote Language

Migran N. Gevorkyan[1], Dmitry S. Kulyabov[1,2,*], Anna V. Korolkova[1] and Leonid A. Sevastianov[1,2]

[1]RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
[2]Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation

Abstract
The algebras of dual quaternions and screws are often opposed to geometric algebra. The purpose of this paper is to describe the algebra of dual quaternions and the algebra of screws, to give a number of examples of the use of dual quaternions to describe the screw motion of points, lines and planes in three-dimensional space. This algebra is very poorly covered in the literature, and the actively used principle of Kotelnikov-Study transfer is apparently forgotten. All calculations were performed using the Asymptote language. Structures were created that implement dual numbers, quaternions, and dual dual quaternions, as well as a set of computational tests to verify these structures.

Keywords
screws, motors, rotations, translations, computer geometry, Asymptote

1. Introduction

In the course of research on the application of analytical projective geometry in the field of computational geometry [2], the authors often came across references to motors, propellers, and dual dual quaternions. All mentions were very brief and basically boiled down to the fact that the mentioned entities are extremely unintuitive and difficult to understand and use. They were often contrasted with geometric algebra methods, which were presented as more understandable and logical [4].

The search for a detailed description of the mathematical apparatus of the algebra of screws and dual quaternions led the authors to works in the field of mechanics of absolutely rigid bodies. It turned out that the theory of screws was developed back in the late 19th and early 20th centuries in the works of R. S. Ball, E. Study [3], A. P. Kotelnikov. The most complete description can be found in the monograph [1]. However, at present this theory is little known and there are practically no software implementations of screw algebras and dual quaternions.

Another methodological problem is the lack of examples of the application of screws and dual quaternions to computer geometry problems. The sources found are mainly focused on the problems of applied mechanics. In this paper, we have tried to at least partially eliminate this shortcoming.

Since all the examples are focused on the application of dual quaternion algebras and screws to geometric problems, the Asymptote language was chosen as the language for implementation. This language allows you to create custom data structures (data types) and overload all the basic

6th International Conference "Computer Algebra", Moscow, June 23–25, 2025
doi 10.22363/12585-2025-6-011
EDN: BNYSRN
*Corresponding author.
gevorkyan-mn@rudn.ru (M. N. Gevorkyan); kulyabov-ds@rudn.ru (D. S. Kulyabov); korolkova-av@rudn.ru (A. V. Korolkova); sevastianov-la@rudn.ru (L. A. Sevastianov)
https://yamadharma.github.io/ (D. S. Kulyabov)
0000-0002-4834-4895 (M. N. Gevorkyan); 0000-0002-0877-7063 (D. S. Kulyabov); 0000-0001-7141-7610 (A. V. Korolkova); 0000-0002-1856-4643 (L. A. Sevastianov)

algebraic operators. This made it possible to create data types for dual numbers, quaternions, and dual dual quaternions. You can also create second-order functions that return other functions that implement screw motion using Roderig's screw formulas and sandwich products of dual quaternions.

2. Dual numbers and quaternions

Here we list the basic concepts used in the construction of dual quaternionic algebra.

A dual number is a parabolic complex number $z = a + b\varepsilon$, where the imaginary unit is a parabolic imaginary unit defined by the equality $\varepsilon^2 = 0$. For these numbers, you can define the same operations as for the usual (elliptical) complex numbers.

For two numbers $z_1 = a_1 + b_1\varepsilon$ and $z_2 = a_2 + b_2\varepsilon$, addition, subtraction, and multiplication can be defined. The formula for multiplication will look like this:

$$z_1 z_2 = a_1 a_2 + (a_1 b_2 + b_1 a_2)\varepsilon,$$

A dual conjugation is a number $\overline{z} = \overline{a + b\varepsilon} = a - b\varepsilon$. The square of the module of the number $|z|^2 = z\overline{z} = a^2$, and the module itself as $|z| = |a|$.

Quaternion is a hypercomplex number of the form $q = q_0 + q_1 i + q_2 j + q_3 k = q_0 + \mathbf{q}$, where the imaginary units i, j, k are determined by the equality $i^2 = j^2 = k^2 = ijk = -1$. From this equality, we can obtain a multiplication table of i, j, k among themselves and define quaternion multiplication, which is most easily expressed in terms of the scalar and vector parts of the quaternion as follows:

$$qp = q_0 p_0 - (\mathbf{q}, \mathbf{p}) + q_0\mathbf{p} + p_0\mathbf{q} + \mathbf{q} \times \mathbf{p}.$$

A pure quaternion is a quaternion without a scalar part q_0. A pure quaternion is associated with a vector, an ordinary quaternion with $q_0 = 1$ is associated with an affine point, and a quaternion with $q_0 \neq 0.1$ is associated with a point mass. The q_0 component plays the role of a weight coordinate in this case.

3. Dual quaternions

A dual quaternion is a dual number with coefficients in the form of quaternions (Cayley-Dickson doubling procedure):

$$Q = q + q^o\varepsilon,$$

where the quaternion q is the main part, and q^o is the moment part. Q can be written as a number with eight components

$$Q = q_0 + q_1 i + q_2 j + q_3 k + q_0^o\varepsilon + q_1^o i\varepsilon + q_2^o j\varepsilon + q_3^o k\varepsilon.$$

Axiomatically, it is assumed that the parabolic imaginary unit ε commutes with elliptical imaginary units i, j, k, that is, $i\varepsilon = \varepsilon i$, $j\varepsilon = \varepsilon j$, $k\varepsilon = \varepsilon k$.

Simplifying somewhat, we will call a screw a dual quaternion, both parts of which are pure quaternions. We will denote the screw in bold: $\mathbf{Q} = \mathbf{q} + \mathbf{q}^o\varepsilon$.

Three different conjugation operations are defined for a dual quaternion

- $Q^* = (q + q^o\varepsilon) = q^* + q^{o*}\varepsilon$ — quaternionic (complex) conjugation;
- $\overline{Q} = \overline{q + q^o\varepsilon} = q - q^o\varepsilon$ — dual conjugation;
- $Q^\dagger = (\overline{q + q^o\varepsilon})^* = q^* - q^{o*}\varepsilon$ is a quaternion dual conjugation.

dual quaternion multiplication can be defined for dual quaternions, and scalar and screw (vector) multiplications for screws.

4. The principle of Kotelnikov–Study transference

The principle of transference in the form in which A. P. Kotelnikov formulated it states. All formulas of the theory of finite rotations and kinematics of motion of a rigid body with one fixed point, when replacing real quantities in them with dual analogues, turn into formulas for finite displacements and kinematics of motion of a free rigid body.

For example, consider the Rodrigues formula for rotating a point P with a radius vector $\mathbf{p}$ around an axis passing through the origin with a guide vector $\mathbf{a}$ by an angle θ:

$$\mathbf{p}' = \cos\theta\mathbf{p} + \sin\theta\mathbf{a} \times \mathbf{p} + (1 - \cos\theta)(\mathbf{a}, \mathbf{p})\mathbf{a}.$$

According to the principle of transfer, the angle θ should be replaced in this formula by the dual angle $\Theta = \theta + \theta^o\varepsilon$, the radius vector $\mathbf{p}$ by the screw $\mathbf{L} = \mathbf{v} + \mathbf{m}\varepsilon$, guiding vector $\mathbf{a}$ onto screw $\mathbf{A} = \mathbf{a} + \mathbf{a}^o\varepsilon$. The scalar and vector product of the vectors will then be replaced by the scalar and screw product of the screws.

5. Dual quaternion formulas for screw motion

If in the unit quaternion λ for the rotation of the vector (pure quaternion) $\mathbf{p}$ around the axis $\mathbf{a}$ is given by the formula $\lambda = \cos\dfrac{\theta}{2} + \sin\dfrac{\theta}{2}\mathbf{a}$, then replacing the angle θ and the vector $\mathbf{a}$ with dual analogues using the principle of transference, then we get a dual quaternion: $\Lambda = \cos\dfrac{\Theta}{2} + \sin\dfrac{\Theta}{2}\mathbf{A}$ which implements the screw movement of points, straight lines and planes.

The sandwich formula for a straight line represented by a screw $\mathbf{L} = \mathbf{v} + \mathbf{m}\varepsilon$ looks like $\mathbf{L}' = \Lambda\mathbf{L}\Lambda^*$. An affine point is represented using a dual quaternion of the following form: $P = 1 + \mathbf{p}\varepsilon$. Planes can also be written as a dual quaternion $\Pi = \mathbf{n} + d\varepsilon$, where $\mathbf{n}$ is the normal vector of the plane, and d is the distance from the plane to the origin. The same formulas work for the screw motion of a point and a plane: $P' = \Lambda P\Lambda^\dagger$ and $\Pi' = \Lambda\Pi\Lambda^\dagger$.

6. Conclusion

The Kotelnikov-Study transference principle is naturally implemented programmatically if the types of dual numbers, quaternions, and dual quaternions are defined, as well as arithmetic and algebraic operators are overloaded, and scalar and vector multiplications are defined. In this case, the calculation of the screw motion is reduced to a compact program code, since all the computational complexity is already implemented in the created data types. And since the implementation of dual quaternions uses ready-made types of dual numbers and quaternions, part of the complexity is transferred to the implementation of these types, thus distributing the overall complexity at different levels. The specific details of the implementation are planned to be outlined in the presentation of the report.

References

[1] F. M. Dimentberg, The screw calculus and its applications in mechanics, 1968.

[2] M. N. Gevorkyan, A. V. Korolkova, D. S. Kulyabov, L. A. Sevastyanov, Implementation of analytic projective geometry for computer graphics, Programming and Computer Software 50 (2024) 153–165. doi:10.1134/S0361768824020075.

[3] F. Klein, Vorlesungen Über Höhere Geometrie, Springer Berlin Heidelberg, 1926. doi:10.1007/978-3-642-49848-0.

[4] E. Lengyel, Mathematics, volume 1, Terathon Software LLC, 2016. URL: http://foundationsofgameenginedev.com.

Investigation of the Periodic Planar Oscillations of a Two-Body System in an Elliptic Orbit Using the Computer Algebra Methods

Sergey A. Gutnik[1,2]

[1]MGIMO University, 76, Prospekt Vernadskogo, 119454, Moscow, Russia
[2]Moscow Institute of Physics and Technology, Institutskii per. 9, Dolgoprudny, 141701 Russi

Abstract
Computer algebra methods are used to investigate the planar oscillations of a system of two bodies connected by a spherical joint that moves along an elliptic orbit under the action of gravitational torque in the plane of the orbit. It is shown that the motion of the two-body system is described by periodic planar oscillations. All the relevant symbolic computations are performed with the help of computer algebra systems.

Keywords
two connected bodies, spherical joint, gravitational torque, elliptic orbit, Lagrange equations, periodic solution, power series, computer algebra

1. Introduction

We study the dynamics of a two-body system (satellite and stabilizer) connected by a spherical joint that moves in gravitational field in the plane of an elliptical orbit using computer algebra methods. The dynamics of various schemes for satellite-stabilizer gravitational orientation systems on a circular orbit was discussed in many papers, some review of them can be found in papers [12, 10, 3].

Since the problem is very complicated, in the previous works we studied the equilibrium orientations of the system on a circular orbit only in the simplest cases when the spherical joint is located at the intersection of the satellite and stabilizer principal central axis of inertia and in the case where the spherical joint is positioned on the line of intersection between two planes formed by the principal central axes of inertia of the satellite and stabilizer [3, 11, 2, 5, 4]. The application of computer algebra makes it possible to find the solutions of this problem. A detailed investigation of the oscillations of a satellite (a rigid body) in the plane of an elliptical orbit and the conditions for their stability were carried out in [15].

The works devoted to the study of planar oscillations of a system of two coupled bodies on a slightly elliptic orbit were carried out only for simple cases, when the centers of mass of the first and second bodies coincide [7], [1]. Here, we consider the planar oscillations of a two-body system on a slightly elliptic orbit in case when the spherical joint is located at the intersection of the first and second body principal central axis of inertia. Applying the perturbation techniques and appropriate symbolic computations with the help of computer algebra system Wolfram Mathematica [14], we construct the periodic solution in the form of a power series in a small parameter.

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-012
EDN: BPZFFL
s.gutnik@inno.mgimo.ru (S. A. Gutnik)
0000-0002-2598-1699 (S. A. Gutnik)

2. Equations of motion

We consider the problem of two bodies connected by a spherical hinge that move on an elliptic orbit. To write the equations of motion of two-body system, we introduce the following right-handed Cartesian coordinate systems: $OXYZ$ is the orbital coordinate system, the OZ axis is directed along the radius vector connecting the Earth center of mass C and the center of mass O of the two-body system, the OX axis is directed along the linear velocity vector of the center of mass O, and the OY axis coincides with the normal to the orbital plane. The axes of coordinate systems $O_1x_1y_1z_1$ and $O_2x_2y_2z_2$, are directed along the principal central axes of inertia of the first and the second body, respectively. The orientation of the coordinate system $O_ix_iy_iz_i$ with respect to the orbital coordinate system is determined by the aircraft angles α_i (pitch), β_i (yaw), and γ_i (roll) (see [3]).

Suppose that (a_i, b_i, c_i) are the coordinates of the spherical hinge P in the body coordinate system $Ox_iy_iz_i$, A_i, B_i, C_i are principal central moments of inertia; $M_1M_2/(M_1 + M_2) = M$; M_i is the mass of the ith body; ω is the angular velocity for the center of mass of the two-body system moving along an elliptic orbit. Then we use the expressions for kinetic energy of the system in the case when $b_1 = b_2 = c_1 = c_2 = 0$ and the coordinates of the spherical hinge P in the body coordinate systems are $(a_i, 0, 0)$ and when the motions of the two-body system are located in the plane of the elliptic orbit ($\alpha_1 \neq 0, \alpha_2 \neq 0, \beta_1 = \beta_1 = 0, \gamma_1 = \gamma_2 = 0, \dot{\alpha_1} = d\alpha_1/dt$, $\dot{\alpha_2} = d\alpha_2/dt$, where t is time) in the form [12]

$$\begin{aligned} T &= 1/2\big(B_1 + Ma_1^2\big)(\dot{\alpha_1} + \omega)^2 + 1/2\big(B_2 + Ma_2^2\big)(\dot{\alpha_2} + \omega)^2 \\ &- Ma_1a_2\cos(\alpha_1 - \alpha_2)(\dot{\alpha_1} + \omega)(\dot{\alpha_2} + \omega). \end{aligned} \tag{1}$$

The force function, which determines the effect of the Earth gravitational field on the system of two connected by a hinge bodies, is given by [12]

$$\begin{aligned} U &= -3\mu/(2\rho^3)\big((A_1 - C_1)\sin^2\alpha_1 + (A_2 - C_2)\sin^2\alpha_2\big) \\ &+ 3/2M\mu/\rho^3\big((a_1\sin\alpha_1 - a_2\sin\alpha_2\big)^2 + M\mu/\rho^3a_1a_2\cos(\alpha_1 - \alpha_2). \end{aligned} \tag{2}$$

Here ρ is a radial distance between the center of mass of the Earth C and center of mass of the system O; $\mu = fM_0$, where f is a gravitational constant, and M_0 is the mass of the Earth; $\omega = \frac{d\vartheta}{dt} = \omega_0(1 + e\cos\vartheta)^2$; $\frac{\mu}{\rho^3} = \omega_0^2(1 + e\cos\vartheta)^3$; ϑ is the true anomaly and e is the orbital eccentricity. On the circular orbit $\omega = \omega_0$, $\frac{\mu}{\rho^3} = \omega_0^2$, $\vartheta = \omega_0 t$.

By using the kinetic energy expression (1) and the expression (2) for the force function, the equations of motion for this system can be written as Lagrange equations of the second kind by applying symbolic differentiation in the Wolfram Mathematica system [14], [13]

$$\frac{d}{dt}\frac{\partial T}{\partial\dot{\alpha_i}} - \frac{\partial T}{\partial\alpha_i} - \frac{\partial U}{\partial\alpha_i} = 0, \quad i = \overline{1,2}, \tag{3}$$

in the form of a system of second-order ordinary differential equations in variables α_1 and α_2 [12].

To find the periodic solutions we assume that the oscillations are small and replace the sine and cosine by their expansions in power series. Doing the substitution $dt = d\vartheta/(\omega_0(1+e\cos\vartheta)^2)$ in Lagrange equations we change the independent variable from t to ϑ and transform the Lagrange system to the form

$$\begin{aligned} &- (1 + e\cos\vartheta)\alpha_2'' + 2e\alpha_2'\sin\vartheta + (B_1 + Ma_1^2)/(Ma_1a_2)\big((1 + e\cos\vartheta)\alpha_1{}'' \\ &- 2e\alpha_1'\sin\vartheta\big) - e(1 + e\cos\vartheta)(\alpha_2' + 1)^2 + e\big(2\sin\vartheta(1 - (B_1 + Ma_1^2)/Ma_1a_2) \\ &+ (4 + 3\big((A_1 - C_1) - Ma_1^2\big)/(Ma_1a_2))\big) = 0, \end{aligned} \tag{4}$$

$$
\begin{aligned}
&- \quad (1+e\cos\vartheta)\alpha_1'' + 2e\alpha_1'\sin\vartheta + (B_2+Ma_2^2)/(Ma_1a_2)\big((1+e\cos\vartheta)\alpha_2{}''\\
&- \quad 2e\alpha_1'\sin\vartheta\big) + e(1+e\cos\vartheta)(\alpha_1'+1)^2 + e\big(2\sin\vartheta(1-(B_2+Ma_2^2)/Ma_1a_2)\\
&+ \quad (2+3\big((A_2-C_2)-Ma_2^2\big)/(Ma_1a_2))\big) = 0.
\end{aligned}
$$

which determine the oscillations of the system in the plane of the elliptic orbit in the orbital coordinate system. The prime in (4) denotes differentiation with respect to ϑ. One can easily check that the system (4) has the stationary solution

$$\alpha_1 = \alpha_2 = 0. \tag{5}$$

Our goal is to obtain the periodic solution of the equations of motion (4) in the form of a power series in a small parameter e ($e \ll 1$) in the neighborhood of the stationary solution (5) with the help of computer algebra system.

3. Periodic oscillations

It is possible to check that a general solution of nonlinear system (4) cannot be found in analytic form. It is convenient for application of the perturbation techniques [6] and symbolic algorithms proposed in paper [8]. However, we can seek for an approximate solution in the form of power series in the small parameter e:

$$\alpha_i(\vartheta) = e{\alpha_i}^{(1)}(\vartheta) + e^2{\alpha_i}^{(2)}(\vartheta) + ..., \tag{6}$$

Computation of unknown functions $\alpha_i(\vartheta)$ in (7) is done in accordance with the techniques proposed in [6] and [8, 9]. requires quite tedious symbolic computations. In this paper symbolic computations are performed using Wolfram Mathematica functions [13]:

Expand, TrigExpand, Series, Normal, Replace, DSolve, NDSolve.

Substituting (6) into (4) and collecting coefficients of equal powers of e, we obtain the set of systems of linear differential equations which can be solved in succession. For example, using in (7) only the first linear elements we obtain the corresponding periodic solutions in the form

$$\alpha_1^{(1)}(\vartheta) = \bar{a}_1\mathrm{sin}(\vartheta) + \bar{b}_1\mathrm{cos}(\vartheta), \quad \alpha_2^{(1)}(\vartheta) = \bar{a}_2\mathrm{sin}(\vartheta) + \bar{b}_2\mathrm{cos}(\vartheta), \tag{7}$$

where the coefficients $\bar{a}_1, \bar{b}_1, \bar{a}_2, \bar{b}_2$ can be defined from the linear algebraic system. The amplitudes of the oscillations of the first and the second bodies have the expressions

$$R_1^2 = (\bar{a}_1^2 + \bar{b}_1^2)e^2 = 4\frac{e^2b^2}{d^2}, \quad R_2^2 = (\bar{a}_2^2 + \bar{b}_2^2)e^2 = 4\frac{e^2\bar{b}^2}{d^2}, \tag{8}$$

where

$$
\begin{aligned}
b &= (B_1 + Ma_1(a_1 - a2))(3(A_2 - C_2) - B_2) - 4Ma_2(a_1B_2 + a_2B_1),\\
\bar{b} &= (B_2 + Ma_2(a_1 - a2))(3(A_1 - C_1) - B_1) - 4Ma_1(a_1B_2 + a_2B_1),\\
d &= (3(A_1 - C_1) - B_1)(3(A_2 - C_2) - B_2) - 4Ma_1^2(3(A_2 - C_2) - B_2)\\
&- \quad 4Ma_2^2(3(A_1 - C_1) - B_1).
\end{aligned} \tag{9}
$$

In the present work, we have considered the first approximation of the planar oscillations of a system of two bodies connected by a spherical hinge that moves along an elliptic orbit. We have found the expressions of the periodic motion of the system in the linear approximation. All the relevant computations in this work are performed with the computer algebra system Wolfram Mathematica. At the next step we will do the quadratic and cubic approximation of the periodic solutions which have very cumbersome expressions.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: I thank the conference organizers for the opportunity to implement this abstract of my report.

References

[1] J. P. C. Clark, Response of a two-body gravity gradient system in a slightly eccentric orbit, J. Spacecraft and Rockets 7(3) (1970) 294–298. doi:10.2514/3.29923.

[2] S. A. Gutnik, V. A. Sarychev, Symbolic investigation of the dynamics of a system of two connected bodies moving along a circular orbit, in: M. England, W. Koepf, T. M. Sadykov, W. M. Seiler, E. V. Vorozhtsov (Eds.), Computer Algebra in Scientific Computing, Springer International Publishing, Cham, 2019, pp. 164–178. doi:10.1007/978-3-030-26831-2_12.

[3] S. A. Gutnik, V. A. Sarychev, Symbolic computations of the equilibrium orientations of a system of two connected bodies moving on a circular orbit around the Earth, Math. Comput. Sci. 15(3) (2021) 407–417. doi:10.1007/s11786-021-00511-6.

[4] S. A. Gutnik, V. A. Sarychev, Computer algebra methods for searching the stationary motions of the connected bodies system moving in gravitational field, Math.Comput. Sci. 16(2) (2022) 1–15. doi:10.1007/s11786-022-00535-6.

[5] S. A. Gutnik, V. A. Sarychev, Symbolic methods for studying the equilibrium orientations of a system of two connected bodies in a circular orbit, Programming and Computer Software 48(2) (2022) 73–79. doi:10.1134/S0361768822020050.

[6] A. H. Nayfeh, Introduction to Perturbation Techniques, John Wiley and Sons, 1993.

[7] V. Penkov, Compensation for eccentric oscillations of a satellite with a gravitational stabilization system, Cosmic Research 15(3) (1977) 376–383.

[8] A. N. Prokopenya, Symbolic computation in studying stability of solutions of linear differential equations with periodic coefficients, Programming and Computer Software 33(2) (2007) 60–66. doi:10.1134/S0361768807020028.

