BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 3–10
ISSN 1024–7696

New Form of the Hidden Logarithm Problem and its Algebraic Support

D. N. Moldovyan

Abstract. The paper introduces a new form of the hidden discrete logarithm problem defined over finite non-commutative associative algebras containing two-sided global unit and sets of local left-sided and right-sided units. The proposed form is characterized in using a new mechanism for masking the finite cyclic group in which the base exponentiation operation is performed. Local units act in frame of subsets of non-invertible vectors and are used as elements of the private key in the proposed post-quantum digital signature scheme. A new 4-dimensional algebra is introduced as algebraic support of the proposed cryptoscheme. Formulas describing units of different types are derived.

Mathematics subject classification: 94A60, 16Z05, 14G50, 11T71, 16S50.
Keywords and phrases: Finite associative algebra, non-commutative algebra, right-sided unit, left-sided unit, local units, discrete logarithm problem, hidden logarithm problem, post-quantum cryptography, digital signature.

1 Introduction

One of current challenges in the area of cryptography relates to the development of the post-quantum public-key cryptoschemes suitable for wide practical application [1, 2]. A cryptographic scheme is called post-quantum if it can resist attacks performed with using hypothetic quantum computers for which there are known algorithms solving the discrete logarithm problem (DLP) and factorization problem in polynomial time [3, 4]. The post-quantum public-key cryptoschemes should be based on computationally difficult problems having superpolynomial complexity when solving them on quantum computers.

Earlier the hidden DLP (HDLP) was proposed as the base primitive for post-quantum public key-agreement protocols [5, 6]. That form of the HDLP has been defined in a finite non-commutative group Γ as follows. Suppose G is a generator of some finite cyclic group having prime order of sufficiently large size. Then the DLP is set as finding a natural number $x < q$ satisfying the equation $Y = G^x$, where the values G and Y are known. In the HDLP [5] the value Y is masked, i. e. instead of the value Y other value Y' is given that is an element of another cyclic group representing a subset of elements of the group Γ.

Recently [7,8] new forms of the HDLP have been proposed, in which both values G and Y are masked in some given values G' and Y' contained in two different finite cyclic groups representing subsets of a finite non-commutative associative algebra

(FNAA). The last forms have been used as the base primitive of the proposed post-quantum digital signature schemes. The last forms of the HDLP have been set in the FNAAs containing large sets of the global single-sided units. Homomorphism maps have been used as the masking mechanism.

The present paper introduces a new form of the HDLP with masking the both values Y and G, which suits well for designing the signature schemes. The proposed form is characterized in using two simple masking mechanisms each of which is performed as one multiplication operation, namely, as the multiplication by the right-sided unit used as the left operand or the multiplication by the left-sided unit used as the right operand. The next section of the paper describes a new 4-dimensional FNAA as an appropriate algebraic support of the proposed form of the HDLP and the intoduced postquantum signature scheme.

2 A 4-dimensional FNAA used as algebraic support

A finite m-dimensional algebra represents a vector space defined over a finite field, for example, over the ground finite field $GF(p)$, in which the vector multiplication operation (distributive relative to the addition operation) is additionally defined. If the multiplication operation (denoted as $\circ$) is non-commutative and associative, then the algebra is FNAA. Suppose $\mathbf{e}_0$, $\mathbf{e}_1$, ... $\mathbf{e}_{m-1}$ are the basis vectors. A vector V is denoted in the following two forms: $A = (a_0, a_1, \ldots, a_{m-1})$ and $A = a_0\mathbf{e}_0 + a_1\mathbf{e}_1 + \cdots + a_{m-1}\mathbf{e}_{m-1}$, where $a_0, a_1, \ldots, a_{m-1} \in GF(p)$.

Usually the multiplication operation of two vectors A and $B = \sum_{i=0}^{m-1} b_i\mathbf{e}_i$ is defined with the formula

$$A \circ B = \sum_{j=0}^{m-1} \sum_{i=0}^{m-1} a_i b_j (\mathbf{e}_i \circ \mathbf{e}_j),$$

in which products of different pairs of basis vectors $\mathbf{e}_i \circ \mathbf{e}_j$ are to be substituted by a single-component vector indicated in the so called basis vector multiplication table (BVMT). Every cell of the BVMT contains a single-component vector $\lambda\mathbf{e}_k$, where $\lambda \in GF(p)$ is called a structural coefficient. If $\lambda = 1$, then the content of the cell is denoted as $\mathbf{e}_k$. One usually assumes that the left operand $\mathbf{e}_i$ defines the row and the right one $\mathbf{e}_j$ defines the column. The intersection of the ith row and jth column defines the cell indicating the value of the product $\mathbf{e}_i \circ \mathbf{e}_j$.

2.1 The BVMT and the invertibility condition

In the case of using the BVMT shown in Table 1 the vector equation defining the value of left-sided units and having the form $X \circ A = A$, where the vector $X = (x_0, x_1, x_2, x_3)$ is an unknown value, can be reduced to the following system of

Table 1. The BVMT setting the 4-dimensional FNAA ($\lambda\sigma \neq 1$).

$\circ$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_0$	$\lambda\mathbf{e}_0$	$\lambda\mathbf{e}_1$	$\mathbf{e}_0$	$\mathbf{e}_1$
$\mathbf{e}_1$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\sigma\mathbf{e}_0$	$\sigma\mathbf{e}_1$
$\mathbf{e}_2$	$\lambda\mathbf{e}_2$	$\lambda\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\sigma\mathbf{e}_2$	$\sigma\mathbf{e}_3$

four linear equations with the unknowns (x_0, x_1, x_2, x_3) :

$$\begin{cases} (\lambda x_0 + x_1)\, a_0 + (x_0 + \sigma x_1)\, a_2 = a_0; \\ (\lambda x_2 + x_3)\, a_0 + (x_2 + \sigma x_3)\, a_2 = a_2; \\ (\lambda x_0 + x_1)\, a_1 + (x_0 + \sigma x_1)\, a_3 = a_1; \\ (\lambda x_2 + x_3)\, a_1 + (x_2 + \sigma x_3)\, a_3 = a_3. \end{cases} \tag{1}$$

Perfoming the following substitution of the variables $u_1 = (\lambda x_0 + x_1)$; $u_2 = (x_0 + \sigma x_1)$; $u_3 = (\lambda x_2 + x_3)$; and $u_4 = (x_2 + \sigma x_3)$ one can get the solution $u_1 = 1$; $u_2 = 0$; $u_3 = 0$; and $u_4 = 1$ that is independent of the value A. From the last solution we get the following two independent systems of two linear equations with the unknowns (x_0, x_1) and (x_2, x_3) respectively:

$$\begin{cases} \lambda x_0 + x_1 = 1; \\ x_0 + \sigma x_1 = 0; \end{cases} \tag{2}$$

$$\begin{cases} \lambda x_2 + x_3 = 0; \\ x_2 + \sigma x_3 = 1. \end{cases} \tag{3}$$

Solving the last two systems we get the following value of the global left-sided unit (it is called global since it acts on all elements of the considered FNAA):

$$E_L = \left(\frac{\sigma}{\lambda\sigma - 1}, \frac{1}{1 - \lambda\sigma}, \frac{1}{1 - \lambda\sigma}, \frac{\lambda}{\lambda\sigma - 1} \right). \tag{4}$$

The vector equation $A \circ X = A$ defining the value of the right-sided units reduces to the following system with the unknowns (x_0, x_1, x_2, x_3) :

$$\begin{cases} (\lambda x_0 + x_2)\, a_0 + (x_0 + \sigma x_2)\, a_1 = a_0; \\ (\lambda x_1 + x_3)\, a_0 + (x_1 + \sigma x_3)\, a_1 = a_1; \\ (\lambda x_0 + x_2)\, a_2 + (x_0 + \sigma x_2)\, a_3 = a_2; \\ (\lambda x_1 + x_3)\, a_2 + (x_1 + \sigma x_3)\, a_3 = a_3. \end{cases} \tag{5}$$

Perfoming the following substitution of the variables $z_1 = (\lambda x_0 + x_2)$; $z_2 = (x_0 + \sigma x_2)$; $z_3 = (\lambda x_1 + x_3)$; and $z_4 = (x_1 + \sigma x_3)$ one can get the solution $z_1 = 1$;

$z_2 = 0$; $z_3 = 0$; and $z_4 = 1$ that is independent of the value A. Computation of the inverse variable substitution gives the following formula for the global right-sided unit

$$E_R = \left(\frac{\sigma}{\lambda\sigma - 1}, \frac{1}{1 - \lambda\sigma}, \frac{1}{1 - \lambda\sigma}, \frac{\lambda}{\lambda\sigma - 1} \right). \tag{6}$$

Comparison of the formulas (4) and (6) shows that the considered algebra contains the unique global two-sided unit $E = E_L = E_R = (e_0, e_1, e_2, e_3)$.

If for a vector A the equation $X \circ A = E$ has a solution, then the vector A is called invertible. The last vector equation reduces to the following two independent systems of two linear equations with the unknowns (x_0, x_1) and (x_2, x_3) correspondingly:

$$\begin{cases} (\lambda a_0 + a_2)\, x_0 + (a_0 + \sigma a_2)\, x_1 = e_0; \\ (\lambda a_1 + a_3)\, x_0 + (a_1 + \sigma a_3)\, x_1 = e_1; \end{cases} \tag{7}$$

$$\begin{cases} (\lambda a_0 + a_2)\, x_2 + (a_0 + \sigma a_2)\, x_3 = e_2; \\ (\lambda a_1 + a_3)\, x_2 + (a_1 + \sigma a_3)\, x_3 = e_3. \end{cases} \tag{8}$$

Each of the last two systems has the same determinant Δ_A :

$$\Delta_A = (\lambda a_0 + a_2)(a_1 + \sigma a_3) - (\lambda a_1 + a_3)(a_0 + \sigma a_2) = (1 - \lambda\sigma)(a_1 a_2 - a_0 a_3) \tag{9}$$

If $\Delta_A \neq 0$, then the vector A is invertible, i. e. we have the following invertibility condition:

$$a_1 a_2 \neq a_0 a_3. \tag{10}$$

2.2 Local units connected with non-invertible vectors

If coordinates of a vector $G = (g_0, g_1, g_2, g_3)$ satisfy the condition $g_1 g_2 = g_0 g_3$, then the vector G is non-invertible. However, some non-invertible vectors can be locally invertible. Such non-invertible vectors are generators of finite cyclic groups contained in the considered FNAA. Besides, to some fixed locally invertible vector G a large set of local left-sided units and a large set of right-sided units may relate. Each of the latter sets contains invertible and non-invertible 4-dimensional vectors.

To derive the formula describing local left-sided units one should consider the solutions of the vector equation $X \circ G = G$ that reduces to the following two independent systems:

$$\begin{cases} (\lambda g_0 + g_2)\, x_0 + (g_0 + \sigma g_2)\, x_1 = g_0; \\ (\lambda g_1 + g_3)\, x_0 + (g_1 + \sigma g_3)\, x_1 = g_1; \end{cases} \tag{11}$$

$$\begin{cases} (\lambda g_0 + g_2)\, x_2 + (g_0 + \sigma g_2)\, x_3 = g_2; \\ (\lambda g_1 + g_3)\, x_2 + (g_1 + \sigma g_3)\, x_3 = g_3. \end{cases} \tag{12}$$

The determinant of each of the latter systems is equal to zero. The auxiliary determinants of the system (11) are

$$\Delta_0 = g_0 (g_1 + \sigma g_3) - g_1 (g_0 + \sigma g_2) = \sigma (g_0 g_3 - g_1 g_2) = 0.$$

$$\Delta_1 = g_1\left(\lambda g_0 + g_2\right) - g_0\left(\lambda g_1 + g_3\right) = g_1 g_2 - g_0 g_3 = 0.$$

For the system (11) we have p solutions described by the formula $x_1 = \frac{g_0 - (\lambda g_0 + g_2)x_0}{g_0 + \sigma g_2}$, where $x_0 = 0, 1, \ldots, p-1$, if $g_0 + \sigma g_2 \neq 0$, or by the formula $x_0 = \frac{g_0 - (g_0 + \sigma g_2)x_1}{\lambda g_0 + g_2}$, where $x_1 = 0, 1, \ldots, p-1$, if $\lambda g_0 + g_2 \neq 0$.

The auxiliary determinants of the system (12) are also equal to zero:

$$\Delta_2 = g_2\left(g_1 + \sigma g_3\right) - g_3\left(g_0 + \sigma g_2\right) = g_1 g_2 - g_0 g_3 = 0.$$

$$\Delta_3 = g_3\left(\lambda g_0 + g_2\right) - g_2\left(\lambda g_1 + g_3\right) = \lambda\left(g_0 g_3 - g_1 g_2\right) = 0.$$

For the system (12) we have p solutions described by the formula $x_3 = \frac{g_2 - (\lambda g_0 + g_2)x_2}{g_0 + \sigma g_2}$, where $x_2 = 0, 1, \ldots, p-1$, if $g_0 + \sigma g_2 \neq 0$, or by the formula $x_2 = \frac{g_2 - (g_0 + \sigma g_2)x_3}{\lambda g_0 + g_2}$, where $x_3 = 0, 1, \ldots, p-1$, if $\lambda g_0 + g_2 \neq 0$.

Thus, for the non-invetible vector G coordinates of which satisfy the condition $g_0 + \sigma g_2 \neq 0$ there exist p^2 different left-sided units $L = (l_0, l_1, l_2, l_3)$ described by the formula

$$L = \left(x_0, \frac{g_0 - \left(\lambda g_0 + g_2\right)x_0}{g_0 + \sigma g_2}, x_2, \frac{g_2 - \left(\lambda g_0 + g_2\right)x_2}{g_0 + \sigma g_2}\right), \tag{13}$$

where $x_0, x_2 = 0, 1, \ldots, p-1$. One can easily show that the set (13) contains $p^2 - p$ invertible and p non-invertible elements of the considered FNAA. The subset of the local left-sided units L' that are non-invertible vectors of the considered 4-dimensinal FNAA is described as follows (for the case $g_0 \neq 0$):

$$L' = \left(x_0, \frac{g_0 - \left(\lambda g_0 + g_2\right)x_0}{g_0 + \sigma g_2}, \frac{g_2}{g_0}x_0, \frac{g_0 g_2 - \left(\lambda g_0 + g_2\right)g_2 x_0}{g_0^2 + \sigma g_0 g_2}\right), \tag{14}$$

where $x_0 = 0, 1, \ldots, p-1$.

