Russian Academy of Sciences
St.Petersburg Department
of Steklov Mathematical Institute
Euler International Mathematical Institute
St.Petersburg Electrotechnical University "LETI"

International Conference

Polynomial Computer Algebra

International Conference on Polynomial Computer Algebra

St. Petersburg, October, 2020

Санкт Петербург
2020

ISBN 978-5-9651-0568-7

Polynomial Computer Algebra '2020
St. Petersburg, Russia
October 12–17, 2020
International Euler Institute

International Conference Polynomial Computer Algebra '2020; St. Petersburg, October 12–17 202, Euler International Mathematical Institute, Ed. by N. N. Vassiliev, SPb.: VVM Publishing, 2020, 200 p.

The book contains short papers, extended abstracts and abstracts of reports presented at the International Conference on Polynomial Computer Algebra 2020, St. Petersburg, October 2020

Supported by Simons Foundation and Russian Foundation for Basic Research (grant 20-01-20034)

International Conference

Polynomial Computer Algebra '2020'

St. Petersburg, Russia, October 12-17, 2020
International Euler Institute

Organizers:
Steklov Institute of Mathematics at St.Petersburg
Joint Institute of Nuclear Research (Dubna)
Saint Petersburg Electrotechnical University

Conference chair: Nikolay Vassiliev

Organizing committee: N. Vassiliev, E. Novikova, T.Vinogradova(secretary), V.Duzhin, Ya.Shibaeva, N.Zalesskaya (secretary), N.Kirshner, S.Pozdnyakov, M.Rybalkin

Scientific committee: B.Buchberger (Schloss Hagenberg, Austria) V. Edneral (Moscow, Russia), V.Gerdt (Dubna, Russia), P.Gianni (Pisa, Italy), D.Grigoriev (Lille, France), A.Khovansky (Toronto,Canada), I.Kotsireas (Wilfrid Laurier University Waterloo, Canada), Yu. Matiyasevich (St. Petersburg, Russia), E.Mayr (Munich, Germany), A.Myllari (Grenada), V.Pan (New York, USA), L.Robbiano (Genova, Italy), D. Stefanescu (Bukharest, Romania), N.Vassiliev (St. Petersburg, Russia), A. Vershik (St. Petersburg, Russia), N.Vavilov (St.Petersburg, Russia), S.Watt (University of Waterloo, Canada)

Topics:

Gröbner bases
Combinatorics of monomial orderings
Differential bases
Involutive algorithms
Computational algebraic geometry
Computational topology
D-modules
Polynomial differential operators
Parallelization of algorithms
Algorithms of tropical mathematics
Quantum computing
Cryptography
Tropical manifolds
Matrix algorithms
Complexity of algorithms

Table of content

On the separability probability of rank-deficient two qubit and qubit-qutrit states

V. Abgaryan, A. Khvedelidze, I. Rogojin and A. Torosyan

We are planning on presenting numerical analysis of separability probability of the rank-deficient random states of qubit-qubit and qubit-qutrit pairs from the so-called Hilbert-Schmidt ensemble. With this aim two methods of generation of a random low rank states of finite-dimensional quantum systems will be formulated and applied to the separability problem. First, we describe a direct method of generation of ensemble of random density matrices exploiting the conditional probability density function. Since this method becomes very cumbersome for composite systems larger than a pair of qubits, we elaborate an alternative method of generation. The latter is based on the recently obtained representation for the distribution functions of eigenvalues of density matrices of the rank-deficient states belonging to the boundary of state space in the form of a special Wishart-Laguerre distribution.

V. Abgaryan
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia
e-mail: vahagnab@googlemail.com

A. Khvedelidze
A Razmadze Mathematical Institute
Iv. Javakhishvili, Tbilisi State University
Tbilisi, Georgia
Institute of Quantum Physics and Engineering Technologies
Georgian Technical University
Tbilisi, Georgia
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia
e-mail: akhved@jinr.ru

I. Rogojin
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia

A. Torosyan
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia
e-mail: astghik@jinr.ru

Rigid body motion symmetries

Semjon Adlaj

Abstract. The Galois axis, which constitutes an axis of "generalized" symmetry of a rigid body, is acted upon by the Klein four-group, generated by reflections across the principal axes, corresponding to extreme values of moments of inertia (minimal and maximal). In particular, the direction of the Galois axis might be reversed, prompting a remarkable duality of critical rigid body motion. Such reversal of direction of the Galois axis might be viewed as a composition of two distinct reflections, each of which would correspond to a time reversal symmetry.

The modular invariant as a symmetric function of the squares of the three principal moments of inertia

A torque free rigid body motion is governed by equations, possessing time and "mirror" symmetries. Several aspects of these symmetries were explored in [2, 3, 4, 5, 6, 12]. In [8], the (three) projections p, q, r of the angular velocity $\boldsymbol{\omega}$ upon the (three) principal axes of inertia, with corresponding moments of inertia A, B and C, are calculated via the Galois alternative elliptic function with corresponding elliptic moduli k_1, k_2 and k_3. The alternating group A_3 was shown to act on the squares k_1^2, k_2^2 and k_3^2 of the (three) elliptic moduli via the transformation $\tau : x \mapsto 1 - 1/x$, as further discussed in [9, 10]. Such a transformation (of order 3) would correspond to a cyclic permutation of the principal moments of inertia $(A\,B\,C)$, that is, putting

$$k^2(A, B, C) = \frac{(A - B)(Ch - m^2)}{(A - C)(Bh - m^2)},$$

with $k_1^2 = k^2(A, B, C)$, we must have $k_2^2 = k^2(B, C, A)$ and $k_3^2 = k^2(C, A, B)$, where the constants h and m^2 represent twice the kinetic energy and the square of the modulus of the angular momentum, respectively. The three elliptic moduli would, of course, correspond to one and the same value of the modular (Klein)

invariant:

$$j = \frac{4(k^2 + 1/k^2 - 1)^3}{27(k^2 + 1/k^2 - 2)} = \frac{4(1 - k_1^2 - 1/k_1^2)(1 - k_2^2 - 1/k_2^2)(1 - k_3^2 - 1/k_3^2)}{27} =$$

$$= 1 + \frac{4(k_1^2 + k_2^2 + k_3^2 - 3/2)^2}{27},$$

which is a symmetric function in the moments of inertia. It vanishes when k^2 coincides with a primitive cube root of -1, that is, a fixed point of the transformation τ. Another special value $j = 1$ is attained with

$$m^2 = \frac{C(A + B) - 2AB}{2C - A - B}\, h,\ Bh - m^2 = \frac{(C - B)(B - A)h}{2C - A - B} \geqslant 0,\ A \leqslant B < C.$$

The rate of precession of a freely rotating rigid body as a symmetric function of its three principal moments of inertia

A formula for the rate of precession, about the (fixed) angular momentum $\boldsymbol{m}$, symmetric in the moments of inertia:

$$\dot{\psi} = \frac{1}{m}\left(h + \frac{(h - m^2/A)(h - m^2/B)(h - m^2/C)}{m^2\omega^2 - h^2}\right),$$

where ω is the angular speed, was presented four years ago at the PCA 2016 conference [13]. The formula demonstrates that the rate of precession is uniform not only for permanent rotations about the axes, corresponding to extreme values of the moments of inertia, but it is also uniform for "permanent rotations" about the axis, corresponding to the intermediate value of the principal moments of inertia, which we label with the letter B. The last quote was taken to emphasize that the uniformity of the rate of precession does not necessarily imply a "usual permanent rotation" but it (the uniformity) holds, as well, for all (critical) solutions, satisfying $h = m^2/B$. Yet, the (improper) integral

$$\int_{-\infty}^{\infty}\left(\dot{\psi} - \frac{h}{m}\right)dt =: 2\theta(A, B, C),$$

taken with respect to time, throughout the critical motion does not vanish! For a triaxial rigid body the multivalued function $\theta(A, B, C)$ represents an angle between the Galois axis and a principal axis, corresponding to an extreme value of the moments of inertia. In fact,

$$\theta(A, B, C) = \sigma\operatorname{Arctan}\left(\sqrt{\frac{A(B - C)}{C(A - B)}}\right),\ \sigma = \begin{cases} 1, & \text{if } A < B < C, \\ -1, & \text{if } A > B > C, \end{cases}^1$$

and the sum $\theta(A, B, C) + \theta(C, B, A)$ matches (modulo π) the angle $\pi/2$. A reflection of the Galois axis across a principal axis, corresponding to an extreme

[1] The sign flip reflects two distinct orientations of a coordinate system, fixed within a rigid body. In particular, an orientation of a coordinate system is changed with reversing the direction of the principal axis, corresponding to the intermediate moment of inertia.

moment of inertia, would rotate it (about the center of mass) by the angle 2θ, as calculated. Thus, such a reflection of the Galois axis corresponds not only to a time reversal but to a "flip" of the intermediate axis of inertia during the critical rigid body motion [11]. The said "mirror" reflection of the Galois axis might also be viewed as a reversal of a given orientation of a body-fixed coordinate system. A right-handed coordinate system is necessarily transformed into a left-handed system and vice versa. Such an observation is relevant to applications when the determination of the orientation must remain error free. A geometric interpretation of the "duality" of the Galois axis was well explained in [15], whereas an advice for correctly determining the (multivalued) angle θ was suggested by E. Mityushov in an email message sent to the author on February 27, 2018. It is based on an observation, concerning the Galois modulus

$$G(A,B,C) := \frac{C(B-A)}{B(C-A)},$$

which, upon ascendingly ordering the principal moments of inertia $A \leqslant B < C$, coincides with the square of the scalar product of two unit vectors, directed (respectively) along the Galois axis and the principal axis, corresponding to the minimal value of inertia (labeled with the letter A). The (ascending) ordering condition turns out being superfluous if one adopts an "invariant" way for defining the Galois modulus as a square of a scalar product of the said unit vectors, where the second is now aligned along a principal axis, corresponding to an extreme value of the moments of inertia, which need not necessarily be minimal. With this simplified (and thus improved) definition, the Galois modulus is seen to be the square of the scalar product of a unit vector along the Galois axis with a unit vector along a principal axis, corresponding to an extreme moment of inertia which we might still label with the letter A (but without further requiring A to be minimal).[2]

As emphasized in [7], the MacCullagh ellipsoid of inertia would "degenerate" to an ellipsoid of revolution whenever two of the principal moments would coincide one with the other. The Galois axis, being an axis orthogonal to the circular sections of such an ellipsoid would then coincide with the axis of (dynamical) symmetry.

Conclusion

The (tip of the) angular momentum pseudovector $\boldsymbol{m}$ might be viewed as a holomorphic (orientation preserving) function, mapping time to a fixed within a rigid body (unit) sphere. We must furthermore distinguish a "right" pseudovector, which coordinates are given with respect to a right-handed body-fixed coordinate system from its "mirror" reflection, that is, a "left" pseudovector which coordinates are given with respect to a (reflected) left-handed body-fixed coordinate system. Establishing such a distinction would

[2] A clarifying formula would be $G(A,B,C) + G(C,B,A) = 1$.

protect us from a conventional (yet erroneous) reversal of the direction of the reflected pseudovector. In other words, a given "parity" of a pseudovector (right or left) cannot be altered by reversing its (preserved) direction, so we risk no confusion since we never alter a given pseudovector aside from subjecting it to transfomations which might (or might not) preserve its parity. Then, the pseudovector, obtained by so reflecting the initial (right) pseudovector $\boldsymbol{m}$ might be regarded as an antiholomorphic (orientation reversing) map. Ignoring such an elementary observation has apparently precluded the practical implementation of exact rigid body motion solution, commonly substituting it with numerical approximations. The duality rather than uniqueness of rigid body critical motion had prompted D. Abrarov to declare, in [1], "the Dzhanibekov's top" as the "classical analogue" of the sought after (in quantum field theory) "massless particle, possessing spin 2", that is, "the graviton"! Dzhanibekov's top appears on a diagonal of the animation in [14]. A "dual top", corresponding to an opposite (initial) rotation appears on the other diagonal. Each of the Galois axes, corresponding to these dual tops are reflections, one of the other across a "mirror", orthogonal to a principal axis, corresponding to one of the two extreme values of the moments of inertia.

This work was partially supported by the Russian Foundation for Basic Research (Project № 19-29-14141).

References

[1] Abrarov D.L. *The exact solvability of model problems of classical mechanics in global L-functions and its mechanical and physical meaning* // International Conference on Mathematical Control Theory and Mechanics, Suzdal', Russia, July 7-11, 2017. Available at `https://www.youtube.com/watch?v=oUGIxkb5gD8&list=PL8StSu9q0Yd7yoOX6oBZ7PDSSLgaT-JzA&index=2`

[2] Adlaj S. *Mirror symmetry in classical mechanics.* International scientific conference "Fundamental and Applied Problems of Mechanics" (FAPM-2019), Moscow, Russia, December 10-12, 2019.

[3] Adlaj S. *Mechanical interpretation and efficient computation of elliptic integrals of the third kind.* The joint MSU-CCRAS Computer Algebra Seminar, Moscow, Russia, November 27, 2019. Available at `http://www.ccas.ru/sabramov/seminar/lib/exe/fetch.php?media=adlaj191127.pdf`

[4] Adlaj S. Lamarche F. *Complex periods, time reversibility and duality in classical mechanics.* Russian Interdisciplinary Temporology Seminar, Moscow, Russia, November 26, 2019. Available at `http://www.chronos.msu.ru/ru/mediatek/video-sem/2019/zasedanie-seminara-26-noyabrya-2019-g` (in Russian)

[5] Adlaj S. *An arithmetic-geometric mean of a third kind!* Lecture Notes in Computer Science, volume 11661: 37-56. Presented on August 30, 2019 at the 21st International Workshop on Computer Algebra in Scientific Computing, Moscow, Russia. Available at `http://semjonadlaj.com/Computer+Algebra+in+Scientific+Computing_37-56.pdf`

[6] Adlaj S. *Modular equations and fundamental problems of classical mechanics.* Available at https://www.youtube.com/playlist?list=PL8StSu9qOYd6vYHtDiS7JNrCbF8CgjfdF (in Russian)

[7] Adlaj S. *The Galois axis.* International scientific conference "Infinite-dimensional analysis and mathematical physics", Moscow, Russia, January 28 - February 1, 2019. Available at http://semjonadlaj.com/GaloisAxis190129.pdf

[8] Adlaj S. *Torque free motion of a rigid body: from Feynman wobbling plate to Dzhanibekov flipping wingnut.* Available at http://semjonadlaj.com/TFRBM.pdf

[9] Adlaj S. *Multiplication and division on elliptic curves, torsion points and roots of modular equations.* Available at http://www.pdmi.ras.ru/znsl/2019/v485.html

[10] Adlaj S. *On the Second Memoir of Évariste Galois' Last Letter.* Computer Tools in Science and Education, 2018 (4): 11–26. Available at http://cte.eltech.ru/ojs/index.php/kio/article/view/1544/1516

[11] Adlaj S. Berestova S. Misyura N. Mityushov E. *Illustrations of Rigid Body Motion Along a Separatrix in the Case of Euler-Poinsot.* Computer Tools in Science and Education, 2018 (2): 5–13. Available at http://ipo.spb.ru/journal/index.php?article/2025/ (in Russian)

[12] Adlaj S. *Dzhanibekov screw.* Available at http://semjonadlaj.com/SScrew.pdf (in Russian)

[13] Adlaj S. *Dzhanibekov's flipping nut and Feynman's wobbling plate.* In: Vassiliev, N.N. (ed.) 9th International Conference on Polynomial Computer Algebra, pp. 10–14. St. Petersburg department of Steklov Institute of Mathematics (2016). Available at http://pca.pdmi.ras.ru/2016/abstracts_files/PCA2016SA.pdf

[14] Misyura N. Mityushov E. Computer animation two dual "Dzhanibekov's tops" (posted on February 24, 2018). Available at https://youtu.be/c0m_yeKeCiQ

[15] Seliverstov A. *Note on circular sections.* Available at http://iitp.ru/upload/publications/8014/PlaneSection12.pdf (in Russian)

Semjon Adlaj
Division of Complex Physical and Technical Systems Modeling
Federal Research Center "Informatics and Control" of the Russian Academy of Sciences
Russia 119333, Moscow, Vavilov Street 40.
e-mail: SemjonAdlaj@gmail.com

Hermitian Finite Elements for Hypercube

Galmandakh Chuluunbaatar, Alexander Gusev, Vladimir Gerdt, Sergue Vinitsky and Lyong Le Hai

Abstract. Algorithm for analytical construction of multivariate Hermite interpolation polynomials in a multidimensional hypercube is presented. In the case of a d-dimensional cube, the basis functions are determined by products of d Hermite interpolation polynomials depending on each of the d variables given explicitly in the analytical form. The efficiency of finite element schemes, algorithms and programs is demonstrated by benchmark calculations of the 4D Helmholtz problem.

Introduction

In this paper we present a new symbolic algorithm implemented in Maple for constructing the Hermitian finite elements or piece-wise multivariate Birkhoff interpolants in a standard d-dimensional cube that generalizes the construction and algorithm proposed for a three and four dimensional cube [1, 2, 3, 4]. Our algorithm realizes recurrence relations [5, 6] and yields explicit expressions in an analytical form for the Hermite interpolation polynomials (HIPs) in opposite the conventional constructions. The basis functions of finite elements are high-order polynomials, determined from a specially constructed set of values of the polynomials themselves and their partial derivatives up to a given order at the vertices of the hypercube. Such a choice of values allows us to construct a piecewise polynomial basis continuous at the boundaries of finite elements together with the derivatives up to a given order. In the case of a d- dimensional cube, it is shown that the basis functions are determined by products of d one-dimensional HIPs depending on each of the d variables given in the analytical form with the derivatives up to a given order continuous at the boundaries of finite elements [6].The efficiency of finite element schemes, algorithms and programs is demonstrated by benchmark calculations of the 4D Helmholtz problem.

1. Algorithm for constructing Hermitian finite elements

The HIPs $\varphi_r^{\kappa}(x) \equiv \varphi_{r_1 \ldots r_i \ldots r_d}^{\kappa_1 \ldots \kappa_i \ldots \kappa_d}(x_1, ..., x_i, ..., x_d)$ of d variables in a d-dimensional parallelepiped element $x = (x_1, ..., x_i, ..., x_d) \in [x_{1;\min}, x_{1;\max}] \times ... \times [x_{d;\min}, x_{d;\max}]$ $= \Delta_q \subset \mathbf{R}^d$ that are obtained on nodes $x_{r_1 \ldots r_i \ldots r_d} = (x_{1r_1}, ..., x_{ir_i}, ..., x_{dr_d})$, $x_{ir_i} = ((p - r_i)x_{i;\min} + r_i x_{i;\max})/p$; $r_i = 0, ..., p$, $i = 1, ..., d$ are determined by relations [1]

$$\varphi_{r_1 \ldots r_i \ldots r_d}^{\kappa_1 \ldots \kappa_i \ldots \kappa_d}(x_{1r'_1}, \ldots, x_{ir'_i}, \ldots, x_{dr'_d}) = \delta_{r_1 r'_1} ... \delta_{r_i r'_i} ... \delta_{r_d r'_d} \delta_{\kappa_1 0} ... \delta_{\kappa_i 0} ... \delta_{\kappa_d 0}, \qquad (1)$$

$$\left.\frac{\partial^{\kappa'_1+\cdots+\kappa'_d}\varphi^{\kappa_1\ldots\kappa_i\ldots\kappa_d}_{r_1\ldots r_i\ldots r_d}(x_1,\ldots,x_i,\ldots,x_d)}{\partial x_1^{\kappa'_1}\cdots\partial x_i^{\kappa'_i}\cdots x_d^{\kappa'_d}}\right|_{(x_1,\ldots,x_i,\ldots,x_d)=(x_{1r'_1},\ldots,x_{ir'_i},\ldots,x_{dr'_d})}$$
$$=\delta_{r_1r'_1}\ldots\delta_{r_ir'_i}\ldots\delta_{r_dr'_d}\delta_{\kappa_1\kappa'_1}\ldots\delta_{\kappa_i\kappa'_i}\ldots\delta_{\kappa_d\kappa'_d}.$$

These HIPs of order $p'=\prod_{s=1}^{d}p'_s$ are calculated as a product of one dimensional HIPs $\varphi^{\kappa_s}_{r_s}(x_s)$: $\varphi^{\kappa}_{r}(x)\equiv\varphi^{\kappa_1\ldots\kappa_i\ldots\kappa_d}_{r_1\ldots r_i\ldots r_d}(x_1,\ldots,x_i,\ldots,x_d)=\prod_{s=1}^{d}\varphi^{\kappa_s}_{r_s}(x_s)$, which are calculated by the following way. For each $z\equiv x_s$ as a set of basis functions, the 1D HIPs $\{\{\varphi^{\kappa}_{r}(z)\}_{r=0}^{p}\}_{\kappa=0}^{\kappa^{\max}_r-1}$ of order $p'=\sum_{r=0}^{p}\kappa^{\max}_r-1$ in a standard interval $z\in[0,1]$ at the nodes z_r, $r=0,\ldots,p$, $z_0=0$, $z_p=1$ are constructed. The values of the functions $\varphi^{\kappa}_{r}(z)\in C^{\kappa_{\max}-1}$ continuous together with their derivatives up to order $(\kappa^{\max}_r-1)$, i.e. $\kappa=0,\ldots,\kappa^{\max}_r-1$, where $\kappa^{\max}_r$ is referred to as the multiplicity [1] of the node z_r, are determined by expressions (1). These 1D HIPs are calculated analytically from the recurrence relations derived in [6]

$$\varphi^{\kappa}_{r}(z)=w_r(z)\sum_{\kappa'=0}^{\kappa^{\max}_r-1}a^{\kappa,\kappa'}_{r}(z-z_r)^{\kappa'},\quad w_r(z)=\prod_{r'=0,r'\neq r}^{p}\left(\frac{z-z_{r'}}{z_r-z_{r'}}\right)^{\kappa^{\max}_{r'}},\qquad(2)$$

$$a^{\kappa,\kappa'}_{r}=\begin{cases}0, & \kappa'<\kappa,\\ 1/\kappa'!, & \kappa'=\kappa,\\ -\sum\limits_{\kappa''=\kappa}^{\kappa'-1}\frac{a^{\kappa,\kappa''}_{r}}{(\kappa'-\kappa'')!}g^{\kappa'-\kappa''}_{r}(z_r), & \kappa'>\kappa,\end{cases}\qquad g^{\kappa}_{r}(z)=\frac{\frac{d^{\kappa}w_r(z)}{dz^{\kappa}}}{w_r(z)}.$$

Below we consider only the HIPs with the nodes of identical multiplicity, $\kappa^{\max}_r=\kappa^{\max}$, $r=0,...,p$, then $p'=\kappa^{\max}(p+1)-1$. For example, at $\kappa_{\max}=2$, $p'=2p+1$ the 1D HIPs take the form:

$$\varphi^{\kappa_s=0}_{r}(z)=\left(1-(z-z_r)\sum_{r'=0,r'\neq r}^{p}\frac{2}{z_r-z'_r}\right)\prod_{r'=0,r'\neq r}^{p}\left(\frac{z-z_{r'}}{z_r-z_{r'}}\right)^2,$$
$$\phi^{\kappa_s=1}_{r}(z)=(z-z_r)\prod_{r'=0,r'\neq r}^{p}\left(\frac{z-z_{r'}}{z_r-z_{r'}}\right)^2,$$

for polynomials $\phi^{\kappa_s=0}_{r}(z)$ or $\phi^{\kappa_s=1}_{r}(z)$ whose value or value of first derivative is equal to 1, respectively.

2. Benchmark Calculations with Hermitian Finite Elements

As benchmark calculations we solve the 4D Helmholtz problem with the edge length π and Neumann boundary conditions. This problem has exact degenerate spectrum: E_m =0 [1] 1 [4] 2 [6] 3 [4] 4 [5] 5 [12]6 [12] 7 [4] 8 [6] 9 [16] 10 [18] 11 [12] 12 [8] 13 [16] 14 [24] 15 [12] ... , where the multiplicity of degeneracy is given in square brackets. The results were calculated on uniform grids using FEM with hypercube LIPs and HIPs of the third order that are obtained by product of four

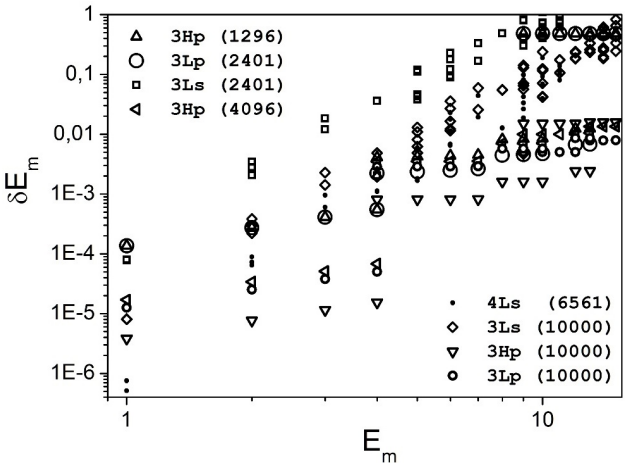


FIGURE 1. The discrepancy $\delta E_m = E_m^h - E_m$ of calculated eigenvalue E_m^h of the Helmholtz problem for a four-dimensional cube with the edge length π. Calculations were performed using FEM with 3rd-order (3Ls) and 4th-order (4Ls) simplex Lagrange elements, and 3rd-order parallelepiped Lagrange (3Lp) and Hermite (3Hp) elements. The dimension of the algebraic problem is given in parentheses.

1D LIPs or four 1D HIPs, respectively. They are compared with simplex LIPs of the third and the fourth order [7].

Figure 1 shows the discrepancy $\delta E_m = E_m^h - E_m$ between the numerical eigenvalues E_m^h and the exact ones E_m. There is a stepwise structure of the discrepancy δE_m calculated with 4D hypercubic LIPs and HIPs, with the steps appearing at the values $E_m = 1, 4, 9, ...$ The structure is also due to the prevalence of approximation errors of eigenfunctions caused by the pure partial derivatives. For the simplex LIPs the oscillating structure of the discrepancy δE_m is due to different contributions the approximation errors caused by the different mixed partial derivatives. The calculation is performed using MAPLE with 12-digit precision. As a consequence, the FEM scheme with the hypercubic LIPs and a large length of the eigenvectors equal to 10000 demonstrates poorer performance than the one with a smaller length due to rounding errors. The above analysis shows the agreement the numerical and theoretical estimations of discrepancy for eigenvalues, $\left|E_m - E_m^h\right| \leq c_m h^{2p'}$, with respect to order p' of FEM schemes with LIPs or HIPs, where h is the step of the uniform grid and $c_m > 0$ are constants independent from h.

Conclusion

The proposed algorithm allows one to construct in analytical form a piecewise polynomial basis continuous on the boundaries of finite elements together with the derivatives up to the given order. It can be used to solve elliptic BVPs as well as other problems with partial derivatives of a high order by means of the high-accuracy finite element method.

The talk was partially supported by RFBR and MECSS, project 20-51-44001.

References

[1] I.S. Berezin, N.P. Zhidkov, *Computing Methods.* v. 1, Pergamon Press, Oxford, 1965.

[2] R.A. Lorentz, *Multivariate Birkhoff interpolation.* Springer-Verlag, Berlin, 1992.

[3] F. Lekien, J. Marsden, *Tricubic interpolation in three dimensions* Int. J. Num. Meth. Eng. 63, 455 (2005)

[4] P. Walker, *Quadcubic interpolation: a four-dimensional spline method*, preprint (2019), available at `http://arxiv.org/abs/1904.09869v1`

[5] G. Chuluunbaatar et al, *Construction of Multivariate Interpolation Hermite Polynomials for Finite Element Method*, EPJ Web of Conferences 226, 02007 (2020).

[6] A.A. Gusev, et al., *Symbolic-numerical solution of boundary-value problems with self-adjoint second-order differential equation using the finite element method with interpolation Hermite polynomials*, Lecture Notes in Computer Sci., 8660, 138–154 (2014)

[7] A.A. Gusev, et al., *Symbolic-numerical algorithm for generating interpolation multivariate Hermite polynomials of high-accuracy finite element method*, Lecture Notes in Computer Sci., 10490, 134–150 (2017)

Galmandakh Chuluunbaatar
Joint Institute for Nuclear Research, Dubna, Russia
RUDN University, Moscow, Russia, 6 Miklukho-Maklaya st, Moscow, 117198

Alexander Gusev
Joint Institute for Nuclear Research, Dubna, Russia
e-mail: `gooseff@jinr.ru`

Vladimir Gerdt
Joint Institute for Nuclear Research, Dubna, Russia
RUDN University, Moscow, Russia, 6 Miklukho-Maklaya st, Moscow, 117198
e-mail: `gerdt@jinr.ru`

Sergue Vinitsky
Joint Institute for Nuclear Research, Dubna, Russia
RUDN University, Moscow, Russia, 6 Miklukho-Maklaya st, Moscow, 117198
e-mail: `vinitsky@theor.jinr.ru`

Lyong Le Hai
Ho Chi Minh city University of Education, Ho Chi Minh city, Vietnam
e-mail: `llhai611987@gmail.com`

Covariance of Parametric Representations of Orthogonal and Symplectic Matrices

Alexander G. Petrov

Abstract. Symplectic matrices are subject to certain conditions that are inherent to the Jacobian matrices of transformations preserving the Hamiltonian form of differential equations. A formula is derived which parameterizes symplectic matrices by symmetric matrices. An analogy is drawn between the obtained formula and the Cayley formula that connects orthogonal and antisymmetric matrices. It is shown that orthogonal and antisymmetric matrices are transformed by the covariant law when replacing the Cartesian coordinate system. Similarly, the covariance of transformations of symplectic and symmetric matrices is proved. From Cayley formulas and their analog, a series of matrix relations is obtained which connect orthogonal and symmetric matrices, together with similar relations connecting symplectic and symmetric matrices.

1. Cayley formulas. Covariance of parameterization of orthogonal matrices.

For orthogonal matrices, the Cayley formulas are known [1]

$$O = (E - K)(E + K)^{-1}, \quad K = (E - O)(E + O)^{-1} \tag{1}$$

They express an orthogonal matrix through an antisymmetric matrix $K^T = -K$.

Here E the unit (identical matrix), the upper index$T$ means the conjugation sign. For any antisymmetric matrix K, the matrix O satisfies the orthogonality conditions $OO^T = O^T O = E$. For any orthogonal matrix O (the eigenvalue is not -1), the matrix K satisfies the antisymmetry conditions.

The transition from one Cartesian system to another can be performed using an orthogonal matrix C, $CC^T = E$. Let the matrix O determine the conversion of a radius vector $\mathbf{r}$ to a radius vector $\mathbf{R} : \mathbf{R} = O\mathbf{r}$. In the new coordinate system we have $\mathbf{r}' = C\,\mathbf{r}$, $\mathbf{R}' = C\,\mathbf{R}$. Find the transformation matrix $\mathbf{R}' = O\mathbf{r}'$. Let's write this transformation in the original coordinate system $C\mathbf{R} = O'C\mathbf{r} \Rightarrow \mathbf{R} =$

The research was carried out within the state assignment of FASO of Russia (state registration № AAAA-A20-120011690138-6)

$C^T O' C\mathbf{r}$. Hence the law of transformation $O = C^T O' C \Rightarrow$

$$O' = COC^T \tag{2}$$

The corresponding transformation law for the matrix K in Cayley's formula follows

$$K' = CKC^T \tag{3}$$

Thus, the transformations of (2) and (3) matrices O and K are both covariant.

2. Covariance of parameterization of symplectic matrices.

A matrix $2n \times 2n$ is called symplectic if it satisfies the relation

$$AIA^T = I, \quad I = \begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix}, \tag{4}$$

where E_n identity matrix $n \times n$.

Symplectic matrices are used in Hamiltonian mechanics. Such a matrix A is the Jacobian transformation matrix of a system of differential equations that preserves the Hamiltonian form [2, 3]. An analogy can be drawn between formulas (1) - (3) and the corresponding formulas for symplectic matrices.

To do this, we must match the orthogonal matrix O – the symplectic matrix and the antisymmetric matrix K to the product of the matrix I and the symmetric matrix $\frac{1}{2}\Psi$. Briefly this correspondence is written as follows $O \to A, \quad K \to \frac{1}{2} I\Psi$.

This analogy is shown in Table 1. On the left of the formula for O and K, on the right for A and $\frac{1}{2}I\Psi$.

$OO^T = O^TO = E \Rightarrow \det(O) = 1$	$A\,I\,A^T = A^T I\,A = I \Rightarrow \quad \det(A) = 1$
$K^T = -K$	$\Psi^T = \Psi$
$O = (E-K)(E+K)^{-1} =$ $= (E+K)^{-1}(E-K) \Rightarrow$ $\det(E+K) = \det(E-K)$	$A = \left(E + \frac{1}{2}I\Psi\right)\left(E - \frac{1}{2}I\Psi\right)^{-1} =$ $= \left(E - \frac{1}{2}I\Psi\right)^{-1}\left(E + \frac{1}{2}I\Psi\right) \Rightarrow$ $\det\left(E + \frac{1}{2}I\Psi\right) = \det\left(E - \frac{1}{2}I\Psi\right)$
$K = (E-O)(E+O)^{-1} =$ $= (E+O)^{-1}(E-O) \Rightarrow$ $\det(E-O)/\det(E+O) = \det(K)$	$\frac{1}{2}I\Psi = (E+A)^{-1}(A-E) =$ $= (A-E)(E+A)^{-1} \Rightarrow$ $\det(A-E)/\det(A+E) = \det\left(\frac{1}{2}I\Psi\right)$
$(E+K)(E+O) = 2E \Rightarrow$	$\left(E - \frac{1}{2}I\Psi\right)(E+A) = 2E \Rightarrow$
$\det(E \pm K)\det(E+O) = 2^{2n}$	$\det\left(E \pm \frac{1}{2}I\Psi\right)\det(E+A) = 2^{2n}$

TABLE 1. Analogy of parameterization of orthogonal and symplectic matrices

Table 2 in the left column shows the output of formulas for converting an orthogonal matrix O, when orthogonal coordinates are replaced with a matrix C. In the right column the same output of the symplectic matrix A transformation formulas for symplectic replacement of coordinates with the matrix B.

$C, \quad CC^T = E$	$B, \quad BIB^T = I$
$\mathbf{R} = O\mathbf{r}$	$\delta\mathbf{R} = A\delta\mathbf{r}$
$\mathbf{r}' = C\mathbf{r},\ \mathbf{R}' = C\mathbf{R}$	$\delta\mathbf{r}' = C\delta\mathbf{r},\ \delta\mathbf{R}' = C\delta\mathbf{R}$
$\mathbf{R}' = O'\mathbf{r}'$	$\delta\mathbf{R}' = A'\delta\mathbf{r}'$
$C\mathbf{R} = O'C\mathbf{r} \Rightarrow \mathbf{R} = C^TO'C\mathbf{r}$	$B\delta\mathbf{R} = O'B\delta\mathbf{r} \Rightarrow \delta\mathbf{R} = B^{-1}A'B\mathbf{r}$
$O = C^TO'C \Rightarrow O' = COC^T$	$A = B^{-1}A'B \Rightarrow A' = BAB^{-1}$

TABLE 2. Formulas for converting an orthogonal O and symplectic A matrices

$O' = COC^{-1}$	$A' = BAB^{-1}$
$O' = (E+K')^{-1}(E-K') =$ $= C(E-K)(E+K)^{-1}C^{-1}$	$A' = \left(E - \frac{1}{2}I\Psi'\right)^{-1}\left(E + \frac{1}{2}I\Psi'\right) =$ $= B\left(E + \frac{1}{2}I\Psi\right)\left(E - \frac{1}{2}I\Psi\right)^{-1}B^{-1}$
$(E-K')C(E+K) =$ $= (E+K')C(E-K)$	$\left(E + \frac{1}{2}I\Psi'\right)B\left(E - \frac{1}{2}I\Psi\right) =$ $\left(E - \frac{1}{2}I\Psi'\right)B\left(E + \frac{1}{2}I\Psi\right)$
$C(E+K-E+K) =$ $= K'C(E-K+E+K)$	$B\left(E - \frac{1}{2}I\Psi - E - \frac{1}{2}I\Psi\right) =$ $= \frac{1}{2}I\Psi'B\left(-E - \frac{1}{2}I\Psi - E + \frac{1}{2}I\Psi\right)$
$2CK = 2K'C \ \Rightarrow K' = CKC^T$	$BI\Psi = I\Psi'B \Rightarrow I\Psi' = BI\Psi B^{-1}$

TABLE 3. Proof of the covariance of the matrix transformation K and $(1/2)I\Psi$

Here $\mathbf{R} = O\mathbf{r}$ is an orthogonal conversion of the radius vector $\mathbf{r}$ to the radius vector $\mathbf{R}$ in the original coordinate system and $\mathbf{R}' = O'\mathbf{r}'$ the same conversion in the other coordinate system. Accordingly $\delta\mathbf{R} = A\delta\mathbf{r}$, $\delta\mathbf{R}' = A'\delta\mathbf{r}'$ are local symplectic transformations in the original and other coordinate systems.

Finally, Table 3 presents in the left column a proof of the covariance of the matrix transformation K that parameterizes the orthogonal transformation. In the right column, there is a similar proof of the covariance of the transformation of the matrix $(1/2)I\Psi$ that parameterizes the symplectic transformation.

Conclusion

Cayley formulas simplify the transformation of orthogonal matrices $O \to O'$. Let the matrix O be expressed as an antisymmetric matrix

$$K = \begin{pmatrix} 0 & k_3 & -k_2 \\ -k_3 & 0 & k_1 \\ k_2 & -k_1 & 0 \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}$$

by Cayley formulas $O = (E-K)(E+K)^{-1}$. For orthogonal substitution with a matrix C , the conversion formula $O' = COC^T$ can be replaced with a vector transformation $\mathbf{k}' = C\mathbf{k}$, to express through vector $\mathbf{k}'$ antisymmetric matrix K' and by Cayley formulas find the transformed orthogonal matrix $O' = (E-K')(E+K')^{-1}$. Similarly, you can transform symplectic matrices $A = \left(E + \frac{1}{2}I\Psi\right)\left(E - \frac{1}{2}I\Psi\right)^{-1}$, expressing them through a symmetric matrix Ψ. Conversion formula $A' = BAB^{-1}$, where B - an arbitrary symplectic matrix can be replaced with the following sequence of transformations $I\Psi' = BI\Psi B^{-1}, \quad A' = \left(E + \frac{1}{2}I\Psi'\right)\left(E - \frac{1}{2}I\Psi'\right)^{-1}$.

Table 3 also implies identities of interest for determinants:

$$\det(E+K) = \det(E-K), \quad \det(E-O)/\det(E+O) = \det(K),$$
$$\det(E \pm K)\det(E+O) = 2^{2n}.$$

The corresponding formulas for symplectic matrices have the form

$$\det\left(E + \tfrac{1}{2}I\Psi'\right) = \det\left(E - \tfrac{1}{2}I\Psi'\right), \quad \det(A-E)/\det(A+E) = \det\left(\tfrac{1}{2}\Psi\right),$$
$$\det\left(E \pm \tfrac{1}{2}I\Psi\right)\det(E+A) = 2^{2n}.$$

In addition, for three-dimensional matrices, we have $\det(E \pm K) = 1 + \mathbf{k}^2$, $\det(E-O) = \det(K) = 0$.
From the last equality, it follows that the eigenvalue of an orthogonal matrix in three-dimensional space is 1 [2, 3].

3. Acknowledgements

The research was carried out within the state assignment of FASO of Russia (state registration № AAAA-A20-120011690138-6)

References

[1] Gantmacher F.R. *Theory of matrices*, Nauka, Moscow, 1967 [in Russian].

[2] Arnold V.I.*Mathematical methods of classical mechanics*, Editorial URSS, Moscow, 2000 [in Russian].

[3] Petrov A.G. *Asymptotic methods for solving the Hamilton equations with the use of a parametrization of canonical transformations*, Dif. Equations 40, 672–685, 2004.

Covariance of Parametric Representations

Alexander G. Petrov
Ishlinsky Institute for Problems in Mechanics RAS (IPMech RAS)
Moscow, Russian Federation
e-mail: petrovipmech@gmail.com

On extensions of canonical symplectic structure from coadjoint orbit of complex general linear group.

Mikhail Babich

Abstract. The Isomonodromic Deformation theory is closely connected with the theory of the phase spaces of the deformation equations. These spaces are the algebraic symplectic spaces constructed from the standard charts. The charts are the coadjoint orbits of $\mathbf{GL}(N,\mathbb{C})$ in the Fuchsian case.

One of the directions of the development of the theory needs to extend the chart, to extend the coadjoint orbit. There are several ways to do it. I introduce an extension method that can be applied to the orbits made from the matrices of arbitrary Jordan type. The method is based on the concept of the flag coordinates on the orbit.

The research was supported by Russian Foundation for Basic Research (RFBR) No. 18-01-00271.

Mikhail Babich
POMI RAN
St.Petersburg, Russia
e-mail: mbabich@pdmi.ras.ru

On invariant coordinate subspaces of normal form of ODE system

Alexander Batkhin

Abstract. A system of ODEs with non-degenerate linear part near its stationary point are considered in two cases: in general case and in Hamiltonian case. Solution of the problem of existence of an invariant coordinate subspace in the coordinates of normal form is proposed as a resonance relation between system's eigenvalues. Algorithms of computer algebra and q-subdiscriminant technique are used for finding such resonance relations.

Introduction

An approach of Poincaré for investigation of systems of nonlinear ordinary differential equations was based on the maximal simplification of the right-hand sides of these equations by invertible transformations. This approach led to the theory of normal forms (NF) of the general system and in particularly of the Hamiltonian ones and was developed in works of G.D. Birkhoff, T.M. Cherry, F.G. Gustavson, C.L. Siegel, J. Moser, A.D. Bruno (see [1]).

The goal of the presented work is to investigate invariant coordinates subspaces of NF of a real Hamiltonian system with non-degenerated linear part. The existence of invariant subspace can reduce the phase flow on the space of less dimension and in some cases can give information about periodic solution of the whole system.

1. Invariant subspaces of normal form of ODE

Consider an analytical system of ODE

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{1}$$

near its stationary point $\mathbf{x} = 0$. Let the linear part

$$\dot{\mathbf{x}} = A\mathbf{x}, \quad A = \partial\mathbf{f}/\partial\mathbf{x}|_{\mathbf{x}=0}\,, \tag{2}$$

of the system (1) be non-degenerated. Let matrix A has n eigenvalues at least one of which is non-zero $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$.

There exists [2] a formal invertible transformation $\mathbf{g} : \mathbf{x} \to \mathbf{y}$, $\mathbf{x} = \mathbf{g}(\mathbf{y})$, represented in the form of power series, which reduces the initial system (1) into its *normal form*

Definition 1. *Normal form* (NF) of the initial system (1) is a system of the form

$$\dot{y}_j = y_j h_j(\mathbf{y}), \quad j = 1, \ldots, n, \tag{3}$$

right-hand sides $y_j h_j(\mathbf{y})$ of which are power series

$$y_j h_j(\mathbf{y}) = y_j \sum_{\mathbf{q}} h_{j\mathbf{q}} \mathbf{y}^{\mathbf{q}}, \quad h_{j0} = \lambda_j, \quad j = 1, \ldots, n, \tag{4}$$

containing only resonant terms with

$$\langle \mathbf{q}, \boldsymbol{\lambda} \rangle = 0. \tag{5}$$

Here $h_{j\mathbf{q}}$ are constant coefficients and in $y_j h_j(\mathbf{y})$ coordinate $q_j \geq -1$, but others $q_k \geq 0$.

Coordinate subspace Let $I = \{i_1, \ldots, i_k\}$ be a set of increasing indices $1 \leq i_1$, $i_k \leq n$, $k \leq n$. By K_I we denote the *coordinate subspace* $K_I = \{\mathbf{y} : y_j = 0 \text{ for all } j \notin I\}$. All non-zero coordinates y_j, $j \in I$, of the subspace K_I we call *internal coordinates* and denote them shortly by $\mathbf{y}_I$, others we call *external coordinates*. The eigenvalues λ_j, $j \in I$, corresponding to the internal coordinates $\mathbf{y}_I$ we call *internal eigenvalues* and denote them by $\boldsymbol{\lambda}_I$. Others λ_j, $j \notin I$, are called *external eigenvalues*.

Problem. Which subspaces K_I are invariant in the normal form (3), (4), (5)?

Theorem 1. *The coordinate subspace K_I of dimension k is invariant in the normal form* (3)–(5) *if each external eigenvalue $\lambda_j \notin \boldsymbol{\lambda}_I$ satisfies the following condition*

$$\lambda_j \neq \langle \mathbf{p}, \boldsymbol{\lambda}_I \rangle, \tag{6}$$

for all integer vectors $\mathbf{p} \geq 0$, $\mathbf{p} \in \mathbb{Z}^k$.

Let consider an analytic Hamiltonian system

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \mathbf{y}}, \quad \dot{\mathbf{y}} = -\frac{\partial H}{\partial \mathbf{x}} \tag{7}$$

with n degrees of freedom near its stationary point $\mathbf{x} = \mathbf{y} = 0$. The eigenvalues of the matrix A can be reordered in a such way: $\lambda_{j+n} = -\lambda_j$, $j = 1, \ldots, n$. Denote by $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$. There exists [3, § 12] a canonical formal transformation which reduces the initial system (7) into its *normal form*

$$\dot{\mathbf{u}} = \partial h / \partial \mathbf{v}, \quad \dot{\mathbf{v}} = -\partial h / \partial \mathbf{u} \tag{8}$$

defined by the normalized Hamiltonian $h(\mathbf{u}, \mathbf{v})$

$$h(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{n} \lambda_j u_j v_j + \sum h_{\mathbf{pq}} \mathbf{u}^{\mathbf{p}} \mathbf{v}^{\mathbf{q}} \tag{9}$$

containing only resonant terms $h_{\mathbf{pq}}\mathbf{u}^{\mathbf{p}}\mathbf{v}^{\mathbf{q}}$ with

$$\langle \mathbf{p}-\mathbf{q}, \boldsymbol{\lambda}\rangle = 0. \tag{10}$$

By L_I we denote the *coordinate subspace* $L_I = \{\mathbf{u}, \mathbf{v} : u_j = v_j = 0 \text{ for all } j \notin I\}$.

Problem. Which subspaces L_I are invariant in the normal form (8), (9), (10)?

Theorem 2. *The coordinate subspace L_I of dimension $2k$ is invariant in the Hamiltonian normal form if each external eigenvalue $\lambda_j \notin \boldsymbol{\lambda}_I$ satisfies the following condition*

$$\lambda_j \neq \langle \mathbf{p}, \boldsymbol{\lambda}_I\rangle\,, \tag{11}$$

for any integer vector $\mathbf{p} \neq 0$, $\mathbf{p} \in \mathbb{Z}^k$.

The principal difference between condition (6) in Theorem 1 and condition (11) in Theorem 2 is that any non-zero vector $\mathbf{p}$ is taken from the lattice $\mathbb{Z}^k$ in the Hamiltonian case but it is taken for $0 \leq \mathbf{p} \in \mathbb{Z}^k$ in the general case.

2. Resonance finding by q-analogue of subdiscriminants

It is immediately follows from (6) and (11) that resonant relations can be determined by eigenvalues of linear part (2). Let consider an important case when all the eigenvalues $\boldsymbol{\lambda}$ are either real or pure imaginary.

Here we propose the following algorithm [4] of searching for resonant relations which essential use technique of q-subdiscriminants [5] and elimination theory. Lets denote by q the commensurability of a pair of eigenvalues: $q = \lambda_i/\lambda_j$.

Step 1: Matrix A of linear system (2) is found and its characteristic polynomial $f_n(\lambda)$ is computed.

Step 2: Compute the sequence of k-th q-subdiscriminants $D_q^{(k)}(f_n)$, $0 \leq k \leq n-2$, which are polynomials in coefficients of f_n and q.

Step 3a: If q-discriminant $D_q(f_n)$ as a polynomial in annulus $\mathbb{Z}[q]$ can be factorized then it is possible to find out all the pairs of resonant eigenvalues.

Step 3b: If the previous step fail then a kind of a brute force algorithm can be applied: for mutually prime pairs (r, s) of integers we check the equality $D_q(f_n) = 0$ for $q = r/s$.

Step 4: Let for a certain value $q^* \in \mathbb{Q}$ it is true that $D_{q^*}(f_n) = 0$. Then it is possible with the help of q-subdiscriminants from the Step 2 to determine the structure of eigenvalues with commensurability q^* and in some cases even to compute them.

Step 5: Having all the commensurable eigenvalues λ_i it is possible to check either conditions (6), (11) take place or no.

3. Example

Let all $\lambda_j/\lambda_1 \notin \mathbb{Z}$, $j = 2, \ldots, n$. Then the normal form has two-dimensional invariant subspace $L_1 = \{u_j = v_j = 0, j \notin I_1\}$, where $I_1 = \{1\}$. On the subspace L_1 Hamiltonian NF (9) induces a NF with one degree of freedom and the normalizing transformation converges.

If $\lambda_1 \neq 0$ and is purely imagine, then for real Hamiltonian system the real subspace L_1 is a family of periodic solutions. This fact was found by A.M. Lyapunov in 1892 and was described with the help of Hamiltonian formalism by C. Siegel [6, §§16, 17].

Acknowledgment

The work is supported by RFBR, Project No. 18-01-00422a.

References

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

[2] A. D. Bruno. Analytical form of differential equations (I). *Trans. Moscow Math. Soc.*, 25:131–288, 1971. = Trudy Moskov. Mat. Obsc. 25 (1971) 119-262 (in Russian).

[3] A. D. Bruno. Analytical form of differential equations (II). *Trans. Moscow Math. Soc.*, 26:199–239, 1972. = Trudy Moskov. Mat. Obsc. 25 (1971) 119-262 (in Russian).

[4] A. B. Batkhin. Invariant coordinate subspaces of normal form of a system of ordinary differential equations. *Preprints of KIAM*, (72), 2020. (in Russian). `https://doi.org/10.20948/prepr-2020-72` `doi:10.20948/prepr-2020-72`.

[5] A. B. Batkhin. Parameterization of a set determined by the generalized discriminant of a polynomial. *Programming and Computer Software*, 44(2):75–85, 2018. `https://doi.org/10.1134/S0361768818020032` `doi:10.1134/S0361768818020032`.

[6] C. L. Siegel and J. K. Moser. *Lectures on Celestial Mechanics.* Springer-Verlag, Berlin, Heidelber, New York, 1971.

Alexander Batkhin
Department of Singular Problems
Keldysh Institute of Applied Mathematics of RAS
Department of Theoretical Mechanics
Moscow Institute of Physics and Technology
Moscow, Russia
e-mail: `batkhin@gmail.com`

The close relation between border and Pommaret marked bases

Cristina Bertone and Francesca Cioffi

Abstract. Given a finite order ideal $\mathcal{O}$, we investigate border and Pommaret marked sets related to this order ideal. We use the framework of reduction structures given in [3].
First, we prove that a marked set B on the border of $\mathcal{O}$ is a basis if and only if the marked set on the Pommaret basis of the complementary ideal of $\mathcal{O}$ contained in B is a basis and generates the same ideal as B.
As a byproduct, using a functorial description of border and Pommaret marked bases, we obtain that the scheme parameterizing marked bases on the border of $\mathcal{O}$ and the scheme parameterizing marked bases on the Pommaret basis of the complementary ideal of $\mathcal{O}$ are isomorphic. We also explicitly construct such an isomorphism.

Introduction

Consider the variables $x_1, \dots, x_n$, with $x_1 < \cdots < x_n$, the set $\mathbb{T}$ containing the terms in the variables $x_1, \dots, x_n$ and the polynomial ring $R_A := A[x_1, \dots, x_n]$, being A a Noetherian algebra over a field K with unit 1_K.

If $\mathcal{O} \subset \mathbb{T}$ is a finite order ideal, we can define a set $F \subset R_A$ of monic marked polynomials whose *head terms* are the *border of* $\mathcal{O}$, $\partial\mathcal{O}$, and study the conditions ensuring that F is a $\partial\mathcal{O}$-marked basis, i.e. $(F) \oplus \langle \mathcal{O} \rangle_A = R_A$, where $\langle \mathcal{O} \rangle_A$ is the A-module generated by $\mathcal{O}$.

Border marked bases (border bases in the literature) were first introduced in [9] and investigated from a numerical point of view because of their stability with respect to perturbation of the coefficients [10, 11]. Border bases have also attracted interest from an algebraic point of view [7, Section 6.4], also because given a finite order ideal $\mathcal{O}$, the border bases on $\mathcal{O}$ parameterize an open subset of a Hilbert scheme (see also [5, 6]).

Since every Artinian monomial ideal in R_K has a *Pommaret basis*, given a finite order ideal $\mathcal{O}$, we can consider the Pommaret basis $\mathcal{P}_\mathcal{O}$ of the monomial ideal

generated by $\mathbb{T} \setminus \mathcal{O}$ and construct monic marked sets and bases whose head terms form $\mathcal{P}_\mathcal{O}$.

Marked bases on strongly stable ideals were first introduced in [4] with the aim to parameterize open subsets of a Hilbert scheme, in order to study it locally. This kind of basis does not need any finiteness assumption on the underlying order ideal. In [1] marked bases were considered only in the case of homogeneous polynomials, but in [2] also non-homogeneous marked bases over monomial ideals having a Pommaret basis were considered, in order to have more efficient computational techniques for the homogeneous case.

The goal of our work is comparing marked sets (and bases) on the border $\partial\mathcal{O}$ of $\mathcal{O}$ and on the Pommaret basis $\mathcal{P}_\mathcal{O}$ of the ideal $(\mathbb{T} \setminus \mathcal{O})$. To this aim we use the framework of reduction structures [3], and a functorial approach to study the schemes parameterizing these two different bases. The monicity of the marked sets we consider is crucial for the use of functors.

Observing that a set B of marked polynomials on $\partial\mathcal{O}$ always contains a set P of marked polynomials on $\mathcal{P}_\mathcal{O}$, we prove that B is a $\partial\mathcal{O}$-marked basis if and only if P is a $\mathcal{P}_\mathcal{O}$-marked basis and generates the same ideal as B.

As a byproduct, using a functorial description of border and Pommaret marked bases, we obtain that the scheme parameterizing $\partial\mathcal{O}$-marked bases and the scheme parameterizing $\mathcal{P}_\mathcal{O}$-marked bases are isomorphic. We also explicitly construct such an isomorphism.

1. Framework

If σ is a term in $\mathbb{T}$, we denote by $\min(\sigma)$ the smallest variable dividing σ. A set $\mathcal{O}$ of terms in $\mathbb{T}$ is called an *order ideal* if for every $\sigma \in \mathbb{T}$ and every $\tau \in \mathcal{O}$, if σ divides τ, then σ belongs to $\mathcal{O}$.
Given a finite order ideal $\mathcal{O}$, *the border* of $\mathcal{O}$ is $\partial\mathcal{O} := \{x_i \cdot \tau \mid \tau \in \mathcal{O}, i \in \{1, \dots, n\}\} \setminus \mathcal{O}$ [7, Definition 6.4.4], and the *Pommaret basis* of $\mathbb{T} \setminus \mathcal{O}$ is $\mathcal{P}_\mathcal{O} = \{\sigma \in \mathbb{T} \setminus \mathcal{O} | \sigma/\min(\sigma) \in \mathcal{O}\}$. Observe that $\mathcal{P}_\mathcal{O} \subset \partial\mathcal{O}$.

Definition 1. [3, Definition 3.1] A *reduction structure* $\mathcal{J}$ in $\mathbb{T}$ is a 3-uple $\mathcal{J} := (\mathcal{H}, \mathcal{L} := \{\mathcal{L}_\alpha \mid \alpha \in \mathcal{H}\}, \mathcal{T} := \{\mathcal{T}_\alpha \mid \alpha \in \mathcal{H}\})$ where: $\mathcal{H} \subseteq \mathbb{T}$ is a *finite set* of terms; for every $\alpha \in \mathcal{H}$, $\mathcal{T}_\alpha \subseteq \mathbb{T}$ is an order ideal, such that $\cup_{\alpha \in \mathcal{H}} \{\tau\alpha | \tau \in \mathcal{T}_\alpha\} = (\mathcal{H})$; for every $\alpha \in \mathcal{H}$, $\mathcal{L}_\alpha$ is a finite subset of $\mathbb{T} \setminus \{\tau\alpha | \tau \in \mathcal{T}_\alpha\}$.

A *marked polynomial* is a polynomial $f \in R_A$ with a specified term of $\mathrm{Supp}(f)$, the *head term* of f, denoted by $\mathrm{Ht}(f)$, which appears in f with coefficient 1_K.

Definition 2. [3, Definitions 4.2 and 4.3] Given a reduction structure $\mathcal{J} = (\mathcal{H}, \mathcal{L}, \mathcal{T})$, a set F of exactly $|\mathcal{H}|$ marked polynomials in R_A is called a *$\mathcal{H}$-marked set* if, for every $\alpha \in \mathcal{H}$, there is $f_\alpha \in F$ with $\mathrm{Ht}(f_\alpha) = \alpha$ and $\mathrm{Supp}(f) \subseteq \mathcal{L}_\alpha$.
Let $\mathcal{O}_\mathcal{H}$ be the order ideal given by the terms of $\mathbb{T}$ outside the ideal generated by $\mathcal{H}$. A $\mathcal{H}$-marked set F is called a *$\mathcal{H}$-marked basis* if $(F) \oplus \langle \mathcal{O}_\mathcal{H} \rangle_A = R_A$.

From now on, let the terms of the border $\partial\mathcal{O}$ be ordered by increasing degree (terms of the same degree are ordered arbitrarily) and labelled coherently: for every $\beta_i, \beta_j \in \partial\mathcal{O}$, if $i < j$ then $\beta_i < \beta_j$.

Definition 3. Let $\mathcal{O} \subset \mathbb{T}$ be a finite order ideal.
The *Pommaret reduction structure* $\mathcal{J}_{\mathcal{P}}$ is the reduction structure with $\mathcal{H} = \mathcal{P}_{\mathcal{O}}$ and, for every $\alpha \in \mathcal{P}_{\mathcal{O}}$, $\mathcal{L}_\alpha = \mathcal{O}$ and $\mathcal{T}_\alpha = \mathbb{T} \cap K[x_1, \ldots, \min(\alpha)]$.
The *border reduction structure* $\mathcal{J}_{\partial\mathcal{O}}$ is the reduction structure with $\mathcal{H} = \partial\mathcal{O}$ and, for every $\beta_i \in \partial\mathcal{O}$, $\mathcal{L}_{\beta_i} = \mathcal{O}$ and $\mathcal{T}_{\beta_i} = \{\mu \in \mathbb{T} \mid \forall j > i,\ \beta_j \text{ does not divide } \beta_i\mu\}$.