[9] A. N. Prokopenya, Construction of a periodic solution to the equations of motion of generalized Atwood's machine using computer algebra, Programming and Computer Software 46(2) (2020) 120–125. doi:10.1134/S0361768820020085.

[10] B. V. Rauschenbakh, M. Y. Ovchinnikov, S. McKenna-Lawlor, Essential Spaceflight Dynamics and Magnetospherics, Kluwer Academic Publishers, 2003. doi:10.1007/0-306-48027-1.

[11] V. A. Sarychev, Relative equilibrium orientations of two bodies connected by a spherical hinge on a circular orbit, Cosmic Research 5 (1967) 360–364.

[12] V. A. Sarychev, Problems of orientation of satellites, Itogi Nauki i Tekhniki. Ser. Space Research, Vol. 11. VINITI, Moscow, 1978. (in Russian).

[13] Wolfram mathematica website, 2025. URL: http://www.wolfram.com/mathematica.

[14] S. Wolfram, An elementary introduction to the Wolfram Language. 2nd ed., Champaign, IL, USA, Wolfram media, 2017.

[15] V. A. Zlatoustov, A. P. Markeev, Stability of planar oscillations of a satellite in an elliptic orbit, Celest. Mech. 7(1) (1973) 31–45. doi:10.1007/BF01243507.

Searching for an Imperfect Palindrome

Georgii A. Khaziev[1,*], Alexandr V. Seliverstov[1] and Oleg A. Zverkov[1]

[1]*Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Bolshoy Karetny per. 19, build.1, Moscow, 127051, Russian Federation*

Abstract

We consider algorithm for optimizing an imperfect palindrome in the input sequence. Our algorithm runs in quadratic time, i. e., it is faster than exhaustive search of admissible palindromes. Also we propose new approach called "trimming" to find long substrings of a sequence, that are closer to palindrome, than sequence itself.

Keywords

palindrome, edit distance, Python, Numba, bioinformatics, computational complexity

1. Introduction

There are many works devoted to the search for perfect as well as imperfect and degenerate palindromes [1]. We consider imperfect palindromes, i. e., sequences with gaps and mismatches in some positions. On the other hand, the edit distance between two sequences can be calculated using dynamic programming[3], but the algorithm does not involve palindromic structures. The connection between the longest common subsequence problem and the Hecke monoid has been rediscovered many times in different forms [4].

Let us consider nucleotide sequences. The nucleotides $\{\mathsf{A}, \mathsf{C}, \mathsf{G}, \mathsf{T}\}$ form two complementary pairs: $\mathsf{c}(\mathsf{A}) = \mathsf{T}$, $\mathsf{c}(\mathsf{T}) = \mathsf{A}$, $\mathsf{c}(\mathsf{C}) = \mathsf{G}$, and $\mathsf{c}(\mathsf{G}) = \mathsf{C}$. Next, $\mathsf{c}()$ denotes the reverse complement, i. e., $\mathsf{c}(xy) = \mathsf{c}(y)\mathsf{c}(x)$. For example, $\mathsf{c}(\mathsf{AACG}) = \mathsf{CGTT}$. In DNA, a perfect palindrome is an inverted sequence repeat, i. e., reverse complement of itself. Let us omit the concatenation symbol. So, all perfect palindromes are of the type $x\mathsf{c}(x)$, where x denotes a sequence. In particular, a sequence of odd length cannot be any perfect palindrome. There are many examples of nucleotide sequences with perfect as well as imperfect palindromes [5]. Regulatory palindromes are typically imperfect [2].

For two sequences x and y, let $\mathrm{dist}(x, y)$ denote the edit distance. Of course,

$$\begin{aligned} \mathrm{dist}(x, y) &= \mathrm{dist}(\mathsf{c}(x), \mathsf{c}(y)) \\ \mathrm{dist}(wx, wy) &= \mathrm{dist}(x, y) \\ \mathrm{dist}(xz, yz) &= \mathrm{dist}(x, y) \end{aligned}$$

2. Results

Let us denote by $|x|$ the length of x. If sequences x and y coincide, then the edit distance in Theorem 1 vanishes for a perfect palindrome. Let us denote by $\mathrm{imp}(x)$ the ratio of the minimum edit distance to the length of the sequence:

$$\mathrm{imp}(x) = \frac{\min\{\mathrm{dist}(x, w\mathsf{c}(w)) | x = wz\}}{|x|}.$$

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*

10.22363/12585-2025-6-013

EDN: CCODUV

*Corresponding author.

georgy.haziev@yandex.ru (G. A. Khaziev)

0009-0003-6466-8595 (G. A. Khaziev); 0000-0003-4746-6396 (A. V. Seliverstov); 0000-0002-8546-364X (O. A. Zverkov)

The ratio shows how imperfect the palindrome is. The correctness of the definition is based on the next theorem.

Theorem 1. *There is a quadratic-time algorithm that computes the* $\mathrm{imp}(x)$ *function as well as optimal partition of the input sequence* x *as a concatenation* $x = wz$ *that minimizes the edit distance between* x *and the palindrome* $w\mathrm{c}(w)$.

Theorem 2. $\mathrm{imp}(x) = \mathrm{imp}(\mathrm{c}(x))$.

Proof. For all $x = wz$, because the prefixes match, $\mathrm{dist}(x, w\mathrm{c}(w)) = \mathrm{dist}(z, \mathrm{c}(w))$. And because the suffixes match, $\mathrm{dist}(x, \mathrm{c}(z)z) = \mathrm{dist}(w, \mathrm{c}(z))$. Of course, since $\mathrm{dist}(a, b)$ is symmetric, it is true that $\mathrm{dist}(z, \mathrm{c}(w)) = \mathrm{dist}(\mathrm{c}(w), z) = \mathrm{dist}(w, \mathrm{c}(z))$. Thus, $\mathrm{dist}(x, w\mathrm{c}(w)) = \mathrm{dist}(x, \mathrm{c}(z)z)$. Next, for all z, $\mathrm{dist}(x, \mathrm{c}(z)z) = \mathrm{dist}(\mathrm{c}(x), \mathrm{c}(z)z)$ because $\mathrm{c}(z)z$ is a palindrome. So, the equality $\mathrm{dist}(x, w\mathrm{c}(w)) = \mathrm{dist}(\mathrm{c}(x), \mathrm{c}(z)z)$ holds. □

Theorem 3. *For all even-length sequences it is true that* $\mathrm{imp}(x) \leq 1/2$. *For all odd-length sequences it is true that* $0 < \mathrm{imp}(x) \leq (1 + 1/|x|)/2$.

Substring is a contiguous sequence of characters within a string. The main idea behind the algorithm to select an imperfect palindrome is checking whether one of the optimal lengths of the prefix w of the input sequence x differs significantly from $|x|/2$. If such case occurs, then the algorithm deletes either prefix or suffix by difference between $|w|$ and $|x|/2$.

The software implementation in Python with the Numba library is available at http://lab6.iitp.ru/-/trimmers. Using Numba generally increases performance of Python code. Performance increase in comparison with base Python is achieved using JIT (just-in-time) compilation. While the base Python interpreter compiles programs into bytecode, which is executed on a virtual machine, Numba recompiles this bytecode into machine code with optimizations tailored to the CPU architecture being used. Additionally, a significant performance benefit is gained through Numba's use of type inference, allowing type-specific optimization during compilation.

All three functions `pref_trimmer`, `suff_trimmer`, and `double_trimmer` take as input nucleotide sequence x, sorted list `optimal_lengths` of optimal prefix lengths $|w|$, and floating-point number `cutoff_condition`. Note that $\min($`optimal_lengths`$)$ and $\max($`optimal_lengths`$)$ are the first and last elements of `optimal_lengths`, respectively.

The `pref_trimmer` function trims first rd symbols of x, where

$$rd = \max(\texttt{optimal\_lengths}) - \left\lfloor \frac{|x|}{2} \right\rfloor,$$

when $rd \geq \mathrm{length}(x) \cdot$ `cutoff_condition` is satisfied.

The `suff_trimmer` function trims last ld symbols of x, where

$$ld = \left\lfloor \frac{|x|}{2} \right\rfloor - \min(\texttt{optimal\_lengths}),$$

when $ld \geq |x| \cdot$ `cutoff_condition` is satisfied.

The `double_trimmer` function initially computes

$$rd = \max(\texttt{optimal\_lengths}) - \left\lfloor \frac{|x|}{2} \right\rfloor$$

and

$$ld = \left\lfloor \frac{|x|}{2} \right\rfloor - \min(\texttt{optimal\_lengths}).$$

Subsequently it checks if

$$rd \geq |x| \cdot \texttt{cutoff\_condition}$$

and

$$ld \geq |x| \cdot \texttt{cutoff\_condition}.$$

If the first inequality is satisfied, the function trims the first rd symbols from the string x. Similarly, if the second inequality is satisfied, the function trims the last ld symbols from x.

Theorem 4. *For perfect palindrome x for any* `cutoff_condition` > 0 *no trimming would be performed.*

Proof. Let x be perfect palindrome. Then it's only optimal partition is precisely at $\left\lfloor \frac{|x|}{2} \right\rfloor$. Then, $ld = \left\lfloor \frac{|x|}{2} \right\rfloor - \left\lfloor \frac{|x|}{2} \right\rfloor = 0$ and $rd = \left\lfloor \frac{|x|}{2} \right\rfloor - \left\lfloor \frac{|x|}{2} \right\rfloor = 0$. For non-zero `cutoff_condition` rd and ld will both be less than $|x| \cdot$ `cutoff_condition` so no trimming would be performed. □

Theorem 5. *For any $n \geq 3$, there exist* `cutoff_condition` > 0 *and x with length of at least n, which satisfies both prefix and suffix trimming conditions such that*

$$\texttt{pref\_trimmer}(x_{suff}) \neq \texttt{suff\_trimmer}(x_{pref}),$$

where x_{suff} and x_{pref} are results of suffix and prefix trimming of x, respectively.

Proof. Let $x = \underbrace{\text{ATA...ATA}}_{\alpha}$ where $\alpha = 2\left\lfloor \frac{n}{2} \right\rfloor + 1$, i. e. minimal odd number greater than or equal to n. For that x optimal partitions would be at $\left\lceil \frac{n}{2} \right\rceil - 1$ and $\left\lceil \frac{n}{2} \right\rceil + 1$. Minimal index of optimal partition corresponds to perfect palindrome $\underbrace{\text{AT...AT}}_{\alpha-1}$ and maximal index of optimal partition corresponds to perfect palindrome $\underbrace{\text{AT...AT}}_{\alpha+1}$. Let us choose `cutoff_condition` $= \frac{1}{10\alpha}$. Since this value would satisfy both prefix and suffix trimming condition, x_{suff} and x_{pref} could be computed and will be $\underbrace{\text{AT...AT}}_{\alpha-1}$ and $\underbrace{\text{T...ATA}}_{\alpha-1}$, respectively. Note, that both sequences are perfect palindromes and thus neither suffix nor prefix trimming would be performed for them. □

3. Conclusion

We are confident that the obtained results will be successfully applied to the prediction of imperfect palindromes in nucleotide sequences. In particular, it is crucial for predicting gene expression regulations as well as RNA structures. The implementation of algorithms in Python will enable a wide range of bioinformaticians to apply them in their work. Moreover, low computational complexity allows efficient processing of large datasets.

Author Contributions: Conceptualization, methodology, programming, writing original draft preparation, Georgii A. Khaziev; Proof of Theorem 1, editing, Alexandr V. Seliverstov; Programming, creating the web page, and editing, Oleg A. Zverkov; All authors have read and agreed to the published version of the manuscript.

Funding: The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation for IITP RAS

Data Availability Statement: Refer to http://lab6.iitp.ru/-/trimmers

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: We thank Vassily A. Lyubetsky for creating friendly and healthy work environment.

References

[1] M. Alzamel, C. Hampson, C. S. Iliopoulos, Z. Lim, S. Pissis, D. Vlachakis, S. Watts, Maximal degenerate palindromes with gaps and mismatches, Theoretical Computer Science 978 (2023) 114182. doi:https://doi.org/10.1016/j.tcs.2023.114182.

[2] R. R. Datta, J. Rister, The power of the (imperfect) palindrome: Sequence-specific roles of palindromic motifs in gene regulation, BioEssays 44 (2022) 2100191. doi:https://doi.org/10.1002/bies.202100191.

[3] S. B. Needleman, C. D. Wunsch, A general method applicable to the search for similarities in the amino acid sequence of two proteins, Journal of Molecular Biology 48 (1970) 443–453. doi:10.1016/0022-2836(70)90057-4.

[4] A. V. Tiskin, The Chvátal–Sankoff problem as a problem of symbolic dynamics, Zap. Nauchn. Sem. POMI 528 (2023) 214–237.

[5] O. A. Zverkov, A. V. Seliverstov, G. A. Shilovsky, Alignment of a hidden palindrome, Mathematical Biology and Bioinformatics 19 (2024) 427–438. doi:10.17537/2024.19.427.

Quantum Mechanics Through the Lens of Finite Groups: Computer Algebra Insights

Vladimir V. Kornyak

Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation

Abstract

The use of non-constructive infinities in physical theories can lead to contradictions and non-physical artifacts. Quantum behavior can be fully described using only finite subgroups of the general unitary group — specifically, the Weyl–Heisenberg group and its extension, the Clifford group. By formulating quantum theory in terms of these groups, we completely eliminate the need for the continuous unitary group, which leads to important empirical consequences. Crucially, this approach provides a natural explanation for the observed absence of quantum entanglement and interference between distinct types of elementary particles. Moving away from continuum-based mathematics also requires redefining the concept of quantum states: the continuous projective Hilbert space should be replaced by some combinatorial structure. Using computer algebra calculations, we study a potential framework for building constructive quantum states, governed by a fixed set of physically motivated criteria.

Keywords

cyclic group, Weyl–Heisenberg group, Clifford group, quantum evolution, constructive quantum states

1. Introduction

In standard quantum mechanics, the evolution of a closed system is described by a continuous one-parameter unitary group generated by a Hamiltonian: $U_t = \mathrm{e}^{-\mathbf{i}\frac{H}{\hbar}t} = \left(\mathrm{e}^{-\mathbf{i}\frac{H}{\hbar}}\right)^t = \mathrm{E}^t$. Any continuous one-parameter group is isomorphic to the unitary group $\mathsf{U}(1)$, usually realized as the unit circle in the complex plane. Without loss of describing physical reality, we can assume that time t is an integer parameter, and the operator E is an element of a representation of a finite cyclic group $\mathbb{Z}_N$, where N is a large natural number. In [1], assuming that time t is given in Planck units, estimates are provided $N \sim \begin{cases} \mathrm{Exp}(\mathrm{Exp}(20)) & \text{for } 1 \text{ cm}^3 \text{ of matter,} \\ \mathrm{Exp}(\mathrm{Exp}(123)) & \text{for the entire Universe.} \end{cases}$

In applications, the one-parameter group $\mathsf{U}(1)$ can be used as a continuum approximation of the discrete group $\mathbb{Z}_N$ as $N \to \infty$. However, this approximation fails to capture certain empirically observed fundamental quantum phenomena that depend on the number-theoretic properties of N. Key example: the primary decomposition of $\mathbb{Z}_N$ (via the Chinese remainder theorem) implies the decomposition of a N-dimensional quantum system into completely decoupled subsystems, i.e. there is nether quantum entanglement nor energy interaction between them.

Hermann Weyl was the first to discover [6] that finite groups $\mathbb{Z}_N$, $\mathbb{Z}_N \times \mathbb{Z}_N$, and a central extension of the latter are needed to describe quantum behavior in a finite-dimensional Hilbert space $\mathcal{H}_N$. Although unitary evolution can be fully described by the cyclic group, the product becomes necessary to incorporate the concept of observation into the theoretical framework. The factors of the product group are associated with *mutually unbiased bases* – a concept introduced by Julian Schwinger [4] that provides a mathematical refinement of Bohr's complementarity principle. Subsequently, the ideas of Weyl and Schwinger have been actively developed in

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*

10.22363/12585-2025-6-014

EDN: CDSESO

vkornyak@gmail.com (V. V. Kornyak)

0000-0002-5712-2960 (V. V. Kornyak)

various fields, including the foundations of quantum theory, quantum information theory [2], and signal processing theory [5].

2. Weyl–Schwinger formalism

Noting that the canonical Heisenberg commutation relation $[\hat{x}, \hat{p}] = \mathbf{i}\hbar\,\mathbb{1}$, – and hence standard quantum mechanics as a whole – can only be realized in an infinite-dimensional Hilbert space, Weyl proved that in physically more meaningful finite-dimensional case, the commutation relation in the Hilbert space $\mathcal{H}_N$ must be of the form $ZX = \omega XZ$, where X is an operator of cyclic permutation of N basis vectors in $\mathcal{H}_N$, Z is the Pontryagin dual of X, ω is a Nth primitive root of unity. Later, Schwinger proved that the eigenbases of operators X and Z in $\mathcal{H}_N$ constitute a mutually unbiased pair. Following 't Hooft's terminology, we will refer to the basis vectors cyclically permuted by the operator X as *ontic* vectors. The pair of operators X and Z generates a projective representation of the group $\mathbb{Z}_N \times \mathbb{Z}_N$ in $\mathcal{H}_N$ — this representation underlies the description of quantum behavior. In short, the main elements of the formalism are as follows.

Let $\mathbb{K}_n$ denote the group of nth roots of unity. The element $\tau = -\,\mathrm{e}^{\pi\mathbf{i}/N}$ generates $\mathbb{K}_N$ if $N = 2k+1$ and $\mathbb{K}_{2N}$ if $N = 2k$. The **Weyl–Heisenberg group** is defined as $\mathrm{WH}(N) = \langle \tau, X, Z \rangle$. The order of $\mathrm{WH}(N)$ is N^3 or $2N^3$ depending on the parity of N. Quantum evolutions are generated by the *displacement operators* $\mathsf{D}_{\mathbf{p}} = \tau^{p_1 p_2} X^{p_1} Z^{p_2}$, $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \in \mathbb{Z}^2$, which form the **projective Weyl–Heisenberg group** $\mathrm{PWH}(N) \cong \mathbb{Z}_N \times \mathbb{Z}_N$ of order N^2. The composition $\mathsf{D}_{\mathbf{p}}\mathsf{D}_{\mathbf{q}} = \tau^{\langle \mathbf{p},\mathbf{q} \rangle}\mathsf{D}_{\mathbf{p}+\mathbf{q}}$ contains the symplectic form $\langle \mathbf{p}, \mathbf{q} \rangle = p_2 q_1 - p_1 q_2$. The symmetry group of this form, the symplectic group $\mathrm{Sp}(2, \mathbb{Z}_N)$, is the outer automorphism group of $\mathrm{WH}(N)$.

Combining inner and outer automorphisms, we arrive at a semidirect product called the **Clifford group** $\mathrm{CL}(N) = \mathrm{Aut}(\mathrm{WH}(N)) \cong \mathrm{WH}(N) \rtimes \mathrm{Sp}(2, \mathbb{Z}_N)$.

Traditionally, the Clifford group is defined as the normalizer of the Weyl–Heisenberg group in the unitary group $\mathsf{U}(N)$. The need for $\mathsf{U}(N)$, which remains a relic of continuous theory, follows neither from the description of quantum evolution by finite cyclic groups nor from Weyl's arguments. We will consider the Clifford group exclusively as the symmetry group of the Weyl–Heisenberg group without resorting to a reference to the continuous group $\mathsf{U}(N)$.

In terms of generators, the Clifford group can be presented as $\mathrm{CL}(N) = \langle X, F, S \rangle$, where F is the Fourier transform matrix and $S = \mathrm{diag}\big(\tau^{i(i+N)}, i = 0, \ldots, N-1\big)$.

The projective Clifford group — the quotient group of $\mathrm{CL}(N)$ by its center $\mathrm{Z}(\mathrm{CL}(N))$ — is generated by the same elements, but matrices that differ only by a phase factor are considered equivalent: $\mathrm{PCL}(N) = \langle X, F, S \rangle / \mathrm{Z}(\mathrm{CL}(N))$.

3. Decomposition of a *N*-dimensional quantum system in subsystems

Let $N = \prod_i n_i$ be a factorization of N into pairwise coprime integers $\{n_i\}$. For concreteness, we assume that all factors take the form $n_i = p_i^{\ell_i}$, where $\left\{p_i^{\ell_i}\right\}$ are prime powers with distinct primes. The cyclic group $\mathbb{Z}_N$ can be decomposed into a direct product of *primary cyclic groups* $\mathbb{Z}_N \cong \prod_i \mathbb{Z}_{n_i}$. This isomorphism provides a natural way to decompose the N-dimensional quantum system into subsystems. Specifically, the *global* Hilbert space admits a decomposition as a tensor product of *local* Hilbert spaces $\mathcal{H}_N = \bigotimes_i \mathcal{H}_{n_i}$. Using the shorthand notation GH for element-wise group action and the tensor product identity $AX \otimes BY = (A \otimes B)\,(X \otimes Y)$, the equivalence class of this decomposition — accounting for the freedom in choosing Hilbert space coordinates — can be formally expressed as:

$$G(N)\,\mathcal{H}_N \simeq \prod_i G(n_i) \bigotimes_i \mathcal{H}_{n_i}, \tag{1}$$

where $G(d)$ denotes the symmetry group acting on a d-dimensional Hilbert space. Since the product of local groups is a subgroup of the global group, $\prod_i G(n_i) \leq G(N)$, description (1) reduces to $G(N)\,\mathcal{H}_N \simeq \bigotimes_i \mathcal{H}_{n_i}$, implying that all decompositions related by the action of $G(N)$ are equivalent. The equivalence class is uniquely fixed by the coprime factorization $N = \prod_i n_i$.

The assumption $G(N) = \mathsf{U}(N)$ can lead to artifacts, since the continuous group $\mathsf{U}(N)$ freely "mixes" states between different components of the tensor product, which would lead to non-observable in nature entanglement between fundamental particles of different types.

The assumption $G(N) = \mathsf{CL}(N)$ does not cause such problems, since in the global Clifford group does not contain transformations that mix states between local Hilbert spaces of coprime dimensions. Mathematically, this is expressed by the fact, proved using the Chinese remainder theorem, that the global Clifford group decomposes into a direct product of local ones: $\mathsf{CL}(N) = \prod_i \mathsf{CL}(n_i)$. The absence of quantum entanglement between subsystems means that during any evolutions of the global system, only quantum states that are tensor products of the states of the subsystems (or their classical combinations, called separable states) are possible.

By applying the Chinese remainder theorem to the eigenvalues of the Hamiltonian for the cyclic evolution of the global system [3], we derive the additive decomposition of energy levels between the global quantum system and its local subsystems:

$$E_{k/N} = \sum_i E_{k_i/n_i}\,, \tag{2}$$

where $E_\nu = h\nu$ represents Planck's energy-frequency relation. Equality (2) demonstrates that:

1. The energy of the global system equals the sum of the energies of its subsystems, and
2. No interaction energies exist between components.