The formula describing the set of the local right-sided units relating to the non-invertible vector G can be derived from the vector equation $G \circ X = X$ that reduces to the following two independent systems of two linear equations

$$\begin{cases} \left(\lambda g_0 + g_1\right)x_0 + \left(g_0 + \sigma g_1\right)x_2 = g_0; \\ \left(\lambda g_2 + g_3\right)x_0 + \left(g_2 + \sigma g_3\right)x_2 = g_2; \end{cases} \tag{15}$$

$$\begin{cases} \left(\lambda g_0 + g_1\right)x_1 + \left(g_0 + \sigma g_1\right)x_3 = g_1; \\ \left(\lambda g_2 + g_3\right)x_1 + \left(g_2 + \sigma g_3\right)x_3 = g_3. \end{cases} \tag{16}$$

The main determinant of each of the systems (15) and (16) is equal to zero. The auxiliary determinants of the system (15) are

$$\Delta_0 = g_0\left(g_2 + \sigma g_3\right) - g_2\left(g_0 + \sigma g_1\right) = \sigma\left(g_0 g_3 - g_1 g_2\right) = 0.$$

$$\Delta_2 = g_2\left(\lambda g_0 + g_1\right) - g_0\left(\lambda g_2 + g_3\right) = g_1 g_2 - g_0 g_3 = 0.$$

For the system (15) we have p solutions described by the formula $x_2 = \frac{g_0-(\lambda g_0+g_1)x_0}{g_0+\sigma g_1}$, where $x_0 = 0, 1, \ldots, p-1$, if $g_0 + \sigma g_1 \neq 0$, or by the formula $x_0 = \frac{g_0-(g_0+\sigma g_1)x_2}{\lambda g_0+g_1}$, where $x_1 = 0, 1, \ldots, p-1$, if $\lambda g_0 + g_1 \neq 0$.

The auxiliary determinants of the system (16) are also equal to zero:

$$\Delta_1 = g_1 (g_2 + \sigma g_3) - g_3 (g_0 + \sigma g_1) = g_1 g_2 - g_0 g_3 = 0.$$

$$\Delta_3 = g_3 (\lambda g_0 + g_1) - g_1 (\lambda g_2 + g_3) = \lambda (g_0 g_3 - g_1 g_2) = 0.$$

For the system (16) we have p solutions described by the formula $x_3 = \frac{g_1-(\lambda g_0+g_1)x_1}{g_0+\sigma g_1}$, where $x_2 = 0, 1, \ldots, p-1$, if $g_0 + \sigma g_1 \neq 0$, or by the formula $x_1 = \frac{g_1-(g_0+\sigma g_1)x_3}{\lambda g_0+g_1}$, where $x_3 = 0, 1, \ldots, p-1$, if $\lambda g_0 + g_1 \neq 0$.

Thus, for the non-invetible vector G coordinates of which satisfy the condition $g_0 + \sigma g_1 \neq 0$ there exist p^2 different right-sided units $R = (r_0, r_1, r_2, r_3)$ described by the formula

$$R = \left(x_0, x_1, \frac{g_0 - (\lambda g_0 + g_1) x_0}{g_0 + \sigma g_1}, \frac{g_1 - (\lambda g_0 + g_1) x_1}{g_0 + \sigma g_1} \right), \tag{17}$$

where $x_0, x_1 = 0, 1, \ldots, p-1$. One can easily show that the set (17) contains $p^2 - p$ invertible and p non-invertible elements of the considered FNAA.

The subset of the local right-sided units R' that are non-invertible vectors of the considered 4-dimensinal FNAA is described as follows (for the case $g_0 \neq 0$):

$$R' = \left(x_0, \frac{g_1}{g_0} x_0, \frac{g_0 - (\lambda g_0 + g_1) x_0}{g_0 + \sigma g_1}, \frac{g_0 g_1 - (\lambda g_0 + g_1) g_1 x_0}{g_0^2 + \sigma g_0 g_1} \right), \tag{18}$$

where $x_0 = 0, 1, \ldots, p-1$.

Let us consider the non-invertible vector G satisfying the conditions $g_0 + \sigma g_2 \neq 0$ and $g_0 + \sigma g_1 \neq 0$. Only one non-invertible vector E' is contained simultaneously in the sets (14) and (18). The value E' can be computed substituting the value

$$x_0 = \frac{g_0^2}{\lambda g_0^2 + g_0 g_1 + g_0 g_2 + \sigma g_1 g_2} \tag{19}$$

in (14) or in (18).

The vector E' is the unit of the cyclic group generated by the vector G. For example, for the fixed values $p = 2q + 1$, where $q = 30894397013$ is a prime, $\lambda = 1234567$, $\sigma = 809$, and $G = (160, 800, 400, 2000)$ computation of the value E' using the formulas (18) and (19) gives the same result as computation of the value

$$G^{p-1} = (52415881640, 14924232092, 7462116046, 37310580230) = E'.$$

3 The proposed form of the HDLP and post-quantum signature scheme

Suppose there are given the 512-bit prime $p = 2q + 1$, where q is also prime, the structural coefficients λ and σ such that $\lambda\sigma \neq 1$. Then one can generate the public key in the form of two vectors Y' and G' as follows:

1. Select at random the non-invertible vector G that is a generator of some finite cyclic group having the order equal to q.

2. Generate two random natural numbers x_0, x_2 and, using the formula (13), compute the local left-sided unit L.

3. Generate two random natural numbers x_0, x_1 and, using the formula (17), compute the local right-sided unit R.

4. Compute the vector $G' = G \circ L$.

5. Generate a random natural number x and compute the vector $Y' = R \circ G^x$.

The private key connected with the public key Y', G' represents the values G, L, R, and x. Finding the private key from the public key represent the proposed form of the HDLP that is put into the base of the following digital signature scheme.

Generation of the signature (v, s) to the electronic document M is to be performed as follows:

1. Select a random integer $k < q$ and compute the vector $U = R \circ G^k \circ L$.

2. Using some specified hash-function F_h compute the hash value v from the document M to which the vector U is cocatenated: $v = F_h(M, U)$. Then compute the value $s = k - xv \bmod q$.

Signature verification procedure is executed as follows:

1. Compute the vector $U^? = Y'^v \circ G'^s$ and the value $v^? = F_h(M, U^?)$.

2. If $e^? = e$, then the signature is accepted as genuine. Otherwise it is rejected.

The signature scheme performs correctly due to the commutativity of the exponentiation operation G^t and the left (right) multiplication by a right-sided (left-sided) unit $R \circ G$ $(G \circ L)$: $\quad (R \circ G)^t = R \circ G^t; \quad (G \circ L)^t = G^t \circ L$.

Correctness proof of the signature scheme is as follows:

$$U^? = Y'^v \circ G'^s = R \circ G^{xv} \circ G \circ L^s = R \circ G^{xv} \circ G^s \circ L = R \circ G^{xv} \circ G^{k-xv} \circ L = \\ = R \circ G^k \circ L = U \Rightarrow F_h(M, U^?) = F_h(M, U) \Rightarrow v^? = v. \tag{20}$$

Thus, the correctly computed signature (v, s) passes the verification procedure as genuine signature.

4 Conclusion

A new form of the HDLP and a post-quantum signature scheme on its base have been introduced. A new 4-dimensional FNAA with two-sided global unit has been considered as algebraic support of the introduced HDLP. The proposed design of the signature scheme can be potentially implemented using different algebraic supports, for example, finite algebra of quaternions and 6-dimensional FNAA described in [9].

Another direction of an independent research can be attributed to combining the masking mechanism of the proposed form of the HDLP with the masking mechanisms described in [8, 10]

References

[1] First NIST standardization conference - April 11–13, 2018.

http://prometheuscrypt.gforge.inria.fr/2018-04- 18.pqc2018.html

[2] *Post-Quantum Cryptography. 9th International Conference, PQCrypto 2018 Proceedings*. Fort Lauderdale, FL, USA, April 9-11, 2018. Lecture Notes in Computer Science. Springer Verlag, 2018, **10786**.

[3] Yan S. Y. *Quantum Computational Number Theory*. Springer, 2015, 252 p.

[4] Yan S. Y. *Quantum Attacks on Public-Key Cryptosystems*. Springer, 2014, 207 p.

[5] Moldovyan D. N. *Non-Commutative Finite Groups as Primitive of Public-Key Cryptoschemes*. Quasigroups and Related Systems, 2010, **18**, No. 2, 165–176.

[6] Kuzmin A. S., Markov V. T., Mikhalev A. A., Mikhalev A. V., Nechaev A. A. *Cryptographic Algorithms on Groups and Algebras*. Journal of Mathematical Sciences, 2017, **223**, No. 5, 629–641.

[7] Moldovyan A. A., Moldovyan N. A. *Post-quantum signature algorithms based on the hidden discrete logarithm problem*. Computer Science Journal of Moldova, 2018, **26**, No. 3(78), 301–313.

[8] Moldovyan N. A., Moldovyan A. A. *Finite Non-commutative Associative Algebras as carriers of Hidden Discrete Logarithm Problem*. Bulletin of the South Ural State University. Ser. Mathematical Modelling, Programming & Computer Software (Bulletin SUSU MMCS), 2019, **12**, No. 1, 66–81.

[9] Moldovyan N. A. *Unified Method for Defining Finite Associative Algebras of Arbitrary Even Dimensions*, Quasigroups and Related Systems, 2018, **26**, No. 2, 263–270.

[10] Moldovyan N. A. *Finite Non-commutative Associative Algebras for Setting the Hidden Discrete Logarithm Problem and Post-quantum Cryptoschemes on its Base*. Buletinul Academiei de Stiinte a Republicii Moldova, Matematica, 2019, No. 1(89), 71–78.

Dmitriy Moldovyan
St. Petersburg Institute for Informatics and Automation
of Russian Academy of Sciences
14-th line 39, 199178, St. Petersburg
Russia
E-mail: *mdn.spectr@mail.ru*

Received February 8, 2019

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 11–23
ISSN 1024–7696

Inequalities of Hermite-Hadamard Type for K-Bounded Modulus Convex Complex Functions

Silvestru Sever Dragomir

Abstract. Let $D \subset \mathbb{C}$ be a convex domain of complex numbers and $K > 0$. We say that the function $f : D \subset \mathbb{C} \to \mathbb{C}$ is called K-bounded modulus convex, for the given $K > 0$, if it satisfies the condition

$$|(1-\lambda) f(x) + \lambda f(y) - f((1-\lambda) x + \lambda y)| \leq \frac{1}{2} K \lambda (1-\lambda) |x-y|^2$$

for any x, $y \in D$ and $\lambda \in [0,1]$. In this paper we establish some new Hermite-Hadamard type inequalities for the complex integral on γ, a smooth path from $\mathbb{C}$, and K-bounded modulus convex functions. Some examples for integrals on segments and circular paths are also given.

Mathematics subject classification: 26D15, 26D10, 30A10, 30A86.
Keywords and phrases: Complex integral, Continuous functions, Holomorphic functions, Hermite-Hadamard inequality, Midpoint inequality, Trapezoid inequality.

1 Introduction

Let $(X; \|\cdot\|_X)$ and $(Y; \|\cdot\|_Y)$ be two normed linear spaces over the complex number field $\mathbb{C}$. Let C be a convex set in X. In the recent paper [1] we introduced the following class of functions:

Definition 1. *A mapping $f : C \subset X \to Y$ is called K-bounded norm convex, for some given $K > 0$, if it satisfies the condition*

$$\|(1-\lambda) f(x) + \lambda f(y) - f((1-\lambda) x + \lambda y)\|_Y \leq \frac{1}{2} K \lambda (1-\lambda) \|x-y\|_X^2 \qquad (1)$$

for any x, $y \in C$ and $\lambda \in [0,1]$. For simplicity, we denote this by $f \in \mathcal{BN}_K(C)$.

We have from (1) for $\lambda = \frac{1}{2}$ the Jensen's inequality

$$\left\| \frac{f(x) + f(y)}{2} - f\left(\frac{x+y}{2} \right) \right\|_Y \leq \frac{1}{8} K \|x-y\|_X^2$$

for any x, $y \in C$.

We observe that $\mathcal{BN}_K(C)$ is a convex subset in the linear space of all functions defined on C and with values in Y.

In the same paper [1], we obtained the following result which provides a large class of examples of such functions.

Theorem 1. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed linear spaces, C an open convex subset of X and $f : C \to Y$ a twice-differentiable mapping on C. Then for any x, $y \in C$ and $\lambda \in [0,1]$ we have*

$$\|(1-\lambda) f(x) + \lambda f(y) - f((1-\lambda) x + \lambda y)\|_Y \leq \frac{1}{2} K \lambda (1-\lambda) \|y - x\|_X^2, \tag{2}$$

where

$$K_{f''} := \sup_{z \in C} \left\| f''(z) \right\|_{\mathcal{L}(X^2;Y)} \tag{3}$$

is assumed to be finite, namely $f \in \mathcal{BN}_{K_{f''}}(C)$.

We have the following inequalities of Hermite-Hadamard type [1]:

Theorem 2. *Let $(X; \|\cdot\|_X)$ and $(Y; \|\cdot\|_Y)$ be two normed linear spaces over the complex number field $\mathbb{C}$ with Y complete. Assume that the mapping $f : C \subset X \to Y$ is continuous on the convex set C in the norm topology. If $f \in \mathcal{BN}_K(C)$ for some $K > 0$, then we have*

$$\left\| \frac{f(x) + f(y)}{2} - \int_0^1 f((1-\lambda) x + \lambda y) \, d\lambda \right\|_Y \leq \frac{1}{12} K \|x - y\|_X^2 \tag{4}$$

and

$$\left\| \int_0^1 f((1-\lambda) x + \lambda y) \, d\lambda - f\left(\frac{x+y}{2} \right) \right\|_Y \leq \frac{1}{24} K \|x - y\|_X^2 \tag{5}$$

for any x, $y \in C$.

The constants $\frac{1}{12}$ and $\frac{1}{24}$ are best possible.

For a monograph devoted to Hermite-Hadamard type inequalities see [3] and the recent survey paper [2].

Let $D \subset \mathbb{C}$ be a convex domain of complex numbers and $K > 0$. Following Definition 1, we say that the function $f : D \subset \mathbb{C} \to \mathbb{C}$ is called K-bounded modulus convex, for the given $K > 0$, if it satisfies the condition

$$|(1-\lambda) f(x) + \lambda f(y) - f((1-\lambda) x + \lambda y)| \leq \frac{1}{2} K \lambda (1-\lambda) |x - y|^2 \tag{6}$$

for any x, $y \in D$ and $\lambda \in [0,1]$. For simplicity, we denote this by $f \in \mathcal{BM}_K(D)$.