For every reduction structure $\mathcal{J} = (\mathcal{H}, \mathcal{L}, \mathcal{T})$ and every $\mathcal{H}$-marked set F, it is possible to define a *reduction relation* on polynomials in R_A, that we denote by $\to_{F\,\mathcal{J}}$. If B (resp. P) is a $\partial\mathcal{O}$-marked set (resp. a $\mathcal{P}_{\mathcal{O}}$-marked set), for every $f \in R_A$ there is $h_B \in \langle\mathcal{O}\rangle_A$ (resp. $h_P \in \langle\mathcal{O}\rangle_A$) such that $f \to_{B\,\mathcal{J}_{\mathcal{O}}} h_B$ (resp. $f \to_{P\,\mathcal{J}_{\mathcal{P}_{\mathcal{O}}}} h_P$). Observe that in general $h_B \neq h_P$. In particular, both the border reduction and the Pommaret reduction structures give Noetherian and confluent reduction relations. These properties ensure that $(B) + \langle\mathcal{O}\rangle_A = (P) + \langle\mathcal{O}\rangle_A = R_A$.

2. Main results

Theorem 4. *Let $\mathcal{O} \subset \mathbb{T}$ be a finite order ideal. Let B be a $\partial\mathcal{O}$-marked set in R_A and we denote by P the $\mathcal{P}_{\mathcal{O}}$-marked set contained in B. Then we have*

$$B \text{ is a } \partial\mathcal{O}\text{-marked basis} \Leftrightarrow P \text{ is a } \mathcal{P}_{\mathcal{O}}\text{-marked basis and } (B) = (P).$$

Definition 5. [3, Appendix A] Let $\mathcal{O} \subset \mathbb{T}$ be a finite order ideal and let $\mathcal{J} = (\mathcal{H}, \mathcal{L}, \mathcal{T})$ be a reduction structure with $(\mathcal{H}) = \mathbb{T} \setminus \mathcal{O}$. We consider the functor

$$\mathcal{M}b_{\mathcal{J}} : \text{ Noeth-}k\text{-Alg} \longrightarrow \text{ Sets},$$

that associates to every Noetherian k-Algebra A the set $\mathcal{M}b_{\mathcal{J}}(A)$ consisting of all the ideals $I \subset R_A$ generated by a $\mathcal{H}$-marked basis, and to every morphism of Noetherian k-algebras $\phi : A \to A'$ the morphism $\mathcal{M}b_{\mathcal{J}}(\phi) : \mathcal{M}b_{\mathcal{J}}(A) \to \mathcal{M}b_{\mathcal{J}}(A')$ that operates in the following natural way:

$$\mathcal{M}b_{\mathcal{J}}(\phi)(I) = I \otimes_A A'.$$

Remark 6. The monicity of marked sets and bases guarantees that marked set and bases are preserved by extension of scalars (see also [3, Lemmas A.1 and A.2]).

If $|\mathcal{O}| = \ell$ and $|\partial\mathcal{O}| = m$, we define $C := \{C_{i,j}\}_{1 \leq i \leq m, 1 \leq j \leq \ell}$. The *generic $\partial\mathcal{O}$-marked set* [8, Definition 3.1] is the set $\mathscr{B}$ of marked polynomials $\{g_1, \ldots, g_m\} \subset R_{K[C]}$ with $g_i = \tau_i - \sum_{j=1}^{\ell} C_{ij}\sigma_j$.

The set $\mathscr{B}$ contains the generic $\mathcal{P}_{\mathcal{O}}$-marked set $\mathscr{P}$. We denote by $\tilde{C}$ the set of parameters not appearing in $\mathscr{P}$. By Buchberger criteria for $\partial\mathcal{O}$-marked bases [7, Proposition 6.4.34] and for $\mathcal{P}_{\mathcal{O}}$-marked bases [2, Proposition 5.6], it is possible to prove that the functor $\mathcal{M}b_{\mathcal{J}_{\partial\mathcal{O}}}$ (resp. $\mathcal{M}b_{\mathcal{J}_{\mathcal{P}_{\mathcal{O}}}}$) is the functor of points of $\mathrm{Spec}(K[C]/\mathfrak{B})$ (resp. $\mathrm{Spec}(K[C \setminus \tilde{C}]/\mathfrak{P})$, where $\mathfrak{B}$ (resp. $\mathfrak{P}$) is generated by a

finite set of polynomials in $K[C]$ (resp. $K[C \setminus \tilde{C}]$) explicitly computed by $\rightarrow_{\mathscr{B},\mathcal{J}_{\partial\mathcal{O}}}$ (resp. $\rightarrow_{\mathscr{P}\,\mathcal{J}_{\mathcal{P}_\mathcal{O}}}$). Thanks to Theorem 4, we can prove the following.

Theorem 7.

1. *The schemes* $\mathrm{Spec}(K[C]/\mathfrak{B})$ *and* $\mathrm{Spec}(K[C \setminus \tilde{C}]/\mathfrak{P})$ *are isomoprhic;*
2. *there is a isomorphism* $\psi : \mathrm{Spec}(K[C]/\mathfrak{B}) \rightarrow \mathrm{Spec}(K[C \setminus \tilde{C}]/\mathfrak{P})$ *defined by computing for every* $\beta \in \partial\mathcal{O} \setminus \mathcal{P}_\mathcal{O}$ *the polynomial* $h_\beta \in \langle\mathcal{O}\rangle_A$ *such that* $\beta \rightarrow_{\mathscr{P}\,\mathcal{J}_{\mathcal{P}_\mathcal{O}}} h_\beta$.

References

[1] C. Bertone, F. Cioffi, P. Lella, and M. Roggero, *Upgraded methods for the effective computation of marked schemes on a strongly stable ideal*, J. Symbolic Comput., 50, (2013).

[2] C. Bertone, F. Cioffi, and M. Roggero, *Macaulay-like marked bases*, J. Algebra Appl., 16, (2017), no.5.

[3] M. Ceria, T. Mora, and M. Roggero, *A general framework for Noetherian well ordered polynomial reductions*, J. Symbolic Comput., 95, (2019).

[4] F. Cioffi, and M. Roggero, *Flat families by strongly stable ideals and a generalization of Gröbner bases*, J. Symbolic Comput., 46, (2011), no.9.

[5] M. E. Huibregtse, *Some syzygies of the generators of the ideal of a border basis scheme*, Collect. Math., 62, (2011), no.3.

[6] M. Lederer, *Gröbner strata in the Hilbert scheme of points*, J. Commut. Algebra, 3, (2011), no.3.

[7] M. Kreuzer, and L. Robbiano, *Computational commutative algebra. 2*, Springer-Verlag, Berlin, (2005).

[8] M. Kreuzer, and L. Robbiano, *Deformations of border bases*, Collect. Math., 59, (2008), no.3.

[9] M.G. Marinari, H.M. Möller, and T. Mora, *Gröbner bases of ideals defined by functionals with an application to ideals of projective points*, Appl. Algebra Engrg. Comm. Comput. 4, (1993), no.2.

[10] B. Mourrain, *A new criterion for normal form algorithms*, Lecture Notes in Comput. Sci.,vol.1719,Springer, Berlin, (1999).

[11] H. J. Stetter, *Numerical polynomial algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, (2004).

Cristina Bertone
Dipartimento di Matematica "G. Peano", Università di Torino, Italy
e-mail: `cristina.bertone@unito.it`

Francesca Cioffi
Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università degli Studi di Napoli Federico II, Italy
e-mail: `cioffifr@unina.it`

The Newest Methods of Celestial Mechanics

Alexander Bruno

Abstract. Here for Hamiltonian systems we describe two of five methods of Celestial Mechanics. Namely: method of normal forms, allowing to study regular perturbations near a stationary solution, near a periodic solution, and method of truncated systems, found with a help of the Newton polyhedrons, allowing to study singular perturbations. Other three methods will be in the full presentation.

1. Normal forms

Here and below vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$ are denoted by boldface font: $\mathbf{x} = (x_1, \ldots, x_n)$.

Let us consider the Hamiltonian system

$$\dot{\xi}_j = \frac{\partial \gamma}{\partial \eta_j}\,, \quad \dot{\eta}_j = -\frac{\partial \gamma}{\partial \xi_j}\,, \quad j = 1, \ldots, n \tag{1}$$

with n degrees of freedom in a vicinity of the stationary solution

$$\boldsymbol{\xi} = \boldsymbol{\eta} = 0. \tag{2}$$

If the Hamiltonian function $\gamma(\boldsymbol{\xi}, \boldsymbol{\eta})$ is analytic in the point (2), then it is expanded into the power series

$$\gamma(\boldsymbol{\xi}, \boldsymbol{\eta}) = \sum \gamma_{\mathbf{pq}} \boldsymbol{\xi}^{\mathbf{p}} \boldsymbol{\eta}^{\mathbf{q}}\,, \tag{3}$$

where $\mathbf{p}, \mathbf{q} \in \mathbb{Z}^n$, $\mathbf{p}, \mathbf{q} \geq 0$, $\boldsymbol{\xi}^{\mathbf{p}} = \xi_1^{p_1} \cdots \xi_n^{p_n}$. Here $\gamma_{\mathbf{pq}}$ are constant coefficients.

As the point (2) is stationary, than the expansion (3) begins from quadratic terms. They correspond to the linear part of the system (1). Eigenvalues of its matrix are decomposed in pairs:

$$\lambda_{j+n} = -\lambda_j\,, \quad j = 1, \ldots, n\,.$$

Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$. The canonical changes of coordinates

$$\boldsymbol{\xi}, \boldsymbol{\eta} \longrightarrow \mathbf{x}, \mathbf{y} \tag{4}$$

preserve the Hamiltonian structure of the system.

Theorem 1 ([1, §12]). *There exists a formal canonical transformation* (4)*, bringing Hamiltonian* (3) *to the normal form*

$$g(\mathbf{x},\mathbf{y}) = \sum g_{\mathbf{pq}}\mathbf{x}^{\mathbf{p}}\mathbf{y}^{\mathbf{q}} \tag{5}$$

contains only resonant terms with scalar product

$$\langle \mathbf{p}-\mathbf{q},\boldsymbol{\lambda}\rangle = 0.$$

If $\boldsymbol{\lambda} \neq 0$, then the system corresponding to the normal form (5) is equivalent to a system with smaller number of degrees of freedom and with additional parameters. The normalizing transformation (4) conserves small parameters and linear automorphisms of the initial system (1)

$$\boldsymbol{\xi},\boldsymbol{\eta} \longrightarrow \widetilde{\boldsymbol{\xi}},\widetilde{\boldsymbol{\eta}},, \quad t \to \widetilde{t}\,.$$

For the real initial system (1), the coefficients $g_{\mathbf{pq}}$ of the complex normal form (5) satisfy to special properties of reality and after a standard canonical linear change of coordinates $\mathbf{x},\mathbf{y} \to \mathbf{X},\mathbf{Y}$ Hamiltonian (5) transforms in a real one [2, Ch. I]. There are several methods of computation of coefficients $g_{\mathbf{pq}}$ of the normal form (5). The most simple method was described in the book [3]. Normal forms near a periodic solution, near an invariant torus and near family of them see in [2, Chs. II, VII, VIII], [4, Part II], [5], [6]. Normal form is useful in study stability, bifurcations and asymptotic behavior of solutions.

2. Truncated Hamiltonian functions

Let $\mathbf{x}$, $\mathbf{y}$ and $\boldsymbol{\mu} = (\mu_1,\dots,\mu_s)$ be canonical variables and small parameters respectively. Let a Hamiltonian function be

$$h(\mathbf{x},\mathbf{y},\boldsymbol{\mu}) = \sum h_{\mathbf{pqr}}\mathbf{x}^{\mathbf{p}}\mathbf{y}^{\mathbf{q}}\boldsymbol{\mu}^{\mathbf{r}} \tag{6}$$

where $h_{\mathbf{pqr}}$ are constant coefficients and $\mathbf{r} \in \mathbb{Z}^s$, $\mathbf{r} \geq 0$. To each term of sum (6) we put in correspondence its vectorial power exponent $Q = (\mathbf{p},\mathbf{q},\mathbf{r}) \in \mathbb{R}^{2n+s}$. Set $\mathbf{S}$ of all points Q with $h_Q \neq 0$ in sum (6) is called as *support* $\mathbf{S} = \mathbf{S}(h)$ of the sum (6). The convex hull $\boldsymbol{\Gamma}(\mathbf{S}) = \boldsymbol{\Gamma}(h)$ of the support $\mathbf{S}$ is called as the *Newton polyhedron* of the sum (6). Its boundary $\partial\boldsymbol{\Gamma}(h)$ consists of vertices $\boldsymbol{\Gamma}_j^{(0)}$, edges $\boldsymbol{\Gamma}_j^{(1)}$ and faces $\boldsymbol{\Gamma}_j^{(d)}$ of dimensions d: $1 < d \leq 2n+s-1$. Intersection $\mathbf{S}\bigcap\boldsymbol{\Gamma}_j^{(d)} = \mathbf{S}_j^{(d)}$ is the *boundary subset* of set $\mathbf{S}$. To each *generalized face* $\boldsymbol{\Gamma}_j^{(d)}$ (including vertices and edges) there correspond:

- *normal cone* $\mathbf{U}_j^{(d)}$ in space $\mathbb{R}_*^{2n+s}$, which is dual to space $\mathbb{R}^{2n+s}$;
- *truncated sum*

$$\hat{h}_j^{(d)} = \sum h_{\mathbf{pqr}}\mathbf{x}^{\mathbf{p}}\mathbf{y}^{\mathbf{q}}\boldsymbol{\mu}^{\mathbf{r}} \;\text{ over }\; Q = (\mathbf{p},\mathbf{q},\mathbf{r}) \in \mathbf{S}_j^{(d)}.$$

The truncated sum is the first approximation to the sum (6), when

$$(\log|x_j|, \log|y_j|, \log|\mu_k|) \to \infty, \quad j = 1, \ldots, n, \quad k = 1, \ldots, s,$$

near the normal cone $\mathbf{U}_j^{(d)}$.

So we can describe the approximate problems by truncated Hamiltonian functions. Example see below in Section 3.

3. Restricted 3-body problem

Let the two bodies $\mathbf{P}_1$ and $\mathbf{P}_2$ with masses $1-\mu$ and μ respectively turn in circular orbits around their common mass center with the period T. The plane circular restricted three-body problem consists in the study of the plane motion of the body $\mathbf{P}_3$ of infinitesimal mass under the influence of the Newton gravitation of bodies $\mathbf{P}_1$ and $\mathbf{P}_2$. In the rotating (synodical) standardized coordinate system the problem is described by the Hamiltonian system with two degrees of freedom and with one parameter μ [2]. The Hamiltonian function has the form

$$h \stackrel{\text{def}}{=} \frac{1}{2}\left(y_1^2 + y_2^2\right) + x_2 y_1 - x_1 y_2 - \frac{1-\mu}{\sqrt{x_1^2 + x_2^2}} - \frac{\mu}{\sqrt{(x_1-1)^2 + x_2^2}} + \mu x_1. \qquad (7)$$

Here the body $\mathbf{P}_1 = \{X, Y : x_1 = x_2 = 0\}$ and the body $\mathbf{P}_2 = \{X, Y : x_1 = 1,\ x_2 = 0\}$, where $X = (x_1, x_2)$, $Y = (y_1, y_2)$. We consider the small values of the mass ratio $\mu \geq 0$. When $\mu = 0$ the problem turns into the two-body problem for $\mathbf{P}_1$ and $\mathbf{P}_3$. But here the points corresponding to collisions of the bodies $\mathbf{P}_2$ and $\mathbf{P}_3$ must be excluded from the phase space. The points of collisions split in parts solutions to the two-body problem for $\mathbf{P}_1$ and $\mathbf{P}_3$. For small $\mu > 0$ there is a singular perturbation of the case $\mu = 0$ near the body $\mathbf{P}_2$. In order to find all the first approximations to the restricted three-body problem, it is necessary to introduce the local coordinates near the body $\mathbf{P}_2$

$$\xi = x_1 - 1,\ \xi_2 = x_2,\ \eta_1 = y_1,\ \eta_2 = y_2 - 1$$

and to expand the Hamiltonian function in these coordinates. After the expansion of $1/\sqrt{(\xi_1+1)^2 + \xi_2^2}$ in the Maclaurin series, the Hamiltonian function (7) takes the form

$$\begin{aligned} h + \frac{3}{2} - 2\mu \stackrel{\text{def}}{=} & \frac{1}{2}(\eta_1^2 + \eta_2^2) + \xi_2\eta_1 - \xi_1\eta_2 - \xi_1^2 + \frac{1}{2}\xi_2^2 + \\ & + f\left(\xi_1, \xi_2^2\right) + \mu\left\{\xi_1^2 - \frac{1}{2}\xi_2^2 - \frac{1}{\sqrt{\xi_1^2 + \xi_2^2}} - f\left(\xi_1, \xi_2^2\right)\right\}, \end{aligned} \qquad (8)$$

where f is the convergent power series, where the terms of order less then three are absent. Let for each term of sum (8) we put

$$p = \operatorname{ord}\xi_1 + \operatorname{ord}\xi_2, \quad q = \operatorname{ord}\eta_1 + \operatorname{ord}\eta_2, \quad r = \operatorname{ord}\mu.$$

Then support $\mathbf{S}$ of the expansion (8) consists of the points

$$(0,2,0), (1,1,0), (2,0,0), (k,0,0), (2,0,1), (-1,0,1), (k,0,1),$$

where $k = 3, 4, 5, \ldots$ The convex hull of the set $\mathbf{S}$ is the polyhedron $\mathbf{\Gamma} \subset \mathbb{R}^3$. The surface $\partial\mathbf{\Gamma}$ of the polyhedron $\mathbf{\Gamma}$ consists of faces $\mathbf{\Gamma}_j^{(2)}$, edges $\mathbf{\Gamma}_j^{(1)}$ and vertices $\mathbf{\Gamma}_j^{(0)}$. To each of the elements $\mathbf{\Gamma}_j^{(d)}$ there corresponds the truncated Hamiltonian $\hat{h}_j^{(d)}$, that is the sum of those terms of Series (8), the points $Q = (p, q, r)$ of which belong to $\mathbf{\Gamma}_j^{(d)}$. Fig. 1 shows the polyhedron $\mathbf{\Gamma}$, which is the semi-infinite trihedral prism with an oblique base. It has four faces and six edges. Let us consider them.

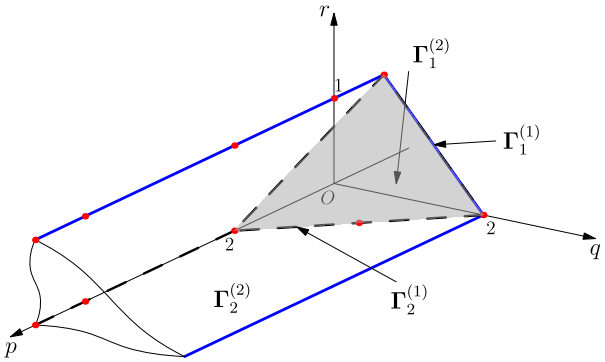


FIGURE 1. The polyhedron $\mathbf{\Gamma}$ for the Hamiltonian function (8) in coordinates p, q, r.

The face $\mathbf{\Gamma}_1^{(2)}$, which is the oblique base of the prism $\mathbf{\Gamma}$, contains vertices $(0, 2, 0)$, $(2, 0, 0)$, $(-1, 0, 1)$ and the point $(1, 1, 0) \in \mathbf{S}$ To the face there corresponds the truncated Hamiltonian function

$$\hat{h}_1^{(2)} = \frac{1}{2}\left(\eta_1^2 + \eta_2^2\right) + \xi_2\eta_1 - \xi_1\eta_2 - \xi_1^2 + \frac{1}{2}\xi_2^2 - \frac{\mu}{\sqrt{\xi_1^2 + \xi_2^2}}. \tag{9}$$

It describes the Hill problem [7], which is a non-integrable one. The canonical power transformation

$$\tilde{\xi}_i = \xi_i\mu^{-1/3}, \quad \tilde{\eta}_i = \eta_i\mu^{-1/3}, \quad i = 1, 2, \tag{10}$$

reduces the Hamiltonian (9) to the Hamiltonian of the form (9), where ξ_i, η_i, μ must be substituted by $\tilde{\xi}_i, \tilde{\eta}_1, 1$ respectively.

The face $\mathbf{\Gamma}_2^{(2)}$ contains points $(0, 2, 0)$, $(1, 1, 0)$, $(2, 0, 0)$ and $(k, 0, 0) \subset \mathbf{S}$. To the face there corresponds the truncated Hamiltonian function $\hat{h}_2^{(2)}$, which is obtained from the function h when $\mu = 0$. It describes the two-body problem for $\mathbf{P}_1$ and $\mathbf{P}_3$, which is an integrable one.

The edge $\mathbf{\Gamma}_1^{(1)}$ includes points $(0, 2, 0)$ and $(-1, 0, 1) \subset \mathbf{S}$. The corresponding truncated Hamiltonian function is

$$\hat{h}_1^{(1)} = \frac{1}{2}(\eta_1^2 + \eta_2^2) - \frac{\mu}{\sqrt{\xi_1^2 + \xi_2^2}}. \tag{11}$$

It describes the two-body problem for $\mathbf{P}_2$ and $\mathbf{P}_3$. The power transformation (10) transforms Hamiltonian (11) into the Hamiltonian function of the form (11), where ξ_i, η_i, μ must be substituted by $\tilde{\xi}_i, \tilde{\eta}_1, 1$ respectively.

The edge $\mathbf{\Gamma}_2^{(1)}$ includes points $(2,2,0), (1,1,0), (0,2,0) \subset \mathbf{S}$. To it there corresponds the truncated Hamiltonian function (9) with $\mu = 0$. It describes the intermediate problem (between the Hill problem and the two-body problem for $\mathbf{P}_1$ and $\mathbf{P}_3$), which is an integrable one. This first approximation was introduced by Hénon [8].

Acknowledgement

This work was supported by RBRF, grant No. 18-01-00422a.

References

[1] A. D. Bruno. Analytical form of differential equations (II). *Trans. Moscow Math. Soc.*, 26:199–239, 1972. = Trudy Moskov. Mat. Obsc. 25 (1971) 119-262 (in Russian).

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

[3] V. F. Zhuravlev, A. G. Petrov, and M. M. Shunderyuk. *Selected Problems of Hamiltonian Mechanics.* LENAND, Moscow, 2015. (in Russian).

[4] A. D. Bruno. *Local Methods in Nonlinear Differential Equations.* Springer–Verlag, Berlin – Heidelberg – New York – London – Paris – Tokyo, 1989.

[5] A. D. Bruno. Normal form of a Hamiltonian system with a periodic perturbation. *Computational Mathematics and Mathematical Physics*, 60(1):36–52, 2020. = Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki, 2020, V. 60(1): 39–56 (in Russian). `https://doi.org/10.1134/S0965542520010066` `doi:10.1134/S0965542520010066`.

[6] A. D. Bruno. Normalization of the periodic Hamiltonian system. *Programming and Computer Software*, 46(2):76–83, 2020. = Programmirovanie, 2020, V. 46(2): 6–13 (in Russian). `https://doi.org/10.31857/S0132347420020053` `doi:10.31857/S0132347420020053`.

[7] G. W. Hill. Researches in the lunar theory. *Amer. J. Math.*, 1:5–26, 129–147, 245–260, 1878.

[8] M. Hénon. Numerical exploration of the restricted problem. V. Hill's case: periodic orbits and their stability. *Astron. & Astrophys.*, 1:223–238, 1969.

Alexander Bruno
Department of Singular Problems
Keldysh Institute of Applied Mathematics of RAS
Moscow, Russia
e-mail: `abruno@keldysh.ru`

Averaged indicator of classicality/quantumness in quasiprobability representations of finite-dimensional quantum systems

Nurlan Abbasly, Vahagn Abgaryan, Martin Bureš, Arsen Khvedelidze, Ilya Rogojin and Astghik Torosyan

Abstract. We discuss measures of classicality/quantumness of states of finite-dimensional quantum systems, which are based on a deviation of quasiprobability distributions from true statistical distributions. Particularly, the dependence of the global indicator of classicality on the assigned geometry of a quantum state space is analysed for a whole family of Wigner quasiprobability representations. General considerations are exemplified by constructing the global indicator of classicality/quantumness for the Hilbert-Schmidt, Bures, Bogoliubov-Kubo-Mori and Wigner-Yanase-Dyson ensembles of qubits and qutrits. In the case of qutrits, by averaging over the one-parameter moduli space (describing a family of unitary non-equivalent Wigner distributions), we construct a mean indicator of classicality/quantumness which gives a representation independent characteristic of classicality.

References

[1] Abbasli, N., Abgaryan, V., Bures, M, Khvedelidze, A., I. Rogojin, A. Torosyan, *On measures of classicality/quantumness in quasiprobability representations of finite-dimensional quantum systems*, Phys. Part. Nuclei 51, 443–447 (2020), preprint `http://arxiv.org/abs/2001.03737`.

[2] Abgaryan, V., Khvedelidze, A., *On the family of Wigner functions for N-level quantum system*, (2020), preprint `https://arxiv.org/abs/1708.05981`

Nurlan Abbasly
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
Institute of Physics
Azerbaijan National Academy of Sciences
Baku, Azerbaijan
e-mail: nurlan.camaloglu@gmail.com

Vahagn Abgaryan
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: vahagnab@googlemail.com

Martin Bureš
Institute of Experimental and Applied Physic
Czech Technical University in Prague
Prague, Czech Republic
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: bures@physics.muni.cz

Arsen Khvedelidze
A Razmadze Mathematical Institute
Iv. Javakhishvili, Tbilisi State University
Tbilisi, Georgia
Institute of Quantum Physics and Engineering Technologies
Georgian Technical University
Tbilisi, Georgia
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: akhved@jinr.ru

Ilya Rogojin
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: virus-atl@inbox.ru

Astghik Torosyan
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: astghik@jinr.ru

Bar Code and involutiveness: Janet and Janet-like divisions

Michela Ceria

Abstract. Involutive monomial divisions have been introduced by Janet as a concept [11, 12, 13, 14], being formally defined by Gerdt and Blinkov [6, 7]. In this talk we focus on two such particular divisions, namely Janet and Janet-like divisions [8, 9], treating them by means of Bar Codes, diagrams representing properties of monomial/semigroup ideals.

Extended abstract

Involutive monomial divisions have been introduced by Janet as a concept [11, 12, 13, 14], being formally defined by Gerdt and Blinkov [6, 7] and used to compute *involutive bases*, Groebner bases particularly efficient to find.

Consider the polynomial ring $\mathcal{P} := \mathbf{k}[x_1, ..., x_n]$ and the semigroup $\mathcal{T} \subset \mathcal{P}$ of terms in $x_1, ..., x_n$. Let $U \subset \mathcal{T}$ be a finite set of terms. For each $t \in U$, Janet defines a set $M(t, U) \subset \{x_1, ..., x_n\}$ of *multiplicative variables*[1] for t and calls *cone* of t the set $C(t, U) = \{tx_1^{\lambda_1} \cdots x_n^{\lambda_n} \,|\, \text{where } \lambda_j \neq 0 \text{ only if } x_j \in M(t, U)\}$; t is called *involutive divisor* for all the terms in $C(t, U)$ and only for them.

All cones are defined to be disjoint and Janet introduces a procedure, called *completion*, to enlarge U to a new set U' so that, called $\mathsf{T}(U)$ the semigroup generated by U,

$$\mathsf{T}(U) = \mathsf{T}(U') = \sqcup_{t \in U'} C(t, U').$$

If $U = U'$, then U is *complete*.

For an ideal $I = (f_1, ..., f_r) \lhd \mathcal{P}$, we consider the set of leading terms for its generators, namely $U = \{\mathsf{T}(f_1), ..., \mathsf{T}(f_r)\}$ and set Janet division on U, supposing U to be complete. A term $t \in \mathcal{T}$ is reducible by means of a generator f_i if and only if $t \in C(\mathsf{T}(f_i), U)$. This reduction procedure is that used to compute Janet involutive bases.

[1] The variables that are not multiplicative for t with respect to U form the set $NM(t, U)$ of non-multiplicative variables.

Janet-like division, defined in [8, 9], is a non-involutive generalization of Janet division, sharing many properties with the latter one. Instead of being based on the concept of multiplicative/non-multiplicative variables, it is based on non-multiplicative powers ($NM(t,U)$): the cone of a term t is given by the set of its multiples that are not divisible by a non-multiplicative power.

Bar Codes, introduced in [1, 2] are diagrams representing finite sets of terms and used in a series of papers in order to study the properties of monomial ideals. They are defined as follows.

Definition 0.1 ([1, 2]). A Bar Code B is a picture composed by segments, called *bars*, superimposed in horizontal rows, which satisfies conditions $a., b.$ below. Denote by

- $\mathsf{B}_j^{(i)}$ the j-th bar (from left to right) of the i-th row (from top to bottom), $1 \leq i \leq n$, i.e. the *j-th i-bar*;
- $\mu(i)$ the number of bars of the i-th row
- $l_1(\mathsf{B}_j^{(1)}) := 1$, $\forall j \in \{1, 2, ..., \mu(1)\}$ the $(1-)$*length* of the 1-bars;
- $l_i(\mathsf{B}_j^{(k)})$, $2 \leq k \leq n$, $1 \leq i \leq k-1$, $1 \leq j \leq \mu(k)$ the *i-length* of $\mathsf{B}_j^{(k)}$, i.e. the number of i-bars lying over $\mathsf{B}_j^{(k)}$

a. $\forall i, j$, $1 \leq i \leq n-1$, $1 \leq j \leq \mu(i)$, $\exists! \bar{j} \in \{1, ..., \mu(i+1)\}$ s.t. $\mathsf{B}_{\bar{j}}^{(i+1)}$ lies under $\mathsf{B}_j^{(i)}$

b. $\forall i_1, i_2 \in \{1, ..., n\}$, $\sum_{j_1=1}^{\mu(i_1)} l_1(\mathsf{B}_{j_1}^{(i_1)}) = \sum_{j_2=1}^{\mu(i_2)} l_1(\mathsf{B}_{j_2}^{(i_2)})$; we will then say that *all the rows have the same length.*

It is possible to associate to a finite set of terms a Bar Code and, on the other side, a finite set of terms to every Bar Code. The association is made so that the exponents of the terms are related to their position in the Bar Code.

In this talk, we will see how to study Janet division and Janet-like division by means of the Bar Code.
In particular, for Janet division, we will see how to compute multiplicative variables, find the involutive divisor of a term and detect whether a given set U is complete with respect to Janet division.

Suppose $x_1 < x_2 < ... < x_n$, let $U \subset \mathcal{T} \subset \mathbf{k}[x_1, ..., x_n]$ be a finite set of terms and B the associated Bar Code.
First perform the following three steps:

a) $\forall 1 \leq i \leq n$, put a star symbol $*$ on the right of t $\mathsf{B}_{\mu(i)}^{(i)}$;

b) $\forall 1 \leq i \leq n-1$, $\forall 1 \leq j \leq \mu(i) - 1$ let $\mathsf{B}_j^{(i)}$ and $\mathsf{B}_{j+1}^{(i)}$ be two consecutive bars not lying over the same $(i+1)$-bar; put a star symbol $*$ between these two bars.

Proposition 0.2. [3] *Let $U \subseteq \mathcal{T}$ be a finite set of terms and let us denote by B_U its Bar Code. For each $t \in U$ x_i, $1 \leq i \leq n$ is multiplicative for t if and only if the i-bar $\mathsf{B}_j^{(i)}$ of B_U, over which t lies, is followed by a star.*

Proposition 0.3. [5] *Let $U \subseteq \mathcal{T}$ be a finite set of terms and B be its Bar Code. Let $t \in U$, $x_i \in NM(t,U)$ and $\mathsf{B}_j^{(i)}$ the i-bar under t.*
Let $s \in U$; it holds $s \mid_J x_i t$ if and only if

1. *$s \mid x_i t$*
2. *s lies over $\mathsf{B}_{j+1}^{(i)}$ and*
3. *for each j' appearing with nonzero exponent in $\frac{x_i t}{s}$ there is a star after the j'-bar under s.*

Theorem 0.4. [5] *Let $U \subseteq \mathcal{T}$ be a finite set of terms and B be its Bar Code. Then U is a complete set if and only if for each $t \in U$ and each $x_i \in NM(t,U)$, called $\mathsf{B}_j^{(i)}$ the i-bar under t, there exists a term $s \in U$ satisfying conditions $1,2,3$ of Proposition 0.3.*

The Bar Code equipped with stars, that we employ to compute multiplicative variables is a reformulation of Gerdt-Blinkov-Yanovich *Janet tree* [10], but in the (equivalent) presentation given by Seiler [15]. We will see more in detail the relation between the two diagrams.

For Janet-like division, we note that non-multiplicative powers are no more than powers of Janet non-multiplicative variables; analogously to Janet division we have

Proposition 0.5 ([4]). *Let $U \subseteq \mathcal{T}$ be a finite set of terms and let us denote by B_U its Bar Code. Let $t \in U$, x_i a Janet-nonmultiplicative variable, $\mathsf{B}_l^{(i)}$ the i-bar under t and t' any term over $\mathsf{B}_{l+1}^{(i)}$. Then for*

$$k_i = \deg_i(t') - \deg_i(t),$$

$x_i^{k_i}$ is a non-multiplicative power for t.

Theorem 0.6 ([4]). *Let $U \subset \mathcal{T}$ be a finite set of terms, B its Bar Code, $t \in U$, $p = x_i^{k_i} \in NMP(t,U)$ a nonmultiplicative power and $\mathsf{B}_j^{(i)}$ the i-bar under t. Let $s \in U$; $s \mid tp$ w.r.t. Janet-like division if and only if the following conditions hold:*

1. *$s \mid pt$*
2. *s lies over $\mathsf{B}_{j+1}^{(i)}$ and*
3. *$\forall j'$ such that $x_{j'} \mid \frac{pt}{s}$ either there is a star after the j'-bar under s or the nonmultiplicative power w.r.t. $x_{j'}$ has greater degree $deg_{j'}(\frac{pt}{s})$.*

Thanks to Theorem 0.6 completeness with respect to Janet-like division can be easily detected, since being complete only means that for each $t \in U$ and for each non-multiplicative power $p = x_i^{k_i}$ for t, there is a term s in U whose cone contains pt. Theorem 0.6 essentially says that we have to look for the term s in the next i-bar with respect that under t and check conditions 1,2,3.

References

[1] Ceria, M., *Bar Code for monomial ideals*, Journal of Symbolic Computation, Volume 91, March - April 2019, DOI: 10.1016/$j.jsc$.2018.06.012, 30-56.

[2] Ceria, M., *Bar code: a visual representation for finite set of terms and its applications.*, accepted for publication in Mathematics in Computer Science, 2019.

[3] Ceria, M., *Bar Code vs Janet tree.* Atti della Accademia Peloritana dei Pericolanti, Classe di Scienze Fisiche, Matematiche e Naturali VOL 97, NO 2 (2019) DOI:http://dx.doi.org/10.1478/AAPP.972A6

[4] Ceria, M., *Bar Code and Janet-like division*, submitted

[5] Ceria, M., *Applications of Bar Code to involutive divisions and a greedy algorithm for complete sets.*, submitted

[6] Gerdt V.P., Blinkov Y.A. *Involutive bases of Polynomial Ideals*, Mathematics and Computers in Simulation 45, DOI: 10.1016/S0378 − 4754(97)00127 − 4 543–560, 1998

[7] Gerdt V.P., Blinkov Y.A. *Minimal involutive bases*, Mathematics and Computers in Simulation 45, DOI: 10.1016/S0378 − 4754(97)00128 − 6 519–541, 1998

[8] Gerdt, Vladimir P., Yuri A. Blinkov. *Janet-like monomial division.*, International Workshop on Computer Algebra in Scientific Computing. DOI:10.1007/11555964_15. Springer, Berlin, Heidelberg, 2005.

[9] Gerdt, Vladimir P., Yuri A. Blinkov. *Janet-like Groebner bases.*, International Workshop on Computer Algebra in Scientific Computing. DOI: 10.1007/11555964_16 Springer, Berlin, Heidelberg, 2005.

[10] Gerdt V., Blinkov Y. and Yanovich D., *Construction of Janet Bases I. Monomial Bases*, DOI: 10.1007/978 − 3 − 642 − 56666 − $0_1$8 in Computer Algebra in Scientific Computing CASC 2001, 233-247.

[11] Janet M., *Sur les systèmes d'équations aux dérivées partelles*, J. Math. Pure et Appl., 3, (1920), 65-151.

[12] Janet M., *Les modules de formes algébriques et la théorie générale des systemes différentiels*, Annales scientifiques de l'École Normale Supérieure, 1924.

[13] Janet M., *Les systèmes d'équations aux dérivées partelles*, Gauthier-Villars, 1927.

[14] Janet M., *Lecons sur les systèmes d'équations aux dérivées partelles* , Gauthier-Villars,1929.

[15] Seiler, W.M., *Involution: The formal theory of differential equations and its applications in computer algebra*, DOI: 10.1007/978 − 3 − 642 − 01287 − 7 Vol.24, 2009, Springer Science & Business Media

Michela Ceria
Department of Computer Science University of Milan Milano, Italy
e-mail: michela.ceria@gmail.com

Groebner bases and error correcting codes: from Cooper Philosophy to Degrobnerization

Michela Ceria, Teo Mora and Massimiliano Sala

Abstract. In the Late Nineties, the classical approach to decode BCH codes based on Berlekamp's *key equation* was upsetted by the application of Gröbner bases to the problem; it appeared a series of papers which terminated with two different proposals: Orsini-Sala general error locator polynomial [14] and Augot *et al.* Newton-Based decoder [1]; both approaches payed not only the hard pre-computation of a Gröbner basis but (mainly) the density of their decoders.

A recent work-in-progress [4, 5, 6, 7] reconsidered the same problem within the frame of *Grobner-free Solving*, an approach aiming to avoid the computation of a Gröbner basis of a (0-dimensional) ideal $J \subset \mathcal{P}$ in favour of combinatorial algorithms, describing instead the structure of the algebra $\mathcal{P}/J$. The consequence is a preprocessing which is quadratic (and a decoding which is linear) on the length of the code.

Extended abstract

In 1990 Cooper [10, 11] suggested to use Gröbner bases' computation in order to decode cyclic codes. Let C be a binary BCH code correcting up to t errors, $\bar{s} = (s_1, \ldots, s_{2t-1})$ be the syndrome vector associated to a received word. Cooper's idea consisted in interpreting the error locations $z_1, \ldots z_t$ of C as the roots of the syndrome equation system: $f_i := \sum_{j=1}^{t} z_j^{2i-1} - s_{2i-1} = 0,\ 1 \leq i \leq t$, and, consequently, the plain error locator polynomial as the monic generator $g(z_1)$ of the principal ideal $\left\{\sum_{i=1}^{t} g_i f_i, g_i \in \mathbb{F}_2(s_1, \ldots, s_{2t-1})[z_1, \ldots, z_t]\right\} \bigcap \mathbb{F}_2(s_1, \ldots, s_{2t-1})[z_1]$, which was computed via the elimination property of lexicographical Gröbner bases.

In a series of papers Chen et al. improved and generalized Cooper's approach to decoding. In particular, for a q-ary $[n, k, d]$ cyclic code, with correction capability t, they made two alternative proposals.

First of all, denoting, for an error with weight μ, $z_1, \ldots, z_\mu$ the error locations, $y_1, \ldots, y_\mu$ the error values and $s_1, \ldots, s_{n-k} \in \mathbb{F}_{q^m}$ the associated syndromes,

they interpreted [8] the coefficients of the plain error locator polynomial as the elementary symmetric functions σ_j and the syndromes as the *Waring functions*, $s_i = \sum_{j=1}^{\mu} y_j z_j^i$. They suggested to deduce the σ_j's from the (known) s_i's via a Gröbner basis computation for the ideal generated by the Newton identities; a similar idea was later developed in [1].

Alternatively, they considered [9] the *syndrome variety*

$$\left\{(s_1,\ldots,s_{n-k},y_1,\ldots,y_t,z_1,\ldots,z_t)\in(\mathbb{F}_{q^m})^{n-k+2t} : s_i=\sum_{j=1}^{\mu} y_j z_j^i,\ 1\le i\le n-k\right\}$$

and proposed to deduce, via a Gröbner basis pre-computation in

$$\mathbb{F}_q[x_1,\ldots,x_{n-k},y_1,\ldots,y_t,z_1,\ldots,z_t],$$

a series of polynomials $g_\mu(x_1,\ldots,x_{n-k},Z),\mu\le t$ such that, for any error with weight μ and associated syndromes $s_1,\ldots,s_{n-k}\in\mathbb{F}_{q^m}$, $g_\mu(s_1,\ldots,s_{n-k},Z)$ in $\mathbb{F}_{q^m}[Z]$ is the plain error locator polynomial.

Their approach was improved in a series of papers which introduced further applications of groebnerian technologies and which culminated with [14] which stated

Theorem 0.1. [14] *In the Gröbner basis of the ideal vanishing in each point of the syndrome variety, there is a unique polynomial, the* general error locator polynomial*, with shape*

$$g = z_t^t + \sum_{l=1}^{t} a_{t-l}(s_1,\ldots,s_{n-k}) z_t^{t-l}.$$

Such polynomial satisfies the following property: given a syndrome vector $s=(s_1,\ldots,s_{n-k})\in(\mathbb{F}_{q^m})^{n-k}$ corresponding to an error with weight $\mu\le t$, its t roots are the μ error locations plus zero counted with multiplicity $t-\mu$.

For a survey of *Cooper Philosophy* see [13], see [3] for Sala-Orsini locator.

Recently the same problem has been reconsidered in a group of papers [4, 6, 5] within the frame of *Grobner-free Solving*, an approach aiming to avoid the Gröbner bases computation for (0-dimensional) ideals.

In particular, given the syndrome variety

$$\mathsf{Z}=\left\{(c+d,c^3+d^3,c,d), c,d\in\mathbb{F}_{2^m}^*, c\neq d\right\}$$

of a BCH $[2^m-1,2]$-code C over $\mathbb{F}_{2^m}$, and denoted $\mathcal{I}(\mathsf{Z})$ the ideal of points of Z, [4] is able with good complexity to produce, via Cerlienco-Mureddu Algorithm [2] and Lazard Theorem, the set $\mathbf{N}:=\mathbf{N}(\mathcal{I}(\mathsf{Z}))$ and proves that the related Gröbner basis has the shape

$$G=(x_1^n-1, g_2, z_2+z_1+x_1, g_4)$$

where (see [14]) $g_2=\frac{x_2^{\frac{n+1}{2}}-x_1^{\frac{n+1}{2}}}{x_2-x_1}=x_2^{\frac{n-1}{2}}+\sum_{i=1}^{\frac{n-1}{2}}\binom{\frac{n-1}{2}}{i}x_1^i x_2^{\frac{n-1}{2}-i}$ and $g_4=z_1^2-\sum_{t\in\mathbf{N}}c_t t$ is Sala-Orsini general error locator polynomial. Such result allowed [4]

to remark (applying Marinari-Mora Theorem) that, for decoding, it is sufficient to compute the polynomial, *half error locator polynomial* (HELP)

$$h(x_1, x_2, z_1) := z_1 - \sum_{t \in \mathbf{H}} c_t t \text{ where } \mathbf{H} := \{x_1^i x_2^j, 0 \leq i < n, 0 \leq j < \frac{n-1}{2}\}$$

which satisfies

$$h(c(1 + a^{2j+1}), c^3(1 + a^{3(2j+1)}), z_1) = z_1 - c, \text{ for each } c \in \mathbb{F}_{2^m}^*, 0 \leq j < \frac{n-1}{2},$$

the other error ca^{2j+1} been computable via the polynomial $z_2 + z_1 + x_1 \in G$ as $z_2 := x_1 - z_1 = (c + ca^{2j+1}) - c = ca^{2j+1}$.

Such polynomial can be easily obtained with good complexity via Lundqvist interpolation formula [12] on the set of points

$$\left\{(c + ca^{2j+1}, c^3 + c^3 a^{3(2j+1)}, c), c \in \mathbb{F}_{2^m}^*, 0 \leq j < \frac{n-1}{2}\right\}.$$

Experiments showed that, in that setting, HELP has a very sparse formula, which has been proved (see [4]):

$$h(x_1, x_2, z_1) = z_1 + \sum_{i=1}^{\frac{n-1}{2}} a_i x_1^{(4-3i) \mod n} x_2^{(i-1) \mod \frac{n-1}{2}}$$

where the unknown coefficient can be deduced by Lundqvist interpolation on the set of points $\{(1 + a^{2j+1}, 1 + a^{3(2j+1)}, 1), 0 \leq j < \frac{n-1}{2}\}$ and on the terms $\{x_1^{(4-3i) \mod n} x_2^{(i-1) \mod \frac{n+1}{2}}, 1 \leq i < \frac{n+1}{2}\}$.

This suggested [6] to consider a binary cyclic code C over $GF(2^m)$, with length $n \mid 2^m - 1$ and *primary* defining set $S_C = \{1, l\}$. Thus it denoted by a a primitive $(2^m - 1)^{\text{th}}$ root of unity so that $\mathbb{F}_{2^m} = \mathbb{Z}_2[a]$, $\alpha := \frac{2^m - 1}{n}$ and $b := a^\alpha$ a primitive n^{th} root of unity, $\mathcal{R}_n := \{e \in \mathbb{F}_{2^m} : e^n = 1\}$ and $\mathcal{S}_n := \mathcal{R}_n \sqcup \{0\}$; considered the following sets of points

$$\begin{aligned}
&\mathsf{Z}_2 := \{(c + d, c^l + d^l, c, d), c, d \in \mathcal{R}_n, c \neq d\}, \#\mathsf{Z}_2^\times = n^2 - n;\\
&\mathsf{Z}_+ := \{(c + d, c^l + d^l, c, d), c, d \in \mathcal{S}_n, c \neq d\}, \#\mathsf{Z}_+^\times = n^2 + n,\\
&\mathsf{Z}_{ns} := \{(c+d, c^l + d^l, c, d), c, d \in \mathcal{S}_n\} \setminus \{(0, 0, c, c), c \in \mathcal{R}_n\}, \#\mathsf{Z}_{ns}^\times = n^2 + n + 1,\\
&\mathsf{Z}_e := \{(c + d, c^l + d^l, c, d), c, d \in \mathcal{S}_n\}, \#\mathsf{Z}_e^\times = (n + 1)^2,
\end{aligned}$$

and denoted, for $* \in \{e, ns, +, 2\}$,

- $J_* := \mathcal{I}(\mathsf{Z}_*)$, the ideal of all polynomials vanishing in Z_*,
- $\mathsf{N}_* := \mathbf{N}(J_*)$ the Gröbner escalier of J_* w.r.t. the lex ordering with $x_1 < x_2 < z_1 < z_2$ and
- $\Phi_* : \mathsf{Z}_* \to \mathsf{N}_*$ a Cerlienco-Mureddu correspondence [2].

Then it assumed to know

(a). the structure of the order ideal N_2, $\#\mathsf{N}_2 = n^2 - n$, i.e. a minimal basis $\{t_1, \ldots, t_r\}, t_i := x_1^{a_i} x_2^{b_i}$, of the monomial ideal $\mathcal{T} \setminus \mathsf{N}_2 = \mathbf{T}(\mathfrak{I}(\mathsf{Z}_2))$,

(b). a Cerlienco Mureddu Correspodence $\Phi_2 : \mathsf{N}_2 \to \mathsf{Z}_2$

and deduced with elementary arguments N_* and Φ_* for $* \in \{e, ns, +\}$.

References

[1] D. Augot, M. Bardet, J.C. Faugere, On formulas for decoding binary cyclic codes, *Proc. IEEE Int. Symp. Information Theory 2007*, (2007) .

[2] L. Cerlienco, M. Mureddu, *From algebraic sets to monomial linear bases by means of combinatorial algorithms*, Discrete Math. **139**, (1995) 73-87.

[3] F. Caruso, E. Orsini, C. Tinnirello and M. Sala *On the shape of the general error locator polynomial for cyclic codes* IEEE Transactions on Information Theory 63.6 (2017): 3641-3657.

[4] M. Ceria, T. Mora, M. Sala, *HELP: a sparse error locator polynomial for BCH codes*, in preparation.

[5] M. Ceria *Half error locator polynomials for efficient decoding of binary cyclic codes*, in preparation.

[6] M. Ceria, *Macaulay, Lazard and the Syndrome Variety*, in preparation.

[7] M.Ceria, T. Mora, M.Sala, *Zech Tableaux as tools for sparse decoding.* accepted for publications in Rendiconti del Seminario Matematico.

[8] X. Chen, I. S. Reed, T. Helleseth, K. Truong, Use of Gröbner Bases to Decode Binary Cyclic Codes up to the True Minimum Distance, *IEEE Trans. on Inf. Th.*, **40** (1994), 1654–1661.

[9] X. Chen, I. S. Reed, T. Helleseth, K. Truong, General Principles for the Algebraic Decoding of Cyclic Codes, *IEEE Trans. on Inf. Th.*, **40** (1994), 1661–1663.

[10] A.B. III Cooper, Direct solution of BCH decoding equations, In E. Arikan (Ed.) *Communication, Control and Singal Processing*, 281–286, Elsevier (1990)

[11] A.B. III Cooper, Finding BCH error locator polynomials in one step *Electronic Letters*, **27** (1991) 2090–2091

[12] Lundqvist S., *Vector space bases associated to vanishing ideals of points.* J. Pure Appl. Algebra **214** (2010), 309-321.

[13] E. Orsini, T. Mora,. *Decoding cyclic codes: the Cooper Philosophy.* in M.Sala et al., *Groebner Bases, Coding, and Cryptography.* Springer (2009), 62–92

[14] E. Orsini, M. Sala, *Correcting errors and erasures via the syndrome variety*, J. Pure Appl. Algebra, **200** (2005), 191–226.

Michela Ceria
Department of Computer Science University of Milan Milano, Italy
e-mail: michela.ceria@gmail.com

Teo Mora
Department of Mathematics University of Genoa Genoa, Italy
e-mail: theomora@disi.unige.it

Massimiliano Sala
Department of Matheatics University of Trento Trento, Italy
e-mail: maxsalacodes@gmail.com

Subexponential–time computation of isolated primary components of a polynomial ideal

Alexander L. Chistov

Let k be a field of arbitrary characteristic with an algebraic closure $\overline{k}$. Let H be a primitive subfield of the field k and $H(t_1, \ldots, t_l)$ be the field of rational functions over H in algebraically independent variables $t_1, \ldots, t_l$ over H. We assume that the field k is a finite separable extention $H(t_1, \ldots, t_l)$ given by its primitive element θ over $H(t_1, \ldots, t_l)$ (a minimal polynomial $\Phi \in H(t_1, \ldots, t_l)[Z]$ of the element θ is also given).

Let $f_1, \ldots, f_m \in k[X_1, \ldots, X_n]$ be polynomials of degree at most d where $d \geqslant 2$ is an integer. Denote by $I \subset \overline{k}[X_1, \ldots, X_n] = A$ the polynomial ideal generated by the polynomials $f_1, \ldots, f_m$. We suggest a simple algorithm for computing all the isolated primary components of the ideal I. At the output of his algorithm they are given up to embedded components.

More precisely, let $\mathfrak{p}$ be an arbitrary isolated associated prime ideal of the ideal I and $I_{\mathfrak{p}}$ be the $\mathfrak{p}$-primary component of the ideal I. Then for each $\mathfrak{p}$ we construct the field of fractions $K_{\mathfrak{p}}$ of the ring $A/\mathfrak{p}$, and the following objects.

1) A polynomial ideal $I'_{\mathfrak{p}} \subset \overline{k}[X_1, \ldots, X_n]$ such that $I_{\mathfrak{p}}$ is a unique isolated primary component of $I'_{\mathfrak{p}}$. The ideal $I'_{\mathfrak{p}}$ is given by its system of generators.
2) A finite dimensional $K_{\mathfrak{p}}$-algebra $K_{\mathfrak{p}} \otimes_A (A/I_{\mathfrak{p}})$. This algebra is given by its basis over $K_{\mathfrak{p}}$ and the multiplication table. Hence $I_{\mathfrak{p}}$ coincides with the kernel of the natural homomorphism $A \to K_{\mathfrak{p}} \otimes_A (A/I_{\mathfrak{p}})$.

Denote by $V = \mathcal{Z}(f_1, \ldots, f_m)$ the algebraic variety of all common zeroes of the polynomials $f_1, \ldots, f_m$ in $\mathbb{A}^n(\overline{k})$. Notice that the homomorphism $A/I \to K_{\mathfrak{p}} \otimes_A (A/I_{\mathfrak{p}})$ for a primary ideal $I_{\mathfrak{p}}$ is an analog of the generic point $\overline{k}[V] \to K_{\mathfrak{p}}$ of the irreducible component $\mathcal{Z}(\mathfrak{p})$ of the algebraic variety V.

The complexity of this algorithm is polynomial in d^{n^2} and the size of the input data. It seems that so far there has not been explicit estimates for the complexity of algorithms for this problem in the considered general situation.

Notice that that the varieties $V(\mathfrak{p})$ might be of distinct dimensions. To substantiate this algorithm we use a non–trivial estimation for the degrees of primary

components of the ideal $\sum_{1\leqslant i\leqslant m} Af_i$ in the case of homogeneous polynomials f_i obtained by the author earlier [1].

References

[1] A.L. Chistov, *Inequalities for Hilbert functions and primary decompositions*, Algebra i Analiz v. 19 (2007) No. 6, p. 143–172 (in Russian) [English transl.: St. Petersburg Math. J., v. 19 (2008) No. 6, p. 975–994].

Alexander L. Chistov
St. Petersburg Department of Steklov Mathematical Institute
of the Academy of Sciences of Russia
Fontanka 27, St. Petersburg 191023, Russia
e-mail: alch@pdmi.ras.ru

Symbolic Dynamics in Analyzing the Complexity of Trajectories of Triple Black Holes.

Ariel Chitan, Aleksandr Mylläri and Shirin Haque

Abstract. The effect of spin and distance unit on Burrau's classic 'problem of three bodies' was studied (Burrau 1913) using computer simulations and symbolic dynamics.In 1995, Valtonen et al. re-analysed this problem by motivation of improved computing technology. Three point masses were placed at the vertices of a 3,4,5 Pythagorean triangle such that the masses of each of the bodies reflected the side of the triangle opposite to it. The mass unit was then varied from 10^5 to 10^9 $M_\odot$ in increasing factors of ten. This set up was recreated and in each of the cases the first black hole was selected to be Kerr with spin vector J= [0.1 0.1 0.1], whilst the other two were kept as non-rotating Schwarzschild black holes. These five cases were each studied with three different distance units: 0.1 parsec, 1 parsec and 10 parsecs. Numerical orbital integration was done on FORTRAN using the algorthimic regularization code, ARCcode provided by Prof. Seppo Mikkola with relativistic corrections (pN) up to 2.5th order (Mikkola, Merritt 2008). Parameters studied were lifetimes of the systems, number of mergers and the number of binary encounters. It was found that with the effect of spin, the lifetimes of the systems decreased as compared to the 1995 study and moved from the 2D planar case to the 3D one. The number of binary encounters decreased as mass increased and the lifetimes decreased as distance unit decreased.

References

[1] C. Burrau, *Numerische Berechnung eines Spezialfalles des Dreikörperproblems.* Astronomical Notes, 1913. https://doi.org/10.1002/asna.19131950602.

[2] M. J., Valtonen and S., Mikkola and H., Pietilä, *Burrau's three-body problem in the post-Newtonian approximation*, The Astronomical Journal, 1995. https://doi.org/10.1093/mnras/273.3.751.

[3] S., Mikkola and D., Merritt , *Implementing Few-Body Algorithmic Regularization with post-Newtonian Terms*, The American Astronomical Society, 2008. https://doi.org/10.1088/0004-6256/135/6/2398.

Ariel Chitan
Physics Department
University of the West Indies
Trinidad, W.I.
e-mail: ariel.chitan@my.uwi.edu

Aleksandr Mylläri
Department of Computers and Technology
St. George's University
Grenada, W.I.
e-mail: amyllari@sgu.edu

Shirin Haque
Physics Department
University of the West Indies
Trinidad, W.I.
e-mail: shirin.haque@sta.uwi.edu

Work distribution between the student and the computer while solving tasks in distant contests.

Sergei Pozdniakov and Anton Chukhnov

Abstract. The work is devoted to distant contests held with the support of interactive environment. While creating a task of this type one always need to choose which activity is left fot the student and which one is left to the computer.

1. Background.

In [?] (based on the role of the tool in the child development [?, ?] and interiorization mechanism [?]) is shown that students' mastering algorithms with mathematical objects goes through several stages. At first, the algorithm is executed consciously, which corresponds to the concept of skill, then the algorithm is convolved, and the student, performing mental actions, actually works with the names of the algorithms, that is, convolutions of the algorithms that are deployed in the process of their application and remain minimized in the process of reasoning built on their basis. The study of the mechanisms of internalization in a computer environment is presented in the works of Papert [?] and his followers.

2. Interactive tasks in distant contests

While holding our contests (Construct, Test, Explore! and the Olympiad in Discrete Mathematics) we use a method of software supported subject tasks [?]. By subject task we mean a task with some understandable real-world statement that does not need any specific knowledge to understand it and to make at least first steps in the solution. By software supported we mean that a task is accompanied with a computer tool, that demonstrates the statement and allows for searching for a solution. Thus, such tasks have to be constructive, and a tool exposes their constructive nature. This tool supporting teh task we call a manipulator.

Within the framework of Olympiad in Discrete Mathematics and Theoretical Informatics we generally use six types of constructive tasks. Each of them is supported by its own manipulator. For CTE contest a new manipulator is developed for any new task.

For example, graph task may be formulated as "Find the minimal graph satisfying the certain conditions". Correctness of the constructed graph is verified automatically. The student can gain additional points for proving the minimality in the text form.

3. Internal and external operations.

While solving research problems, the student is faced with new concepts that, at the time of the occurrence of troubles, may not have representations known by the student. Creating a computer environment for operating with a subject environment allows you to hide (convolve) part of the algorithms that become a part of the environment and are executed automatically. At the same time, other operations are controlled by the student interacting with the environment. Thus, it becomes possible to control the mechanism of interiorization through carrying out the certain actions into the external environment.

4. Example of task. Network repair.

Ten buttons are somehow connected by elastic bands. Each button is removed from this web in turn along with the elastic bands connecting it to the rest of the network. The resulting ten meshes with nine buttons are shown at the top of the screen. Your goal is to reconstruct the original network.

This task is related to reconstruction conjecture which states that every graph with at least three vertices can be reconstucted by the multiset of its vertex-deleted subgraphs (called deck).

The program interface allows not only to clearly demonstrate the solution, but also to examine the deck obtained from its own graph along with the original deck, track the matching subgraphs of the decks, and arrange the elements of the deck in a convenient manner. It provides a meaningful process of experiments, not limited to simple guessing.

In this task the concept of isomorphism plays an auxiliary role, making it possible to quickly verify compliance with the requirements for the constructed solution. The participant has the opportunity to use the idea of isomorphism, moving the vertices of the subgraphs to better understand the reaction of the program. Programs are distinguished by a tool for working with graphs. Thus, the use of the concept of isomorphism in this task without its mathematical definition is compensated by the fact that the participant is given a tool that allows experimentally checking isomorphism, turning one representation of the graph into another.

This possibility of convolution of some skills and «deployment» of others, besides the possibilities of using the mechanism of interiorization, opens up possibilities for a more independent presentation of various ideas of mathematics.

5. Questions

How to properly divide the student's research activities into external operations performed by the student and internal operations performed by the computer?

Should all the algorithms that are convolved be brought outside the human mind?