This result implies that subsystems of coprime dimensions are completely decoupled, showing no quantum entanglement or energy exchange, and thus permit fully independent investigation.

In prime-dimensional quantum systems, quantum interference can occur. However, they cannot exhibit quantum entanglement due to absence of proper subsystems. Entanglement is possible only in a system of non-trivial prime power dimension.

Consequently, the most physically significant systems for study are:

1. Systems of prime dimensions ($N = p$), only quantum interference is possible, and
2. Systems of prime power dimensions ($N = p^\ell,\ \ell > 1$) where quantum entanglement emerges.

4. Constructive quantum states

In continuous quantum mechanics, the set of pure states in a N-dimensional Hilbert space is the complex projective space $\mathbb{P}(\mathcal{H}_N) = \mathbb{CP}^{N-1}$, which is a homogeneous space of the unitary group $\mathsf{U}(N)$, i.e., $\mathbb{CP}^{N-1}$ is the orbit of an arbitrary unit vector, e.g., $|0\rangle$, under the action of the unitary group: $\mathbb{CP}^{N-1} \cong \mathsf{Orb}_{\mathsf{U}(N)}(|0\rangle) = \mathsf{U}(N)\,|0\rangle$.

In our approach, the group of symmetries of quantum systems is the finite Clifford group, which acts on the set of quantum states non-transitively, splitting it into disjoint orbits. Replacing $\mathsf{U}(N)$ with $\mathsf{CL}(N)$ as the group of symmetries, we assume that the constructive set of pure quantum states, which we denote as $\mathsf{CQS}(N)$, consists of elements of the form $|a\rangle = \sum_{i=0}^{N-1} \varphi_i \alpha_i\,|i\rangle$, where $\alpha_i \in \mathbb{R}$, $\sum_{i=0}^{N-1} \alpha_i^2 = 1$, $\varphi_i \in \mathsf{Z}(\mathsf{CL}(N))$, i.e., the phase factors belong to the center of the Clifford group $\mathsf{CL}(N)$. The set $\mathsf{CQS}(N)$ must 1) be $\mathsf{CL}(N)$-invariant; 2) contain ontic vectors; 3) consist only of elements with rational Born probabilities of transitions between themselves; 4) contain all superpositions of vectors with phase factors from $\mathsf{Z}(\mathsf{CL}(N))$ that satisfy the rationality requirement.

To study the properties of the quantum states corresponding to these requirements, we implemented a procedure to sequentially construct the orbits that make up the set $\mathsf{CQS}(N)$. At

first, the orbit $\mathcal{O}_1 = \mathrm{Orb}_{\mathrm{CL}(N)}(|0\rangle)$ is constructed. In dimensions $N = p^\ell$ this orbit consists of $N(N+1)$ vectors, forming a complete set of $N+1$ mutually unbiased bases and, in particular, contains the ontic basis. Next, we construct other orbits using superpositions of already existing elements. The results of computer experiments possibly indicate the formation of a dense subset in $\mathbb{CP}^{N-1}$.

4.1. Computations in dimensions 2 and 3

Table 1
Generators, centers, and sizes of the Clifford groups in dimensions 2 and 3, $\omega = -\frac{1}{2} + \mathbf{i}\frac{\sqrt{3}}{2}$.

N	X	F	S	$\mathrm{Z}(\mathrm{CL}(N))$	$\lvert\mathrm{CL}(N)\rvert$
2	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$	$\begin{pmatrix}1&0\\0&\mathbf{i}\end{pmatrix}$	$\mathbb{K}_8$	192
3	$\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1&1&1\\1&\omega&\omega^2\\1&\omega^2&\omega\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&\omega^2&0\\0&0&\omega^2\end{pmatrix}$	$\mathbb{K}_{12}$	2592

For quantum states $|a\rangle, |b\rangle \in \mathrm{CQS}(N)$, we define the distance function $\mathrm{Dist}(a,b) = 1 - \mathbf{P}(a,b) \equiv \sin^2 \mathrm{D_{FS}}(a,b)$, where $\mathbf{P}(a,b) = |\langle a \mid b\rangle|^2$ is the Born probability of transition between states, and $\mathrm{D_{FS}}(a,b)$ is the length of the geodesic line in the natural for $\mathbb{CP}^{N-1}$ Fubini–Study metric.

To estimate the density of states for a subset $S \subset \mathrm{CQS}(N)$ in the complex projective space, we define the function $\Delta(S) = \max_{a\in S} \min_{b\in S\setminus\{a\}} \mathrm{Dist}(a,b)$ that computes, for each point in S, the distance to its nearest neighbor and then takes the maximum of these minimal distances.

4.1.1. *N* = 2

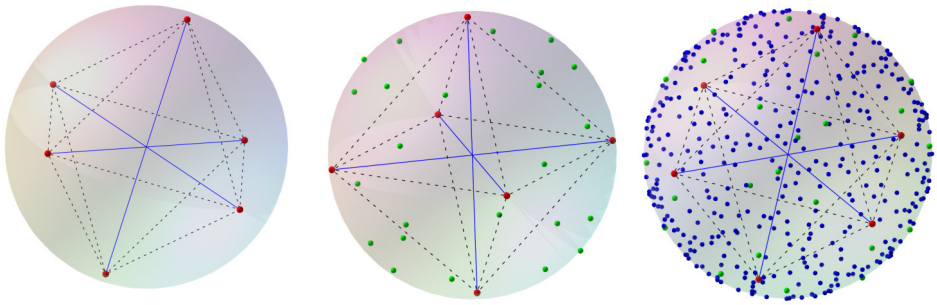
The results permit geometric visualization: pure states reside in the complex projective line $\mathbb{CP}^1$, which is isomorphic to the Riemann sphere (Bloch sphere). The projective Clifford group $\mathrm{PCL}(2) = \mathrm{CL}(2)/\mathbb{K}_8$ has order 24. We calculated — initial steps are shown in Fig. 1 — a subset of states $S \subset \mathrm{CQS}(2)$, which is the union of 986 orbits with a total number of elements 23646. The maximum distances between neighboring states for the initial orbit and for the entire computed set of states are, respectively, $\Delta(\mathcal{O}_1) = 1/2$ and $\Delta(S) = 1/1515 \approx 10^{-3}$.

The initial orbit $\mathcal{O}_1 = \mathrm{Orb}_{\mathrm{CL}(2)}(|0\rangle)$ consists of six vectors, orthogonal pairs of which form a complete set of three mutually unbiased bases:

$$\mathcal{O}_1 = \left\{ |0\rangle, |1\rangle;\ \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \frac{|0\rangle - |1\rangle}{\sqrt{2}};\ \frac{|0\rangle + \mathbf{i}\,|1\rangle}{\sqrt{2}}, \frac{|0\rangle - \mathbf{i}\,|1\rangle}{\sqrt{2}} \right\}. \tag{3}$$

4.1.2. *N* = 3

The order of the projective group $\mathrm{PCL}(3) = \mathrm{CL}(3)/\mathbb{K}_{12}$ is 216. The initial orbit forms a complete set of four mutually unbiased bases: $\mathcal{O}_1 = \left\{ |0\rangle, |1\rangle, |2\rangle; \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega\\\omega^2\end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega^2\\\omega\end{pmatrix}; \frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega^2\\\omega^2\end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\\omega\end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega\\1\end{pmatrix}; \frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega\\\omega\end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\\omega^2\end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega^2\\1\end{pmatrix} \right\}$. The set of states $S \subset \mathrm{CQS}(3)$ that we calculated consists of 169 orbits that contain a total of 27237 vectors. The maximum distances between neighboring states for the initial orbit and for the entire calculated set of states are, respectively, $\Delta(\mathcal{O}_1) = 2/3$ and $\Delta(S) = 1/99 \approx 10^{-2}$.



(a) (b) (c)

Figure 1: Initial steps in generating constructive quantum states in dimension 2:
(a) the vectors of orbit (3) form the vertices of an octahedron, whose spatial diagonals represent the three mutually unbiased bases;
(b) pairwise interferences of the vectors in (a) with rational transition probabilities add one orbit of size 24;
(c) pairwise interferences of the vectors in (b) add 16 orbits of size 24 to the set of constructive quantum states.

References

[1] T. Banks, Finite deformations of quantum mechanics, arXiv:2001.07662 [hep-th] (2020).
[2] I. Bengtsson, K. Zyczkowski, Geometry of Quantum States: An Introduction to Quantum Entanglement, Cambridge University Press, 2006.
[3] V. V. Kornyak, Constructive quantum mechanics and calculations with finite groups, Programming and Computer Software 51 (2025) 6–16.
[4] J. Schwinger, Unitary operator bases, Proc. Nat. Acad. Sci. U. S. A. 46 (1960) 570–579.
[5] S. F. D. Waldron, An Introduction to Finite Tight Frames, Birkhaüser, NY, 2018.
[6] H. Weyl, The Theory of Groups and Quantum Mechanics, Dover Publications, NY, 1931.

Moshinsky Atom as a Test for FEM on Hypercubes

Oleg O. Kovalev[1,2], Balt Batgerel[3], Alexander A. Gusev[1,2,4,*], Luong Le Hai[5], Ochbadrakh Chuluunbaatar[1,3,4] and Sergue I. Vinitsky[1,2,6]

[1]*Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*
[2]*Dubna State University, 19 Universitetskaya St, Dubna, 141980, Russian Federation*
[3] *Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Ulaanbaatar, 13330, Mongolia*
[4] *School of Applied Sciences, Mongolian University of Science and Technology, Ulaanbaatar, 14191, Mongolia*
[5]*Ho Chi Minh city University of Education, Ho Chi Minh City, 72711, Viet Nam*
[6]*RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

Abstract
Third- and fourth-order FEM schemes with multivariate Hermite interpolation polynomials of a d-dimensional hypercube for solving boundary value problems (BVPs) on hyperparallelepipedal meshes are elaborated. An exactly solvable model of a system of several identical particles with pair oscillator interaction known as the Moshinsky atom is used as a test example. To describe the energy spectra of symmetric and antisymmetric bound states, the 2-,3-,4-, and 5- dimensional BVPs with Dirichlet and Neumann boundary conditions on a nonrectangular domain are formulated. To generate new FEM schemes with mixed partial derivatives, additional affine coordinate transformations are applied. Benchmark calculations of the BVPs confirm the order of declared FEM schemes.

Keywords
multivariate Hermite interpolation polynomials, hyperparallelepipedal mesh, finite element method, Moshinsky atom

1. Introduction

At present, in connection with the study of energy spectra and electromagnetic transitions of collective quadrupole-octupole models of heavy and superheavy atomic nuclei in a self-consistent relativistic mean field parameterization, it is relevant to develop computational schemes and programs for solving multidimensional (up to six-dimensional) boundary value problems (BVPs) for elliptic equations with mixed partial derivatives [1]. Their specific feature is that the calculation of the tabulated coefficients of differential equations even with restricted accuracy takes quite a long time. So, the problem of constructing nonstructured grids does not arise, and it is sufficient to restrict ourselves to constructing FEM schemes on hyperparallelepipedal grids. To save computer resources, the BVP is solved not on a parallelepiped, but on a complex domain composed of parallelepipeds [2]. It was tested by solving $d = 2$ BVPs for a system of elliptic differential equations by means of the 2DFEM program implemented in Wolfram Mathematica [1].

In the present communication, we continue elaborating FEM schemes and programs for solving the multidimensional BVPs. As a test example, we choose a BVP with known degenerate spectrum, the so-called Moshinsky atom describing a system of $A \geq 3$ one-dimensional identical particles with pair oscillatory interaction. In the center-of-mass system in symmetrized coordinates [3], it is reduced to a BVP of dimension $d = A - 1$ with Dirichlet or Neumann boundary conditions on a non-rectangular domain. After an affine transformation, the problem is

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-015
EDN: CFJIKG
*Corresponding author.
okovalev@jinr.ru (O. O. Kovalev); batgerel@mas.ac.mn (B. Batgerel); gooseff@jinr.ru (A. A. Gusev); haill@hcmue.edu.vn (L. L. Hai); chuka@jinr.ru (O. Chuluunbaatar); vinitsky@theor.jinr.ru (S. I. Vinitsky)

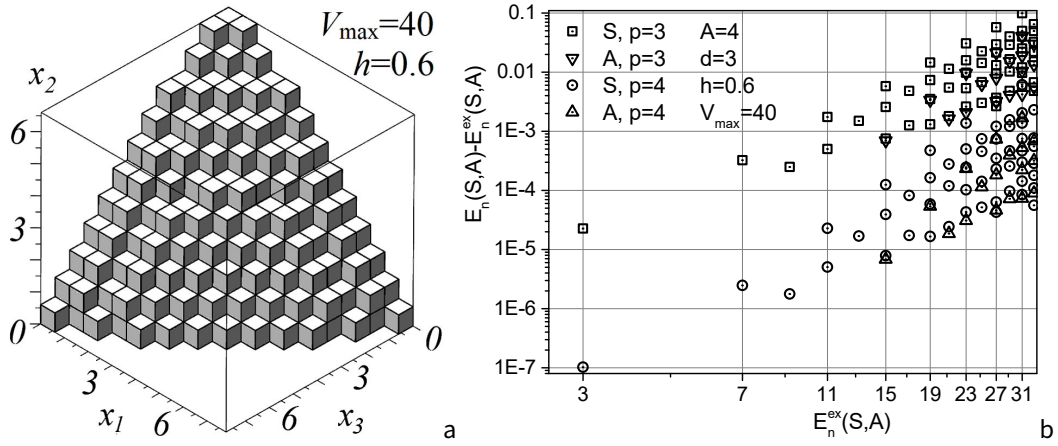


Figure 1: The FEM mesh (a) and the differences $E_n - E_n^{\text{ex}}$ plotted vs their exact degenerate values E_n^{ex} (b) for the FEM schemes of order p approximating the three-dimensional BVP of the relative motion of four identical particles

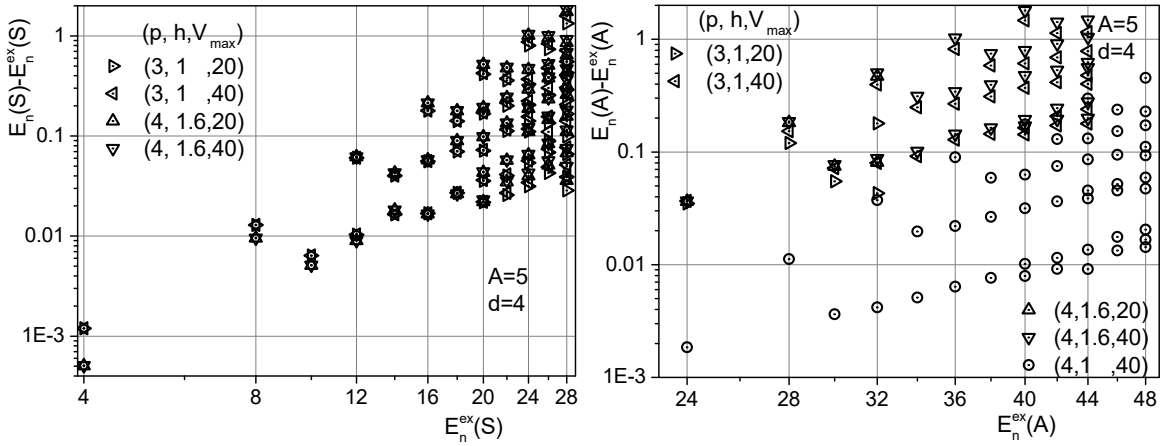


Figure 2: The differences $E_n - E_n^{\text{ex}}$ plotted vs their exact degenerate values E_n^{ex} for the FEM schemes of order p approximating the four-dimensional BVP of the relative motion of five identical particles

reduced to a BVP on a rectangular domain with mixed derivatives. Some large-scale calculations were performed on the Multifunctional Information and Computing Complex MLIT JINR.

2. Formulation of BVP and FEM scheme

Consider a self-adjoint BVP for the d-dimensional elliptic differential equation [1]

$$(T+V(x)-E_n)\,\Phi_n(x)=0, \quad T=-\frac{1}{g_0(x)}\sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}g_{ij}(x)\frac{\partial}{\partial x_j}, \quad g_0(x)>0, \quad g_{ij}(x)=g_{ji}(x) \qquad (1)$$

in a bounded domain $\Omega(x)$, $x=(x_1,...,x_d)\in\Omega(x)\subset\mathcal{R}^d$ with mixed Dirichlet and Neumann boundary conditions, where $V(x)$ is the potential energy, E_n is an eigenenergy. For the principal part coefficients of Eq. (1), the condition of uniform ellipticity holds in the bounded domain $\Omega(x)$ of the Euclidean space $\mathcal{R}^d$, i.e., the constants $\mu>0$, $\nu>0$ exist such that $\mu\xi^2 \le \sum_{i,j=1}^{d}\xi_i\xi_j \le \nu\xi^2$, $\xi^2 \le \sum_{i=1}^{d}\xi_i^2$, $\forall\xi_i\in\mathcal{R}$. The left-hand side of this inequality expresses the requirement of ellipticity, while the right-hand side expresses the boundedness of the coefficients $g_{ij}(x)$.

To solve the BVP, third- and fourth-order FEM schemes with rectangular finite elements were used. Local functions are products of one-dimensional third- (p=3) or fourth-order (p=4)

Table 1
The dimension of the algebraic problem $D_p^{S,A}$ and the number of nonzero elements $N_p^{S,A}$ (in 10^6) for FEM schemes of order p approximating a three-dimensional BVP of the relative motion of four identical particles. Calculated on a mesh of hyperparallelepipeds with equal step h in all variables forming cells Δ_j, in which the potential $V(x)$ satisfies the condition $V(x \in \Delta_j) < V_{\max}$, $h = 0.6$

$V_{\max}$	D_3^S	N_3^S	D_4^S	N_4^S	D_3^A	N_3^A	D_4^A	N_4^A
40	6888	0.56	20153	2.8	5539	0.41	17293	2.2
20	3136	0.23	8825	1.1	2373	0.16	7234	0.84

Table 2
The dimension of the algebraic problem $D_p^{S,A}$ and the number of nonzero elements $N_p^{S,A}$ (in 10^6) for the FEM schemes of order p approximating the four-dimensional BVP of the relative motion of five identical particles. Here $h = 1$ for $p = 3$ and in two last columns, and $h = 1.6$ for $p = 4$

$V_{\max}$	D_3^S	N_3^S	D_4^S	N_4^S	D_3^A	N_3^A	D_4^A	N_4^A	D_4^A	N_4^A
40	16272	5.3	17683	10.7	9901	2.5	11107	4.8	41593	23.4
20	7264	2.0	8065	4.2	3969	0.75	4624	1.5	15934	7.2

Hermite interpolation polynomials $\varphi_i^3(x)$ or $\varphi_i^4(x)$, which in the interval $x \in [0,1]$ have the form

$$\varphi_i^3(x{=}0){=}\delta_{i1}, \quad \left.\frac{d\varphi_i^3(x)}{dx}\right|_{x=0}{=}\delta_{i2}, \quad \varphi_i^3(x{=}1){=}\delta_{i3}, \quad \left.\frac{d\varphi_i^3(x)}{dx}\right|_{x=1}{=}\delta_{i4},$$
$$\varphi_1^3(x){=}(x{-}1)^2(1{+}2x), \quad \varphi_2^3(x){=}(x{-}1)^2x, \quad \varphi_3^3(x){=}x^2(3{-}2x), \quad \varphi_4^3(x){=}x^2(x{-}1),$$
$$\varphi_i^4(x{=}0){=}\delta_{i1}, \quad \left.\frac{d\varphi_i^4(x)}{dx}\right|_{x=0}{=}\delta_{i2}, \quad \varphi_i^4(x{=}1/2){=}\delta_{i3}, \quad \varphi_i^4(x{=}1){=}\delta_{i4}, \quad \left.\frac{d\varphi_i^4(x)}{dx}\right|_{x=1}{=}\delta_{i5},$$
$$\varphi_1^4(x){=}(4x{+}1)(1{-}2x)(x{-}1)^2, \quad \varphi_2^4(x){=}x(1{-}2x)(x{-}1)^2, \quad \varphi_3^4(x){=}16x^2(x{-}1)^2,$$
$$\varphi_4^4(x){=}x^2(1{-}2x)(4x{-}5), \quad \varphi_5^4(x){=}x^2(2x{-}1)(x{-}1).$$

Algorithms for constructing FEM schemes are given in Ref. [2].

3. Moshinsky atom: BVP in symmetrized and affine coordinates

The Moshinsky atom equation describing the relative motion dynamics of $A \geq 3$ identical particles with pair oscillatory interaction in the center-of-mass system in symmetrized coordinates $X{=}(X_1, ..., X_d){\in}\mathcal{R}^d$, $d{=}A{-}1$ has the form [3]

$$\left(\sum_{i=1}^{A-1}\left(-\frac{\partial^2}{\partial X_i^2}+X_i^2\right)-E_n^{\mathrm{S,A}}\right)\Phi_n^{\mathrm{S,A}}(X)=0.$$

To construct states that are symmetric (S) or antisymmetric (A) with respect to permutations of pairs of particles, the problem should be solved in one of the regions bounded by planes

$$\frac{2-\sqrt{A}}{\sqrt{A}-1}X_1+\frac{1}{\sqrt{A}-1}\sum_{i=2}^{A-1}X_i=0, \quad X_{j-1}-X_j=0, \quad j=2,...,A-1$$

with Dirichlet or Neumann boundary conditions, respectively. The degenerate spectrum is

$$\begin{aligned} E_n^{\mathrm{S}} &= A-1, A+3, A+5, A+7, A+9, ..., \\ E_n^{\mathrm{A}} &= A^2-1, A^2+3, A^2+5, A^2+7, A^2+9, ... \end{aligned} \tag{2}$$

Table 3
The first 6 energy levels E_n^{S} for the 3rd order FEM scheme approximating the five-dimensional BVP of the relative motion of six identical particles, where E_4^{S}, E_5^{S} are double degenerate, E_6^{S} is fourfold ones.

	E_1^{S}	E_2^{S}	E_3^{S}	E_4^{S}	E_5^{S}	E_6^{S}		D_3^S
exact	5	9	11	13	15	17		N_3^S
$V_{\max}$=12, h=1	5.003	9.029	11.012	13.014	15.003	16.963	17.035	21312
p=3				13.109	15.062	17.054	17.273	23.1
$V_{\max}$=18, h=1	5.003	9.029	11.014	13.020	15.023	17.028	17.047	34528
p=3				13.116	15.080	17.103	17.321	41.6

In affine coordinates $x{=}(x_1,...,x_d){\in}\mathcal{R}^d$, $d{=}A{-}1$ of the center-of-mass system of A particles:

$$x_1 = \frac{2-\sqrt{A}}{\sqrt{A}-1}X_1 + \frac{1}{\sqrt{A}-1}\sum_{i=2}^{A-1} X_i, \quad x_j = X_{j-1} - X_j, \quad j = 2,...,A-1,$$

the BVP has the form (1) with

$$g_0(x) = 1, \quad g_{ij}(x) = \{2, i=j; -1, i=j\pm 1; 0 \text{ otherwise}\}, \quad i,j = 1,...,A-1$$

and potential energy $V(x) = \sum_{i=1}^{A-1}(X_i(x))^2$ and the spectrum (2). The BVP is solved in a rectangular domain $\Omega(x)$ bounded by planes $x_i \in [0,+\infty)$, $i = 1,...,A-1$.