All the above results can be translated for complex functions defined on convex subsets $D \subset \mathbb{C}$.

In the following, in order to obtain several inequalities for the complex integral, we need the following facts.

Suppose γ is a smooth path from $\mathbb{C}$ parametrized by $z(t)$, $t \in [a,b]$ and f is a complex function which is continuous on γ. Put $z(a) = u$ and $z(b) = w$ with u, $w \in \mathbb{C}$. We define the integral of f on $\gamma_{u,w} = \gamma$ as

$$\int_\gamma f(z) \, dz = \int_{\gamma_{u,w}} f(z) \, dz := \int_a^b f(z(t)) \, z'(t) \, dt.$$

We observe that the actual choice of parametrization of γ does not matter.

This definition immediately extends to paths that are piecewise smooth. Suppose γ is parametrized by $z(t)$, $t \in [a,b]$, which is differentiable on the intervals $[a,c]$ and $[c,b]$, then assuming that f is continuous on γ we define

$$\int_{\gamma_{u,w}} f(z)\,dz := \int_{\gamma_{u,v}} f(z)\,dz + \int_{\gamma_{v,w}} f(z)\,dz$$

where $v := z(s)$ for some $s \in (a,b)$. This can be extended for a finite number of intervals.

We also define the integral with respect to arc-length

$$\int_{\gamma_{u,w}} f(z)\,|dz| := \int_a^b f(z(t))\left|z'(t)\right| dt$$

and the length of the curve γ is then

$$\ell(\gamma) = \int_{\gamma_{u,w}} |dz| = \int_a^b \left|z'(t)\right| dt.$$

Let f and g be holomorphic in D, an open domain and suppose $\gamma \subset D$ is a piecewise smooth path from $z(a) = u$ to $z(b) = w$. Then we have the *integration by parts formula*

$$\int_{\gamma_{u,w}} f(z)\,g'(z)\,dz = f(w)\,g(w) - f(u)\,g(u) - \int_{\gamma_{u,w}} f'(z)\,g(z)\,dz. \tag{7}$$

We recall also the *triangle inequality* for the complex integral, namely

$$\left|\int_\gamma f(z)\,dz\right| \le \int_\gamma |f(z)|\,|dz| \le \|f\|_{\gamma,\infty}\,\ell(\gamma) \tag{8}$$

where $\|f\|_{\gamma,\infty} := \sup_{z\in\gamma} |f(z)|$.

We also define the p-norm with $p \ge 1$ by

$$\|f\|_{\gamma,p} := \left(\int_\gamma |f(z)|^p\,|dz|\right)^{1/p}.$$

For $p = 1$ we have

$$\|f\|_{\gamma,1} := \int_\gamma |f(z)|\,|dz|.$$

If p, $q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then by Hölder's inequality we have

$$\|f\|_{\gamma,1} \le [\ell(\gamma)]^{1/q}\,\|f\|_{\gamma,p}.$$

Motivated by the above results, in this paper we establish some new Hermite-Hadamard type inequalities for the complex integral on γ, a smooth path from $\mathbb{C}$, and K-bounded modulus convex functions. Some examples for integrals on segments and circular paths are also given.

2 Integral Inequalities

We have:

Theorem 3. *Let $D \subset \mathbb{C}$ be a convex domain of complex numbers and $K > 0$. Assume that f is holomorphic on D and $f \in \mathcal{BM}_K(D)$. If $\gamma \subset D$ parametrized by $z(t)$, $t \in [a,b]$ is a piecewise smooth path from $z(a) = u$ to $z(b) = w$ and $v \in D$, then*

$$\left| \int_\gamma f(z)\,dz - \left[f(v) + f'(v)\left(\frac{w+u}{2} - v \right) \right](w-u) \right| \leq \frac{1}{2} K \int_\gamma |z-v|^2 \,|dz|. \tag{9}$$

In particular, we have for $v = \frac{w+u}{2}$ that

$$\left| \int_\gamma f(z)\,dz - f\left(\frac{w+u}{2} \right)(w-u) \right| \leq \frac{1}{2} K \int_\gamma \left| z - \frac{w+u}{2} \right|^2 |dz|. \tag{10}$$

Proof. Let x, $y \in D$. Since $f \in \mathcal{BM}_K(D)$, then we have

$$|f((1-\lambda)x + \lambda y) - f(x) + \lambda [f(x) - f(y)]| \leq \frac{1}{2} K \lambda (1-\lambda) |x-y|^2$$

that implies that

$$\left| \frac{f(x + \lambda(y-x)) - f(x)}{\lambda} + f(x) - f(y) \right| \leq \frac{1}{2} K (1-\lambda) |x-y|^2$$

for $\lambda \in (0,1)$.

Since f is holomorphic on D, then by letting $\lambda \to 0+$, we get

$$\left| f'(x)(y-x) + f(x) - f(y) \right| \leq \frac{1}{2} K |x-y|^2$$

that is equivalent to

$$\left| f(y) - f(x) - f'(x)(y-x) \right| \leq \frac{1}{2} K |y-x|^2 \tag{11}$$

for all x, $y \in D$.

We have

$$\begin{aligned} &\int_\gamma \left[f(z) - f(v) - f'(v)(z-v) \right] dz \qquad (12) \\ &= \int_\gamma f(z)\,dz - f(v) \int_\gamma dz - f'(v) \left(\int_\gamma z\,dz - v \int_\gamma dz \right) \\ &= \int_\gamma f(z)\,dz - f(v)(w-u) - f'(v) \left[\frac{1}{2}\left(w^2 - u^2 \right) - v(w-u) \right] \\ &= \int_\gamma f(z)\,dz - \left[f(v) + f'(v)\left(\frac{w+u}{2} - v \right) \right](w-u) \end{aligned}$$

for any $v \in D$.

By using (11) we get

$$\left|\int_\gamma f(z)\,dz - \left[f(v) + f'(v)\left(\frac{w+u}{2} - v\right)\right](w-u)\right|$$
$$\leq \int_\gamma \left|f(z) - f(v) - f'(v)(z-v)\right||dz| \leq \frac{1}{2}K\int_\gamma |z-v|^2\,|dz|$$

for any $v \in D$, which proves the inequality (9). □

If the path γ is a segment $[u,w] \subset G$ connecting two distinct points u and w in G then we write $\int_\gamma f(z)\,dz$ as $\int_u^w f(z)\,dz$.

Corollary 4. *With the assumptions of Theorem 3, suppose $[u,w] \subset D$ is a segment connecting two distinct points u and w in D and $v \in [u,w]$. Then for $v = (1-s)u + sw$ with $s \in [0,1]$, we have*

$$\left|\int_u^w f(z)\,dz - f((1-s)u+sw)(w-u) - f'((1-s)u+sw)\left(\frac{1}{2}-s\right)(w-u)^2\right|$$
$$\leq \frac{1}{6}K|w-u|^3\left[(1-s)^3 + s^3\right]. \quad (13)$$

In particular, we have, see also (5),

$$\left|\int_u^w f(z)\,dz - f\left(\frac{w+u}{2}\right)(w-u)\right| \leq \frac{1}{24}K|w-u|^3. \quad (14)$$

Proof. It follows by Theorem 3 by observing that

$$\int_u^w |z-v|^2\,|dz| = |w-u|\int_0^1 |(1-t)u + tw - (1-s)u - sw|^2\,dt$$
$$= |w-u|\int_0^1 |(1-t)u + tw - (1-s)u - sw|^2\,dt$$
$$= |w-u|^3\int_0^1 (t-s)^2\,dt = \frac{1}{3}|w-u|^3\left[(1-s)^3 + s^3\right]$$

for $s \in [0,1]$. □

Theorem 5. *Let $D \subset \mathbb{C}$ be a convex domain of complex numbers and $K > 0$. Assume that f is holomorphic on D and $f \in \mathcal{BM}_K(D)$. If $\gamma \subset D$ parametrized by $z(t)$, $t \in [a,b]$ is a piecewise smooth path from $z(a) = u$ to $z(b) = w$ and $v \in D$, then*

$$\left|\frac{1}{2}\left[f(w)(w-v) + f(u)(v-u) + f(v)(w-u)\right] - \int_\gamma f(z)\,dz\right|$$

$$\leq \frac{1}{4}K\int_{\gamma}|z-v|^{2}\,|dz|\,. \quad (15)$$

In particular, we have for $v=\frac{w+u}{2}$ *that*

$$\left|\frac{1}{2}\left[\frac{f(w)+f(u)}{2}+f\left(\frac{w+u}{2}\right)\right](w-u)-\int_{\gamma}f(z)\,dz\right| \leq \frac{1}{4}K\int_{\gamma}\left|z-\frac{w+u}{2}\right|^{2}|dz|\,. \quad (16)$$

Proof. By using (11) we get

$$\int_{\gamma}\left|f(v)-f(z)-f'(z)(v-z)\right||dz| \leq \frac{1}{2}K\int_{\gamma}|v-z|^{2}\,|dz| \quad (17)$$

for $v\in D$.

By the complex integral properties, we have

$$\left|\int_{\gamma}\left[f(v)-f(z)-f'(z)(v-z)\right]dz\right| \leq \int_{\gamma}\left|f(v)-f(z)-f'(z)(v-z)\right||dz| \quad (18)$$

for $v\in D$.

Using integration by parts, we get

$$\begin{aligned}
&\int_{\gamma}\left[f(v)-f(z)-f'(z)(v-z)\right]dz \\
&\qquad = f(v)\int_{\gamma}dz-\int_{\gamma}f(z)\,dz-\int_{\gamma}f'(z)(v-z)\,dz \\
&\qquad = f(v)(w-u)-\int_{\gamma}f(z)\,dz-\left[f(z)(v-z)\big|_{u}^{w}+\int_{\gamma}f(z)\,dz\right] \\
&= f(v)(w-u)-\int_{\gamma}f(z)\,dz-f(w)(v-w)+f(u)(v-u)-\int_{\gamma}f(z)\,dz \\
&\qquad = f(w)(w-v)+f(u)(v-u)+f(v)(w-u)-2\int_{\gamma}f(z)\,dz,
\end{aligned}$$

which implies that

$$\frac{1}{2}\left[f(w)(w-v)+f(u)(v-u)+f(v)(w-u)\right]-\int_{\gamma}f(z)\,dz = \frac{1}{2}\int_{\gamma}\left[f(v)-f(z)-f'(z)(v-z)\right]dz \quad (19)$$

for $v\in D$.

By utilising (17)-(19) we get the desired result (15). □

We have:

Corollary 6. *With the assumptions of Theorem 3, suppose $[u,w] \subset D$ is a segment connecting two distinct points u and w in D and $v \in [u,w]$. Then for $v = (1-s)u + sw$ with $s \in [0,1]$, we have*

$$\left|\frac{1}{2}\left[(1-s)f(w) + sf(u) + f((1-s)u+sw)\right](w-u) - \int_u^w f(z)\,dz\right| \leq \frac{1}{12}K|w-u|^3\left[(1-s)^3 + s^3\right]. \quad (20)$$

In particular, we have for $v = \frac{w+u}{2}$ that

$$\left|\frac{1}{2}\left[\frac{f(w)+f(u)}{2} + f\left(\frac{w+u}{2}\right)\right](w-u) - \int_u^w f(z)\,dz\right| \leq \frac{1}{48}K|w-u|^3. \quad (21)$$

We observe that, if f is holomorphic on D and $K = \sup_{z\in D}|f''(z)|$ is finite, then by (9) and (10) we have

$$\left|\int_\gamma f(z)\,dz - \left[f(v) + f'(v)\left(\frac{w+u}{2} - v\right)\right](w-u)\right| \leq \frac{1}{2}\sup_{z\in D}|f''(z)|\int_\gamma |z-v|^2\,|dz| \quad (22)$$

for all $v \in D$. In particular,

$$\left|\int_\gamma f(z)\,dz - f\left(\frac{w+u}{2}\right)(w-u)\right| \leq \frac{1}{2}\sup_{z\in D}|f''(z)|\int_\gamma \left|z - \frac{w+u}{2}\right|^2 |dz|. \quad (23)$$

From (15) and (16) we get

$$\left|\frac{1}{2}\left[f(w)(w-v) + f(u)(v-u) + f(v)(w-u)\right] - \int_\gamma f(z)\,dz\right| \leq \frac{1}{4}\sup_{z\in D}|f''(z)|\int_\gamma |z-v|^2\,|dz|. \quad (24)$$

for all $v \in D$. In particular,

$$\left|\frac{1}{2}\left[\frac{f(w)+f(u)}{2} + f\left(\frac{w+u}{2}\right)\right](w-u) - \int_\gamma f(z)\,dz\right| \leq \frac{1}{4}\sup_{z\in D}|f''(z)|\int_\gamma \left|z - \frac{w+u}{2}\right|^2 |dz|. \quad (25)$$

The inequalities (22)-(25) provide many examples of interest as follows.

If we consider the function $f(z)=\exp z$, $z\in\mathbb{C}$ and $\gamma\subset\mathbb{C}$ parametrized by $z(t)$, $t\in[a,b]$ is a piecewise smooth path from $z(a)=u$ to $z(b)=w$ then by (22)-(25) we have by the inequalities

$$\left|\exp w-\exp u-\left(1+\frac{w+u}{2}-v\right)(w-u)\exp v\right| \leq \frac{1}{2}\sup_{z\in D}|\exp z|\int_\gamma |z-v|^2\,|dz| \quad (26)$$

for all $v\in\mathbb{C}$. In particular,

$$\left|\exp w-\exp u-\exp\left(\frac{w+u}{2}\right)(w-u)\right| \leq \frac{1}{2}\sup_{z\in D}|\exp z|\int_\gamma \left|z-\frac{w+u}{2}\right|^2|dz|. \quad (27)$$

We also have

$$\left|\frac{1}{2}\left[(w-v)\exp w+(v-u)\exp u+(w-u)\exp v\right]-\exp w+\exp u\right| \leq \frac{1}{4}\sup_{z\in D}|\exp z|\int_\gamma |z-v|^2\,|dz|. \quad (28)$$

for all $v\in\mathbb{C}$. In particular,

$$\left|\frac{1}{2}\left[\frac{\exp w+\exp u}{2}+\exp\left(\frac{w+u}{2}\right)\right](w-u) -\exp w+\exp u\right| \leq \frac{1}{4}\sup_{z\in D}|\exp z|\int_\gamma \left|z-\frac{w+u}{2}\right|^2|dz|. \quad (29)$$

Consider the function $F(z)=\operatorname{Log}(z)$ where $\operatorname{Log}(z)=\ln|z|+i\operatorname{Arg}(z)$ and $\operatorname{Arg}(z)$ is such that $-\pi<\operatorname{Arg}(z)\leq\pi$. Log is called the *"principal branch"* of the complex logarithmic function. F is analytic on all of $\mathbb{C}\backslash\{x+iy:x\leq 0,\ y=0\}$ and $F'(z)=\frac{1}{z}$ on this set.