The classical view on the formation of skills (in other words, the ability to use a set of algorithms that are not explicitly formulated, that is, in a formal algorithmic language) statess that when building a new skill based on the previous ones, the students should first master the previous skills until they are convolved mechanically.

At the same time you can use other methods of accessing basic entities. The latter can be immediately perceived by the learner as a real object that has understandable properties, and then on these entities you can build skills (algorithms) of a higher level, without reducing them to basic ones.

This happens in programming, almost none of the students can imagine the work of the algorithm in terms of electrical signals or even in assembler instructions but they still can use a computer and write programs.

6. Example of task. Steiner tree problem.

The task is about connecting a set of vertices on the plane by edges with possible adding another vertices. The goal is to minmize the total length of edges.

This task could be divided in two parts: the geometrical and the topological or graph one. The geometrical part of the solution is that the minimum is reached when every vertice added is a Torricelli point of a triangle connected with its vertices. The second part of the task has no effective algorithm to solve.

In our version of this task the geometrical part of the solution was performed by the computer: the student should just build the optimal topological configuration. It measn that two topologically equivalent solutions would be evaluated equally.

7. Conclusions

In case of graph tasks the possibility of combining algorithms for working with data in mathematical models is shown so that some operations are displayed externally, that is, they are interface elements and are controlled by the student, while others are presented in a minimized form, that is, these actions are performed automatically by modeling program.

Using various interpretations of mathematical concepts in simulation computer models, it becomes possible to form the skills of working with some algorithms. This makes it possible to more freely organize educational material when operations with complex mathematical concepts can be completed before the components of these concepts are studied in detail.

The use of the convolved algorithms in research problems with computer support allows us not to focus on the insignificant details of the plots and to pose new research problems that form a new mathematical intuition and which could not be posed without using computer support.

The work was supported by the Russian Foundation for Basic Research (Project No 18-013-01130).

References

[1] Maytarattanakon A., Posov I. A.: *Automation of distance contests based on research problems in mathematics and informatics.* Computer Tools in Education, vol. 6, 45–51, 2014. (in Russian).

[2] Papert S.: *Mindstorms: Children, Computers, and Powerful Ideas.* Basic Books, Inc., Publishers / New York, 1980.

[3] Akimushkin V. A., Pozdniakov S. N., Chukhnov A. S.: *Constructive Problems in the Structure of the Olympiad in Discrete Mathematics and Theoretical Informatics* Olympiads in Informatics, Vol. 11, pp. 3–18, 2017.

[4] Pozdniakov S., Posov I, Akimushkin V., Maytarattanakon A.: *The bridge from science to school* 10th IFIP World Conference on Computers in Education / WCCE 2013 Torun, 25 July 2013.

[5] Vygotsky L.: *Thought and Language* (Alex Kozulin, Trans. & Ed.). Cambridge, MA: MIT Press. , 1986. (Original work published in 1934)

[6] Vygotsky, L., Luria, A.: *Tool and symbol in child development* (in The Vygotsky Reader, edited by Jaan Valsiner and Rene van der Veer.), 1930

[7] Shapiro S. I.: *From algorithms to judgement.* M., Soviet radio, 288 p. (rus). (1973)

[8] Leontiev, A. N: *The Development of Mind*, a reproduction of the Progress Publishers 1981 edition, plus "Activity and Consciousness", originally published by Progress Publishers, 1977, published by Erythrospress, see Erythrospress.com 1977

Sergei Pozdniakov
Dept. of Algorithmic Mathematics
Saint Petersburg Electrothecnical University
Saint Petersburg, Russia
e-mail: pozdnkov@gmail.com

Anton Chukhnov
Dept. of Algorithmic Mathematics
Saint Petersburg Electrothecnical University
Saint Petersburg, Russia
e-mail: septembreange@gmail.com

Study of the Liénard Equation by Means of the Method of Normal Form

Victor F. Edneral

Abstract. The purpose of this report is the demonstration of searching the first integral of motion by the method of normal form. For the object of this demonstration, we chose the Liénard equation. We represented the equation as a dynamical system and parameterized it. After a calculation of the normal forms near stationary points, we found parameter values at which the condition of local integrability is satisfied for all stationary points simultaneously. We found two such sets of parameters. For each of them, the global integrability takes place.

Introduction

We use the approach based on the local analysis. It uses the resonance normal form calculated near stationary points [1]. In the paper [2] it was suggested the method for searching the values of parameters at which the dynamical system is locally integrable in all stationary points simultaneously. Satisfying the such local integrability conditions is a necessary condition of a global integrability. For global integrability of autonomous planar system, it is enough to have one global integral of motion. From its expression, you can get the solution of the system in quadratures. That is the integrability always leads to solvability. Note, that the converse is not true. Note also that a record of solutions via corresponding integrals is often more compact than an expression of the solution itself.

Problem

We will check our method on the example of the Liénard equation [3]

$$\ddot{x} = f(x)\dot{x} + g(x) = 0, \tag{1}$$

The publication has been prepared with the support of the "RUDN University Program 5-100".

which can be rewritten as a dynamical system. Let $f(x)$ be a quadratic polynomial and $g(x)$ is polynomial of fourth order. Then equation (1) is equivalent to the system

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

here parameters $a_0, a_1, a_2, b_1, b_2, b_3, b_4 \in \mathbb{R}, b_1 \neq 0$.

Method

The main idea of the discussed method is a search of conditions on the system parameters at which this system is locally integrable near its stationary points. The local integrability means we have enough number (one here) of the local integrals which are meromorphic near each stationary point. Local integrals can be different for each such point, but for the existence of the global integral, the local integrals should exist in all stationary points. This is a necessary condition. We have an algebraic condition for local integrability. It is the condition **A** [1, 2]. We look for sets of parameters at which the condition **A** is satisfied at all stationary points simultaneously. Such sets of parameters are good candidates for the existence of global integrals. These integrals we look for by other methods.

Condition of Local Integrability

The condition **A** is some infinite sequence of polynomial equations in coefficients of the system. Near each of the stationary points is its own equations. The condition of global integrability is a unification of these infinite systems of polynomial equations. Any part of this infinite system is a necessary condition of the integrability. We solve a finite subset of these equations. The condition **A** is formulated in terms of the normal form of the system. We calculated the resonance normal form for the system (2) near the stationary point in the origin using the MATHEMATICA 11 system and the program [4] till the 8^{th} order. After that, we wrote down the lowest equations (till the order of the eight) of the local integrability condition **A**. We solved this finite subsystem and got three sets of parameters.

1. $a_0 = 0, a_1 = 0, a_2 = 0$;
2. $a_0 = 0, a_1 b_2 = a_2 b_1, b_3 = 0, b_4 = 0$;
3. $a_0 = 0, a_2 = 0, b_2 = 0, b_4 = 0$.

Then we checked the condition of local integrability **A** near other stationary points. The third set above does not satisfy the local condition near some of the stationary points, so this is not a candidate for the global integrability.

First Integrals of Motion

For searching for global integrals, we divided the left and right sides of equations (2) into each other for each from the sets above. In result we had the first-order differential equations for $x(y)$ or $y(x)$. Then we solved them by the MATHEMATICA solver and got cumbersome solutions $y(x)$. After that we calculated the integrals from these solutions extracting the integration constants $I(x(t), y(t)) = const$. For the first set of parameters above, we got

$$I_1(x(t), y(t)) = 30\, y(t)^2 - 30\, b_1 x^2(t) - 20\, b_2 x^3(t) - 15\, b_3 x^4(t) - 12\, b_4 x^5(t). \quad (3)$$

Its time derivative $dI_1(t)/dt = 0$ along the system (2) over all phase space. So, it is the first integral.

For the second set we got

$$I_2(x(t), y(t)) = 6\, b_1^2\, \log[b_1 + a_1 y(t)] - 6\, a_1 b_1 y(t) + a_1^2\, x^2(t)\, [3\, b_1 + 2\, b_2 x(t)], \quad b_1 \neq 0. \quad (4)$$

The time derivative $dI_2(t)/dt = 0$ along the system (2). The limitation on the positivity of the argument of the logarithm can be eliminated by the representation of integral I_2 in the form $I = \exp(I_2)$. Of course, later additional studying analytical properties of this integral and the phase picture of the system should be carried . Nevertheless, we have here the first integral.

Scheme

The proposed method is intended for a search of the values of parameters at which some a polynomial autonomous dynamical system is integrable. The main steps of the method are:

- calculation of the normal form at stationary points of the system ;
- calculation finite subsets of equations of **A** condition at these points;
- the solution of these subsets of the equations in system parameters at the one (or more) stationary points;
- verifying the fulfillment of the condition **A** for the found parameter sets at the rest stationary points.
- the found parameter sets are used for searching integrals of the system by other methods.

Of course, there are different tactics. For example, in the current paper, we solved the set of equations near the origin, then tried found integrals, and only after that checked the condition **A** at other stationary points. Our practice work with planar autonomous systems demonstrates that at such values of parameters corresponding integrals exist for most sets of solutions of **A**.

Note, the proposed method of searching suitable parameters has no limitation on a system dimension. The limitations arise on the stage of searching integrals of motion.

Result

We found integrability at two sets of parameters. Set (1) corresponds to equation (1) in the form

$$\ddot{x} = b_1 x + b_2 x^2 + b_3 x^3 + b_4 x^4.$$

The corresponding integral is (3).
Set (2) corresponds to equation (1) in the form

$$\ddot{x} = x\,(b_1 + b_2 x)\,(1 + a\,x),$$
or
$$\ddot{x} = b_1 x + b_2 x^2 + a\,(b_1 x + b_2 x^2)\,y,$$
$$a \equiv a_1/b_1, \quad b_1 \neq 0.$$

The corresponding integral is (4).

Conclusion

We represented the Liénard equation as a dynamical system and parameterized it in a polynomial form. For this system, we found two sets of parameters at which it has the first integrals of motion and solvable. Both cases are trivial from a point of view of studying of Liénard's equation (1). The first case corresponds to equality $f(x) = 0$, the second one to $f(x) \sim g(x)$. But the workability of the method was illustrated.

The proposed method of searching suitable parameters has no limitation on a system dimension. The limitations arise on the stage of searching integrals of motion.

References

[1] A.D. Bruno, *Analytical form of differential equations (I,II).* Trudy Moskov. Mat. Obsc. **25**, 119–262 (1971), **26**, 199–239 (1972) (in Russian) = Trans. Moscow Math. Soc. **25**, 131–288 (1971), **26**, 199–239 (1972) (in English) A.D. Bruno, *Local Methods in Nonlinear Differential Equations.* Nauka, Moscow 1979 (in Russian) = Springer-Verlag, Berlin (1989) P.348.

[2] V.F. Edneral, *About integrability of the degenerate system.* Computer Algebra in Scientific Computing (CASC 2019), M. England et al. Lecture Notes in Computer Science, **11661**. Springer International Publishing, Springer Nature, Switzerland AG (2019)140–151.

[3] A. Liénard, Etude des oscillations entretenues, Revue générale de l'électricité **23** (1928) 901–912 and 946–954.

[4] V.F. Edneral, R. Khanin, *Application of the resonance normal form to high-order nonlinear ODEs using Mathematica.* Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **502**(2–3) (2003) 643–645.

Method of Normal Form

Victor F. Edneral
Skobeltsyn Institute of Nuclear Physics,
Lomonosov Moscow State University
Moscow, Russian Federation
Peoples′ Friendship University of Russia (RUDN University)
Moscow, Russian Federation
e-mail: edneral@theory.sinp.msu.ru; edneral_vf@pfur.ru

Implementation of algebraic algorithms for approximate pattern matching on compressed strings

Maria Fedorkina and Alexander Tiskin

1. Pattern matching and the LCS problem

Approximate matching is a natural generalization of classical (exact) pattern matching, allowing for some character differences between the pattern and a matching substring of the text. Given a pattern string p of length m and a text string t of length $n \geq m$, approximate pattern matching asks for all the substrings of the text that are similar to the pattern. We consider the classical approach to string comparison based on the following numerical measure of string similarity:

Definition 1.1. *Let a, b be strings. The* longest common subsequence (LCS) score *$lcs(a, b)$ is the length of the longest string that is a subsequence of both a and b. Given strings a, b, the* LCS problem *asks for the LCS score $lcs(a, b)$.*

Definition 1.2. *Given strings a, b, the* semi-local LCS problem *asks for the LCS scores as follows:*

- *the whole a against every substring of b* (string-substring LCS)
- *every prefix of a against every suffix of b* (prefix-suffix LCS)
- *every suffix of a against every prefix of b* (suffix-prefix LCS)
- *every substring of a against the whole b* (substring-string LCS)

In particular, string-substring LCS is closely related to approximate pattern matching, where a short fixed pattern string is compared to various substrings of a long text string.

2. LCS and the sticky braid monoid

The algebraic approach to the semi-local LCS problem is based on the monoid of sticky braids.

Definition 2.1. *The* sticky braid monoid *(a.k.a. the 0-Hecke Monoid of the symmetric group) of order n, denoted T_n is the monoid generated by the identity element ι and $n-1$ generators $g_1, \ldots, g_{n-1}$ defined by the relations:*

- $g_i^2 = g_i$ *for all i (idempotence)*
- $g_i g_j = g_j g_i$ *for all i, j, $j - i \geq 2$ (far commutativity)*
- $g_i g_j g_i = g_j g_i g_j$ *for all i, j, $j - i = 1$ (braid relations)*

where $i, j \in [1 : n-1]$.

The sticky braid monoid consists of $n!$ elements, which can be represented canonically by permutations of order n. An algorithm for multiplication of sticky braids (in permutation form) in time $O(n \log n)$ was given by the second author [4].

Intuitively speaking, the behavior of the LCS score under concatenation of strings is isomorphic to monoid multiplication of sticky braids. Therefore, it is possible to calculate the semi-local LCS of two strings a, b by partitioning one of the strings into two substrings and calling the algorithm recursively to obtain the semi-local LCS scores for each substring against the other string. The semi-local LCS scores in the subproblem and in the main problem are represented implicitly by sticky braids (in permutation form), and the subproblems are composed by sticky braid multiplication.

3. Grammar-compressed strings

Nowadays nearly all data used in science and technology are compressed. From an algorithmic viewpoint, it is natural to ask whether compressed strings can be processed efficiently without decompression. Early examples of such algorithms were given e.g. by Amir et al. [1] and by Rytter [3]; for a recent survey on the topic, see Lohrey [2]. Efficient algorithms for compressed strings can also be applied to achieve speedup over ordinary string processing algorithms for plain strings that are highly compressible.

The following generic compression model is well-studied, and covers many data compression formats used in practice.

Definition 3.1. *Let t be a string of length n. String t is said to be* grammar-compressed *if it is generated by a context-free grammar. A* context-free grammar *of length $\bar{n}$ is a sequence of $\bar{n}$ statements. A statement numbered k, $1 \leq k \leq \bar{n}$, has either the form $t_k = \alpha$ where α is an alphabet character, or the form $t_k = t_i t_j$ for some i, j, $1 \leq i, j < k$. For convenience we will also allow statements of the form $t_k = \epsilon$, where ϵ is the empty string.*

We will be discussing algorithms for the comparison of a plain (uncompressed) pattern string p of length m and a text string t of length n, compressed by a context-free grammar of length $\bar{n}$. The algorithm of Section 2 can be applied to perform approximate pattern matching efficiently in this setting.

The recursive nature of grammar compression makes it natural to apply the sticky braid approach. Since a statement produces a string that is the concatenation of two strings produced by previous statements, the calculation of the implicit semi-local LCS scores for the statement requires only multiplying the sticky braids corresponding to the previous statements using our implementation of the algorithm. The resulting algorithm takes $O(m\bar{n} \log m)$ time, as it makes $\bar{n}$ calls of the sticky multiplication subroutine, each running in time $O(m \log m)$.

4. Results

We have implemented the algorithm of [4]; to the best of our knowledge, it is the first existing implementation of this rather intricate algorithm. We have also implemented the algorithm of Section 3 calculating the semi-local LCS scores of a pattern p of length m and a text t compressed by a context-free grammar of length $\bar{n}$, and examined its performance on several examples of grammar-compressed strings.

Example 4.1. *The $\bar{n}$-th* Fibonacci string *is generated by the following context-free grammar:*

$$t_1 = B \quad t_2 = A \quad t_3 = t_2 t_1 \quad t_4 = t_3 t_2 \quad \ldots \quad t_{\bar{n}} = t_{\bar{n}-1} t_{\bar{n}-2}$$

E.g. the seventh Fibonacci string is "ABAABABAABAAB".

The length n of the $\bar{n}$-th Fibonacci string grows exponentially in $\bar{n}$. This suggests that our algorithm, running in time $O(m\bar{n} \log m)$ independent of n, should be substantially faster than the standard dynamic programming algorithm for calculating the LCS of two strings, running in time $O(mn)$.

We ran several experiments to examine the performance of our algorithm. In our experiments we generated the pattern strings randomly, drawing each character independently and equiprobably from the subset of letters of the Latin alphabet {'A', 'B', 'C'}. Using our algorithm, we calculated the LCS score for the pattern against a grammar-compressed Fibonacci string. We also calculated the LCS score for the pattern and the uncompressed Fibonacci string using dynamic programming and compared the resulting running times.

LCS calculation times (ms)				
Pattern length	Compressed text length	Uncompressed text length	Sticky braids (plain v. compressed)	Dynamic programming (plain v. uncompressed)
4	16	987	1	0
16	16	987	6	1
64	16	987	23	6
256	16	987	74	18
4	24	46368	2	15
16	24	46368	7	55
64	24	46368	29	211
256	24	46368	116	819
4	32	2178309	2	673
16	32	2178309	10	2550
64	32	2178309	38	9778
256	32	2178309	173	40124

We can see that even though dynamic programming performs better on short strings, the sticky braid algorithm starts to perform faster on longer strings. Additionally, the sticky braid algorithm keeps working even on larger Fibonacci strings that do not fit into the computer's memory uncompressed.

Fibonacci strings are an artificial construct that is not often used in practice. We now consider a more natural type of compression: the classical compression schemes LZ78 and LZW by Ziv, Lempel and Welch [6, 5].

Example 4.2. *The LZ78 and LZW compression schemes can both be represented a context-free grammar consisting of three sections:*

- *in the first section, all statements are of the form* $t_k = \alpha$*;*
- *in the second section, the first statement is of the form* $t_k = \epsilon$ *and all the following statements are of the form* $t_k = t_i t_j$*, where statement* i*,* $i < k$*, is from the second section, and statement* j *is from the first section;*
- *in the third section, the first statement is of the form* $t_k = \epsilon$ *and all the following statements are of the form* $t_k = t_{k-1} t_j$*, where statement* $k-1$ *is from the third section, and statement* j *is from the second section.*

We call context-free grammars of this form LZ-grammars.

The LZ-grammar corresponding to LZ78 or LZW compression might not be substantially shorter than the length of an uncompressed string. We construct a class of LZ-grammars corresponding to LZ78 compression that generates strings of length n growing quadratically in the grammar's length $\bar{n}$.

Example 4.3. *The* LZ78max-grammar *of length* $\bar{n} = 3r + 3$ *is an LZ-grammar defined as follows:*

$$\begin{array}{lll} t_0 = \epsilon & u_0 = \epsilon & v_0 = \epsilon \\ t_1 = \alpha_1 & u_1 = u_0 t_1 & v_1 = v_0 u_1 \\ t_2 = \alpha_2 & u_2 = u_1 t_2 & v_2 = v_1 u_2 \\ \dots & \dots & \dots \\ t_r = \alpha_n & u_r = u_{r-1} t_r & v_r = v_{r-1} u_r \end{array}$$

where α_k *is the* k*-th character of the alphabet.*
E.g. the LZ78max-grammar of length $18 = 3 \cdot 5 + 3$ *generates the string "AABABCABCDABCDE" of length* 15.

We ran several experiments to examine the performance of our algorithm. In our experiments we generated the pattern strings randomly, drawing each character independently and equiprobably from the uppercase letters of the Latin alphabet. Using our algorithm, we calculated the LCS score for the pattern against an LZ78max-grammar. We also calculated the LCS score for the pattern and the uncompressed LZ78max-grammar string using dynamic programming and compared the resulting running times.

LCS calculation times (ms)				
Pattern length	Compressed text length	Uncompressed text length	Sticky braids (plain v. compressed)	Dynamic programming (plain v. uncompressed)
4	195	2145	15	0
16	195	2145	48	2
64	195	2145	169	8
256	195	2145	687	35
4	1539	131841	94	39
16	1539	131841	339	142
64	1539	131841	1376	532
256	1539	131841	5774	2242
4	12291	8394753	763	2594
16	12291	8394753	2807	9599
64	12291	8394753	12282	35108
256	12291	8394753	49187	153342

Again, we see that even though dynamic programming performs better on short strings, the sticky braid algorithm starts to perform faster on longer strings.

5. Conclusion and future work

Our experiments demonstrate that the algebraic string comparison approach of [4] is not only of theoretical interest, but can also give substantial speedups on problems of practical significance, such as approximate matching on grammar-compressed strings, which includes the classical LZ78 and LZW compression schemes as a special case. Further work may involve generalising our implementation to deal with scoring schemes other than LCS (e.g. edit distance matching), and using it for efficient approximate pattern matching on large compressed datasets.

References

[1] A Amir, G Benson, and M Farach. Let Sleeping Files Lie: Pattern Matching in Z-Compressed Files. *Journal of Computer and System Sciences*, 52(2):299–307, 1996.

[2] M Lohrey. Algorithmics on SLP-compressed strings: a survey. *Groups Complexity Cryptology*, 4(2):241–299, 2012.

[3] W Rytter. Algorithms on Compressed Strings and Arrays. In *Proceeedings of SOFSEM*, volume 1725 of *Lecture notes in Computer Science*, pages 48–65, 1999.

[4] Alexander Tiskin. Fast Distance Multiplication of Unit-Monge Matrices. *Algorithmica*, 71:859–888, 2015.

[5] T A Welch. A Technique for High-Performance Data Compression. *Computer*, 17(6):8–19, 1984.

[6] G Ziv and A Lempel. Compression of individual sequences via variable-rate coding. *IEEE Transactions on Information Theory*, 24:530–536, 1978.

Maria Fedorkina
St. Petersburg School of Physics, Mathematics, and Computer Science
Higher School of Economics
St. Petersburg, Russia
e-mail: `s17b2_fedorkina@179.ru`

Alexander Tiskin
Dept. of Mathematics and Computer Science
St. Petersburg University
St. Petersburg, Russia
e-mail: `a.tiskin@spbu.ru`

On the Moments of Squared Binomial Coefficients

Nikita Gogin and Mika Hirvensalo

Abstract. Explicit recurrent formulas for ordinary and alternated power moments of the squared binomial coefficients are derived in this article. Every such moment proves to be a linear combination of the previous ones via a coefficient list of the relevant Krawtchouk polynomial.

Introduction

In this article, we study the sums of form

$$\mu_r^{(n)} = \sum_{m=0}^{n} m^r \binom{n}{m}^2 \text{ and } \nu_r^{(n)} = \sum_{m=0}^{n} (-1)^m m^r \binom{n}{m}^2, n \geq 0 \tag{1}$$

and refer them as the r-th order (where $r \geq 0$ and $0^0 = 1$ by convention) ordinary and alternating moments of the squared binomial coefficients.

Such sums emerge very often in different theoretical and applied mathematical areas, for example, see A000984, A002457, A037966, A126869, A100071, and A294486 in Sloane's database OEIS of integer sequences.

Unfortunately, up to the date (February 2020) not many explicit and closed formulas for these sums are known and moreover all such formulas are limited in order by $r \leq 4$, see [2], [3], [6], and A074334.

In the present article we prove two main theorems and their corollaries (in the Sections 3 and 4), providing the explicit recurrent formulas for obtaining the closed forms for the aforesaid moments of any order $r \geq 0$. We also give some examples of applications of these results.

1. Introduction

In this article, we study the sums of form

$$\mu_r^{(n)} = \sum_{m=0}^{n} m^r \binom{n}{m}^2 \text{ and } \nu_r^{(n)} = \sum_{m=0}^{n} (-1)^m m^r \binom{n}{m}^2, n \geq 0 \tag{2}$$

and refer them as the r-th order (where $r \geq 0$ and $0^0 = 1$ by convention) ordinary and alternating moments of the squared binomial coefficients.

Such sums emerge very often in different theoretical and applied mathematical areas, for example, see A000984, A002457, A037966, A126869, A100071, and A294486 in Sloane's database OEIS of integer sequences.

Unfortunately, up to the date (February 2020) not many explicit and closed formulas for these sums are known and moreover all such formulas are limited in order by $r \leq 4$, see [2], [3], [6], and A074334.

In the present article we prove two main theorems and their corollaries (in the Sections 3 and 4), providing the explicit recurrent formulas for obtaining the closed forms for the aforesaid moments of any order $r \geq 0$. We also give some examples of applications of these results.

2. Preliminaries

2.1. Hadamard transform, Krawtchouk polynomials and McWilliams duality formula

In this section, we remind to the reader some definitions and results from algebraic coding theory. The proofs of the claims can be found for example in [4].

For any integer $n \geq 1$, let $\mathbb{F}_2^n$ be an n-dimensional vector space over the binary field $\mathbb{F}_2 = \{0, 1\}$ and let V_n be the 2^n-dimensional Euclidean vector space of all functions $f\colon \mathbb{F}_2^n \to \mathbb{R}$ equipped with the usual scalar product $(f, g) = \sum_{\boldsymbol{v} \in \mathbb{F}_2^n} f(\boldsymbol{v}) g(\boldsymbol{v})$

Every additive character of the space $\mathbb{F}_2^n$ can be written in a form $\chi_{\boldsymbol{u}}(\boldsymbol{v}) = \chi_{\boldsymbol{v}}(\boldsymbol{u}) = (-1)^{\boldsymbol{v} \cdot \boldsymbol{u}}$, where the dot product is defined as $\boldsymbol{u} \cdot \boldsymbol{v} = \sum_{i=1}^{n} u_i v_i = |\boldsymbol{u} \cap \boldsymbol{v}|$, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are interpreted (in the obvious way) as subsets of indices $\{1, 2, ..., n\}$. Note that for the vector $\mathbf{1} = (1, 1, \ldots, 1) \in V_n$ the scalar product $(\boldsymbol{v}, \mathbf{1}) = |\boldsymbol{v} \cap \{1, 2, ..., n\}|$ is $|\boldsymbol{v}|$, the Hamming weight of $\boldsymbol{v}$.

It is a well-known fact that the full set of characters forms an orthogonal basis in the space V_n for which

$$(\chi_{\boldsymbol{u}}, \chi_{\boldsymbol{v}}) = 2^n \delta_{\boldsymbol{u}, \boldsymbol{v}} = 2^n \begin{cases} 1 & \text{if } \boldsymbol{u} = \boldsymbol{v} \\ 0 & \text{if } \boldsymbol{u} \neq \boldsymbol{v} \end{cases} \tag{3}$$

Definition 1. *For every function $f \in V_n$ we define the Hadamard transform as*

$$\widehat{f}(\boldsymbol{u}) = (f, \chi_{\boldsymbol{u}}) = \sum_{\boldsymbol{v} \in V_n} (-1)^{|\boldsymbol{v} \cap \boldsymbol{u}|} f(\boldsymbol{v}). \tag{4}$$

Note that $(\widehat{f}, \widehat{g}) = 2^n (f, g)$ and $\widehat{\widehat{f}} = 2^n f$ (unitary and involutory properties of Hadamard transform).

Define now for every $0 \le r \le n$ the r-th weight-function $\psi_r(\boldsymbol{v}) = \begin{cases} 1 & \text{if } |\boldsymbol{v}| = r \\ 0 & \text{if } |\boldsymbol{v}| \neq r \end{cases}$ and let $\mathfrak{H}_r^{(n)}$ be the r-th Hamming sphere in $\mathbb{F}_2^n$: $\mathfrak{H}_r^{(n)} = \{\boldsymbol{v} : \psi_r(\boldsymbol{v}) = 1\}$.

Definition 2. *For any $f \in V_n$ the $(n+1)$-tuple of real numbers $\mathfrak{S}(f) = (S_0(f), S_1(f), \ldots S_n(f))$, where $S_r(f) = (f, \psi_r) = \sum_{\boldsymbol{v} \in \mathfrak{H}_r^{(n)}} f(\boldsymbol{v})$ is said to be the weight spectrum of f while $\mathfrak{S}(\widehat{f})$ is referred as a dual weight spectrum of f .*

Due to the unitary and involutory properties of the Hadamard transform one has:

$$S_r(\widehat{f}) = (\widehat{f}, \psi_r) = 2^{-n}(\widehat{\widehat{f}}, \widehat{\widehat{\psi}}_r) = (f, \widehat{\psi}_r). \tag{5}$$

Functions $\widehat{\psi}_r(\boldsymbol{v})$ depend obviously only on $|\boldsymbol{v}|$ and this gives rise to the following definition:

Definition 3. *Function*

$$\widehat{\psi}_r(\boldsymbol{v}) = \widehat{\psi}_r(|\boldsymbol{v}|) = K_r^{(n)}(x) = \sum_{i=0}^{r} (-1)^i \binom{n-x}{r-i}\binom{x}{i} \tag{6}$$

where $x = |\boldsymbol{v}|$ is being called as the r-th Krawtchouk polynomial of order n; $0 \le r = \deg K_r^{(n)} \le n$.

Now we can easily rewrite (5) to obtain the famous MacWilliams formula for dual spectrae: $S_r(\widehat{f}) = \sum_{i=0}^{n} K_r^{(n)}(i) S_i(f)$.

2.2. Auxiliary lemmata

We will use these results in the sequel, but they may have independent combinatorial interest, as well.

Lemma 1. *For any nonnegative integer d, $d \ge 1$ let $g(\boldsymbol{v}) = \binom{d}{|\boldsymbol{v}|}$, where $\boldsymbol{v} \in V_n$. Then $\widehat{g}(\boldsymbol{u}) = K_d^{(n+d)}(|\boldsymbol{u}|), \quad \boldsymbol{u} \in V_n$*

Lemma 2. *For every integer $k, 0 \le k \le d$ the following identity is valid:*

$$\sum_{i=0}^{n} K_i^{(n)}(k) K_d^{(n+d)}(i) = 2^n \binom{d}{k}. \tag{7}$$

Corollary 1.

$$(-2)^n \sum_{i=0}^{n} \binom{\frac{i-1}{2}}{n} K_i^{(n)}(k) = \binom{n}{k}. \tag{8}$$

Lemma 3.

$$\sum_{m=0}^{n}\binom{n}{m}^2 K_j^{(n)}(m)=\binom{n}{j}K_n^{(2n)}(j) \tag{9}$$

for every nonnegative integer j.

Remark 1. *For $j=0$ we get from (9) the classical identity*

$$\mu_0^{(n)}=\sum_{m=0}^{n}\binom{n}{m}^2=\binom{2n}{n}. \tag{10}$$

3. Recurrent formula for the moments $\mu_j^{(n)}$ and its applications

For a fixed nonnegative integer j let $\left(\kappa_r^{(n,j)}\right)_{0\le r\le j}$ be a coefficient list of the polynomial $K_j^{(n)}$ and let $\boldsymbol{\mu}=(\mu^s)_{s\ge 0}$ be an *umbral* variable with a rule $\mu^s \rightleftarrows \mu_s^{(n)}, s\ge 0$.

Remark 2. *Descriptions of umbral calculus can be found in many sources. We recommend the reader to consult* [8]*, for instance. A more developed and formalized treatises are available at* [5] *and* [7].

In this article, we need only the elementary notion of umbral variable as a linear functional $\to$ on $\mathbb{C}[[\mu]]$ (the formal power series over μ) defined as $\to(\mu^s)=\mu_s^{(n)}$, $s\ge 0$.

Theorem 1.

$$K_j^{(n)}(\boldsymbol{\mu})=\binom{n}{j}K_n^{(2n)}(j). \tag{11}$$

Corollary 2. *(Recurrent formula for $\mu_j^{(n)}$)*

$$\frac{(-2)^j}{j!}\mu_j^{(n)}=\binom{n}{j}K_n^{(2n)}(j)-\sum_{r=0}^{j-1}\kappa_r^{(n,j)}\mu_r^{(n)},\quad j\ge 1, \mu_0^{(n)}=\binom{2n}{n}. \tag{12}$$

For example, for $j=6$ one can consecutively applying Corollary 2 find

$$\mu_6^{(n)}=\frac{n^3\left(n^6+3n^5-13n^4-15n^3+30n^2+8n-2\right)}{8(2n-1)(2n-3)(2n-5)}\binom{2n}{n}.$$

4. Recurrent formula for the alternating moments $\nu_j^{(n)}$ and their applications

The case of the alternated moments is in general similar to the previous one but is more subtle: Let $\boldsymbol{\nu}=(\nu^s)_{s\ge 0}$ be an umbral variable with a rule $\nu^s \rightleftarrows \nu_s^{(n)}, s\ge 0$.

Theorem 2.

$$K_j^{(n)}(\boldsymbol{\nu}) = \binom{n}{j} K_n^{(2n)}(n-j). \tag{13}$$

Corollary 3. *(Recurrent formula for $\nu_j^{(n)}$)*

$$\frac{(-2)^j}{j!}\nu_j^{(n)} = \binom{n}{j} K_n^{(2n)}(n-j) - \sum_{r=0}^{j-1} \kappa_r^{(n,j)} \nu_r^{(n)}, \quad j \geq 1, \tag{14}$$

$$\nu_0^{(n)} = \begin{cases} (-1)^{\frac{n}{2}} \binom{n}{n/2} & \textit{if } n \equiv 0 \pmod 2 \\ 0 & \textit{if } n \equiv 1 \pmod 2 \end{cases}. \tag{15}$$

For example, for $j = 6$, one can consecutively applying corollary 3 find:

$$\nu_6^{(n)} = \begin{cases} (-1)^{\frac{n+2}{2}} \frac{n^3(n+1)(3n-1)}{8} \binom{n}{n/2} & \text{if } n \equiv 0 \pmod 2 \\ \frac{(-1)^{\frac{n-1}{2}} n^2(n+1)(n^3+n^2-9n+3)}{8} \binom{n}{(n+1)/2} & \text{if } n \equiv 1 \pmod 2 \end{cases}. \tag{16}$$

References

[1] Gogin N., Hirvensalo M.: *Recurrent Construction of MacWilliams and Chebyshev Matrices*. Fundamenta Informaticae 116:1–4, pp. 93–110 (2012). DOI: 10.3233/FI-2012-671

[2] Gould H. W.: *Combinatorial Identities: A Standardized Set of Tables Listing 500 Binomial Coefficient Summations*. Morgantown Printing and Binding Co (1972).

[3] Graham R.L., Knuth D. E., Patashnik O.: *Concrete Mathematics*. Addison-Wesley (1994).

[4] MacWilliams F.J., Sloane N. J. A.: *The Theory of Error-Correcting Codes*. North Holland (1977).

[5] Licciardi, S.: *Umbral Calculus, A Different Mathematical Language*. `https://arxiv.org/abs/1803.03108` (2018) (retrieved February 15, 2020).

[6] Riordan J.: *Combinatorial Identities*. John Wiley & Sons Inc. (1968).

[7] Nigel Ray: *Universal Constructions in Umbral Calculus*. In: Bruce E. Sagan, Richard P. Stanley (Eds.): Mathematical Essays in honor of Gian-Carlo Rota, pp. 343–357. Springer (1996).

[8] Roman, S.M., Rota G.-C.: *The Umbral Calculus*. Advances in Mathematics 27, pp. 95–188 (1978).

Nikita Gogin

e-mail: `ngiri@list.ru`

Mika Hirvensalo
Department of Mathematics and Statistics
University of Turku
Turku, Finland
e-mail: `mikhirve@utu.fi`

Average number of solutions and mixed symplectic volume

B. Kazarnovskii

Abstract. The famous Koushnirenko-Bernstein theorem, also known as theorem BKK, asserts that the number of solutions of a polynomial system is equal to the mixed volume of Newton polyhedra of polynomials. This theorem creates the interaction platform for algebraic and convex geometries, which is useful in both directions. Here some statement is given that we view as a smooth version of the BKK theorem.

Introduction

Let X be an n-dimensional manifold, $V_1, \dots, V_n$ be finite dimensional vector subspaces in $C^\infty(X, \mathbb{R})$, and let V_i^* be their dual vector spaces. We consider the systems of equations

$$f_1 - a_1 = \dots = f_i - a_i = \dots = f_n - a_n = 0, \tag{1}$$

$f_i \in V_i$, $f_i \neq 0$, $a_i \in \mathbb{R}$. Let $H_i = \{v^* \in V_i^* \mid v^*(f_i) = a_i\}$ be an affine hyperplane in V_i^*, corresponding to equation $f_i - a_i = 0$, and $H = (H_1, \dots, H_n)$ be a tuple of hyperplanes, corresponding to system (1). For a measure Ξ on the set of tuples $(H_1, \dots, H_n)$, we define *the average number of solutions* as an integral of a number of zeroes of (1) with respect to Ξ.

We fix smooth Banach metrics in the spaces V_i. Further, we assume that the unit balls of these metrics are smooth and strictly convex bodies. We use these metrics for, firstly, to construct the measure Ξ and, secondly, to construct the Banach convex bodies $\mathcal{B}_i$ in X. Banach convex body or B-body in X is a collection $\mathcal{B} = \{\mathcal{B}(x) \subset T_x^*X\}$ of centrally symmetric convex bodies in the fibers of the cotangent bundle of X. We define the mixed symplectic volume of B-bodies and prove that the average number of zeroes is equal to the mixed symplectic volume of B-bodies $\mathcal{B}_1, \dots, \mathcal{B}_n$.

Main theorems

We choose the measure Ξ on the space of systems (1) equal to a product of Crofton measures in spaces V_i^*. Recall that a translation invariant measure on the Grassmanian of affine hyperplanes in Banach space is called a Crofton measure, if a measure of a set of hyperplanes, crossing any segment, equals to its length. Under certain smoothness conditions there exists a unique such measure.

The symplectic volume of $\bigcup_{x\in X}\mathcal{B}(x)\subset T^*X$ we call the volume of Banach body $\mathcal{B}=\{\mathcal{B}(x)\}$. Using Minkowski sum and homotheties, we consider linear combinations of convex sets with non-negative coefficients. The linear combination of B-bodies is defined by $(\sum_i\lambda_i\mathcal{B}_i)(x)=\sum_i\lambda_i\mathcal{B}_i(x)$. The volume of a linear combination $\mathrm{vol}(\lambda_1\mathcal{B}_1+\ldots+\lambda_n\mathcal{B}_n)$ is a homogeneous polynomial of degree n in $\lambda_1,\ldots,\lambda_n$.

Definition 1. *The coefficient of polynomial* $\mathrm{vol}(\lambda_1\mathcal{B}_1+\ldots+\lambda_n\mathcal{B}_n)$ *at* $\lambda_1\cdot\ldots\cdot\lambda_n$ *divided by* $n!$ *is called the mixed volume of* B*-bodies and is denoted by* $\mathrm{vol}(\mathcal{B}_1,\ldots,\mathcal{B}_n)$.

B-body corresponding to Banach space V of smooth functions on X appears as follows. Let $B\subset V$ be a unit Banach ball. Define the mapping $\theta\colon X\to V^*$, as $\theta(x)\colon f\mapsto f(x)$. Let $d_x\theta$ be a differential of θ at $x\in X$, and let $d_x^*\theta\colon V\to T_x^*$ be an adjoint operator. So we get a Banach body $\mathcal{B}=\{\mathcal{B}(x)=d_x^*\theta(B)\}$.

Let $U\subset X$ be an open set with compact closure. Denote by $\mathfrak{M}(U)$ the average number of solutions of (1) in U. Let $\mathcal{B}_i$ be a B-body in X, corresponding to the space of functions V_i, and let $\mathcal{B}_i^U$ be a restriction of $\mathcal{B}_i$ to U.

Theorem 1.

$$\mathfrak{M}(U)=\frac{n!}{2^n}\,\mathrm{vol}(\mathcal{B}_1^U,\ldots,\mathcal{B}_n^U),$$

where $\mathrm{vol}(\mathcal{B}_1^U,\ldots,\mathcal{B}_n^U)$ *is a mixed volume of* B*-bodies.*

The proof of the theorem is based on the calculations in the ring of normal densities constructed in [AK18].

However, the Crofton measure in Banach space is quite exotic. For example, if the unit ball in Banach space is not a zonoid, then the Crofton measure is not everywhere positive; see [SC06, K20]. Recall that the zonotope is a polyhedron, represented as the Minkowski sum of segments, and the zonoid is a limit of a sequence of zonotopes converging with respect to the Hausdorff metric. If the unit ball of the Banach metric is a zonoid, then we call this metric a *zonoid metric*. All ellipsoids are zonoids and, respectively, Euclidean metrics are zonoid metrics.

The non-positivity of the measure Ξ reduces the validity of the notion of average number of solutions. For this reason, in order to avoid the non-positivity of Crofton meaure, we consider the averaging process under the general families of positive measures on the manifolds of affine hyperplanes.

Theorem 2. *Let* V *be a finite dimensional vector space, and let* μ *be a translation invariant countably additive smooth positive measure on the manifold of affine hyperplanes in* V^**. Then there is the unique Banach metric* $\|\ \|^\mu$ *in* V*, such that* μ *is a Crofton measure of the dual Banach metric in* V^*.

Therefore, applying Theorem 1 for metrics corresponding to an arbitrary tuple of measures $\mu_1, \dots, \mu_n$ mentioned above, we obtain a version of the BKK *theorem for a general tuple of positive smooth measures.*

Remarks

Remark 1. *The metric* $\| \|^\mu$ *from Theorem* 2 *is a zonoid metric, and the corresponding* B*-body is a family of zonoids.*

Remark 2. *The case of Euclidean metrics in the spaces* V_i *was previously considered; see* [AK18, ZK14]. *In this case, the Banach bodies are ellipsoid families.*

Remark 3. *From Nash embedding theorem it follows that any smooth collection of ellipsoids in the fibers of* T^*X *can be obtained as* B*-body, corresponding to some Euclidean space of functions. If* X *is a compact manifold (may be with boundary), then from the Banach analogue of Nash theorem proved in* [BI94] *it follows that any collection* $\{\mathcal{B}(x)\colon\ x \in T^*_x X\}$ *of smooth strongly convex centrally symmetric bodies* $\mathcal{B}(x)$ *is a* B*-body, corresponding to some Banach space of functions on* X.

Remark 4. *Let* $\mu_1, \dots, \mu_n$ *be smooth translation invariant (not necessarily positive) measures on the manifolds of affine hyperplanes in* $V_1^*, \dots, V_n^*$ *respectively. Then the corresponding average number of solutions is equal to the mixed symplectic volume of some uniquely defined virtual Banach zonoids. Note that an arbitrary smooth centrally symmetric convex body is a virtual zonoid.*

References

[AK18] D. Akhiezer, B. Kazarnovskii, *Average number of zeros and mixed symplectic volume of Finsler sets*, Geom. Funct. Anal., vol. 28 (2018), pp.1517–1547.

[BI94] D. Yu. Burago, S. Ivanov, *Isometric embeddings of Finsler manifolds*, Algebra i Analiz 5 (1993), no. 1, 179 – 192 (Russian, with Russian summary); English transl., St. Petersburg Math. J. 5 (1994), no. 1, 159 – 169.

[K20] B. Kazarnovskii, *The average number of solutions.* To appear in Funct. Anal. Appl.

[SC06] Rolf Schneider, *Crofton measures in projective Finsler Spaces*, pp. 67–98, in Integral Geometry and Convexity, Wuhan, China, 18 – 23 October 2004, Edited By: Eric L Grinberg, Shougui Li, Gaoyong Zhang and Jiazu Zhou

[ZK14] D. Zaporozhez, Z. Kabluchko, *Random determinants, mixed volumes of ellipsoids, and zeros of Gaussian random fields*, Journal of Math. Sci., vol. 199, no.2 (2014), pp. 168–173.

B. Kazarnovskii
Institute for Information Transmission Problems
Moscow, Russia
e-mail: kazbori@gmail.com

Eigenvalues and eigenvectors for the composition of Lorentz boosts in concise form

Mikhail Kharinov

Abstract. This paper considers the formal eigenvalue/eigenvector problem for Lorentz transformation $\mathcal{L}$ in the real four-dimensional spacetime $\mathbb{R}^4$. According to the problem statement, it is required to find a quartet of linearly independent eigenvectors for the composition $\mathcal{L} = L_1 L_2$ of the boosts L_1 and L_2. To analytically find the eigenvalues, a fourth-degree polynomial characteristic equation is obtained and solved. The a priori expected concise expressions for the eigenvectors are presented.

Introduction

This work completes the phase of the study of general Lorentz transformations, begun in [1] and continued in [2]. In [1], a special case of Lorentz boost composition was not considered. In [2], the expression for the eigenvectors of the boost composition turned out to be too cumbersome. The latter disadvantage is overcome in this paper.

The Lorentz transformations $\mathcal{L}$ are defined as a linear homogeneous transformation of the spacetime vectors u, v that preserves the real inner product $(u, \bar{v})$ of one *conjugated* vector $\bar{v}$
$textrmuiv2(v, i_0) - v$ by another vector u:

$$(\mathcal{L}\{u\}, \overline{\mathcal{L}\{v\}}) = (u, \bar{v}),$$

where i_0 is the *unit* vector of unit length $\sqrt{(i_0, i_0)} \equiv 1$ along the time axis.

For brevity, only one option of $\pm\mathcal{L}\{u\}$ and $\pm\mathcal{L}\{\bar{u}\}$ is treated.

The transformation $\mathcal{L}$ involves Lorentz boost L as *self-adjoint* transform i.e. in an inner product L is transferred from one vector to another. So, for any u, v $(L\{u\}, v) = (u, L\{v\})$.

The problem is to obtain the quartet of eigenvectors c_k for the transformation $L_1 L_2$:

$$L_1 L_2\{c_k\} = \xi_k c_k \Leftrightarrow L_2\{c_k\} = \xi_k L_1^{-1}\{c_k\}, \tag{1}$$

where ξ_k is the real eigenvalue and the eigenvector serial number k ranges from 0 to 3. For simpler calculations, it is better to treat the equation located in (1) on the right.

General statements

To solve the equation (1), the easily provable general considerations are very useful. They are as follows:

1. The desired quartet of eigenvectors is always exist.
2. If ξ is an eigenvalue, then $\frac{1}{\xi}$ is also an eigenvalue.
3. If an eigenvalue ξ is different from 1, then it corresponds to a *lightwise* eigenvector c having a zero pseudo–length: $\xi \neq 1 \Rightarrow (c, \bar{c}) = 0$.

From these statements it is easy to establish without paper calculations that the quartet of eigenvalues consists of two units and a pair of mutually inverse values: 1, 1, ξ and $\frac{1}{\xi}$.

The characteristic equation for ξ is:

$$(\xi - 1)^2(\xi^2 - 2\xi \cosh \chi + 1) = 0, \tag{2}$$

where the scalar parameter χ is defined in accordance with famous cosine rule:

$$\cosh \frac{\chi}{2} = \cosh \frac{\theta_1}{2} \cosh \frac{\theta_2}{2} + (n_1, n_2) \sinh \frac{\theta_1}{2} \sinh \frac{\theta_2}{2} \tag{3}$$

and the scalar parameters θ_1 and θ_2 are the *rapidities*, such that the velocities v_1, v_2 divided by scalar speed of light c are expressed as $\mathrm{v}_1/\mathrm{c} = n_1 \tanh \theta_1$, $\mathrm{v}_2/\mathrm{c} = n_2 \tanh \theta_2$.

Note that (3) refers to the half hyperbolic angles $\frac{\chi}{2}$, $\frac{\theta_1}{2}$ and $\frac{\theta_2}{2}$, while the famous velocity addition is expressed via whole hyperbolic angles θ, θ_1 and θ_2 [4, 3].

From the above and concomitant considerations, we can conclude that the eigenvectors c_0, c_1, c_2, c_3 form a system of *pseudo-orthogonal* vectors, such that $(c_0, \bar{c}_0) = (c_0, \bar{c}_2) = (c_0, \bar{c}_3) = (c_1, \bar{c}_1) = (c_1, \bar{c}_2) = (c_1, \bar{c}_3) = (c_2, \bar{c}_3) = 0$.

Eigenvectors

The eigenvectors c_0, c_1, c_2, c_3 for the composition $L_1 L_2$ of Lorentz boosts L_1, L_2 and the corresponding eigenvalues are listed in Table 1.

In Table 1 n_1 and n_2 are the unit *spatial* vectors along the considered intersecting velocities, such that $(n_1, n_1) = (n_2, n_2) = 1$ and $(n_1, i_0) = (n_2, i_0) = 0$. The cross product $[n_1, n_2]$ is directed along the Wigner rotational axis ν [5], so that $[n_1, n_2] = \nu\sqrt{1 - (n_1, n_2)^2}$. The spatial part of the eigenvectors c_0 and c_1 depends

Notation	Eigenvector	Eigenvalue
c_0	$i_0 - d\vert_{\xi=e^{\chi}}$	e^{χ}
c_1	$i_0 - d\vert_{\xi=e^{-\chi}}$	$e^{-\chi}$
c_2	$i_0 - n_1\frac{\coth\frac{\theta_1}{2}+(n_1,n_2)\coth\frac{\theta_2}{2}}{1-(n_1,n_2)^2} + n_2\frac{\coth\frac{\theta_2}{2}+(n_1,n_2)\coth\frac{\theta_1}{2}}{1-(n_1,n_2)^2}$	1
c_3	$[n_1, n_2]$	1

TABLE 1. Eigenvectors for the composition of Lorentz boosts L_1L_2

on the eigenvalue ξ and, up to the sign, coincides with the unit vector d_ξ that is defined as a function of eigenvalue ξ in the form:

$$d_\xi = \frac{n_1\sqrt{\xi}\sinh\frac{\theta_1}{2} + n_2\sinh\frac{\theta_2}{2}}{\sqrt{\xi}\cosh\frac{\theta_1}{2} - \cosh\frac{\theta_2}{2}}. \tag{4}$$

The spatial parts $d|_{\xi=e^{\chi}}$ and $d|_{\xi=e^{-\chi}}$ of the eigenvectors c_0 and c_1 are obtained by substituting into (4) the values $\xi = e^{\chi}$ and $\xi = e^{-\chi}$, respectively.

Thus, in the context of the eigenvalue/eigenvector problem, the composition of Lorentz boosts is as *elementary* as a single Lorentz boost. In both cases, the solution boils down to stretching of one basis eigenvector and reverse decreasing of the second basis eigenvector with the remaining basis eigenvectors unchanged.

A ready-made solution of the eigenvalue/eigenvector problem for the composition of any rotation with a boost, as well as the expressions for representing the composition L_1L_2 of the boosts L_1, L_2 as a composition of the Wigner rotation and boost is given in [1, 2].

Conclusion

The relations (2)–(4) seem perfectly concise and quite simple to be widely presented in reference books to all whom it may concern. But these are missing. The only obstacle to obtaining the above formulae is cumbersome calculations as in [6]. Two things are important to overcome this obstacle. To simplify the formulae it is useful, firstly, to use hyperbolic geometry, as prescribed in [3, 4], and secondly, to carry out calculations in terms of quaternions [7].

This paper presents a solution to the eigenvalue/eigenvector problem in the vein of [3, 4] in coordinate–free way using the conventional cross product of four-dimensional vectors. In fact, the formulae (2)–(4) turned out to be easier to obtain in terms of the quaternion algebra equipped with quaternionic multiplication [8].

It's remarkable that a modern cross product of vectors is best defined in quaternions and generalized to the case of three arguments [9]. A triple cross product is especially convenient for describing Lorentz transformations, which are represented by a linear combination of orthogonal transforms and are described by triple products of variable vector and constant vector parameters [1, 2].

Among normalized algebras with a multiplicative unit, quaternions expand to octonions. In this case, a cross product of vectors is also generalized to the eight-dimensional case [8, 9, 10, 11]. Along the way, a generalization of Lorentz transformations to eight-dimensional spacetime is anticipated. Probably in the future it will be extremely interesting to generalize the laws of motion from the conditions of their invariance with respect to generalized Lorentz transformations.

The concise representation and description of the Lorentz transformations via eigenvectors may be useful for researchers who will be engaged in the mentioned generalization.

References

[1] M. Kharinov, *Sketch on quaternionic Lorentz transformations*, Int. Conf. on Polynomial Computer Algebra (PCA'2019), St.Petersburg, April 15–20, 2019, 71–74.

[2] M. Kharinov, *The Eigenvector Quartet for the Composition of Lorentz Boosts in the Quaternionic Space*, Proc. of the 15-th Int. Conf. on Finsler Extensions of Relativity Theory-2019 (FERT-2019), Moscow, October 24–27, 2019, 195–200.

[3] A.A. Ungar, *Hyperbolic geometry*, Fifteenth International Conference on Geometry, Integrability and Quantization, Varna, Bulgaria, 2013, 259–282, DOI: 10.7546/giq-15-2014-259-282.

[4] J.F. Barrett, *Minkovski Space-Time and Hyperbolic Geometry*, MASSEE International Congress on Mathematics MICOM-2015, 2015, 14 pp., available at `https://eprints.soton.ac.uk/397637/2/J\_F\_Barrett\_MICOM\_2015\_2018\_revision\_.pdf`.

[5] C. Møller, *The Theory of Relativity*, Oxford University Press, 1955, 386 pp.

[6] R. Ferraro, M. Thibeault, *Generic composition of boosts: an elementary derivation of the Wigner rotation*, European Journal of Physics, **20**(3), 1999, 143–151, DOI:10.1088/0143-0807/20/3/003.

[7] I.L. Kantor, A.S. Solodovnikov, *Hypercomplex numbers: an elementary introduction to algebras*, Springer, 1989, 166 pp.

[8] M. Kharinov, *Product of three octonions*, Springer Nature, Adv. Appl. Clifford Algebras, **29**(1), 2019, 11–26, DOI:10.1007/s00006-018-0928-x.

[9] T. Dray, C.A. Manogue, *The octonionic eigenvalue problem*, Springer Nature,, Adv. Appl. Clifford Algebras, **8**(2), 1998, 341–364.

[10] Z.K. Silagadze, *Multi-dimensional vector product*, Journal of Physics A: Mathematical and General **35**(23), Institute of Physics Publishing, 2002, 4949–4953.

[11] S. Okubo, *Triple products and Yang–Baxter equation. I. Octonionic and quaternionic triple systems*, Journal of mathematical physics **34**(7), 3273–3291 (1993)

Mikhail Kharinov
The Federal State Institution of Science
St. Petersburg Institute for Informatics and Automation
of the Russian Academy of Sciences (SPIIRAS)
St. Petersburg, Russia
e-mail: `khar@iias.spb.su`

Emergence of geometry in quantum mechanics based on finite groups

Vladimir V. Kornyak

Abstract. In the framework of constructive quantum mechanics, we consider the emergence of geometry from entanglement in composite quantum systems. We specify the most general structure of the symmetry group of a quantum system with geometry. We show that the 2nd Rényi entanglement entropy may be useful in applying polynomial computer algebra to model metric structures in quantum systems with geometry.

1. Introduction

In [1, 2, 3] we proposed a constructive modification of quantum mechanics that replaces the unitary group in a Hilbert space over the field $\mathbb{C}$ with the unitary representation of a finite group in a Hilbert space over an abelian extension of $\mathbb{Q}$ which is a dense subfield of $\mathbb{R}$ or $\mathbb{C}$ depending on the structure of the group. T. Banks recently [4] analyzed this modification from the point of view of real physics and cosmology and came to the conclusion that it "can probably be a model of the world we observe."

In short, constructive quantum mechanics boils down to the following. We start with the set $\Omega = \{e_1, \ldots, e_{\mathcal{N}}\} \cong \{1, \ldots, \mathcal{N}\}$ of "types" of primary ("ontic") objects on which a permutation group G acts (T. Banks showed that it suffices to assume that $G = \mathsf{S}_{\mathcal{N}}$ in order to "encompass finite dimensional approximations to all known models of theoretical physics"). Let n_i be the number of instances of ontic objects of the ith type. Then the set of all objects can be described by the vector

$$|n\rangle = (n_1, \ldots, n_{\mathcal{N}})^{\mathsf{T}}. \tag{1}$$

These "ontic" vectors form the semimodule H_Ω over the semiring of natural numbers $\mathbb{N} = \{0, 1, 2, \ldots\}$.

The action of G on Ω determines the *permutation representation* $\mathcal{P}(G)$ in the semimodule H_Ω. For $g \in G$, the matrix of the permutation representation has the form $\mathcal{P}(g)_{i,j} = \delta_{ig,j}$. Using standard mathematical procedures, the semiring $\mathbb{N}$ can be extended to a field $\mathcal{F}$ which is a *splitting field* for the group G. The field $\mathcal{F}$ is a subfield of ℓth *cyclotomic field*, where ℓ is the *exponent* of the group G. Depending on the structure of G, the field $\mathcal{F}$ is a dense subfield of either $\mathbb{R}$ or $\mathbb{C}$, i.e., $\mathcal{F}$ is physically indistinguishable from these continuous fields. The extension of $\mathbb{N}$ to $\mathcal{F}$ induces the extension of the ontic semimodule H_Ω to the

Hilbert space $\mathcal{H}_\Omega$. The inner product in this Hilbert space is a natural extension of the *standard inner product* in the ontic semimodule: $\langle m \mid n \rangle = \sum_{i=1}^{\mathcal{N}} m_i n_i$, where $|m\rangle = (m_1, \ldots, m_\mathcal{N})^\intercal$ and $|n\rangle = (n_1, \ldots, n_\mathcal{N})^\intercal$ are ontic vectors. The standard inner product is invariant under the representation $\mathcal{P}(G)$)

Since $\mathcal{F}$ is a splitting field, we can decompose the Hilbert space $\mathcal{H}_\Omega$ into irreducible subspaces that are invariant with respect to the representation $\mathcal{P}(G)$:

$$\mathcal{H}_\Omega = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots \oplus \mathcal{H}_K \,.$$

This decomposition can be constructed algorithmically by calculating the complete set of mutually orthogonal invariant projectors: $B_1, B_2, \ldots, B_K$.[1] An arbitrary invariant subspace $\mathcal{H}_\alpha \leq \mathcal{H}_\Omega$ is a direct sum of irreducible ones:

$$\mathcal{H}_\alpha = \bigoplus_{k' \in \alpha} \mathcal{H}_{k'}, \quad \alpha \subseteq \{1, \ldots, K\} \,.$$

Accordingly, the projection operator in $\mathcal{H}_\alpha$ has the form $B_\alpha = \sum\limits_{k' \in \alpha} B_{k'}$.

In any invariant subspace $\mathcal{H}_\alpha$, an independent quantum system can be constructed, since the results of both unitary evolutions and projective measurements applied to any vector belonging to the subspace $\mathcal{H}_\alpha$ will remain in this subspace.

The inner product for the projections $|\varphi\rangle = B_\alpha |m\rangle$ and $|\psi\rangle = B_\alpha |n\rangle$ of ontic vectors takes the form $\langle \varphi \mid \psi \rangle_\alpha = \langle m | B_\alpha | n \rangle$. In terms of ontic vectors, a pure state in the subspace $\mathcal{H}_\alpha$ can be represented as the unit vector $|\psi\rangle = \frac{B_\alpha |n\rangle}{\sqrt{\langle n | B_\alpha | n \rangle}}$ or as the density matrix $\rho = \frac{B_\alpha |n\rangle\langle n| B_\alpha}{\langle n | B_\alpha | n \rangle}$. Operators of unitary evolution in the subspace $\mathcal{H}_\alpha$ have the form $U_{\alpha,g} = B_\alpha \mathcal{P}(g)$.