4. Results

As an example, Tables 1 and 2 present the dimensions $D_p^{S,A}$ of the algebraic eigenvalue problem and the number of nonzero elements $N_p^{S,A}$ (in 10^6) for FEM schemes of order p, approximating 3D- and 4D- BVPs for the relative motion of four and five identical particles, calculated on reduced domain meshes with hyperparallelepiped cells Δ_j in which the potential satisfies the condition $V(x{\in}\Delta_j){<}V_{\max}$ (see, e.g., Fig. 1a for $d{=}3$). The differences between the exact and approximate eigenvalues are shown in Figs. 1b and 2. The eigenvalues obtained by solving the 5D BVP are given in Table 3. The main results of the benchmark calculations of the BVPs presented in the above tables and the visible slopes of straight lines connecting the dots in the above figures confirm the declared order of the developed FEM schemes.

Author Contributions: All authors contributed equally to the work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: This publication was supported by the Peoples' Friendship University of Russia (RUDN) Strategic Academic Leadership Program (project No. 021934-0-000). This research is funded by Ho Chi Minh City University of Education Foundation for Science and Technology (grant No. CS. 2024.19.86)

References

[1] B. Batgerel, O. Chuluunbaatar, V. Derbov, A. Gusev, L. Hai, A. Deveikis, P. Hess, E. Mardyban, M. Mardyban, I. Vinitsky, P. Wen, Symbolic-numeric solving boundary value problems: Collective models of atomic nuclei, Lecture Notes in Computer Science 14938 (2024) 63–81.

[2] G. Chuluunbaatar, A. Gusev, O. Chuluunbaatar, S. Vinitsky, Hermite interpolation polynomials on parallelepipeds and fem applications, in: S.A. Abramov, A.B. Batkhin, L.A. Sevastyanov (Ed.), Computer algebra: 5th International Conference Materials, KIAM, Moscow, 2023, pp. 49–52. URL: http://www.ccas.ru/ca/_media/ca-2023.pdf.

[3] A. Gusev, S. Vinitsky, O. Chuluunbaatar, V. Rostovtsev, L. L. Hai, V. Derbov, A. Góźdź, E. Klimov, Symbolic-numerical algorithm for generating cluster eigenfunctions: identical particles with pair oscillator interactions, Lecture Notes in Computer Science 8136 (2013) 155–168.

Existence of Liouvillian Solutions in the Problem of Motion of a Heavy Gyrostat Under the Action of Gyroscopic Forces

Alexander S. Kuleshov*, Anton D. Skripkin

Department of Mechanics and Mathematics, Lomonosov Moscow State University, Main building of MSU, Leninskie Gory, Moscow, 119234, Russian Federation

Abstract

We study the problem of motion of a gyrostat about a fixed point under the action of gravity and gyroscopic forces in the Hess integrability case. It is shown, that the solution of the problem is reduced to the integration of the second -- order linear differential equation with rational coefficients. Using the Kovacic algorithm, we obtain the conditions on the parameters of the problem under which we can find the general solution of the corresponding second order linear differential equation in explicit form.

Keywords

gyrostat with a fixed point, Hess case, linear second order differential equation, Liouvillian solutions, Kovacic algorithm

1. Problem formulation

Let us consider the problem of motion of a gyrostat with a fixed point under the action of gravity and gyroscopic forces in the Hess case of integrability. Equations of motion of the gyrostat can be written as follows [1]:

$$\begin{gathered}
A_1\dot{\omega}_1 + (A_3 - A_2)\,\omega_2\omega_3 - s_2\omega_3 = -Mgx_2\gamma_3 + \lambda x_2\omega_3\,(\gamma_1 x_1 + \gamma_2 x_2)\,, \\
A_2\dot{\omega}_2 + (A_1 - A_3)\,\omega_1\omega_3 + s_1\omega_3 = Mgx_1\gamma_3 - \lambda x_1\omega_3\,(\gamma_1 x_1 + \gamma_2 x_2)\,, \\
A_3\dot{\omega}_3 + (A_2 - A_1)\,\omega_1\omega_2 + s_2\omega_1 - s_1\omega_2 = Mg\,(x_2\gamma_1 - x_1\gamma_2) + \lambda\,(x_1\omega_2 - x_2\omega_1)\,(\gamma_1 x_1 + \gamma_2 x_2)\,; \\
\dot{\gamma}_1 = \omega_3\gamma_2 - \omega_2\gamma_3, \quad \dot{\gamma}_2 = \omega_1\gamma_3 - \omega_3\gamma_1, \quad \dot{\gamma}_3 = \omega_2\gamma_1 - \omega_1\gamma_2.
\end{gathered} \tag{1}$$

Here ω_1, ω_2, ω_3 are the components of angular velocity vector of the gyrostat, γ_1, γ_2, γ_3 are the components of the unit vector of upward vertical, A_1, A_2, A_3 are the moments of inertia of the gyrostat with respect to principal axes of inertia at the fixed point, x_1, x_2 are the components of the radius – vector from the fixed point to the center of mass of the gyrostat, s_1, s_2 are the components of the gyrostatic momentum vector, M is the mass of the gyrostat, g is the acceleration due to gravity, λ is a parameter of gyroscopic forces. The components x_1, x_2 are connected with the moments of inertia A_1, A_2, A_3 by the following condition:

$$A_2\,(A_3 - A_1)\,x_2^2 = A_1\,(A_2 - A_3)\,x_1^2, \quad A_2 > A_3 > A_1; \tag{2}$$

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-016
EDN: CFVTKQ
*Corresponding author.
kuleshov@mech.math.msu.su (A. S. Kuleshov); antohaskripkin@gmail.com (A. D. Skripkin)
0000-0003-1128-8364 (A. S. Kuleshov)

Equations (1) possess the following first integrals:

$$\frac{1}{2}\left(A_1\omega_1^2+A_2\omega_2^2+A_3\omega_3^2\right)+Mg\left(x_1\gamma_1+x_2\gamma_2\right)=E,\quad \gamma_1^2+\gamma_2^2+\gamma_3^2=1,$$

$$\left(A_1\omega_1+s_1\right)\gamma_1+\left(A_2\omega_2+s_2\right)\gamma_2+A_3\omega_3\gamma_3+\frac{\lambda}{2}\left(x_1\gamma_1+x_2\gamma_2\right)^2=k,$$

$$A_1\omega_1x_1+A_2\omega_2x_2+\frac{A_1x_1\left(s_2x_1-s_1x_2\right)}{\left(A_3-A_1\right)x_2}=0. \tag{3}$$

Let us introduce the following new variables:

$$L_1=\frac{A_1\omega_1x_1+A_2\omega_2x_2}{\sqrt{x_1^2+x_2^2}},\quad L_2=\frac{A_2\omega_2x_1-A_1\omega_1x_2}{\sqrt{x_1^2+x_2^2}},\quad L_3=A_3\omega_3,$$

$$\nu_1=\frac{\gamma_1x_1+\gamma_2x_2}{\sqrt{x_1^2+x_2^2}},\quad \nu_2=\frac{\gamma_2x_1-\gamma_1x_2}{\sqrt{x_1^2+x_2^2}},\quad \nu_3=\gamma_3,\quad k_1=\frac{s_1x_1+s_2x_2}{\sqrt{x_1^2+x_2^2}},$$

$$k_2=\frac{s_2x_1-s_1x_2}{\sqrt{x_1^2+x_2^2}},\quad \Gamma=Mg\sqrt{x_1^2+x_2^2},\quad \Lambda=\lambda\left(x_1^2+x_2^2\right),$$

$$a=\frac{A_2x_1^2+A_1x_2^2}{A_1A_2\left(x_1^2+x_2^2\right)},\quad b=\frac{\left(A_1-A_2\right)x_1x_2}{A_1A_2\left(x_1^2+x_2^2\right)},\quad c=\frac{1}{A_3}.$$

Using the variables L_1, L_2, L_3, ν_1, ν_2, ν_3, we can rewrite equations (1) as follows:

$$\begin{aligned}
\dot L_1 &= -bL_3\left(L_1-\frac{ck_2}{b}\right),\quad \dot L_2=\left(a-c\right)L_1L_3+bL_2L_3-ck_1L_3-\Lambda c\nu_1L_3+\Gamma\nu_3,\\
\dot L_3 &= -\left(a-c\right)L_1L_2+bL_1^2-bL_2^2+\left(k_1b-k_2a\right)L_1+\left(k_1c-k_2b\right)L_2+\\
&+ \Lambda\left(bL_1+cL_2\right)\nu_1-\Gamma\nu_2,\\
\dot\nu_1 &= cL_3\nu_2-\left(cL_2+bL_1\right)\nu_3,\quad \dot\nu_2=\left(aL_1+bL_2\right)\nu_3-cL_3\nu_1,\\
\dot\nu_3 &= -\left(aL_1+bL_2\right)\nu_2+\left(cL_2+bL_1\right)\nu_1.
\end{aligned} \tag{4}$$

From the first equation of the system (4) we can obtain the first integral (3). It has the form

$$L_1=\frac{ck_2}{b}. \tag{5}$$

Under conditions (2), (5) equations (4) are noticeably simplified and take the form:

$$\begin{gathered}
\dot{\overline{L}}_2=b\overline{L}_2L_3+\left(F-Gc\right)L_3-\Lambda c\nu_1L_3+\Gamma\nu_3,\quad \dot L_3=-b\overline{L}_2^2-\left(F-Gc\right)\overline{L}_2+\Lambda c\overline{L}_2\nu_1-\Gamma\nu_2,\\
\dot\nu_1=cL_3\nu_2-c\overline{L}_2\nu_3,\quad \dot\nu_2=-cL_3\nu_1+b\overline{L}_2\nu_3+F\nu_3,\quad \dot\nu_3=c\overline{L}_2\nu_1-b\overline{L}_2\nu_2-F\nu_2.
\end{gathered} \tag{6}$$

Here we introduce the following notations:

$$\overline{L}_2=L_2+k_2,\quad F=\frac{\left(ac-b^2\right)k_2}{b},\quad G=\frac{ck_2}{b}+k_1.$$

Equations (6) possess the following first integrals

$$\frac{c}{2}\left(\overline{L}_2^2+L_3^2\right)+\Gamma\nu_1=E;\quad \overline{L}_2\nu_2+L_3\nu_3+G\nu_1+\frac{\Lambda}{2}\nu_1^2=k;\quad \nu_1^2+\nu_2^2+\nu_3^2=1. \tag{7}$$

2. Reduction to the second order linear differential equation

Let us introduce now the dimensionless variables y and z

$$y = \overline{L}_2\sqrt{\frac{c}{\Gamma}}, \qquad z = L_3\sqrt{\frac{c}{\Gamma}},$$

the dimensionless time τ:

$$t = \frac{\tau}{\sqrt{\Gamma c}},$$

the dimensionless constants of the first integrals

$$h = \frac{E}{\Gamma}, \qquad p_1 = k\sqrt{\frac{c}{\Gamma}}$$

and the dimensionless parameters

$$d_1 = \frac{b}{c}, \quad Q = \Lambda\sqrt{\frac{c}{\Gamma}}, \quad A = \frac{F}{\sqrt{\Gamma c}}, \quad B = G\sqrt{\frac{c}{\Gamma}}.$$

Now we can rewrite equations (6) and the first integrals (7) in dimensionless form:

$$\begin{gathered}\frac{dy}{d\tau} = d_1yz + (A-B)\,z - Q\nu_1 z + \nu_3, \quad \frac{dz}{d\tau} = -d_1y^2 - (A-B)\,y + Q\nu_1 y - \nu_2,\\ \frac{d\nu_1}{d\tau} = z\nu_2 - y\nu_3, \quad \frac{d\nu_2}{d\tau} = d_1y\nu_3 - z\nu_1 + A\nu_3, \quad \frac{d\nu_3}{d\tau} = -d_1y\nu_2 + y\nu_1 - A\nu_2,\end{gathered} \tag{8}$$

$$\frac{y^2+z^2}{2} + \nu_1 = h, \quad y\nu_2 + z\nu_3 + B\nu_1 + \frac{Q}{2}\nu_1^2 = p_1, \quad \nu_1^2+\nu_2^2+\nu_3^2 = 1.$$

In the system (8) we make the following change of variables

$$y = x\cos\varphi, \quad z = x\sin\varphi.$$

Then for x and φ we have the system of two differential equations

$$\begin{gathered}x\frac{dx}{d\tau} = -\frac{\sqrt{P_8(x)}}{8},\\ x^2\frac{d\varphi}{d\tau} = -d_1x^3\cos\varphi - \frac{3Q}{8}x^4 + \frac{1}{2}(Qh+B-2A)\,x^2 - \frac{1}{2}\left(Qh^2+2Bh-2p_1\right),\end{gathered} \tag{9}$$

$$P_8(x) = -Q^2x^8 + 8\left(BQ+Q^2h-2\right)x^6 + 8\left(8h-2B^2-6BQh-3Q^2h^2+2p_1Q\right)x^4 +$$

$$+32\left(2-2h^2-2p_1B-2p_1Qh+2B^2h+3QBh^2+Q^2h^3\right)x^2 - 16\left(Qh^2+2Bh-2p_1\right)^2.$$

From this system it is easy to find the differential equation for $\varphi = \varphi(x)$:

$$\frac{d\varphi}{dx} = \frac{8d_1x^2}{\sqrt{P_8(x)}}\cos\varphi + \frac{3Qx^4 - 4(Qh+B-2A)\,x^2 + 4\left(2p_1 - Qh^2 - 2Bh\right)}{x\sqrt{P_8(x)}}. \tag{10}$$

Using the change of variables

$$w = \tan\frac{\varphi}{2},$$

we reduce this equation to the Riccati equation

$$\frac{dw}{dx} = f_2w^2 + f_0, \tag{11}$$

$$f_2 = \frac{3Qx^4 - 8d_1x^3 + 4\left(2A - B - Qh\right)x^2 + 4\left(2p_1 - Qh^2 - 2Bh\right)}{2x\sqrt{P_8\left(x\right)}},$$
$$f_0 = \frac{3Qx^4 + 8d_1x^3 + 4\left(2A - B - Qh\right)x^2 + 4\left(2p_1 - Qh^2 - 2Bh\right)}{2x\sqrt{P_8\left(x\right)}}. \tag{12}$$

It is well known from the general theory of ordinary differential equations [3], that if the Riccati equation has the form:

$$\frac{dw}{dx} = f_2w^2 + f_0,$$

then the substitution of the form

$$u\left(x\right) = \exp\left(-\int f_2wdx\right)$$

reduces it to the second order linear differential equation

$$\frac{d^2u}{dx^2} + a\left(x\right)\frac{du}{dx} + b\left(x\right)u = 0, \quad a\left(x\right) = -\frac{1}{f_2}\frac{df_2}{dx}, \quad b\left(x\right) = f_0f_2. \tag{13}$$

Taking into account that the functions f_2 and f_0 have the form (12), we can conclude, that the coefficients of the linear second order differential equation (13) are rational functions of x. Thus, we have the following theorem.

Theorem 1. *The solution of the problem of motion of a gyrostat with a fixed point under the action of gravity and gyroscopic forces in the integrable Hess case is reduced to solving the second order linear differential equation* (13) *with rational coefficients.*

Direct application of the Kovacic algorithm [2] to the differential equation (13) gives the following result.

Theorem 2. *Let $Q \neq 0$ (gyroscopic forces are present) and $d_1 \neq 0$ (the mass distribution of the gyrostat does not correspond to the Lagrange integrable case). Then the second order linear differential equation* (13) *admits a general solution expressed in terms of Liouvillian functions under the condition*

$$A = \frac{\left(d_1^2 + 1\right)}{Q}.$$

Funding: This work was supported by the grant of the Russian Science Foundation No. 24–11–20009

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] A. A. Kosov, On analogues of the Hess case for a gyrostat under the action of the moment of gyroscopic and circular forces, Mechanics of Solids 57 (2022) 1848–1861. doi:10.3103/S0025654422080192.

[2] J. J. Kovacic, An algorithm for solving second order linear homogeneous differential equations, Journal of Symbolic Computation 2 (1986) 3–43. doi:10.1016/S0747-7171(86)80010-4.

[3] A. D. Polyanin, V. F. Zaitsev, Handbook of Exact Solutions for Ordinary Differential Equations, CRC Press, Boca Raton–New York, 2003.

Symbolic Investigation of a Difference Scheme for Wave Processes in Coaxial Elastic Cylindrical Shells Filled with an Incompressible Viscous Fluid

Artem V. Mesyanzhin[1], Yury A. Blinkov[2,*]

[1]SC Design Bureau of Industrial Automatics, 239 B. Sadovaya St, Saratov, 410005, Russian Federation
[2]Saratov State University, 83 Astrakhanskaya St, Saratov, 410012, Russian Federation

Abstract

This study investigates difference schemes for modelling wave processes in coaxial elastic cylindrical shells filled with an incompressible viscous fluid. Using symbolic computation methods (Gröbner bases) and numerical methods, the authors derive a Crank–Nicolson-type scheme for a coupled nonlinear system of equations analogous to the Korteweg–de Vries (KdV) equation. The first differential approximation method is employed to analyze the scheme's accuracy, supported by exact soliton solutions and numerical experiments.

Keywords

Gröbner bases, difference schemes, first differential approximation, Korteweg–de Vries equation, nonlinear shells, viscous fluid

The problem of geometrically and physically nonlinear coaxial shells filled with a viscous incompressible fluid is of great importance for acoustic diagnostics and non-destructive material testing systems.

We assume that the radii of the shells are small compared to the characteristic wavelength of longitudinal deformation. The fluid flow is described by hydrodynamic lubrication theory, and there is no dependence on the Reynolds number. Transitioning to dimensionless variables and combining the method of multiple scales with the method of matched asymptotic expansions leads to the following system [3, 5]:

$$\begin{aligned} &\phi_t^{(1)} + 6\sigma_0\phi^{(1)}\phi_\eta^{(1)} + \phi_{\eta\eta\eta}^{(1)} 6\sigma_1 {\phi^{(1)}}^2 \phi_\eta^{(1)} + \phi^{(1)} - \phi^{(2)} = 0, \\ &\phi_t^{(2)} + 6\sigma_0\phi^{(2)}\phi_\eta^{(2)} + \phi_{\eta\eta\eta}^{(2)} 6\sigma_1 {\phi^{(2)}}^2 \phi_\eta^{(2)} + \phi^{(2)} - \phi^{(1)} - \sigma\phi^{(2)} = 0. \end{aligned} \tag{1}$$

We use Gröbner bases of the difference ideal to construct a Crank–Nicolson-type scheme, following the works [7, 4]. Represent the system of equations (1) in integral form:

$$\begin{aligned} &\oint_{\partial\Omega} \left(-3\sigma_0{\phi^{(1)}}^2 - 2\sigma_1{\phi^{(1)}}^3 - \phi_{\eta\eta}^{(1)}\right) dt + \phi^{(1)}\, d\eta + \iint_\Omega \left(\phi^{(1)} - \phi^{(2)}\right) dt\, d\eta = 0, \\ &\oint_{\partial\Omega} \left(-3\sigma_0{\phi^{(2)}}^2 - 2\sigma_1{\phi^{(2)}}^3 - \phi_{\eta\eta}^{(2)}\right) dt + \phi^{(2)}\, d\eta + \iint_\Omega \left(\phi^{(2)} - \phi^{(1)} - \sigma\phi^{(2)}\right) dt\, d\eta = 0. \end{aligned} \tag{2}$$

To transition to a discrete formulation, let ${u^{(i)}}_j^n = \phi^{(i)}(t_n, \eta_j)$ and choose the basic contour shown in Figure 1.

Supplement the system (2) with integral relations connecting discrete functions and their discrete derivatives, written in integral form:

$$\int_{\eta_j}^{\eta_{j+1}} {u^{(i)}}_\eta\, d\eta = u^{(i)}(t, \eta_{j+1}) - u^{(i)}(t, \eta_j), \quad \int_{\eta_j}^{\eta_{j+2}} {u^{(i)}}_{\eta\eta}\, d\eta = {u^{(i)}}_\eta(t, \eta_{j+2}) - {u^{(i)}}_\eta(t, \eta_j). \tag{3}$$

6th International Conference "Computer Algebra", Moscow, June 23–25, 2025
10.22363/12585-2025-6-017
EDN: CKRCFH
*Corresponding author.
a.v.mesyanzhin@gmail.com (A. V. Mesyanzhin); blinkovua@sgu.ru (Y. A. Blinkov)
0000-0002-7984-2168 (A. V. Mesyanzhin); 0000-0001-7340-0919 (Y. A. Blinkov)

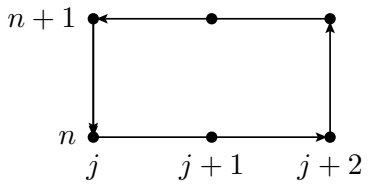


Figure 1: Basic difference stencil with integration direction along the contour

Using the trapezoidal rule for time integration and the first derivative with respect to η, and the mean value formula for the second derivative with respect to η, and setting $t_{n+1} - t_n = \tau$, $\eta_{j+1} - \eta_j = h$, rewrite the relations (2), (3) as:

$$\begin{gathered}
\left(-3\sigma_0\left({u^{(1)}}_j^{2n} + {u^{(1)}}_j^{2n+1} - {u^{(1)}}_{j+2}^{2n} - {u^{(1)}}_{j+2}^{2n+1}\right) -\right.\\
-2\sigma_1\left({u^{(1)}}_j^{3n} + {u^{(1)}}_j^{3n+1} - {u^{(1)}}_{j+2}^{3n} - {u^{(1)}}_{j+2}^{3n+1}\right) -\\
\left.-\left(u^{(1)n}_{\eta\eta\,j} + u^{(1)n+1}_{\eta\eta\,j} - u^{(1)n}_{\eta\eta\,j+2} - u^{(1)n+1}_{\eta\eta\,j+2}\right)\right)\cdot\frac{\tau}{2} + \left({u^{(1)}}_{j+1}^{n+1} - {u^{(1)}}_{j+1}^{n}\right)\cdot 2h +\\
+\left(\left({u^{(1)}}_{j+1}^{n+1} + {u^{(1)}}_{j+1}^{n}\right) - \left({u^{(2)}}_{j+1}^{n+1} + {u^{(2)}}_{j+1}^{n}\right)\right)\cdot h\tau = 0,\\
\left(-3\sigma_0\left({u^{(2)}}_j^{2n} + {u^{(2)}}_j^{2n+1} - {u^{(2)}}_{j+2}^{2n} - {u^{(2)}}_{j+2}^{2n+1}\right) -\right.\\
-2\sigma_1\left({u^{(2)}}_j^{3n} + {u^{(2)}}_j^{3n+1} - {u^{(2)}}_{j+2}^{3n} - {u^{(2)}}_{j+2}^{3n+1}\right) -\\
\left.-\left(u^{(2)n}_{\eta\eta\,j} + u^{(2)n+1}_{\eta\eta\,j} - u^{(2)n}_{\eta\eta\,j+2} - u^{(2)n+1}_{\eta\eta\,j+2}\right)\right)\cdot\frac{\tau}{2} + \left({u^{(2)}}_{j+1}^{n+1} - {u^{(2)}}_{j+1}^{n}\right)\cdot 2h +\\
+\left(\left({u^{(2)}}_{j+1}^{n+1} + {u^{(2)}}_{j+1}^{n}\right) - \left({u^{(1)}}_{j+1}^{n+1} + {u^{(1)}}_{j+1}^{n}\right) - \sigma\left({u^{(2)}}_{j+1}^{n+1} + {u^{(2)}}_{j+1}^{n}\right)\right)\cdot h\tau = 0,\\
({u^{(i)}_\eta}_{j+1}^{n} + {u^{(i)}_\eta}_{j}^{n})\cdot\tfrac{h}{2} = {u^{(i)}}_{j+1}^{n} - {u^{(i)}}_{j}^{n}, \quad u^{(i)n}_{\eta\eta\,j+1}\cdot 2h = {u^{(i)}_\eta}_{j+2}^{n} - {u^{(i)}_\eta}_{j}^{n}.
\end{gathered} \tag{4}$$

Since the PyGInv package for working with difference ideals operates only in the case of linear difference ideals, and the original differential equation (1) is nonlinear, we replace the nonlinear part by introducing an additional function $F^{(i)} = 3\sigma_0 {u^{(i)}}^2 + 2\sigma_1 {u^{(i)}}^3$. By choosing an admissible ordering such that $u^{(1)} \succ u^{(2)} \succ \ldots \succ F^{(1)} \succ F^{(1)}$, and then by variables n, j, the nonlinear part will not enter the leading monomials of the system when constructing the Gröbner basis, and the structure of the basis will allow checking the membership of the desired difference scheme.