If we consider $f:D\to\mathbb{C}$, $f(z)=\frac{1}{z}$ where $D\subset\mathbb{C}\backslash\{x+iy:x\leq 0,\ y=0\}$, then F is a primitive of f on D and if $\gamma\subset D$ parametrized by $z(t)$, $t\in[a,b]$ is a piecewise smooth path from $z(a)=u$ to $z(b)=w$, then

$$\int_\gamma f(z)\,dz=\operatorname{Log}(w)-\operatorname{Log}(u).$$

For $D\subset\mathbb{C}\backslash\{x+iy:x\leq 0,\ y=0\}$, define $d:=\inf_{z\in D}|z|$ and assume that $d\in(0,\infty)$. By the inequalities (22)-(25) we then have

$$\left|\operatorname{Log}(w)-\operatorname{Log}(u)-\left[\frac{1}{v}-\frac{1}{v^2}\left(\frac{w+u}{2}-v\right)\right](w-u)\right| \leq \frac{1}{d^3}\int_\gamma |z-v|^2\,|dz| \quad (30)$$

for all $v \in D$. In particular,

$$\left|\operatorname{Log}(w)-\operatorname{Log}(u)-\left(\frac{w+u}{2}\right)^{-1}(w-u)\right| \leq \frac{1}{d^3}\int_\gamma \left|z-\frac{w+u}{2}\right|^2 |dz|. \quad (31)$$

We also have

$$\left|\frac{1}{2}\left(\frac{w-v}{w}+\frac{v-u}{u}+\frac{w-u}{v}\right)-\operatorname{Log}(w)+\operatorname{Log}(u)\right| \leq \frac{1}{2d^3}\int_\gamma |z-v|^2\,|dz|. \quad (32)$$

for all $v \in D$. In particular,

$$\left|\frac{1}{2}\left[\frac{u+w}{2wu}+\left(\frac{w+u}{2}\right)^{-1}\right](w-u) - \operatorname{Log}(w)+\operatorname{Log}(u)\right| \leq \frac{1}{2d^3}\int_\gamma \left|z-\frac{w+u}{2}\right|^2 |dz|. \quad (33)$$

3 Examples for Circular Paths

Let $[a,b] \subseteq [0,2\pi]$ and the circular path $\gamma_{[a,b],R}$ be centered at 0 and with radius $R>0$

$$z(t) = R\exp(it) = R(\cos t + i\sin t),\ t \in [a,b].$$

If $[a,b]=[0,\pi]$ then we get a half circle while for $[a,b]=[0,2\pi]$ we get the full circle.

Since

$$\left|e^{is}-e^{it}\right|^2 = \left|e^{is}\right|^2 - 2\operatorname{Re}\left(e^{i(s-t)}\right)+\left|e^{it}\right|^2 = 2-2\cos(s-t) = 4\sin^2\left(\frac{s-t}{2}\right)$$

for any t, $s \in \mathbb{R}$, then

$$\left|e^{is}-e^{it}\right|^r = 2^r\left|\sin\left(\frac{s-t}{2}\right)\right|^r \quad (34)$$

for any t, $s \in \mathbb{R}$ and $r>0$. In particular,

$$\left|e^{is}-e^{it}\right| = 2\left|\sin\left(\frac{s-t}{2}\right)\right|$$

for any t, $s \in \mathbb{R}$.

For $s = a$ and $s = b$ we have

$$\left|e^{ia} - e^{it}\right| = 2\left|\sin\left(\frac{a-t}{2}\right)\right| \text{ and } \left|e^{ib} - e^{it}\right| = 2\left|\sin\left(\frac{b-t}{2}\right)\right|.$$

If $u = R\exp(ia)$ and $w = R\exp(ib)$ then

$$\begin{aligned} w - u &= R\left[\exp(ib) - \exp(ia)\right] = R\left[\cos b + i\sin b - \cos a - i\sin a\right] \\ &= R\left[\cos b - \cos a + i\left(\sin b - \sin a\right)\right]. \end{aligned}$$

Since

$$\cos b - \cos a = -2\sin\left(\frac{a+b}{2}\right)\sin\left(\frac{b-a}{2}\right)$$

and

$$\sin b - \sin a = 2\sin\left(\frac{b-a}{2}\right)\cos\left(\frac{a+b}{2}\right),$$

hence

$$\begin{aligned} w - u &= R\left[-2\sin\left(\frac{a+b}{2}\right)\sin\left(\frac{b-a}{2}\right) + 2i\sin\left(\frac{b-a}{2}\right)\cos\left(\frac{a+b}{2}\right)\right] \\ &= 2R\sin\left(\frac{b-a}{2}\right)\left[-\sin\left(\frac{a+b}{2}\right) + i\cos\left(\frac{a+b}{2}\right)\right] \\ &= 2Ri\sin\left(\frac{b-a}{2}\right)\left[\cos\left(\frac{a+b}{2}\right) + i\sin\left(\frac{a+b}{2}\right)\right] \\ &= 2Ri\sin\left(\frac{b-a}{2}\right)\exp\left[\left(\frac{a+b}{2}\right)i\right]. \end{aligned}$$

We also have

$$z'(t) = Ri\exp(it) \text{ and } \left|z'(t)\right| = R$$

for $t \in [a, b]$.

In what follows we assume that f is defined on a domain containing the circular path $\gamma_{[a,b],R}$ and that f is holomorphic on that domain.

Consider the circular path $\gamma_{[a,b],R}$ and assume that $v = R\exp(is) \in \gamma_{[a,b],R}$ with $s \in [a, b]$. Then by using the inequality (9) we get

$$\begin{aligned} &\left|Ri\int_a^b f\left(R\exp(it)\right)\exp(it)\,dt \right. \\ &\quad - \left[f\left(R\exp(is)\right) + f'\left(R\exp(is)\right)\left(\frac{R\exp(ib) + R\exp(ia)}{2} - R\exp(is)\right)\right] \\ &\qquad \left. \times 2Ri\sin\left(\frac{b-a}{2}\right)\exp\left[\left(\frac{a+b}{2}\right)i\right]\right| \\ &\leq \frac{1}{2}\sup_{t\in[a,b]}\left|f''\left(R\exp(it)\right)\right| R\int_a^b \left|R\exp(it) - R\exp(is)\right|^2 dt \end{aligned}$$

$$= \frac{1}{2} \sup_{t\in[a,b]} \left|f''\left(R\exp\left(it\right)\right)\right| R^3 \int_a^b 4\sin^2\left(\frac{s-t}{2}\right) dt$$

$$= 2 \sup_{t\in[a,b]} \left|f''\left(R\exp\left(it\right)\right)\right| R^3 \int_a^b \sin^2\left(\frac{s-t}{2}\right) dt,$$

which is equivalent to

$$\left| \int_a^b f\left(R\exp\left(it\right)\right)\exp\left(it\right) dt \right.$$
$$- 2R\left[f\left(R\exp\left(is\right)\right) + f'\left(R\exp\left(is\right)\right)\left(\frac{\exp\left(ib\right)+\exp\left(ia\right)}{2} - \exp\left(is\right)\right)\right]$$
$$\left. \times \sin\left(\frac{b-a}{2}\right)\exp\left[\left(\frac{a+b}{2}\right)i\right]\right|$$
$$\leq 2 \sup_{t\in[a,b]} \left|f''\left(R\exp\left(it\right)\right)\right| R^2 \int_a^b \sin^2\left(\frac{s-t}{2}\right) dt \quad (35)$$

for $s \in [a,b]$.

Since

$$\sin^2\left(\frac{s-t}{2}\right) = \frac{1-\cos\left(s-t\right)}{2},$$

hence

$$\int_a^b \sin^2\left(\frac{s-t}{2}\right) dt$$
$$= \int_a^b \frac{1-\cos\left(s-t\right)}{2} dt = \frac{1}{2}\left[b-a-\sin\left(b-s\right)-\sin\left(s-a\right)\right]$$
$$= \frac{1}{2}\left[b-a-2\sin\left(\frac{b-a}{2}\right)\cos\left(\frac{a+b}{2}-s\right)\right]$$
$$= \frac{b-a}{2} - \sin\left(\frac{b-a}{2}\right)\cos\left(\frac{a+b}{2}-s\right)$$

for $s \in [a,b]$.

Therefore by (35) we get

$$\left| \int_a^b f\left(R\exp\left(it\right)\right)\exp\left(it\right) dt \right.$$
$$- 2R\left[f\left(R\exp\left(is\right)\right) + f'\left(R\exp\left(is\right)\right)\left(\frac{\exp\left(ib\right)+\exp\left(ia\right)}{2} - \exp\left(is\right)\right)\right]$$
$$\left. \times \sin\left(\frac{b-a}{2}\right)\exp\left[\left(\frac{a+b}{2}\right)i\right]\right|$$
$$\leq 2R^2 \sup_{t\in[a,b]} \left|f''\left(R\exp\left(it\right)\right)\right| \left[\frac{b-a}{2} - \sin\left(\frac{b-a}{2}\right)\cos\left(\frac{a+b}{2}-s\right)\right] \quad (36)$$

for $s \in [a,b]$.

In particular, for $s = \frac{a+b}{2}$, we obtain from (36) the best possible inequality

$$\begin{aligned}\left| \int_a^b f\left(R\exp\left(it\right)\right)\exp\left(it\right)dt \right. \\ - 2R\left[f\left(R\exp\left(\frac{a+b}{2}i\right)\right) + f'\left(R\exp\left(\frac{a+b}{2}i\right)\right)\right. \\ \times \left.\left(\frac{\exp\left(ib\right)+\exp\left(ia\right)}{2} - \exp\left(\frac{a+b}{2}i\right)\right)\right] \\ \times \sin\left(\frac{b-a}{2}\right)\exp\left[\left(\frac{a+b}{2}\right)i\right]\Bigg| \\ \leq 2R^2 \sup_{t\in[a,b]}\left|f''\left(R\exp\left(it\right)\right)\right|\left[\frac{b-a}{2} - \sin\left(\frac{b-a}{2}\right)\right]. \quad (37)\end{aligned}$$

By utilising the inequality (24) for the circular path $\gamma_{[a,b],R}$ and $v = R\exp\left(is\right) \in \gamma_{[a,b],R}$ with $s \in [a,b]$, we also get

$$\begin{aligned}\Bigg| f\left(R\exp\left(ib\right)\right)\sin\left(\frac{b-s}{2}\right)\exp\left[\left(\frac{s+b}{2}\right)i\right] \\ + f\left(R\exp\left(ia\right)\right)\sin\left(\frac{s-a}{2}\right)\exp\left[\left(\frac{a+s}{2}\right)i\right] \\ + f\left(R\exp\left(is\right)\right)\sin\left(\frac{b-a}{2}\right)\exp\left[\left(\frac{a+b}{2}\right)i\right] \\ - \int_a^b f\left(R\exp\left(it\right)\right)\exp\left(it\right)dt\Bigg| \\ \leq R^2 \sup_{t\in[a,b]}\left|f''\left(R\exp\left(it\right)\right)\right|\left[\frac{b-a}{2} - \sin\left(\frac{b-a}{2}\right)\cos\left(\frac{a+b}{2} - s\right)\right]. \quad (38)\end{aligned}$$

In particular, for $s = \frac{a+b}{2}$, we get from (38) best possible inequality

$$\begin{aligned}\Bigg| f\left(R\exp\left(bi\right)\right)\sin\left(\frac{b-a}{4}\right)\exp\left[\left(\frac{a+3b}{4}\right)i\right] \\ + f\left(R\exp\left(ia\right)\right)\sin\left(\frac{b-a}{4}\right)\exp\left[\left(\frac{3a+b}{4}\right)i\right] \\ + f\left(R\exp\left(\frac{a+b}{2}i\right)\right)\sin\left(\frac{b-a}{2}\right)\exp\left[\left(\frac{a+b}{2}\right)i\right] \\ - \int_a^b f\left(R\exp\left(it\right)\right)\exp\left(it\right)dt\Bigg| \\ \leq R^2 \sup_{t\in[a,b]}\left|f''\left(R\exp\left(it\right)\right)\right|\left[\frac{b-a}{2} - \sin\left(\frac{b-a}{2}\right)\right]. \quad (39)\end{aligned}$$

References

[1] Dragomir S. S. *Integral inequalities of Hermite-Hadamard type for K-bounded norm convex mappings,* Ukrainian Mathematical Journal, 2017, **68**, No. 10, 1530–1551.

[2] Dragomir S. S. *Ostrowski type inequalities for Lebesgue integral: a survey of recent results,* Australian J. Math. Anal. Appl.,Volume **14**, Issue 1, Article 1, pp. 1-287, 2017. [Online http://ajmaa.org/cgi-bin/paper.pl?string=v14n1/V14I1P1.tex].

[3] Dragomir S. S., Pearce C. E. M. *Selected Topics on Hermite-Hadamard Inequalities and Applications*, RGMIA Monographs, 2000. [Online

http://rgmia.org/monographs/hermite_hadamard.html].

S. S. Dragomir
College of Engineering & Science Victoria University
PO Box 14428 Melbourne City, MC 8001
Australia.

E-mail: *sever.dragomir@vu.edu.au*

Received September 11, 2019

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 24–29
ISSN 1024–7696

Commutative Weakly Tripotent Group Rings

Peter V. Danchev

Abstract. Very recently, Breaz and Cîmpean introduced and examined in Bull. Korean Math. Soc. (2018) the class of so-called *weakly tripotent rings* as those rings R whose elements satisfy at leat one of the equations $x^3 = x$ or $(1-x)^3 = 1-x$. These rings are generally non-commutative. We here obtain a criterion when the commutative group ring RG is weakly tripotent in terms only of a ring R and of a group G plus their sections.
Actually, we also show that these weakly tripotent rings are *strongly invo-clean rings* in the sense of Danchev in Commun. Korean Math. Soc. (2017). Thereby, our established criterion somewhat strengthens previous results on commutative strongly invo-clean group rings, proved by the present author in Univ. J. Math. & Math. Sci. (2018). Moreover, this criterion helps us to construct a commutative strongly invo-clean ring of characteristic 2 which is *not* weakly tripotent, thus showing that these two ring classes are different.