2. Symmetry Group of Composite Quantum System

The Hilbert space of an N-component quantum system has the form

$$\widetilde{\mathcal{H}} = \bigotimes_{x \in X} \mathcal{H}_x \,. \tag{2}$$

where $X \cong \overline{N} = \{1, \ldots, N\}$. A Hilbert space that can be decomposed into a tensor product of spaces of smaller dimensions is a special case of a general Hilbert space, so it is natural to assume that structures like (2) arise as approximations. This is consistent with the general "holistic" view that the partition of the system as a whole into subsystems is always conditional and approximate.

We make the following assumptions:

- The set X of indices of "local" Hilbert spaces $\mathcal{H}_x$ has symmetries that form the group G.
- The local Hilbert spaces are isomorphic, i.e., $\mathcal{H}_x \cong \mathcal{H}$ for any $x \in X$, where $\mathcal{H}$ is a representative of the equivalence class of spaces $\mathcal{H}_x$.

[1] We have developed and implemented an efficient algorithm for such calculations [5].

- In the local space $\mathcal{H}$, the unitary representation acts, which is a subrepresentation of the permutation representation of the group F acting on the set $V \cong \overline{M} = \{1, \ldots, M\}$, that is, the set V is the basis of the permutation representation.

The set X can be interpreted as a "geometric space", and the group G as a group of "spatial" symmetries. The group F is interpreted as a group of "local" symmetries.

Based on the natural properties that a geometric space must have, we can show that the group $\widetilde{W}$, which combines spatial and local symmetries, belongs to an equivalence class of group extensions of the form

$$\begin{array}{ccccccccc} & & & & \widetilde{W} & & & & \\ & & & \nearrow & \downarrow & \searrow & & & \\ \mathbf{1} & \longrightarrow & F^X & & \Phi & & G & \longrightarrow & \mathbf{1} \\ & & & \searrow & \downarrow & \nearrow & & & \\ & & & & \widetilde{W}' & & & & \end{array} , \tag{3}$$

where F^X is a group of F-valued functions on the space X, and $\Phi : \widetilde{W} \to \widetilde{W}'$ is a group isomorphism that provides the commutativity of the diagram.

The set of elements of $\widetilde{W}$ can be identified with the Cartesian product of the sets F^X and G, i.e., the elements of $\widetilde{W}$ can be represented as pairs $(f(x), g)$, where $f(x) \in F^X$, $g \in G$. Explicit calculations lead to the following:

- The equivalence classes of extensions (3) are parameterized by *antihomomorphisms* of the space group, that is, by functions $\mu : G \to G$ such that $\mu(ab) = \mu(b)\,\mu(a)$ for any $a, b \in G$.
- An isomorphism of equivalent extensions has the form

$$\Phi : (f(x), g) \mapsto (f(x\varphi(g)), g),$$

  where $\varphi : G \to G$ is an *arbitrary* function.
- The main group operations have the following explicit form:

$$v(x)\,(f(x), g) = v(x\mu(g))\, f(x\varphi(g)), \tag{4}$$

$$(f(x), g)\,(f'(x), g') = \Big(f\Big(x\varphi(gg')^{-1}\mu(g')\,\varphi(g)\Big)\, f'\Big(x\varphi(gg')^{-1}\varphi(g')\Big), gg'\Big), \tag{5}$$

$$(f(x), g)^{-1} = \Big(f\Big(x\varphi\big(g^{-1}\big)^{-1}\mu(g)^{-1}\,\varphi(g)\Big)^{-1}, g^{-1}\Big), \tag{6}$$

  where (4) is the action of $(f(x), g) \in \widetilde{W}$ on the function $v(x) \in V^X$,
  (5) is the group multiplication in $\widetilde{W}$, and (6) is the group inversion.

There are two universal (i.e., existing for any group, regardless of its specific properties) antihomomorphisms: $\mu(g) = \mathbf{1}$ and $\mu(g) = g^{-1}$. The choice of $\mu(g) = \mathbf{1}$ leads to the trivial extension, i.e., to the direct product $\widetilde{W} \cong F^X \times G$. The antihomomorphism (in fact, antiisomorphism) $\mu(g) = g^{-1}$ leads to a semidirect product

of the groups F^X and G, which is called the *wreath product* of the groups F and G:

$$\widetilde{W} = F \wr G \cong F^X \rtimes G. \tag{7}$$

As for the arbitrary function φ, we use two options in the implementation of our algorithms : $\varphi(g) = g^{-1}$ and $\varphi(g) = \mathbf{1}$. In these cases, expressions (4) – (6) for group operations are more or less compact:

	$\varphi(g) = g^{-1}$	$\varphi(g) = \mathbf{1}$
$v(x)\,(f(x)\,,g) =$	$v\left(xg^{-1}\right) f\left(xg^{-1}\right)$	$v\left(xg^{-1}\right) f(x)$
$(f(x)\,,g)\,(f'(x)\,,g') =$	$(f(x)\,f'(xg)\,,gg')$	$\left(f\left(xg'^{-1}\right) f'(x)\,,gg'\right)$
$(f(x)\,,g)^{-1} =$	$\left(f\left(xg^{-1}\right)^{-1},g^{-1}\right)$	$\left(f(xg)^{-1},g^{-1}\right)$

The unitary representations of the group (7) in the whole Hilbert space (2) describe the quantum properties of the system as a whole. To calculate invariant projectors and decompose permutation representations of wreath products into irreducible components, we developed an algorithm [6], whose C implementation splits representations having dimensions and ranks up to 10^{16} and 10^9, respectively.

3. Emergence of Geometry From Entanglement

The natural idea is to determine the distances between points in the space X in terms of quantum correlations: the greater the correlation, the less the distance. Quantitatively, quantum correlations are described by *measures of entanglement*. The problems of constructing metrics and topology in entangled quantum systems are considered, in particular, in [7, 8, 9].

Denote by $\mathcal{D}\left(\widetilde{\mathcal{H}}\right)$ the set of all states (density matrices) in the Hilbert space (2). The set of *separable states* $\mathcal{D}_{\mathrm{S}}\left(\widetilde{\mathcal{H}}\right)$ consists of states $\rho \in \mathcal{D}\left(\widetilde{\mathcal{H}}\right)$ that can be represented as weighted sums of tensor products of states of components:

$$\rho = \sum_k w_k \bigotimes_{x\in X} \rho_x^k, \quad w_k \geq 0, \quad \sum_k w_k = 1, \quad \rho_x^k \in \mathcal{D}(\mathcal{H}_x)\,.$$

The set of *entangled states* $\mathcal{D}_{\mathrm{E}}\left(\widetilde{\mathcal{H}}\right)$ is defined as the complement of $\mathcal{D}_{\mathrm{S}}\left(\widetilde{\mathcal{H}}\right)$ in the set of all states: $\mathcal{D}_{\mathrm{E}}\left(\widetilde{\mathcal{H}}\right) = \mathcal{D}\left(\widetilde{\mathcal{H}}\right) \setminus \mathcal{D}_{\mathrm{S}}\left(\widetilde{\mathcal{H}}\right)$.

Let ρ_{AB} denote the density matrix for a composite quantum system consisting of components A and B. The statistics of observations of subsystem A are reproduced by the *reduced density matrix* $\rho_A = \mathrm{tr}_B\,(\rho_{AB})$, where the *partial trace* tr_B over subsystem B is defined by the relation

$$\mathrm{tr}_B\,(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) = |a_1\rangle\langle a_2|\,\mathrm{tr}\,(|b_1\rangle\langle b_2|)\,,$$

which must hold for any vectors $|a_1\rangle, |a_2\rangle \in \mathcal{H}_A$ and $|b_1\rangle, |b_2\rangle \in \mathcal{H}_B$.

The constructions considered below depend on "the quantum state of the universe"

$$\rho_X \in \widetilde{\mathcal{H}}\,. \tag{8}$$

There are a variety of entanglement measures [10]. A typical measure of entanglement for the pair of points $\{x, y\} \subseteq X$ is the *mutual information*

$$I(x,y) = S(\rho_x) + S(\rho_y) - S(\rho_{xy})\,, \tag{9}$$

where $\rho_x = \mathrm{tr}_{X\setminus\{x\}}\,\rho_X$, $\rho_y = \mathrm{tr}_{X\setminus\{y\}}\,\rho_X$, and $\rho_{xy} = \mathrm{tr}_{X\setminus\{x,y\}}\,\rho_X$. The function $S(\rho)$ is called *entanglement entropy*. The entanglement entropy is usually defined as the *von Neumann entropy* $S(\rho) = -\,\mathrm{tr}(\rho\log\rho)$, which is the quantum version of the *Shannon entropy*

$$H(p_1,\ldots,p_n) = -\sum_{k=1}^{n} p_k \log p_k\,, \tag{10}$$

where $p_1,\ldots,p_n$ is a probability distribution.

From a general point of view, entropy is a function on probability distributions that satisfies some natural postulates. A. Rényi proved [11] that such functions form the following family

$$H_q(p_1,\ldots,p_n) = \frac{1}{1-q}\log\sum_{k=1}^{n} p_k^q\,, \tag{11}$$

where $q \geq 0$ and $q \neq 1$. The function H_q is called the Rényi entropy of order q.

The Shannon entropy (10) is a limiting case of (11): $H \equiv H_1 = \lim\limits_{q\to 1} H_q$. Note that the Shannon entropy has better statistical properties compared to the Renyi entropies with $q \neq 1$, for which, in particular, expression (9) can take negative values. The entropy $H_2(p_1,\ldots,p_n) = -\log\sum_{k=1}^{n} p_k^2$ is called the *collision entropy*.

The *quantum Rényi entropy* is the quantum analogue of (11):

$$S_q(\rho) = \frac{1}{1-q}\log\mathrm{tr}(\rho^q)\,.$$

We will use the 2nd quantum Rényi entropy (quantum collision entropy)

$$S_2(\rho) = -\log\mathrm{tr}\big(\rho^2\big) \tag{12}$$

as the entanglement entropy for the following reasons.

Gleason's theorem provides a one-to-one correspondence between probability measures on subspaces of a Hilbert space and quantum states in this space. More specifically, the most general expression for the Born probability has the form $\mathbb{P} = \mathrm{tr}(\rho_O\rho_S)$, where ρ_O and ρ_S are quantum states of the "observer" and the "observed system", respectively. Since the Born rule is the only fundamental source of probability in quantum theory, it is natural to associate a single state ρ with some Born probability. The probability $\mathbb{P} = \mathrm{tr}\big(\rho^2\big)$ – "the system observes itself" – is such a choice, and its logarithm is precisely the 2nd Rényi entropy (12).

In models of emergent space, the geodesic distance between local quantum subsystems is determined by a certain monotonic function of the entanglement

measure [9]. Such a "scaling" function should, at least approximately, tend to zero for maximally entangled pairs of local subsystems, tend to infinity for separable pairs, and satisfy the usual distance properties, such as the triangle inequality, etc. Using the 2nd Rényi entropy as the entanglement entropy, we can get rid of the logarithms in computer algebra calculations by replacing the mutual information (9) with the expression

$$P(x,y) = \exp(-I(x,y)) = \frac{\operatorname{tr}\left(\rho_{xy}^2\right)}{\operatorname{tr}\left(\rho_x^2\right)\operatorname{tr}\left(\rho_y^2\right)}\,. \tag{13}$$

For a separable pair $\{x,y\}$, we have $\rho_{xy} = \rho_x \otimes \rho_y$ and, therefore, $P(x,y) = 1$. For a maximally entangled pair $P(x,y) = (\dim \mathcal{H})^2$, where $\mathcal{H}$ is the local Hilbert space. $\rho_{xy} \neq \rho_x \otimes \rho_y$ implies $P(x,y) \neq 1$, so expression (13) can quantify the quantum correlation between x and y. For the pure state (8), expression (13) is a combination of polynomials in the coordinates of the ontic vector (1).

References

[1] Kornyak V.V. *Quantum models based on finite groups.* IOP Conf. Series: Journal of Physics: Conf. Series **965**, 012023, 2018. arXiv:1803.00408 [physics.gen-ph]

[2] Kornyak V.V. *Modeling Quantum Behavior in the Framework of Permutation Groups.* EPJ Web of Conferences **173**, 01007, 2018. arXiv:1709.01831 [quant-ph]

[3] Kornyak V.V. *Mathematical Modeling of Finite Quantum Systems.* In: Adam G. *et al* (eds) MMCP 2011. LNCS, **7125**. Springer, 2012. arXiv:1107.5675 [quant-ph]

[4] Banks T. *Finite Deformations of Quantum Mechanics.* arXiv:2001.07662 [hep-th], 20 p., 2020.

[5] Kornyak V.V. *An Algorithm for Decomposing Representations of Finite Groups Using Invariant Projections.* J. Math. Sci. **240**, 651–664, 2019.

[6] Kornyak V.V. *An Algorithm for Computing Invariant Projectors in Representations of Wreath Products.* LNCS, **11661**, 300–314, Springer, 2019.

[7] Van Raamsdonk M. *Building up spacetime from quantum entanglement.* Gen. Relativ. Grav. **42**, 2323–2329, 2010.

[8] Maldacena J., Susskind L. *Cool horizons for entangled black holes.* Fortsch. Phys. **61** (9), 781–811, 2013.

[9] Cao C., Carroll S.M., Michalakis S. *Space from Hilbert space: Recovering geometry from bulk entanglement.* Phys. Rev. D **95**, 024031, 2017.

[10] Plenio M.B., Virmani S. *An introduction to entanglement measures.* Quant. Inf. Comput. **7**, 1–51, 2007. arXiv:0504163 [quant-ph]

[11] Rényi A. *On measures of information and entropy.* Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability, 547–561, 1960.

Vladimir V. Kornyak
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: `vkornyak@gmail.com`

Teleportation of the Bell states on IBM Q 5 Yorktown quantum computer

Vladimir P. Gerdt and Ekaterina A. Kotkova

Abstract. In this talk, we present realization of a protocol of two-qubit state teleportation on the 5-qubit quantum IBM quantum computer Yorktown and compare its results with those obtained with the Maple-based simulator called Feynman. We compare error rates for implementations on the IBM Q Yorktown achieved in our previous studies to actual error rates for the same quantum circuits and the circuits reduced after IBM Q Yorktown modification. Our experiments show the technical improvement of IBM quantum hardware over the past year.

Noisy Intermediate-Scale Quantum (NISQ) technology is rapidly developing over last years. The near-term quantum computers with 50-100 qubits are able to perform tasks which surpass the capabilities of today's classical digital computers. However, noisy qubits and quantum gates lead to limitation of the size of quantum circuits that can be executed reliably. In the given talk we present our results partially taken from [1] on implementation on the IBM Q 5 Yorktown quantum computer (Fig. 1), and other 5-qubit computers accessible via the IBM cloud (https://www.ibm.com/quantum-computing/technology/experience/), the protocol [2] of quantum teleportation of Bell (EPR) states. We adopt the original version of this protocol to set of gates built-in IBM Q which includes one-qubit gate

$$U2(\phi,\lambda)=\begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{\exp i\lambda}{\sqrt{2}}\\ \frac{\exp i\phi}{\sqrt{2}} & \frac{\exp i(\lambda+\phi)}{\sqrt{2}}\end{pmatrix},\quad \oplus=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \lambda,\phi\in[0,2\pi],\quad H=U2(0,\pi),$$

the measurement gate and the 2-qubit control-$\oplus$ (CNOT) gate (Fig. 2).

The Bell states are the maximally entangled 2-qubit states. They are created in the lower left corner of the circuit if the input qubits q_3 and q_4 are in the classical states

$$|q_0\rangle=|x\rangle\,,|q_1\rangle=|y\rangle\longrightarrow|\beta_{xy}\rangle=\frac{1}{\sqrt{2}}\left(|0,y\rangle+(-1)^x\,|1,1-y\rangle\right),\quad x,y\in\{0,1\}.$$

The bell states $|\beta_{xy}\rangle$ are transported to the states of qubits q_1 and q_2. Then they are measured in the classical basis on the output. The results of the measurement

shown in Fig. 3. Because of the noise, the obtained probabilities are substantially different of the expected ones shown in Fig. 4. These expected probability values were computed by using FEYNMAN [3], the classical simulator written in MAPLE. Since it does the related computations exactly, i.e. without noise, its output 2-qubit state is exactly equal to the input Bell state.

Thus, for applicability of IBM quantum computers even to small (w.r.t. the number of qubits needed) problems, one has to decrease the hardware errors.

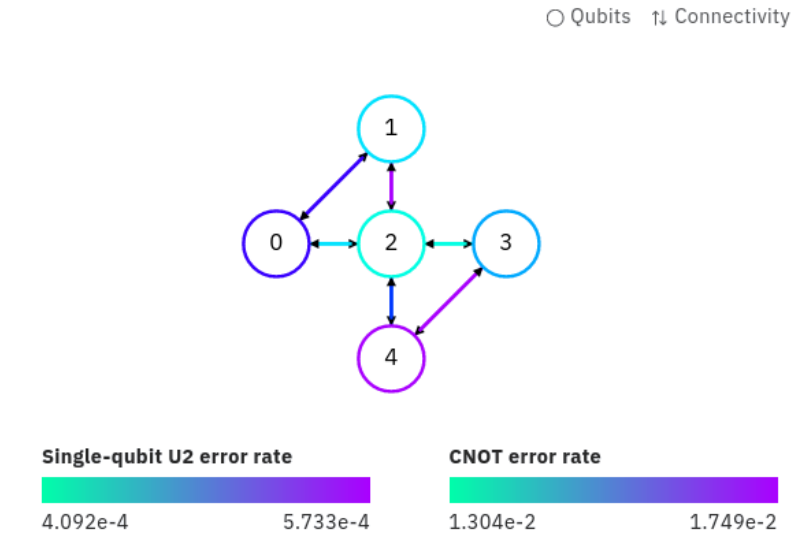


FIGURE 1. Quantum computer IBM Q 5 Yorktown

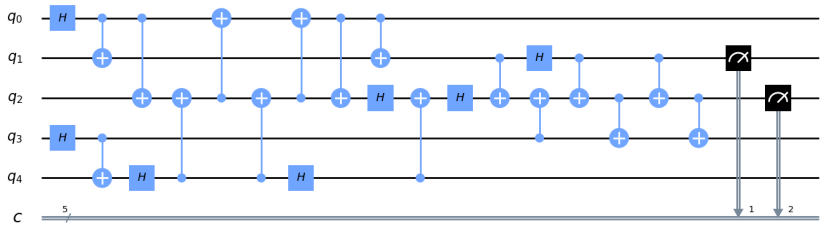


FIGURE 2. Quantum circuit

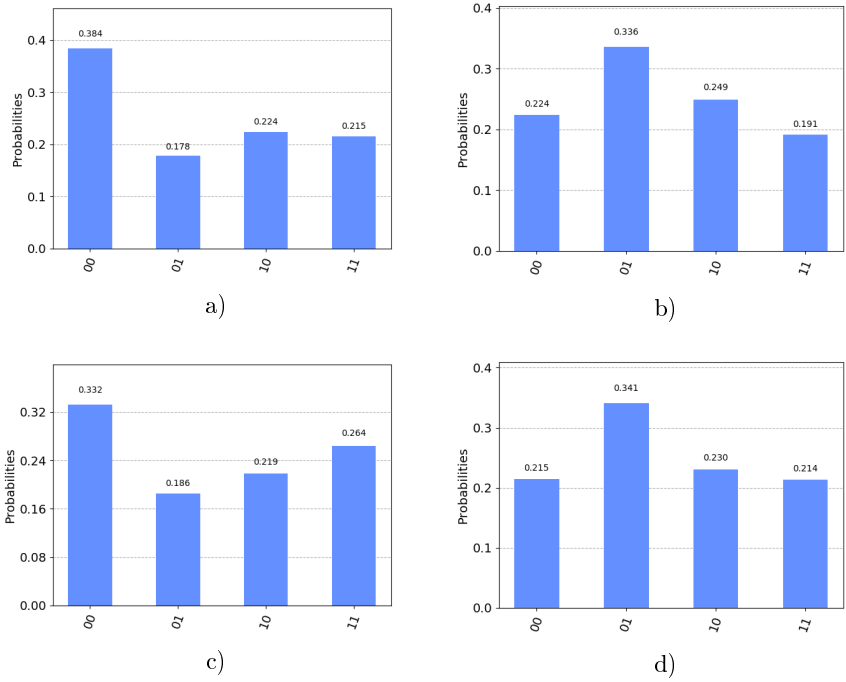


FIGURE 3. Results of teleportation of the Bell states on the IBM Q Yorktown with readout in the classical basis. a) state $|\beta_{00}\rangle$, b) state $|\beta_{01}\rangle$, c) state $|\beta_{10}\rangle$, d) state $|\beta_{11}\rangle$.

References

[1] V.P.Gerdt, E.A.Kotkova and V.V.Vorob'ev. *The Teleportation of the Bell States Has Been Carried Out on the Five-Qubit Quantum IBM Computer.* Physics of Particles and Nuclei Letters, 2019, Vol. 16, No. 6, pp. 975–984.

[2] V.N.Gorbachev and A.I.Trubilko. *Quantum teleportation of an Einstein-Podolsky-Rosen pair using an entanglement three-particle state.* Journal of Experimental and Theoretical Physics, 118, 1036–1040 (2000); arXiv:9906110 [quant-ph].

[3] T.Radtke and S.Fritzsche, *Simulation of n-qubit quantum systems. I. Quantum registers and quantum gates.* Computer Physics Communications, 173, 91–113 (2005); *II. Separability and entanglement.* 175, 145–166 (2006); *III. Quantum operations.* 176, 617–633 (2007); *IV. Parametrization of quantum states.* 179, 647–664 (2008); *V. Quantum measurements.* 181, 440–453 (2010).

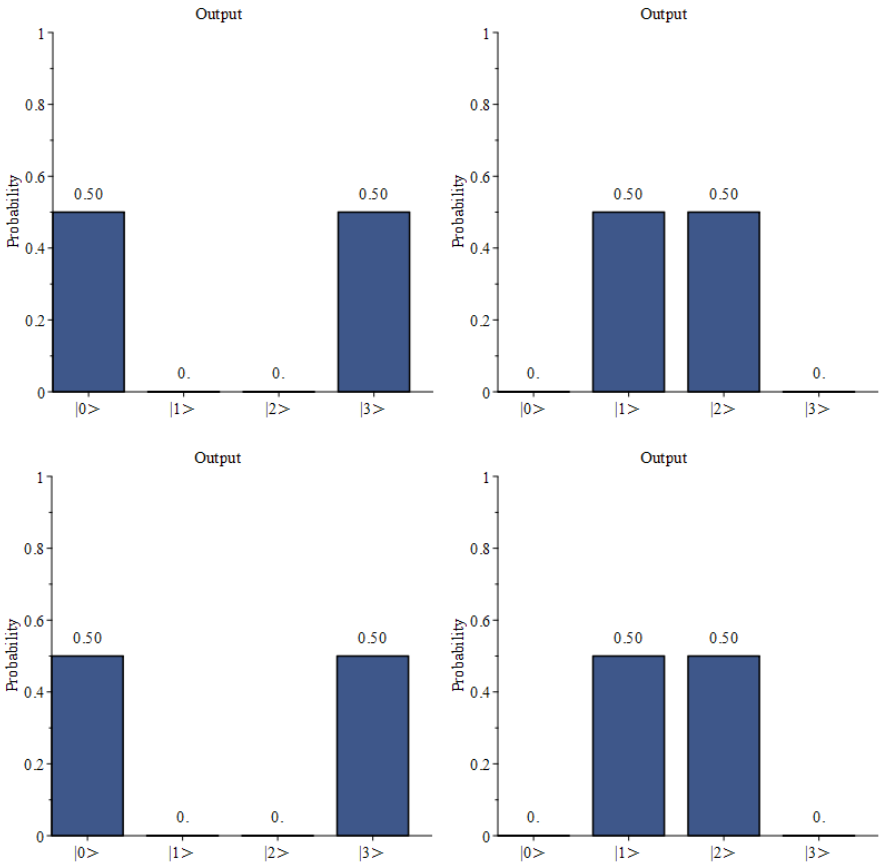


FIGURE 4. Classical simulation with FEYNMAN

Vladimir P. Gerdt
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: gerdt@jinr.ru

Ekaterina A. Kotkova
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: ekaterina.a.kotkova@gmail.com

Tropical algebra solution of a project scheduling problem

N. Krivulin and M. Petrakov

Introduction

Time-constrained project scheduling problems constitute an integral part of project management. These problems are to find an optimal schedule for a project that consists of a set of activities operating in parallel under various temporal constraints, including start-start, start-finish, finish-start, release time, deadline, due-dates and other constraints. As optimization criteria to minimize, one can take the project makespan, the maximum deviation from due dates, the maximum flow-time, the maximum deviation of start or finish times [1, 2].

Many time-constrained scheduling problems can be formulated as linear, integer, or mixed-integer linear programs, graph and network optimization problems, and then solved using appropriate computational algorithms. This approach usually allows one to obtain a numerical solution of the problem, but cannot provide a complete analytical solution in an explicit form.

In this paper, we consider a project, in which activities are performed under temporal constraints in the form of start-start precedence relationships, release start and release end times. The scheduling problem of interest is to find the start times of activities to provide the minimum deviation of start times. Such an optimality criterion can arise when the schedule has to provide a common start time for all activities in the project.

We represent the problem in terms of tropical mathematics, which deals with the theory and applications of algebraic systems with idempotent operations [3, 4, 5]. To solve the project scheduling problem, we apply methods and results of tropical optimization [6, 7, 8], and then obtain a new complete solution, which provides the result in an explicit analytical form, ready for further analysis and numerical implementation.

This work was supported in part by the Russian Foundation for Basic Research, Grant No. 20-010-00145.

1. A time-constrained project scheduling problem

Consider a project that consists of n activities operating under start-start, release start and release end temporal constraints. For each activities $i = 1, \ldots, n$ we denote the start time by x_i. Let g_i and h_i be release start and release end times, which specify the earliest and latest allowed time for activity i to start. Let b_{ij} be the minimum allowed time lag between the start of activity i and the start of j.

Suppose that the optimal schedule has to minimize the maximum deviation of start times x_i over all activities i. The project scheduling problem is formulated as follows: given the parameters g_i, h_i and b_{ij}, find the start times x_i for all $i = 1, \ldots, n$ to solve the minimization problem

$$\begin{aligned} &\max_{1 \le i \le n} x_i + \max_{1 \le i \le n} (-x_i) \to \min, \\ &\max_{1 \le j \le n} (x_j + b_{ij}) \le x_i, \\ &g_i \le x_i \le h_i, \quad i = 1, \ldots, n. \end{aligned} \tag{1}$$

2. Elements of tropical algebra

Let $\mathbb{X}$ be a set endowed with two associative and commutative operations: $\oplus$ (addition) and $\otimes$ (multiplication), and equipped with additive and multiplicative neutral elements: $\mathbb{0}$ (zero) and $\mathbb{1}$ (unit). Addition is idempotent, which yields $x \oplus x = x$ for each $x \in \mathbb{X}$. Multiplication distributes over addition and is invertible to provide each nonzero $x \in \mathbb{X}$ with its inverse x^{-1} such that $x^{-1} \otimes x = \mathbb{1}$. The algebraic structure $(\mathbb{X}, \mathbb{0}, \mathbb{1}, \oplus, \otimes)$ is normally called the idempotent semifield.

Let $\mathbb{X}^{m \times n}$ be the set of matrices consisting of m rows and n columns with elements from $\mathbb{X}$. Matrix addition and multiplication and multiplication by scalars are performed according to the usual rules with replacement of arithmetic addition and multiplication by the operations $\oplus$ and $\otimes$.

Consider the set $\mathbb{X}^{n \times n}$ of square matrices of order n. A matrix with $\mathbb{1}$ on the diagonal and $\mathbb{0}$ elsewhere is the identity matrix denoted $\boldsymbol{I}$. The power notation with nonnegative integer exponents serves to represent iterated products of matrices as follows: $\boldsymbol{A}^0 = \boldsymbol{I}$ and $\boldsymbol{A}^p = \boldsymbol{A}^{p-1}\boldsymbol{A}$ for any matrix $\boldsymbol{A}$ and integer $p > 0$.

The trace of a matrix $\boldsymbol{A} = (a_{ij})$ is calculated as $\operatorname{tr} \boldsymbol{A} = a_{11} \oplus \cdots \oplus a_{nn}$.

Furthermore, we introduce the function

$$\operatorname{Tr}(\boldsymbol{A}) = \operatorname{tr} \boldsymbol{A} \oplus \cdots \oplus \operatorname{tr} \boldsymbol{A}^n.$$

If $\operatorname{Tr}(\boldsymbol{A}) \le \mathbb{1}$, we define a matrix, which is usually called the Kleene star matrix

$$\boldsymbol{A}^* = \boldsymbol{I} \oplus \boldsymbol{A} \oplus \cdots \oplus \boldsymbol{A}^{n-1}.$$

Let $\mathbb{X}^n$ denote the set of column vectors of dimension n. A vector containing all elements as $\mathbb{0}$ is the zero vector. A vector without zero components is called regular. A vector with all elements equal to $\mathbb{1}$ is denoted by $\mathbf{1} = (\mathbb{1}, \ldots, \mathbb{1})^T$.

For any nonzero vector $\boldsymbol{x} = (x_i) \in \mathbb{X}^n$, its multiplicative conjugate transpose is the row vector $\boldsymbol{x}^- = (x_i^-)$, where $x_i^- = x_i^{-1}$ if $x_i^- > \mathbb{0}$, and $x_i^- = \mathbb{0}$ otherwise.

An example of the idempotent semifield under consideration is the real semifield $\mathbb{R}_{\max,+} = (\mathbb{R} \cup \{-\infty\}, \max, +, -\infty, 0)$, in which the addition $\oplus$ is defined as maximum, and the multiplication $\otimes$ is as ordinary addition, with the zero $\mathbb{0}$ given by $-\infty$, and the identity $\mathbb{1}$ by 0. Each number $x \in \mathbb{R}$ has the inverse x^{-1} equal to the opposite number $-x$ in the conventional notation. For all $x, y \in \mathbb{R}$, the power x^y is well-defined and coincides with the arithmetic product xy.

In the algebraic expressions below, the multiplication sign $\otimes$ is omitted to save writing: $x \otimes y = xy$.

3. Representation and solution of project scheduling problem

Consider problem (1) and represent it in terms of the semifield $\mathbb{R}_{\max,+}$. The constraints in the problem take the form

$$\bigoplus_{j=1}^{n} b_{ij} x_j \le x_i, \qquad g_i \le x_i \le h_i, \qquad i = 1, \ldots, n.$$

Next, we introduce the matrix-vector notation $\boldsymbol{B} = (b_{ij})$, $\boldsymbol{g} = (g_i)$ and $\boldsymbol{h} = (h_i)$ to represent the constraints in the vector form

$$\begin{aligned} \boldsymbol{B}\boldsymbol{x} &\le \boldsymbol{x}, \\ \boldsymbol{g} &\le \boldsymbol{x} \le \boldsymbol{h}, \end{aligned}$$

and note that the inequalities $\boldsymbol{B}\boldsymbol{x} \le \boldsymbol{x}$ and $\boldsymbol{g} \le \boldsymbol{x}$ are equivalent to the inequality $\boldsymbol{B}\boldsymbol{x} \oplus \boldsymbol{g} \le \boldsymbol{x}$.

In terms of the semifield $\mathbb{R}_{\max,+}$, the objective function becomes

$$\bigoplus_{1 \le i \le n} x_i \bigoplus_{1 \le j \le n} x_j^{-1} = \mathbf{1}^{\mathrm{T}} \boldsymbol{x} \boldsymbol{x}^- \mathbf{1} = \boldsymbol{x}^- \mathbf{1}\mathbf{1}^{\mathrm{T}} \boldsymbol{x}.$$

By combining the objective function with constraints, we obtain the tropical optimization problem

$$\begin{aligned} &\boldsymbol{x}^- \mathbf{1}\mathbf{1}^{\mathrm{T}} \boldsymbol{x} \to \min, \\ &\boldsymbol{B}\boldsymbol{x} \oplus \boldsymbol{g} \le \boldsymbol{x}, \\ &\boldsymbol{x} \le \boldsymbol{h}. \end{aligned} \tag{2}$$

The following result offers a complete solution to the problem.

Theorem 1. *Let $\boldsymbol{B}$ be a matrix, $\boldsymbol{g}$ be a vector, and $\boldsymbol{h}$ be a regular vector such that $\mathrm{Tr}(\boldsymbol{B}) \oplus \boldsymbol{h}^- \boldsymbol{B}^* \boldsymbol{g} \le \mathbb{1}$. Then the minimum in problem (2) is equal to*

$$\theta = \bigoplus_{i=0}^{n-1} \mathbf{1}^{\mathrm{T}} \boldsymbol{B}^i (\boldsymbol{I} \oplus \boldsymbol{g}\boldsymbol{h}^-)(\boldsymbol{I} \oplus \boldsymbol{B})^{n-1-i} \mathbf{1}, \tag{3}$$

and all regular solutions are given by

$$\boldsymbol{x} = (\theta^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}} \oplus \boldsymbol{B})^{*}\boldsymbol{u}, \quad \boldsymbol{g} \le \boldsymbol{u} \le (\boldsymbol{h}^{-}(\theta^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}} \oplus \boldsymbol{B})^{*})^{-}. \tag{4}$$

If $\mathrm{Tr}(\boldsymbol{B}) \oplus \boldsymbol{h}^{-}\boldsymbol{B}^{*}\boldsymbol{g} > \mathbb{1}$*, then there are no regular solutions.*

4. A numerical example

Let us examine a project that involves $n = 4$ activities under constraints given by the matrix and the vectors

$$\boldsymbol{B} = \begin{pmatrix} 0 & 2 & 3 & 1 \\ \mathbb{0} & 0 & 6 & 7 \\ \mathbb{0} & \mathbb{0} & 0 & 3 \\ \mathbb{0} & \mathbb{0} & \mathbb{0} & 0 \end{pmatrix}, \qquad \boldsymbol{g} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \qquad \boldsymbol{h} = \begin{pmatrix} 20 \\ 15 \\ 10 \\ 10 \end{pmatrix}.$$

We start with the verification of existence conditions for regular solutions in Theorem 2. We obtain

$$\boldsymbol{B}^{*} = \begin{pmatrix} 0 & 2 & 8 & 11 \\ \mathbb{0} & 0 & 6 & 9 \\ \mathbb{0} & \mathbb{0} & 0 & 3 \\ \mathbb{0} & \mathbb{0} & \mathbb{0} & 0 \end{pmatrix}, \qquad \boldsymbol{h}^{-}\boldsymbol{B}^{*}\boldsymbol{g} = -2, \qquad \mathrm{Tr}\,\boldsymbol{B} = 0.$$

Since $\mathrm{Tr}\,\boldsymbol{B} \oplus \boldsymbol{h}^{-}\boldsymbol{B}^{*}\boldsymbol{g} = 0$, the problem has regular solutions. As the next step, we find the minimum value θ by application of (3). We have

$$\theta = 11.$$

To describe the solution set defined by (4), we obtain

$$(\theta^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}} \oplus \boldsymbol{B})^{*} = \begin{pmatrix} 0 & 2 & 8 & 11 \\ -2 & 0 & 6 & 9 \\ -8 & -6 & 0 & 3 \\ -11 & -9 & -3 & 0 \end{pmatrix}, \qquad (\boldsymbol{h}^{-}(\theta^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}} \oplus \boldsymbol{B})^{*})^{-} = \begin{pmatrix} 17 \\ 15 \\ 9 \\ 6 \end{pmatrix}.$$

With (4), all solutions $\boldsymbol{x}$ to the problem are given by

$$\boldsymbol{x} = \begin{pmatrix} 0 & 2 & 8 & 11 \\ -2 & 0 & 6 & 9 \\ -8 & -6 & 0 & 3 \\ -11 & -9 & -3 & 0 \end{pmatrix} \boldsymbol{u}, \qquad \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \le \boldsymbol{u} \le \begin{pmatrix} 17 \\ 15 \\ 9 \\ 6 \end{pmatrix}.$$

References

[1] E. L. Demeulemeester, W. S. Herroelen. *Project Scheduling*, volume 49 of *International Series in Operations Research and Management Science. Springer, New York,*, 2002.

[2] K. Neumann and C. Schwindt and J. Zimmermann. *Project Scheduling with Time Windows and Scarce Resources.*. Springer, Berlin, 2003.

[3] V.N. Kolokoltsov and V.P. Maslov. *Idempotent analysis and its applications*, volume 401 of *Mathematics and its Applications*. Kluwer Acad. Publ., Dordrecht, 1997.

[4] J.S. Golan. *Semirings and Affine Equations Over Them*, volume 556 of *Mathematics and its Applications*. Kluwer Acad. Publ., Dordrecht, 2003.

[5] B. Heidergott, G. J. Olsder, and J. van der Woude. *Max-plus at Work*. Princeton Series in Applied Mathematics. Princeton Univ. Press, Princeton, NJ, 2006.

[6] N. Krivulin. Extremal properties of tropical eigenvalues and solutions to tropical optimization problems. *Linear Algebra ans its Applications*, 468:211-232, 2015.

[7] N. Krivulin. Tropical optimization problems with application to project scheduling with minimum makespan. *Annals of Operations Research*, 256(1):75-92, 2015.

[8] N. Krivulin. A multidimensional tropical optimization problem with nonlinear objective function and linear constraints. *Optimization*, 64(5):1107–1129, 2015.

N. Krivulin
Faculty of Mathematics and Mechanics
Saint Petersburg State University
Saint Petersburg, Russia
e-mail: nkk@math.spbu.ru

M. Petrakov
Faculty of Mathematics and Mechanics
Saint Petersburg State University
Saint Petersburg, Russia
e-mail: petrakov.9611@gmail.com

Surface electromagnetic waves

Oleg Bikeev, Oleg Kroytor and Mikhail Malykh

Abstract. In the report we discuss surface electromagnetic waves propagating along the boundary of isotropic and anisotropic media. We show how these waves can be investigated in CAS Sage.

In the 1980s, surface waves were discovered that propagate along the interface between two dielectrics without loss [1, 2]. In [3, 4] the first analytical expressions were obtained manually for solutions that are waves propagating along the interface of an anisotropic medium with permittivity

$$\epsilon = \mathrm{diag}(\epsilon_o, \epsilon_o, \epsilon_e).$$

and isotropic medium with constant permittivity ϵ.

For definiteness, let the plane $x = 0$ serve as the interface. The field in the anisotropic medium ($x < 0$) is sought in the form

$$\vec{E} = \left(a_o \vec{E}_o e^{p_o x} + a_e \vec{E}_e e^{p_e x}\right) e^{ik_y y + ik_z z - i\omega t},$$
$$\vec{H} = \left(a_o \vec{H}_o e^{p_o x} + a_e \vec{H}_e e^{p_e x}\right) e^{ik_y y + ik_z z - i\omega t}.$$

Here ω is the circular frequency of the wave, $k_0 = \omega/c$ is the wave number, $\vec{k}_\perp = (0, k_y, k_z)$ is its wave vector, a_o, a_e is the amplitude of two partial waves, and positive numbers p_o, p_e characterize the rate of wave decay in the anisotropic medium. Maxwell's equations give

$$p_o^2 = k_y^2 + k_z^2 - \epsilon_o k_0^2$$
$$p_e^2 = k_y^2 + \frac{\epsilon_e}{\epsilon_o} k_z^2 - \epsilon_e k_0^2$$

and for the vectors $\vec{E}_o, \ldots, \vec{H}_e$, explicit expressions are obtained, which we will not present here.

For the isotropic medium ($x > 0$) the field is described by similar formulas

$$\vec{E} = \left(b_o \vec{E}'_o + b_e \vec{E}'_e\right) e^{-px} e^{ik_y y + ik_z z - i\omega t},$$
$$\vec{H} = \left(b_o \vec{H}'_o + b_e \vec{H}'_e\right) e^{-px} e^{ik_y y + ik_z z - i\omega t},$$

but now the constant p, which characterizes the decrease in the field in the isotropic medium, turns out to be the same:

$$p^2 = k_y^2 + k_z^2 - \epsilon k_0^2.$$

The conditions for matching electromagnetic fields at the interface lead to a system of homogeneous linear equations for the amplitudes a_o, a_e, b_o, b_e. The condition of zero determinant of this system gives the equation

$$\left((k_z^2 - \epsilon k_0^2)p_o + (k_z^2 - \epsilon_o k_0^2)p\right)\left((k_z^2 - \epsilon k_0^2)\epsilon_o p_e + (k_z^2 - \epsilon_o k_0^2)\epsilon p\right) = (\epsilon_o - \epsilon)^2 k_y^2 k_z^2 k_0^2 \tag{1}$$

Thus, we manage to reduce the study of the existence of surface waves to a purely algebraic problem: if in the domain of variation of five variables $k_y k_z p_o p_e p$, in the region specified by the inequalities

$$p_o > 0, \quad p_e > 0, \quad p > 0,$$

the system of algebraic equations

$$\begin{cases} p_o^2 = k_y^2 + k_z^2 - \epsilon_o k_0^2 \\ p_e^2 = k_y^2 + \dfrac{\epsilon_e}{\epsilon_o} k_z^2 - \epsilon_e k_0^2 \\ p^2 = k_y^2 + k_z^2 - \epsilon k_0^2 \end{cases}$$

together with Eq. (1) has a solution, then this solution corresponds to a field satisfying Maxwell's equations, matching conditions at the interface, and exponentially decreasing with distance from the interface.

It is possible to eliminate k_0 from this system by assuming

$$p = k_0 q, \quad p_o = k_0 q_o, \quad p_e = k_0 q_e$$

and

$$k_y = k_0 \beta, \quad k_z = k_0 \gamma.$$

Then the system of equations is written in the form

$$\begin{cases} q_o^2 = \beta^2 + \gamma^2 - \epsilon_o \\ q_e^2 = \beta^2 + \dfrac{\epsilon_e}{\epsilon_o}\gamma^2 - \epsilon_e \\ q^2 = \beta^2 + \gamma^2 - \epsilon \\ \left((\gamma^2 - \epsilon)q_o + (\gamma^2 - \epsilon_o)q\right)\left((\gamma^2 - \epsilon)\epsilon_o q_e + (\gamma^2 - \epsilon_o)\epsilon q\right) = (\epsilon_o - \epsilon)^2 \beta^2 \gamma^2. \end{cases} \tag{2}$$

This system defines a curve in the space $\beta\gamma q q_o q_e$ and we are interested to know whether this curve falls within the region

$$q > 0, \quad q_o > 0, \quad q_e > 0.$$

The following observation allowed us to move forward: this curve can be described relatively simply if we consider its projection not on the β, γ plane,

which we tried to do first of all, but on the $q_o q_e$ plane. It turned out that the curve has genus zero and all variables are expressed in radicals through the value

$$t = q_e/q_o.$$

The talk will be devoted to the study of this curve in the Sage computer algebra system.

We believe that the use of computer algebra methods will make it possible to investigate the essentially algebraic question of the existence of surface waves as completely as it deserves due to its obvious applied significance [4].

References

[1] F. N. Marchevskii, V. L. Strizhevskii, and S. V. Strizhevskii. Singular electromagnetic waves in bounded anisotropic media // Sov. Phys. Solid State 26, 857 (1984).

[2] M. I. Dyakonov. New type of electromagnetic wave propagating at an interface // Sov. Phys. JETP 67, 714 (1988).

[3] O. N. Bikeev, L. A. Sevastianov. Surface Electromagnetic Waves at the Interface of Two Anisotropic Media // RUDN journal. Seria MIF. V. 25, No 2, 2017. P. 141–148. DOI: 10.22363/2312-9735-2017-25-2-141-148

[4] O. N. Bikeev et al. Electromagnetic surface waves guided by a twist discontinuity in a uniaxial dielectric with optic axis lying in the discontinuity plane // Journal of Electromagnetic Waves and Applications. Volume 33, 2019 - Issue 15. Pages 2009-2021

Oleg Bikeev
Institute of Physical Research and Technology
Peoples' Friendship University of Russia,
Moscow, Russia

e-mail: bikeev_on@pfur.ru

Oleg Kroytor
Department of Applied Probability and Informatics
Peoples' Friendship University of Russia,
Moscow, Russia

e-mail: kroytor_ok@pfur.ru

Mikhail Malykh
Department of Applied Probability and Informatics
Peoples' Friendship University of Russia,
Moscow, Russia

e-mail: malykh_md@pfur.ru

On the calculation of a generalized inverse matrix in a domain

Gennadi Malaschonok and Ihor Tchaikovsky

Abstract. We examined the well-known algorithms for calculating the generalized inverse matrix and proposed another algorithm that avoids the accumulation of errors when rounding numbers. Therefore, it is attractive for large matrices.

Introduction

The Moore-Penrose generalized inverse matrix [1]-[3] has many applications in physics, computer science and other fields.

the matrix A^+ is called the generalized inverse of the matrix A if the following 4 equalities hold:

$$A^+ = A^+AA^+,\ A = AA^+A,\ (A^+A)^T = A^+A,\ (AA^+)^T = AA^+. \qquad (1)$$

Let a matrix $A \in F^{n\times m}$ be decomposed as follows: $A = B \cdot C$, $B \in F^{n\times k}$, $C \in F^{k\times m}$, $rank(A) = rank(B) = rank(C)$. It is easy to check that matrix

$$A^+ = C^T(CC^T)^{-1}(B^TB)^{-1}B^T, \qquad (2)$$

is the generalized inverse matrix for the matrix A. In the case of complex numbers we have to use in (1) and (2) the operation of conjugation: $A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$.

There are many possibilities to obtain the decomposition (2,). For example, you can used the QR decomposition or LU decomposition. This idea was first expressed by Vera Kublanovskaya in 1965 [4].

1. SVD algorithm

We can evaluate the complexity of Kublanovskaya algorithm. In total, 5 matrix multiplications, two matrix inversions, and one more decomposition are required. The total number of operations does not exceed $\sim 8max(n,m)^3$.

Unfortunately, classical Gaussian inversion and LU decomposition are not numerically stable for large matrices due to the accumulation of rounding errors.

Today the most popular known computational method used singular value decomposition. This method consists of two stages. In the first stage, due to the Householder reflections (or Givens rotations)[5-7], an initial matrix is reduced to the upper bidiagonal form (the Golub-Kahan bidiagonalization algorithm).

The second stage is known as the Golub-Reinsch algorithm [8]. This is an iterative procedure which with the help of the Givens rotations generates a sequence of bidiagonal matrices converging to a diagonal form. This allows to obtain an iterative approximation to the singular value decomposition of the bidiagonal matrix.

In the paper [9] was presented a new finite recursive numerical algorithm for obtaining explicit rational expressions for the generalized inverse of bidigonal matrix. This rational algorithm has less number of operations. But the problem of stability is the main problem here.

2. New approach

We propose a different approach for calculating the generalized inverse matrix. It guarantees the stability of rational computing. Our approach is based on LDU matrix decomposition [10],[11]. As in the LU decomposition, we can use equality (2). The main difference is that at each step we operate with some elements of the commutative ring, which are the minors of the original matrix. Therefore, all operations are performed accurately.

Moreover, we can refuse to use the expression (2) if we use the algorithm [11], which, in addition to the factors L, D, U, calculates their inverse matrices M and W: We denote by D the weighted truncated permutation matrix. The truncated permutation matrix E can be obtain from the matrix D by replacing each non-sero element by unit element. I_D - is a diagonal matrix with unit elements which stand at the non-zero rows of matrix D. J_D - is a diagonal matrix with unit elements which stand at the non-zero columns of matrix D. The generalized inverse matrix D^+ can be obtain by transposition of matrix D and inverse each of the non-zero element of D^T. Each of these matrices has rank the same as matrix A. We denote by the sign $\hat{}$ the full rank matrices which obtained by replacing zero block by unit block. So we can write the equalities: $I_D\hat{D} = D$, $DJ_D = D$ and like these.

With these denotes we can write generalized inverse of matrix A as follows:

$$A^+ = U^{-1}D^+L^{-1} = W\hat{E}^T\hat{D}D^+\hat{D}\hat{E}^TM = W\hat{E}^TD\hat{E}^TM$$

$$AA^+A = LDD^+DU = A,$$

$$A^+AA^+ = W\hat{E}^T\hat{D}D^+DD^+\hat{D}\hat{E}^TM = A^+$$

$AA^+ = LDD^+L^{-1} = LT_DL^{-1} = I_D$ - is a diagonal matrix with unit elements which stand at the non-zero rows of matrix D.

$A^+A = U^{-1}D^+DU = U^{-1}J_DU = J_D$ - is a diagonal matrix with unit elements which stand at the non-zero columns of matrix D.

Everybody can use Math Partner cloud mathematical service [12],[13]. This is a program, which can demonstrate our algorithm :

$$SPACE = Z[]; A = [[2,3],[4,6]]; R = \backslash LDUWK(A); f = \backslash elementOf(R);$$

$$W = f\_\{4\}; D = f\_\{2\}; M = f\_\{6\}; GI = W * D * M; Ch = A * GI * A;$$

Matrix GI is the generalise inverse matrix for A and matrix Ch must be equals A.

3. Algorithm of dichotomious LDU decomposition

This is a short dexcription of dichotomious block-recursive algorithm in commutative domain:

$$(L, D, U, M, W, \hat{I}_D, \hat{J}_D, det) = \mathbf{LDU}(A, \alpha).$$

Let matrix A has size $n \times n$, $n = 2^p$. If n=1 then if(A=0) return $(1,0,1,1,1,1,1,\alpha)$ else return $(A, A^{-1}, A, \alpha, \alpha, A^{-1}, A^{-1}, A)$.

For the case $n \geq 2$ we can get 4 equal blocks of matrix: $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$,

Let us compute decomposition of the block A_{11}:

$$(L_{11},, D_{11}U_{11}, M_{11}, W_{11}, \hat{I}_{D11}, \hat{J}_{D11}, det_{11}) = \mathbf{LDU}(A_{11}, \alpha),$$

then compute the blocks A''_{12}, A''_{21} and obtain their decomposition

$A''_{12} = \bar{I}_{11}M_{11}A_{12}$, $A''_{21} = A_{21}W_{11}\bar{J}_{11}$,

$$(L_{12},, D_{12}, U_{12}, M_{12}, W_{12}, \hat{I}_{D12}, \hat{J}_{D12}, det_{12}) = \mathbf{LDU}(A''_{12}, det_{11})$$

$$(L_{21},, D_{21}, U_{21}, M_{21}, W_{21}, \hat{I}_{D21}, \hat{J}_{D21}, det_{21}) = \mathbf{LDU}(A''_{21}, det_{11})$$

$A'_{22} = A_{22} - \alpha^{-2}A_{21}W_{11}E_{11}(\bar{I}_{11} + \hat{I}_{D11}M_{11})A_{12}$

$\alpha_s = det_{12}det_{21}/det_{11}$, $A'''_{22} = \alpha_s\bar{I}_{21}M_{21}A'_{22}W_{12}\bar{J}_{12}$

And now we calculate the decomposition of the block A'''_{22}:

$$(L_{22},, D_{22}U_{22}, M_{22}, W_{22}, \hat{I}_{D22}, \hat{J}_{D22}, det_{22}) = \mathbf{LDU}(A'''_{22}, \alpha_s)$$

Then $L = \begin{pmatrix} L_{11}L_{12} & 0 \\ L_3 & L_{21}L_{22} \end{pmatrix}$, $D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$, $U = \begin{pmatrix} U_{21}U_{11} & U_2 \\ 0 & U_{22}U_{12} \end{pmatrix}$,
with $L_3 = A_{21}W_{11}I_{11} + A_{22}W_{12}I_{12}$ and $U_2 = J_{21}M_{21}A_{22} + J_{11}M_{11}A_{12}$,

$\hat{I}_D = \begin{pmatrix} \hat{I}_{D12}\hat{I}_{D11} & 0 \\ 0 & \hat{I}_{D22}\hat{I}_{D21} \end{pmatrix}$, $M = \begin{pmatrix} M_{12}M_{11} & 0 \\ M_3 & M_{22}M_{21} \end{pmatrix}$

$W = \begin{pmatrix} W_{11}W_{21} & W_2 \\ 0 & W_{12}W_{22} \end{pmatrix}$, $\hat{J}_D = \begin{pmatrix} \hat{J}_{D11}\hat{J}_{D21} & 0 \\ 0 & \hat{J}_{D12}\hat{J}_{D22} \end{pmatrix}$

With $W_2 = -(W_{21}J_{D21}M_{21}A'_{22} + W_{11}J_{D11}M_{11}A_{12})W_{12}W_{22}$
and $M_3 = -M_{22}M_{21}(A_{21}W_{11}\hat{I}_{D11}M_{11} + A'_{22}W_{12}\hat{I}_{D12}M_{12})$.

References

[1] B. Penrose A generalized inverse for matrices. Proe. Cambridge Philos. Soc, 51 (1955).

[2] A. Ben-Israel and T. N. E. Greville, Generalized Inverses. Theory and Applications, 2nd ed. Springer, New York, (2003).

[3] A. S. Hauscholder . The theory of matrices in numerical analysis. Boston, Ginn and Co., (1963).

[4] V.N.Kublanovskaya. Evaluation of a generalized inverse matrix and projector. USSR Computational Mathematics and Mathematical Physics, Volume 6, Issue 2, 179-188, (1966)

[5] G. H. Golub and C. Reinsch, Singular Value Decomposition and Least Squares Solutions. Numer. Math. 14, 403–420, (1970).

[6] G. H. Golub and Ch. F. van Loan. Matrix Computations, 3rd ed. The John Hopkins University Press, (1996).

[7] J. W. Lewis, Invertion of tridiagonal matrices. Numer.Math., 38, 333–345 (1982).

[8] G. H. Golub and W. Kahan., Calculating the Singular Values and Pseudo-Inverse of a Matrix. SIAM J. Num. Anal., 2, 205–224. (1965)

[9] Yu. Hakopian. Computing the Moor-Penrose inverse for bidiagonal matrices.Mohyla Mathematical Journal. National University of Kyiv-Mohyla Academy. V.2, 11-23, (2019)

[10] G. Malashonok, Generalized Bruhat decomposition in commutative domains. Computer Algebra in Scientific Computing, CASC'2013, LNCS 8136, Springer, Heidelberg, 2013, 231–242, (2013)

[11] G. Malashonok, A. Scherbinin. Triangular Decomposition of Matrices in a Domain, Computer Algebra in Scientific Computing, LNCS 9301, Springer, Switzerland, 2015, 290–304, (2015)

[12] G.I. Malaschonok. MathPartner Computer Algebra, *Programming and Computer Software*, V. 43, No. 2, 112–118, (2017)

[13] Math Partner Cloud Mathematical Service. http://mathpar.ukma.edu.ua

Gennadi Malaschonok
National University of Kyiv-Mohyla Academy
Kiev, Ukraine
e-mail: malaschonok@gmail.com

Ihor Tchaikovsky
Lviv Polytechnic National University
Institute of Telecommunications, Radioelectronics and Electronic Engineering
Lviv, Ukraine
e-mail: ihortch@yahoo.com

Dynamic systems with quadratic integrals

Ali Baddour, Mikhail Malykh, Leonid Sevastianov and Yu Ying

Abstract. In the report we discuss the problems of constructing difference schemes that mimic the properties of dynamic systems. We show how these problems can be solved in systems with quadratic integrals and how a many-body problem can be reduced to such systems.

One of the most widespread mathematical models is a dynamic system described by an autonomous system of ordinary differential equations, i.e., the system of the form

$$\frac{dx_i}{dt} = f_i(x_1, \ldots, x_n), \quad i = 1, 2, \ldots n, \tag{1}$$

where t is an independent variable, commonly interpreted as time, and the variables $x_1, \ldots, x_n$ depending on it as coordinates of a point of several points. In applications the sight-hand sides f_i are often rational or algebraic functions of the coordinates $x_1, \ldots, x_n$ or can be reduced to such form using a certain change of variables. As a rule, from physical reasons a few integrals of motion are known, but they are not sufficient to reduce the system of differential equations to Abel quadratures.

For example, the classical problem of n bodies [1] consists in finding solutions to the autonomous system of ordinary differential equations

$$m_i \ddot{\vec{r}}_i = \sum_{j=1}^{n} \gamma \frac{m_i m_j}{r_{ij}^3} \left(\vec{r}_j - \vec{r}_i\right), \quad i = 1, \ldots, n \tag{2}$$

Here $\vec{r}_i$ is the radius vector of the i-th body and r_{ij} is the distance between the i-th and j-th body. This dynamic system is a Hamiltonian system of the order $2 \cdot 3 \cdot n$. For reducing it to quadratures using the Liouville method it is necessary to find $3 \cdot n$ algebraic integrals of motion in involution [2]. At the time of Liouville, only ten independent algebraic integrals of the many-body problem were known, which were called classical. In the 1880s, Bruns proved that every other algebraic integral of this problem is expressed in terms of these ten [2, 3]. This means that the many-body problem cannot be reduced to quadratures by the Liouville method. The question of whether it can be reduced to Abelian quadratures in another way was formulated by Bruns himself and resolved negatively [3, n. 23].

Classical explicit difference schemes, including explicit Runge-Kutta schemes, do not preserve these integrals. However, among implicit difference schemes there are schemes that preserve some classes of integrals of motion. The most studied are symplectic Runge-Kutta schemes that preserve all quadratic integrals of motion. For example, for a linear oscillator or a system of several coupled oscillators, these schemes allow organizing the calculation of the approximate solution in such a way that all the integrals of this system are preserved. In this case, the approximate solution mimics the periodicity of the exact solution, for example, you can choose a time step so that the approximate solution is a periodic sequence [4].

The construction of such mimetic schemes in the case of nonlinear dynamical systems is complicated by the appearance of non-quadratic integrals. For example, in the classical many-body problem by the Bruns theorem, there are 10 independent algebraic integrals, of which 9 are quadratic and therefore are preserved using any symplectic scheme. The first finite-difference scheme for the many-body problem, preserving all classical integrals of motion, was proposed in 1992 by Greenspan [5, 6] and independently in somewhat different form by J.C. Simo and O. González [7, 8]. The Greespan scheme is a kind of combination of the midpoint method and discrete gradient method.

In other site the standard symplectic schemes will preserve all integrals if we introduce the new variables such a way that all classical integrals are quadratic with respect of new variables. This approach is close to the invariant energy quadratization method (IEQ method) which was first proposed by Yang et al. [9] and used by Hong Zhang et al. [10] to conserve the energy at discretization of Hamiltonian systems including Kepler two-body problem. We applicate the same idea in many body problem.

First of all, we get rid of irrationality by introducing new variables r_{ij}, related to the coordinates by the equation

$$r_{ij}^2 - (x_i - x_j)^2 - (y_i - y_j)^2 - (z_i - z_j)^2 = 0$$

Then we eliminate the denominators in the energy integral by introducing new variables ρ_{ij}, related to the already introduced ones by the equations

$$r_{ij}\rho_{ij} = 1.$$

Note that this relation is quadratic again, so that after introducing additional variables this relation will turn into an additional quadratic integral.