As a result, we obtain the following difference scheme for the system of equations (1), analogous to the Crank–Nicolson scheme for the heat equation:

$$\begin{gathered}
\frac{{u^{(1)}}_j^{n+1} - {u^{(1)}}_j^{n}}{\tau} + 3\sigma_0\frac{({u^{(1)}}_{j+1}^{2n+1} - {u^{(1)}}_{j-1}^{2n+1}) + ({u^{(1)}}_{j+1}^{2n} - {u^{(1)}}_{j-1}^{2n})}{4h} +\\
+2\sigma_1\frac{({u^{(1)}}_{j+1}^{3n+1} - {u^{(1)}}_{j-1}^{3n+1}) + ({u^{(1)}}_{j+1}^{3n} - {u^{(1)}}_{j-1}^{3n})}{4h} + (({u^{(1)}}_{j+2}^{n+1} - 2{u^{(1)}}_{j+1}^{n+1} + 2{u^{(1)}}_{j-1}^{n+1} - {u^{(1)}}_{j-2}^{n+1}) +\\
+ ({u^{(1)}}_{j+2}^{n} - 2{u^{(1)}}_{j+1}^{n} + 2{u^{(1)}}_{j-1}^{n} - {u^{(1)}}_{j-2}^{n}))/4h^3 + \frac{{u^{(1)}}_j^{n+1} + {u^{(1)}}_j^{n}}{2} - \frac{{u^{(2)}}_j^{n+1} + {u^{(2)}}_j^{n}}{2} = 0,\\
\frac{{u^{(2)}}_j^{n+1} - {u^{(2)}}_j^{n}}{\tau} + 3\sigma_0\frac{({u^{(2)}}_{j+1}^{2n+1} - {u^{(2)}}_{j-1}^{2n+1}) + ({u^{(2)}}_{j+1}^{2n} - {u^{(2)}}_{j-1}^{2n})}{4h} +\\
+2\sigma_1\frac{({u^{(2)}}_{j+1}^{3n+1} - {u^{(2)}}_{j-1}^{3n+1}) + ({u^{(2)}}_{j+1}^{3n} - {u^{(2)}}_{j-1}^{3n})}{4h} + (({u^{(2)}}_{j+2}^{n+1} - 2{u^{(2)}}_{j+1}^{n+1} + 2{u^{(2)}}_{j-1}^{n+1} - {u^{(2)}}_{j-2}^{n+1}) +\\
+ ({u^{(2)}}_{j+2}^{n} - 2{u^{(2)}}_{j+1}^{n} + 2{u^{(2)}}_{j-1}^{n} - {u^{(2)}}_{j-2}^{n}))/4h^3 + \frac{{u^{(2)}}_j^{n+1} + {u^{(2)}}_j^{n}}{2} - \frac{{u^{(1)}}_j^{n+1} + {u^{(1)}}_j^{n}}{2} -
\end{gathered}$$

$$- \sigma \frac{{u^{(2)}}_j^{n+1} + {u^{(2)}}_j^n}{2} = 0. \quad (5)$$

The work [9] formulates the method of differential approximations of a difference scheme. Technically, it was applied only to evolutionary-type equations by constructing a Gröbner basis by replacing time derivatives with spatial derivatives in the series terms.

For the first difference scheme (5), this yields 6 terms for h^2 and 133 terms for τ^2:

$$6\sigma_0\phi_1\phi_{1x} + 6\sigma_1{\phi_1}^2\phi_{1x} + \phi_1 + \phi_{1t} + \phi_{1xxx} - \phi_2 + h^2\Big(\sigma_0\phi_1\phi_{1xxx} + 3\sigma_0\phi_{1xx}\phi_{1x} + \sigma_1{\phi_1}^2\phi_{1xxx} + \\ + 6\sigma_1\phi_1\phi_{1xx}\phi_{1x} + 2\sigma_1{\phi_{1x}}^3 + \frac{\phi_{1xxxxx}}{4}\Big) + \tau^2\Big(-\frac{\sigma^2\phi_2}{12} - \frac{\sigma\sigma_0\phi_1\phi_{2x}}{2} - \dots \\ \dots + \frac{\phi_{1xxxxxxxxx}}{12} + \frac{\phi_{1xxxxx}}{4} + \frac{\phi_{1xxx}}{2} - \frac{\phi_2}{3} - \frac{\phi_{2xxxxx}}{4} - \frac{\phi_{2xxx}}{2}\Big) + \dots = 0. \quad (6)$$

Similarly, for the second scheme, there are 6 terms for h^2 and 155 terms for τ^2:

$$\sigma\phi_2 + 6\sigma_0\phi_2\phi_{2x} + 6\sigma_1{\phi_2}^2\phi_{2x} - \phi_1 + \phi_2 + \phi_{2t} + \phi_{2xxx} + h^2\Big(\sigma_0\phi_2\phi_{2xxx} + 3\sigma_0\phi_{2xx}\phi_{2x} + \sigma_1{\phi_2}^2\phi_{2xxx} + \\ + 6\sigma_1\phi_2\phi_{2xx}\phi_{2x} + 2\sigma_1{\phi_{2x}}^3 + \frac{\phi_{2xxxxx}}{4}\Big) + \tau^2\Big(\frac{\sigma^3\phi_2}{12} + \frac{7\sigma^2\sigma_0\phi_2\phi_{2x}}{2} + \dots \\ \dots + \frac{\phi_{2xxxxxxxxx}}{12} + \frac{\phi_{2xxxxxx}}{4} + \frac{\phi_{2xxx}}{2}\Big) + \dots = 0. \quad (7)$$

The works [2] propose transitioning from lexicographic ordering to total degree reverse lexicographic ordering. This reduces computational volume in the case of nonlinear equations. In our case, for the first scheme, there are 15 terms for h^2 and one term for τ^2:

$$6\sigma_0\phi_1\phi_{1x} + 6\sigma_1{\phi_1}^2\phi_{1x} + \phi_1 + \phi_{1t} + \phi_{1xxx} - \phi_2 + h^2\Big(3\sigma_0^2{\phi_1}^2\phi_{1x} + 6\sigma_0\sigma_1{\phi_1}^3\phi_{1x} + \\ + \frac{\sigma_0{\phi_1}^2}{2} + \frac{\sigma_0\phi_1\phi_{1t}}{2} - \frac{\sigma_0\phi_1\phi_2}{2} - \frac{3\sigma_0\phi_{1xx}\phi_{1x}}{2} + 3\sigma_1^2{\phi_1}^4\phi_{1x} + \frac{\sigma_1{\phi_1}^3}{2} + \frac{\sigma_1{\phi_1}^2\phi_{1t}}{2} - \frac{\sigma_1{\phi_1}^2\phi_2}{2} - \\ - 3\sigma_1\phi_1\phi_{1xx}\phi_{1x} - \sigma_1{\phi_{1x}}^3 - \frac{\phi_{1txx}}{4} - \frac{\phi_{1xx}}{4} + \frac{\phi_{2xx}}{4}\Big) + \tau^2\Big(-\frac{\phi_{1ttt}}{12}\Big) + \dots = 0. \quad (8)$$

Similarly, for the second scheme, there are 18 terms for h^2 and one term for τ^2:

$$\sigma\phi_2 + 6\sigma_0\phi_2\phi_{2x} + 6\sigma_1{\phi_2}^2\phi_{2x} - \phi_1 + \phi_2 + \phi_{2t} + \phi_{2xxx} + h^2\Big(\frac{\sigma\sigma_0{\phi_2}^2}{2} + \frac{\sigma\sigma_1{\phi_2}^3}{2} - \\ - \frac{\sigma\phi_{2xx}}{4} + 3\sigma_0^2{\phi_2}^2\phi_{2x} + 6\sigma_0\sigma_1{\phi_2}^3\phi_{2x} - \frac{\sigma_0\phi_1\phi_2}{2} + \frac{\sigma_0{\phi_2}^2}{2} + \frac{\sigma_0\phi_2\phi_{2t}}{2} - \frac{3\sigma_0\phi_{2xx}\phi_{2x}}{2} + 3\sigma_1^2{\phi_2}^4\phi_{2x} - \\ - \frac{\sigma_1\phi_1{\phi_2}^2}{2} + \frac{\sigma_1{\phi_2}^3}{2} + \frac{\sigma_1{\phi_2}^2\phi_{2t}}{2} - 3\sigma_1\phi_2\phi_{2xx}\phi_{2x} - \sigma_1{\phi_{2x}}^3 + \frac{\phi_{1xx}}{4} - \frac{\phi_{2txx}}{4} - \frac{\phi_{2xx}}{4}\Big) + \\ + \tau^2\Big(-\frac{\phi_{2ttt}}{12}\Big) + \dots = 0 \quad (9)$$

The works [6] develop a method for verifying the quality of difference schemes using exact solutions.

The equations of the system (1) coincide when $\sigma = 0$. This corresponds to the absence of fluid in the inner shell. Consider the exact solution in the form of a soliton with a pedestal, which generalizes the Slyunyayev–Pelinovski solution [8] for $D = 0$:

$$\phi^{(1)} = \phi^{(2)} = D + \frac{k^2}{F + \sqrt{F^2 \mp k^2\sigma_1}\cosh(\xi)}. \quad (10)$$

Here, $\xi = k\left(\eta - (6D(\mp D\,\sigma_1 + \sigma_0) + k^2)t\right), F = \mp 2D\,\sigma_1 + \sigma_0$.

Substituting into the first differential approximation yields the following expression for the term at h^2:

$$
\begin{aligned}
&-k^3\left(-2\sqrt{A}B^3Cs_0\left(BD+k^2\right)+\sqrt{A}B^2\left(BD+k^2\right)^2\left(-2Cs_1+12s_0^2\right)+\right.\\
&\quad+4\sqrt{A}s_1\left(-Ak^6\sinh^2(\xi)+3k^4\left(BD+k^2\right)\left(\sqrt{A}B\cosh(\xi)-2A\sinh^2(\xi)\right)+\right.\\
&\quad\left.+3s_1\left(BD+k^2\right)^4\right)+B\left(6\sqrt{A}k^4s_0\left(\sqrt{A}B\cosh(\xi)-2A\sinh^2(\xi)\right)+\right.\\
&\quad+24\sqrt{A}s_0s_1\left(BD+k^2\right)^3-BCk^2\left(-\sqrt{A}B^2+2\sqrt{A}\left(\sqrt{A}B\cosh(\xi)-3A\sinh^2(\xi)\right)+\right.\\
&\quad\left.\left.\left.+4AB\cosh(\xi)\right)\right)\right)\sinh(\xi)/(4B^6) \qquad (11)
\end{aligned}
$$

Here, $A=\mp k^2\sigma_1+(\mp 2D\sigma_1+\sigma_0)^2$, $B=\mp 2D\sigma_1+\sigma_0+\sqrt{A}\cosh(\xi)$, and $C=6D(\mp D\sigma_1+\sigma_0)+k^2$. The term at τ^2 takes the following form:

$$
-C^3k^5\cdot\left(6A^{\frac{3}{2}}\sinh^2(\xi)+\sqrt{A}B^2-6AB\cosh(\xi)\right)\sinh(\xi)/(12B^4). \qquad (12)
$$

All constructions of the difference scheme itself and the first differential approximation used PyGInv (a version of GInv [1] written in pure Python using SymPy. For the abbreviated notation (11) and (12), SymPy and Python's text replacement capabilities were used.

Numerical experiments using numerical differentiation compared with the exact formulas (11) and (12) showed good agreement. Thus, a symbolic-numerical verification of the obtained scheme (5) was performed algorithmically, as in the works [6].

References

[1] Y. A. Blinkov, V. P. Gerdt, Specialized computer algebra system GINV, Programming and Computer Software 34 (2008) 112–123. doi:10.1134/S0361768808020096.

[2] Y. A. Blinkov, A. Y. Rebrina, Investigation of difference schemes for two-dimensional Navier–Stokes equations by using computer algebra algorithms, Programming and Computer Software 49 (2024) 26–31. doi:10.1134/S0361768823010024.

[3] Y. A. Blinkov, L. I. Mogilevich, A. V. Mesyanzhin, Wave occurrences computer modeling in two geometrically non-linear elastic coaxial cylindrical shells, containing viscous incompressible liquid, in: 2016 Dynamics of Systems, Mechanisms and Machines (Dynamics), 2016, pp. 1–4. doi:10.1109/Dynamics.2016.7818982.

[4] Y. A. Blinkov, V. P. Gerdt, K. B. Marinov, Discretization of quasilinear evolution equations by computer algebra methods, Programming and Computer Software 43 (2017) 84–89. doi:10.1134/S0361768817020049.

[5] Y. A. Blinkov, A. V. Mesyanzhin, L. I. Mogilevich, Nonlinear wave propagation in coaxial shells filled with viscous liquid, Computational Continuum Mechanics 10 (2017) 172–186. doi:10.7242/1999-6691/2017.10.2.15.

[6] Y. A. Blinkov, Symbolic and numerical investigation of finite difference schemes using first differential approximation: The case of the Korteweg–de Vries equation, Mathematical Modeling, Computational and Experimental Studies in Natural Sciences 4 (2024) 13. doi:10.24412/2541-9269-2024-4-20-32.

[7] V. P. Gerdt, Y. A. Blinkov, V. V. Mozzhilkin, Gröbner bases and generation of difference schemes for partial differential equations, Symmetry, Integrability and Geometry: Methods and Applications. 2 (2006) 26. doi:10.3842/SIGMA.2006.051.

[8] A. V. Slyunyaev, E. N. Pelinovski, Dynamics of large-amplitude solitonss, Journal of Experimental and Theoretical Physics 89 (1999) 173–181. doi:10.1134/1.558966.

[9] N. N. Yanenko, Y. I. Shokin, The first differential approximation to finite-difference schemes for hyperbolic systems of equations, Siberian Mathematical Journal 10 (1969) 868–880. doi:10.1007/BF00971662.

Combinatorics of Lyndon-Shirshov Words and Algorithms of Symbolic Computation in Lie Superalgebras

Alexander A. Mikhalev

Faculty of Mechanics and Mathematics, M.V. Lomonosov Moscow State University, Leninskie Gory, Moscow, 119234, Russian Federation

Abstract

We consider main combinatorial properties of Lyndon-Shirshov words and their applications in algorithms of symbolic computation in free Lie (p-)superalgebras and in Lie (p-)superalgebras given by generators and defining relations.

Keywords

Lyndon-Shirshov words, Lie superalgebras, computer algebra

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*

doi 10.22363/12585-2025-6-018

EDN: CPQGFH

aamikhalev@mail.ru (A. A. Mikhalev)

FEM Calculations of Coulomb Two Center Problem

Azalia I. Muratova[1], Alexander A. Gusev[2,3], Andrey G. Zorin[2] and Sergue I. Vinitsky[2,4,*]

[1]Tver State University, 33 Zhelyabova St, Tver, 170100, Russian Federation
[2]Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation
[3]Dubna State University, 19 Universitetskaya St, Dubna, 141980, Russian Federation
[4]RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

Abstract
An algorithm for solving of the boundary value problem for the discrete spectrum of a two-center Coulomb system is presented. The energy and separation constant eigenvalues and the corresponding eigenfunctions are calculated by the secant method on a suitable grid of the parameter, the distance between two Coulomb charges. The eigensolutions at each step of the secant method are calculated using KANTBP 5M program, which implements the finite element method in Maple.

Keywords
Coulomb two center problem, discrete spectrum, finite element method

1. Introduction

The boundary value problem (BVP) of two Coulomb centers (discrete spectrum) allows separation of variables in a prolate spheroidal coordinate system and is characterized by two eigenvalues, the energy $E(R)$ and the separation constant $\lambda(R)$, depending on a real parameter, the distance R between the centers of Coulomb charges Z_1 and Z_2. Discretization of the problem is traditionally carried out using expansions of eigenfunctions in functional series with constant coefficients, depending on R. This approach is applied in programs implemented in FORTRAN, see, e.g., [2] or in Wolfram Mathematica [3]. Of interest is an alternative discretization of this problem by the finite element method (FEM), in which polynomials of a lower degree of the order of 10 are used, in contrast to 100–200 in traditional Galerkin-type expansions. Using lower-order polynomials ensures greater resistance of the FEM computational scheme to rounding errors, i.e. it does not require the use high bit depth for arithmetic operations with floating point with an increase in the numbers of the sought eigenvalues.

The aim of the work is to develop a method and algorithm for solving the problem of two Coulomb centers with real eigenvalues of the discrete energy spectrum and the separation constant on the grid of the values of the real parameter R, the distance between the centers of Coulomb charges, using the secant method and the finite element method implemented in Maple by means of the KANTBP 5M program [1].

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-019
EDN: CUFETW
*Corresponding author.
azaliamuratova628@gmail.com (A. I. Muratova); gooseff@jinr.ru (A. A. Gusev); andrey.zorin@inbox.ru (A. G. Zorin); vinitsky@theor.jinr.ru (S. I. Vinitsky)

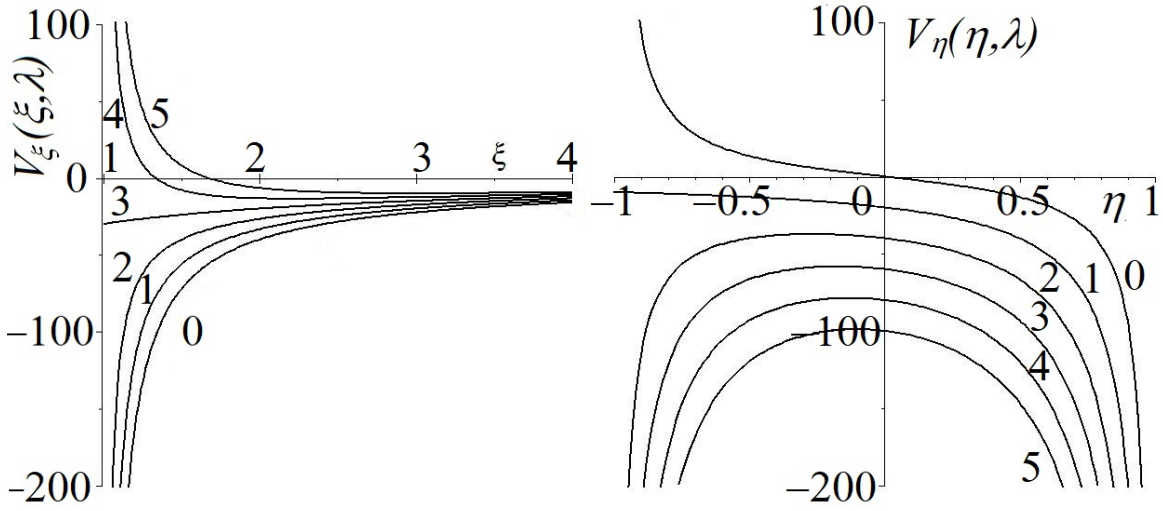


Figure 1: Potentials $V_\xi(\xi,\lambda)$ and $V_\eta(\eta,\lambda)$ at 0: λ=0; 1: λ=20; 2: λ=40; 3: λ=60; 4: λ=80; 5: λ=100.