Mathematics subject classification: 16S34; 16U99; 20C07.
Keywords and phrases: Tripotent rings, weakly tripotent rings, strongly invo-clean rings, group rings.

1 Introduction and Background

Throughout the text of the current article all rings into consideration are assumed to be associative, possessing the identity element 1 which differs from the zero element 0. Our terminology and notation in both ring and group theories are mainly standard and some additional notions will be specified in the sequel. For instance, for a ring R, the symbol $U(R)$ denotes the set of all units in R, $Id(R)$ the set of all idempotents in R, $Nil(R)$ the set of all nilpotents in R which, in the commutative case, coincides with the nil-radical $N(R)$, and $J(R)$ the Jacobson radical of R. Likewise, for an abelian group G, the letter G^2 stands for the subgroup of G consisting of all elements of the type $\{g^2 \mid g \in G\}$. As usual, RG designates the group ring of G over R with augmentation ideal $I(RG;G)$ generated by elements of the sort $\{g-1 \mid g \in G\}$.

An element t of a ring R is called *tripotent* if the equality $t^3 = t$ holds. If each element of R is with this property, R is said to be *tripotent* as well. The complete description of such rings is well-known as a subdirect product (= a special subring of a direct product) of a family of copies of the fields $\mathbb{Z}_2$ and $\mathbb{Z}_3$.

On the other side, generalizing the aforementioned concept, in [1] were explored the so-called *weakly tripotent* rings that are rings in which at least one of the elements t or $1-t$ is a tripotent. It is immediate that weakly tripotent rings of characteristic

3 are themselves tripotent as $(1-t)^3 = 1-t^3 = 1-t$ yields that $t^3 = t$. Interestingly, for any element y from the Jacobson radical $J(R)$ of such a ring R, it must be that $y^2 = 2y$. To look at this, we have that $y^3 = y$ or $(1-y)^3 = 1-y$. In the first version, $y(1-y^2) = 0$ gives that $y = 0$ as $1-y^2 \in U(R)$, so that $y^2 = 0 = 2y$. In the second one, we observe that $(1-y)^2 = 1$ as $1-y \in U(R)$ and hence $y^2 = 2y$, as promised. This substantiates that $y^2 = 2y$ is always true. Replacing $y \to -y$ allows us to get that $4y = 0$.

These rings are, in general, non-commutative and there is no a complete description of their structure yet. However, a complete characterization theorem for a commutative ring R to be weakly tripotent is [1, Theorem 14]. However, since the direct product of an arbitrary family of Boolean rings is still a Boolean ring, we will state below this theorem in an equivalent form, but in a manner which is slightly more transparent and convenient for direct applications, as follows: *A commutative ring R is weakly tripotent if, and only if, $R \cong R_1 \times R_2$, where $R_2 = \{0\}$ or $x^3 = x$ holds $\forall\ x \in R_2$ with $3R_2 = \{0\}$, and $R_1 = \{0\}$ or R_1 is a subdirect product of $R_{11} \times R_{12}$ with $2^k R_1 = \{0\}$ for some $k = 1, 2, 3$ (or, in other words, $3 \in U(R_1)$), where R_{11} and R_{12} are rings such that either $R_{11} = \{0\}$ or $R_{11}/J(R_{11}) \cong \mathbb{Z}_2$ with $y^2 = 2y\ \forall\ y \in J(R_{11})$, and R_{12} is a Boolean ring.* Resultantly, it readily follows that each commutative weakly tripotent ring of even characteristic (i.e., of characteristic 2^k for $k = 1, 2, 3$) is *nil-clean* in the sense that the quotient-ring $R/N(R)$ is Boolean or, equivalently, the factor-ring $R/J(R)$ is Boolean with nil $J(R)$. Besides, an pretty easy consequence is that (compare also with [1, Corollary 9]) a ring R is weakly tripotent such that $3 \in U(R)$ and $Id(R) = \{0, 1\}$ if, and only if, $R/J(R) \cong \mathbb{Z}_2$ with $z^2 = 2z$ for any $z \in J(R)$. A further characterization of arbitrary weakly tripotent rings (possibly non-commutative) is given in [5].

Our major motivation to write up this paper is to use the cited above theorem in order to deduce a full criterion for a commutative group ring to be weakly tripotent only in terms of the coefficient ring and the former group plus their divisions. This will be successfully done in the next section.

2 Main Results

Let us recall that a ring R is said to be *strongly invo-clean* in [2], provided for any element $r \in R$ the existence of an idempotent $e \in R$ and a unit $v \in R$ of order at most 2 such that $r = e + v$ with $ev = ve$. These ring were completely classified in [2, Corollary 2.17] with the aid of structural results from [6].

The next relationship considerably strengthens [1, Corollary 4].

Proposition 1. *Every weakly tripotent ring is strongly invo-clean.*

Proof. For such a ring R, we have $r^3 = r$ or $(1-r)^3 = 1-r$ whenever $r \in R$. In the first case, one writes that $r = (1-r^2) + (r^2 + r - 1)$. A direct manipulation shows that $1 - r^2 \in Id(R)$ as $r^2 \in Id(R)$ and that $(r^2 + r - 1)^2 = 1$ observing elementarily that these two elements commute, as required.

Dealing with the other equality $(1-r)^3 = 1-r$, the replacement $r \to 1-r$ with the above trick at hand lead to this that $1-r = f+w$ for some commuting idempotent $f = 1-(1-r)^2 = 2r-r^2$ and an involution $w = 1-3r+r^2$. Therefore, $r = (1-f)+(-w)$, where $(1-f)^2 = 1-f$ and $(-w)^2 = w^2 = 1$, as required. □

We will demonstrate now that the converse implication is totally untrue. Before doing that, we need the following technicality.

Lemma 1. *Let B be a boolean ring and let G be an abelian group. Then the group ring BG is weakly tripotent if, and only if, $B \cong \mathbb{Z}_2$ and $G^2 = \{1\}$.*

Proof. "**Necessity.**" Given $g \in G$, it follows that $g^3 = g$ or $(1-g)^3 = 1-g$. The first equality yields that $g^2 = 1$. The second equality can be written as $g^3 - g^2 + g = 0$ since $2 = 0$ in both B and BG. If we assume that $g^2 \neq 1$, then one sees that $g^3 \neq g$, $g^3 \neq g^2$ and $g^2 \neq g$. This assumption, however, leads to a contradiction since then $g^3 - g^2 + g$ is an element in BG written in canonical form. Finally, it must be in the second case that $g^2 = 1$ as well, as expected.

Next, for every $b \in B$, we consider the element $b+g \in BG$, where $1 \neq g \in G$. Therefore, $(b+g)^3 = b+g$ or $[(1-b)-g]^3 = (1-b)-g$. In the first possibility, we derive that $b+bg+b+g = b+g$, i.e., that $b+bg = 0$. This forces at once that $b = 0$. Since $1-b$ is again an idempotent, the second possibility guarantees in a way of similarity that $1-b = 0$, that is, $b = 1$. Hence $B = \{0,1\} \cong \mathbb{Z}_2$, as asserted.

"**Sufficiency.**" Each (possibly non-zero) element x of BG has take the form $x = g_1 + \cdots + g_n$ for some $n \in \mathbb{N}$. It is clear that $x^2 = n$, whence $x^2 = 1$ if $n = 2k+1$ or $x^2 = 0$ if $n = 2k$. In the first situation, one infers that $x^3 = x$, as wanted. In the second situation, $(1-x)^2 = 1+x^2 = 1$ so that $(1-x)^3 = 1-x$, as desired. □

So, we are in a position to exhibit the concrete construction.

Example 1. There exists a commutative strongly invo-clean ring of characteristic 2 which is *not* weakly tripotent.

In fact, consider the group ring $K = (\mathbb{Z}_2 \times \mathbb{Z}_2)G$ with $G^2 = \{1\}$. Utilizing Lemma 1 this group ring is manifestly non-weakly tripotent. Nevertheless, we claim that such a ring is strongly invo-clean. To show this, we apply the chief result from [9] to deduce that K is nil-clean in the sense that each element $a \in K$ is writable as $a = q+e$, where $q \in Nil(K) = N(K)$ and $e \in Id(K)$. On the other hand, in view of [8] or [7], one finds that $N(K) = I(KG;G)$ because $N(\mathbb{Z}_2 \times \mathbb{Z}_2) = \{0\}$. It is obvious now that $N(K)$ has an index of nilpotence not exceeding 2 as char $(K) = 2$. And since $a = (q+1)+(1+e)$ with $(q+1)^2 = 1$ and $(1+e)^2 = 1+e$, we are set. This concludes our initial claim in the example.

We are now ready to proceed by proving with our necessary and sufficient condition for a commutative group ring to be weakly tripotent, thus generalizing Lemma 1 listed above.

Theorem 1. *Suppose R is a commutative ring and G is an abelian group. Then the group ring RG is weakly tripotent if, and only if, exactly one of the next two points holds:*

(i) $G = \{1\}$ and R is weakly tripotent.

(ii) $G \neq \{1\}$ with $G^2 = \{1\}$, R is weakly tripotent such that $R \cong R_1 \times R_2$, where both R_1 and R_2 are weakly tripotent rings with $R_1 = \{0\}$ or R_1 a subdirect product of $L \times \mathbb{Z}_2$ for either $L = \{0\}$ or $L/J(L) \cong \mathbb{Z}_2$ having $y^2 = 2y$ for all $y \in J(L)$, and $R_2 = \{0\}$ or $\mathrm{char}(R_2) = 3$ with $x^3 = x$ for all $x \in R_2$, and either

(ii.1) $|G| = 2$, $2d^2 = 2d$ for all $d \in L$ (and hence $4L = \{0\}$);

or

(ii.2) $|G| > 2$, $\mathrm{char}(L) = 2$.

Proof. As $RG \cong R$ whenever G is the trivial group, we shall hereafter assume that it is non-trivial.

"**Necessity.**" Each element x in RG satisfies one of the equations $x^3 = x$ or $(1-x)^3 = 1-x$. Thus, for any $1 \neq g \in G$, one has that $g^3 = g$ or $(1-g)^3 = 1-g$. The first equality gives that $g^2 = 1$. The second one assures that $g^3 - 3g^2 + 2g = 0$. Assume in a way of contradiction that $g^2 \neq 1$. Since then $g \neq g^3 \neq g^2 \neq g$ and since the equation $g^3 - 3g^2 + 2g = 0$ is a canonical record, we will obtain a contrary to our assumption. Consequently, one extracts in both cases that $g^2 = 1$. Even something more: in the second equality we have that $3g - 3 = 0$ implying $3 = 0$. Thus the equation $(1-x)^3 = 1-x$ is tantamount to $x^3 = x$, as expected.

Furthermore, as $R \subseteq RG$, or even there is an epimorphism (= a surjective homomorphism) $RG \to R$ defined by the augmentation map, we employ [1, Lemma 1] to get that R is weakly tripotent, too. We, therefore, may write in conjunction with the listed above theorem for commutative weakly tripotent rings that $R \cong R_1 \times R_2$, where both R_1 and R_2 are weakly tripotent rings with $R_1 = \{0\}$ or R_1 a subdirect product of $L \times B$ for some Boolean ring B and either $L = \{0\}$ or $L/J(L) \cong \mathbb{Z}_2$ having $z^2 = 2z$ for all $z \in J(L)$, and $R_2 = \{0\}$ or $\mathrm{char}(R_2) = 3$. It is then an easy technical matter to check that $RG \cong R_1G \times R_2G$, where R_1G, R_2G both remain weakly tripotent. We will study these two direct factors separately:

About R_1G: Here R_1G is a subdirect product of $LG \times BG$, where LG and BG are both weakly tripotent rings with $2 \in J(L)$ for the first ring. In the latter case, Lemma 1 enables us that $B \cong \mathbb{Z}_2$, as stated. Concentrating now on LG, we shall consider two possibilities on the cardinality of the basis group G. In fact, firstly suppose that $|G| = 2$. Then, for every $r \in L$, we consider the element $r(1-g) \in J(LG)$, where $1 \neq g \in G$ with $g^2 = 1$ (see, for instance,[7]). Since $[r(1-g)]^2 = 2r(1-g)$, one inspects by simple manipulations that $(2r^2 - 2r) - (2r^2 - 2r)g = 0$ implying immediately the desired equality $2r^2 = 2r$. Secondly, suppose that $|G| > 2$. Then there are two different elements $g, h \in G \setminus \{1\}$ with $g^2 = h^2 = 1$. Since both $g - 1 \in J(LG)$ and $1 - h \in J(LG)$, it follows directly that $g - h = (g-1) + (1-h) \in J(LG)$ and so $(g-h)^2 = 2(g-h)$ yielding that

$2gh + 2g - 2h - 2 = 0$. But as $gh \neq 1$ because $g \neq h$ as well as $gh \neq g$ and $gh \neq h$ because $h \neq 1$ and $g \neq 1$, we conclude after all that $2 = 0$ in L, as required.

About R_2G: As it was easily observed above, all weakly tripotent rings having characteristic 3 are obviously tripotent. Since then the non-zero ring R_2 of characteristic 3 has to be tripotent, we are done.

"**Sufficiency.**" Writing $RG \cong R_1G \times R_2G$, we must explore both direct factors R_1G and R_2G.

Dealing with the non-zero variant of the first direct factor R_1G, we detect that R_1G is a subring of $LG \times \mathbb{Z}_2G$. The application of Lemma 1 is a guarantor that $\mathbb{Z}_2G$ is weakly tripotent bearing in mind that $G^2 = \{1\}$. We now assert that $J(LG) = J(L)G + I(LG; G)$. Indeed, seeing that $2 \in J(L)$, one verifies in virtue of [7] (see also [8]) that $J(LG) = J(L)G + \langle r(g-1) \mid r \in L, g \in G\rangle \subseteq J(L)G + I(LG; G)$. To derive the converse inclusion, we differ two possible cases for the cardinality of G: if $|G| = 2$, then $g(g-1) = -(g-1)$ as $g^2 = 1$. If now $|G| > 2$ with $G^2 = \{1\}$ and $2 = 0$ in L, we are observing that $I^2(LG; G) = \{0\}$ and so $I(LG; G) \subseteq J(LG)$ since $I(LG; G)$ is a nil ideal, giving the pursued equality, because $J(L)G \subseteq J(LG)$ holds always (cf.[7]). That is why, $LG/J(LG) = LG/[J(L) + I(LG; G)] \cong L/J(L) \cong \mathbb{Z}_2$, as needed. What suffices to prove in order to complete this point is that $z^2 = 2z$ for any $z \in J(LG)$. This, however, follows directly by the same token as in the proof of [3, Theorem 2.3]. Finally, we arrive at the fact that LG is a weakly tripotent ring, whence so does R_1G in accordance with [1, Lemma 1].