For the sake of brevity let us denote the velocity components of the i-th body as $\dot{x}_i = u_i$, $\dot{y}_i = v_i$, and $\dot{z}_i = w$ and combine them into vector $\vec{v}_i$. From the many-body problem we pass to a system that consists of three coupled subsystems, namely, the system for coordinates

$$\dot{\vec{r}}_i = \vec{v}_i, \quad i = 1, \ldots, n \tag{3}$$

the system for velocities

$$m_i \dot{\vec{v}}_i = \sum_{j=1}^{n} \gamma \frac{m_i m_j \rho_{ij}}{r_{ij}^2} \left(\vec{r}_j - \vec{r}_i\right), \quad i = 1, \ldots, n \tag{4}$$

the system for distances

$$\dot{r}_{ij} = \frac{1}{r_{ij}}(\vec{r}_i - \vec{r}_j) \cdot (\vec{v}_i - \vec{v}_j), \quad i, j = 1, \ldots, n;\ i \neq j. \tag{5}$$

and the system for inverse distances

$$\dot{\rho}_{ij} = -\frac{\rho_{ij}}{r_{ij}^2}(\vec{r}_i - \vec{r}_j) \cdot (\vec{v}_i - \vec{v}_j), \quad i, j = 1, \ldots, n;\ i \neq j. \tag{6}$$

This system possesses 10 classical integrals of the many-body problem and additional integrals

$$r_{ij}^2 - (x_i - x_j)^2 - (y_i - y_j)^2 - (z_i - z_j)^2 = \text{const}, \quad i \neq j \tag{7}$$

and

$$r_{ij}\rho_{ij} = \text{const}, \quad i \neq j. \tag{8}$$

The autonomous system of differential equations (3)-(6), involving $n(n-1)$ additional variables r_{ij} and ρ_{ij}, has the following properties:

1. this system has quadratic integrals of motion (7) and (8), that allow expressing the additional variables r_{ij} and ρ_{ij} in terms of the coordinates of the bodies,
2. if the constants in these integrals are chosen such that

   $$r_{ij}^2 - (x_i - x_j)^2 - (y_i - y_j)^2 - (z_i - z_j)^2 = 0 \quad \text{and} \quad r_{ij}\rho_{ij} = 1,$$

   the solutions of the new system coincide with the solutions of the original one,
3. the new system has quadratic integrals of motion, which, with the relation between the additional variables and the coordinates of the bodies taken into account, turn into 10 classical integrals of the many-body problem.

Since all the classical integrals of the many-body problem, as well as the additional integrals in the new variables are quadratic, any symplectic Runge-Kutta difference scheme, including the simplest of them, the midpoint scheme, preserves all these integrals for sure.

Moreover, the autonomous system of differential equations (3)-(6) preserves the symmetry of the original problem with respect to permutations of bodies and time reversal, as, for example, the midpoint scheme.

At each step of the midpoint scheme, new values will be determined not only for the coordinates and velocities of the bodies, but also for auxiliary quantities r_{ij} and ρ_{ij}. If at the initial moment of time only the coordinates and velocities were specified, and the auxiliary variables were defined by equalities

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}, \quad \rho_{ij} = \frac{1}{r_{ij}},$$

then these equalities are preserved exactly (maybe up to the radical's signs) due to the fact that the auxiliary integrals (7) and (8) are quadratic and are preserved exactly when using the midpoint scheme; therefore, the quantities r_{ij} and ρ_{ij} do

not lose their original meaning of the distances between the bodies and the inverse distances between the bodies.

Therefore, the midpoint scheme written for the system (3) - (6) preserves all its algebraic integrals exactly and is invariant under permutations of bodies and time reversal.

The report will present the results of numerical experiments with a midpoint scheme with an emphasis on its mimetic character, see also [11].

It should also be emphasized that the many-body problem has been reduced to the problem, all of whose integrals are quadratic. Another classical mechanical problem, the gyroscope rotation problem, has the same properties. The same problems arise when introducing classical transcendental functions elliptic and Abelian. Therefore, we intend to investigate in more detail dynamical systems with quadratic integrals.

References

[1] *Marshal Ch.* The three-body problem. Elsivier, 1990.

[2] *Whittaker E.T.* A Treatise on the Analytical Dynamics of Particles and Rigid Bodies. 4 ed. Cambridge University Press., 1989.

[3] *Bruns H.* Über die Integrale des Vielkörper-problems // Acta Mathematica. — 1887 — Vol. 11. — P. 25-96.

[4] *Gerdt V.P. et al.* On the properties of numerical solutions of dynamical systems obtained using the midpoint method // Discrete and Continuous Models and Applied Computational Science. — 2019 — Vol. 27, no. 3. — P. 242-262. DOI: 10.22363/2658-4670-2019-27-3-242-262 .

[5] *Greenspan D.* Completely Conservative and Covariant Numerical Methodology for N-Body Problems With Distance-Dependent Potentials. — 1992. — Technical Report no. 285 at http://hdl.handle.net/ 10106/2267.

[6] *Greenspan D.* N-Body Problems and Models. — World Scientific, 2004.

[7] *Simo J. C., González Mónica Alegre.* Assessment of Energy-momentum and Symplectic Schemes for Stiff Dynamical Systems // American Society of Mechanical Engineers. — 1993.

[8] *Graham E., Jelenić G., Crisfield M. A.* A note on the equivalence of two recent time-integration schemes for N-body problems // Communications in Numerical Methods in Engineering. — 2002. — Vol. 18. — P. 615–620.

[9] *Yang Xiaofeng, Ju Lili.* Efficient linear schemes with unconditional energy stability for the phase field elastic bending energy model // Computer Methods in Applied Mechanics and Engineering. — 2016. — 11. — Vol. 315.

[10] *Zhang Huarong, Qian Xu, Song Songhe.* Novel high-order energy-preserving diagonally implicit Runge-Kutta schemes for nonlinear Hamiltonian ODEs // Appl. Math. Lett. — 2020. — Vol. 102. — P. 106091.

[11] *Gerdt V.P. et al.* On conservative difference schemes for the many-body problem // ArXiv. 2007.01170.

Ali Baddour
Department of Applied Probability and Informatics
Peoples' Friendship University of Russia,
Moscow, Russia

e-mail: alibddour@gmail.com

Mikhail Malykh
Department of Applied Probability and Informatics
Peoples' Friendship University of Russia,
Moscow, Russia

e-mail: malykh_md@pfur.ru

Leonid Sevastianov
Department of Applied Probability and Informatics
Peoples' Friendship University of Russia,
Moscow, Russia

e-mail: sevastianov_la@pfur.ru

Yu Ying
Department of Algebra and Geometry,
Kaili University,
Kaili, China

e-mail: yuying05720062@sina.com

Growth in groups and the number of curves and knots

Andrei Malyutin

Abstract. We use results of Vershik, Nechaev, and Bikbov on growth of random heaps to improve known lower bounds on the rate of growth of the number of knots with respect to the crossing number.

We study the structure and statistical characteristics of the set of classical knots (see [M15, M18, M18b, M19, BM19] and references therein). A particular point of this study is the growth rate of the number of knots with respect to various complexity measures on the set of knots. Historically, the crossing number is considered as the most natural knot complexity measure. The growth rate of the number of knots with respect to the crossing number is studied in particular in [ES87, W92, STh98, Th98, St04, Ch18]. Being applied to the sequence $(K_1, K_2, K_3, ...)$, where K_n denote the number of knots of n crossings,

results of [ES87] imply that[1]

$$2.13... \quad \leq \quad \liminf_{n\to\infty} \sqrt[n]{K_n};$$

results of [W92] imply that

$$\limsup_{n\to\infty} \sqrt[n]{K_n} \quad \leq \quad 13.5;$$

results of [STh98] imply that

$$\limsup_{n\to\infty} \sqrt[n]{K_n} \quad \leq \quad \frac{101+\sqrt{21001}}{20} = 12.29...;$$

The research was partially supported by the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS" and by RFBR according to the research project n. 20-01-00070.

[1]The lower bound $2.68 \leq \liminf_{n\to\infty} \sqrt[n]{K_n}$ given in [W92] as an interpretation of results obtained in [ES87] seems to be a typo.

and results of [St04] imply that

$$\limsup_{n\to\infty} \sqrt[n]{K_n} \quad \leq \quad \frac{91+\sqrt{13681}}{20} = 10.39...$$

Thus, the record asymptotic estimates, presented in the literature, for the growth rate of K_n are

$$2.13... \quad \leq \quad \liminf_{n\to\infty} \sqrt[n]{K_n} \quad \leq \quad \limsup_{n\to\infty} \sqrt[n]{K_n} \quad \leq \quad 10.39...$$

For some reasons explained in [M18] it would be useful to find bounds a and b such that

$$a \quad \leq \quad \liminf_{n\to\infty} \sqrt[n]{K_n} \quad \leq \quad \limsup_{n\to\infty} \sqrt[n]{K_n} \quad \leq \quad b$$

and

$$a^3 \quad > \quad b^2.$$

It turns out that a new lower bound

$$4 \quad \leq \quad \liminf_{n\to\infty} \sqrt[n]{K_n}$$

is implied by the results of [V00, VNB00] on the growth rate of locally free semigroups (heaps). To obtain this bound, we construct embeddings of locally free semigroups into the set of knots. Furthermore, passing to more complex semigroups with weighted elements, we show that

$$4.45 \quad \leq \quad \liminf_{n\to\infty} \sqrt[n]{K_n}.$$

Moreover, we have the same lower bound for the case of alternating prime knots.

References

[BM19] Belousov, Yu. and A. Malyutin. "Hyperbolic knots are not generic." (2019): preprint arXiv:1908.06187.

[Ch18] Chapman, H. "On the structure and scarcity of alternating knots." (2018): preprint arXiv:1804.09780.

[ES87] Ernst, C. and D. W. Sumners. "The growth of the number of prime knots." *Math. Proc. Cambridge Philos. Soc.* 102, no. 2 (1987): 303–315.

[M15] Malyutin, A. V. "Satellite knots strike back." *Abstracts of the International Conference "Polynomial Computer Algebra '15"* (2015): 65–66.

[M18] Malyutin, A. V. "On the question of genericity of hyperbolic knots." *Int. Math. Res. Not.* (2018). `https://doi.org/10.1093/imrn/rny220`.

[M18b] Malyutin, A. V. "What does a random knot look like?" *Abstracts of the International Conference "Polynomial Computer Algebra '18"* (2018): 69–71.

[M19] Malyutin, A. V. "Hyperbolic links are not generic." (2019): preprint arXiv:1907.04458.

[St04] Stoimenow, A. "On the number of links and link polynomials." *Q. J. Math.* 55, no. 1 (2004): 87–98.

[STh98] Sundberg, C. and M. B. Thistlethwaite. "The rate of growth of the number of prime alternating links and tangles." *Pacific J. Math.* 182, no. 2 (1998): 329–358.

[Th98] Thistlethwaite, M. B. "On the structure and scarcity of alternating links and tangles." *J. Knot Theory Ramifications* 7, no. 7 (1998): 981–1004.

[V00] Vershik, A. M. "Dynamic theory of growth in groups: Entropy, boundaries, examples." *Uspekhi Mat. Nauk* 55, no. 4(334) (2000): 59–128; *Russian Math. Surveys* 55, no. 4 (2000): 667–733.

[VNB00] Vershik, A. M., S. Nechaev, and R. Bikbov. "Statistical properties of locally free groups with applications to braid groups and growth of random heaps." *Comm. Math. Phys.* 212, no. 2 (2000): 469–501.

[W92] Welsh, D. J. A. "On the number of knots and links." In *Sets, Graphs and Numbers (Proceedings of 1991 Budapest conference)*, 713–718. Colloq. Math. Soc. János Bolyai 60. Amsterdam: North-Holland, 1992.

Andrei Malyutin
St. Petersburg Department of
Steklov Institute of Mathematics
St. Petersburg State University
St. Petersburg, Russia
e-mail: malyutin@pdmi.ras.ru

Visualization uniform discrete subgroup of the $SO(3)$ group

Mityushov E. and Misyura N.

Abstract. The problem of generating uniform deterministic samples on a rotation group $SO(3)$ is fundamental for many fields, such as computational structural biology, robotics, computer graphics, astrophysics. In this paper, we propose a method for constructing uniform discrete subgroups of a group $SO(3)$ using the coordination of vertices of regular four-dimensional polyhedra. In particular, an algorithm for finding a uniform distribution of 60 points in a ball of radius π identified with elements uniform discrete subgroup of a group $SO(3)$ is presented.

1. Equations

As is known, the group of unit quaternions $Sp(1)$ covers the rotation group $SO(3)$ in two sheets. Each quaternion $q = [q_0, q_1, q_2, q_3] \in Sp(1)$ corresponds to a point on the sphere $S^3 \subset R^4$:

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1.$$

Since the same element of the group $SO(3)$ corresponds to quaternions q and $-q$, discarding the diametrically opposite points of the sphere S^3, we can find the coordinates of all elements of the group $SO(3)$.

The obtaining uniform discrete subgroup of a group $Sp(1)$ is proved by the existence of five centrosymmetric regular four-dimensional polyhedra inscribed in the unit hypersphere $S^3 \subset R^4$. These polyhedra are (the number of vertices is indicated in parentheses): tesseract (16), 16 - cell (8), 24 – cell (24), 120 - cell (600), 600 - cell (120).

The Cartesian coordinates of the vertices of regular polyhedra in are determined by the following groups of permutations:

Tesseract

The coordinates of its 16 peaks are all sorts of permutations $(\pm\frac{1}{2}; \pm\frac{1}{2}; \pm\frac{1}{2}; \pm\frac{1}{2})$.

16 - cell

The coordinates of its 8 peaks are all sorts of permutations $(\pm 1, 0, 0, 0)$.

24 - cell

The coordinates of its 8 peaks are all sorts of permutations $(\pm 1, 0, 0, 0)$.

The coordinates of the 16 peaks are all sorts of permutations $(\pm\frac{1}{2}; \pm\frac{1}{2}; \pm\frac{1}{2}; \pm\frac{1}{2})$.

120 - cell

The coordinates of 24 of its vertices are all sorts of permutations The coordinates of the 16 peaks are all sorts of permutations $(0; 0 \pm \frac{1}{\sqrt{2}}; \pm\frac{1}{\sqrt{2}})$

The coordinates of 64 vertices are all sorts of permutations $(\pm\frac{1}{\sqrt{8}}; \pm\frac{1}{\sqrt{8}}; \pm\frac{1}{\sqrt{8}}; \pm\sqrt{\frac{5}{8}})$

The coordinates of 64 vertices are all sorts of $(\pm\frac{\Phi^{-2}}{\sqrt{8}}; \pm\frac{\Phi}{\sqrt{8}}; \pm\frac{\Phi}{\sqrt{8}}; \pm\frac{\Phi}{\sqrt{8}})$, where $\Phi = \frac{1+\sqrt{5}}{2}$ is the ratio of the golden section.

The coordinates of 64 vertices are all sorts of permutations $(\pm\frac{\Phi^{-1}}{\sqrt{8}}; \pm\frac{\Phi^{-1}}{\sqrt{8}}; \pm\frac{\Phi^{-1}}{\sqrt{8}}; \pm\frac{\Phi^{-2}}{\sqrt{8}})$.

The coordinates of 96 vertices are all sorts of even permutations $(0; \pm\frac{\Phi^{-2}}{\sqrt{8}}; \frac{1}{\sqrt{8}}; \pm\frac{\Phi^{2}}{\sqrt{8}})$.

The coordinates of 96 vertices are all sorts of even permutations $(0; \pm\frac{\Phi^{-1}}{\sqrt{8}}; \frac{\Phi}{\sqrt{8}}; \pm\sqrt{\frac{5}{8}})$.

The coordinates of the remaining 192 peaks - all sorts of clear permutations $(\pm\frac{\Phi^{-1}}{\sqrt{8}}; \pm\frac{1}{\sqrt{8}}; \pm\frac{\Phi}{\sqrt{8}}; \pm\frac{1}{\sqrt{2}})$.

600 - cell

The coordinates of its 8 peaks are all sorts of permutations $(\pm 1, 0, 0, 0)$.

The coordinates of the 16 peaks are all sorts of permutations $(\pm\frac{1}{2}; \pm\frac{1}{2}; \pm\frac{1}{2}; \pm\frac{1}{2})$.

The coordinates of the remaining 96 vertices are all sorts of even regular polyhedra permutations $(\pm\frac{\Phi}{2}; \pm\frac{1}{2}; \pm\frac{\Phi^{-1}}{2}; 0)$.

By discarding diametrically opposite vertices, we find that four-dimensional regular polyhedra generate five uniform finite subgroups of a group $SO(3)$ with four, eight, twelve, sixty and three hundred elements.

As an example, consider the construction uniform subgroup of a group $SO(3)$ generated by a 600-cell, The corresponding elements of this group, expressed in terms of the coordinates of individual quaternions, are given in Table 1.

To visualize a uniform discrete subgroup of group $SO(3)$, we use the

$$f : S^3 \to D, S^3 \subset R^4, D : |r| \leq \pi, r \subset R^3.$$

The law of mapping into a ball with radius π defined by the equalities

$$x_1^{(k)} = \frac{2q_1^{(k)} arccos(q_0^{(k)})}{\sqrt{1 - \left(q_0^{(k)}\right)^2}}, x_2^{(k)} = \frac{2q_2^{(k)} arccos(q_0^{(k)})}{\sqrt{1 - \left(q_0^{(k)}\right)^2}}, x_3^{(k)} = \frac{2q_3^{(k)} arccos(q_0^{(k)})}{\sqrt{1 - \left(q_0^{(k)}\right)^2}}.$$

Examination shows that part of the points goes beyond the ball (Fig. 1).

For elements that go beyond the radius, the sign is changed for the corresponding quaternions. As a result, all points are inside the ball (Fig. 2).

An analysis of the distribution of point patterns of elements uniform discrete subgroup of a group $SO(3)$ inside a ball of radius shows that the corresponding

points are located on four spheres with radii $\frac{\pi}{5}$, $\frac{\pi}{3}$, $\frac{2\pi}{5}$ and π. These points are located at the vertices inscribed in these spheres of the icosahedrons, dodecahedron, and icosododecahedron (Fig. 3).

2. Tables and Figures

Table 1. Element uniform subgroup of a group generated by a 600-cell.

$q^{(1)}=[1,0,0,0]$	$q^{(2)}=[0,1,0,0]$	$q^{(3)}=[0,0,1,0]$	$q^{(4)}=[0,0,0,1]$	$q^{(5)}=\left[\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right]$	$q^{(6)}=\left[\frac{1}{2},-\frac{1}{2},\frac{1}{2},\frac{1}{2}\right]$
$q^{(7)}=\left[\frac{1}{2},\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right]$	$q^{(8)}=\left[\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right]$	$q^{(9)}=\left[\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right]$	$q^{(10)}=\left[\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right]$	$q^{(11)}=\left[\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right]$	$q^{(12)}=\left[\frac{1}{2},-\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right]$
$q^{(13)}=\frac{1}{2}\left[\varphi,1,\frac{1}{\varphi},0\right]$	$q^{(14)}=\frac{1}{2}\left[\varphi,-1,\frac{1}{\varphi},0\right]$	$q^{(15)}=\frac{1}{2}\left[\varphi,0,1,\frac{1}{\varphi}\right]$	$q^{(16)}=\frac{1}{2}\left[1,\varphi,0,\frac{1}{\varphi}\right]$	$q^{(17)}=\frac{1}{2}\left[1,\frac{1}{\varphi},\varphi,0\right]$	$q^{(18)}=\frac{1}{2}\left[1,0,\frac{1}{\varphi},\varphi\right]$
$q^{(19)}=\frac{1}{2}\left[\frac{1}{\varphi},\varphi,1,0\right]$	$q^{(20)}=\frac{1}{2}\left[\frac{1}{\varphi},1,0,\varphi\right]$	$q^{(21)}=\frac{1}{2}\left[\frac{1}{\varphi},0,\varphi,1\right]$	$q^{(22)}=\frac{1}{2}\left[0,\varphi,\frac{1}{\varphi},1\right]$	$q^{(23)}=\frac{1}{2}\left[0,1,\varphi,\frac{1}{\varphi}\right]$	$q^{(24)}=\frac{1}{2}\left[0,\frac{1}{\varphi},1,\varphi\right]$
$q^{(25)}=\frac{1}{2}\left[\varphi,-1,\frac{1}{\varphi},0\right]$	$q^{(26)}=\frac{1}{2}\left[\varphi,\frac{1}{\varphi},0,-1\right]$	$q^{(27)}=\frac{1}{2}\left[\varphi,0,-1,\frac{1}{\varphi}\right]$	$q^{(28)}=\frac{1}{2}\left[-1,\varphi,0,\frac{1}{\varphi}\right]$	$q^{(29)}=\frac{1}{2}\left[-1,\frac{1}{\varphi},\varphi,0\right]$	$q^{(30)}=\frac{1}{2}\left[-1,0,\frac{1}{\varphi},\varphi\right]$
$q^{(31)}=\frac{1}{2}\left[\frac{1}{\varphi},\varphi,-1,0\right]$	$q^{(32)}=\frac{1}{2}\left[\frac{1}{\varphi},-1,0,\varphi\right]$	$q^{(33)}=\frac{1}{2}\left[\frac{1}{\varphi},0,\varphi,-1\right]$	$q^{(34)}=\frac{1}{2}\left[0,\varphi,\frac{1}{\varphi},-1\right]$	$q^{(35)}=\frac{1}{2}\left[0,-1,\varphi,\frac{1}{\varphi}\right]$	$q^{(36)}=\frac{1}{2}\left[0,\frac{1}{\varphi},-1,\varphi\right]$
$q^{(37)}=\frac{1}{2}\left[\varphi,1,-\frac{1}{\varphi},0\right]$	$q^{(38)}=\frac{1}{2}\left[\varphi,-\frac{1}{\varphi},0,1\right]$	$q^{(39)}=\frac{1}{2}\left[\varphi,0,1,-\frac{1}{\varphi}\right]$	$q^{(40)}=\frac{1}{2}\left[1,\varphi,0,-\frac{1}{\varphi}\right]$	$q^{(41)}=\frac{1}{2}\left[1,-\frac{1}{\varphi},\varphi,0\right]$	$q^{(42)}=\frac{1}{2}\left[1,0,-\frac{1}{\varphi},\varphi\right]$
$q^{(43)}=\frac{1}{2}\left[-\frac{1}{\varphi},\varphi,1,0\right]$	$q^{(44)}=\frac{1}{2}\left[-\frac{1}{\varphi},1,0,\varphi\right]$	$q^{(45)}=\frac{1}{2}\left[-\frac{1}{\varphi},0,\varphi,1\right]$	$q^{(46)}=\frac{1}{2}\left[0,\varphi,-\frac{1}{\varphi},1\right]$	$q^{(47)}=\frac{1}{2}\left[0,1,\varphi,-\frac{1}{\varphi}\right]$	$q^{(48)}=\frac{1}{2}\left[0,-\frac{1}{\varphi},1,\varphi\right]$
$q^{(49)}=\frac{1}{2}\left[\varphi,-1,-\frac{1}{\varphi},0\right]$	$q^{(50)}=\frac{1}{2}\left[\varphi,-\frac{1}{\varphi},0,-1\right]$	$q^{(51)}=\frac{1}{2}\left[\varphi,0,-1,-\frac{1}{\varphi}\right]$	$q^{(52)}=\frac{1}{2}\left[-1,\varphi,0,-\frac{1}{\varphi}\right]$	$q^{(53)}=\frac{1}{2}\left[-1,-\frac{1}{\varphi},\varphi,0\right]$	$q^{(54)}=\frac{1}{2}\left[-1,0,-\frac{1}{\varphi},\varphi\right]$
$q^{(55)}=\frac{1}{2}\left[-\frac{1}{\varphi},\varphi,-1,0\right]$	$q^{(56)}=\frac{1}{2}\left[-\frac{1}{\varphi},-1,0,\varphi\right]$	$q^{(57)}=\frac{1}{2}\left[-\frac{1}{\varphi},0,\varphi,-1\right]$	$q^{(58)}=\frac{1}{2}\left[0,\varphi,-\frac{1}{\varphi},-1\right]$	$q^{(59)}=\frac{1}{2}\left[0,-1,\varphi,-\frac{1}{\varphi}\right]$	$q^{(60)}=\frac{1}{2}\left[0,-\frac{1}{\varphi},-1,\varphi\right]$

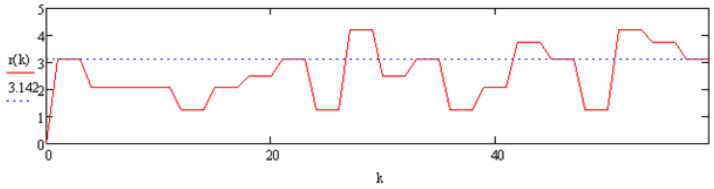


FIGURE 1. Distribution of distances from the center of the ball to point images of elements uniform subgroup of the group $SO(3)$.

Mityushov E.
Ural Federal University
Ekaterinburg, Russia
e-mail: mityushov-e@mail.ru

Misyura N.
Ural Federal University
Ekaterinburg, Russia
e-mail: n_misura@mail.ru

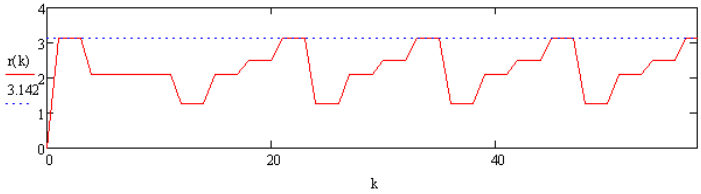


FIGURE 2. Distribution of distances from the center of the ball to point images of elements uniform discrete subgroup of a group $SO(3)$ inside a ball of radius π.

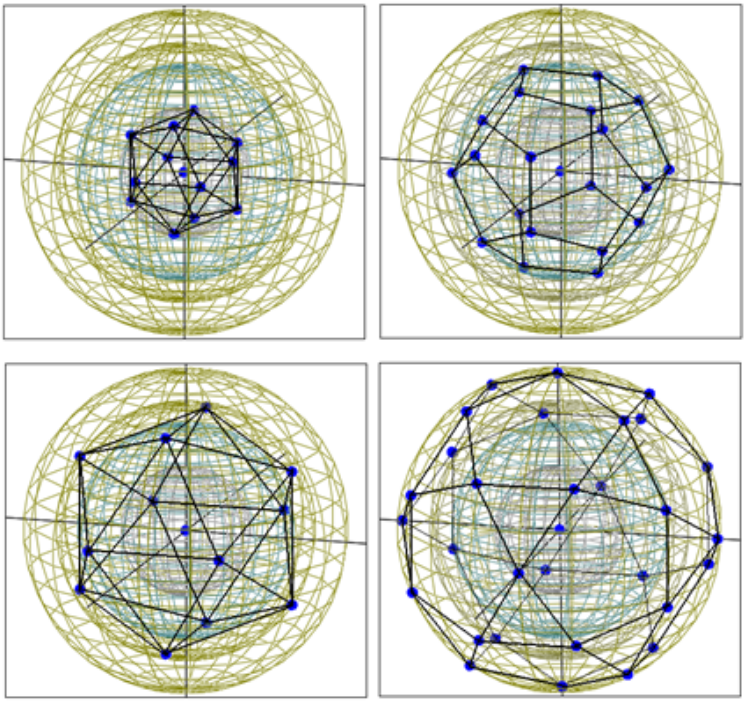


FIGURE 3. The diametrically opposite points at the vertices of the icosododecahedron define the same elements of group $SO(3)$.

Numerical Symbolic Dynamics and Complexity of Individual Trajectories

Mylläri Aleksandr, Mylläri Tatiana, Myullyari Anna and Vassiliev Nikolay

Abstract. We analyze complexity of trajectories in the free-fall equal-mass three-body problem using finite symbolic sequences received as a result of the numerical integration of equations of motion. Binary approaches are used to construct symbolic sequences. Different methods and measures of complexity of finite sequences are used and compared: Shannon and Markov entropies and Kolmogorov complexity. Selection of initial conditions for orbits with high values of the entropies was done. Examples of orbits from this selection are presented.

Introduction

Description of the problem and some history can be found in [6], [7] and [8]. Here, we study complexity of individual trajectories by analysing complexity of finite sequences received as a result of the numerical integration of equations of motion. Binary approaches are used to construct symbolic sequences. We use Shannon entropy, family of Markov entropies (see e.g. [3]) and Kolmogorov complexity to analyse complexity of the sequences constructed.

The equal mass free-fall three-body problem is convenient for study since it reduces drastically dymension of the initial conditions space and allows easy visualization of initial configuration: if we place two bodies in the points $(-0.5; 0)$ and $(0.5; 0)$, then all possible configurations will be covered if we place the third body inside region D bounded by two straight line segments and with the arc of the unit circle centered at $(-0.5, 0)$ (Fig. 1) [1]. This region is used in the following visualizations.

We used symplectic code by Seppo Mikkola (Tuorla Observatory, University of Turku) [5] for numerical simulations.

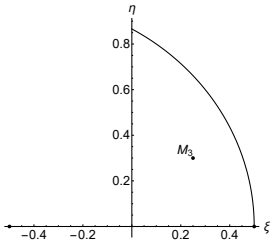


FIGURE 1. Region D of initial conditions.

1. Complexity of finite sequences

We construct symbolic sequences using binary encounters. We detect the minimum distance between two bodies, and the corresponding symbol is the number of the distant body. Some systems disrupt fast, forming a binary moving in one direction while third body escapes in the opposite direction. In this case end of the sequence consists of repeating symbol - number of the escaping body. Thus, our symbols are from the alphabet $\{1, 2, 3\}$. To have a reasonable computing time, we constructed symbolic sequences of length 256. In previous study, we were interested in the analysis of active three-body interactions, so as one approach we considered sub-sequences of each of these sequences, increasing the length step-by-step, calculating Shannon entropy for each of these sub-sequences, and finding maximum value of these entropies. Maximum values (and moment of time/length of the sub-sequence) correspond to the stage of active interaction between bodies [2]. Here, we analyse a whole sequence (making it cyclic when needed - e.g., for calculating Markov entropies). We used the length of the archive of the sequence as an estimate of Kolmogorov complexity. Family of Markov entropies can be defined as ([3])

$$H^l = -\sum_i p_i \sum_j q_{ij}^{(l)} \ln q_{ij}^{(l)}$$

where p_i is frequency of symbol i in the sequence, $q_{ij}^{(l)}$ - frequency of transitions from i to j with lag l. We calculate values of H^l for each sequence with $l = 1, 2, \ldots 255$ and chooze maximum value (we call it Markov Max entropy). We also consider standard Markov entropy ($l = 1$):

$$H^1 = -\sum_i p_i \sum_j q_{ij} \log q_{ij}$$

where q_{ij} is frequency of transitions from i to j.

2. Trajectories with high complexity.

It is no surprise that four measures of complexity of the symbolic sequences that we are using demonstrate high correlation (see Table 1).

	Sh	H^1	H^l_{Max}	K
Sh	1.	0.857774	0.997759	0.873823
H^1	0.857774	1.	0.870351	0.970332
H^l_{Max}	0.997759	0.870351	1.	0.886194
K	0.873823	0.970332	0.886194	1.

TABLE 1. Correlation of the complexity measures.

Shannon entropy (Sh) and Markov Max entropy (H^l_{Max}), and Markov H^1 entropy and Kolmogorov complexity (K) have higher correlations. It can be seen also on the histograms (Fig. 2).

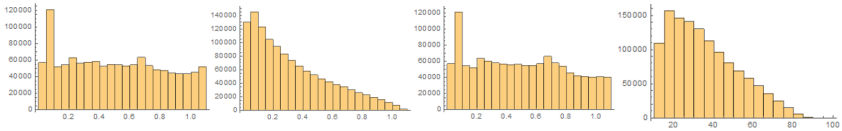


FIGURE 2. Histograms of complexity measures. Left to right: Shannon entropy, Markov H^1 entropy, Markov Max entropy and Kolmogorov complexity.

For each of the four complexity measures we consider, we have selected appr. 5500 orbits with highest values of the corresponding measure. Initial conditions for these selections are shown in Fig. 3.

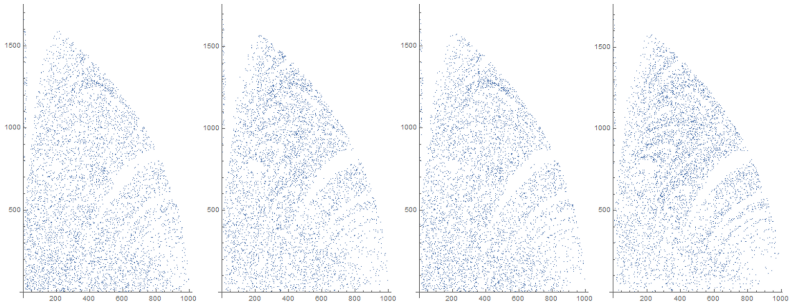


FIGURE 3. Initial conditions of selected orbits with high complexity. Left to right: Shannon entropy, Markov H^1 entropy, Markov Max entropy and Kolmogorov complexity.

Comparing Fig. 3 with life-time of the systems (Fig. 4) one can see that selected orbits avoid regions for short-living systems. Out of 14076 individual orbits

selected this way, there are 416 trajectories that were selected by all four criteria. Initial conditions of these trajectories are shown in Fig. 5.

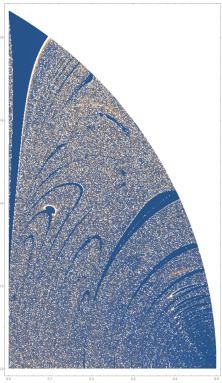


FIGURE 4. Life-time of the systems. Blue corresponds to the short-living systems, yellow - long-living.

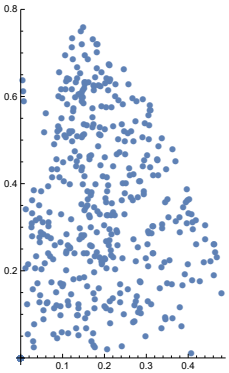


FIGURE 5. Initial conditions of orbits with high complexity selected by all four criteria.

We traced early evolution of some of the orbits that were selected by all four entropy measures. Orbit demonstrate typical behaviour of metastable systems - after some active interactions they approach neighborhood of some periodic (Figure Eight, Shubart, etc.) orbit, spend some time in this neighborhood, then move to the neighborhood of another periodic orbit, and so on [4].

3. Acknowledgements

Authors are thankful to Prof. Seppo Mikkola (Tuorla Observatory, University of Turku) for allowing the use of his code. This work was carried out as part of a project supported by an RFBR grant 17-01-00433.

References

[1] Agekian, T.A. and Anosova, J.P. 1967, Astron. Zh., 44, 1261

[2] Chase, T., Mylläri, A., Mylläri, T. and Vassiliev, N. 2019, in International Conference Polynomial Computer Algebra 2019, St. Petersburg department of Steklov Institute of Mathematics, RAS, pp. 142 - 146

[3] Luo, L. F. 1993, Il Nuovo Cimento vol. 108 B, N. 1, pp. 80 - 92

[4] Martynova A.I., Orlov V.V., Rubinov A.V.,2003, MNRAS, 344, 1091

[5] Mikkola, S. and Tanikawa, K. 1999, Celest. Mech. Dyn. Astron., 74, 287-295.

[6] Mylläri, A., Orlov, V., Chernin, A., Martynova, A. and Mylläri, T. 2016, Baltic Astronomy, vol. 25, 254

[7] Mylläri, A., Mylläri, T., Myullyari, A. and Vassiliev, N. 2018, in International Conference Polynomial Computer Algebra 2018, St. Petersburg department of Steklov Institute of Mathematics, RAS, pp. 82 - 84

[8] Mylläri, A., Mylläri, T., Myullyari, A. and Vassiliev, N. 2019, in International Conference Polynomial Computer Algebra 2019, St. Petersburg department of Steklov Institute of Mathematics, RAS, pp. 103 - 111

Mylläri Aleksandr
Dept. of Computers & Technology, SAS
St.George's University
St.George's, Grenada, West Indies
e-mail: amyllari@sgu.edu

Mylläri Tatiana
Dept. of Computers & Technology, SAS
St.George's University
St.George's, Grenada, West Indies
e-mail: tmyllari@sgu.edu

Myullyari Anna
QiO, Miami,
Florida, USA
e-mail: anna.myullyari@qio.io

Vassiliev Nikolay
V.A. Steklov Institute of Mathematics
of the Russian Academy of Sciences
St. Petersburg, Russia

Tensor Networks for Quantum Systems

Yuri Palii

Description of quantum many body systems involves tensors of high rank and large dimensions. Approximation of a tensor by a set of lower rank tensors called tensor network (TN) makes calculations feasible [1] . In the last three decades lots of numerical algorithms based on the TN-methods were developed in condensed matter, quantum chemistry, high energy physics. Nowadays lattice gauge theories (LGT) are a general approach to nonlinear phenomena in quantum systems [2]. TN-methods were applied to low dimensional LGT with remarkable success [3]. Quantum entanglement lies in the core of TN-representation of the state vector and Hamiltonian for a system. Moreover every quantum circuits represent a kind of a tensor network [4]. On the other side it seems that quantum computers and simulators are the most suitable devices for the implementation of the TN-algorithms [5].

The use of the notion of Tensor Networks in study of quantum systems we can see on the simplest example of two one-half spins. The state vector $|\Psi\rangle$ of this system has the explicit form

$$|\Psi\rangle = \sum_{s_1,s_2=0}^{1} T_{s_1,s_2}|s_1,s_2\rangle = t_{00}|0,0\rangle + t_{01}|0,1\rangle + t_{10}|1,0\rangle + t_{11}|1,1\rangle \qquad (1)$$

where $|0,0\rangle$ represent a state with both spins down, $|0,1\rangle$ represent a state with the first spin down and the second spin up and so on.

We prepare the singular value decomposition (SVD) of the matrix T_{s_1,s_2}

$$T_{s_1,s_2} = \sum_{a,a'=1}^{\chi} U_{s_1,a}\Lambda_{a,a'}V^*_{a',s_2}, \qquad (2)$$

where the matrices U and V are unitary. The real matrix Λ is diagonal with ordered elements $\lambda_1 > \lambda_2 > \ldots > \lambda_\chi$. The positive integer number χ is called the Schmidt rank of the matrix T and $\lambda_a, a = 1,\ldots,\chi$ are called the Schmidt numbers. The product of three matrices U, V and Λ used for the representation of the matrix T_{s_1,s_2} is an example of the tensor networks. The indices a, a' do not relate to the Hilbert space $\mathcal{H}$ in contrast to the indices s_1, s_2. In the new basis formed by the

vectors

$$|u\rangle_a = \sum_{s_1=0}^{1} U_{s_1 a}|s_1\rangle, \qquad |v\rangle_a = \sum_{s_2=0}^{1} V^*_{a s_2}|s_2\rangle, \tag{3}$$

the state vector $|\Psi\rangle$ can be written as a linear combination

$$|\Psi\rangle = \sum_{a=1}^{\chi} \lambda_a |u\rangle_a |v\rangle_a. \tag{4}$$

From the point of view of quantum information science, the Schmidt decomposition (4) exposes quantum correlations between two spins, s_1 and s_2, or quantum entanglement between them. The so called entanglement entropy

$$S = -2\sum_{a=1}^{\chi} \lambda_a^2 \ln \lambda_a, \qquad \sum_{a=1}^{\chi} \lambda_a^2 = 1, \tag{5}$$

serves as a quantitative measure of the entanglement.

On the language of linear algebra, the Schmidt decomposition (4) of a state is nothing then singular value decomposition (SVD) of a matrix. The Schmidt numbers λ_a form the singular value spectrum and the dimension of the spectrum is the rank of the matrix. The SVD gives the optimal lower rank approximation of a matrix. Namely, with a given matrix M of the rank χ, the task is to find a rank $\tilde{\chi}$ matrix M' that minimizes the distance $\mathcal{D}$ in the matrix space

$$\mathcal{D} = |M - M'| = \sqrt{\sum_{ss'} (M_{ss'} - M'_{ss'})^2}. \tag{6}$$

The optimal solution, for which we use TN-algorithms, is given by the SVD

$$M'_{ss'} = \sum_{a=0}^{\chi'-1} U_{s,a} \Lambda_{a,a'} V^*_{s',a}. \tag{7}$$

The error of the approximation of M by M' is given by

$$\varepsilon = \sqrt{\sum_{a=\chi'}^{\chi-1} \lambda_a}.$$

Using the operation of reshaping of tensors [6] one can implement the truncation procedure for a tensor of the arbitrary rank. This become indispensible when one has deal with tensors of high rank and large ranges of tensor indices.

We generalize the example above to present a sample of TN which is especially effective for quantum many-body systems spreaded in one space dimension (for example, spin chains). Let us consider a system of N spins described by the state vector $|\Psi\rangle$ in the Hilbert space $\mathcal{H}$. We define $|\Psi\rangle$ using the N-th order tensor T

given by its coefficients $T_{s_1,\ldots,s_N}$ in a chosen basis (the product of eigenstates of each individual spin),

$$|\Psi\rangle = \sum_{s_1,\ldots,s_N=1}^{d_1,\ldots,d_N} T_{s_1,\ldots,s_N}|s_1,\ldots,s_N\rangle, \tag{8}$$

where each (physical) index s_i runs from 1 to $d_i, i = 1,\ldots,N$. The numbers d_i are called the bond dimensions and are equal $2S_i + 1$ where S_i is the spin of the i-th component of the system. If we neglect completely the entanglement property in the system then the tensor T can be presented as a direct product of the small rank tensors called Matrix Product State (MPS):

$$T_{s_1,\ldots,s_N} = \sum_{a_1,\ldots,a_N} A^{[1]}_{s_1,a_1} A^{[2]}_{s_2,a_1a_2} \ldots A^{[N-1]}_{s_{N-1},a_{N-2}a_{N-1}} A^{[N]}_{s_N,a_{N-1}a_N} \tag{9}$$

with open boundary conditions (OBC), or with periodic boundary conditions (PBC):

$$T_{s_1,\ldots,s_N} = \sum_{a_1,\ldots,a_N} A^{[1]}_{s_1,a_Na_1} A^{[2]}_{s_2,a_1a_2} \ldots A^{[N-1]}_{s_{N-1},a_{N-2}a_{N-1}} A^{[N]}_{s_N,a_{N-1}a_N} \tag{10}$$

The indices a_i are called the geometrical or virtual indices. A graphical representations of these MPS is given in the figure 1 taken from [4] (left - OBC, right - PBC).

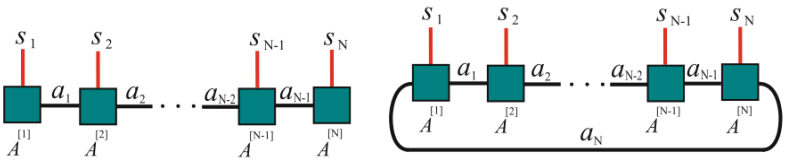


FIGURE 1. Matrix Product State.

References

[1] Simone Montangero. Introduction to Tensor Network Methods Numerical Simulations of Low-Dimensional Many-Body Quantum Systems. Springer Nature Switzerland AG 2018.

[2] Heinz J. Rothe. Lattice gauge theories. 4-th ed. World Scientific 2012.

[3] Mari Carmen Bañuls, Krzysztof Cichy, J. Ignacio Cirac, Karl Jansen, Stefan Kühn. Tensor Networks and their use for Lattice Gauge Theories. Proceedings of Science (LATTICE2018) 022.

[4] Shi-Ju Ran, Emanuele Tirrito, Cheng Peng, Xi Chen, Luca Tagliacozzo, Gang Su, Maciej Lewenstein. Tensor Network Contractions Methods and Applications to Quantum Many-Body Systems. Springer Open, 2020.

[5] M.C. Bañuls, R. Blatt, J. Catani, A. Celi, J.I. Cirac, M. Dalmonte, L. Fallani, K. Jansen, M. Lewenstein, S. Montangero, C.A. Muschik, B. Reznik, E. Rico, L. Tagliacozzo, K. Van Acoleyen, F. Verstraete, U.-J. Wiese, M. Wingate, J. Zakrzewski, and P. Zoller. Simulating Lattice Gauge Theories within Quantum Technologies, arXiv:1911.00003v1 [quant-ph].

[6] Andrzej Cichocki, Namgil Lee, Ivan Oseledets, Anh-Huy Phan, Qibin Zhao, Danilo P. Mandic. Tensor Networks for Dimensionality Reduction and Large-Scale Optimization. Part 1 Low-Rank Tensor Decompositions. Foundations and Trends® in Machine Learning, Vol. 9, No. 4-5 (2016) 249–429.

Yuri Palii
Laboratory of Quantum Photonics, Institute of Applied Physics,
str. Academiei 5, Chisinau, MD-2028 Republic of Moldova
and
Laboratory of Information Technologies,
Joint Institute for Nuclear Research,
str. Joliot-Curie 6, Dubna, 141980 Russia
e-mail: palii@jinr.ru

Fair and envy-free necklace splittings

Gaiane Panina (by a joint work with Duško Jojić and Rade T. Živaljević.)

Assume that r thieves have stolen a necklace and wish to divide it by cutting into pieces and distributing the pieces fairly.

There exist two natural ways to formalize what a "fair division" is:

(1) Assume that the necklace (= a segment) carries n probability measures describing the distribution of n kinds of precious gemstones. Each thief should be treated fairly and receive an equal value of the necklace, as evaluated by each of the measures.

The Splitting Necklace Theorem of Noga Alon states that $n(r-1)$ cuts is sufficient for a fair partition. This is one of the best known early results of topological combinatorics where the methods of equivariant algebraic topology were applied with great success.

(2) The first approach does not take into account personalities of the thieves. Assume now that there are no measures, but for each partition of the necklace each of the thieves prefers one or several of the pieces. Different thieves may prefer different pieces. We wish to divide the necklace such that each thief can have his(her) most preferred piece. That is, no person should be envious of another's share.

David Gale's theorem states that envy-free division with $r-1$ cuts is always possible.

We shall discuss the combination of these two settings: there are n measures on the necklace that should be divided evenly, and also each of the thieves has his own preferences.

The research is partially Supported by the RFBR grant 20-01-00070 "Geometry of metric spaces and its applications to the dynamical systems theory and topology".

Gaiane Panina (by a joint work with Duško Jojić and Rade T. Živaljević.)
St. Petersburg Department of Steklov Institute of Mathematics

Russia
e-mail: gaiane-panina@rambler.ru

Polynomial coefficients as traces and applications to graph colorings

Alexey Gordeev, Zhiguo Li, Fedor Petrov and Zeling Shao

Abstract. Let $G = C_n \times C_m$ be a toroidal grid (that is, 4-regular graph), where nm is even. We prove that this graph G is 3-choosable. We also prove some more general results about list colorings of direct products. The proofs are algebraic, the starting point is Alon–Tarsi application of Combinatorial Nullstellensatz, and the main difficulty is to prove that the corresponding coefficient of the graph polynomial is non-zero.

Let $\mathbb{F}$ be a field, $\mathbf{x} = (x_1, \ldots, x_n)$ a set of variables. For $A \subset \mathbb{F}$ and $a \in A$ denote

$$D(A, a) := \prod_{b \in A \setminus a} (a - b).$$

For a multi-index $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{\geqslant 0}$ denote $|\mathbf{d}| = d_1 + \ldots + d_n$, $\mathbf{x}^{\mathbf{d}} = \prod_{i=1}^n x_i^{d_i}$. For a polynomial $f \in \mathbb{F}[\mathbf{x}]$ denote by $[\mathbf{x}^{\mathbf{d}}]f$ the coefficient of monomial $\mathbf{x}^{\mathbf{d}}$ in polynomial f.

Choose arbitrary subsets $A_i \subset \mathbb{F}$, $|A_i| = d_i + 1$ for $i = 1, \ldots, n$. Denote $A = A_1 \times A_2 \times \ldots \times A_n$.

Recall the formula version of Combinatorial Nullstellensatz (it appeared in this form in quite recent papers [6, 8, 11], but essentially already in [5], see [7] for a modern exposition of the algebraic geometry behind this formula):

$$\left[\mathbf{x}^{\mathbf{d}}\right] f = \sum_{\mathbf{a}=(a_1,\ldots,a_n) \in A} \frac{f(\mathbf{a})}{\prod_{i=1}^n D(A_i, a_i)} \tag{1}$$

for any polynomial $f \in \mathbb{F}[\mathbf{x}]$ such that $\deg f \leqslant |\mathbf{d}|$.

In particular, if $[\mathbf{x}^{\mathbf{d}}]f \neq 0$, then (1) yields the existence of $\mathbf{a} \in A$ for which $f(\mathbf{a}) \neq 0$. This is Combinatorial Nullstellensatz [1], which has numerous applications.

Alon and Tarsi [2] suggested to use it for list graph colorings. Namely, if $G = (V, E)$ is a non-directed graph with the vertex set $V = \{v_1, \ldots, v_n\}$ and the

edge set E, we define its graph polynomial in n variables $x_1, \dots, x_n$ as

$$F_G(\mathbf{x}) = \prod_{(i,j)\in E} (x_j - x_i).$$

Here each edge corresponds to one linear factor $x_j - x_i$, so the whole F_G is defined up to a sign. Assume that each vertex v_i has a list A_i consisting of $d_i + 1$ colors, which are real numbers. A *proper list coloring of G subordinate to lists $\{A_i\}_{1\leqslant i\leqslant n}$* is a choice of colors $\mathbf{a} = (a_1, \dots, a_n) \in A_1 \times \ldots \times A_n = A$ for which neighbouring vertices have different colors: $a_i \neq a_j$ whenever $(i, j) \in E$. In other words, a proper list coloring is a choice of $\mathbf{a} \in A$ for which $F_G(\mathbf{a}) \neq 0$. If $|\mathbf{d}| = |E|$, the existence of a proper list coloring follows from $[\mathbf{x}^{\mathbf{d}}]F_G \neq 0$.

Define the *chromatic number* $\chi(G)$ of the graph G as the minimal m such that there exists a proper list coloring of G subordinate to equal lists of size m: $A_i = \{1, \dots, m\}$. Define the *list chromatic number* $\mathrm{ch}(G)$ of the graph G as the minimal m such that for arbitrary lists A_i, $|A_i| \geqslant m$, there exists a proper list coloring of G subordinate to these lists. Define the *Alon–Tarsi number* $AT(G)$ of the graph G as the minimal k for which there exists a monomial $\mathbf{x}^{\mathbf{d}}$ such that $\max(d_1, \dots, d_n) = k - 1$ and $[\mathbf{x}^{\mathbf{d}}]F_G \neq 0$.

From above we see that the list chromatic number does not exceed the Alon–Tarsi number:

$$\mathrm{ch}(G) \leqslant \mathrm{AT}(G). \tag{2}$$

Further we consider the Alon–Tarsi numbers for the graphs which are direct products $G_1 \square G_2$ of simpler graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$. Recall that the vertex set of $G_1 \square G_2$ is $V_1 \times V_2$ and two pairs (v_1, v_2) and (u_1, u_2) are joined by an edge if and only if either $v_1 = u_1$ and $(v_2, u_2) \in E_2$ or $v_2 = u_2$ and $(v_1, u_1) \in E_1$.

It is well known (Lemma 2.6 in [10]) that $\chi(G_1 \square G_2) = \max(\chi(G_1), \chi(G_2))$. Much less is known about the list chromatic number (and the Alon–Tarsi number) of the Cartesian product of graphs. Borowiecki, Jendrol, Král, and Miškuf [3] gave the following bound:

Theorem 1 ([3]). *For any two graphs G and H,*

$$\mathrm{ch}(G \square H) \leqslant \min(\mathrm{ch}(G) + \mathrm{col}(H), \mathrm{col}(G) + \mathrm{ch}(H)) - 1.$$

Here $\mathrm{col}(G)$ is the *coloring number* of G, i.e. the smallest integer k for which there exists an ordering of vertices $v_1, \dots, v_n$ of G such that each vertex v_i is adjacent to at most $k - 1$ vertices among $v_1, \dots, v_{i-1}$.

Our first result [9] concerns the toroidal grid $C_n \square C_m$ (here C_n is a simple cycle with n edges)

Theorem 2. $\mathrm{AT}(C_n \square C_{2k}) = 3$.

[9] the right hand side of (1) in the necessary case was treated as a trace of the $(2k)$-th power of a certain matrix which for some lucky choice of the sets A_i's appeared to be Hermitian that almost immediately yields the result. This last phenomenon looks bit mysterious for us. We do not know whether it works

for other interesting classes of graphs. The different way to work with these traces was proposed in [4]. It allowed to prove the following rather technical but general result.

Definition. We call a coefficient $\left[\mathbf{x}^{\xi}\right] F_G(\mathbf{x})$ of the graph polynomial F_G *central*, if $\xi_i = \deg_G(v_i)/2$ for all i, and *almost central*, if $|\xi_i - \deg_G(v_i)/2| \leqslant 1$ for all i.

Theorem 3. *Let G be a graph, all vertices in which have even degree. Suppose that the graph polynomial F_G has at least one non-zero almost central coefficient. Then for $H = G \square C_{2k}$ the central coefficient is non-zero. In particular, H is* $(\deg_H/2 + 1)$*-choosable and*

$$\mathrm{ch}(H) \leqslant \mathrm{AT}(H) \leqslant \frac{\Delta(H)}{2} + 1 = \frac{\Delta(G)}{2} + 2.$$

Note that Theorem 1 gives the bound $\mathrm{ch}(H) \leqslant \min(\mathrm{ch}(G) + 2, \mathrm{col}(G) + 1)$ under the same conditions. When $\mathrm{ch}(G)$ (or $\mathrm{col}(G)$) is small, this bound is stronger. But it can also be weaker when $\mathrm{ch}(G)$ and $\mathrm{col}(G)$ are close to $\Delta(G)$. For example, if $G = C_{2l+1}$ is an odd cycle, then F_G obviously has a non-zero almost central coefficient, so, by Theorem 3, $\mathrm{ch}(C_{2l+1} \square C_{2k}) \leqslant 3$ (so this reproves the main result of [9] by a different argument). On the other hand, Theorem 1 gives only $\mathrm{ch}(C_{2l+1} \square C_{2k}) \leqslant 4$.

The talk was partially supported by RFBR grant 19-31-90081.

References

[1] N. Alon, *Combinatorial Nullstellensatz*, Combinatorics, Probability and Computing **8** (1999), no. 1-2, 7–29.

[2] N. Alon and M. Tarsi, *Colorings and orientations of graphs*, Combinatorica **12** (1992), no. 2, 125–134. MR1179249

[3] M. Borowiecki, S. Jendrol, D. Král, and J. Miškuf, *List coloring of Cartesian products of graphs*, Discrete Mathematics **306** (2006), no. 16, 1955–1958.

[4] A. Gordeev and F. Petrov, *Alon – Tarsi numbers of direct products*, arXiv preprint arXiv:2007.07140 (2020).

[5] K. G. Jacobi, *Theoremata nova algebraica circa systema duarum aequationum inter duas variabiles propositarum*, J. Reine Angew. Math. **14** (1835), 281–288.

[6] R. N. Karasev and F. V. Petrov, *Partitions of nonzero elements of a finite field into pairs*, Israel J. Math. **192** (2012), no. 1, 143–156.

[7] E. Kunz and M. Kreuzer, *Traces in strict Frobenius algebras and strict complete intersections*, J. Reine Angew. Math. **381** (1987), 181–204.

[8] M. Lasoń, *A Generalization of Combinatorial Nullstellensatz*, The Electronic Journal of Combinatorics **17** (2010), no. 1.

[9] Z. Li, Z. Shao, F. Petrov, and A. Gordeev, *The Alon–Tarsi Number of A Toroidal Grid*, arXiv preprint arXiv:1912.12466 (2019).

[10] G. Sabidussi, *Graphs with given group and given graph-theoretical properties*, Canadian Journal of Mathematics **9** (1957), 515–525.

[11] U. Schauz, *Algebraically Solvable Problems: Describing Polynomials as Equivalent to Explicit Solutions*, The Electronic Journal of Combinatorics **15** (2008), no. 1.

Alexey Gordeev
Euler International Mathematical Institute
St. Petersburg Department of Steklov Mathematical Institute of Russian Academy of Sciences
St. Petersburg, Russia

Zhiguo Li
School of Science
Hebei University of Technology
Tianjin, China
e-mail: zhiguolee@hebut.edu.cn

Fedor Petrov
Department of Mathematics and Computer Science
St. Petersburg State University
St. Petersburg, Russia
e-mail: f.v.petrov@spbu.ru

Zeling Shao
School of Science
Hebei University of Technology
Tianjin, China
e-mail: zelingshao@163.com

The interaction of algorithms and proofs in the discrete mathematics course for future engineers

Sergei Pozdniakov and Elena Tolkacheva

Abstract. The paper discusses the possibility of basing the introduction of new concepts, the deducing of their properties and the proof of theorems, based on an analysis of the algorithms associated with these concepts and theories. In this case, activity with an object comes first, which is one of the essential component of technical thinking (it can be considered as conceptual-shapely-active thinking in accordance with the work of T. V. Kudryavtsev [1]), which is not sufficiently taken into account in teaching mathematics in technical universities. It is shown that Papert's thesis to base teaching mathematics on a student's personal thinking can be applied to computer use not only at school, but also at a technical university. An example is given of two topics ("Diophantine equations" and "Continuous fractions"), which can be studied as a single section, considering different interpretations of the extended Euclidean algorithm. Based on the theoretical analysis of the given example, it is shown that when setting out the course of mathematics in technical universities it is advisable to focus on the algorithmic representation of the material. This will naturally connect the material with the activities of the programmer and thereby increase the applied character of teaching mathematics. The work was supported by the RFBR grant No. 19-29-14141

Introduction

One of the urgent problems of teaching mathematics in technical universities is the harmonization of methods of teaching mathematics with the goals of training engineers and taking into account changes in the information environment both in the student's educational environment and in the structure of the engineer's professional activity. The most entrenched tradition of building a mathematics course in a technical university is to copy the style of teaching mathematics to future mathematicians. An indicative is how most of lecturers of technical universities

see the role of examples in reading a course of mathematics. Such a lecturer will first give an abstract definition of a concept, then he will prove its formal properties, then he will prove theorems and ONLY AT THE END will give an example linking a new concept with existing ideas, well-known concepts and applications. Thus, instead of giving a tool for work ("like an ax for a carpenter" according to academician Krylov [2]), the teacher builds a magnificent building of mathematics, demonstrating all its small details and admiring the logical beauty of the structure. At the same time, teachers of mathematics who work with future engineers unanimously note that the presentation of material through algorithms for actions with subject objects meets an incomparably greater response from the audience. Critics of this approach to the course of mathematics will first criticize it for the lack of a strictly logical structure and neglect of evidence. Here are a few arguments that justify this approach and show the inconsistency of such comments.

1. Substantiation of concepts

We consider one of the important arguments presented in the articles [3] by Semour Papert about changing the object basis of ideas that are formed in people's brains under the influence of the information environment. He denies the uniqueness of basing the modern mathematical culture of schoolchildren on such traditional objects as numbers and fractions and shows how studying the control algorithms for a turtle and other computer objects allows not only to form concepts using other basic ideas, but also to use them to prove statements. We will try to show that the analysis of simple algorithms can provide no less proofness than traditional sequence of theorems not related to algorithms.