2. Algorithm for solving the boundary value problem by the secant method and the KANTBP 5M program

In the input data file of KANTBP 5M program for solving BVPs, the ordinary second-order differential equation with variable coefficients is specified in the form [1]:

$$\left[-\frac{1}{f_B(z)}\frac{d}{dz}f_A(z)\frac{d}{dz}+V(z)-E\right]\Phi(z)=0. \tag{1}$$

Therefore, we represent the equations of the two-center Coulomb problem in prolate spheroidal coordinates for the eigenfunctions $F_{n_\xi m}(\xi;R)$ and $\Phi_{n_\eta m}(\eta;R)$ in the required form

$$\left[-\frac{1}{\xi^2-1}\frac{d}{d\xi}(\xi^2-1)\frac{d}{d\xi}-\epsilon_{n_\xi}(R)+V_\xi(\xi,\lambda(R))\right]F_{n_\xi m}(\xi;R)=0, \tag{2}$$

$$\left[-\frac{1}{1-\eta^2}\frac{d}{d\eta}(1-\eta^2)\frac{d}{d\eta}-\epsilon_{n_\eta}(R)+V_\eta(\eta,\lambda(R))\right]\Phi_{n_\eta m}(\eta;R)=0. \tag{3}$$

Here $\epsilon_{n_\xi}(R)=-p^2(R)$ and $\epsilon_{n_\eta}(R)=-p^2(R)$ are the eigenvalues, $V_\xi(\xi,\lambda_{n_\xi}(R))$ and $V_\eta(\eta,\lambda_{n_\eta}(R))$ are the potentials with parameters $a=(Z_1+Z_2)R$ and $b=(Z_2-Z_1)R$, $Z_2\geq Z_1$:

$$V_\xi(\xi,\lambda(R))=+\frac{\lambda(R)-a\xi}{\xi^2-1}+\frac{m^2}{(\xi^2-1)^2},\quad V_\eta(\eta,\lambda(R))=-\frac{\lambda(R)+b\eta}{1-\eta^2}+\frac{m^2}{(1-\eta^2)^2}, \tag{4}$$

depending on the separation constant $\lambda_{n_\xi}(R)=\lambda(R)$ and $\lambda_{n_\eta}(R)=\lambda(R)$ as a parameter (see Fig. 1). Note that problem (2) always has both a Coulomb discrete ($p^2>0$) and a continuous ($p^2<0$) spectrum, whereas problem (3) has only a discrete spectrum. Here the sign of λ is opposite to that of $\bar{\lambda}$ accepted in the ARSENY program [2], i.e., $\lambda(0)\geq 0$ and $\lambda=-\bar{\lambda}$. The asymptotic behavior of the solution is $F_{n_\xi m}(\xi;R)\sim(\xi^2-1)^{m/2}$ and $\Phi_{n_\eta m}(\eta;R)\sim(1-\eta^2)^{m/2}$. For zero azimuthal quantum number m=0, the eigenfunctions obey the Neumann condition:

$$\lim_{\xi\to 1}(\xi^2-1)\frac{dF_{n_\xi m}(\xi;R)}{d\xi}=0,\quad \lim_{\xi\to\xi_{max}}(\xi^2-1)\frac{dF_{n_\xi m}(\xi;R)}{d\xi}=0,\quad \lim_{\eta\to\mp 1\pm 0}(1-\eta^2)\frac{d\Phi_{n_\eta m}(\eta;R)}{d\eta}=0,$$

while for $m\neq 0$ the eigenfunctions obey the following Dirichlet and Neumann conditions:

$$F_{n_\xi m}(\xi=1;R)=0,\quad \lim_{\xi\to\xi_{max}}(\xi^2-1)\frac{dF_{n_\xi m}(\xi;R)}{d\xi}=0,\quad \lim_{\eta\to\mp 1\pm 0}(1-\eta^2)\Phi_{n_\eta m}(\eta;R)=0.$$

and the orthogonality and normalization conditions

$$\int_0^{\xi_{\max}} F_{n_\xi m}(\xi;R)F_{n'_\xi m}(\xi;R)(\xi^2-1)d\xi=\delta_{n_\xi n'_\xi},\ \int_{-1}^{1}\Phi_{n_\eta m}(\eta;R)\Phi_{n_\eta m}(\eta;R)(1-\eta^2)d\eta=\delta_{n_\eta n'_\eta}.$$

Note that the energy eigenvalues corresponding to $\epsilon_{n_\xi}(R)=-p^2(R)$ or $\epsilon_{n_\eta}(R)=-p^2(R)$ monotonically increase with increasing separation constant $\lambda(R)$ for Eq. (2) and monotonically decrease for Eq. (3), i.e., the differences $\epsilon_{n_\xi}(R)-\epsilon_{n_\eta}(R)$ monotonically increase with increasing $\lambda(R)$. This follows from Eq. (4) and is illustrated in Fig. 1. The BVP potentials have Coulomb asymptotic behavior $V_\xi(\xi\to\infty,\lambda)\to-1/\xi$, $V_\xi(\xi\to1+0,\lambda)\sim-1/(\xi-1)$, $V_\eta(\eta\to1-0,\lambda)\sim-1/(1-\eta)$, $V_\eta(\eta\to-1+0,\lambda)\sim-1/(1+\eta)$. This circumstance allows us to classify the solutions of the original problem by the number of zeros n_ξ and n_η in the variables ξ and η, respectively. Note that the discrete spectrum $\epsilon_{n_\xi}(R)=-p^2(R)$ of Eq. (2) is countable and lies in the interval $\epsilon_{n_\xi}(R)\in(\epsilon_{\min},0)$. This allows solving the problem for a given set of quantum numbers (n_ξ,n_η,m) or $(N=n_\xi+n_\eta+m+1,l=n_\eta-m,m)$ for a fixed $m\geq0$ using the initial approximation $E_{n_\xi,n_\eta,m}(0)=-(Z_1+Z_2)^2/N^2$ and $\lambda_{n_\xi,n_\eta,m}(0)=l(l+1)$ by means of the following algorithm that implements the secant method for solving equation $f(x)=0$:

$$x^{(s+1)}=[f(x^{(s)})x^{(s-1)}-f(x^{(s-1)})x^{(s-1)}]/[f(x^{(s)})-f(x^{(s-1)})],\quad s=1,2,...$$

with initial values $x^{(1)}$ and $x^{(0)}$ at $f(x)=\epsilon_{n_\xi}(\lambda;R)-\epsilon_{n_\eta}(\lambda;R)$, $x=\lambda$.

Algorithm SECANT for calculating eigenvalues and eigenfunctions of two Coulomb centers by the secant method and the KANTBP 5M program for solving BVP for Eqs. (2) and (3)
Input: Z_1, Z_2 are the problem parameters
N is the maximum number,
Ω_ξ and Ω_η are the grids for BVPs for Eq. (2) and for Eq. (3)
Output: $\lambda(R)$, $-p^2(R)$, $F_{n_\xi m}(\xi;R)$, and $\Phi_{n_\eta m}(\eta;R)$ are the solutions of BVPs for Eqs. (2) and (3)
Cycle over $n_\xi=0,...,N_{\max}$ and $n_\eta=0,...,N_{\max}-n_\xi-m-1$
Step 1 Initial approximation of interval boundaries $\lambda\in[\lambda_0,\lambda_1]$ for the first program run
Step 2 Loop by parameter $R=\{R_{\min}(\delta R)R_{\max}\}$ with step δR
Step 2.1 Reevaluation $a=(Z_1+Z_2)R$ and $b=(Z_2-Z_1)R$
Step 2.2 Calculation of the eigenvalue $\epsilon^{(0)}_{n_\xi}\equiv\epsilon_{n_\xi}$ of the BVP for Eq. (2) at $\lambda=\lambda_0$
Step 2.3 Calculation of the eigenvalue $\epsilon^{(1)}_{n_\xi}\equiv\epsilon_{n_\xi}$ of the BVP for Eq. (2) at $\lambda=\lambda_1$
Step 2.4 Calculation of the eigenvalue $\epsilon^{(0)}_{n_\eta}\equiv\epsilon_{n_\eta}$ of the BVP for Eq. (3) at $\lambda=\lambda_0$
Step 2.5 Calculation of the eigenvalue $\epsilon^{(1)}_{n_\eta}\equiv\epsilon_{n_\eta}$ of the BVP for Eq. (3) at $\lambda=\lambda_1$
Step 2.6 Calculate the energy differences for the first two approximations
$\epsilon_0:=\epsilon^{(0)}_{n_\xi}-\epsilon^{(0)}_{n_\eta}$ and $\epsilon_1:=\epsilon^{(1)}_{n_\xi}-\epsilon^{(1)}_{n_\eta}$
Step 2.7 New approximation for λ and energy difference using the formula of the secant method $\lambda=(\epsilon_1\lambda_0-\epsilon_0\lambda_1)/(\epsilon_1-\epsilon_0)$ and $\delta\epsilon=\epsilon_1-\epsilon_0$
Step 2.8 secant method: loop is executed until $|\delta\epsilon|>\delta$
Step 2.8.1 Calculation of the eigenvalue ϵ_{n_ξ} of the BVP for Eq. (2) at given λ
Step 2.8.2 Calculation of the eigenvalue ϵ_{n_η} of the BVP for Eq. (3) at given λ
Step 2.8.3 Calculating the energy difference $\epsilon:=\epsilon_{n_\xi}-\epsilon_{n_\eta}$
Step 2.8.4 Selecting the initial approximations $(\epsilon_0,\lambda_0)=(\epsilon,\lambda)$ or $(\epsilon_1,\lambda_1)=(\epsilon,\lambda)$ (they will also be used when moving to the new R)
Step 2.8.5 New approximation for λ and energy difference using the formula of the secant method $\lambda=(\epsilon_1\lambda_0-\epsilon_0\lambda_1)/(\epsilon_1-\epsilon_0)$ and $\delta\epsilon=\epsilon_1-\epsilon_0$
Step 2.8 End of loop of secant method
Step 2. End of loop by R
Step 3. Write OUTPUT: R, λ, $\epsilon_{n_\xi}\approx-p^2(R)$, $\epsilon_{n_\eta}\approx-p^2(R)$, $F_{n_\xi m}(\xi;R)$, and $\Phi_{n_\eta m}(\eta;R)$
End of cycle over n_ξ and n_η

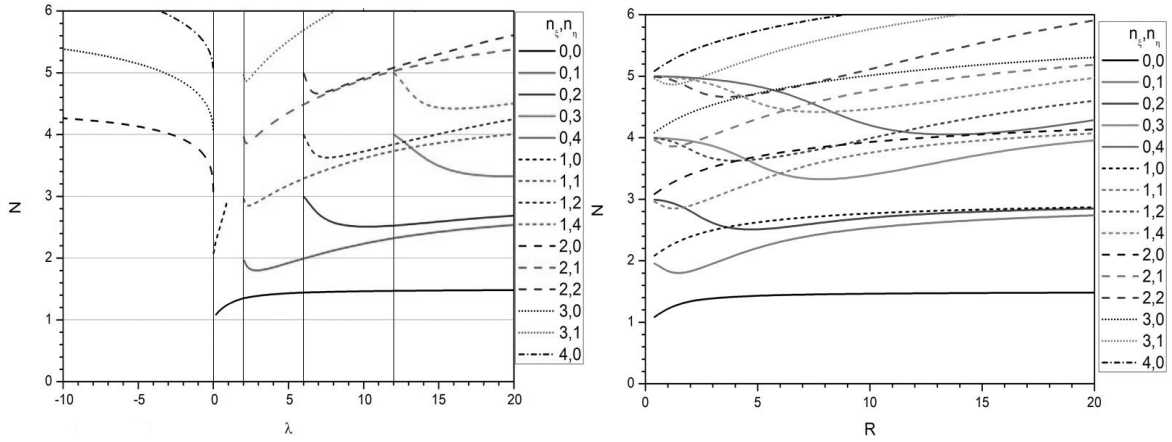


Figure 2: Effective quantum number $N \equiv N_{n_\xi,n_\eta,m}(\lambda_{n_\xi,n_\eta,m}(R))$ versus the parameter $\lambda \equiv \lambda_{n_\xi,n_\eta,m}(R)$, and $N \equiv N_{n_\xi,n_\eta,m}(R)=(Z_1+Z_2)/\sqrt{-2E_{n_\xi,n_\eta,m}(R)}$ versus the parameter R for $Z_1=1$, $Z_2=2$, and $m=0$.

As an example, we present results of the *Algorithm SECANT* calculations of a set of eigenvalues of energy $E_{n_\xi,n_\eta,m}(R)$ and parameter $\lambda_{n_\xi,n_\eta,m}(R)$ for $Z_1=1$, $Z_2=2$, $m=0$, and $N_{\max}=10$ on grids $\Omega_\xi=\{\xi_0=1, ..., \xi_i=\xi_{i-1}+0.08/\xi_{i-1}, ..., \xi_{n-1}=187.4197583, \xi_n=202.4133390\}$ and $\Omega_\eta=\{-1(0.1)1\}$ by means of program KANTBP 5M with 5-th order Hermite interpolation polynomials, which agree within 6-7 digits with those calculated by means of program ARSENY [2]. Figure 2 plots the effective quantum number $N_{n_\xi,n_\eta,m}(\lambda_{n_\xi,n_\eta,m}(R))$ as a function of the parameter λ, and $N_{n_\xi,n_\eta,m}(R)=(Z_1+Z_2)/\sqrt{-2E_{n_\xi,n_\eta,m}(R)}$ as a function of the parameter R for $Z_1=1$, $Z_2=2$, and $m=0$. Note that $N_{n_\xi,n_\eta,m}(0)=(Z_1+Z_2)/\sqrt{-2E_{n_\xi,n_\eta,m}(0)}$ at $R=0$, is equal to the value of the principal quantum number $N(0)=N=n_\xi+n_\eta+1$ of the united atom with the Coulomb charge $Z=Z_1+Z_2$, for the asymptotic energy eigenvalues $E_{Nlm}(R=0)$, and the asymptotic separation constant values $\bar{\lambda}_{lm}(R=0)=l(l+1)$, $l=n_\eta$ at $m=0$, which can be seen in the figures. Moreover, due to the separation of variables, the number of zeros n_ξ and n_η of the eigenfunctions $F_{m,n_\xi}(\xi;R)$ and $\Phi_{m,n_\eta}(\eta;R)$ is preserved for all values of the parameter R.

Note, for solving a continuous spectrum problem at a fixed value $E>0$, it is sufficient to solve eigenvalue problem for Eq. (3) and substitute a calculated eigenvalue $\lambda_{n_\xi m}$ to Eq. (2) and to solve the corresponding BVP with the mixed Neumann (or Dirichlet) and Robin boundary conditions. The algorithm SECANT can be also applied to calculate the series of branching points R_c sought for in the complex plane of distance R and the hidden crossings of complex energy curves $E_{n_\xi,n_\eta,m}(R)$ following the corresponding algorithms of ARSENY program [2].

Author Contributions: All authors contributed equally to the work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: This publication has been supported by the Peoples' Friendship University of Russia (RUDN) Strategic Academic Leadership Program (project No. 021934-0-000).

References

[1] G. Chuluunbaatar, A. Gusev, V. Derbov, S. Vinitsky, O. Chuluunbaatar, L. Hai, V. Gerdt, A Maple implementation of the finite element method for solving boundary-value problems

for systems of second-order ordinary differential equations, Communications in Computer and Information Science 1414 (2021) 152–162. doi:10.1007/978-3-030-81698-8_11.

[2] A. Gusev, E. Solov'ev, S. Vinitsky, ARSENY: A program for computing inelastic transitions via hidden crossings in one-electron atomic ion–ion collisions with classical description of nuclear motion, Computer Physics Communications 286 (2023) 108662. doi:10.1016/j.cpc.2023.108662.

[3] Wolfram mathematica, 2025. URL: https://www.wolfram.com/mathematica/?source=nav.

Symbolic Computing in the Constructing Problem of Controlled Compartmental Model

Alexey A. Petrov[1], Olga V. Druzhinina[2,*] and Olga N. Masina[1]

[1]Bunin Yelets State University, 28 Kommunarov St, Yelets, 399770, Russian Federation
[2]Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, 44/2 Vavilov St, Moscow, 119333, Russian Federation

Abstract

The results of developing a toolkit that enables symbolic computations in the process of solving problems of constructing controlled compartmental models of dynamic systems are presented. A controlled compartmental of epidemiology model SIRHU which is a generalization of the SIRH model to a controlled case is constructed and studied. Control actions are specified in the form of rules, the transition intensity in which is a changing parameter. When constructing this controlled model the implementation of a domain-specific language is used to construct models based on the interaction scheme between compartments and on the developed algorithm for the simulation implementation of models taking into account control. Computational experiments are conducted to simulate a controlled system of epidemic spread related to the SIRHU type. The analysis of the trajectory dynamics of the simulation model and the corresponding differential model is performed. The obtained results can be used to solve problems of modeling dynamic processes based on one-step interactions.

Keywords

compartmental model, one-step processes, scheme of interactions, computer algebra, dynamic systems, controlled models

1. Introduction

Modern research in the field of computer algebra is aimed at solving current problems of symbolic representation of dynamic systems models [9, 2]. However, key problems related to the development of effective algorithms and specialized software for the analysis of such models remain unresolved. In particular, an open problem is the automation of the construction and parameterization of controlled compartmental models including epidemiological systems.

The Julia programming language [1] provides powerful tools for symbolic computations, numerical modeling, and high-performance calculations making it a promising platform for solving computer algebra problems. In [4], a software package in Julia is proposed that allows constructing compartmental models based on a description of component interactions. However, issues of efficient symbolic reduction of complex models, automatic derivation of control equations, and optimization of their structure remain unexplored.

Of theoretical and applied interest are epidemiological models such as SIR, SEIR, SAPHIRE and SIDARTHE [3, 7, 12, 6, 10, 13] which require expansion by including control parameters and additional factors influencing the dynamics of infection spread. The significance of changes in the dynamics of asymptomatic morbidity taken into account in the SAPHIRE and SIDARTHE models should be noted.

Despite the active study of controlled models [5, 10] the question of developing universal symbolic analysis algorithms for such systems remains open including:

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-020
EDN: EFEJYM
*Corresponding author.
xeal91@yandex.ru (A. A. Petrov); ovdruzh@mail.ru (O. V. Druzhinina); olga121@inbox.ru (O. N. Masina)
0000-0002-3011-9278 (A. A. Petrov); 0000-0002-9242-9730 (O. V. Druzhinina); 0000-0002-0934-7217 (O. N. Masina)

- automatic generation of equations based on structural interaction diagrams;
- parametric identification under conditions of incomplete data;
- symbolic methods for global optimization of control actions.

Solving these problems requires new approaches at the intersection of computer algebra, machine learning, and optimization methods. In particular, the following issues remain poorly understood:

- integration of symbolic and numerical methods for analyzing controlled dynamic systems;
- development of domain-specific languages (DSL) for compact representation of models;
- application of macros and metaprogramming (for example, in Julia) for automating code generation.

This paper examines the algorithmic and software aspects of symbolic computations for controlled compartmental models. The basic focus is on:

- formalization of the SIRHU model representation using computer algebra methods;
- development of algorithms for symbolic generation of model equations;
- creation of specialized modules in Julia, including DSL and macros for automation of model construction;
- comparative analysis of trajectory dynamics in controlled systems.

In [11] the results of development of the software package, which implements the functionality for modeling compartmental systems based on the schemes of interactions between compartments, are presented. Julia language with the use of scientific computing libraries is chosen as the programming language. In the indicated article some models of epidemic spread are constructed using the developed algorithmic and software. The purpose of present paper is to develop and test the approach to construction of compartmental models by means of computer algebra. Construction and analysis of the controlled SIRHU model using symbolic computing is of both theoretical and applied interest.

2. Model description

The classic epidemic model is the SIR model proposed by W. Kermack and A. McKendrick [8]. This model distinguishes three states of individuals: S – susceptible, I – infected, R – recovered (or non-susceptible). The most famous modification of the SIR model is the SEIR model which takes into account the number of asymptomatic carriers of E.

We consider an extension of the compartmental SIR model, called SIRHU. In addition to compartments S, I, and R, this model takes into account an additional group of individuals: unregistered infected (H).

The interaction diagram for the SIRHU model has the form shown in Fig. 2. α, β, δ, γ denote the intensity of interaction between compartments, and u denotes the intensity of the control action to identify asymptomatic patients.

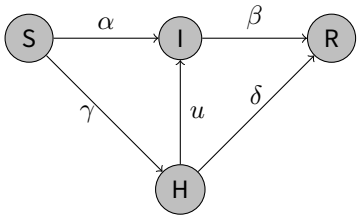


Figure 1: Interaction diagram for the SIRHU model

The interaction scheme for the SIRHU model is described by the following rules:

$$\begin{aligned} &i + s \xrightarrow{\alpha} 2i, \\ &i \xrightarrow{\beta} r, \\ &h + s \xrightarrow{\gamma} 2h, \\ &h \xrightarrow{u} i, \\ &h \xrightarrow{\delta} r. \end{aligned} \tag{1}$$

Rules of the form (1) are used for a meaningful description of the model in a software package for symbolic computing with an implementation based on DSL.

The software package provides for automatic generation of differential equations of the model. For the SIRHU model we have

$$\dot{K} = \begin{bmatrix} -hs\gamma - is\alpha \\ hu - i\beta + is\alpha \\ h\delta + i\beta \\ -hu - h\delta + hs\gamma \end{bmatrix}, K = [s, i, r, h]^{\mathsf{T}}. \tag{2}$$

Equations (2) correspond to the continuous form of the SIRHU dynamic model.

3. Results of experiments

Numerical experiments were conducted for the SIRHU model using the simulation algorithm described in [11]. Fig. 3 shows the dynamics of the numbers of SIRHU compartments for the case that corresponds to a constant intensity of the control action.

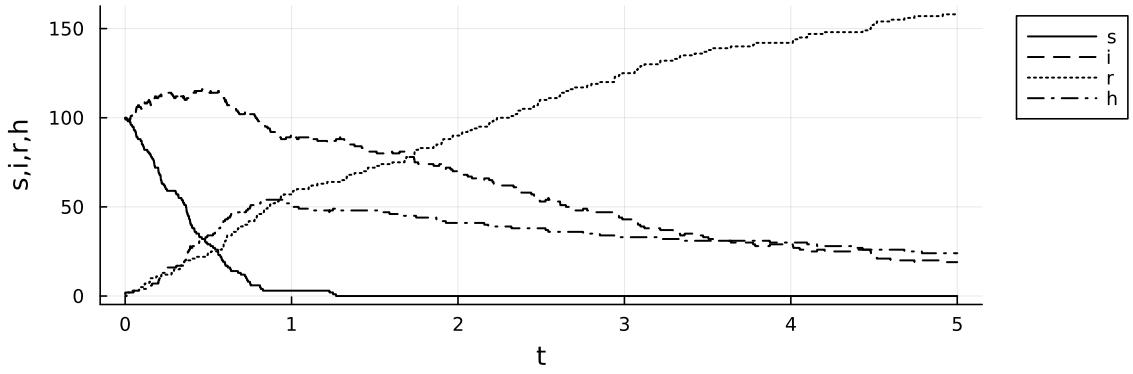


Figure 2: Dynamics of the SIRHU compartment numbers for $\alpha, \beta, \gamma, \delta = (0.01, 0.5, 0.1, 0.1)$ and $u = 0.1$

The results of the experiments showed the consistency of the simulation modeling for SIRHU with the results of the numerical solution of differential equations (2). In the framework of this paper, calculations are also carried out taking into account the dynamic change of the parameter u.

4. Conclusion

In this paper, an approach to creating algorithmic support for symbolic analysis of controlled compartmental models of dynamic systems is developed. The results demonstrate the capabilities of symbolic methods in epidemic modeling problems and highlight promising areas for further research in computer algebra. Among the promising areas for the development of the

paper we can note the modification of DSL taking into account the composite coefficients of interaction between compartments, as well as additional variables describing the quantitative composition of population groups.

References

[1] J. Bezanson, S. Karpinski, V. Shah, A. Edelman, Julia: A fast dynamic language for technical computing, 2012. URL: https://arxiv.org/abs/1209.5145.

[2] B. Bittner, K. Sreenath, Symbolic Computation of Dynamics on Smooth Manifolds, Springer International Publishing, Cham, 2020, pp. 336–351. doi:10.1007/978-3-030-43089-4_22.

[3] F. Brauer, Mathematical epidemiology: Past, present, and future, Infectious Disease Modelling 2 (2017) 113–127.

[4] A. Demidova, O. Druzhinina, O. Masina, A. Petrov, Constructing compartmental models of dynamic systems using a software package for symbolic computation in Julia, Programming and Computer Software 50 (2024) 138–146.