Further, concerning the second direct factor R_2G, since R_2 is either zero or a tripotent ring of characteristic 3, one plainly obtains that R_2G is also a tripotent ring by taking into account that $G^2 = \{1\}$. We, finally, appeal to [1, Proposition 6] to conclude that RG is weakly tripotent, as claimed. □

We close our work with the following challenging question.

Proposition 2. *Find a criterion for a non-commutative group ring to be weakly tripotent.*

Funding. The work in this article is partially supported by the Bulgarian National Science Fund under Grant KP-06 N 32/1 of Dec. 07, 2019

References

[1] Breaz S., Cîmpean A. *Weakly tripotent rings.* Bull. Korean Math. Soc., 2018, **55**, No. 4, 1179–1187.

[2] Danchev P. V. *Invo-clean unital rings.* Commun. Korean Math. Soc., 2017, **32**, 19–27.

[3] Danchev P. V. *Commutative invo-clean group rings.* Univ. J. Math. & Math. Sci., 2018, **11**, 1–6.

[4] Danchev P. V. *Commutative nil-clean and π-regular group rings.* Uzbek Math. J., 2019, 33–39.

[5] Danchev P. V. *A characterization of weakly tripotent rings.* Submitted preprint.

[6] Danchev P. V., Lam T.-Y. *Rings with unipotent units*. Publ. Math. (Debrecen), 2016, **88**, 449–466.

[7] Karpilovsky G. *The Jacobson radical of commutative group rings*. Arch. Math. (Basel), 1982, **39**, 428–430.

[8] May W. L. *Group algebras over finitely generated rings*. J. Algebra, 1976, **39**, 483–511.

[9] McGovern W. Wm., Raja S., Sharp A. *Commutative nil clean group rings*. J. Algebra Appl., 2015, **14**.

Peter V. Danchev
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
"Acad. G. Bonchev", str., bl. 8, 1113 Sofia
Bulgaria

E-mail: *danchev@math.bas.bg; pvdanchev@yahoo.com*

Received November 18, 2019

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 30–43
ISSN 1024–7696

Isohedral tilings for hyperbolic translation group of genus two which admit additional isometries

Elizaveta Zamorzaeva

Abstract. The paper continues articles [3, 4] where we derived all the Delone classes of isohedral tilings of the hyperbolic plane with disks for the translation group of genus two. The total number of these Delone classes is corrected to 118. The attention is drawn to the parametric diversity of these tilings. The emphasis is made on more symmetric tilings that are promising for further research and application. Isohedral tilings that admit additional isometries are analyzed, their adjacency diagrams are shown in figures.

Mathematics subject classification: Primary 52C20; Secondary 05B45.
Keywords and phrases: Isohedral tiling; hyperbolic plane; translation group of genus two; Delone class; adjacency symbol; adjacency diagram.

1 Introduction

Hyperbolic two-dimensional discrete isometry groups with compact fundamental domains were classified by A. M. Macbeath [1], their number is infinite. Further J. H. Conway [2] proposed for the notation of such groups an orbifold symbol, which is equivalent to Macbeath's symbol and is shorter. Any hyperbolic plane translation group is characterized by its genus and has the orbifold symbol by Conway $\circ\circ\cdots\circ$ where the number of circles is equal to the genus.

In the author's works [3, 4] all Delone classes of isohedral tilings of the hyperbolic plane with disks were derived for the translation group of genus two, with the orbifold symbol $\circ\circ$. Further these results can be used in several directions. Firstly they can be applied for obtaining some body-transitive tilings of the three-dimensional hyperbolic space. Secondly the factorization of the hyperbolic plane by this translation group yields Riemann surfaces of genus two with different metrics. A Riemann surface also can be obtained from a fundamental domain by gluing its boundary with isometries of the group. The most symmetric Riemann surfaces are known as regular maps [5].

Taking in consideration the above mentioned, in the present paper we are going to describe the isohedral tilings for the hyperbolic translation group of genus two in more details. In our investigation we basically adhere to ideas and methods which B. N. Delone (and coauthors) used for obtaining isohedral tilings of the Euclidean plane in [6, 7]. We consider tilings with disks as in the monograph [8], which is a more general setting. The number of the obtained Delone classes is corrected.

Remark that the method of obtaining fundamental isohedral tilings of the hyperbolic plane with compact fundamental domains developed by Z. Lučić and E. Molnár [9, 10] gives only a part of isohedral tilings for a hyperbolic translation group. Here we will compare some results obtained by application of the two methods. Also there exists one more method, based on Delauney–Dress symbols, of obtaining k-isohedral tilings of all three 2-dimensional spaces of constant curvature, which was developed by D. H. Huson (and coauthor) in [11, 12].

We are going to discuss parameters of the polygonal tiles and conditions the tiles must obey. Tilings that admit additional isometries are more appropriate for applications. We will draw attention to such tilings and will describe them more detailed.

2 Basic concepts and methods

A set W of closed topological disks in the plane is called a tiling of the plane if every point of the plane belongs to at least a disk and no two disks have an inner point in common. The disks of a tiling are called tiles.

A non-empty component of the intersection of two or more different tiles is called a vertex of the tiling or an edge of the tiling depending on whether it is a single point or not (then it is a curve). The boundary of a tile is divided by vertices of the tiling into curves that are edges of the tiling, so any tile may be considered as a curvilinear polygon.

Definition 1. Let W be a tiling of the hyperbolic plane with disks, G be a discrete isometry group of the hyperbolic plane with a compact (bounded) fundamental domain. The tiling W is called isohedral with respect to the group G if the group G maps the tiling W onto itself and G acts transitively on the set of the tiles.

Definition 2. Consider all possible pairs (W, G) where W is a tiling of the hyperbolic plane with disks which is isohedral with respect to a discrete hyperbolic isometry group G with a bounded fundamental domain. Two pairs (W, G) and (W', G') are said to belong to the same Delone class if there exists homeomorphic transformation φ of the plane which maps the tiling W onto the tiling W' and the relation $G = \varphi^{-1} G' \varphi$ holds.

Homeomeric types in [8] and equivariantly equivalent tilings in [11] are other terms used with the same meaning.

A Delone class (W, G) is called fundamental if the group G acts simply transitively (one time transitively) on the set of tiles of W. Any translation group admits only fundamental Delone classes (or fundamental tilings).

In articles [3, 4] we solved the problem of finding all fundamental Delone classes of isohedral tilings of the hyperbolic plane with disks for the translation group of genus two.

We followed the scheme of work [6]. First we solved Diophantine equations obtained from Euler theorem and get possible sets of valencies. We used the fact

that our isometry group is a group of translations and therefore its orbifold is a manifold without boundary and without singular points. Then we formed ordered cycles of valencies $(\alpha_1, \alpha_2, \ldots, \alpha_k)$, which are the same for each tile in an isohedral tiling.

Now we describe shortly the method of adjacency symbols proposed in [6] (see also [7] or [13]). Let (W, G) be a fundamental isohedral tiling of the plane with disks. Choose a tile and label consecutively all its edges with letters a, b,.... Then apply the group G to the labeled tile resulting in all the tiles being labeled. Now we form an adjacency symbol as follows. The letter which labels the first chosen edge stands first, the letter which labels the adjacent edge of the neighbor tile stands next, then the lower index indicates the valency of the end vertex of the first edge, after that we pass to the second consecutive edge, and so on. If the orientations of the initial and neighbor tiles are opposite, it is indicated with a bar over the second letter. In our case we deal only with translations, they preserve orientation, and no bars are needed.

An adjacency diagram is a polygonal tile where the vertices are labeled with their valencies and the paired edges are connected with arcs.

We generate all possible adjacency symbols for each appropriate equivalence class of ordered cycles. For each candidate in adjacency symbol we check if the condition of transition around a vertex is satisfied, for every vertex equivalence class. Further we must choose one representative among equivalent adjacency simbols. It is adjacency diagrams that help us to determine visually whether two adjacent symbols correspond to the same Delone class.

3 Correction of the number of all Delone classes and comparison of results obtained by two methods

Examine two Delone classes of isohedral tilings of the hyperbolic plane with 16-gons denoted in [4] as $16A14, 1$ and $16A14, 2$. Their adjacency symbols are $(ae_3bl_3cg_3dm_3ea_4fn_3gc_3hk_3io_4jp_3kh_3lb_3md_3nf_4oi_3pj_4)$ and, respectively, $(ai_3bk_3cm_3dg_3eo_4fp_3gd_3hl_3ia_4jn_3kb_3lh_3mc_3nj_4oe_3pf_4)$. Their adjacency diagrams are shown in Fig. 1. A given adjacency symbol can be obtained from the respective adjacency diagram as follows: the letter a labels the right bottom edge and we go round counter-clockwise.

In the right diagram a straight line is drawn connecting the left bottom vertex of valency 3 with the opposite vertex. Applying the reflection in this line to the right diagram, which corresponds to Delone class $16A14, 2$, we obtain just the left diagram, which corresponds to Delone class $16A14, 1$. It implies that two diagrams as well as two adjacency symbols determine the same Delone class and we denote this Delone class by $16A14$.

Thus we have shown that Delone classes $16A14, 1$ and $16A14, 2$ from the article [4] coincide. Now we can give the corrected formulation of Theorem 1 as follows.

Theorem 1. *For translation group of genus two there exist 118 Delone classes of*

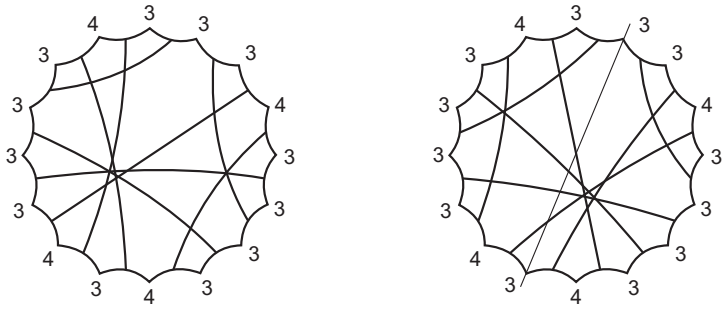


Figure 1. Two adjacency diagrams corresponding to the Delone class of isohedral tilings $16A14$

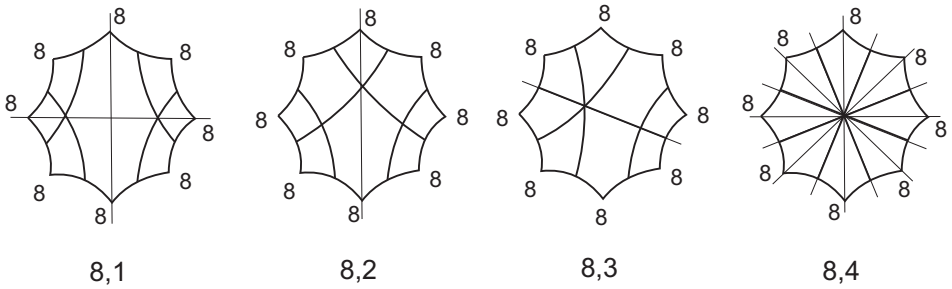


Figure 2. Adjacency diagrams for the Delone classes of isohedral tilings $8,1$, $8,2$, $8,3$, and $8,4$, with reflection axes

isohedral tilings of the hyperbolic plane with disks, and namely 4 classes with 8-gons, 18 classes with 10-gons, 31 classes with 12-gons, 39 classes with 14-gons, 20 classes with 16-gons, and 6 classes with 18-gons.

The results for 8-, 10- and 12-gons have been described in [3], the results for 14-, 16- and 18-gons have been described in [4] and the present article. The adjacency symbols have been given for all 118 Delone classes, the adjacency diagrams have been depicted for a part of them. An adjacency symbol contains information both on generators and relations of the group, it fully determines a Delone class and allows us to restore a tiling.

Now compare our results with results which can be obtained using the method developed in [9, 10]. Examine the case of the minimal number of edges (and vertices) of a fundamental domain for group $\circ\circ$, then the minimal number is 8. In this case the only fundamental domain can be obtained by cutting along the canonical graph with two loops, which gives an 8-gon with canonical pairings. It determines a Delone class that we denoted in [3] by $8,1$ (Fig. 2). However using our method we obtained 3 more Delone classes denoted by $8,2$, $8,3$ and $8,4$ (Fig. 2), the last one being well known (see [5]). Thus the methods [9, 10] yield only a part of possible Delone classes of isohedral tilings of the hyperbolic plane with disks for the translation group of genus two.

4 Parameters and conditions for polygonal tiles

We discuss the 'freedom' of a hyperbolic isometry group and its fundamental domains. The parameter space of a hyperbolic isometry group corresponds to the Teichmüller space of the group, which has the dimension $n = 6g-6+t+2r$, where g is the genus of the orbifold, t is the number of boundary components, r is the number of rotation centers [14]. According to the formula, the hyperbolic translation group of genus two $\circ\circ$ has the parameter space of dimension 6.

We state our results in the most general form, however tilings with convex polygons seem to be more interestig for further research and application. Assuming that isohedral tilings of the hyperbolic plane with convex polygons are considered, we determine which metrical conditions tiles and tilings must satisfy. These conditions are obtained from the facts that adjacent edges have the same length and the sum of angles at a vertex in the hyperbolic plane is equal to 2π. The pairs of adjacent edges appear directly in adjacent symbols. Equivalent vertices can be obtained using adjacency diagrams.