2. About algorithm analysis as proof

This problem is especially interesting from the point of view of the potential ability to base reasoning not so much on formal premises as on algorithm Most theorems of mathematics are formulated in a constructive form, thereby they already give an algorithm, often far from the most effective, but which students can realize using a simple example for protocol or program for general cases. As a rule, constructive proofs are associated with cyclic (recursive or iterative) algorithms. In this case, introducing the concept of an invariant of a cycle, we can write an algorithm as a special form of writing evidence by the method of mathematical induction, and the proof of the correctness of the algorithm will actually be determined by its structure. As an example, we consider the use of the Euclidean algorithm for decomposing an ordinary fraction into a continuous one and then deriving the properties of convergents. This example is interesting in that the topic "Continuous fractions" is usually set out separately and requires a certain lecture time, while the algorithm for generating convergents is only distinguished by signs from the intermediate operations of the extended extended Euclidean algorithm. This allows

not only more compactly presenting the topic, but, most importantly, showing how the same algorithm solves different problems, increasing the degree of connectivity of the material presented. The Euclidean algorithm can be written as $r_{n-1} = r_n \cdot q_{n+1} + r_{n+1}$ or $r_{n+1} = r_{n-1} - r_n \cdot q_{n+1}$ where $r_{-2} = a, r_{-1} = b$, and $d = GCD(a;b) = r_n$ for n such as $r_{n+1} = 0$. It can also be written in the form $r_{n-1}/r_n = q_{n+1} + 1/(r_n/r_{n+1})$, which indicates the possibility of representing the fraction a/b as an ordered set of quotients $[q_0; q_1, ..., q_n]$ which called continuous fraction. The extended Euclidean algorithm finds a particular solution to the equation $a \cdot x + b \cdot y = d$. It can be naturally obtained from the previous algorithm, considering recurrence as a vector formula $R_{n+1} = R_{n-1} - R_n \cdot q_{n+1}$, where $R_n = (x_n; y_n), R_{-2} = (1;0), R_{-1} = (0;1)$. The invariant of the cycle is the condition $a \cdot x_k + b \cdot y_k = r_k$. At the nth step, the $GCD(a;b)$ and its linear representation will be calculated: $a \cdot x_n + b \cdot y_n = r_n = d$. The next step is two numbers x' and y': $a \cdot x' + b \cdot y' = r_{n+1} = 0$, from which we get $a/b = -y'/x'$. Thus, the extended Euclidean algorithm can be considered as an algorithm for "folding" a continuous fraction $[q_0; q_1, ..., q_n]$ - converting it to a regular fraction $-y'/x'$. It is easy to notice and prove the alternation of signs in the sequence $(x_n; y_n)$, that is, in the formula $R_{n+1} = R_{n-1} - R_n \cdot q_{n+1}$, the addition of either positive or negative numbers always occurs. This allows the "folding" algorithm by substituting the subtraction in the extended Euclidean algorithm for addition: $F_{n+1} = F_{n-1} + F_n \cdot q_{n+1}, F_n = (Q_n; P_n), F_{-2} = (1;0), F_{-1} = (0;1)$. The fractions P_n/Q_n are called convergents for the fraction a/b. The extended Euclidean algorithm in terms of convergents will look like this: $P_{n+1} = P_{n-1} + P_n \cdot q_{n+1}, Q_{n+1} = Q_{n-1} + Q_n \cdot q_{n+1}, P_{-2} = 0, P_{-1} = 1, Q_{-2} = 1, Q_{-1} = 0$. Let us prove that all convergents are irreducible. From the equality$a \cdot x_n + b \cdot y_n = r_n = d$, where $d = GCD(a;b)$, it follows that xn and yn are coprime, that is, the fraction P_n/Q_n is irreducible. But then the previous convergent P_{n-1}/Q_{n-1} will be irreducible, since it can be considered as the penultimate one in the decomposition of Pn / Qn into a continuous fraction. Thus, we have shown that the extended Euclidean algorithm can be considered as an algorithm for reducing fractions. As mentioned above, the signs xn and yn alternate, which can be written exactly as $x_n = (-1)^n \cdot Q_n,\ y_n = (-1)^{n+1} \cdot P_n$. We apply the extended Euclidean algorithm to P_{n+1}/Q_{n+1} : since P_n and Q_n are coprime, we obtain the equality $P_{n+1} \cdot x_n + Q_{n+1} \cdot y_n = 1$ or $P_{n+1} \cdot (-1)^n Qn + Q_{n+1} \cdot (-1)^{n+1} P_n = 1$, which is equivalent to $P_{n+1} \cdot Q_n - Q_{n+1} \cdot P_n = (-1)^n$, whence the formula for the difference of neighboring convergents is obtained $P_{n+1}/Q_{n+1} - P_n/Q_n = (-1)^n/(Q_n \cdot Q_{n+1})$. Other properties of convergents are obtained in the usual way from this formula. increases.

3. About the role of examples

Papert's book [3] draws attention to the term "personal thinking". We will interpret it as "relying on those ideas that the learner owns". Why does solving problems (not exercises) have such a positive effect on mathematical development?

Because this is a direct path to initiating the student's own judgments based on his OWN IDEAS. How to make the presentation understandable to everyone? Examples are a good tool for that. Firstly, they connect two different interpretations of a new idea (in this case, formal with a concrete one), and as you know, images based on internal connections are stored in memory. Secondly, they open the way for independent activity. Finally, and most importantly, they are a way of using the mechanism of internalization [4]. The independent "discovery" by the student of the various patterns outlined above can be supported by the structuring of his activities in the process of constructing and analyzing the protocols of the algorithm.

Conclusion

The report shows that when setting out the course of mathematics in technical universities it is advisable to focus on the algorithmic representation of the material. This will naturally connect the material with the activities of the programmer and thereby increase the applied character of the presentation. It is also shown that an analysis of algorithms can become an adequate replacement for the traditions of presenting material in a non-constructive style in the form of a series of theorems.

References

[1] Kudryavtsev T.V. *Psychology of technical thinking: the process and methods of solving technical problems.* - Moscow: Pedagogy, 1975 (rus)

[2] Krylov A.N. *The importance of mathematics for a shipbuilder. "Shipbuilding"* (No. 7 [43], July 1935) and in the "Bulletin of the USSR Academy of Sciences" (No. 7-8, 1938) (rus)

[3] Papert, S. (1996). *An Exploration in the Space of Mathematics Educations.* International Journal of Computers for Mathematical Learning, Vol. 1, No. 1, pp. 95-123, in 1996.

[4] Leontiev, A. N.*The Development of Mind*, a reproduction of the Progress Publishers 1981 edition, plus *"Activity and Consciousness"*, originally published by Progress Publishers, 1977, published by Erythrospress, see Erythrospress.com (1977)

Sergei Pozdniakov
Algorithmic Mathematics Department
Saint-Petersburg Electrotechnic University LETI
Saint-Petersburg, Russia
e-mail: pozdnkov@gmail.com

Elena Tolkacheva
Algorithmic Mathematics Department
Saint-Petersburg Electrotechnic University LETI
Saint-Petersburg, Russia
e-mail: eatolkacheva@etu.ru

On some matrices whose entries are character sums

N. V. Proskurin

Abstract. Given a finite field $\mathbb{F}_q$ of order q, we define some $(q+1)\times(q+1)$-matrices whose entries are classical character sums. That are Kloosterman, twisted Kloosterman, Salie, Birch sums. All the matrices considered are either symmetric unitary or hermitian unitary or real symmetric orthogonal ones.

1. Preliminaries

Given prime p, let $\mathbb{F}_q$ be the finite field with $q=p^l$ elements and with prime subfield $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$. We write $\mathbb{F}_q^\star$ for the multiplicative group of $\mathbb{F}_q$. Let $e_q\colon \mathbb{F}_q\to\mathbb{C}^\star$ be a non-trivial additive character. By multiplicative characters of $\mathbb{F}_q$ we mean homomorphisms $\chi\colon \mathbb{F}_q^\star\to\mathbb{C}^\star$ extended to $\mathbb{F}_q$ by setting $\chi(0)=0$. To each such character χ one attach the Gauss sum

$$G(\chi)=\sum_{x\in\mathbb{F}_q^\star}\chi(x)e_q(x).$$

For basic of the theory see [1], [2]. Consider the set $\mathbb{F}_q\cup\{o\}$ of $q+1$ elements obtained by adding some point o to the field $\mathbb{F}_q$ (do not mix this o with zero 0). All the matrices we deal with in this paper are $(q+1)\times(q+1)-$ matrices whose rows and columns are numerated by elements of the set $\mathbb{F}_q\cup\{o\}$. We write det for the determinant and E for the identity matrix.

2. Kloosterman sums

Given $c\in\mathbb{F}_q$, let

$$Kl(c)=\sum_{x\in\mathbb{F}_q^\star}e_q(x^{-1}+cx).$$

That are classical Kloosterman sums. Let $Z_q = \big(z_{c,d}\big)$ be the matrix whose entries $z_{c,d}$ with $c, d \in \mathbb{F}_q \cup \{o\}$ are as follows:

$$z_{0,0} = 1 - \frac{1}{q} \quad \text{and} \quad z_{c,d} = \frac{1}{q}\, Kl(cd) \quad \text{for all other pairs } \ c, d \in \mathbb{F}_q,$$

$$z_{o,o} = 0 \quad \text{and} \quad z_{c,o} = z_{o,d} = \frac{1}{\sqrt{q}} \quad \text{for all } \ c, d \in \mathbb{F}_q.$$

Theorem. *The matrix Z_q is a real symmetric orthogonal matrix. The trace of Z_q is equal to the number of square roots of -1 in $\mathbb{F}_q$. One has $Z_q^2 = E$. The eigenvalues of Z_q are just ± 1.*

3. Twisted Kloosterman sums

Given $c \in \mathbb{F}_q$ and a non-trivial multiplicative character χ of $\mathbb{F}_q$, let

$$Kl(c; \chi) = \sum_{x \in \mathbb{F}_q^\star} \chi(x) e_q(x^{-1} + cx).$$

That are twisted Kloosterman sums. Let $V_q = \big(v_{c,d}\big)$ be the matrix whose entries $v_{c,d}$ with $c, d \in \mathbb{F}_q \cup \{o\}$ are as follows:

$$v_{o,o} = \frac{1}{q}\, \overline{G(\chi)}, \qquad v_{c,d} = \frac{1}{q}\, Kl(cd; \chi),$$

$$v_{c,o} = \frac{1}{\sqrt{q}}\, \bar{\chi}(c), \quad \text{and} \quad v_{o,d} = \frac{1}{\sqrt{q}}\, \bar{\chi}(d) \quad \text{for all } \ c, d \in \mathbb{F}_q.$$

Theorem. *The matrix V_q is a symmetric unitary matrix. One has $V_q \overline{V}_q = E$, $|\det V_q| = 1$ and $|t| = 1$ for each eigenvalue t of V_q.*

4. Salie sums

Assuming $q \equiv 1 \pmod 2$, denote by κ a unique quadratic multiplicative character of $\mathbb{F}_q$. The Salie sums

$$Sl(c) = \sum_{x \in \mathbb{F}_q^\star} \kappa(x) e_q(x^{-1} + cx), \quad c \in \mathbb{F}_q,$$

are just the twisted Kloosterman sums $Kl(c; \kappa)$, see [3], [4]. We have the complex symmetric unitary matrix V_q attached to these sums as above. In the meantime, the quadratic case differs significantly from all others and we can change the general definition of V_q to obtain real matrix. The Salie sums can be evaluated explicitly as follows:

$$Sl(0) = G(\kappa),$$

$$Sl(c) = 0, \ \text{ if } \ \kappa(c) = -1, \ \ c \in \mathbb{F}_q^\star,$$

$$Sl(c) = \{e_q(2r) + e_q(-2r)\}\, G(\kappa), \ \text{ if } \ c = r^2, \ \ r \in \mathbb{F}_q^\star.$$

In particular, we see that $Sl(c)/G(\kappa) \in \mathbb{R}$ for all $c \in \mathbb{F}_q$. The quadratic Gauss sum $G(\kappa)$ equals either $\pm\sqrt{q}$ or $\pm i\sqrt{q}$ according to $\kappa(-1)$ equals either 1 or -1. Let us consider the matrix

$$X_q = \frac{1}{\sqrt{q}}\,\big(x_{c,d}\big),$$

where the entries $x_{c,d}$ with $c, d \in \mathbb{F}_q \cup \{o\}$ are as follows:

$$\begin{aligned} &x_{o,o} = 1\,, \quad x_{c,d} = Sl(cd)/G(\kappa), \\ &x_{c,o} = \kappa(c), \quad \text{and} \quad x_{o,d} = \kappa(d) \quad \text{for all } c, d \in \mathbb{F}_q. \end{aligned}$$

Theorem. *The matrix X_q is a real symmetric orthogonal matrix. Its trace is equal to 0. The eigenvalues of X_q are just ± 1. One has $X_q^2 = E$ and $\det X_q = (-1)^{(q+1)/2}$.*

5. The cubic case

Assume $q \equiv 1 \,(\mathrm{mod}\ 3)$ and denote by ψ someone of two cubic multiplicative characters of $\mathbb{F}_q$. It was shown by H. Iwaniec and W. Duke [5] that the cubic Kloosterman sum can be represented by the Birch sum. It is convenient to write their result as in [7]. That is

$$Kl(h;\psi) = \psi(h)\,C(h) \quad \text{with} \quad C(h) = \sum_{z \in \mathbb{F}_q} e_q\Big(\frac{z^3}{h} - 3z\Big) \in \mathbb{R}, \quad h \in \mathbb{F}_q^\star.$$

Let $U_q = \big(u_{c,d}\big)$ be the matrix whose entries $u_{c,d}$ with $c, d \in \mathbb{F}_q \cup \{o\}$ are as follows:

$$\begin{aligned} u_{o,o} &= u_{0,0} = 0, \\ u_{0,o} &= \frac{1}{q}\,G(\psi), \quad u_{o,0} = \frac{1}{q}\,\overline{G(\psi)}, \\ u_{c,d} &= \frac{1}{q}\,C(cd) \quad \text{for all } c, d \in \mathbb{F}_q^\star, \\ u_{o,n} &= u_{n,0} = \frac{1}{\sqrt{q}}\,\psi(n), \\ u_{n,o} &= u_{0,n} = \frac{1}{\sqrt{q}}\,\bar{\psi}(n) \quad \text{for } n \in \mathbb{F}_q^\star. \end{aligned}$$

Theorem. *The matrix U_q is an hermitian and unitary matrix. Its trace is equal to the number of square roots of 3 in $\mathbb{F}_q$. One has $U_q^2 = E$. The eigenvalues of U_q are just ± 1.*

6. Commentary

Given multiplicative characters μ, ν of $\mathbb{F}_q$ and $a, b \in \mathbb{F}_q^\star$, let $\rho = \mu\bar{\nu}$ and $c = a - b$. One has the formula

$$\frac{1}{q} \sum_{n \in \mathbb{F}_q^\star} Kl(an; \mu)\, \overline{Kl(bn; \nu)}\, \rho(n) + \frac{1}{q}\, G(a; \mu)\, \overline{G(b; \nu)} = G(c; \rho)$$

proved by the author [6]. Its special case with trivial characters μ, ν is given in [8]. Here G denotes the Gauss sum,

$$G(m; \chi) = \sum_{x \in \mathbb{F}_q^\star} \chi(x) e_q(mx)$$

for all multiplicative characters χ of $\mathbb{F}_q$ and all $m \in \mathbb{F}_q$. With this formula, we evaluate inner products of the rows of our matrices. We choose the (c, o) and (o, d) entries to reach orthogonality of the rows.

Just by the definitions of the matrices above, we have character sums expressions for their traces. Say, for the cubic case, we have

$$\mathrm{trace}(U_q) = \frac{1}{q} \sum_{m \in \mathbb{F}_q^\star} C(m^2)$$

and the computation is as follows. The sum over m equals

$$\sum_{m \in \mathbb{F}_q^\star} \sum_{z \in \mathbb{F}_q} e_q\Big(\frac{z^3}{m^2} - 3z\Big).$$

Contribution of the terms with $z = 0$ equals $q - 1$, so that

$$\mathrm{trace}(U_q) = 1 - \frac{1}{q} + \frac{1}{q}\, R \quad \text{with} \quad R = \sum_{z \in \mathbb{F}_q^\star} e_q(-3z) \sum_{m \in \mathbb{F}_q^\star} e_q\Big(\frac{z^3}{m^2}\Big).$$

We take $m = z/h$ with $h \in \mathbb{F}_q^\star$ and continue the computation as follows.

$$R = \sum_{z \in \mathbb{F}_q^\star} e_q(-3z) \sum_{h \in \mathbb{F}_q^\star} e_q(zh^2) = \sum_{h \in \mathbb{F}_q^\star} \sum_{z \in \mathbb{F}_q^\star} e_q\big(z(h^2 - 3)\big).$$

If $h^2 = 3$, then the sum over z equals $q - 1$. Otherwise, it equals -1. It follows,

$$\mathrm{trace}(U_q) = \sharp\,\{h \in \mathbb{F}_q^\star \mid h^2 = 3\},$$

as desired. We find further that the trace equals 1 if $\mathbb{F}_q$ is a field of characteristic 2, and it equals either 2 or 0 for all other characteristics.
The computation of traces of other matrices is similar but a little bit more involved.

Say, for the matrix X_q, the computation leads to

$$\mathrm{trace}(X_q) = 1 + \sum_{y \in \mathbb{F}_q} \kappa(y^2 + 1),$$

and then we find (see [2], ch. 1, §2) that the right-hand side equals 0. A similar formula, involving χ instead of κ, takes place for $\mathrm{trace}(V_q)$.

References

[1] K. Ireland, M. Rosen, *A classical introduction to modern number theory*, Second eq. Grad. Texts in Math., 84. Springer-Verlag, 1990.

[2] S. A. Stepanov, *Arithmetic of algebraic curves*, Moscow, 1991 (in Russian). English translation: Springer-Verlag, 1995.

[3] H. Salié, *Über die Kloostermanschen Summen $S(u, v; q)$*, Math. Z. **34** (1931), 91–109.

[4] K. S. Williams, *Note on Salié's sum*, Proc. AMS **30**, No. 2 (1971), 393–394.

[5] W. Duke, H. Iwaniec, *A relation between cubic exponential and Kloosterman sums*, Contemporary Mathematics **143** (1993), 255–258.

[6] N. V. Proskurin, *On twisted Kloosterman sums*, Zap. Nauchn. semin. POMI **302**, 96-106, 2003 (in Russian). English translation: Journal of Mathematical Sciences **129** (3), 3868-3873, 2005.

[7] N. V. Proskurin, *On cubic exponential sums and Gauss sums* Zap. Nauchn. semin. POMI **458**, 159-163, 2017 (in Russian). English translation: Journal of Mathematical Sciences **234**, 697–700, 2018.

[8] D. H. Lehmer, E. Lehmer, *The cyclotomy of Kloosterman sums*, Acta Arithm. **12**, No. 4 (1967), 385–407.

N. V. Proskurin
St. Petersburg Department of Steklov Institute of Mathematics RAS
191023, Fontanka 27, St. Petersburg, Russia

Some problems on character sums in finite fields

N. V. Proskurin

Abstract. For polynomial character sums in finite fields, it is constructed some analogue to the known conjecture on uniform distribution of the Kloosterman sums values with respect to the Sato-Tate measure.

1. Kloosterman sums and Sato-Tate measure

Given prime p, consider the field $\mathbb{Z}/p\mathbb{Z}$ of order p and some non-trivial additive character $e_p\colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}^\star$. Then

$$Kl_p(c) = \sum_{t\in\mathbb{Z}/p\mathbb{Z}\setminus\{0\}} e_p(t^{-1}+ct) \quad \text{with} \quad c\in\mathbb{Z}$$

are *Kloosterman sums*. According to Weil (1948), one has $|Kl_p(c)| \le 2\sqrt{p}$ and one may look on distribution of the points

$$\frac{Kl_p(c)}{2\sqrt{p}} \quad \text{in the interval} \quad [-1,1]\subset\mathbb{R}.$$

On this interval, one has *Sato-Tate probability measure*

$$[u,v] \;\mapsto\; \frac{2}{\pi}\int\limits_u^v \sqrt{1-t^2}\,dt$$

for all $[u,v]\subset[-1,1]$. It is expected, that

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\,\sharp\Big\{p\le x \;\Big|\; \frac{Kl_p(c)}{2\sqrt{p}}\in[u,v]\Big\} = \frac{2}{\pi}\int\limits_u^v \sqrt{1-t^2}\,dt$$

for all $[u,v]\subset[-1,1]$, $c\in\mathbb{Z}$. Hereafter $\pi(x)$ denotes the number of all prime $p\le x$.

Writing $\vartheta(p,c)$ for a unique number in $[0,\pi]$ with $Kl_p(c) = 2\sqrt{p}\,\cos\vartheta(p,c)$, we get an equivalent conjecture:

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\,\sharp\Big\{p\le x\;\Big|\;\theta(p,c)\in[u,v]\Big\} = \frac{2}{\pi}\int\limits_u^v \sin^2 t\,dt$$

for all $[u,v]\subset[0,\pi]$ and $c\in\mathbb{Z}$. For this problem we refer to J.-P. Serre (Asterisque 41–42, 1977) and to N. M. Katz (Ann. of Math. St. 116, 1988).

2. Elliptic curves

Originally, the Sato-Tate measure is related to elliptic curves. By Hasse theorem, if E_p is an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$, then the number of its points equals

$$p+1+R_p \quad\text{with}\quad R_p \quad\text{satisfying}\quad |R_p|\le 2\sqrt{p}.$$

Now, let E be an elliptic curve defined over $\mathbb{Q}$. For almost all p, its reduction E_p modulo p is an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$. The *Sato-Tate conjecture* states

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\,\sharp\Big\{p\le x\;\Big|\;\frac{R_p}{2\sqrt{p}}\in[u,v]\Big\} = \frac{2}{\pi}\int\limits_u^v \sqrt{1-t^2}\,dt$$

for all $[u,v]\subset[-1,1]$. (For exceptional p, take $R_p=0$.) After R. Taylor (2007), one knows the conjecture holds for non-CM curves with a non-integral j-invariants.

3. Polynomial character sums

Given polynomials a, b over $\mathbb{Z}$ and a multiplicative character χ_p of $\mathbb{Z}/p\mathbb{Z}$ (extended by $\chi_p(0)=0$), let

$$S_p = \sum_{t\in\mathbb{Z}/p\mathbb{Z}} \chi_p(a(t))\,e_p(b(t)). \tag{1}$$

We intend to consider possible analogues of the Sato-Tate type conjectures for the polynomial character sums like (1).

Notice that the Kloosterman sums with $p\ne 2$ can be written as in (1) with a unique quadratic character χ_p of $\mathbb{Z}/p\mathbb{Z}$.

Under some general assumptions, the Artin L-function attached to a, b, χ_p, S_p can be written as the product

$$L(z;a,b) = \prod_{j=1}^{k}(1-\omega_j z) \quad\text{with}\quad \omega_1,\dots,\omega_k\in\mathbb{C}, \quad\text{and}$$

$$S_p = -\sum_{j=1}^{k}\omega_j, \quad\text{where}\quad k = n+m-1,$$

n is the degree of $b \bmod p$, and radical of $a \bmod p$ is a polynomial of degree m.

For the zeros of the Artin L-functions we have $|\omega_j| = \sqrt{p}$ for all j. That is an analoq of the Riemann hypotheses. It has been proved by Weil (1948).
As a consequence, we have the fundamental estimate for the character sums:

$$|S_p| \le (n+m-1)\sqrt{p}. \tag{2}$$

We refer to J.-P. Serre (Asterisque 41–42, 1977) for review of general theory and to S. A. Stepanov for monography on arithmetic of algebraic curves (1991).

4. Choice of characters

To study distribution of the sums (1) in dependence of p and to state something like the Sato-Tate conjecture, we need a natural and fruitful agreement on the choice of characters χ_p and e_p. There is no problem with choice of additive characters e_p. Just take any non-zero $f \in \mathbb{Z}$ and take $e_p(x) = \exp(2\pi i f x/p)$ for all $x \in \mathbb{Z}$. Also, there is no problem with sums S_p attached to trivial and quadratic characters χ_p. That is so, because of uniqueness of such characters for each prime p. The case of higer order characters χ_p is entirely different. To deal with the sums S_p attached to the characters χ_p of order $h \ge 3$ we suggest the following construction.

Fix some 'auxiliary' $l \in \mathbb{Z}$ and $w \in \mathbb{C}$, which is degree h primitive root of 1.

Let Ω_l be the set of all prime $p \equiv 1 \pmod h$ under the condition: there exists a unique order h character χ_p of the field $\mathbb{Z}/p\mathbb{Z}$, such that $\chi_p(l) = w$.

Here $p \equiv 1 \pmod h$ is a necessary and sufficient condition for existence of order h characters χ_p. The uniqueness can be restated as follows: h is coprime with the index of l relative to some generator of the multiplicative group of $\mathbb{Z}/p\mathbb{Z}$. In particular, for prime h, that means l is not h-th degree in $\mathbb{Z}/p\mathbb{Z}$.

Let $\pi_l(x) = \sharp\left\{ p \in \Omega_l \;\middle|\; p \le x \right\}$ for any $x \in \mathbb{R}$. Assume, the sums S_p are majorized as in (2). Let $D \subset \mathbb{C}$ be the circle of radius $n+m-1$ with center at 0.

We suggest the following statement as an adequate analog or generalization of the conjecture above for the Kloosterman sums.

For any (good) measurable set $V \subset D$, we expect

$$\lim_{x\to\infty} \frac{1}{\pi_l(x)} \sharp\left\{ p \in \Omega_l \;\middle|\; p \le x,\ S_p/\sqrt{p} \in V \right\} = \int_V C(z)\,dz,$$

where $C\colon D \to \mathbb{R}$ is a measurable function, depending on parameters a, b, w, l only. This function should be determined. In the case $S_p \in \mathbb{R}$ for all $p \in \Omega_l$, it is reasonable to replace D with the interval $D \cap \mathbb{R}$ and to treat V as intervals $\subset D \cap \mathbb{R}$.

N. V. Proskurin
St. Petersburg Department of Steklov Institute of Mathematics RAS
191023, Fontanka 27, St. Petersburg, Russia
e-mail: np@pdmi.ras.ru

On the Minimal Orbit Intersection Distance between two elliptical orbits

Rosaev A.E.

Abstract. A method for calculating the minimum distance between two almost overlapping elliptical orbits is proposed. There was obtained an expression for the longitude of the minimum distance point as a function of the orbital elements. There were considered applications of the method for determining the age of pairs of asteroids in close-by orbits.

Introduction

Let λ_1, λ_2 – longitude of the asteroids. Denote λ – longitude of the points of intersection of orbits or point of minimal orbits intersection distance Δ (MOID). Obviously, when forming a pair, it must be true: $\lambda_1 = \lambda_2 = \lambda$. If an encounter occurs at another longitude, this cannot be origination of a pair. This is a strong criterion: if we have a difference in λ about 10^{-3} degree, then the distance between asteroids in the inner asteroid belt will be about 6 thousand kilometres. On the other hand, if take place $\lambda_1 = \lambda_2$. but not $\lambda_1 = \lambda$, we have Δ (MOID) large than Hill sphere radius except very close (identical!) orbits.

Thus, we can conclude that at the time of the appearance of the pair is true:

$$\lambda_1(T) \approx \lambda_2(T) = \lambda \tag{1}$$

$$r_1 \approx r_2 = r \tag{2}$$

Here r is a heliocentric distance.

1. Method of MOID calculations

There are few methods of analytic MOID search (Kholshevnikov,Vassiliev,1999), (Gronchi,2005) but they are not practically applicable in the case almost overlapping orbits. The most simple and useful way to determine the longitude of orbits intersection points λ is described in (Rosaev,2019). Previously, described method

of minimal orbit intersection distance (MOID) was applied to near Earth objects (Rosaev,2001). Now some developments are made and new applications are considered. This method allows calculate λ with uncertainties of few degrees. To reduce uncertainties, we may apply numeric calculation in small vicinity λ.

Presented analytic calculation of λ has an additional advantage. If we have theoretic dependences of orbital elements on time, we can build the respect dependence for λ in case nominal orbit. Finally, we can try to solve the inverse problem and obtain some restrictions on initial conditions, satisfying equality $\lambda_1 = \lambda_2 = \lambda$.

The problem of determining the minimum distance between elliptical orbits reduces to solving the system of equations (in rectangular x, y, z coordinates):

$$\Delta^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \tag{3}$$

$$x_j = r_j \left(\cos(u_j)\cos(\Omega_j) - \sin(u_j)\sin(\Omega_j)\cos(i_j)\right) \tag{4}$$

$$y_j = r_j \left(\cos(u_j)\sin(\Omega_j) + \sin(u_j)\cos(\Omega_j)\cos(i_j)\right) \tag{5}$$

$$z_j = r_j \sin(u_j)\sin(i_j) \tag{6}$$

$$u_j = \lambda_j - \Omega_j, \lambda_j = \nu_j + \Omega_j + \omega_j \tag{7}$$

where ω_j, Ω_j , e_j , i_j - are the perihelion argument, the longitude of the ascending node, eccentricity and the inclination of the first and second objects. We assume that conditions (1) and (2) are satisfied. The expression (3) can be rewritten:

$$\Delta_s^2 = r^2 \sum_{i=1}^{3} \left(c_i \cos(\lambda) + b_i \sin(\lambda)\right)^2 \tag{8}$$

$$\begin{aligned} c_1 = {} & \cos^2(\Omega_1) + \sin^2(\Omega_1)\cos(i_1) - \\ & - \cos^2(\Omega_2) + \sin^2(\Omega_2)\cos(i_2) \end{aligned} \tag{9}$$

$$\begin{aligned} b_2 = {} & \sin^2(\Omega_1) + \cos^2(\Omega_1)\cos(i_1) - \\ & - \sin^2(\Omega_2) + \cos^2(\Omega_2)\cos(i_2) \end{aligned} \tag{10}$$

$$\begin{aligned} c_2 = {} & \sin(\Omega_1)\cos(\Omega_1)\left(1 - \cos(i_1)\right) - \\ & - \sin(\Omega_2)\cos(\Omega_2)\left(1 - \cos(i_2)\right) \end{aligned} \tag{11}$$

$$\begin{aligned} b_1 = {} & \sin(\Omega_1)\cos(\Omega_1)\left(1 - \cos(i_1)\right) - \\ & - \sin(\Omega_2)\cos(\Omega_2)\left(1 - \cos(i_2)\right) \end{aligned} \tag{12}$$

$$c_3 = -\sin(\Omega_1)\sin(i_1) + \sin(\Omega_2)\sin(i_2) \tag{13}$$

$$b_3 = \cos(\Omega_1)\sin(i_1) - \cos(\Omega_2)\sin(i_2) \tag{14}$$

The condition for critical points:

$$\frac{\mathrm{d}\Delta^2}{\mathrm{d}\lambda} = 2\sum_{j=1}^{3}\left(c_j\cos\lambda + b_j\sin\lambda\right)\left(b_j\cos\lambda - c_j\sin\lambda\right) = 0 \tag{15}$$

Solution in the assumption negligible eccentricity:

$$\lambda_s = \frac{1}{2}\arctan\frac{2\sum c_i b_i}{\sum\left(c_i^2 - b_i^2\right)} \tag{16}$$

There are eight roots for λ in interval $[-\pi...\pi]$. And there are no more than eight critical points exists the problem. Additional critical points can appear in the case of an exact orbits coincidence ($\Delta = 0$) the satisfaction of the equality $\sum c_i^2 - b_i^2 = 0$, both cases are very rare.

2. Exact intersection conditions

In this paper we continue the development of the method.In actual cases of neighbour orbits only one root is respected minimum distance. This root can be found by derivative of Δ in equation(8).In addition, we add the condition of exact intesection of two elliptic orbits:

$$D = (\beta\gamma)^2 - \left(\gamma^2 - e_2^2\sin^2(\delta\varpi)\right)\left(\beta + e_2^2\sin^2(\delta\varpi)\right) > 0 \tag{17}$$

$$-1 < \cos\upsilon = \frac{-\beta\gamma/2 \pm \sqrt{D}}{\left(\beta^2 + e_2^2\sin^2(\delta\varpi)\right)} < 1 \tag{18}$$

where:

$$\beta = e_2\cos\delta\varpi + (\gamma - 1)\,e_1 \tag{19}$$

$$\gamma = 1 - \frac{a_2(1-e_2^2)}{a_1(1-e_1^2)} \tag{20}$$

Conclusion

A method for calculating the minimum distance between two almost overlapping elliptical orbits is proposed. There was obtained an expression for the longitude of the minimum distance point as a function of the orbital elements. There were considered applications of the method for determining the age of pairs of asteroids in close-by orbits.

References

[1] Gronchi G., An algebraic method to compute the critical points of the distance function between two keplerian orbits, Celestial Mechanics and Dynamical Astronomy, 93,295-329,(2005)

[2] Kholshevnikov, K.V. and Vassiliev, N.N., On the distance function between two Keplerian elliptic orbits, Celest. Mech. Dyn. Astr., 75(2), (1999)

[3] Rosaev A.E, On the relationship between asteroids, fireballs and meteorites, Proceeding of the Meteoroids 2001 Conference. Swedish Institute of Space Physics, Kiruna, Sweden, 6-10 August 2001, ESA SP-495,477-481 (2001)

[4] Rosaev A.E, Application minimal orbit intersection distance calculation to studying young asteroid pairs., Astrophys Space Sci V.364 p.209 (2019)

Rosaev A.E.
Cultural and Educational Centre named after Valentina Tereshkova
Yaroslavl, Russia
e-mail: hegem@mail.ru

An Effectively Computable Projective Invariant

Alexandr V. Seliverstov

Abstract. We consider a projective invariant of hypersurfaces over a field of characteristic zero. The invariant can be computed in polynomial time with generalized register machines. It has been computed for certain low-dimensional hypersurfaces. One can effectively recognize some plane cubic curves as well as some cubic surfaces. Our method allows to recognize some cubic hypersurfaces with reducible Hessian.

The aim of this work is to introduce an effectively computable projective invariant of hypersurfaces over the field of complex numbers. The simplest case of hypersurface is a plane curve. Every plane cubic curve is projectively equivalent to a curve whose affine part is given by a Weierstrass equation $y^2 = x^3 + px + q$. This curve is singular iff the discriminant of the right-hand univariate polynomial vanishes, that is, $-4p^3 - 27q^2 = 0$. Classification of cubic surfaces is more complicated. A cubic surface is cyclic when there exists a Galois cover of degree 3 over projective plane. A cyclic cubic surface is projectively equivalent to a surface defined by a form of the type $x_0^3 + x_1^3 + x_2^3 + x_3^3 + px_1x_2x_3$, where p is a parameter [1]. The general cubic surface depends on four parameters. It can be defined by the Emch normal form [2]. But this normal form has been found earlier [3, 4].

Let us consider generalized register machines over a field of characteristic zero $(\mathbb{K}, 0, 1, +, -, \times)$. Each register contains an element of $\mathbb{K}$. There exist index registers containing nonnegative integers. The running time is said polynomial, when the total number of operations performed before the machine halts is bounded by a polynomial in the number of registers occupied by the input. Initially, this number is placed in the zeroth index register [5]. If a polynomial serves as an input, then its coefficients are written into registers.

For $n \geq 2$, let us consider a square-free form $f(x_0, \ldots, x_n)$ of degree $d \geq 2$. Let us fix a point U with homogeneous coordinates $(u_0 : \cdots : u_n)$. Every straight line passing through the point U consists of points with homogeneous coordinates $((x_0 - u_0)t + u_0 s : \cdots : (x_n - u_n)t + u_n s)$, where $(s : t)$ are homogeneous coordinates inside the line. The restriction of the form f is a binary form denoted by $r(s, t)$. Let us denote by $D[f, U]$ the discriminant of the binary form $r(s, t)$. If $x_0 = 1$, then the discriminant is a inhomogeneous polynomial in affine coordinates x_k. In the

general case, its degree is equal to $d^2 - d$. If a straight line either is tangent to the hypersurface or passes through a singular point, then the discriminant of the form $r(t, s)$ vanishes. So, if the point U is not any singular point of the hypersurface, then the polynomial $D[f, U](x_1, \ldots, x_n)$ defines a cone with U as a vertex. If U is singular, then $D[f, U]$ vanishes identically.

The set of polynomials of the type $D[f, U]$ for all points U generates a linear subspace W_f of the ambient linear space of all inhomogeneous polynomials of degree $d^2 - d$ in n variables. The dimension of the ambient linear space is equal to

$$w(n, d) = \frac{(n + d^2 - d)!}{n!(d^2 - d)!}.$$

For every irreducible form f, the dimension $\dim W_f$ is projectively invariant. If $d \geq 3$ and n is sufficiently large, then $\dim W_f < w(n, d)$, that is, W_f is a proper subspace of the ambient linear space. If the rank of a quadratic form f is equal to n, then the equality $\dim W_f = w(n, 2)$ holds. For given n and d, the dimension $\dim W_f$ considered as a function of coefficients of f is lower semi-continuous [6]. Thus, if there exists a form $f(x_0, \ldots, x_n)$ of degree d such that $\dim W_f = w(n, d)$, then for almost every form $f(x_0, \ldots, x_n)$ of degree d, the equality $\dim W_f = w(n, d)$ holds too.

Let be given a square-free polynomial $f(x_1, \ldots, x_n)$. In accordance with [6], in the expansion of the polynomial $D[f, U]$ in powers of coordinates of the point U, each coefficient belongs to the linear subspace W_f. These polynomials in variables $x_1, \ldots, x_n$ span whole linear subspace W_f. Thus, there exists a polynomial time algorithm to compute the dimension of the linear subspace W_f.

It is sufficient to calculate the rank of a matrix whose order equals $w(n, d)$. It requires $O(w^\omega)$ multiplications, where ω denotes the matrix multiplication exponent [7, 8]. In small dimensions, the rank can be calculated with computer algebra system software.

We have computed $\dim W_f$ for certain plane curves ($n = 2$). In this case, the linear subspace W_f can be improper. But it is small for the Fermat type curves, where $F_2 = x_0^d + x_1^d + x_2^d$.

d	2	3	4	5	6	7	8	9	10
$w(2, d)$	6	28	91	231	496	946	1653	2701	4186
$\dim W_{F_2}$	6	26	82	207	446	856	1506	2477	3862

If $f(x_0, x_1, x_2)$ defines a singular curve, then the strict inequality $\dim W_f < w(2, d)$ holds. For almost every f of degree $d \leq 7$, the equality $\dim W_f = w(2, d)$ holds. For all $d \leq 7$, the equality $\max\limits_{f(x_0, x_1, x_2)} \dim W_f = w(2, d)$ holds. In particular, the equality holds at forms of the type $f = x_0^d + x_1^d + x_2^d + (x_0 + x_1 + x_2)^d$. We guess that it holds for every larger degree too.

Let us consider cubic forms of the Fermat type $F_n = x_0^3 + \cdots + x_n^3$. The polynomial $D[F_n, U](x_1, \ldots, x_n)$ is equal to the discriminant of a binary form of the type $at^3 + bt^2 s + pts^2 + qs^3$ whose coefficients are sums of univariate polynomials,

that is, $a = a_1(x_1)+\cdots+a_n(x_n)$, $b = b_1(x_1)+\cdots+b_n(x_n)$, $p = p_0+p_1x_1+\cdots+p_nx_n$, and q is a constant. So, every monomial of $D[F_n, U]$ depends on at most four variables. Thus, $\dim W_{F_n} = O(n^4)$.

For $n \le 9$, the equation $\dim W_{F_n} = \frac{1}{4}n^4 + \frac{5}{6}n^3 + \frac{9}{4}n^2 + \frac{8}{3}n + 1$ holds.

We have also computed $\dim W_f$ for certain cubic hypersurfaces. For $n \le 3$, this result found by symbolic computations with parameters, where every parameter can be considered as a transcendental number.

For $n \ge 4$, $\dim W_f$ was only computed for certain cubic forms. They provide the lower bound on the maximum value of $\dim W_f$ for given n. For cubic forms $f(x_0, \ldots, x_n)$, we guess that the maximum dimension is

$$\max_f \dim W_f = n + \dim W_{F_n} = \frac{1}{12}(n+1)(3n^3 + 7n^2 + 20n + 12)$$

n	2	3	4	5	6	7	8	9
$w(n,3)$	28	84	210	462	924	1716	3003	5005
$\max_f \dim W_f$	28	75	≥ 169	≥ 336	≥ 608	≥ 1023	≥ 1625	
$\dim W_{F_n}$	26	72	165	331	602	1016	1617	2455

Let us consider cubic curves. In the general case, $\dim W_f = 28$ except the Fermat type curves and all singular curves. We computed the determinant of a matrix composed by coefficients of polynomials generating the linear space W_f. For the Weierstrass normal form $f = x_2^2x_0 + x_1^3 + px_1x_0^2 + qx_0^3$, the determinant is proportionate to the expression $p^4(4p^3+27q^2)^8$. If $p = 0$ and $q \ne 0$, then the curve is projectively equivalent to a curve of the Fermat type. If $4p^3+27q^2 = 0$, then the curve is singular, else it is smooth. For the Fermat cubic curve, $\dim W_{F_2} = 26$. In this case, the Hessian curve is the union of three straight lines. For an irreducible cubic curve with a node, $\dim W_f = 25$. For a cubic curve with a cusp, $\dim W_f = 21$. Therefore, one can distinguish between nodal and cuspidal curves.

For the general cubic surface, $\dim W_f = 75$. For the general cyclic cubic surface, $\dim W_f = 74$. For the Fermat cubic surface, $\dim W_{F_3} = 72$. So, if the Hessian surface contains a plane, then $\dim W_f$ is small. These results found by symbolic computations with parameters, where every parameter can be considered as a transcendental number. For some singular surfaces, the equality $\dim W_f = 75$ holds too. For example, it holds for $f = x_0^3 + px_0^2x_1 + x_1^3 + x_0x_2^2 + (x_0^2 + x_1^2 + x_2^2)x_3$, where p is transcendental; the point $(0 : 0 : 0 : 1)$ is singular. Therefore, the approach does not allow one to decide whether a given cubic surface is smooth.

If $f = x_0^3 + px_0^2x_1 + x_1^3 + x_0x_2^2 + x_1x_2x_3$, where p is transcendental, then $\dim W_f = 73$; the point $(0 : 0 : 0 : 1)$ is singular too. If $f = x_0^3 + px_0^2x_1 + x_1^3 + x_0x_2^2 + x_0^2x_3$, where p is transcendental, then $\dim W_f = 48$; the point $(0 : 0 : 0 : 1)$ is singular too.

Conjecture. *For every cubic form g with reducible Hessian, the inequality holds* $\dim W_g < \max_f \dim W_f$. *Moreover, the more factors exists in Hessian, the more gap is between two values* $\dim W_g$ *and* $\max_f \dim W_f$.

The computational results show that one can easily verify smoothness of almost every plane quartic curve as well as almost every quartic surface in $\mathbb{P}^3$ by

means of computing $\dim W_f$. The method is also applicable to other plane curves. On the other hand, the same problem for cubic surfaces is hard enough because the proposed projective invariant is useless in this case. Nevertheless, one can recognize singularities of some types. We also assume that our method allows to recognize cubic hypersurfaces with reducible Hessian in deterministic polynomial time.

Acknowledgments. The reported study was funded by RFBR according to the research project no. 18-29-13037.

References

[1] Dolgachev I., Duncan A. Automorphisms of cubic surfaces in positive characteristic. *Izvestiya: Mathematics*, 2019, vol. 83, no. 3, pp. 424–500. https://doi.org/10.1070/IM8803

[2] Emch A. On a new normal form of the general cubic surface. *American Journal of Mathematics*, 1931, vol. 53, no. 4, pp. 902–910. https://doi.org/10.2307/2371234

[3] Reznick B. Some new canonical forms for polynomials. *Pacific Journal of Mathematics*, 2013, vol. 266, no. 1, pp. 185–220. https://doi.org/10.2140/pjm.2013.266.185

[4] Wakeford E.K. On canonical forms. *Proceedings of the London Mathematical Society*, 1920, vol. 18, no. 1, pp. 403–410.

[5] Neumann E., Pauly A. A topological view on algebraic computation models. *Journal of Complexity*, 2018, vol. 44, pp. 1–22. https://doi.org/10.1016/j.jco.2017.08.003

[6] Seliverstov A.V. On tangent lines to affine hypersurfaces. *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, 2017, vol. 27, no. 2, pp. 248–256 [in Russian]. https://doi.org/10.20537/vm170208

[7] Schönhage A. Unitäre Transformationen großer Matrizen. *Numerische Mathematik*, 1973, vol. 20, pp. 409–417. https://doi.org/10.1007/BF01402563

[8] Malaschonok G., Gevondov G. Quick triangular orthogonal decomposition of matrices. In: Vasilev N.N. (ed.) *Polynomial Computer Algebra 2019*, St. Petersburg, VVM, 2019. https://www.elibrary.ru/item.asp?id=41320907

Alexandr V. Seliverstov
Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute)
Moscow, Russia
e-mail: slvstv@iitp.ru

On Natural Transformations in Compact Closed Categories with Generating Unit

Sergei Soloviev
IRIT, University of Toulouse-3,
118, route de Narbonne, 31062, Toulouse, France,
soloviev@irit.fr

1 What has been done.

The idea that the arbitrary natural transformations of superpositions of distinguished functors (such as tensor product $\otimes : K \times K \to K$ and internal hom-functor $\multimap : K^{op} \times K \to K$) can be described if tensor unit I is a generating object in the category K was first explored in [5] and studied further in [1].

More precisely, the question is formulated as follows. Let K be some category with additional structure including distinguished functors and natural transformations (distinguished objects may be seen as constant functors). Canonical natural transformations are those obtained from distinguished natural transformations by application of functors and composition.[1] Is it possible to describe arbitrary natural transformations between superpositions of distinguished functors in terms of canonical natural transformations with parameters?

The main results below are obtained in the situation when tensor unit I is a generator. One of typical examples where all these results hold is the category of finitely generated projective modules over a commutative ring I with unit.

1.1 SMC and CC Categories

How we may proceed is illustrated by some of the results of [5].

For example, the structure of a symmetric monoidal closed (SMC) category K contains two distinguished functors $\otimes : K \times K \to K$ and $\multimap : K^{op} \times K \to K$. (Tensor product and internal *hom*-functor in typical cases.)

There are also the distinguished object I (tensor unit), the distinguished natural isomorphisms $a_{XYZ} : (X \otimes Y) \otimes Z \to X \otimes (Y \otimes Z)$, $b_X : X \otimes I \to X$, $c_{XY} : X \otimes Y \to Y \otimes X$, and the (generalized) natural transformations

[1] Composition of natural transformations and (since the category of functors and natural transformations is 2-category) composition with distinguished functors as well. The latter may be seen as substitution of functorial expressions for variables, e.g., if commutativity of tensor product $c_{XY} : X \otimes Y \to Y \otimes X$ is distinguished natural transformation (thus it is canonical) then $c_{(X \otimes Y)Z} : (X \otimes Y) \otimes Z \to Z \otimes (X \otimes Y)$ is canonical as well.

$d_{XY} : X \to Y \multimap X \otimes Y$, $e_{XY} : (X \multimap Y) \otimes X \to Y$. They must satisfy certain equations. Using these natural transformations, also the *adjunctions* $\pi_{XYZ} : Hom(X \otimes Y, Z) \to Hom(X, Y \multimap Z))$ and its inverse π^{-1} may be defined.[2] (See [4], [5] for detailed definitions.)

One may recall that $I \in Ob(K)$ is a generator iff for any $f \neq g : X \to Y \in Mor(K)$ there exists $h : I \to X$ such that $f \circ h \neq g \circ h$. For example, in the category of sets any non-empty set is a generator. In the category of modules over a ring I, I is a generator.

The key technical lemma in [5] (lemma 2.1) stated that if I is a generator in K then an arbitrary generalized natural transformation $f_X : (X \multimap I) \otimes X \to I$ (in the category of functors over K) is the composition

$$(X \multimap I) \otimes X \stackrel{e_{XI}}{\to} I \stackrel{h}{\to} I$$

for some endomorphism $h : I \to I$.

Let F, G be superpositions of distinguished functors represented by formulas in appropriate syntax. Some variables in F, G may be identified; the schema of this identification was called *graph* in [4]. For each occurrence in F, G its variance is defined as usual. The expression $F \to G$ is called the type of a natural transformation $f : F \to G$. For ordinary natural transformations, each variable occurs once in F and once in G (with the same variance). For generalized natural transformations (see, e.g., [4] or [5]) two more cases are admitted: a variable may occur either twice in F or twice in G, with opposite variances (cf. e_{XY} and d_{XY} above). The type $F \to G$ is balanced iff each variable occurs exactly twice (with the same variances when its occurrences lie at the opposite sides of the arrow, and with opposite variances otherwise).

An SMC category K is compact closed (CC) category if there are two more distinguished natural isomorphisms: $s_X : (X \multimap I) \multimap I \to X$ and $t_{XY} : X \multimap Y \to (X \multimap I) \otimes Y$ (their inverses are canonical natural transformations of SMC category, in CC-case they must be isomorphisms).

Let K be compact closed. If the type $F \to G$ is balanced and contains variables $X_1, ..., X_n$ then every generalized natural transformation $g : F \to G$ in K can be obtained from some natural transformation

$$f(X_1 ... X_n) : ((X_1 \multimap I) \otimes X_1) \otimes ... ((X_n \multimap I) \otimes X_n) \to I$$

by composition with canonical natural transformations and applications of adjunctions π and π^{-1}. (For the sake of certainty we assume that $\otimes$, as well as $\oplus$ below, associate to the left.) It may be described as application of some operator Φ depending only on type $F \to G$ to f, i.e., $g = \Phi(f)$. Using lemma 2.1 of [5] we prove that f is equal to the following composition

$$(X_1 \multimap I) \otimes X_1 \otimes ... (X_n \multimap I) \otimes X_n \stackrel{e_{X_1 I} \otimes ... e_{X_n I}}{\longrightarrow} I \otimes ... I \stackrel{b_{I \otimes ... I}}{\longrightarrow} ... \stackrel{b_I}{\to} I \stackrel{f_0}{\to} I$$

and $g(X_1 ... X_n) = \Phi(f_0 \circ b_I \circ ... \circ b_{I \otimes ... I} \circ (e_{X_1 I} \otimes ... e_{X_n I}))$. (Cf. Th. 3.10 of [5].)

[2] It is possible also other way round: if π and π^{-1} are considered as basic, e and d may be derived.

Remark 1.1 *Further modifications of this representation are possible. For* $f : X \to Y$*, the* linear multiplication *by an endomorphism* $h : I \to I$ *(denoted* $h \cdot f$*) is defined as the composition*

$$X \xrightarrow{b_X^{-1}} I \otimes X \xrightarrow{h \otimes f} I \otimes Y \xrightarrow{b_Y} Y.$$

Using adjunctions and isomorphisms in a slightly different way, we may show that $g = \Phi'(f)$ *where* $f : X_1 \otimes ... \otimes X_n \to X_1 \otimes ... \otimes X_n$*. Moreover* $f = h \cdot 1_{X_1 \otimes ... \otimes X_n}$*. If we take* $X_1 = ... = X_n = I$*, there exist unique isomorphisms* $\phi : I \to F(I, ..., I), \psi^{-1} : G(I, ..., I) \to I$*, and* $h = \psi^{-1} \circ g(I, ..., I) \circ \phi$*. (We shall note this composition* $g^*(I)$*.)*

It implies also genericity in another sense. When the type $F \to G$ is balanced the equality of two natural transformations $g_1, g_2 : F \to G$ may be checked on I only: $g_1 = g_2 \iff g_1^*(I) = g_2^*(I)$. (Th. 3.12 of [5].)

In [5] the case of CC categories with biproduct $\oplus$ was studied as well, and some results about arbitrary natural transformations $f : F \to G$ where F, G may contain $\oplus$ were proved. The description, however, was partial. The results were subject to some constraints concerning the structure of F, G. For example these results did not cover the case of tensor powers that were not balanced, as $X \otimes X \to X \otimes X$.

1.2 SM Categories with Biproduct

In a more recent work [1] our aim was to obtain full description of natural transformations for arbitrary superpositions of $\otimes$ and $\oplus$ (without constraints such as balancedness).

That is, Symmetric Monoidal (SM) Categories K with $\otimes$ (tensor) and $\oplus$ (biproduct called also direct sum) as distinguished functors were considered. Respectively, there were 2 distinguished objects, tensor unit I and zero-object 0. Distinguished natural transformations for $\otimes$ were the same as before. Distinguished natural transformations for $\oplus$ included canonical projections (for $\oplus$ as product), injections (for $\oplus$ as sum), natural transformations that characterize 0 as zero-object, diagonal and codiagonal maps $\Delta, \bigtriangledown$. Distributivity isomorphisms for $\otimes$ over $\oplus$ in this setting are derived.

Notice that the sum $f + g$ of morphisms $f, g : X \to X$ can be defined as

$$X \xrightarrow{\Delta} X \oplus X \xrightarrow{f \oplus g} X \oplus X \xrightarrow{\bigtriangledown} X$$

(in fact K is *semi-additive*).

In absence of internal *hom*-functor a slightly modified notion of generation called $tensor - generation$ was necessary. The unit I was called $tensor - generator$ iff given any pair of unequal maps

$$X_1 \otimes ... \otimes X_n \underset{g}{\overset{f}{\rightrightarrows}} X$$

there is a map $k_i : I \to X_i$ such that

$$X_1 \otimes ... \otimes I ... \otimes X_n \xrightarrow{1 \otimes ... k_i \otimes ... 1} X_1 \otimes ... \otimes X_i ... \otimes X_n \underset{g}{\overset{f}{\rightrightarrows}} X$$

are also unequal. That is $1 \otimes ... \otimes k_i ... \otimes 1$ distinguishes f from g. (If internal *hom*-functor $\multimap$ is present as well, I is tensor-generator iff I is generator in ordinary sense because of adjointness of $\otimes$ and $\multimap$.)

Now (proposition 2.3 of [1]) every natural transformation

$$f(X_1, ..., X_n) : X_1 \otimes ... \otimes X_n \to X_{\sigma(1)} \otimes ... \otimes X_{\sigma(n)}$$

may be represented as $f^*(I) \cdot \sigma$ where

$$\sigma : X_1 \otimes ... \otimes X_n \to X_{\sigma(1)} \otimes ... \otimes X_{\sigma(n)}$$

is the canonical natural transformation determined by the permutation σ (i.e. obtained from natural associativity and commutativity of $\otimes$) and $f^*(I)$ is

$$I \xrightarrow{b_I^{-1}} I \otimes I \xrightarrow{b_{I\otimes I}^{-1}} ...I \otimes ... \otimes I \xrightarrow{f(I,...,I)} I \otimes ... \otimes I ... \xrightarrow{b_{I\otimes I}} I \otimes I \xrightarrow{b_I} I.$$

Let F, G be tensor powers (tensor products of variables). The type $F \to G$ was called in [1] multibalanced if the number of occurrences of each variable in F is the same as in G. In K as described above (with biproducts and 0) *each natural transformation $f : F \to G$ where $F \to G$ is not multibalanced is zero.* (See [1], proposition 4.10.)

Let now $F \to G$ be some type. The type $F' \to G'$ is called its *generalization* if $F \to G$ may be obtained from $F' \to G'$ by identification of some variables. Let $F_1 \to G_1, ..., F_k \to G_k$ be balanced generalizations of a multibalanced type $F \to G$, and $\tau_1, ..., \tau_k$ denote the corresponding identifications of variables (substitutions), i.e., $\tau_i(F_i \to G_i) = F \to G$.

Theorem 1.2 *In the conditions described above, with $F \to G$ multibalanced, every natural transformation $f : F \to G$ is the sum $\tau_1(f_1) + ... + \tau_k(f_k)$ where $f_1, ..., f_k$ are some natural transformations of $F_1 \to G_1, ..., F_k \to F_k$ respectively. If $F \to G$ is not multibalanced then f is zero.*

This theorem is a slightly reformulated version of the extraction theorem (theorem 4.2) of [1]. And f_i when I is tensor-generating are of the form

$$F_i \xrightarrow{h_i \cdot \sigma_i} G_i$$

with σ_i canonical natural transformations determined by $F_i \to G_i$ and $h_i : I \to I$ endomorphisms of I. Moreover, $h_i = f_i^*(I)$ (it can be computed using the I-component of f_i).

Example 1.3 *Let $f : X \otimes X \to X \otimes X$. There are two balanced types $X \otimes Y \to Y \otimes X$ and $X \otimes Y \to X \otimes Y$ that produce $X \otimes X \to X \otimes X$ by identification of variables. Then $f = h_1 \cdot c_{XX} + h_2 \cdot 1_{X \otimes X}$. If K is the CC category of finitely dimensional vector spaces, h_1 and h_2 may be seen merely as scalar coefficients.*

Remark 1.4 *Let X^k denote $X \otimes ... \otimes X$ (k times). Up to natural associativity and commutativity each multibalanced type $F \to G$ without $\oplus$ may be seen as $X_1^{k_1} \otimes ... \otimes X_l^{k_l} \to X_1^{k_1} \otimes ... \otimes X_l^{k_l}$. If we "deidentify" the variables of each cluster $X_i^{k_i}$ in a standard way (take at the left $X_{i1} \otimes ... \otimes X_{ik_i}$ and at the right all possible permutations of the same variables) we obtain altogether $n = k_1! \cdot ... \cdot k_l!$ different balanced types that produce $F \to G$ by identification. Then f is equal to the following composition:*

$$F \to X_1^{k_1} \otimes ... \otimes X_l^{k_l} \stackrel{\sum_{i=1}^{n} h_i \cdot \tau(\sigma_i)}{\longrightarrow} X_1^{k_1} \otimes ... \otimes X_l^{k_l} \to G.$$

(Here τ merely identifies back all deidentified variables like in the example above, so it is the same for all i.)

If $\oplus$ occurs in $F \to G$ then by distributivity there exist natural isomorphisms $\phi : F_1 \oplus ... \oplus F_k \to F$ and $\psi^{-1} : G \to G_1 \oplus ... \oplus G_l$ where $F_1, ..., F_k, G_1, ..., G_l$ are tensor products of variables. Thus $f : F \to G$ is $\psi \circ g \circ \phi^{-1}$ for some $g : F_1 \oplus ... \oplus F_k \to G_1 \oplus ... \oplus G_l$. Because $\oplus$ is biproduct, g may be represented by the matrix (g_{ij}) where $g_{ij} = p_j \circ g \circ q_i : F_i \to G_j$ $(1 \leq i \leq k, 1 \leq j \leq l)$, and each natural transformation $g_{ij} : F_i \to G_j$ is described as in theorem 1.2 and remark 1.4.

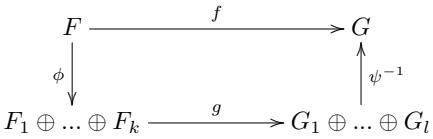
2 New Advancements

R. Houston proved [3] that if a compact closed category has finite products or finite coproducts then it in fact has finite biproducts, and so is semi-additive. This result shows that in fact CC categories with finite biproducts are much more common than we expected when [5] and [1] were written and incites us to consider more closely what can be obtained if we combine the results of our earlier works.