[5] D. Gromov, I. Bulla, E. Romero-Severson, O. Serea, Numerical optimal control for HIV prevention with dynamic budget allocation, Mathematical Medicine and Biology 35 (2016) 469–491.

[6] X. Hao, S. Cheng, D. Wu, T. Wu, X. Lin, C. Wang, Reconstruction of the full transmission dynamics of COVID-19 in Wuhan, Nature 584 (2020) 420–424. doi:10.1038/s41586-020-2554-8.

[7] S. Kabanikhin, O. Krivorotko, Optimization methods for solving inverse immunology and epidemiology problems, Computational Mathematics and Mathematical Physics 60 (2020) 580–589.

[8] W. Kermack, A. McKendrick, Contributions to the mathematical theory of epidemics, Bulletin of Mathematical Biology 53 (1991) 33–55. doi:10.1007/BF02464423.

[9] A. Korol'kova, M. Gevorkyan, A. Fedorov, K. Shtepa, D. Kulyabov, Symbolic studies of Maxwell's equations in space-time algebra formalism, Programming and Computer Software 50 (2024) 166–171.

[10] V. Kotin, A. Sychugina, Optimization of software-based control of vaccination procedures, Biomedical Radioelectronics (2016) 25–30.

[11] A. Petrov, O. Druzhinina, O. Masina, A. Demidova, Development of algorithmic and software support for symbolic computations in problems of constructing controlled compartmental models of dynamic systems, Programming and Computer Software 51 (2025) 21–31. doi:10.1134/S0361768824700580.

[12] S. Purkayastha, R. Bhattacharyya, R. Bhaduri, R. Kundu, X. Gu, M. Salvatore, D. Ray, S. Mishra, B. Mukherjee, A comparison of five epidemiological models for transmission of SARS-CoV-2 in India, BMC Infectious Diseases 21 (2021) 533. doi:10.1186/s12879-021-06077-9.

[13] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (2020) 270–273. doi:10.1038/s41586-020-2012-7.

Symbolic Computation in Studying the Dynamics of Two-Body System with Variable Masses

Alexander Prokopenya[1,*], Mukhtar Minglibayev[2,3] and Aigerim Ibraimova[2]

[1]Warsaw University of Life Sciences, Nowoursynowska 166, 02-787, Warsaw, Poland
[2]Al-Farabi Kazakh National University, Al-Farabi av. 71, 050040, Almaty, Kazakhstan
[3]Fesenkov Astrophysical Institute, Observatoriya 23, 050020, Almaty, Kazakhstan

Abstract
We discuss symbolic calculations in studying the classical two-body problem with variable masses. The Gyldén equation describing relative motion of the bodies is not integrable, in general, and so the problem is solved in the framework of the perturbation theory. Two approaches to reduce the Gyldén equation to the form appropriate for application of the perturbation methods are discussed. It is shown that although differential equations of the perturbed motion are different the both models demonstrate similar behavior of the secular perturbations of the orbital elements. All the relevant symbolic calculations are performed with the computer algebra system Wolfram Mathematica.

Keywords
two-body problem, variable mass, Gyldén equation, perturbation methods, evolutionary equations, computer algebra

1. Introduction

The classical two-body problem describes motion of two particles P_0, P_1 of constant masses m_0, m_1, respectively, attracting each other according to Newton's law of universal gravitation. Its general solution is well known and is used as a first approximation in a wide variety of orbital motion problems (see [8]). In practice, attraction of other bodies disturbs this motion and enforces its orbital parameters to change. However, application of the perturbation theory enables to investigate the system motion accurately (see, for example, [5]).

Dependence of the masses on time complicates the two-body problem substantially and its general solution cannot be written in symbolic form, in general. In the simplest case, the bodies can lose their masses isotropically and the only effect of the loss of mass is the decrease in the attraction between the two bodies. Then, the relative motion of one body around the other may be described by the following differential equation:

$$\frac{d^2\vec{r}}{dt^2} + Gm(t)\frac{\vec{r}}{r^3} = 0, \tag{1}$$

where G is the gravitational constant, $\vec{r}$ is the radius-vector of the body P_1 relative to P_0, and $m(t) = m_0(t) + m_1(t)$ is the total mass of the system, which is a given function of time. Equation (1) was proposed first by Gyldén [1] and is known now as *Gyldén's problem*.

In the case of constant masses a general solution to Equation (1) is known and describes motion of the bodies on conic sections the orbital elements of which depend on the total mass $m(t) = const$ (see [8]). Therefore, it appears reasonable that variability of masses may disturb this motion and enforce the orbital elements to change with time.

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-021
EDN: EMTDVZ
*Corresponding author.
alexander_prokopenya@sggw.edu.pl (A. Prokopenya); minglibayev@gmail.com (M. Minglibayev); ibraimova@aphi.kz (A. Ibraimova)
0000-0001-9760-5185 (A. Prokopenya); 0000-0002-8724-2648 (M. Minglibayev); 0000-0002-6998-8323 (A. Ibraimova)

The main purpose of this talk is to present two models which are used most often to reduce (1) to the form being appropriate for application of the perturbation theory and to derive the differential equations determining evolution of the orbital elements. Derivation of such equations involves quite cumbersome symbolic calculations which may be performed efficiently with the aid of the computer algebra system Wolfram Mathematica [9].

2. Models description

One model was proposed in [2, 6], where the isotropic loss of mass was interpreted as additional force given by

$$\vec{F} = \frac{\dot{\gamma}(t)}{2\gamma(t)}\frac{d\vec{r}}{dt}. \tag{2}$$

Here the dot means derivation with respect to time and the function $\gamma(t)$ is assumed to be differentiable and given by

$$\gamma(t) = \frac{m_{00} + m_{10}}{m_0(t) + m_1(t)}, \tag{3}$$

where $m_{00} = m_0(t_0)$, $m_{10} = m_1(t_0)$ are the masses of the bodies P_0, P_1, respectively, at the initial instant of time t_0.

Adding to the both sides of (1) the same term $\vec{F}$, we rewrite it in the form

$$\frac{d^2\vec{r}}{dt^2} + Gm(t)\frac{\vec{r}}{r^3} + \frac{\dot{\gamma}(t)}{2\gamma(t)}\frac{d\vec{r}}{dt} = \vec{F}. \tag{4}$$

Note that (4) becomes integrable in the case of $\vec{F} = 0$ and its exact solution describes aperiodic motion of the body P_1 on a conic section (see [7]). The nonzero perturbing force (2) disturbs this motion and its orbital elements must necessarily vary with the time. Derivation of the differential equations determining evolution of the orbital elements is given in [3], where the corresponding symbolic calculations are described in detail. Finally, we obtain the following system of differential equations determining the dependence of the orbital elements on time:

$$\frac{da}{dt} = \frac{a(1 + e\cos E)}{1 - e\cos E}\frac{\dot{\gamma}(t)}{\gamma(t)}, \quad \frac{de}{dt} = \frac{(1 - e^2)\cos E}{1 - e\cos E}\frac{\dot{\gamma}(t)}{\gamma(t)}, \tag{5}$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1 - e^2}\sin E}{e(1 - e\cos E)}\frac{\dot{\gamma}(t)}{\gamma(t)}, \quad \frac{di}{dt} = 0, \quad \frac{d\Omega}{dt} = 0, \tag{6}$$

$$\frac{dE}{dt} = \frac{\sqrt{\kappa}}{a^{3/2}(1 - e\cos E)\sqrt{\gamma(t)}} - \frac{\sin E}{e(1 - e\cos E)}\frac{\dot{\gamma}(t)}{\gamma(t)}, \tag{7}$$

where $\kappa = G(m_{00} + m_{10})$ and $a, e, i, \Omega, \omega, E$ are the analogues of the Keplerian orbital elements (see [8, 4]).

The Gyldén equation (1) may be also rewritten in the form

$$\frac{d^2\vec{r}}{dt^2} + Gm(t)\frac{\vec{r}}{r^3} - \frac{\ddot{\gamma}(t)}{\gamma(t)}\vec{r} = \vec{F}, \tag{8}$$

where the function $\gamma(t)$ is defined by (3) and the perturbing force $\vec{F}$ is given by

$$\vec{F} = -\frac{\ddot{\gamma}(t)}{\gamma(t)}\vec{r}. \tag{9}$$

This approach was exploited in [4] in studying the three-body problem with variable masses. Equation (8) also becomes integrable in the case of $\vec{F} = 0$ and its exact solution describes aperiodic motion of the body P_1 on a quasi-conic section. In the framework of this model the

differential equations of the perturbed motion are obtained in terms of the osculating elements of the aperiodic motion on quasi-conic section (see [4, 3, 7]) and are given by

$$\frac{da}{dt} = -\frac{2a^{5/2}e\sin E}{\sqrt{\kappa}}\gamma(t)\ddot{\gamma}(t), \quad \frac{de}{dt} = -\frac{a^{3/2}(1-e^2)\sin E}{\sqrt{\kappa}}\gamma(t)\ddot{\gamma}(t), \tag{10}$$

$$\frac{d\omega}{dt} = \frac{a^{3/2}(\cos E - e)\sqrt{1-e^2}}{e\sqrt{\kappa}}\gamma(t)\ddot{\gamma}(t), \quad \frac{di}{dt} = 0, \quad \frac{d\Omega}{dt} = 0, \tag{11}$$

$$\frac{dE}{dt} = \frac{\sqrt{\kappa}}{a^{3/2}(1-e\cos E)\gamma^2(t)} + \frac{a^{3/2}}{e\sqrt{\kappa}}\left(2e - (1+e^2)\cos E\right)\gamma(t)\ddot{\gamma}(t). \tag{12}$$

Note that the masses of the bodies vary quite slowly and the time derivatives $\dot{\gamma}(t)$, $\ddot{\gamma}(t)$ in the right-hand sides of equations (5)–(7), (10)–(12) are small. Therefore, the perturbing forces (2), (9) are also small and results in slow variation of the orbital parameters. As the forces (2), (9) act in the orbital plane, the inclination i of the orbital plane and the longitude of the ascending node Ω remain constant (see (6), (11)).

It should be emphasized that in the case of constant masses ($\gamma(t) = 1$) the right-hand sides of equations (5)–(6), (10)–(11) are equal to zero while equations (7), (12) determining the eccentric anomaly E as function of time in two models coincide. Therefore, the orbital elements a, e, i, Ω, ω become constants and coincide with the corresponding results in the two-body problem of constant mass (see [8]).

Choosing some realistic laws of the masses $m_0(t), m_1(t)$ variation, one can solve equations (5)–(7), (10)–(12) numerically and investigate the dynamics of the two-body system with variable masses. From the other side, averaging these equations over the eccentric anomaly, one can derive the evolutionary equations determining the secular perturbations of the orbital elements (see [7]). Note that the obtained results may be generalized to the case of three and more interacting bodies of variable masses. The calculations become much more complicated but application of the modern computer algebra system like Wolfram Mathematica enables to solve such problems successfully.

Author Contributions: Conceptualization, methodology, writing—review and editing, Alexander Prokopenya and Mukhtar Minglibayev; software, validation, investigation, Alexander Prokopenya and Aigerim Ibraimova. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] H. Gyldén, Die Bahnbewegungen in einem Systeme von zwei Körpern in dem Falle, dass die Massen Veränderungen unterworfen sind, Astronomische Nachrichten 109(1) (1884) 1–6. doi:https://doi.org/10.1002/asna.18841090102.

[2] J. D. Hadjidemetriou, Two-body problem with variable mass: A new approach, Icarus 2 (1963) 440–451. doi:https://doi.org/10.1016/0019-1035(63)90072-1.

[3] A. T. Ibraimova, M. Z. Minglibayev, A. N. Prokopenya, Study of secular perturbations in the restricted three-body problem of variable masses using computer algebra, Computational Mathematics and Mathematical Physics 63(1) (2023) 115–125. doi:10.1134/S0965542523010098.

[4] M. Z. Minglibayev, Dynamics of gravitating bodies with variable masses and sizes, LAMBERT Academic Publ., Saarbrucken, 2012. (in Russian).

[5] C. D. Murray, S. F. Dermott, Solar System Dynamics, Cambridge University Press, 2000. doi:10.1017/CBO9781139174817.
[6] T. B. Omarov, Two-body motion with corpuscular radiation, Soviet Astronomy 7(5) (1964) 707–711. URL: https://adsabs.harvard.edu/full/1964SvA.....7..707O.
[7] A. N. Prokopenya, M. Z. Minglibayev, A. T. Ibraimova, Perturbation methods in solving the problem of two bodies of variable masses with application of computer algebra, Applied Sciences 14 (2024) 11669. doi:10.3390/app142411669.
[8] A. E. Roy, Orbital Motion, 4th ed., IoP: Bristol, UK, 2005.
[9] S. Wolfram, An elementary introduction to the Wolfram Language, 2nd ed., Wolfram Media, New York, 2016.

The Generic-Case Complexity of Finding a Binary Solution to a System of Linear Equations

Alexandr V. Seliverstov

Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), 19, Bolshoy Karetny per., Moscow, 127051, Russian Federation

Abstract

Using the Schwartz–Zippel lemma, we estimate the generic-case complexity of finding a binary solution to a system of linear equations over an infinite field. Compared to previous works, the new algorithm is applicable to systems with a smaller number of linear equations, but it requires more running time. The algorithm works over a field of any characteristic. If the system has many binary solutions, then our algorithm is not applicable. So, in the worst case, there is no polynomial upper bound on the computational complexity.

Keywords

boolean programming, knapsack, matrix rank, polynomial, heuristics, computational complexity

1. Introduction

Let us consider the recognition problem whether there is a $\{0,1\}$-solution to a system of linear equations. The problem is NP-complete not only over the ring of integers, but also over the field of residues modulo any odd prime. Over the ring of integers, under a constraint on the size of the coefficients, a heuristic polynomial-time algorithm is known based on finding the shortest nonzero vector in an integer lattice [5]. The problem is also related to some optimization problems known as the multidimensional knapsack problem and Boolean programming, refer to the review [3] and recent article [1]. Constraints for dimensionality reducing of the problem by means of projection onto a coordinate hyperplane are known [2]. On the other hand, some systems of quadratic equations have been considered recently [4].

Over an arbitrary field K of characteristic $\mathrm{char}(K) \neq 2$, for almost all systems having at least $n - \sqrt{2n - o(n)}$ linear equations in n variables, a heuristic polynomial-time algorithm had been proposed several years ago [9]. In this work, the restriction on the number of equations is significantly relaxed, although the computational complexity increases. We consider systems over an arbitrary computable field. The computational complexity is estimated by the number of algebraic operations over the field.

For a recognition problem, let us assume three possible answers: the input may not only be accepted or rejected, but also an explicit notification of uncertainty of the choice is possible. In any case, the answer must be obtained in a finite time and without errors, and if an easily verifiable condition is met, then the notification of uncertainty can be issued only for a small fraction of inputs among all inputs of a given length. Such algorithms are called *generic* [6] or *errorless heuristics*. It is known that an NP-complete problem can be split into several subproblems that are also NP-complete [8]. Generic algorithms for an NP-complete problem can be considered for determining NP-complete subproblems for which the generic algorithm gives an uncertain answer.

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-022
EDN: EZKBKJ
slvstv@iitp.ru (A. V. Seliverstov)
0000-0003-4746-6396 (A. V. Seliverstov)

To estimate the number of inputs of a given size on which the algorithm quickly makes the correct decision, we use the Schwartz–Zippel lemma [7].

Lemma 1. *Given a non-constant polynomial $f(x_1, \ldots, x_n)$ of degree d over a field K. If random variables ξ_1, ..., ξ_n are independent and uniformly distributed on a finite set $S \subseteq K$ of cardinality $|S|$, then the inequality*

$$\text{Prob}\left[f(\xi_1, \ldots, \xi_n) = 0\right] \leq \frac{d}{|S|}$$

holds, where Prob$[\cdot]$ *denotes the probability of the condition indicated in square brackets.*

2. Results

Let us consider a system of m linear equations in $n > m$ variables:

$$\begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n + 1 = 0 \\ \cdots \\ \alpha_{m1}x_1 + \cdots + \alpha_{mn}x_n + 1 = 0 \end{cases}.$$

Multiplying each linear equation by each of the variables and taking into account the equalities $x_k^2 = x_k$, which are satisfied with $\{0,1\}$-solutions, we obtain mn new equations of the second degree. Discarding the terms depending only on one variable, we obtain a set of mn bilinear forms, the coefficients of which form a matrix W. The rows correspond to the bilinear forms, and the columns correspond to monomials of the form x_jx_k for $j < k$.

Lemma 2. *Let the matrix W be computed for m linear equations in n variables over a purely transcendental extension of the field K, where all coefficients α_{ij} are algebraically independent of each other. The rank of the matrix satisfies the inequality*

$$\text{rank}(W) \geq mn - \frac{m(m+1)}{2}.$$

Example 1. For $n = 3$ and $m = 1$, the 3×3 matrix

$$W = \begin{pmatrix} \alpha_{12} & \alpha_{13} & 0 \\ \alpha_{11} & 0 & \alpha_{13} \\ 0 & \alpha_{11} & \alpha_{12} \end{pmatrix}$$

is degenerate over a field of characteristic $\text{char}(K) = 2$ because $\det(W) = -2\alpha_{11}\alpha_{12}\alpha_{13}$. Next, for $n = 5$ and $m = 2$, the 10×10 matrix W is degenerate over any field because $\text{rank}(W) \leq 9$. For $n = 7$ and $m = 3$, the 21×21 matrix W is also degenerate over any field because $\text{rank}(W) \leq 18$. (The rank is computed with SymPy.)

Let the number of equations m be such that $mn \geq \text{rank}(W) + n - m$. This inequality holds for $m > n/2$. But a smaller number m is sufficient because the rank of W is small. In the general case, n linearly independent linear equations can be derived from resulting quadratic equations as well as the initial linear equations. In particular, there are new linear equations. Next, using these n linearly independent linear equations, one can find a solution and check whether it consists of zeros and ones. Of course, the method is not applicable when the system has many $\{0,1\}$-solutions. Thus, we have a polynomial upper bound on the generic-case complexity, but we have nothing in the worst case.

Example 2. Let us consider a linear equation in two variables $\alpha x_1 + \beta x_2 + 1 = 0$. Multiplying this equation by each of the variables and taking into account the equalities $x_k^2 = x_k$, which are satisfied with $\{0,1\}$-solutions, we obtain two equations:

$$\begin{cases} \beta x_1x_2 + (1+\alpha)x_1 = 0 \\ \alpha x_1x_2 + (1+\beta)x_2 = 0 \end{cases}.$$

This yields the linear equation $\alpha(1+\alpha)x_1 = \beta(1+\beta)x_2$. For $\alpha = \beta = -1$, this equation turns into the identity. But in the general case, it is a new linear equation, which is linearly independent of original one. So, one can either find the $\{0,1\}$-solution or prove its absence.

The probability of success is equal to the probability that the determinant of an $n \times n$ matrix does not vanish. A bound can be obtained using the Schwartz–Zippel lemma, i. e., Lemma 1. Let K denote a field and ε denote a positive real parameter.

Theorem. *For our algorithm, there is an univariate function $f(n)$ so that if n is even, $m \geq n/2$, and the coefficients α_{ij} are uniformly and independently distributed on the set $S \subset K$ of cardinality $\lceil f(n)/\varepsilon \rceil$, then the upper bound on the probability of the uncertain answer equals ε. The generic-case complexity is equal to the complexity of finding the rank of W.*

Proof. Let us consider the special system of exactly $n/2$ linear equations, where the k-th equation depends on two variables x_{2k-1} and x_{2k}:

$$\begin{cases} \beta_1 x_1 & +\beta_2 x_2 + 1 & = & 0 \\ \cdots & \cdots & & \\ \beta_{2k-1} x_{2k-1} & +\beta_{2k} x_{2k} + 1 & = & 0 \\ \cdots & & & \\ \beta_{n-1} x_{n-1} & +\beta_n x_n + 1 & = & 0 \end{cases}.$$

As in Example 2, new linear equations are

$$\begin{cases} \beta_1(1+\beta_1)x_1 & -\beta_2(1+\beta_2)x_2 & = & 0 \\ \cdots & \cdots & & \\ \beta_{2k-1}(1+\beta_{2k-1})x_{2k-1} & -\beta_{2k}(1+\beta_{2k})x_{2k} & = & 0 \\ \cdots & \cdots & & \\ \beta_{n-1}(1+\beta_{n-1})x_{n-1} & -\beta_n(1+\beta_n)x_n & = & 0 \end{cases}.$$

They together compose a system of n linear equations in n variables. Let us denote by M the $n \times n$ matrix of linear term coefficients. The determinant is a polynomial in coefficients β_1, ..., β_n of the initial system:

$$\det(M) = \pm \prod_{k=1}^{n/2} \beta_{2k-1}\beta_{2k}(2 + \beta_{2k-1} + \beta_{2k}),$$

where the sign depends on the order of equations. If all coefficients of the initial system are nonzero and inequalities $\beta_{2k-1} + \beta_{2k} \neq -2$ holds, then $\det(M)$ does not vanish.

In fact, the matrix M is not unique. Its entries are rational functions in β_1, ..., β_n. But one can compute M so that all entries are polynomials.

In the general case, using the first $n/2$ linear equations, our algorithm produces a nondegenerate system of n linear equations in n variables. It has unique solution. Let us consider the $n \times n$ matrix M of linear term coefficients. Its entries are rational functions in coefficients α_{ij} of the initial system. Without loss of generality, let the entries be polynomials. Thus, $\det(M)$ is a polynomial too. Moreover, the polynomial does not vanish identically. Let the degree upper bound be the desired function $f(n)$. In accordance with Lemma 1, the determinant does not vanish for almost all coefficients, i. e., our algorithm fails with probability at most ε.

3. Discussion

In accordance with Lemma 2, in the general case, such bounds based on the Schwartz–Zippel lemma cannot be significantly improved without increasing runtime. However, such an improvement is possible for sparse systems of equations with a fixed arrangement of nonzero coefficients.

The algorithm can be useful over a finite field too, although the Schwartz–Zippel lemma requires sufficiently many elements depending on the number of variables. Of course, if K is infinite, then a sufficiently large set $S \subset K$ exists for all n and ε.

Of course, if the number of variables is sufficiently large, then the worst-case computational complexity remains high. Nevertheless, in accordance with our result, the Merkle–Hellman cryptosystem based on the subset sum problem can be broken in almost all cases by means of a broadcast attack against it, refer to [5]. Many related problems can also be reduced to the problem under consideration. Our algorithm can also be considered as method to compute the Gröbner basis of some zero-dimensional ideal in the ring of multivariate polynomial.

Author Contributions: Conceptualization and writing, Alexandr V. Seliverstov. The author has read and agreed to the published version of the manuscript.

Funding: The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation for IITP RAS.

Data Availability Statement: No new data were created or analyzed during this study. Data sharing is not applicable.

Conflicts of Interest: The author declares no conflict of interest.

Acknowledgments: We thank the conference organizers.