We begin with tilings with 8-gons, which belong to 4 Delone classes (Fig. 2). The sum of all 8 angles of a tile is equal to 2π, for each of 4 classes. For the Delone class $8,1$ with the adjacency symbol $(ac_8bd_8ca_8db_8eg_8fh_8ge_8hf_8)$, the conditions on edges are $a = c$, $b = d$, $e = g$, $f = h$. For the Delone class $8,2$ with the adjacency symbol $(ac_8be_8ca_8dg_8eb_8fh_8gd_8hf_8)$, the conditions on edges are $a = c$, $b = e$, $d = g$, $f = h$. For the Delone class $8,3$ with the adjacency symbol $(ac_8bf_8ca_8dg_8eh_8fb_8gd_8he_8)$, the conditions on edges are $a = c$, $b = f$, $d = g$, $e = h$. For the Delone class $8,4$ with the adjacency symbol $(ae_8bf_8cg_8dh_8ea_8fb_8gc_8hd_8)$, the conditions on edges are $a = e$, $b = f$, $c = g$, $d = h$. Particular cases of $8,1$ and $8,4$ are known with regular 8-gons, i.e. with each angle being equal to $\pi/4$ and the conditions on edges being $a = b = c = d = e = f = g = h$. However some less symmetric variants of these Delone classes can also be realized due to different choices of parameters of the group. Besides, with parameters of the group being fixed, there exists the choice of a fundamental domain.

For tilings by 10-gons with the set of valencies $10A : 3337777777$, all 6 Delone classes of isohedral tilings of the hyperbolic plane with disks are given both with their adjacency symbols and adjacency diagrams in [3]. For each of the Delone classes $10A1$, $10A2$, $10A3$, $10A4,1$, $10A4,2$, and $10A4,3$, the sum of 3 angles of valency 3 is equal to 2π, as well as the sum of 7 angles of valency 7 is equal to 2π. The pairs of edges with equal length can easily be read from their adjacency symbols.

For the set of valencies $10B : 4444666666$, the 6 Delone classes $10B1$, $10B2$, $10B3$, $10B4,1$, $10B4,2$, and $10B5$ are given with adjacency symbols in [3]. Adjacency diagrams can easily be obtained from the corresponding adjacency symbols. For each of these Delone classes, both the sum of 4 angles of valency 4 and the sum of 6 angles of valency 6 are equal to 2π. All the pairs of edges with equal length can be taken from their adjacency symbols.

For the set of valencies $10C : 5555555555$, the 6 Delone classes $10C,1$, $10C,2$, $10C,3$, $10C,4$, $10C,5$, and $10C,6$ are given with adjacency symbols in [3]. For each

of these Delone classes, 10 vertices fall into 2 equivalence classes, and the sum of 5 angles of valency 5 at vertices of the same equivalence class is equal to 2π. All the pairs of edges with equal length can be taken from the adjacency symbols.

For tilings by 12-gons with the set of valencies $12A$: 333333666666, 12 Delone classes of isohedral tilings are given with their adjacency symbols in [3]. For each Delone class, 6 vertices of valency 3 fall into 2 equivalence classes and all 6 vertices of valency 6 belong to one equivalence class. So the sums of 2 triples of angles at equivalent vertices of valency 3 and the sum of all 6 angles of valency 6 are equal to 2π. All the pairs of edges with equal length can be seen in adjacency symbols.

For the set of valencies $12B$: 333444455555, 13 Delone classes are given with their adjacency symbols in [3]. For each Delone class, the sum of 3 angles of valency 3 is equal to 2π, the sum of 4 angles of valency 4 is equal to 2π and the sum of 5 angles of valency 5 is equal to 2π. All the pairs of edges with equal length can be seen in adjacency symbols.

For the set of valencies $12C$: 444444444444, 6 Delone classes are given with their adjacency symbols in [3]. For each Delone class, all 12 vertices of valency 4 fall into 3 equivalence classes. So the 3 sums of 4 angles at equivalent vertices are equal to 2π. All the pairs of edges with equal length can be seen in adjacency symbols.

For tilings by 14-gons with the set of valencies $14A$: 33333333355555, 13 Delone classes of isohedral tilings are given with their adjacency symbols in [4]. For each Delone class, 9 vertices of valency 3 fall into 3 equivalence classes and all 5 vertices of valency 5 belong to one equivalence class. So the sums of 3 triples of angles at equivalent vertices of valency 3 and the sum of all 5 angles of valency 5 are equal to 2π. All the pairs of edges with equal length can be taken from adjacency symbols.

For the set of valencies $14B$: 33333344444444, 26 Delone classes are given with their adjacency symbols in [4]. For each Delone class, 6 vertices of valency 3 fall into 2 equivalence classes, as well as 8 vertices of valency 4 fall into 2 equivalence classes. So the two sums of 3 angles at equivalent vertices of valency 3 and the two sums of 4 angles at equivalent vertices of valency 4 are equal to 2π. All the pairs of edges with equal length can be taken from adjacency symbols.

For tilings by 16-gons with the set of valencies $16A$: 3333333333334444, where the letter A is used for convenience, 20 Delone classes of isohedral tilings are given with their adjacency symbols in [4]. For each Delone class, 12 vertices of valency 3 fall into 4 equivalence classes and all 4 vertices of valency 4 belong to one equivalence class. So the 4 sums of 3 angles at equivalent vertices of valency 3 and the sum of all 4 angles of valency 4 are equal to 2π. All the pairs of edges with equal length can be taken from adjacency symbols.

For tilings by 18-gons with the set of valencies 18 : 333333333333333333, 6 Delone classes of isohedral tilings are given with their adjacency symbols in [4]. For each Delone class, all 18 vertices of valency 3 fall into 6 equivalence classes. So the sums of 6 triples of angles at equivalent vertices are equal to 2π. All the pairs of edges with equal length can be taken from adjacency symbols.

5 Tiles admitting additional isometries

More symmetric tilings are of special interest for applications. Here we analyze adjacency diagrams and determine which tiles can admit additional isometries with taking into account their pairing isometries. It will be the first step in the direction of finding symmetric tilings among our list of tilings.

A thorough examination has been done yielding the following results. First we enumerate tilings where a tile can admit an isometry group (in prospect a stabilizer) of order greater than 2, with indication of adjacency diagrams. Also we list tilings where a tile can admit an isometry group of order 2. Each isohedral tiling is given with its adjacency symbol.

For tilings with 8-gons, their adjacency symbols have been given in Section 4. Each of the 4 Delone classes admits additional isometries if suitable parameters are chosen. For the Delone class $8,1$ a tile admits the isometry group of order 4 generated by a rotation of order 2 and a reflection, with suitable choice of parameters. For the Delone class $8,4$ a tile admits the isometry group of order 16 generated by a rotation of order 8 and a reflection if the tile is a regular 8-gon (with angle of $\pi/4$). For both Delone classes $8,2$ and $8,3$ tiles admit the isometry group of order 2 generated by respective reflections, with suitable choice of parameters. Adjacency diagrams for all the 4 Delone classes are given in Fig. 2, with the indication of admissible reflection axes.

Among tilings by 10-gons with the set $10A$, the 5 Delone classes of isohedal tilings $10A1$ with the adjacency symbol $(ad_3be_7cf_3da_7eb_3fc_7gi_7hj_7ig_7jh_7)$, $10A3$ with symbol $(ad_3bg_7ch_3da_7ei_7fj_7gb_3hc_7ie_7jf_7)$, $10A4,1$ with the adjacency symbol $(ae_3bg_7cj_7dh_3ea_7fi_7gb_3hd_7if_7jc_7)$, $10A4,2$ with the adjacency symbol $(ah_3bd_7ci_7db_3eg_7fj_7ge_3ha_7ic_7jf_7)$, and $10A4,3$ with the adjacency symbol $(ah_3bd_7cj_7db_3eg_7fi_7ge_3ha_7if_7jc_7)$ admit tiles with the isometry group of order 2 generated by respective reflection, for suitable choice of parameters.

Among tilings with the set $10B$, the Delone class $10B3$ with the adjacency symbol $(ac_4bg_4ca_6di_6ej_6fh_4gb_4hf_6id_6je_6)$ admits tiles with the isometry group of order 4 generated by a rotation of order 2 and a reflection if suitable parameters are chosen (Fig. 3). For the 4 Delone classes $10B1$ with the adjacency symbol $(ac_4be_4ca_6df_4eb_4fd_6gi_6hj_6ig_6jh_6)$, $10B2$, which has the adjacency symbol $(ac_4bf_4ca_6di_6eg_4fb_4ge_6hj_6id_6jh_6)$, $10B4,2$ with the adjacency symbol $(ad_4bh_6cf_4da_6ei_4fc_6gj_6hb_4ie_6jg_6)$, and $10B5$ with the adjacency symbol $(ag_4bh_6ci_4df_6ej_6fd_4ga_6hb_4ic_6je_6)$, tiles admit the isometry group of order 2 generated by respective reflection, for suitable choice of parameters.

Among tilings with the set $10C$, the 2 Delone classes $10C,1$ with the adjacency symbol $(ac_5bd_5ca_5db_5ej_5fh_5gi_5hf_5ig_5je_5)$ and $10C,4$ with the adjacency symbol $(ad_5be_5ch_5da_5eb_5fi_5gj_5hc_5if_5jg_5)$ admit tiles with the isometry group of order 4 generated by a rotation of order 2 and a reflecton, with suitable choice of parameters (Fig. 3). For the Delone class $10C,5$ with symbol $(ad_5bi_5cf_5da_5eh_5fc_5gj_5he_5ib_5jg_5)$ a tile admits the isometry group of order 10 generated by a rotation of order 5 and a reflection if the tile is a regular 10-gon (with angle of $2\pi/5$) (Fig. 3). For the Delone

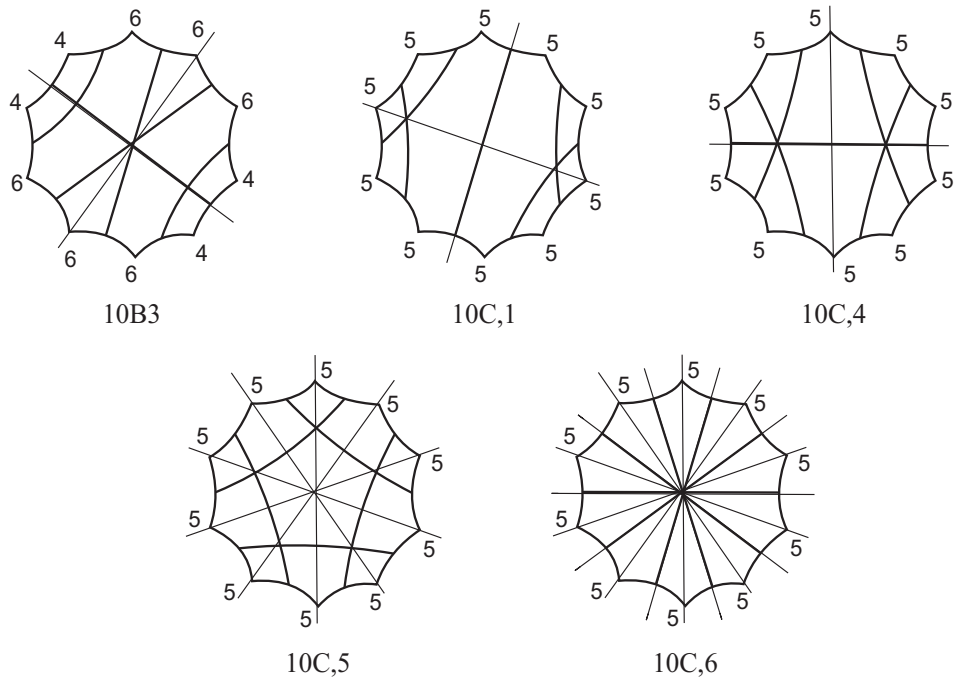


Figure 3. Adjacency diagrams for the Delone classes of isohedral tilings $10B3$, $10C,1$, $10C,4$, $10C,5$, and $10C,6$, with reflection axes

class $10C,6$ with symbol $(af_5bg_5ch_5di_5ej_5fa_5gb_5hc_5id_5je_5)$ a tile admits the isometry group of order 20 generated by a rotation of order 10 and a reflection if the tile also is a regular 10-gon (with angle of $2\pi/5$) (Fig. 3). The Delone class $10C,3$ with symbol $(ac_5bf_5ca_5dh_5ej_5fb_5gi_5hd_5ig_5je_5)$ admits a tile with the isometry group of order 2 generated by a reflection, for suitable choice of parameters.

Among tilings by 12-gons with the set $12A$, the 2 Delone classes $12A3$ with the adjacency symbol $(ad_3bh_3ci_3da_6ek_6fl_6gj_3hb_3ic_3jg_6ke_6lf_6)$ and $12A9,1$ with symbol $(ad_3be_6cf_3da_6eb_3fc_6gj_3hk_6il_3jg_6kh_3li_6)$ admit tiles with the isometry group of order 4 generated by a rotation of order 2 and a reflection, with suitable choice of parameters (Fig. 4). For the Delone class $12A9,4$ with the adjacency symbol $(aj_3be_6cl_3dg_6eb_3fi_6gd_3hk_6if_3ja_6kh_3lc_6)$ a tile admits the isometry group of order 12 generated by a rotation of order 6 and a reflection if the tile is a semiregular 12-gon with alternating angles of $\pi/3$ and $2\pi/3$ (Fig. 4). For the 6 Delone classes $12A1$ with symbol $(ad_3bf_3cg_3da_6eh_3fb_3gc_3he_6ik_6jl_6ki_6lj_6)$, $12A2$ with the adjacency symbol $(ad_3bg_3ch_3da_6ek_6fi_3gb_3hc_3if_6jl_6ke_6lj_6)$, $12A6$ with the adjacency symbol $(ae_3bg_3cj_6dh_3ea_6fk_3gb_3hd_6il_6jc_3kf_6li_6)$, $12A8$ with the adjacency symbol $(ae_3bh_3ck_6di_3ea_6fj_6gl_3hb_3id_6jf_6kc_3lg_6)$, $12A9,2$ with the adjacency symbol $(ad_3bg_6ch_3da_6ej_3fk_6gb_3hc_6il_3je_6kf_3li_6)$, and $12A9,3$ with the adjacency symbol $(aj_3be_6ch_3dk_6eb_3fi_6gl_3hc_6if_3ja_6kd_3lg_6)$, tiles admit the isometry group of order 2 generated by respective reflection, with suitable choice of parameters.