Let us consider first an arbitrary natural transformation $f : F \to G$ where F, G do not contain $\oplus$. If $F \to G$ is not multibalanced then f is zero. If it is, using composition with canonical natural isomorphisms of the CC-structure and adjunctions π, π^{-1}, f may be represented as $\Psi(f_0)$ where $f_0 : X_1^{k_1} \otimes ... \otimes X_i^{k_i} \to X_1^{k_1} \otimes ... \otimes X_i^{k_i}$. The natural transformation f_0 may be described as in section 1.2, i.e., it is either zero or the sum of $f_0^*(I) \cdot \sigma$.

In CC categories with biproduct not only $\otimes$ distributes over $\oplus$ but also there exists canonical isomorphism $X \oplus Y \multimap I \leftrightarrow (X \multimap I) \oplus (Y \multimap I)$. Using canonical isomorphisms, for any natural transformation $f : F \to G$ one obtains

the following commutative diagram



where vertical arrows represent appropriate canonical isomorphisms and $F_1, ..., F_k$, $G_1, ..., G_l$ do not contain $\oplus$.

The natural transformation g may be represented by the matrix (g_{ij}) where $g_{ij} : F_i \to G_j$. In its turn, g_{ij} is either the sum $\sum_{m=1}^{n_{ij}} \tau_{ij}(g_{ijm})$ if $F_i \to G_j$ is multibalanced or zero otherwise, as in section 1.2. The proof uses the extraction technique of [1] based on properties of biproducts.

The $g_{ijm} : F_{im} \to G_{jm}$ are natural transformations of balanced generalizations of $F_i \to G_j$ obtained by deindentification of variables ($1 \leq m \leq n$, where n is given by factorial expression similar to that considered in remark 1.4). Each g_{ijm} may be described as $h_{ijm} \cdot \sigma_{ijm}$ where $h_{ijm} = g^*_{ijm}(I) : I \to I$ and σ_{ijm} is unique canonical natural transformation of the type $F_{im} \to G_{jm}$ (here the results of [5] are used). The operator τ_{ij} identifies back the variables.

References

[1] Cockett, R., Hyland, M. and Soloviev, S. (2001). Natural transformation between tensor powers in the presence of direct sums. Preprint, 01-12-R, IRIT, Université Paul Sabatier, Toulouse, july 2001.

[2] Di Cosmo, R. (1995) *Isomorphisms of types: from lambda-calculus to information retrieval and language design.* Birkhauser.

[3] Houston, R. Finite products are biproducts in a compact closed category. *Journal of Pure and Applied Algebra* 212, 2 (2008), 394 - 400.

[4] Kelly, G.M. and Mac Lane, S. (1971) Coherence in closed categories. *J.of Pure and Applied Algebra*, 1(1), 97-104.

[5] Soloviev, S. (1987) On natural transformatioms of distinguished functors and their superpositions in certain closed categories. *J.of Pure and Applied Algebra*, 47, 181-204

Power geometry in solving system of nonlinear polynomial equations

Akhmadjon Soleev

Abstract. Here we present basic ideas and algorithms of Power Geometry and give a survey of some of its applications. We present a procedure enabling us to distinguish all branches of a space curve near the singular point and to compute parametric of them with any degree of accuracy. Here for a specific example we show how this algorithm works.

Introduction

Many problems in mathematics, physics, biology, economics and other sciences are reduced to nonlinear polynomial equations or to systems of such equations. The solutions of these equations and systems subdivide into regular and singular ones. Near a regular solution the implicit function theorem or its analogs are applicable, which gives a description of all neighboring solutions. Near a singular solution the implicit function theorem is inapplicable, and until recently there had been no general approach to analysis of solutions neighboring the singular one. Although different methods of such analysis were suggested for some special problems.
We offer an algorithm for calculating branches of nonlinear polynomial systems of equations based on Power Geometry [**?**, **?**, **?**]. Here we will consider only to compute local and asymptotic expansions of solutions to nonlinear equations of algebraic classes. As well as to systems of such equations.But it can also be extended to other classes of nonlinear equations for such as differential, functional, integral, integro-differential, and so on [**?**, **?**, **?**].

1. Ideas and algorithms

are common for all classes of equations. Computation of asymptotic expansions of solutions consists of 3 following steps (we describe them for one equation f=0).

1. Isolation of truncated equations $\hat{f}_j^{(d)} = 0$ by means of generalized faces of the convex polyhedron $\Gamma(\mathrm{f})$ which is a generalization of the Newton polyhedron. The first term of the expansion of a solution to the initial equation $f = 0$ is a solution to the corresponding truncated equation $\hat{f}_j^{(d)} = 0$.

2. Finding solutions to a truncated equation $\hat{f}_j^{(d)} = 0$ which is quasi homogenous. Using power and logarithmic transformations of coordinates we can reduce the equation $\hat{f}_j^{(d)} = 0$ to such simple form that can be solved. Among the solutions found we must select appropriate ones which give the first terms of asymptotic expansions.

3. Computation of the tail of the asymptotic expansion. Each term in the expansion is a solution of a linear equation which can be written down and solved.

Elements of plane Power Geometry were proposed by Newton for algebraic equation (1670). Space Power Geometry for a nonlinear autonomous system of ODEs were proposed by Bruno (1962) [**?**].

2. System of algebraic equations [?, ?]

Let an algebraic curve F be defined in C^n by the system of polynomial equations

$$f_i(X) \stackrel{\text{def}}{=} \sum a_{iQ} X^Q = 0, \quad Q = (q_1, \ldots, q_n) \in D_i, \quad i = 1, \ldots, n-1, \tag{1}$$

where $D_i \stackrel{\text{def}}{=} D(f_i) = \{Q : a_{iQ} \neq 0\}$. Let $X = (x_1, \ldots, x_n) = 0$ be a singular point of F, i.e. all $f_i(0) = 0$ and $\operatorname{rank}(\partial f_i / \partial x_j) < n-1$ in $X = 0$. Then several branches of F pass through the $X = 0$. Each branch has its own local uniformization of the form

$$x_i = \sum_{k=1}^{\infty} b_{ik} t^{p_{ik}}, \quad i = 1, \ldots, n \tag{2}$$

where exponents p_{ik} are integers, $0 > p_{ik} > p_{ik+1}$, and coefficients b_{ik} are complex numbers, series converge for large $|t|$, i.e. $X \to 0$ for $t \to \infty$. We propose an algorithm for finding any initial parts of the expansion (2) for all branches of F.

3. Objects and algorithms of Power Geometry

Let us consider the finite sum of monomials

$$f(X) = \sum a_Q X^Q \tag{3}$$

without similar terms and $a_Q \in C$.

The set $D = \{Q : a_Q \neq 0\}$ is called the support of f. We assume that $D \subset \mathbf{Z}^n$, and we enumerate all points of D as $Q_1, \ldots, Q_l$. At first with a help of Newton polyhedrons and its normal cones of polynomials f_i, we find a list of truncated systems [**?**]

$$\widehat{f}_i(X) \stackrel{\text{def}}{=} \sum a_{iQ} X^Q = 0, \quad Q \in D_{ij}^{(d_i)}(f_i), \quad i = 1, \ldots, n-1. \tag{4}$$

Each of them is the first approximation of (1). By the power transformation

$$y_i = x_1^{\alpha_{i1}} \dots x_n^{\alpha_{in}}, \quad i = 1, \dots, n \tag{5}$$

we reduce the number of variables in the truncated system(3). The power transformation (4) resolves (only partly) the singularity $X = 0$ of the system (1). In the transformed system (1), we find all points Y^0 corresponding to the point $X = 0$ of F. We translate each Y^0 into the origin and repeat the procedure described above. After a finite number of such steps, we come to a system having unique local branch and we uniforms it by means of the Implicit Function Theorem. Returning to the initial coordinate X by inverse transformations we receive the branch in the form (2). Analogously we uniforms all other branches of the curve F near the origin $X = 0$ and all branches going to infinity and real branches of a real curve as well.

4. Computation of branches of solutions of the specific system (1) consists of the following 8 stages:

1. For each coordinate singular point X^0 we do parallel-transfer $X - X^0$, write the system in the form (1) and make following stages for each such system separately. We shall describe them for system (1).

2. For each f_i compute the Newton polyhedron M_i, all its faces $\Gamma_{ij}^{(d_i)}$, and normal cones $N_{ij}^{(d_i)}$ and sets $D_{ij}^{(d_i)}$.

3. Find all nonempty intersections $N_{1j}^{(d_1)} \cap \dots \cap N_{n-1j}^{(d_{n-1})}$ with all $d_i > 0$ and, for each of them, write the corresponding truncated system (9).

4. For each such system (9), compute vectors T_i and the matrix α by Theorem 3.1 and make corresponding transformations of (1) and (3) into (17) and (16).

5. Find all roots of (4) and, by computation of the matrix $G = (\partial g_i/\partial y_j)$ separate simple roots Y^0 of (17).

6. By Implicit Function Theorem, compute an initial part of expansions for the branch corresponding to the simple root Y^0 of (17).

7. For each non-simple root Y^0 of (17), compute the new system (19) and repeat the procedure until a full isolation of all branches.

8. By inverse transformations, write all branches in initial coordinates X.

Stages 1–4 were programmed in PC. Stages 5, 6, and 8 are essentially nonlinear but can be done by standard programs.

Here we want to note that the complexity of the truncated system (9) is defined to be the $(n-1)$-dimensional Minkowski mixed volume of the corresponding parallel-transfer faces.

So we got the following result: If we perform calculations for 1-4 using this procedure , then at each step we find all the roots of the corresponding truncated system of equations , and find all the curves of the roots of the truncated system of equations, we obtain a local description of each component in the small neighborhood of the starting point $X = 0$, in the form of expansions (3).

References

[1] A.D. Bruno, *Power Geometry in Algebraic and Differential Equations*, orth-Holland Mathematical Library, N, V.57, Elsevier, 2000.

[2] A.D. Bruno, A.S. Soleev, *Local uniformization of branches of a space curve and Newton polyhedra*. Algebra and Analiz, 1991. Vol. 3, no. 1. P. 67-102.

[3] A.S. Soleev, *Algorithm of local resolution of singularities of space curve*, LNCS 3718, pp.405-415. Springer-Verlag, 2005.

[4] A.S. Soleev, N.A.Soleeva *Power Geometry for Finding Periodic Solutions in One System of ODE*. Malaysian Journal of Mathematical Sciences, No 2, 2014.

[5] A.S. Soleev, *Singling out branches of an algebraic curve and Newton polyhedra*, Dokl. Akad. Nauk SSSR 268 (1983), 1305-1307; (R) = Soviet Math Dokl. 27 (1983) (E).

[6] A.D. Bruno, A.B. Batkhin *Asymptotic solution of an algebraic equation*, DAN 440:3 (2011) 295-300 (R) = Doklady Mathematics 84:2 (2011) 634-639 (E).

Akhmadjon Soleev
Samarkand State University, Uzbekistan
e-mail: `asoleev@yandex.ru`

Forecasting Bitcoin-US Dollar Trend using ANN

K. S. Senthilkumar1 and Naresh Gopal2

Abstract. Bitcoin has gained an amazing popularity and much attention in various research fields. Due to the self-regulation of the availability of the coin, the price of the Bitcoin is mostly dependent on direct factors like lag price, volatility, and volume. This paper aims on predicting the Bitcoin's trend for the next day based on the variables including Bitcoins trading volume, M2 (Money Supply), Price Volatility and Bitcoin lag price using an Artificial Neural Network. Artificial Neural Networks (ANN), one of the data mining technique widely accepted in the business arena due to its ability to learn and detect relationships among nonlinear variables.

Introduction

Bitcoin is a cryptocurrency currently used for Electronic Financial Transactions all over the world, created in 2008 by Satoshi Nakamoto. It is traded by individuals with cryptographic keys and the records are not managed by a bank or agency, but all transactions are recorded in the blockchain and contain records of each transaction that takes place. Bitcoin affects the world economic constancy and has experienced a tremendous volatility over the last few years hence a precise prediction of bitcoin exchange rate with respect to the US dollar has become vital. The Bitcoin value has reached $\$20,000$ on 16th of December 2017 and then it has seen a steep decline at the beginning of 2018. Measurement, prediction, and modelling of BCC price volatility found an important area of research in recent days. Artificial Neural Networks are highly complex composite functions providing the ability of computing non-linear and non-stationary problems like this.

1. Literature

There is dearth of research on the valuation of financial assets and techniques adopted. However, Bitcoin is generally labelled as a virtual currency, which suggests both that it is intangible and innovative. It is therefore necessary to review

the literature on the techniques adopted, its validation and their determinants. Currency exchange rates and stock prices are time series data that is a series of discrete data in time order. Many experimental researches combining technical analysis with computational intelligent techniques have been conducted to improve the modelling and predicting performance for different application in time series data. In the recent past notable amounts of research have been devoted to develop an efficient predictive model using machine learning to assist the traders in making wise investment decisions, including the Simple Moving Average (SMA) for classification tasks [3, 2]. The speculators favour always price fluctuations to profit by predicting which direction prices are headed. Generally, such trading activity has an irrational component translated into prices. The unpredictability of asset prices is a cause for concern because of the adverse effects it has on the traders. To gauge this volatility, GARCH class of models is the most appropriate model to estimate the volatility of the returns of groups of stocks with large number of observations. The analysis of ARCH and GARCH models and their many extensions catered many theories of asset pricing and portfolio analysis [5]. Sean et al. used more complex Neural Network structures such as Recurrent Neural Networks and Long Short-Term Memory Networks. They have shown Long Short-Term Memory Network performed the best achieving a prediction accuracy of 52% [4]. Khuat et al. investigated and compared the effectiveness of Fuzzy logic and ANN, to tackle the financial time series stock forecasting problem. The proposed approaches were tested on the historical price data collected from Yahoo Finance with different companies [1].

2. Proposed Model

It has been proven that a Multi-layer feed forward fully connected network is able to approximate any non-linear function and therefore should be a suitable tool to solve this problem. The input data was normalized in order to avoid over ?tting, improve the prediction performance, and correct scaling since mean square error (MSE) training is used here. The formula used for normalizing the data is given in equation (1).

$$\text{Scalled data} = \frac{x - min(x)}{max(x) - min(x)} \tag{1}$$

The formula above changes the values of all inputs SxS from R to [0,1]. In the forward propagation phase, we walk through the network and compared the received results with those we expected and are essential to prove how good the model is. In the backpropagation phase, we aim to make the network learn from its mistake by starting with the end in mind and working through the system in reverse order. This feedback process gives new, more accurate weights to the initial layer of artificial neurons. Repeating the cycle back and forth number of times (epochs), makes the prediction accurate enough that it can't be improved any further. The network is initialized with randomly chosen weights. The gradient of the error

function is computed and used to correct the initial weights. In a network, using Bias (b) the output of the neuron is shifted by B. Every node in the neural network has a weight $(w_1, \ldots, w_n)$ associated with it. If the input to a neuron is $x_1, \ldots, x_n$, then the output it produces is given in the following equation (2).

$$z = b + \sum_{k=1}^{n} w_k x_k \tag{2}$$

An activation function enables the neuron to convert an input to an output value. Activation Function is also referred to as Transfer Function. Activation functions are an extremely important feature of the artificial neural networks. They basically decide whether a neuron should be activated or not. We used logistic sigmoid function (Range: (0,1); slow convergence) for activation function and is given in the equation (3).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

The most significant part of an ANN is the dataset. Therefore, in this study, the historical daily prices of Bitcoin from July 2010 to Feb 2019 is selected and used for the experiment. The data contains 446 samples. The dependent variable is the daily price of Bitcoin and the independent variables are trading volume, M2 (Money Supply), volatility and Lag price (previous day price). In this study, 350 samples were used for training the model, 96 samples were used for validating and 50 samples were used for testing. After every iteration the cost function of the network is used in the update process. The cost function is application dependent. We used MSE as our cost function in this model.

$$MSE = \frac{\sum_{i=1}^{n} (\text{Actual output-Target output})^2}{n} \tag{4}$$

Where n is the number of samples, actual output is the network output and target output is the known value in the data set. After choosing the weights of the network randomly, the backpropagation algorithm is used to compute the necessary corrections. The algorithm is stopped when the value of the error function becomes sufficiently small.

3. Results and Discussion

We have done some experiments to find the best specification for our model. The graphs depicted in figure 1a, 1b and figure 2a, 2b shows the comparison results for having single hidden layer and two hidden layers. We have chosen two hidden layers with nine neurons. The algorithm was stopped at 25000 epochs and the value of the error function was sufficiently small 0.00039419. For the validation purpose we used 96 unknown samples and for testing purposes we used 50 samples. The actual output from the network and the target output comparison graph for validation is depicted in figure 3a and for testing depicted in figure 3b. Accuracy of the network in the validation phase was 0.86048 and testing phase was 0.94521.

Conclusion

In this work, we proposed a multi-layer Neural Network model for forecasting the Bitcoin price so that the traders could benefit. The computing platform for the entire experiments was done using MATLAB (MATrix LABoratory) program. The accuracy of the ANN is convincing. Therefore, Bitcoin traders could look at the indicators like Volatility, Money Supply, Trading Volume and the Bitcoin lag prices to forecast the Bitcoin price. We would like to extend this work by introducing different activation function with hybrid techniques in the ANN model in future.

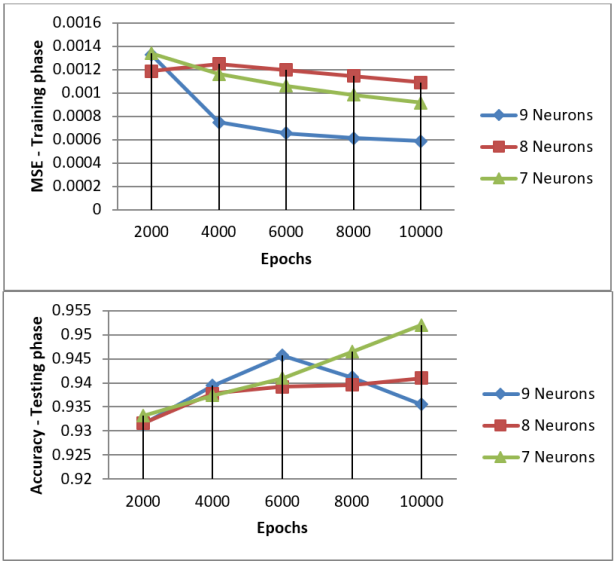


FIGURE 1. Single hidden layer with different number of neurons

References

[1] T. T. Khuat, and M. H. Le, *international Journal on Informatics Visualization*, vol. 1(2), 40-49, 2017.

[2] S. Lauren and S. D. Harlili, Proceeding of the IEEE conference in Advanced Informatics: Concept, Theory and Application (ICAICTA), 135-139, 2014.

[3] S. Sathe, S.M. Purandare, P. D. Pujari, S. D. Sawant, *International Education and Research Journal*, vol. 2(3):74-75, 2016.

[4] Sean McNally, Jason Roche, and Simon Caton, proceedings of the 26th Euro micro International Conference on Parallel, Distributed and Network-based Processing (PDP), 339-343, 2018.

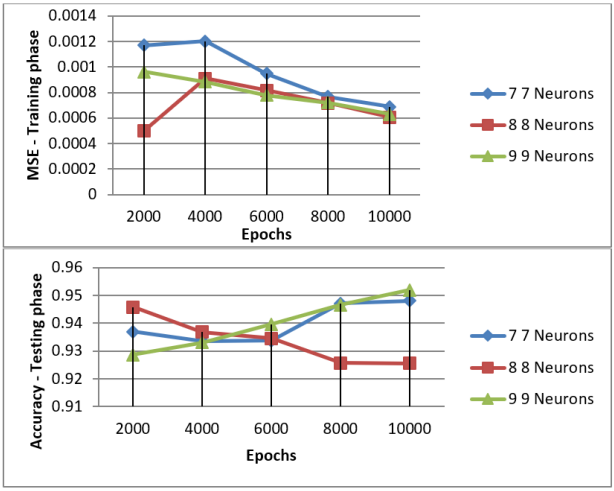
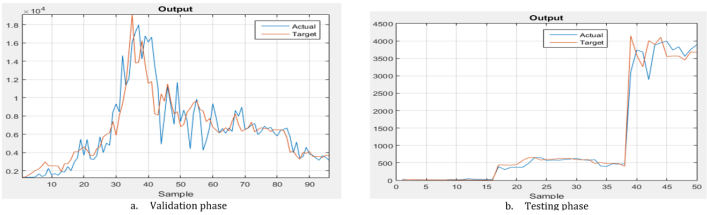


FIGURE 2. Two hidden layers with different number of neurons



a. Validation phase b. Testing phase

FIGURE 3. Comparioson graph of actual and target outputs

[5] S. Thiyagarajan, G. Naresh, and S. Mahalakshmi. *Global Business and Finance Review*, 20, 95-104, 2015.

K. S. Senthilkumar1
Department of Computers and Technology
St. George's University
St. George's, Grenada, W.I
e-mail: ssomasun@sgu.edu

Naresh Gopal2
Indian Institute of Management Ranchi,
Jharkhand, 834008, India
e-mail: kgnaresh@gmail.com

Integer divisibility on $\mathbb{Q}$, quantifier elimination and one Weispfenning's remark

Mikhail R. Starchak

Abstract. In 1999 V. Weispfenning presented a quantifier elimination procedure for the elementary theory of the structure $\langle \mathbb{R}; 0, 1, +, -, [\,], =, <, \{n \mid\}_{n\in\mathbb{N}} \rangle$, where $[\,]$ is the unary integer part operation, and therefore proved decidability of this theory. For the integer divisibility relation $x \mid y \Leftrightarrow \exists z(Int(z) \wedge y = z \cdot x)$ on $\mathbb{R}$, he proved undecidability of the elementary theory of the structure $\langle \mathbb{R}; 0, 1, +, -, [\,], =, <, | \rangle$ and that the theory does not admit quantifier elimination. As a remark, Weispfenning asked whether the positive existential theory of the same structure is decidable.

A decidability proof for this existential theory is the first result of this note. We also sketch a proof of the fact that for every positive existential formula of the first-order language with the signature $\left\langle 0, 1, +, -, \{c\cdot\}_{c\in\mathbb{Q}}, =, \neq, \perp \right\rangle$ there is an equivalent in the rationals $\mathbb{Q}$ quantifier-free formula of the same language. Here $c\cdot$ is a unary functional symbol for multiplication by a rational constant c and $x \perp y \Leftrightarrow Int(x) \wedge Int(y) \wedge GCD(x, y) = 1$.

Introduction

Let L_{PrA} be the first-order language of the signature $\langle 0, 1, +, -, =, <, 2 \mid, 3 \mid, 4 \mid, ... \rangle$. V. Weispfenning [4] considered a natural generalization of Presburger Arithmetic (PrA) and proved that after adjoining the unary integer part operation $[\,]$ to the signature of L_{PrA} (this extended language was named L'), for every positive existential formula we can construct an equivalent in the real numbers $\mathbb{R}$ positive quantifier-free formula [4, Theorem 3.1]. As a corollary, we get decidability of the elementary theory of the structure $\langle \mathbb{R}; 0, 1, +, -, [\,], =, <, 2 \mid, 3 \mid, 4 \mid, ... \rangle$ and also a characterization of the relations, definable in this structure.

If we introduce unary functional symbols $c\cdot$ for multiplication by rational constants c, we get a quantifier elimination procedure for the elementary theory of the structure $\left\langle \mathbb{R}; 0, 1, +, -, [\,], \{c\cdot\}_{c\in\mathbb{Q}}, =, < \right\rangle$. The corresponding language was named L'' and let σ'' be the signature of this language. Then V. Weispfenning

writes: «By way of contrast, quantifier elimination definitely breaks down if one admits scalar multiplication by a real parameter or integer divisibility in the language. In the latter case the elementary theory of real is in fact undecidable». Simultaneously with the integer divisibility $x \mid y \Leftrightarrow \exists z(Int(z) \wedge y = z \cdot x)$ it was also considered the relation $x \parallel y \rightleftharpoons Int(x) \wedge Int(y) \wedge x \mid y$. For the structures $\langle \mathbb{R}; 0, 1, +, -, [\,], =, | \rangle$ and $\langle \mathbb{R}; 0, 1, +, -, [\,], =, \| \rangle$ he proved undecidablity of the elementary theories and decidability of the existential theory of the first structure (it follows from the the Bel'tyukov-Lipshitz theorem [1, 2]). After this proof there is a remark saying that «We do not know whether a corresponding theorem holds in the analogous language L'_{div}», where L'_{div} is the first-order language of the signature $\langle 0, 1, +, -, [\,], =, | \rangle$. We prove that the theory is decidable in section 1.

If we assume that $x \perp y \rightleftharpoons \mathrm{GCD}(x, y) = 1$, then for rational numbers x and y their coprimeness means that these numbers are coprime integers. The elementary theory of the structure $\langle \mathbb{Q}; \sigma \rangle$ admits quantifier elimination (see [4, Corollary 3.5]). Extend σ'' by the coprimeness relation $\perp$ and dis-equality $\neq$; exclude the order relation and the integer part operation. Denote the resulting signature $\sigma_\perp$. In section 2 we sketch the proof of the fact that for every positive existential $L_{\sigma_\perp}$-formula there is an equivalent in $\mathbb{Q}$ quantifier-free $L_{\sigma_\perp}$-formula. Note that $\langle \mathbb{Q}; \sigma_\perp \rangle$ has undecidable elementary theory as a corollary of the undecidability result for the elementary theory of the structure $\langle \mathbb{Z}; 0, 1, +, -, =, \perp \rangle$ proved by D. Richard in [3].

1. One Weispfenning's remark

Theorem 1. *The existential theory of the structure* $\langle \mathbb{R}; 0, 1, +, -, [\,], =, <, | \rangle$ *is decidable.*

Proof. To prove the theorem we reduce it to the decidable positive existential theory of the structure $\langle \mathbb{Q}; 0, 1, +, -, =, <, | \rangle$. Its decidability follows from Bel'tyukov-Lipshitz theorem on decidability of $\exists \mathrm{Th}\langle \mathbb{Z}; 1, +, <, | \rangle$. In the first step of the proof we apply some syntactic transformations of a given formula. For example, using the formula $y = \left[\frac{y}{x}\right] x + \left\{\frac{y}{x}\right\} x$ we can define $x \nmid y$ by a positive existential formula in $\langle \mathbb{R}; 0, 1, +, -, =, <, | \rangle$. Then we have to prove that this formula is true in $\mathbb{R}$ iff it is true in $\mathbb{Q}$.

Let the formula

$$\varphi(\overline{x}) \rightleftharpoons \bigwedge_{i=1..k} g_i(\overline{x}) = 0 \wedge \bigwedge_{i=k+1..l} f_i(\overline{x}) \mid g_i(\overline{x}) \wedge \bigwedge_{i=l+1..m} g_i(\overline{x}) < 0,$$

be satisfiable in $\mathbb{R}$, where $\overline{x}$ is a list of variables $x_1, ..., x_n$; $g_i(\overline{x})$ for $i \in [1..m]$ and $f_j(\overline{x})$ for $j \in [k+1..l]$ are linear polynomials with integer coefficients.

Suppose this formula is true for some real values $\alpha_1, ..., \alpha_n$. Then let for $i = k+1..k'$ we have $g_i(\alpha_1, ..., \alpha_n) = 0$ and $g_j(\alpha_1, ..., \alpha_n) \neq 0$ for every $j \in [k'+1..l]$.

Now define the formula

$$\varphi'(\overline{x}) \rightleftharpoons \bigwedge_{i=1..k'} g_i(\overline{x}) = 0 \wedge \bigwedge_{i=k'+1..l} f_i(\overline{x}) \mid g_i(\overline{x}) \wedge \bigwedge_{i=k'+1..l} \sigma_i \cdot g_i(\overline{x}) < 0 \wedge \bigwedge_{i=l+1..m} g_i(\overline{x}) < 0,$$

where $\sigma_i = 1$ if $g_i(\alpha_1, ..., \alpha_n) < 0$ and $\sigma_i = -1$ if $g_i(\alpha_1, ..., \alpha_n) > 0$ for $i = k'+1..l$.

Consider the system of linear equations with integer coefficients $\bigwedge_{i=1..k'} g_i(\overline{x}) = 0$. Let $A\overline{y} + b$ be a solution set of the system for some rational matrix A, rational vector b and fresh variables $\overline{y} = y_1, ..., y_t$. Substitute $A\overline{y} + b$ for $\overline{x}$ and get an equisatisfiable over the reals system of linear inequalities and divisibilities with rational coefficients

$$\varphi''(\overline{y}) \rightleftharpoons \bigwedge_{i=k'+1..l} \widetilde{f}_i(\overline{y}) \mid \widetilde{g}_i(\overline{y}) \wedge \bigwedge_{i=k'+1..l} \sigma_i \cdot \widetilde{g}_i(\overline{y}) < 0 \wedge \bigwedge_{i=l+1..m} \widetilde{g}_i(\overline{y}) < 0,$$

such that for every rational solution of $\varphi''(\overline{y})$ we can get a rational solution of $\varphi'(\overline{x})$ and thus of $\varphi(\overline{x})$.

Let $\beta_1, ..., \beta_t$ be some real satisfying assignment of $\varphi''(\overline{y})$ Let also the real numbers $\{1, \gamma_1, ..., \gamma_s\}$ for some $s \leq t$ be a basis of the linear space over $\mathbb{Q}$ generated by the reals $\{1, \beta_1, ..., \beta_t\}$. Each element β_i is uniquely represented as $c_{i,0} \cdot 1 + c_{i,1} \cdot \gamma_1 + ... + c_{i,s} \cdot \gamma_s$ for $i = 1..t$, where all $c_{i,j} \in \mathbb{Q}$. Define $\chi_i(z_1, ..., z_s) = c_{i,0} + c_{i,1} z_1 + ... + c_{i,s} z_s$ for $i = 1..t$, substitute $\chi_i(z_1, ..., z_s)$ for y_i in $\varphi''(\overline{y})$ and get a new formula

$$\psi(\overline{z}) = \varphi''(\chi_1(\overline{z}), ..., \chi_t(\overline{z})).$$

Thus for every rational satisfying assignment of the formula $\psi(\overline{z})$ one can get a rational satisfying assignment of $\varphi''(\overline{y})$, and moreover $\psi(\gamma_1, ..., \gamma_s)$ holds.

Rewrite $\psi(\overline{z})$ in the following form:

$$\bigwedge_{i=1..l'} \widetilde{\widetilde{f}}_i(\overline{z}) \mid \widetilde{\widetilde{g}}_i(\overline{z}) \wedge \bigwedge_{i=1..m'} \widetilde{\widetilde{g}}_i(\overline{z}) < 0$$

for some $l' \leq m'$. Consider independently each divisibility $\widetilde{\widetilde{f}}(\overline{z}) \mid \widetilde{\widetilde{g}}(\overline{z})$ in $\psi(\overline{z})$ for $\widetilde{\widetilde{f}}(\overline{z}) = a_0 + a_1 z_1 + ... + a_s z_s$ and non-zero polynomial $\widetilde{\widetilde{g}}(\overline{z}) = b_0 + b_1 z_1 + ... + b_s z_s$. We will show that, actually, $\widetilde{\widetilde{g}}(\overline{z})$ is an integer multiple of $\widetilde{\widetilde{f}}(\overline{z})$ and thus the divisibility holds for every values of $\overline{z}$.

For some integer w we have $w \cdot f(\gamma_1, ..., \gamma_s) = g(\gamma_1, ..., \gamma_s)$. Let $\gamma_0 = 1$, then assuming that $w \cdot a_i \gamma_i \neq b_i \gamma_i$ for some $i \in [0..s]$, we get that $\gamma_i(w \cdot a_i - b_i) = \sum_{j=0..s \wedge j \neq i} \gamma_j (b_j - w \cdot a_j)$. But this is impossible since $1, \gamma_1, ..., \gamma_s$ are linearly independent over $\mathbb{Q}$.

Thus every solution of the subsystem of linear inequalities $\bigwedge_{i=1..m'} \widetilde{\widetilde{g}}_i(\overline{z}) < 0$ with rational coefficients is also a solution of $\psi(\overline{z})$, and since the system is consistent in $\mathbb{R}$, there is some rational solution. $\square$

2. Integer divisibility on $\mathbb{Q}$ and quantifier elimination

Theorem 2. *For every positive existential $L_{\sigma_\perp}$-formula one can construct an equivalent in $\mathbb{Q}$ quantifier-free $L_{\sigma_\perp}$-formula.*

As $\mathrm{GCD}(x,y) = d \Leftrightarrow \frac{x}{d} \perp \frac{y}{d}$, we can consider linear polynomials with rational coefficients in expressions of the form $\mathrm{GCD}(f(\overline{x}), g(\overline{x})) = d$, $f(\overline{x}) = 0$ and $f(\overline{x}) \neq 0$. Elimination of an existential quantifier is based on the following lemma.

Lemma 1. *For the system $\bigwedge\limits_{i\in[1..m]} GCD(a_i, b_i + x) = d_i$ with $a_i, b_i, d_i \in \mathbb{Q}$ and $a_i \neq 0$, $d_i > 0$ for every $i \in [1..m]$, we define for every prime p the integer $M_p = \max\limits_{i\in[1..m]} v_p(d_i)$ and the index sets $J_p = \{i \in [1..m] : v_p(d_i) = M_p\}$ and $I_p = \{i \in J_p : v_p(a_i) > M_p\}$. Then the system has a solution in $\mathbb{Q}$ iff the following conditions simultaneously hold:*

(i) $\bigwedge\limits_{i\in[1..m]} d_i \mid a_i$

(ii) $\bigwedge\limits_{i,j\in[1..m]} GCD(d_i, d_j) \mid b_i - b_j$

(iii) $\bigwedge\limits_{i,j\in[1..m]} GCD(a_i, d_j, b_i - b_j) \mid d_i$

(iv) For every prime $p \leq m$ and every $I \subseteq I_p$ such that $|I| = p$ there are such $i, j \in I$, $i \neq j$ that $v_p(b_i - b_j) > M_p$.

In our case in place of a_i and b_i there will be some linear polynomials with rational coefficients.

As a corollary, we get that the relation $x \not\perp y$ is not positively existentially definable in this structure as otherwise the theory $\mathrm{Th}\langle\mathbb{Q}; 0, 1, +, -, =, \perp\rangle$ is decidable.

Conclusion

It is natural to ask for the following generalization of both Weispfenning's main theorem and Theorem 2. How the signature $\sigma = \left\langle 0, 1, +, -, [\,], \{c\cdot\}_{c\in\mathbb{Q}}, =, <, \perp \right\rangle$ can be extended with some predicates, positively existentially definable in $\langle\mathbb{Q}; \sigma\rangle$, such that for every positive existential formula there is some equivalent in this structure quantifier-free formula?

References

[1] A. Bel'tyukov, *Decidability of the universal theory of natural numbers with addition and divisibility (in Russian)*, Zapiski Nauchnyh Seminarov LOMI, vol. 60, 1976, pp. 15-28.

[2] L. Lipshitz, *The Diophantine problem for addition and divisibility*, Trans. Amer. Math. Soc., vol. 235, 1978, pp. 271-283.

[3] D. Richard, *Definability in Terms of the Successor Function and the Coprimeness Predicate in the Set of Arbitrary Integers*, The Journal of Symbolic Logic, vol. 54, no. 4, 1989, pp. 1253-1287.

[4] V. Weispfenning, *Mixed real-integer linear quantifier elimination*, International Symposium on Symbolic and Algebraic Computation (ISSAC), ACM Press, 1999, pp. 129-136.

Mikhail R. Starchak
Dept. of Informatics
Saint-Petersburg State University
St. Petersburg, Russia
e-mail: mikhstark@gmail.com

Symmetric polynomials, exterior power of the polynomial ring in one variable

Timur R. Seifullin

Abstract. We consider the r-th exterior power of the polynomial ring in one variable as a module over a ring of symmetric polynomials in r variables. It was obtained expliit expressions of symmetric polynomials via elementary symmetric polynomials.

Mathematics Subject Classification (2000). 15A69,15A75,15A72,13B25.

Keywords. Grassmann algebra, symmetric polynomial, elementary symmetric polynomial.

Let $\mathbf{R}$ be a commutative ring, x be a variable. $\mathbf{R}[x]^{\otimes r}$ is a commutative algebra over $\mathbf{R}$ as tensor product of commutative algebras over $\mathbf{R}$.

There is an isomorphism of algebras over $\mathbf{R}$

$\nu:\ \mathbf{R}[x_i|_{i=1,r}] \to \mathbf{R}[x]^{\otimes r},\ x_i \mapsto 1^{\otimes(i-1)}\otimes x\otimes 1^{\otimes(r-i)}$ for $i{=}1,r$.

The isomorphism ν induces the isomorphisms of algebras over $\mathbf{R}$

$$\mathsf{ts}^r(\mathbf{R}[x_i|_{i=1,r}]) \to \mathsf{TS}^r(\mathbf{R}[x]).$$

Here

$\mathsf{ts}^r(\mathbf{R}[x_i|_{i=1,r}])$ is the set of all symmetric polynomials in $\mathbf{R}[x_i|_{i=1,r}]$,

$\mathsf{TS}^r(\mathbf{R}[x])$ is the set of all symmetric tensors in $\mathbf{R}[x]^{\otimes r}$.

$\bigwedge^r(\mathbf{R}[x])$ *is a modules over* $\mathsf{TS}^r(\mathbf{R}[x])$.

Then by the isomorphism ν $\bigwedge^r(\mathbf{R}[x])$ *is a module over* $\mathsf{ts}^r(\mathbf{R}[x_i|_{i=1,r}])$,

Elementary symmetric polynomials $\sigma_p(x_i|_{i=1,r})$ are coefficients of the polynomial

$$\prod_{i=1}^{r}(x{-}x_i) = \sum_{p=0}^{r}\sigma_p(x_i|_{i=1,r})\cdot x^{r-p}.$$

Theorem 1. *Let* $\Delta{=}2$, $r{=}3$, $d{=}r{-}1$,

$$h_i(x) = \sum_{\delta=0}^{d+\Delta} h_{i,\delta}{\cdot}x^{\delta} \text{ for } i{=}1,r,$$

$$S(x_1,x_2,x_3) = \det\begin{Vmatrix} \sigma_0 & & h_{1,4} & h_{2,4} & h_{3,4} \\ \sigma_1 & \sigma_0 & h_{1,3} & h_{2,3} & h_{3,3} \\ \sigma_2 & \sigma_1 & h_{1,2} & h_{2,2} & h_{3,2} \\ \sigma_3 & \sigma_2 & h_{1,1} & h_{2,1} & h_{3,1} \\ & \sigma_3 & h_{1,0} & h_{2,0} & h_{3,0} \end{Vmatrix}.$$

Then

$$S(x{\otimes}1{\otimes}1, 1{\otimes}x{\otimes}1, 1{\otimes}1{\otimes}x){\cdot}(x^2{\wedge}x^1{\wedge}x^0) = h_1(x){\wedge}h_2(x){\wedge}h_3(x).$$

The last is equivalent to

$$S(x_1,x_2,x_3){\cdot}\det\begin{Vmatrix} x_1^2 & x_1^1 & x_1^0 \\ x_2^2 & x_2^1 & x_2^0 \\ x_3^2 & x_3^1 & x_3^0 \end{Vmatrix} = \det\begin{Vmatrix} h_1(x_1) & h_2(x_1) & h_3(x_1) \\ h_1(x_2) & h_2(x_2) & h_3(x_2) \\ h_1(x_3) & h_2(x_3) & h_3(x_3) \end{Vmatrix}.$$

Theorem 2. *Let* $\Delta{=}2$, $r{=}3$, $d{=}r{-}1$,

$$S(x_1,x_2,x_3) \in \mathsf{ts}^r(\mathbf{R}[(x_i)^{\leq\Delta}|_{i=1,r}]).$$

If

$$S(x{\otimes}1{\otimes}1, 1{\otimes}x{\otimes}1, 1{\otimes}1{\otimes}x){\cdot}(x^2{\wedge}x^1{\wedge}x^0) = \textstyle\sum(h_1^q(x){\wedge}h_2^q(x){\wedge}h_3^q(x)|q{\in}Q),$$

where

$$h_i^q(x) = \sum_{\delta=0}^{d+\Delta} h_{i,\delta}^q{\cdot}x^{\delta} \text{ for } i{=}1,r, \text{ for } q{\in}Q,$$

then

$$S(x_1,x_2,x_3) = \sum_{q\in Q} \det\begin{Vmatrix} \sigma_0 & & h_{1,4}^q & h_{2,4}^q & h_{3,4}^q \\ \sigma_1 & \sigma_0 & h_{1,3}^q & h_{2,3}^q & h_{3,3}^q \\ \sigma_2 & \sigma_1 & h_{1,2}^q & h_{2,2}^q & h_{3,2}^q \\ \sigma_3 & \sigma_2 & h_{1,1}^q & h_{2,1}^q & h_{3,1}^q \\ & \sigma_3 & h_{1,0}^q & h_{2,0}^q & h_{3,0}^q \end{Vmatrix}.$$

The first is equivalent to

$$S(x_1,x_2,x_3){\cdot}\det\begin{Vmatrix} x_1^2 & x_1^1 & x_1^0 \\ x_2^2 & x_2^1 & x_2^0 \\ x_3^2 & x_3^1 & x_3^0 \end{Vmatrix} = \sum_{q\in Q} \det\begin{Vmatrix} h_1^q(x_1) & h_2^q(x_1) & h_3^q(x_1) \\ h_1^q(x_2) & h_2^q(x_2) & h_3^q(x_2) \\ h_1^q(x_3) & h_2^q(x_3) & h_3^q(x_3) \end{Vmatrix}.$$

Thus $\mathsf{ts}^r(\mathbf{R}[(x_i)^{\leq\Delta}|_{i=1,r}]) \subseteq \mathsf{ts}^r(\mathbf{R}[(x_i)^{\leq 1}|_{i=1,r}])^{\Delta}$.

Theorem 3. *Let* $\Delta{=}3$, $r{=}3$, $d{=}3$,

$$h_i(x) = \sum_{\delta=0}^{d+\Delta} h_{i,\delta}{\cdot}x^\delta \text{ for } i{=}1,r.$$

Then

$$h_1(x){\wedge}h_2(x){\wedge}h_3(x)$$

$$= \det \left\| \begin{array}{cccccccccc} \sigma_0 & & & & & & & h_{1,6} & h_{2,6} & h_{3,6} \\ \sigma_1 & -1 & & & & & & h_{1,5} & h_{2,5} & h_{3,5} \\ \sigma_2 & & -1 & & & \sigma_0 & & h_{1,4} & h_{2,4} & h_{3,4} \\ \sigma_3 & & & -1 & & \sigma_1 & \sigma_0 & h_{1,3} & h_{2,3} & h_{3,3} \\ & & & & -1 & \sigma_2 & \sigma_1 & h_{1,2} & h_{2,2} & h_{3,2} \\ & & & & & \sigma_3 & \sigma_2 & h_{1,1} & h_{2,1} & h_{3,1} \\ & & & & & & \sigma_3 & h_{1,0} & h_{2,0} & h_{3,0} \\ \wedge & & x^3 & x^2 & x^1 & x^0 & & & & \end{array} \right\|.$$

Thus $\bigwedge^r(\mathbf{R}[x]^{\le d+\Delta}) \subseteq \mathsf{ts}^r(\mathbf{R}[(x_i)^{\le\Delta}|_{i=1,r}])\cdot \bigwedge^r(\mathbf{R}[x]^{\le d})$.

Denote by $\nabla(x_1,x_2;x)_*$ a map $\mathbf{R}[x] \to \mathbf{R}[x_1,x_2] \simeq \mathbf{R}[x]{\otimes}\mathbf{R}[x]$ such that

$$(x_1{-}x_2){\cdot}(\nabla(x_1,x_2;x)_*\,F(x)) = F(x_1){-}F(x_2).$$

Then $\nabla(x_1,x_2;x)_*$ is a cocommutative and coassociative coproduct. By a coassociativity of the the coproduct $\nabla(x_1,x_2;x)_*$ it determines the coproduct

$$\nabla(x_i|_{i=1,r};x)_*\colon \mathbf{R}[x] \to \mathbf{R}[x_i|_{i=1,r}] \simeq \mathbf{R}[x]^{\otimes r}.$$

Lemma 1. *Let* x *be a variable,* $F(x)$ *be a polynomial. Then*

1. $(\nabla(x_i|_{i=1,r};x)_*\,F(x)){\cdot}\det \left\|x_i^{r-j}\right\|_{j=1,r}^{i=1,r} = \det \left\|F(x_i) \quad x_i^{r-j}\right\|_{j=2,r-1}^{i=1,r}$.

2. $\det \left\|x_i^{r-j}\right\|_{j=1,r}^{i=1,r} = \prod\limits_{k=1}^{r-1}\Big(\prod\limits_{i=k+1}^{r}(x_i{-}x_k)\Big) = \prod\limits_{i,k:i>k}(x_i{-}x_k)$.

Let $\mathbf{x}_i = 1^{\otimes(i-1)}{\otimes}x{\otimes}1^{\otimes(r-i)}$ for $i=1,r$, then *1* of lemma 1 is equivalent to

$$(\nabla(\mathbf{x}_i|_{i=1,r};x)_*\,F(x)){\cdot}\Big(\bigwedge_{j=1}^{r} x^{r-j}\Big) = F(x){\wedge}\Big(\bigwedge_{j=2}^{r} x^{r-j}\Big).$$

Let d=3, r=4, Δ=r−1,

$$\mathbf{x}_1 = x\otimes 1\otimes 1\otimes 1,\ \mathbf{x}_2 = 1\otimes x\otimes 1\otimes 1,\ \mathbf{x}_3 = 1\otimes 1\otimes x\otimes 1,\ \mathbf{x}_4 = 1\otimes 1\otimes 1\otimes x,$$

$$F(x) = \sum_{\delta=0}^{d+\Delta} F_\delta \cdot x^\delta,$$

then *1* of lemma 1 is equivalent to

$$(\nabla(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4; x)_*\, F(x))\cdot(x^3\wedge x^2\wedge x^1\wedge x^0) = F(x)\wedge x^2\wedge x^1\wedge x^0,$$

and by of theorem 1

$$\nabla(x_1, x_2, x_3, x_4; x)_*\, F(x)$$

$$= \det\begin{Vmatrix} \sigma_0 & & & F_6 & & & \\ \sigma_1 & \sigma_0 & & F_5 & & & \\ \sigma_2 & \sigma_1 & \sigma_0 & F_4 & & & \\ \sigma_3 & \sigma_2 & \sigma_1 & F_3 & & & \\ \sigma_4 & \sigma_3 & \sigma_2 & F_2 & 1 & & \\ & \sigma_4 & \sigma_3 & F_1 & & 1 & \\ & & \sigma_4 & F_0 & & & 1 \end{Vmatrix} = \det\begin{Vmatrix} \sigma_0 & & & F_6 \\ \sigma_1 & \sigma_0 & & F_5 \\ \sigma_2 & \sigma_1 & \sigma_0 & F_4 \\ \sigma_3 & \sigma_2 & \sigma_1 & F_3 \end{Vmatrix}.$$

Timur R. Seifullin
V. M. Glushkov Institute of Cybernetics
National Academy of Sciences of Ukraine
e-mail: timur_sf@mail.ru

On probability distributions for the boundary states of the Hilbert-Schmidt ensemble of qudits

Vahagn Abgaryan, Arsen Khvedelidze, Ilya Rogojin and Astghik Torosyan

Let $\mathcal{H}$ be an $n-$dimensional Hilbert space and unitary group $U(\mathcal{H})$ acts on it preserving the standard Hermitian product. The set $\mathfrak{P}_n$ of density states ϱ of an $n-$dimensional quantum system is distinguished in the cone of non-negatively defined operators on $\mathcal{H}$ by the equation $\mathrm{Tr}(\varrho) = 1$ and can be regarded as embedded in the dual $\mathfrak{u}^*(n)$ of the Lie algebra $\mathfrak{u}(n)$.

It is known that the space $\mathfrak{P}_n$ of quantum states is not a differential manifold with smooth boundary. It is a stratified space, the union of different strata $\mathfrak{P}_n^k$ labelled by the rank $k = 1, 2, \ldots n$ of the quantum state, $\dim(\mathfrak{P}_n^k) = 2nk - k^2 - 1$ [1]. However, from the standpoint of dynamics of closed quantum system, it is often important to consider decomposition of the state space $\mathfrak{P}_n$ according to the unitary evolution. In this case a natural decomposition of the state space $\mathfrak{P}_n$ based on the coadjoint action of the unitary group $U(\mathcal{H})$ is relevant. The (co)adjoint action of the group $U(\mathcal{H})$ in $\mathfrak{u}^*(n)$ induces a corresponding non-transitive action on $\mathfrak{P}_n$ and thus different dimension if $k > 1$. Since the interior of the state space $\mathfrak{P}^n$ is a submanifold of the affine subspace of Hermitian operators of a unit trace, non-trivial differential structures in this stratification pattern appear only for the boundary $\partial\mathfrak{P}_n$, consisting of those density states ϱ for which $\det(\varrho) = 0$.

In the present report, we will describe generic features of a geometry of the boundary $\partial\mathfrak{P}_n$, particularly, its Riemannian characteristics in relation with the probability distributions of random states from the Hilbert-Schmidt ensemble of n-dimensional states, i.e., qudits.

An introduction of the notion of a distance between quantum states allows one to endow a quantum state space $\mathfrak{P}$ with a metric structure and thus consider $\mathfrak{P}$ as a Riemannian manifold. This enables us to relate geometrical concepts to a physical ones and use these concepts for studies of statistical properties of quantum systems [2]. The aim of our studies is a derivation of the probability distributions on the boundary $\partial\mathfrak{P}_n$, starting from the flat Hilbert-Schmidt metric on $\mathfrak{u}^*(n)$ [3]. It will be shown that an inherited metric on subsets $\mathfrak{P}_n^k$ gives rise to the joint probability density of a random rank-deficient states $\varrho \in \mathfrak{P}_n^k$, $k < n$, with real

eigenvalues $1 > \lambda_1 > \lambda_2 > \cdots > \lambda_k > 0$ of the so-called β- Wishart-Laguerre ensemble [4, 5]:

$$P_n^{\alpha,\beta}(\lambda_1, \lambda_2, \ldots, \lambda_k) = C_{n,\alpha,\beta} |\Delta_k(\{\lambda\})|^\beta \prod_{s=1}^{k} \lambda_s^\alpha e^{-\frac{1}{2}\beta\lambda_s} ,$$

where $C_{n,\alpha,\beta}$ is a normalization factor, $\Delta_k(\{\lambda\})$ is the Vandermonde determinant. The equation defining parameters α and β as function of the stratum $\mathfrak{P}_n^k$ is derived considering the induced metric on the degenerate unitary orbits $\mathcal{O}_\varrho \in \mathfrak{P}_n^k$.

References

[1] J. Grabowski, G. Marmo and M. Kus, *Geometry of quantum systems: Density states and entanglement*, J. Phys. A **38**, 10217, 2005.

[2] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer-Verlag, 2008.

[3] I. Bengtsson, K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press, 2017.

[4] M.L. Mehta, *Random Matrices*, 3rd edn. Academic Press, New York, 2004.

[5] P.J. Forrester, *Log-Gases and Random Matrices*, Princeton University Press, 2010.

Vahagn Abgaryan
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia
e-mail: vahagnab@googlemail.com

Arsen Khvedelidze
A Razmadze Mathematical Institute
Iv. Javakhishvili, Tbilisi State University
Tbilisi, Georgia
Institute of Quantum Physics and Engineering Technologies
Georgian Technical University
Tbilisi, Georgia
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia
e-mail: akhved@jinr.ru

Ilya Rogojin
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia

Probability distributions on boundary

Astghik Torosyan
Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980 Dubna, Russia
e-mail: astghik@jinr.ru

Waring Problem as an Issue of Polynomial Computer Algebra

Nikolai Vavilov

Abstract. In its original XVIII century form the classical Waring problem consisted in finding for each natural k the smallest such $s = g(k)$ that all natural numbers n can be written as sums of s non-negative k-th powers, $n = x_1^k + \ldots + x_s^k$. In the XIX century the problem was modified as the quest of finding such minimal $s = G(k)$ that *almost all* n can be expressed in this form. In the XX century this problem was further specified, as for finding such $G(k)$ *and* the precise list of exceptions. In the present talk I sketch the key steps in the solution of this problem, with a special emphasis on algebraic and computational aspects. I describe various connections of this problem, and its modifications, such as the rational Waring problem, the easier Waring problem, etc., with the current research in polynomial computer algebra, especially with identities, symbolic polynomials, etc. and promote several outstanding computational challenges.

Introduction

In this talk I plan to describe the status of the classical Waring problem, its versions and variants. The XVIII century Waring problem has been mostly solved. Not by Hilbert in 1909, of course, as many people misguidedly believe, but mostly by Dickson in 1936 (the outstanding small cases $k = 6, 5, 4$ were then settled in 1940, 1964 and 1984, respectively, see § 1 and § 6 for details). But already its XIX century version suggested by Jacobi, not to say all other major XIX and XX century variations, are widely open, as of today.

My objective is to attract attention to some algebraic and computational aspects of the Waring problem in the spirit of reconnecting with the goddess Namakkal, as described in [40]. Here I focus mostly on the related polynomial and rational identities, conjectural answers, and explicit lists of exceptions, many more details and further aspects can be found in [41, 42].

The present work was supported by the RFBR project N.19-29-14141.

1. Waring problem

Here we take a quick glance at some facets of what is known as the Waring problem. There are many further aspects to be featured in a more systematic treatment, as well as oodles of various generalisations and related problems, some of them mentioned towards the end of the present abstract and discussed in [41, 42].

1.1. Original Waring problem

Guided by the analogy with Lagrange's four squares theorem and scarce numerical evidence in 1770–1772 Waring and J. A. Euler (= Euler jr.) proposed what later became known as the [classical] Waring problem, see [14].

• **Waring problem**. Find for each natural k the smallest $s = g(k)$ such that every natural number n can be expressed as the sum of k-th powers of *non-negative* integers

$$n = x_1^k + \ldots + x_s^k,$$

with s summands.

Actually, Waring conjectured that $g(3) = 9$ and $g(4) = 19$, while J. A. Euler made similar prediction for *all* values of $g(k)$:

$$g(k) = 2^k + q + 2,$$

where $3^k = q \cdot 2^k + r$, $1 \leqslant r \leqslant 2^k - 1$, = the **ideal Waring theorem**.

In this form Waring problem was *essentially* solved in 1909–1984.

▪ In 1909 Wieferich [48] established that $g(3) = 9$, gaps in his proof were later filled up by Kempner [24] in 1912 and by Dickson in 1927.

▪ For $k \geqslant 7$ the problem was solved by Dickson [15, 16] and Pillai in 1936, modulo the **Pillai conjecture** that $q + r \leqslant 2^k$. They also compute the precise value of $g(k)$ when Pillai conjecture fails. But there is every reason to believe that Pillai conjecture holds for all k. Firstly, it may fail at most for finitely many values of k. Secondly, it holds for all $k < 5 \cdot 10^8$. And there is much more compelling evidence than that.

▪ The three remaining values $g(6) = 73$, $g(5) = 37$ and $g(4) = 19$ were computed by Pillai [28] in 1940, by Chen Jing-run [5] in 1964, and by Balasubramanian, Deshouillers and Dress [1] in 1984, respectively.

1.2. Asymptotic Waring problem

However, in XIX–XX centuries this problem was remodeled as follows.

• **Asymptotic Waring problem**. Find for each natural k the smallest s such that *almost all* natural numbers n can be expressed as the sum of k-th powers of *non-negative* integers $n = x_1^k + \ldots + x_s^k$, with s summands.

Clearly, the specific purport of this problem depends on the precise meaning of the expression *almost all*. The two most common interpretations are as follows:

▪ **According to Jacobi** as "all, except a finite number" = "all starting from a certain value". The corresponding minimal s is denoted by $G(k)$.

▪ **According to Hardy—Littlewood** in the sense of *natural density*. There can be inifinitely many exceptions, but they become progressively more rare, their number grows as $o(n)$. The corresponding minimal s is denoted by $G^+(k)$.

However apart from the case of squares $G^+(2) = G(2) = g(2) = 4$ that was already known to Lagrange, and a few more values such as $G(4) = 16$, $G^+(4) = 15$, as of today, the asymptotic Waring problem is very far from being solved in either sense.

1.3. Algorithmic Waring problem

However, with the advent of computers this problem was reformulated once again as something terribly much more ambitious.

• **Waring problem, XX century version**. Find $G(s)$ as above and the explicit list of exceptions. Construct an algorithm that for a given n finds a shortest expression of n as the sum of k-th powers (or, preferably, all such expressions).

In this form the problem seems to be quite recalcitrant. The *only* non-trivial case, for which Waring problem is fully *solved* in this form is that of biquadrates, see § 6. The only other case, for which the problem is fully *stated* in this form is that of cubes. However, for cubes we are nowhere near its solution and even the statement itself required thumping calculations, see § 5. For fifth powers it seems we are not even close to being able to *state* the problem in this form, see § 7.

1.4. Easier Waring problem

In the 1930-ies several mathematicians started to systematically consider the following version of Waring problem, which turned out to be *much* harder than the original Waring problem and is still unsolved even today.

• **Easier Waring problem**. Find for each natural k the smallest $s = v(k)$ such that all natural numbers n can be expressed as sums/differences of k-th powers of integers

$$n = \pm x_1^k \pm x_2^k \pm \ldots \pm x_s^k.$$

This is what Hardy and Wright call "sums affected with signs" and what Habsieger renamed **signed Waring problem**. They prove an obvious bound $v(k) \leqslant 2^k + (k!)/2$, [21], Theorems 400 and 401. There is a much better upper bound $v(k) \leqslant G(k) + 1$, of course. However, the explicit value of $v(k)$ is not known even for $k = 3$.

1.5. Rational Waring problem

Actually, there are further versions of Waring problem, also known since the early XIX century.

• **Rational Waring problem**. Find for each natural k the smallest $s = \rho(k)$ such that every rational number x can be expressed as sums/differences of k-th powers of rational numbers $n = \pm x_1^k \pm \ldots \pm x_s^k$.

• **Positive rational Waring problem**. Find for each natural k the smallest s such that every positive rational number x can be expressed as the sum of k-th powers of non-negative rational numbers $x = x_1^k + \ldots + x_s^k$.

These problems are closely related to another classical problem.

• **Waring problem at zero**. Find for each natural k the smallest $s = \theta(k)$ such that 0 can be *non-trivially* expressed as sums/differences of k-th powers of integers $\pm x_1^m \pm x_2^k \pm \ldots \pm x_s^k = 0$.

The existence of Pythagorean triples implies that $\theta(2) = 3$. The great Fermat theorem is the claim that $\theta(k) \geqslant 4$ for all $k \geqslant 3$. Using the geometry of elliptic curves, Fermat and Euler have proven that indeed Fermat equations $x^3 + y^3 = z^3$ and $x^4 + y^4 = z^4$ have no non-trivial solutions, and thus $\theta(3) = 4$ (Plato's cubes!) and $\theta(4) \geqslant 4$. Euler even made a much stronger claim that $\theta(k) \geqslant k + 1$, but that turned out to be both wrong and false. In particular, already $\theta(4) = 4$. Similarly, $4 \leqslant \theta(5) \leqslant 5$, but it is unknown, whether the precise value is 4 or 5.