References

[1] G. Alonistiotis, A. Antonopoulos, N. Melissinos, A. Pagourtzis, S. Petsalakis, M. Vasilakis, Approximating subset sum ratio via partition computations, Acta Informatica 61 (2024) 101–113. doi:10.1007/s00236-023-00451-7.

[2] A. A. Boykov, A. V. Seliverstov, On a cube and subspace projections, Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki 33 (2023) 402–415. doi:10.35634/vm230302.

[3] V. Cacchiani, M. Iori, A. Locatelli, S. Martello, Knapsack problems — an overview of recent advances. part II: Multiple, multidimensional, and quadratic knapsack problems, Computers and Operations Research 143 (2022) 1–14. doi:10.1016/j.cor.2021.105693.

[4] Y. G. Evtushenko, A. A. Tret'yakov, Exact formula for solving a degenerate system of quadratic equations, Computational Mathematics and Mathematical Physics 64 (2024) 365–369. doi:10.1134/S0965542524030072.

[5] Y. Pan, F. Zhang, Solving low-density multiple subset sum problems with SVP oracle, Journal of Systems Science and Complexity 29 (2016) 228–242. doi:10.1007/s11424-015-3324-9.

[6] A. N. Rybalov, Generic polynomial algorithms for the knapsack problem in some matrix semigroups, Siberian Electronic Mathematical Reports 20 (2023) 100–109. URL: http://semr.math.nsc.ru/v20/n1/p100-109.pdf. doi:10.33048/semi.2023.20.009.

[7] J. T. Schwartz, Fast probabilistic algorithms for verification of polynomial identities, J. ACM 27 (1980) 701–717. doi:10.1145/322217.322225.

[8] L. Zhang, M. Quweider, F. Khan, H. Lei, Splitting NP-complete sets infinitely, Information Processing Letters 186 (2024) 1–7. doi:doi.org/10.1016/j.ipl.2024.106472.

[9] O. A. Zverkov, A. V. Seliverstov, Effective lower bounds on the matrix rank and their applications, Programming and Computer Software 49 (2023) 441–447. doi:10.1134/S0361768823020160.

On the Summation of Fourier Series over the Roots of Transcendental Equations Using Annihilation

Elizaveta D. Volchkova[1], Ksaverii Yu. Malyshev[1,2,*]

[1]RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
[2]Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, 1 Leninskiye Gory St, Moscow, 119991, Russian Federation

Abstract
Trigonometric series in the roots of transcendental equations are considered. Similar series arise in problems of mathematical physics in the case of boundary conditions of the third kind. For this class of series, sufficient conditions for their representability in the finite form of elementary functions are indicated. We generalize to this case our previous works based on the use of finite expressions for divergent Fourier series. In this case, we obtain a simple method for determining the elementarity of the ζ-function of N.S. Koshlyakov and its generalizations. These considerations are used to create new functions of symbolic summation in computer algebra systems.

Keywords
series summation, roots of transcendental equations, divergent series, computer algebra

1. Introduction

Solving problems in mathematical physics requires working with series in eigenfunctions of differential operators [9]. The simplest example is trigonometric Fourier series, located along the sines and cosines of multiple arcs. In this case, the natural number is both the summation index (harmonic number) and the square root of the eigenvalue of the differential operator. However, a situation often arises when the eigenvalue is not the square of a natural number, but is determined from some transcendental equation, although the series remains trigonometric [6], [7]. A. N. Krylov called them as "series similar to Fourier series" [2]. To this class of series, under certain conditions on the Fourier coefficients, one can apply Krylov's scheme for accelerating convergence, which requires the finite expressions of some "remarkable" series. The said "remarkable" series must be summed in a finite form, and this form becomes dependent on the transcendental equation that determines the eigenvalue. We propose a simple way to obtain these expressions in computer algebra systems that also allows summation in finite form of an important class of Fourier series. This requires generalizing our previous work [5] on ordinary Fourier series.

2. Problem formulation

For simplicity, we consider the Fourier series of the eigenfunctions of the problem with the Dirichlet condition at the right end and the Robin condition (3-rd kind) at the second end:

$$y'' + \lambda^2 y = 0, \quad x \in (0,1)$$

$$y(0) = 0, \quad y'(1) + hy(1) = 0, \quad h > 0.$$

6th International Conference "Computer Algebra", Moscow, June 23–25, 2025
10.22363/12585-2025-6-023
EDN: FLROZO
*Corresponding author.
liza.volchkova.s7@gmail.com (E. D. Volchkova); kmalyshev08102@mail.ru (K. Yu. Malyshev)
0000-0001-8823-9136 (K. Yu. Malyshev)

Here h is a fixed parameter of the Robin condition. Eigenfunctions of this Sturm–Liouville problem are:

$$y_\lambda = \sin \lambda x.$$

The eigenvalues are the positive roots of the transcendental equation:

$$\lambda \cos \lambda + h \sin \lambda = 0.$$

We consider the infinite sums over all positive roots of this equation. The square of the norm of an eigenfunction has the form

$$\|y_\lambda\|^2_{L_2(0,1)} = \frac{1}{2}\frac{h^2 + h + \lambda^2}{h^2 + \lambda^2}.$$

It was for this spectral problem that N.S. Koshliakov constructed a generalization of Bernoulli polynomials and obtained their generating function in the work [1].

Let us pose the problem of finding the sum of the Fourier series for the specified eigenfunctions, provided that the expansion coefficients are rational functions of the eigenvalue (cf. [5], [3]), i.e. the problem of calculating in final form a sum

$$u = \sum_\lambda a_\lambda \frac{h^2 + \lambda^2}{h^2 + h + \lambda^2} \sin \lambda x, \qquad x \in [0, 1], \qquad a_\lambda \in \mathbb{R}(\lambda).$$

3. Summation using annihilation

By analogy with work [5], it is not difficult to construct a differential expression L, the formal application of which to the Fourier series under consideration yield a new series with polynomial coefficients. A differential equation for u has the form

$$Lu = \sum_\lambda A_\lambda \frac{h^2 + \lambda^2}{h^2 + h + \lambda^2} \sin \lambda x, \qquad x \in [0, 1], \qquad A_\lambda \in \mathbb{R}[\lambda].$$

The right side contains a divergent series. It can be shown that if the polynomial $A(\lambda)$ resulting from the annihilation is an odd function, then the right side of the resulting equation can be expressed through derivatives of the Dirac δ-function. Then the desired sum is expressed through the corresponding Green function. Thus, it is easy to prove analogs of Theorems 7 and 8 from the work [5]. Note that in this way it is possible to obtain expressions for Koshliakov [1] polynomials of odd degree.

Note that simultaneously with the considered sine series it would be natural to consider cosine series, however this generalization is not so trivial due to the boundary conditions that functions $\cos \lambda x$ satisfy.

At present, we do not know whether this reasoning can be generalized to the case of an arbitrary polynomial $A(\lambda)$ that does not satisfy the oddity condition, thereby obtaining at least an analogue of Theorem 9 [5] without involving Koshliakov's transcendental functions. For example, we do not know the final expression for the series

$$\sum_\lambda \frac{h^2 + \lambda^2}{h^2 + h + \lambda^2} \sin \lambda x, \quad x \in [0, 1].$$

At $h = 0$, this series represents the generalized function [8] $\widehat{\cot}$.

4. Concluding remarks

We present a generalization of our works on the summation of trigonometric series to new systems of eigenfunctions. A significant part of the results can be transferred to new classes of Fourier series. Note that the presented method of summation can be used, for example, in the summation problems posed in the work [7]. The described considerations can easily be transformed into new functions for our "Kryloff for Sage" package [4]. Note that not all previous results can be transferred even to the simplest case of trigonometric series by the roots of transcendental equations. Probably, the gaps in the summation that arise can be eliminated by using other methods. This issue be considered further.

Author Contributions: Conceptualization—K.M., writing—review and editing, K.M and E.V.; methodology, writing—original draft preparation, E.V and K.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the RUDN University Strategic Academic Leadership Program, project No. 021934-0-000.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: The authors thanks to M.D. Malykh and L.A. Sevastyanov and A.V. Seliverstov for valuable discussions.

References

[1] N. S. Koshliakov, An extension of Bernoulli's polynomials, Matematicheskii Sbornik 42 (1935) 425–434. URL: https://www.mathnet.ru/eng/sm/v42/i4/p425.

[2] A. N. Krylov, On some differential equations of mathematical physics, GITTL, Moscow-Leningrad, 1950.

[3] M. D. Malykh, K. Y. Malyshev, On summation of Fourier series in finite form, Discrete and Continuous Models and Applied Computational Science 32 (2024) 406–413. URL: https://journals.rudn.ru/miph/article/view/43669. doi:10.22363/2658-4670-2024-32-4-406-413.

[4] K. Y. Malyshev, M. D. Malykh, Implementation of A.N. Krylov series convergence acceleration in the CAS Sage, International Conference Polynomial Computer Algebra'2023 (2023) 78–79. URL: https://pca-pdmi.ru/2023/program.

[5] K. Y. Malyshev, M. D. Malykh, L. A. Sevastyanov, A. V. Zorin, On summation of Fourier series in finite form using generalized functions, Mathematics 13(3) (2025) 538. doi:10.3390/math13030538.

[6] F. O. Naydyuk, V. L. Pryadiev, S. M. Sitnik, Laguerre polynomials in the forward and backward wave profile description for the wave equation on an interval with the Robin condition or the attached mass condition, Math. Notes 115 (2024) 789–799. doi:10.1134/S0001434624050146.

[7] V. D. Repnikov, Y. P. Huhryanskii, E. V. Bordakov, On some function series containing roots of a transcendental equation, Differ. Equ. 32(4) (1996) 575–577. URL: https://geodesic-test.mathdoc.fr/item/DE_1996_32_4_a19/.

[8] L. Schwartz, Méthodes mathématiques pour les sciences physiques, Hermann, Paris, 1961.

[9] A. G. Sveshnikov, A. N. Bogolyubov, V. V. Kravtsov, Lectures on Mathematical Physics, Nauka, Moscow, 2004.

On the Determination of the Volumetric Heat Capacity and the Thermal Conductivity of a Substance in the Three-Dimensional Case

Vladimir Zubov

Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, Russia

Abstract

The problem of simultaneous identification the temperature-dependent thermal conductivity and volumetric heat capacity of the investigated substance in the three-dimensional case is considered. The consideration is based on the first boundary value problem for a three-dimensional unsteady heat equation. The inverse coefficients problem was reduced to a variational problem. The optimization problem was solved using gradient methods of functional minimization. Based on the fast automatic differentiation technique exact values of the cost functional gradient components were obtained.

Keywords

thermal conductivity, inverse coefficient problems, three-dimensional unsteady heat equation, fast automatic differentiation

1. Statement of the problem

When solving optimal control problems of complex dynamic processes, gradient methods of minimizing the cost function are often used. If the gradient of the functional is calculated approximately, then it is often impossible to obtain a solution to the optimization problem. In [2], an approach (fast automatic differentiation technique) was proposed that makes it possible to determine the gradient of a cost function with machine precision. Using this approach has allowed us to solve a number of theoretically interesting and practically important problems.

One of such problem is the problem of identifying the volumetric heat capacity and thermal conductivity of a substance based on the results of observing the dynamics of the temperature field. The results of the study of this problem in the one-dimensional spatial case are presented in [1]. The present paper provides a solution to the same problem in the three-dimensional case.

A sample of the test substance in the form of a straight parallelepiped of length X, width Y and height Z is considered. The points of the parallelepiped form the domain $Q = \{(0, X) \times (0, Y) \times (0, Z)\}$ with border $\Gamma = \partial Q$. The distribution of the temperature field in the parallelepiped at each moment of time is determined by solving the following direct problem:

$$C(T(s,t))\frac{\partial T(s,t)}{\partial t} = div_s(K(T(s,t))\nabla_s T(s,t)), \qquad (s,t) \in \{Q \times (0,\Theta]\}, \tag{1}$$

$$T(s,0) = w_0(s), \qquad s \in \overline{Q}, \tag{2}$$

$$T(s,t) = w_\Gamma(s,t), \qquad s \in \Gamma, \quad 0 \le t \le \Theta. \tag{3}$$

Here s is the point of the parallelepiped with Cartesian coordinates (x, y, z); t is the time; $T(s,t)$ is the temperature of the substance at the point s at the moment of time t; $C(T)$ is the volumetric

6^{th} *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-024
EDN: FMRNAJ
vladimir.zubov@mail.ru (V. Zubov)
0000-0001-5207-8414 (V. Zubov)

heat capacity of the substance, $K(T)$ is the thermal conductivity; $w_0(s)$ is the temperature set of the parallelepiped at the initial moment of time; $w_\Gamma(s,t)$ is the temperature set at the boundary of the domain.

One of possible formulations of the problem of identifying model parameters is the following: to find a dependence of the volumetric heat capacity of a substance and the thermal conductivity on the temperature at which the temperature field $T(s,t)$ obtained as a result of solving the formulated direct problem (1)-(3) differs little from the field $P(s,t)$, obtained experimentally. A measure of the deviation of these functions can be the value

$$\Phi(K(T),C(T)) = \int_0^\Theta \int_Q \mu(s,t)\cdot[T(s,t)-P(s,t)]^2\,dsdt, \tag{4}$$

where $\mu(s,t)\geq 0$ is the specified weight function. Thus, the optimal control problem consists in determining the optimal control and the corresponding optimal solution $T(s,t)$ of the problem (1)-(3), at which the functional (4) reaches the minimum value.

It should be noted that in the presented formulation, the optimal control problem always has a non-unique solution. Indeed, if the solution $\{C^*(T),K^*(T)\}$ to the formulated inverse problem is found, then for any real number λ, the control $\{\lambda C^*(T),\lambda K^*(T)\}$ is also the solution to this problem. To identify the unique solution to the formulated identification problem, it is necessary to use an additional condition. Such a condition can be, for example, a condition $K(T_*)=K_*$ with the specified numbers T_* and K_*. You can also define a relationship $\chi(T)=\frac{C(T)}{K(T)}$ when solving a problem.

2. On the numerical solution to the inverse problem

To solve the problem numerically, we introduce a time grid $\{t^j\}_{j=0}^J$, $t^0=0$, $t^J=\Theta$, and a spatial grid $\{x_n\}_{n=0}^N$, $x_0=0$, $x_N=X$, $\{y_i\}_{i=0}^I$, $y_0=0$, $y_I=Y$, $\{z_l\}_{l=0}^L$, $z_0=0$, $z_L=Z$.

The segment $[a,b]$ on which the functions $K(T)$ and $C(T)$ are restored was defined as the set of values of the given functions $w_0(s)$ and $w_\Gamma(s,t)$. This segment was divided by the points $\widetilde{T}_0=a$, $\widetilde{T}_1,\widetilde{T}_2,\ldots,\widetilde{T}_M=b$ into M parts. The numbers $c_m=C(\widetilde{T}_m)$ and $k_m=K(\widetilde{T}_m)$ were assigned to each of the points $\widetilde{T}_m$ $(m=0,\ldots,M)$. The desired functions $K(T)$ and $C(T)$ were approximated by continuous piecewise linear functions with support nodes at points $\left\{(\widetilde{T}_m,k_m)\right\}_{m=0}^M$ and $\left\{(\widetilde{T}_m,c_m)\right\}_{m=0}^M$.

The algorithm for the numerical solution of the direct problem is based on the use of the heat balance equation for the computational cell, which consists in the fact that the change in the heat content of a substance in a volume V_{nil} over a fixed period of time $\tau^j=t^{j+1}-t^j$ is equal to the amount of heat passing through the surface S_{nil} of the volume V_{nil} over the same period of time. A locally one-dimensional scheme was used to approximate equation (1) in time.

The fast automatic differentiation technique made it possible to formally construct a system of conjugate equations for the used approximation of the optimal control problem. The components of cost function gradient with respect to vectors $(k_0,k_1,\ldots,k_M)$ and $(c_0,c_1,\ldots,c_M)$ were also determined in accordance with the fast automatic differentiation technique. Note that the obtained values of the gradient components are exact for the chosen approximation of the optimal control problem.

To verify the performance of the proposed algorithm, we solved a large number of test problems. Some of the results obtained are presented here.

The locally one-dimensional scheme chosen in this work for approximating the heat equation is stable, so the time step can be fairly large. Nevertheless, research concerning the choice of a time grid has to be performed for each spatial grid used. In the numerical experiments below the domain $Q\times(0,\Theta)=(0,1)\times(0,1)\times(0,1)\times(0,1)$, $N=I=L=30$, $J=100$, the segment

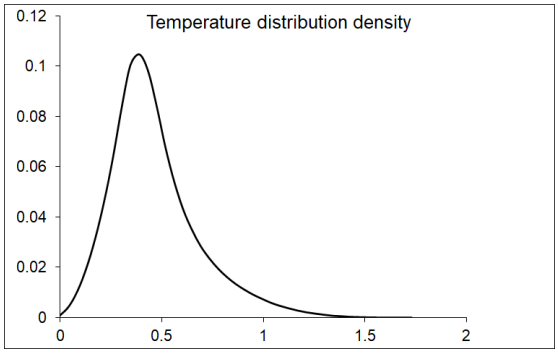


Figure 1: Distribution of the function $W(T)$.

$[a, b]$ was divided into 40 intervals, i.e. $M = 40$. It was assumed that in the cost functional the weight function is $\mu(x, y, z, t) \equiv 1$, $(x, y, z, t) \in Q \times (0, \Theta)$.

The performed numerical experiments showed that the quality of the recovered thermal conductivity and volumetric heat capacity depend strongly on the distribution of the "experimental" temperature field. It may happen that some subintervals of $[a, b]$ contain data insufficient to identify the coefficients. Therefore, the distribution of experimental data on subintervals of the temperature interval of interest has to be analyzed in each particular case.

In **the first series of calculations**, numerical experiments were based on the function $\Lambda(x, y, z, t) = x + y + z + 3t + 0.5$. This function is the solution of equation (1) for $C(T) \equiv 1$ and $K(T) = T$. The traces of function $\Lambda(x, y, z, t)$ on the parabolic boundary of the domain under consideration were chosen as the boundary values (see (2)-(3)). The "analytical" field was used as the experimental temperature field: $P^j_{nil} = \Lambda(x_n, y_i, z_l, t^j) = x_n + y_i + z_l + 3t^j + 0.5$. The temperature in the closure of the domain $Q \times (0, \Theta)$ varies from $a = 0.5$ to $b = 6.5$.

To obtain a unique solution to the inverse problem when performing calculations, at the point $T_* = a = 0.5$ neither the thermal conductivity ($K(0.5) = 0.5$) nor the volumetric heat capacity ($C(0.5) = 1.0$) changed. As a result of the minimization, the cost functional decreased from 3.39×10^{-03} at the initial approximation ($C(T) = T$ for $T \neq a$, C(a)=1.0; $K(T) = 4.5$ for $T \neq a$, $K(a) = a$) to 9.35×10^{-30}, the maximum relative deviation of the obtained temperature field from the "experimental" field changed from 3.46×10^{-02} to 2.65×10^{-14}, the modulus of the gradient of the cost function decreased by 15 orders of magnitude. The thermal conductivity and volumetric heat capacity in this case are restored with almost machine precision.

The goal of the first series of runs was not only to check the performance of the proposed algorithm, but also to verify the software codes implementing this algorithm.

In **the second series of calculations**, the "experimental" temperature field $P(x, y, z, t)$ was constructed by solving the direct problem (1)-(3) with special input data: the temperature at the points of the parabolic boundary of the domain was equal to the trace at these points of the function $\Lambda(x, y, z, t) = \sqrt{\frac{x^2+y^2+z^2}{9-8t}}$; the parameters of substance were set by the equalities $C(T) = T$ and $K(T) = T^2$.

The analysis of the temperature field $P(x, y, z, t)$ made it possible to determine the range of temperature variation ($a = 0.0$,$b = 1.732$), and to obtain the distribution of experimental data along the temperature segment $[0.0, 1.732]$ (figure 1). Figure 1 shows that there are too few experimental data outside the segment $[0.05, 1.2]$. Therefore, it can be assumed that there will be difficulties with restoring the sought parameters $C(T)$ and $K(T)$.

When performing calculations, at the point $T_* = \frac{a+b}{2}$ neither the thermal conductivity ($K(T_*) = (T_*)^2$) nor the volumetric heat capacity ($C(T_*) = T$) changed. The initial distributions of the sought functions are shown in figures 2, 3 by dashed lines. As a result of minimization, the cost functional decreased from 2.19×10^{-4} at the initial approximation to 4.41×10^{-12},

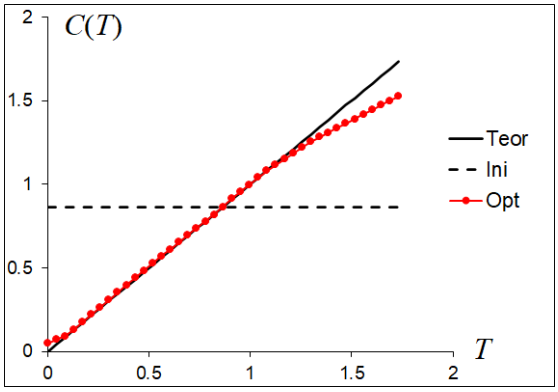


Figure 2: Volumetric heat capacity $C(T)$.

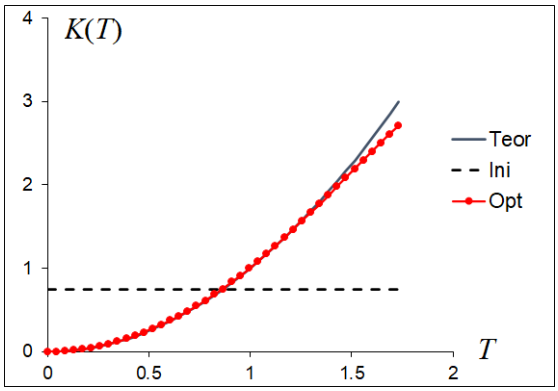


Figure 3: Thermal conductivity $K(T)$.

the modulus of the cost function gradient decreased by 7 orders of magnitude. The obtained distributions of the functions $C(T)$ and $K(T)$ are marked in figures 2, 3 by the symbol "Opt".

As expected, the functions $C(T)$ and $K(T)$ are not restored where there is too little experimental data, i.e. outside the interval $[0.05, 1.2]$.

References

[1] A. F. Albu, A. Y. Gorchakov, V. I. Zubov, On the simultaneous identification of the volumetric heat capacity and the thermal conductivity of a substance, in: Optimization and Applications: 13th International Conference, OPTIMA 2022, Petrovac, Montenegro, September 26–30, 2022, Revised Selected Papers, Springer-Verlag, Berlin, Heidelberg, 2022, pp. 207–220. doi:10.1007/978-3-031-22543-7_15.

[2] Y. G. Evtushenko, Computation of exact gradients in distributed dynamic systems, Optimization Methods and Software 9 (1998) 45–75. doi:10.1080/10556789808805686.

Author index