2. Polynomial identities in the classical Waring problem

Here we display some assorted classical identities used to estimate $g(k)$. In my view, they deserve a serious further scrutiny, and with the tools of polynomial computer algebra we can now start a systematic search for new such identities.

2.1. Tardy type identities

Already in Euclid's "Elements" one can find the identity $4xy = (x+y)^2 - (x-y)^2$, later reproduced by Diophantus. Gauss generalised it to cubes

$$24xyz = (x+y+z)^3 - (x+y-z)^3 - (x-y+z)^3 + (x-y-z)^3.$$

In 1851 Tardy observed a similar identity for biquadrates

$$\begin{aligned} 192xyzw = {} & (x+y+z+w)^4 - (x+y+z-w)^4 - (x+y-z+w)^4 \\ & - (x-y+z+w)^4 + (x+y-z-w)^4 + (x-y+z-w)^4 \\ & + (x-y-z+w)^4 - (x-y-z-w)^4 \end{aligned}$$

and all further powers, and thus gave the first solution of the [cheap] *rational* Waring problem. This was clearly the starting point for Liouville and all subsequent development (Tardy was his student in Paris). Tardy identities were then rediscovered by Boutin in 1910.

2.2. Liouville type identities

The first non-trivial estimate for $g(k)$ for any $k \geqslant 3$ in Waring problem was obtained by Liouville some time before 1859, who proved that $g(4) \leqslant 53$. His

proof begins with the following identity. Let $2n = x^2 + y^2 + z^2 + w^2$, then

$$6n^2 = x^4 + y^4 + z^4 + w^4 + \Big((x+y+z+w)/2\Big)^4 + \Big((x+y+z-w)/2\Big)^4 + \Big((x+y-z+w)/2\Big)^4 + \Big((x+y-z-w)/2\Big)^4 + \Big((x-y+z+w)/2\Big)^4 + \Big((x-y+z-w)/2\Big)^4 + \Big((x-y-z+w)/2\Big)^4 + \Big((x-y-z-w)/2\Big)^4.$$

Later, Hurwitz and Venkov gave an interpretation of this identity in terms of integral quaternions, whereas Lucas has rewritten it in the form

$$6\big(x_1^2 + x_2^2 + x_3^2 + x_4^2\big)^2 = \sum (x_i + x_j)^4 + \sum (x_i - x_j)^4,$$

where both sums in the right-hand-side are taken over all $1 \leqslant i < j \leqslant 4$. Clearly, Lucas identity readily generalises:

$$6\big(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2\big)^2 = \sum (x_i + x_j)^4 + \sum (x_i - x_j)^4 - 2\sum x_h^4,$$

$$6\big(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2\big)^2 = \sum (x_i + x_j)^4 + \sum (x_i - x_j)^4 - 4\sum x_h^4,$$

and similarly for any m, where the sums are taken over $1 \leqslant i < j \leqslant m$, $1 \leqslant h \leqslant m$.

2.3. Maillet and Wieferich identities

To give a first non-trivial estimate of $g(3)$ Maillet used the following identity:

$$6x(x^2+y^2+z^2+w^2) = (x+y)^3 + (x-y)^3 + (x+z)^3 + (x-z)^3 + (x+w)^3 + (x-w)^3.$$

He himself derived this identity differently, but retrospectively, it is simply the derivative of the Liouville identity in Lucas form. Later, Linnik has used a more general identity

$$4(x_1^3 + y_1^3 + x_2^3 + y_2^3 + x_3^3 + y_3^3) = (x_1 + y_1)^3 + (x_2 + y_2)^3 + (x_3 + y_4)^3 + 3\big((x_1 + y_1)(x_1 - y_1)^2 + (x_2 + y_2)(x_2 - y_2)^2 + (x_3 + y_3)(x_3 - y_3)^2\big).$$

in his proof of the seven cube theorem.

Later, Maillet obtained a similar estimate for *fifth* powers, and Wieferich [47] explicitly produced the corresponding identity

$$2x\Big(2^2 \cdot 3 \cdot 5(43x^2 + y^2 + z^2 + w^2)^2 - 2^2 \cdot 1579x^4\Big) = (8x+y)^5 + (8x-y)^5 + (8x+z)^5 + (8x-z)^5 + (8x+w)^5 + (8x-w)^5 + (x+y+z+w)^5 + (x+y+z-w)^5 + (x+y-z+w)^5 + (x-y+z+w)^5 + (x+y-z-w)^5 + (x-y+z-w)^5 + (x-y-z+w)^5 + (x-y-z-w)^5.$$

In the same paper Wieferich used also a similar identity for *seventh* powers, which we do not reproduce here.

2.4. Fleck, Hurwitz and Schur identities

In 1907 Fleck came up with a similar identity for the 6-th powers,

$$60(x^2+y^2+z^2+w^2)^3 = 36(x^6+y^6+z^6+w^6)+$$
$$2\big((x+y)^6+(x-y)^6+\ldots+(z+w)^6+(z-w)^6\big)+$$
$$(x+y+z)^6+(x-y+z)^6+(x+y-z)^6+(x-y-z)^6+\ldots+(y-z-w)^6,$$

there are 12 summands in the second line (the choice of a pair, and a sign), and 16 summands in the third line (the choice of a triple and two *independent* choices of signs), 32 summands in total.

The same year Hurwitz has discovered the identity for 8-th powers,

$$5040(x^2+y^2+z^2+w^2)^4 = 6\big((2x)^8+(2y)^8+(2z)^8+(2w)^8\big)+$$
$$60\big((x+y)^8+(x-y)^8+\ldots+(z+w)^8+(z-w)^8\big)+$$
$$(2x+y+z)^8+(2x-y+z)^8+(2x+y-z)^8+(2x-y-z)^8+\ldots+(-y-z+2w)^{10}+$$
$$6\big((x+y+z+w)^8+(x+y+z-w)^8+\ldots+(x-y-z-w)^8\big).$$

and conjectured the existence of such similar identities expressing [some multiple of] $(x^2+y^2+z^2+w^2)^k$ as the sum of $2k$-th powers of linear forms in x,y,z,w for all k. The next such identity was indeed constructed the same year by Schur,

$$22680(x^2+y^2+z^2+w^2)^5 = 9\big((2x)^{10}+(2y)^{10}+(2z)^{10}+(2w)^{10}\big)+$$
$$180\big((x+y)^{10}+(x-y)^{10}+\ldots+(z+w)^{10}+(z-w)^{10}\big)+$$
$$(2x+y+z)^{10}+(2x-y+z)^{10}+(2x+y-z)^{10}+(2x-y-z)^{10}+\ldots+(-y-z+2w)^{10}+$$
$$9\big((x+y+z+w)^{10}+(x+y+z-w)^{10}+\ldots+(x-y-z-w)^{10}\big),$$

Observe that these identities have 12 summands in the second line (the choice of a pair and a sign), 48 summands in the third line (the choice of one position out of four for the coefficient 2, the choice of one of the three remaining positions for the coefficient 0 and two *independent* choices of signs), and, finally, 8 summands in the last line (three *independent* choices of signs for all positions other than the first one), 72 summands in total.

2.5. Hilbert type identities

In 1909 Hilbert [22] solved a cheap version of the classical Waring problem = mere finiteness of $g(k)$, without computing the actual value, or actually providing *any* estimate of $g(k)$. As part of his solution, Hilbert verified the above Hurwitz conjecture. In fact, he has proven that there *exist* identities expressing k-th power of the sum of m squares as *positive* linear combinations of $q=\binom{2k+1}{m}$ expressions which are $(2k)$-th powers of *linear forms* :

$$a(x_1^2+\ldots+x_m^2)^k = a_1(b_{11}x_1+\ldots+b_{1m}x_m)^{2k}+\ldots+a_q(b_{q1}x_1+\ldots+b_{qm}x_m)^{2k},$$

where $a,a_i\in\mathbb{N}$ and $b_{ij}\in\mathbb{Z}$, for $1\leqslant i\leqslant q$, $1\leqslant j\leqslant m$.

Actually in his solution of the cheap Waring problem Hilbert only used the identities for $m = 5$, but his method is quite general and allows to prove the *existence* of similar identities **Hilbert identities** for arbitrary m and k. His proof is a pure existence proof and, in its original form, does not give any estimate on the size of the coefficients.

A posteriori, many further such identities were explicitly written. Say, by Kürschak [25] in 1911, for $k = 2$ and $m \equiv 1 \pmod 3$:

$$60\big(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2\big)^2 = \sum (x_i \pm x_j \pm x_h)^4,$$

$$672\big(x_1^2 + x_2^2 + x_3^2 + \ldots + x_9^2 + x_{10}^2\big)^2 = \sum (x_i \pm x_j \pm x_h \pm x_l)^4,$$

etc. By Kempner [24] in 1912, for $m = 4$ and $k = 6, 7$. Note also the next Fleck identity

$$60\big(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2\big)^3 = \sum (x_i \pm x_j \pm x_h)^6 + 36 \sum x_l^6,$$

and the like. Some *estimates* on the size of coefficients in Hilbert identities were later produced by Rieger, Pollack and Nesterenko [35, 29, 27] in the process of effectivisation of Hilbert's proof, but they have not attempted to come up with the actual coefficients.

The following problem seems to be extremely significant not just as a direct mathematical and computational challenge, but also as a methodological, historical and philosophical issue.

Problem 1. *Can one solve the original Waring problem with Hilbert's approach*?

If mathematics is what we think it is, this should be possible. Personally, I would feel very disappointed should Wieferich [48] proof of the equality $g(3) = 9$ and the estimate $g(4) \leqslant 30$ Dress [17] remain the best and only partial solutions obtained along these lines.

However, it will be by no means easy.

Problem 2. *Implement a systematic computer search of Hilbert type and similar identities with small coefficients.*

3. Polynomial identities in the easier and rational Waring problems

Identities that allow to give estimates of $v(k)$ and $\rho(k)$ are shorter and [in a sense] easier, but much less understood, than the identities used to estimate $g(k)$.

3.1. Richmond identity, Norrie identity, and beyond

Actually, Tardy identities and all such further uniform series of identities give vastly exaggerated upper bounds for $\rho(k)$ in the rational Waring problem. So far, getting the best possible estimate required separate clever identities in each individual case.

Below we reproduce two such classical identities, stemming from 1920-ies, **Richmond identity** for cubes

$$x = \left(\frac{x^3 - 3^6}{3^2x^2 + 3^4x + 3^6}\right)^3 + \left(\frac{-x^3 + 3^5x + 3^6}{3^2x^2 + 3^4x + 3^6}\right)^3 + \left(\frac{3^3x^2 + 3^5x}{3^2x^2 + 3^4x + 3^6}\right)^3$$

and **Norrie identity** for biquadrates

$$x = \left(\frac{a^2(a^8 - b^8 + 2x)}{2(a^8 - b^8)}\right)^4 - \left(\frac{a^2(a^8 - b^8 - 2x)}{2(a^8 - b^8)}\right)^4 + \left(\frac{2a^4x - b^4(a^8 - b^8)}{2ab(a^8 - b^8)}\right)^4 - \left(\frac{2a^4x + b^4(a^8 - b^8)}{2ab(a^8 - b^8)}\right)^4.$$

There were similar identities for $k = 5, 6, 7, 8, 9$, but they are constructed ad hoc, and there is no clear pattern as to their shape. Compare, in particular, Choudhry or Reynya [6, 7, 8, 33, 34].

3.2. Rao and Vaserstein identities, and beyond

However, the works by Habsieger [19, 20] give some hope. Imitating the classical Rao identity for *sixth* powers,

$$\begin{aligned}12abcd(c^4 - d^4)(a^{24} - b^{24})x = \\ (a^5c + bdx)^6 + (a^5d - bcx)^6 + (b^5c - adx)^6 + (b^5d + acx)^6 \\ - (a^5c - bdx)^6 - (a^5d + bcx)^6 - (b^5c + adx)^6 - (b^5d - acx)^6.\end{aligned}$$

Vaserstein [38] discovered a similar identity for *eighth* powers. Habsieger [20] has rewritten Vaserstein identity in the following more symmetric form:

$$\begin{aligned}16(uvw)^6(u^{48}v^{64} + v^{48}w^{64} + w^{48}u^{64} - u^{48}w^{64} - v^{48}u^{64} - w^{48}v^{64})y = \\ (u^7v^{10} + u^5w^6y)^8 + (u^7w^{10} - u^5v^6y)^8 + (v^7w^{10} + v^5u^6y)^8 \\ + (u^7u^{10} - v^5w^6y)^8 + (w^7u^{10} + w^5v^6y)^8 + (w^7v^{10} - w^5u^6y)^8 \\ - (u^7v^{10} - u^5w^6y)^8 - (u^7w^{10} + u^5v^6y)^8 - (v^7w^{10} - v^5u^6y)^8 \\ - (u^7u^{10} + v^5w^6y)^8 - (w^7u^{10} - w^5v^6y)^8 - (w^7v^{10} + w^5u^6y)^8\end{aligned}$$

At this point the link to the representation theory of finite groups becomes obvious, and Habsieger [19, 20] is able to construct many similar **symmetric identities**.

Problem 3. *Is it possible to construct series of rational identities of all degrees that would give correct bound in the easier Waring problem and in the rational Waring problem?*

3.3. Becker type identities

In 1979 Becker [2] constructed analogues of Hilbert identities

$$(x_1^l + \ldots + x_m^l)^k = f_1(x_1, \ldots, x_m)^{lk} + \ldots + f_q(x_1, \ldots, x_m)^{lk}$$

for arbitrary k, l, m. However, for $l \geqslant 4$ the f_j's here have to be rational functions rather than polynomials. If they had been polynomials, they are bound to be linear forms, which immediately leads to a contradiction.

I plan to demonstrate some such explicit identities in my talk.

3.4. Frolov type identities

There are another type of identities that were used in the easier Waring problem, which come from the solution of Prouhet—Tarry—Escott problem, and which oftentimes lead to better bounds for $v(k)$, than the bounds obtained via the above symmetric identities.

Recall that, given natural numbers s and k with $s > k$, the Prouhet—Tarry—Escott problem (or simply PTE for short) asks, whether there are distinct multisets of integers, say $X = [x_1, \ldots, x_s]$ and $Y = [y_1, \ldots, y_s]$, such that

$$x_1^i + \ldots + x_s^i = y_1^i + \ldots + y_s^i, \qquad j = 1, \ldots, k.$$

Hosts of special/partial solutions to this problem were constructed in the late XIX century and in the early XX century.

The relevance of PTE resides in the fact that every such solution leads to the corresponding **Frolov identity**

$$(t + x_1)^k + \ldots + (t + x_s)^k = (t + y_1)^k + \ldots + (t + y_s)^k.$$

These and similar identities were extensively used by Demianenko, Revoy [31, 32] and others to obtain sharper bounds in the easier Waring problem.

4. Vinogradov's method

In the early 1920-ies Hardy and Littlewood considered the generating function

$$f_k(z) = 1 + z^{1^k} + z^{2^k} + z^{3^k} + \ldots$$

Then the coefficient $r_{k,s}(n)$ of z^n in the series

$$f_k(z)^s = 1 + \sum_{n=1}^{\infty} r_{k,s}(n) z^n$$

equals the number of representations of n as the sum of k-th powers of s non-negative integers. In particular, the original Waring conjecture is equivalent to the claim that $r_{3,9}(n) \neq 0$, that $r_{4,19}(n) \neq 0$, that $r_{5,37}(n) \neq 0$, etc., for all natural n.

Side remark. Actually, Hardy and Littlewood considered a slightly different generating function, namely $f_k(z) = 1 + 2z^{1^k} + 2z^{2^k} + 2z^{3^k} + \ldots$ But this is pure fetishism, explained by the fact that for $k = 2$ such a choice of the generating function leads to the Jacobi theta-function, and explicit computation of $r_{2,s}(n)$. We do not know, what could be a correct choice of the coefficients in the generating function that would produce a similar theory for higher degrees. If we do not attempt to calculate explicit values, but are interested only in the **asymptotic**

behaviour of $r_{k,s}(n)$, the specific choice of the generating function does not play any role anyway.

As a function of the complex variable $z \in \mathbb{C}$ this series converges inside the unit disk, but the circle $|z| = 1$ consists entirely of singular points. The idea of the **circle method** is to use the Cauchy formula

$$r_{k,s}(n) = \frac{1}{2\pi i}\int_C \frac{f_k(z)^s}{z^{n+1}}dz,$$

where C is the circle of radius $0 < \rho < 1$, and then to *estimate* this integral when $\rho \longrightarrow 1$, using the character of singularities on the unit circle.

In the late 1920-ies Vinogradov proposed a radical simplification of this method. Namely, he noticed that if we are interested in the number of representations of a *specific* n as the sum of s non-negative k-th powers, then we do not have to look at the whole generating function, as Hardy and Littlewood did. In fact, the whole infinite tail of the generating function does not play any role, we can limit ourselves with the *polynomial*

$$f_{k,N}(z) = 1 + z^{1^k} + z^{2^k} + \ldots + z^{N^k}.$$

Then the coefficient $r^N_{k,s}(n)$ of z^n in the *polynomial*

$$f_{k,N}(z)^s = 1 + \sum_{n=1}^{sN} r^N_{k,s}(n)z^n$$

equals the number of representations of n as the sum of k-th powers of $\leqslant s$ integers $1 \leqslant m \leqslant N$.

Clearly, the integers m such that $m^k > n$ cannot occur in such a representation. Thus, for any $N \geqslant \sqrt[k]{n}$ one has $r^N_{k,s}(n) = r_{k,s}(n)$. Thus, in Vinogradov's method the passage to limits still occurs, but now we can from the onset assume that $\rho = 1$ and calculate the limit as $N \longrightarrow \infty$, which is a dramatic technical simplification.

It was precisely this idea that allowed to improve bounds on $G(k)$ from exponential in k to polynomial in k (and, eventually, to almost linear in k). It was precisely the huge gap between the expected exponential bound for $g(k)$ and the polynomial bound for $G(k)$ that allowed to apply Dickson's ascent.

Problem 4. *Can one solve the original Waring problem as a problem of polynomial computer algebra by directly verifying that for any k and n and any $N \geqslant \sqrt[k]{n}$ there exists an s such that $r^N_{k,s}(n) \neq 0$?*

The idea is to try to explicitly process $f_{k,N}$ as symbolic polynomials in the fashion of Steven Watt [43, 44, 45, 46].

5. Algorithmic Waring problem for cubes

In 1909 Landau proved by the methods of *elementary* analytic number theory that $G(3) \leqslant 8$, in other words, almost all positive integers are sums of $\leqslant 8$ positive

cubes. Indeed, in 1939 Dickson established that the only positive integers that require 9 cubes are 23 and 239. In 1943 Linnik proved his famous seven cubes theorem asserting that $G(3) \leqslant 7$. A few years ago this result was made explicit.

5.1. Experimental evidence for cubes.

Based on extensive computer calculations, asymptotics in the Hardy—Littlewood theory, and probabilistic trials Romani stated the following conjectures [36].

▪ **Problem of seven cubes.** There are exactly 15 natural numbers that can be expressed as sums of *eight*, but not of *seven* non-negative cubes, the largest of them being 454.

▪ **Problem of six cubes.** There are exactly 121 natural numbers that can be expressed as sums of *seven*, but not of *six* non-negative cubes, the largest of them being 8042.

▪ **Problem of five cubes.** There are exactly 3922 natural numbers that can be expressed as sums of *six*, but not of *five* non-negative cubes, the largest of them being 1290740.

As a further evidence for that in 1999 Bertault, Ramaré and Zimmerman [3] established that all integers between 1290740 and $3.375 \cdot 10^{12}$ can be expressed as sums of five cubes, which by Dickson's ascent implies that all integers between 455 and $2.5 \cdot 10^{26}$ can be expressed as sums of seven cubes. The same year Deshouillers, Hennecart and Landreau [11] extended these calculations to 10^{16}.

The largest natural number known today that requires exactly five cubes is 7373170279850. In 1999 Deshouillers, Hennecart and Landreau [11] stated the following conjecture (l. c., Conjectures 1 and 2):

▪ **Problem of four cubes.** There are exactly 113936676 natural numbers that can be expressed as sums of *five*, but not of *four* non-negative cubes, the largest of them being 7373170279850.

5.2. Problem of seven cubes.

In 2005 Ramaré published yet another effectivisation of Linnik's theorem: all integers

$$n \geqslant e^{205000} \approx 2.3377074809 \cdot 10^{89030}.$$

can be expressed as sums of seven cubes. The improvement was based on the **Bombieri identity**

$$\begin{aligned} 2(u^6v^6 + u^6w^6 + v^6w^6)a^3 + 6au^2v^2w^2(x^2 + y^2 + z^2) = \\ (u^2v^2a + wx)^3 + (u^2v^2a - wx)^3 + (u^2w^2a + vy)^3 + \\ (u^2w^2a - vy)^3 + (v^2w^2a + uz)^3 + (v^2w^2a - uz)^3, \end{aligned}$$

In 2007 Ramaré [30] further dramatically improved the bound to

$$n \geqslant e^{524} \approx 3.71799 \cdot 10^{227},$$

after which it became clear that a complete solution was close.

In 2008–2009 Boklan and Elkies [4] proved the seven cube conjecture for numbers divisible by 4, and in 2010 Elkies [18] proved it for all even integers. These results essentially used both the Ramaré upper bound, and the Deshouillers—Hennecart—Landreau lower bound. Finally, in 2015 Siksek announced a complete solution of the problem, which was published in 2016 in [37]. The only numbers which cannot be presented in such a form are

15, 22, 23, 50, 114, 167, 175, 186, 212, 231, 238, 239, 303, 364, 420, 428, 454.

Among other things, this work relies on dozens of thousands hours of computer time.

However, the problems of six, five and four cubes are still wi[l]d[e]ly open!

6. Algorithmic Waring problem for biquadrates

Dickson's estimate $g(4) \leqslant 35$ has not been improved for almost 40 years. However, in 1970–1971 Dress had a happy idea to return to the elementary approach with new techniques. In particular, using new polynomial identities that occurred in the solution of the easier Waring problem, and some computer calculations, he improved the bound to $g(4) \leqslant 30$ *by elementary methods*. After that things accelerated, see [41] for a detailed description.

6.1. Nineteen biquadrates.

In 1985 Deshouillers announces a complete solution of the original Waring problem in the last remaining case of biquadrates. Observe the $\geqslant 125$ year gap between the Liouville breakthrough (who proved not mere *finiteness* of $g(4)$, but established a realistic estimate!), and the final solution of the Waring problem $g(4) = 19$, as stated by Waring himself.

In 1985 Balasubramanian, Deshouillers and Dress [1] announce the general plan of such a solution. In [1] it is claimed that all integers $n \geqslant 10^{367}$ are sums of 19 biquadrates, the details were then published in [9]. Moreover in [1] the authors describe a calculation that shows that all natural numbers $n \leqslant 10^{378}$ are also sums of 19 biquadrates. Later in [10] this computation is even extended to $n \leqslant 10^{448}$. Thus, the upper and lower domains overlap by 80 orders of magnitude!

6.2. Sixteen biquadrates.

In 1939 Davenport has proven that $G(4) = 16$. Now we know that 13792 is the largest integer that requires more than 16 biquadrates, all $n \geqslant 13793$ are in fact sums of 16 biquadrates. This was shown in 1999–2005 by Deshouillers, Hennecart, Kawada, Landreau and Wooley.

Namely, in [13] it is proven that all integers $n \geqslant 10^{216}$ not divisible by 16, are sums of 16 biquadrates. The proof of this result uses new polynomial identities. Also, the authors had to rework the estimates in and around the circle method from scratch and *with explicit constants*. On the other hand, in 2000 Deshouillers, Hennecart and Landreau [12] established that all $13793 \leqslant n \leqslant 10^{245}$ are sums of 16

biquadrates. Thus, again the upper and lower domains overlap and for biquadrates we can give a *complete* answer to Waring problem. There are exactly 96 natural numbers that are not sums of 16 biquadrates, here they are:

47, 62, 63, 77, 78, 79, 127, 142, 143, 157, 158, 159, 207, 222, 223, 237, 238, 239, 287, 302, 303, 317, 318, 319, 367, 382, 383, 397, 398, 399, 447, 462, 463, 477, 478, 479, 527, 542, 543, 557, 558, 559, 607, 622, 623, 687, 702, 703, 752, 767, 782, 783, 847, 862, 863, 927, 942, 943, 992, 1007, 1008, 1022, 1023, 1087, 1102, 1103, 1167, 1182, 1183, 1232, 1247, 1248, 1327, 1407, 1487, 1567, 1647, 1727, 1807, 2032, 2272, 2544, 3552, 3568, 3727, 3792, 3808, 4592, 4832, 6128, 6352, 6368, 7152, 8672, 10992, 13792

for each one of them it is very easy to determine, whether it requires 17, 18 or 19 biquadrates.

7. The big computational challenge

As we've seen above, $k = 4$ is *the only* case (apart from that of $k = 2$, known to Lagrange back in 1770), when Waring problem has been completely solved in the XX century sense. Even in the case $k = 3$ there is a *huge* uncertainty $4 \leqslant G(3) \leqslant 7$ as to the actual value of $G(3)$ — not to say the explicit list of exceptions!

To give some idea of the computational immensity of the problem, below we reproduce the table of values of $g(k)$, $5 \leqslant k \leqslant 15$, as confronted with the *conjectural* values of $G(k)$ — with the known upper *estimates* of $G(k)$, coming mostly from the work of Vaughan and Wooley (see, for instance, [39]) somewhere in between.

k	5	6	7	8	9	10	11	12	13	14	15
$g(k)$	37	73	143	279	548	1079	2132	4223	8384	16673	33203
$G(k) \leqslant$	17	24	33	42	50	59	67	76	84	92	100
$G(k) =$	6	9	8	32	13	12	12	16	14	15	16

TABLE 1. Conjectured values of $G(k)$ for $5 \leqslant k \leqslant 15$

It would be a rather ambitious project simply to repeat with the use of computers what Dickson has accomplished *by hand* back in the 1930-ies. But of course, today we should set much higher goals, namely, to try to document the explicit lists of exceptions that require more than $G(k)$ non-negative k-th powers.

Can we do this? Say for the cases $5 \leqslant k \leqslant 20$, with which Dickson *started*? For instance, $g(5) = 37$, while $G(5) = 6$, as everybody believes, so that we have to verify one by one all values $s = 37, 36, \ldots, 7$ and towards the end of this list the possible exceptions are bound to occur well into 10^{hundreds}. So here is the warm up problem, which would show, where we are, as far as the computational power.

Problem 5. *Compute for each $s = 37, \dots, 7$ the explicit list of natural n which can be expressed as sums of s non-negative fifth powers, and cannot be expressed as shorter such sums.*

If we can do this, about what I have some doubts, we could proceed to higher powers, and see where we have to stop. It seems to me, that the XX century form of Waring problem is well beyond our current grasp — or what's the metaphor.

Conclusion

Poincaré used to say "Il n'y a pas de problèmes résolus, il n'y a que des problèmes plus ou moins résolus". Waring problem is certainly one of the kind. Despite the egregious efforts of many generations of mathematicians, even the XVIII century Waring problem is only 99.9999% solved, and in the meantime we were able to *fully* solve the XIX–XX century forms of the problem (with an explicit list of exceptions) for a single new case, $G(4) = 16$.

Here are my principles — well, problems — if you don't like them, I have other[s]. One can ask the same questions for other fields and rings, in particular, for number rings other than $\mathbb{Z}$, for polynomial rings, fields of rational fractions, etc. (compare the recent papers by Im Bo-Hae, Larsen, and Nguyen Dong Quan Ngoc [23, 26] for a whole new look at Waring type problems, in the context of algebraic groups). There are simultaneous sums of powers, Euler problem, taxicab numbers, PTE and variants, etc. Not to say, the mixed Waring problems, the restricted Waring problems, the Waring—Goldbach problems of all sorts, the Kamke type problems, etc. And, of course, we are still not anywhere close to doing for cubes what Jacobi has done for squares, the explicit formulas for the number of representations.

References

[1] R. Balasubramanian, J. M. Deshouillers, F. Dress, *Problème de Waring pour les bicarrés.* I. *Schéma de la solution.* C. R. Acad. Sci. Paris Sér. I Math. **303** (1986), no. 4, 85–88; II. *Résultats auxiliaires pour le théorème asymptotique.* C. R. Acad. Sci. Paris Sér. I Math. **303** (1986), no. 5, 161–163.

[2] E. Becker, *Summen n-ter Potenzen in Körpern.* J. Reine Angew. Math. **307** (1979), 8–30.

[3] F. Bertault, O. Ramaré, P. Zimmermann, *On sums of seven cubes.* Math. Comp. **68** (1999), 1303–1310.

[4] K. D. Boklan, N. D. Elkies, *Every multiple of* 4 *except* 212*,* 364*,* 420*, and* 428 *is the sum of seven cubes*, arXiv:0903.4503v1 [math.NT] 26 Mar 2009, 1–8.

[5] Chen Jing-run, *Waring's problem for $g(5) = 37$* (Chinese). Sci. Sinica, **13** (1964), 1547–1568; translated as Chinese Math. Acta, **6** (1965), 105–127.

[6] A. Choudhry, *Representation of every rational number as an algebraic sum of fifth powers of rational numbers.* Enseign. Math. (2) **35** (1989), no. 1–2, 19–20.

[7] A. Choudhry, *On sums of eighth powers.* J. Number Theory **39** (1991), no. 1, 104–107.

[8] A. Choudhry, *On sums of seventh powers.* J. Number Theory **81** (2000), no. 2, 266–269.

[9] J. M. Deshouillers, F. Dress, *Sums of* 19 *biquadrates*: *on the representation of large integers.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **19** (1992), no. 1, 113–153.

[10] J. M. Deshouillers, F. Dress, *Numerical results for sums of five and seven biquadrates and consequences for sums of* 19 *biquadrates.* Math. Comp. **61** (1993), 195–207.

[11] J.-M. Deshouillers, F. Hennecart, B. Landreau 7373170279850, Math. Comp., **69** (2000), 421–439; With an appendix by I. Gusti Putu Purnaba.

[12] J.-M. Deshouillers, F. Hennecart, B. Landreau *Waring's problem for sixteen biquadrates — numerical results.* Colloque International de Théorie des Nombres (Talence, 1999). J. Théor. Nombres Bordeaux **12** (2000), no. 2, 411–422.

[13] J.-M. Deshouillers, K. Kawada, T. D. Wooley, *On sums of sixteen biquadrates.* Mém. Soc. Math. Fr. (N.S.), no. 100 (2005), 120p.

[14] L. E. Dickson, *History of the theory of numbers, vol.* II. *Diophantine analysis*, Chelsea, 1952, 803p.

[15] L. E. Dickson, *Proof of the ideal Waring theorem for exponents* 7–180, Amer. J. Math. **58** (1936), 521–529.

[16] L. E. Dickson, *Solution of Waring's problem.* Amer. J. Math. **58** (1936), 530–535.

[17] F. Dress, *Amélioration de la majoration de* $g(4)$ *dans le problème de Waring*: $g(4) \leqslant 30$. Acta Arith., **22** (1973), 137–147.

[18] N. D. Elkies, *Every even number greater than* 454 *is the sum of seven cubes,* arXiv:1009.3983 [math.NT] [v1] Tue, 21 Sep 2010, 1–9.

[19] L. Habsieger, *Représentations des groupes et identités polynomiales.* Sém. Théor. Nombres Bordeaux (2) **3** (1991), no. 1, 1–11.

[20] L. Habsieger, *Applications of group representation theory to the easier Waring problem.* J. Number Theory **45** (1993), no. 1, 92–111.

[21] G. H. Hardy, E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford University Press, Oxford, 1979.

[22] D. Hilbert D., *Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl* n^{ter} *Potenzen* (*Waringsches problem*). Dem Andenken an Hermann Minkowski gewidmet. Gött. Nachr., (1909), 17–36; Abdruck mit Veränderungen und Zusätzen: Math. Ann **67** (1909), 281–300.

[23] Im Bo-Hae, M. Larsen, *Waring's problem for rational functions in one variable.* Q. J. Math. **71** (2020), no. 2, 439–449.

[24] A. Kempner, *Bemerkungen zum Waringschen Problem.* Math. Ann. **72** (1912), 387–399.

[25] J. Kürschák, *Über die Liouvillesche Identität.* Arch. Math. Phys. (3) **18** (1911), 242–243.

[26] M. Larsen, Nguyen Dong Quan Ngoc, *Waring's problem for unipotent algebraic groups.* Ann. Inst. Fourier (Grenoble) **69** (2019), no. 4, 1857–1877.

[27] Yu. V. Nesterenko, *On Waring's problem* (*elementary methods*). J. Math. Sci. (N.Y.) **137** (2006), no. 2, 4699–4715.

[28] S. S. Pillai, *On Waring's problem*: $g(6) = 73$, Proc. Indian Acad. Sci., Sect. A **12** (1940), 30–40.

[29] P. Pollack, *On Hilbert's solution of Waring's problem.* Cent. Eur. J. Math. **9** (2011), no. 2, 294–301.

[30] O. Ramaré, *An explicit result of the sum of seven cubes,* Manuscripta Math. **124** (2007), no. 1, 59–75.

[31] Ph. Revoy *Sur les sommes de quatre cubes.* Enseign. Math. (2) **29** (1983), no. 3–4, 209–220.

[32] Ph. Revoy *Le problème facile de Waring.* Enseign. Math. **37** (1991), 223–234.

[33] M. A. Reynya, *Representations of a rational number as a sum of odd powers.* Int. J. Number Theory **12** (2016), no. 4, 903–911.

[34] M. A. Reynya, *Representation of a rational number as a sum of ninth or higher odd powers.* Funct. Approx. Comment. Math. **58** (2018), no. 1, 79–87

[35] G. J. Rieger, *Zur Hilbertschen Lösung des Waringschen Problems*: *Abschätzung von* $g(n)$. Arch. Math. **4** (1953), 275–281.

[36] F. Romani, *Computations concerning Waring's problem for cubes,* Calcolo, **19** (1982), 415–431.

[37] S. Siksek, *Every integer greater than* 454 *is the sum of at most seven positive cubes.* Algebra Number Theory **10** (2016), no. 10, 2093–2119.

[38] L. N. Vaserstein, *Every integer is a sum or difference of* 28 *integral eighth powers.* J. Number Theory, **28** (1988), 66–68.

[39] R. C. Vaughan, T. D. Wooley, *Waring's problem: a survey.* In Number Theory for the Millennium, III (Urbana, IL, 2000) (2002), 301–340. https://www.researchgate.net/publication/2842101_Waring's_Problem_A_Survey

[40] N. A. Vavilov *Computers as novel mathematical reality*: I. *Personal account.* Computer Instruments in Education, 2020, 1–20 (in Russian).

[41] N. A. Vavilov *Computers as novel mathematical reality*: II. *Waring problem.* Computer Instruments in Education, 2020, 1–47 (in Russian).

[42] N. A. Vavilov *Computers as novel mathematical reality*: III. *Easier Waring problem.* Computer Instruments in Education, 2021, 1–42 (in Russian).

[43] S. M. Watt, *Making Computer Algebra More Symbolic*, 43–49, Proc. Transgressive Computing 2006: A conference in honor or Jean Della Dora, (TC 2006), April 24–26 2006, Granada, Spain.

[44] S. M. Watt, *Two Families of Algorithms for Symbolic Polynomials,* 193–210, in Computer Algebra 2006: Latest Advances in Symbolic Algorithms, Proceedings of the Waterloo Workshop, World Scientific 2007.

[45] S. M. Watt, *Symbolic Polynomials with Sparse Exponents,* 91–97, Proc. Milestones in Computer Algebra: a Conference in Honour of Keith Geddes' 60th Birthday, (MICA 2008), May 1–3 2008, Stonehaven Bay, Trinidad and Tobago, University of Western Ontario

[46] S. M. Watt, *Functional Decomposition of Symbolic Polynomials,* 353–362, Proc. International Conference on Computational Sciences and its Applications, (ICCSA 2008), June 30–July 3 2008, Perugia, Italy, IEEE Computer Society.

[47] A. Wieferich *Zur Darstellung der Zahlen als Summen von fünften und siebenten Potenzen positiver ganzer Zahlen.* Math. Ann. **67** (1909), 61–75.

[48] A. Wieferich, *Beweis des Satzes, daß sich eine jede ganze Zahl als Summe von höchstens neun positiven Kuben darstellen läßt.* Math. Ann. **66** (1909), 95–101.

Nikolai Vavilov
Dept. of Mathematics and Computer Science
St Petersburg State University
St Petersburg, Russia
e-mail: nikolai-vavilov@yandex.ru

Numerations of the partially ordered sets and generalized Coxeter groups

Anatoly Vershik

We discuss a project in which combinatorics of the posets are studied together with some class of groups
Distributive lattices $L(P)$ of ideals of poset P. Hasse diagram of $L(G)$, Example: Young graph. Finite (countable) Coxeter group G(P) associated with finite (resp.countable) poset (P) Problem: When G(P) is isomorphic to a classical Coxeter group? A classification of the posets dependently of classification of the central measures on $L(P)$. Concrete problems and conjectures.

This work was carried out as part of a project supported by an RFBR grant 17-01-00433

Anatoly Vershik
St. Petersburg Department of Steklov Institute of Mathematics,on
St. Petersburg State University,
Institute for Information Transmission Problems
St.Petersburg, Russia
e-mail: avershik@gmail.com

Nearly Optimal Univariate Polynomial Root-finding: Old and New Algorithms

Victor Y. Pan

keywords: Polynomial root-finding, Subdivision, Sparse polynomials, Root-counting, Functional iterations, Deflation, Real roots
2000 Math. Subject Classification: 65H05, 26C10, 30C15

Extended Abstract

We first review the State of the Art and then outline our progress and state some major research challenges. Further details can be found in arXiv:1805.12042

1. The problem and three celebrated approaches. Univariate polynomial root-finding has been the central problem of mathematics and computational mathematics for four millennia, since Sumerian times (see [5], [10], [11]). Interest to it has been revived due to the advent of modern computers and applications to signal processing, control, financial mathematics, geometric modeling, and computer algebra. The problem remains the subject of intensive research. Hundreds of efficient polynomial root-finders have been proposed, and new ones keep appearing (see [6], [7]).

Two known root-finders are nearly optimal. The algorithm of [9] and [13], proposed in 1995 and extending the previous progress in [15] and [8], first computes numerical factorization of a polynomial into the product of its linear factors and then readily approximate the roots.[1] In the case of inputs of large size the algorithm solves both problems of numerical factorization and root-finding in record low and nearly optimal Boolean time, that is, it approximates all linear factors of a polynomial as well as all its roots, respectively, almost as fast as one can access the input coefficients with the precision required for these tasks.[2] The algorithm, however, is quite involved and has never been implemented.

[1] Numerical polynomial factorization has various important applications to modern computations, besides root-finding, in particular to time series analysis, Wiener filtering, noise variance estimation, co-variance matrix computation, and the study of multichannel systems.

[2] For an input polynomial of degree d the bounds on the required input precision and Boolean time are greater by a factor of d for root-finding than that for numerical factorization.

Recently Becker et al in [1] proposed the second nearly optimal polynomial root-finder, by extending the previous advances of [14] and [12] for the classical subdivision iterations. The algorithm has been implemented in 2018 and promises to become practical, but so far the root-finder of user's choice is the package MPSolve, devised in 2000 [2] and revised in 2014 [3]. It implements Ehrlich's iterations of 1967, rediscovered by Aberth in 1973. Currently subdivision root-finder performs slightly faster than MPSolve of [3] for root-finding in a disc on the complex plain containing a small number of roots but is noticeably inferior for the approximation of all roots of a polynomial.[3]

2. Representation of an input polynomial. The algorithms of [9], [13], and [1] involve the coefficients of an input polynomial $p = p(x)$, relying on its representation in monomial basis:

$$p(x) = \sum_{i=0}^{d} p_i x^i = p_d \prod_{j=1}^{d} (x - x_j), \ p_d \neq 0, \tag{1}$$

where we may have $x_k = x_l$ for $k \neq l$. In contrast Ehrlich's and various other functional root-finding iterations such as Newton's and Weierstrass's can be applied to a more general class of *black box polynomials* – those represented by a black box subroutine for their evaluation, e.g., those represented in Bernstein's bases and sparse polynomials such as Mandelbrot's (cf. [2, Eqn.16]).

3. Our progress. Having reviewed the State of the Art, we significantly accelerate subdivision and Ehrlich's iterations by means of properly combining them with known and novel root-finding techniques. Moreover we extend subdivision iterations to black box polynomials, enabling their dramatic acceleration in the case of sparse input polynomials. Next we itemize our progress.

- We dramatically accelerate *root-counting* for a polynomial in a disc on the complex plain, which is a basic ingredient of subdivision iterations.[4]
- Even stronger we accelerate *exclusion test*: it verifies that a disc contains no roots and is the other key ingredient of subdivision iterations.
- We extend our fast exclusion test to *proximity estimation*, that is, estimation of the distance from a complex point to a closest root of $p(x)$.[5]
- We accelerate subdivision iterations by means of decreasing the number of required exclusion tests,
- We accelerate subdivision iterations by means of deflation of small degree factors whose root sets are well-isolated from the other roots of p.
- We accelerate real polynomial root-finding by means of nontrivially extending all our progress with subdivision iterations.

[3]The computational cost of root-finding in [9], [13], and [1] decreases at least proportionally to the number of roots in a region of interest such as a disc on the complex plain, while MPSolve approximates the roots in such regions almost as slow as all complex roots.
[4]We count m times a root of multiplicity m.
[5]Proximity estimation for $p'(x)$ is critical in path-lifting polynomial root-finders [4].

- Our simple but novel deflation algorithm supports accurate approximation of all roots of a polynomial of extremely high degree.
- We accelerate Ehrlich's iterations by means of incorporation of the Fast Multipole Method (*FMM*).

4. Further details of root-counting and exclusion test. The previous acceleration of the known root-counting in [1] was justly claimed to be their major algorithmic novelty versus their immediate predecessors of [14] and [12], but we enhance that progress: our root-counting is performed at a smaller computational cost under milder assumptions about the isolation of the boundary circle of a disc from the roots of $p(x)$. We can counter the decrease of the root isolation by a factor of f by means of increasing the number of evaluation points just by a factor of $\log(f)$. Compared to the common recipe of root-squaring this has similar arithmetic cost but avoids coefficient growth. Even if we do not know how well the boundary circle is isolated from the roots we just recursively double the number of evaluation points until correctness of the root count is confirmed. For heuristic confirmation we can stop where the computed root count approximates an integer, and we propose additional verification recipes. The same algorithm enables fast exclusion test for a fixed disc, but by perfroming some simple low cost computations we decrease the need for exclusion tests.

5. Three major research challenges. We hope that our work will motivate further research effort towards synergistic combination of some efficient techniques, both well- and less-known for polynomial root-finding.

Devising practical and nearly optimal algorithms for numerical factorization of a polynomial is still a challenge – both Ehrlich's and subdivision iterations are slower for that task by at least a factor of d than the nearly optimal solution in [9] and [13], which is quite involved and not practically competitive.

Our root-finders accelerate the known nearly optimal ones and promise to become user's choice. Their implementation, testing and refinement are major challenges. This work, just initiated, already shows that our improvement of the known algorithms *is for real*.

References

[1] Becker, R., Sagraloff, M., Sharma, V., Yap, C.: A near-optimal subdivision algorithm for complex root isolation based on the Pellet test and Newton iteration. J. Symb. Comput. **86**, 51–96 (2018)

[2] Bini, D.A., Fiorentino, G.: Design, analysis, and implementation of a multiprecision polynomial rootfinder. Numer. Algs. **23**, 127–173 (2000)

[3] Bini, D.A., Robol, L.: Solving secular and polynomial equations: a multiprecision algorithm. J. Comput. Appl. Math. **272**, 276–292 (2014)

[4] Blum, L., Cucker, F., Shub, M., Smale, S.: Complexity and Real Computation. Springer Verlag (1998).

[5] Boyer, C.A.: A History of Mathematics. Wiley, New York (1991)

[6] McNamee, J.M.: Numerical Methods for Roots of Polynomials, Part I, XIX+354 pages. Elsevier (2007) ISBN13: 9780444527295

[7] McNamee, J.M., Pan, V.Y.: Numerical Methods for Roots of Polynomials, Part II, XXI+728 pages. Elsevier (2013) ISBN: 9780444527301

[8] Neff, C.A., Reif, J.H.: An $o(n^{1+\epsilon})$ algorithm for the complex root problem. In: Proc. 35th Ann. IEEE Symp. on Foundations of Comput. Sci. (FOCS '94), pp. 540–547. doi: 10.1109/SFCS.1994.365737

[9] Pan, V.Y.: Optimal (up to polylog factors) sequential and parallel algorithms for approximating complex polynomial zeros. In: Proc. 27th Ann. ACM Symp. on Theory of Computing (STOC'95), 741–750. ACM Press, New York (1995) doi: 10.1145/225058.225292

[10] Pan, V.Y.: Solving a polynomial equation: Some history and recent progress. SIAM Review **39**(2), 187–220 (1997)

[11] Pan, V.Y.: Solving polynomials with computers. American Scientist **86**, 62–69 (1998) doi: 10.1511/1998.1.62

[12] Pan, V.Y.: Approximation of complex polynomial zeros: modified quadtree (Weyl's) construction and improved Newton's iteration. J. Complexity **16**(1), 213–264 (2000) doi: 10.1006/jcom.1999.

[13] Pan, V.Y.: Univariate polynomials: nearly optimal algorithms for factorization and rootfinding. J. Symb. Comput. **33**(5), 701–733 (2002) doi: 10.1006/jsco.2002.0531

[14] Renegar, J.: On the worst-case arithmetic complexity of approximating zeros of polynomials. J. Complex. **3**(2), 90–113 (1987)

[15] Schönhage, A.: The fundamental theorem of algebra in terms of computational complexity. University of Tübingen, Germany (1982)

Victor Y. Pan
Department of Computer Science
Lehman College of the City University of New York

e-mail: `victor.pan@lehman.cuny.edu`

Compact Monomial Involutive Bases

Vladimir P. Gerdt and Yury A. Blinkov

Abstract. Based on the minimal Gröbner basis G of a monomial ideal $\mathcal{I}$ in the commutative polynomial ring $\mathcal{K}[x_1, x_2, \ldots, x_n]$ over a field $\mathcal{K}$ and a total monomial ordering $\succ$, we define another monomial ordering $\succ_G$ such that pairwise involutive partition of variables $\{x_1, \ldots, x_n\}$ for monomials in $\mathcal{I}$ generated by $\succ_G$ yields more compact involutive basis than that generated by $\succ$. In particular, for $\succ_{\text{alex}}$, the antigraded lexicographic ordering, the involutive basis for $\succ_{\text{alex}_G}$ and $n \gg 1$ is much more compact then involutive basis for $\succ_{\text{alex}}$. We illustrate this by computer experiments.

The notion of involutive monomial division introduced in our paper [1] is a cornerstone of theory of involutive bases and their algorithmic construction. The basic idea behind this notion goes back to Janet [2] and consists in a proper partition of variables for every element in a finite monomial set into the two subsets called multiplicative and nonmultiplicative. Given a polynomial set and an admissible monomial order, the partition of variables is defined in terms of the leading monomial set. Each such partition generates a monomial division [3] called involutive, if it is defined for an arbitrary monomial set and satisfies the axioms given in Definition 1 [1]. For more definitions and proofs see [3] and book [4]).

Definition 1. [9] An *involutive division* $\mathcal{L}$ is defined on $\mathcal{M}$ if for any nonempty set $U \subset \mathcal{M}$ and for any $u \in U$ a subset $M_{\mathcal{L}}(u, U) \subseteq X$ is defined that generates submonoid $\mathcal{L}(u, U) \subset \mathcal{M}$ of power products in $M_{\mathcal{L}}(u, U)$ and the following holds

1. $v \in U \ \wedge \ u\mathcal{L}(u, U) \cap v\mathcal{L}(v, U) \neq \emptyset \Longrightarrow u \in v\mathcal{L}(v, U) \ \vee \ v \in u\mathcal{L}(u, U)$,
2. $v \in U \ \wedge \ v \in u\mathcal{L}(u, U) \Longrightarrow \mathcal{L}(v, U) \subseteq \mathcal{L}(u, U)$ (transitivity),
3. $u \in V \wedge \ V \subseteq U \Longrightarrow \mathcal{L}(u, U) \subseteq \mathcal{L}(u, V)$ (filter axiom).

Variables in $M_{\mathcal{L}}(u, U)$ are *$\mathcal{L}$-multiplicative* for u and those in $NM_{\mathcal{L}}(u, U) = X \setminus M_{\mathcal{L}}(u, U)$ are *$\mathcal{L}$-nonmultiplicative*. If $w \in u\mathcal{L}(u, U)$, then u is *$\mathcal{L}$-(involutive) divisor* of w (denotation: $u \mid_{\mathcal{L}} w$).

In an involutive algorithm the nonmultiplicative variables of a polynomial are used for its prolongation, that is, for the multiplication by these variables, whereas the multiplicative variables of other polynomials in the set are used for reduction of the nonmultiplicative prolongations. An involutive basis is a polynomial set such that all its nonmultiplicative prolongations are multiplicatively reducible to

zero. If an involutive algorithm terminates it outputs an involutive basis which is a Gröbner basis of the special structure determined by properties of underlying involutive division. In our approach, a reduced Gröbner basis is always a well defined subset of the involutive basis and can be extracted from the last one without any extra computation [3].

In the talk we consider pair divisions introduced in [5] which are pairwise generated by total monomial orderings and studied in [6] - [9]. They are called $\prec$ $-$*divisions*, where $\prec$ is a total monomial ordering compatible with multiplication, i.e. $a \succ b \to m \cdot a \succ m \cdot b$ for all m. In [9], from this class of divisions we singled out the $\succ_{\rm alex}$-division generated the antigraded lexicographic ordering $\succ_{\rm alex}$ and shown, by computer experimentation, that in the vast majority of cases $\succ_{\rm alex}$-division yields much more compact monomial involutive bases than Janet division which is pairwise generated by the pure lexicographic ordering $\succ_{\rm lex}$.

Definition 2. [9]. Let U be a finite set of monomials in $\mathcal{K}[x_1, \ldots, x_n]$, $\prec$ a total monomial ordering compatible with multiplication and σ a permutation of variables $x_1, \ldots, x_n$. Then a (pairwise) $\succ$-*division* is defined as

$$(\forall u \in U) \ [\ NM_{\mathcal{L}}(u, U) = \bigcup_{v \in U \setminus \{u\}} NM_{\mathcal{L}}(u, \{u, v\}) \ ] , \tag{1}$$

where

$$NM_{\mathcal{L}}(u, \{u, v\}) := \begin{cases} \text{if } u \succ v \text{ or } (u \prec v \wedge v \mid u) \text{ then } \emptyset \\ \text{else } \{x_{\sigma(i)}\}, \ i = \min\{j \mid \deg_{\sigma(j)}(u) < \deg_{\sigma(j)}(v)\} . \end{cases} \tag{2}$$

Definition 3. For a monomial $u \in U$ and a total monomial ordering $\succ$, the element $v \in G$ where $G(U)$ is the reduced Gröbner basis of U is said to be an *ancestor* of u in U w.r.t. $\succ$ (denotation: $v = \mathrm{anc}_{\succ}(u)$) if

$$v := \max_{\succ} \{ \, w \in G(U) \ \mid \ w \mid u \, \} .$$

Given a $\prec$ $-$division defined in (1)-(2) and a finite monomial set U, one can further compactify its involutive basis if to define the total ordering $\succ_G$ of elements in the monomial ideal $\mathcal{I}$ generated by U as follows

$$u \succ_{{\rm alex}_G} v \ \textbf{if} \ \mathrm{anc}_{\succ}(u) \succ \mathrm{anc}_{\succ}(v) \ \textbf{or} \ (\mathrm{anc}_{\succ}(u) = \mathrm{anc}_{\succ}(v) \ \textbf{and} \ u \succ w) \tag{3}$$

and to use Eqs. (1)-(2) for the involutive completion of G.

Another possibility of the compactification of $\prec$ $-$divisions is to use the total orderings

$$\begin{aligned} u \succ_G v \qquad & \textbf{if} \ \deg(\mathrm{anc}_{\succ}(u)) \succ \deg(\mathrm{anc}_{\succ}(v)) \\ & \textbf{or} \ (\deg(\mathrm{anc}_{\succ}(u)) = \deg(\mathrm{anc}_{\succ}(v)) \ \textbf{and} \ u \succ w) . \end{aligned} \tag{4}$$

For several pairwise divisions, we generated randomly monomial sets for different numbers of variables and averaged the cardinalities of their involutive bases over the permutations σ of variables occurring in Eq. (2). Clearly, Gröbner bases for Eqs. (3)-(4) are much more compact and computed much faster then those for $\prec$ $-$divisions.

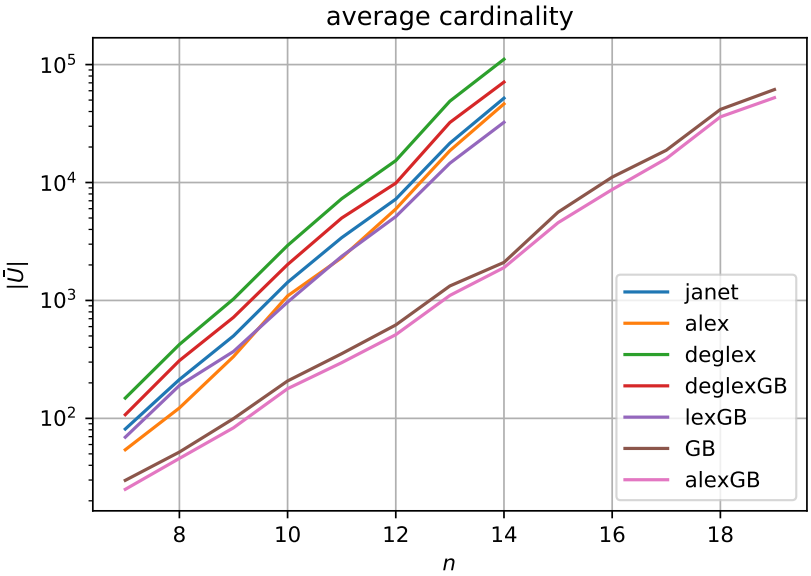


FIGURE 1. Cardinality growth with the number of variables

References

[1] V.P.Gerdt and Yu.A.Blinkov. *Involutive Bases of Polynomial Ideals*. Mathematics and Computers in Simulation, 45, 519–542, 1998; *Minimal Involutive Bases*, ibid. 543–560.

[2] M.Janet. *Leçons sur les Systèmes d'Equations aux Dérivées Partielles*. Cahiers Scientifiques, IV, Gauthier-Villars, Paris, 1929.

[3] V.P.Gerdt. *Involutive Algorithms for Computing Gröbner Bases*. Computational Commutative and Non-Commutative Algebraic Geometry. IOS Press, Amsterdam, 2005, pp.199–225.

[4] W.M.Seiler. *Involution: The Formal Theory of Differential Equations and its Applications in Computer Algebra*. Algorithms and Computation in Mathematics, 24, Springer, 2010.

[5] V.P.Gerdt. *Involutive Division Technique: Some Generalizations and Optimizations*. Journal of Mathematical Sciences, 108, 6, 1034–1051, 2002.

[6] A.S.Semenov. *On Connection Between Constructive Involutive Divisions and Monomial Orderings*. In: "Computer Algebra in Scientific Computing CASC 2006", LNCS 4194, Springer-Verlag, Berlin, 2007, pp. 261-278.

[7] A.S.Semenov. *Constructivity of Involutive Divisions*. Programming and Computer Software, 32, 2, 96–102, 2007.

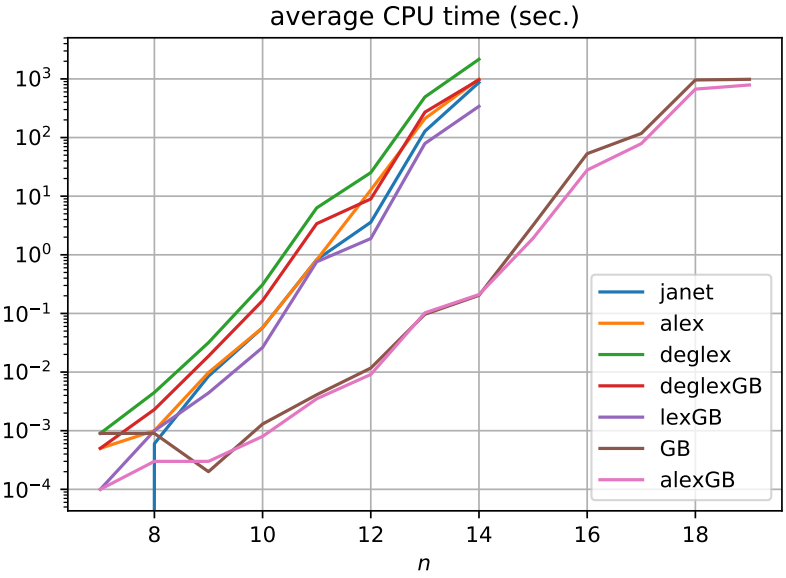


FIGURE 2. CPU time growth with the number of variables

[8] A.S.Semenov and P.A.Zyuzikov. *Involutive Divisions and Monomial Orderings*. Programming and Computer Software, 33, 3, 139–146, 2007; *Involutive Divisions and Monomial Orderings: Part II*. Programming and Computer Software, 34, 2, 107–111.

[9] V.P.Gerdt and Yu.A.Blinkov. *Involutive Division Generated by an Antigraded Monomial Ordering*. In: "Computer Algebra in Scientific Computing / CASC 2011", V.P.Gerdt, W.Koepff, E.W.Mayr, E.V.Vorozhtsov (Eds.), LNCS 6885, Springer-Verlag, Berlin, 2011, pp.158–174.

Vladimir P. Gerdt
Laboratory of Information Technologies, Joint Institute for Nuclear Research
141980 Dubna, Russia
e-mail: gerdt@jinr.ru

Yury A. Blinkov
Department of Mathematics and Mechanics, Saratov State University
410071 Saratov, Russia
e-mail: blinkovua@info.sgu.ru

Parallel Computation of Involutive and Gröbner Bases Using the Tableau Representation of Polynomials

Denis A. Yanovich

Abstract. For the work with polynomials such data representations as lists of terms, geobuckets, and heaps are usually used. In this talk an attempt for using new representation of polynomials for parallel computing involutive and Gröbner bases of systems of nonlinear polynomial equations will be made. Using the proposed data structure makes it possible to compute complex and memory-hungry tasks on the cluster of computers utilizing MPI technologie. In-depth explanation of the new table-based data structure and various benchmarks of parallel and sequental computations will be presented.

Denis A. Yanovich
Laboratory of Information Technologies
Joint Institute for Nuclear Research
Dubna, Russia
e-mail: yan@jinr.ru

International Conference
Polynomial Computer Algebra ‘2020;
St. Petersburg, October 12–17, 2020,
Euler International Mathematical Institute,
Ed. by N. N. Vassiliev

VVM Pubishing, 2020
Order 1298
Circulation is 200 copies