Among tilings with the set $12B$, the 7 Delone classes $12B1$ with the adja-

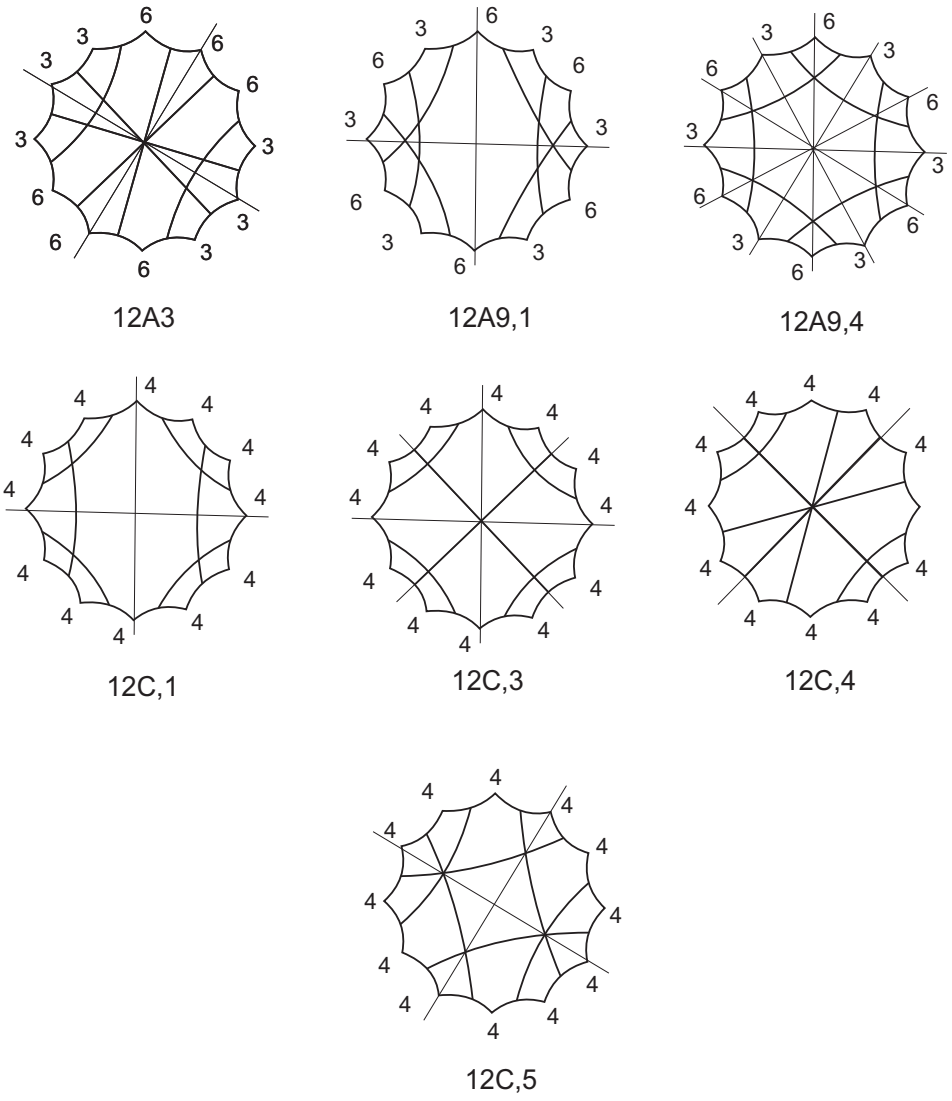


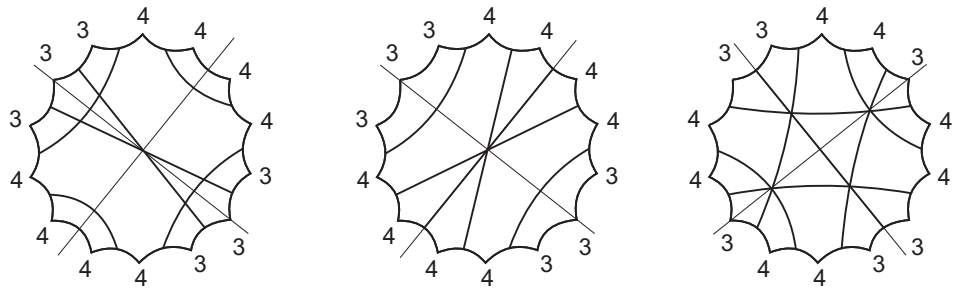
Figure 4. Adjacency diagrams for the Delone classes of isohedral tilings $12A3$, $12A9,1$, $12A9,4$, $12C,1$, $12C,3$, $12C,4$, and $12C,5$, with reflection axes

cency symbol $(ad_3be_4cf_3da_4eb_3fc_4gl_5hj_5ik_5jh_5ki_5lg_4)$, $12B6$ with the adjacency symbol $(ae_3bg_4ck_5dh_3ea_5fj_4gb_3hd_5il_4jf_5kc_4li_5)$, $12B7$ with the adjacency symbol $(ae_3bi_4cf_5dj_3ea_5fc_4gk_5hl_4ib_3jd_5kg_4lh_5)$, $12B8,2$ with the adjacency symbol $(aj_3be_4ck_5dh_4eb_3fi_5gl_4hd_5if_3ja_5kc_4lg_5)$, $12B9$ with the adjacency symbol $(ae_3bk_4ch_5dl_3ea_5fi_4gj_5hc_4if_5jg_4kb_3ld_5)$, $12B11$ with the adjacency symbol $(ad_3be_5cf_3da_5eb_3fc_5gi_4hk_4ig_5jl_4kh_4lj_5)$, and $12B12$ with the adjacency symbol $(ad_3bh_5ci_3da_5eg_4fk_4ge_5hb_3ic_5jl_4kf_4lj_5)$ admit tiles with the isometry group of order 2 generated by respective reflection if suitable parameters are chosen.

Among tilings with the set $12C$, the 3 Delone classes $12C,1$ with the adjacency symbol $(ac_4be_4ca_4df_4eb_4fd_4gi_4hk_4ig_4jl_4kh_4lj_4)$, $12C,4$ with the adjacency symbol $(ac_4bh_4ca_4dj_4ek_4fl_4gi_4hb_4ig_4jd_4ke_4lf_4)$ and $12C,5$ with the adjacency symbol $(ad_4bf_4ck_4da_4ei_4fb_4gj_4hl_4ie_4jg_4kc_4lh_4)$ admit tiles with the isometry group of order 4 generated by a rotation of order 2 and a reflection if suitable parameters are chosen (Fig. 4). For the class $12C,3$ with symbol $(ac_4bh_4ca_4df_4ek_4fd_4gi_4hb_4ig_4jl_4ke_4lj_4)$ a tile admits the isometry group of order 8 generated by a rotation of order 4 and a reflection if suitable parameters are chosen (Fig. 4). For the 2 Delone classes $12C,2$ with the adjacency symbol $(ac_4bf_4ca_4dj_4eg_4fb_4ge_4hk_4il_4jd_4kh_4li_4)$ and $12C,6$ with symbol $(ad_4bi_4cf_4da_4ej_4fc_4gk_4hl_4ib_4je_4kg_4lh_4)$ a tile admits the isometry group of order 2 generated by respective reflection, with suitable choice of parameters.

Among tilings by 14-gons with the set $14A$, the 7 Delone classes $14A1$ with the adjacency symbol $(aj_3be_3cg_3dh_3eb_3fi_3gc_3hd_3if_3ja_5km_5ln_5mk_5nl_5)$, $14A6$ with symbol $(ad_3bf_3cg_3da_5eh_3fb_3gc_3he_5il_3jm_5kn_3li_5mj_3nk_5)$, $14A7$ with the adjacency symbol $(ad_3bh_3ci_3da_5el_3fm_5gj_3hb_3ic_3jg_5kn_3le_5mf_3nk_5)$, $14A10,1$ with adjacency symbol $(ai_3bm_3cf_3dj_5ek_3fc_3gl_5hn_3ia_5jd_3ke_5lg_3mb_3nh_5)$, $14A10,2$ with adjacency symbol $(ak_3bf_3cm_3dh_5en_3fb_3gj_5hd_3il_5jg_3ka_5li_3mc_3ne_5)$, $14A10,3$ with adjacency symbol $(ak_3bm_3cf_3dh_5ei_3fc_3gl_5hd_3ie_5jn_3ka_5lg_3mb_3nj_5)$, and $14A11$ with adjacency symbol $(af_3bh_3ck_3dm_5ei_3fa_5gl_3hb_3ie_5jn_3kc_3lg_5md_3nj_5)$ admit tiles with the isometry group of order 2 generated by respective reflection if suitable parameters are chosen.

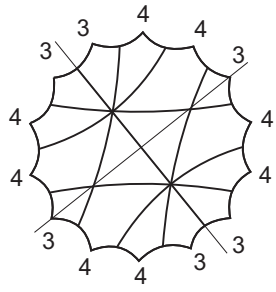
Among tilings with the set $14B$, the 9 Delone classes $14B3,1$ with the adjacency symbol $(ad_3bi_3cj_3da_4eg_4fm_4ge_4hk_3ib_3jc_3kh_4ln_4mf_4nl_4)$, $14B3,2$ with symbol $(ad_3bi_3cj_3da_4el_4fm_4gn_4hk_3ib_3jc_3kh_4le_4mf_4ng_4)$, $14B10,1$ with the adjacency symbol $(af_3bi_3cl_4dg_4ej_3fa_4gd_4hm_3ib_3je_4kn_4lc_3mh_4nk_4)$, $14B10,2$ with the symbol $(af_3bi_3cl_4dn_4ej_3fa_4gk_4hm_3ib_3je_4kg_4lc_3mh_4nd_4)$, $14B10,3$ with adjacency symbol $(am_3bi_3ce_4dg_4ec_3fh_4gd_4hf_3ib_3jl_4kn_4lj_3ma_4nk_4)$, $14B10,4$ with the adjacency symbol $(am_3bi_3ce_4dk_4ec_3fh_4gn_4hf_3ib_3jl_4kd_4lj_3ma_4ng_4)$, $14B10,5$ with the symbol $(am_3bi_3ce_4dn_4ec_3fh_4gk_4hf_3ib_3jl_4kg_4lj_3ma_4nd_4)$, $14B13,1$ with adjacency symbol $(ad_3be_4cf_3da_4eb_3fc_4gn_4hk_3il_4jm_3kh_4li_3mj_4ng_4)$, and $14B13,5$ with the symbol $(ak_3be_4cm_3dh_4eb_3fj_4gn_4hd_3il_4jf_3ka_4li_3mc_4ng_4)$ admit tiles with the isometry group of order 4 generated by a rotation of order 2 and a reflection if suitable parameters are chosen (Fig. 5). For the 5 Delone classes $14B1$ with the adjacency symbol $(ad_3bf_3cg_3da_4eh_3fb_3gc_3he_4ik_4jm_4ki_4ln_4mj_4nl_4)$, $14B2$ with symbol $(ad_3bg_3ch_3da_4el_4fi_3gb_3hc_3if_4jm_4kn_4le_4mj_4nk_4)$, $14B11,2$ with adjacency symbol $(af_3bi_4cm_3dg_4ej_3fa_4gd_3hl_4ib_3je_4kn_4lh_3mc_4nk_4)$, $14B11,4$ with adjacency symbol $(am_3bi_4cf_3dg_4eh_3fc_4gd_3he_4ib_3jl_4kn_4lj_3ma_4nk_4)$, and $14B12$ with adjacency symbol



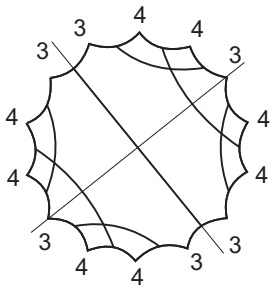
14B3,1

14B3,2

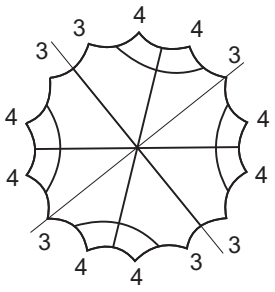
14B10,1



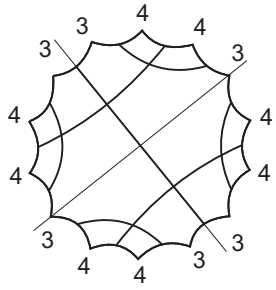
14B10,2



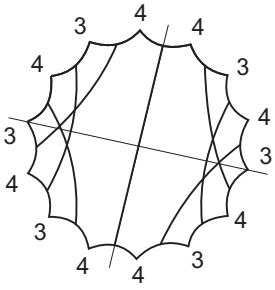
14B10,3



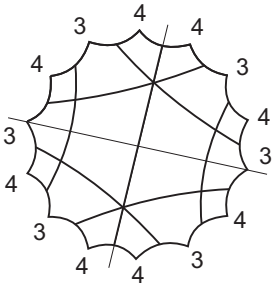
14B10,4



14B10,5



14B13,1



14B13,5

Figure 5. Adjacency diagrams for the Delone classes of isohedral tilings $14B3,1$, $14B3,2$, $14B10,1$, $14D10,2$, $14D10,3$, $14B10,4$, $14B10,5$, $14B13,1$, and $14B13,5$, with reflection axes

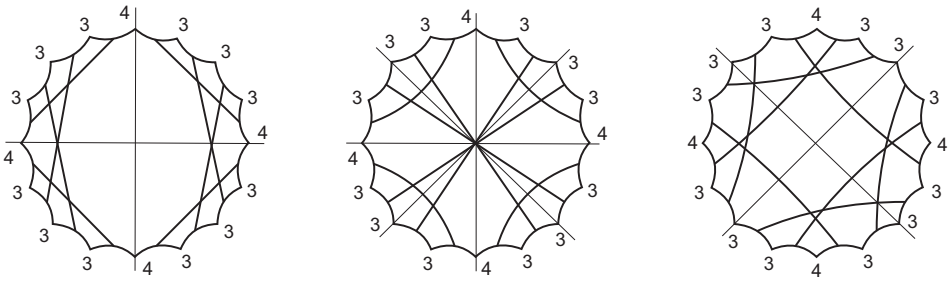


Figure 6. Adjacency diagrams for the Delone classes of isohedral tilings $16A16,1$, $16A16,2$ and $16A16,3$, with reflection axes

$(am_3be_4ck_3dg_4eb_3fl_4gd_3hj_4in_4jh_3kc_4lf_3ma_4ni_4)$, a tile admits the isometry group of order 2 generated by respective reflection, for suitable choice of parameters. For the Delone class $14B13,6$ with symbol $(ak_3bl_4cf_3dh_4ei_3fc_4gn_4hd_3ie_4jm_3ka_4lb_3mj_4ng_4)$ a tile admits the isometry group of order 2 generated by a rotation of order 2, with suitable choice of parameters.

Among tilings by 16-gons, the 2 Delone classes $16A16,1$ with the adjacency symbol $(ad_3bf_3cg_3da_4eh_3fb_3gc_3he_4il_3jn_3ko_3li_4mp_3nj_3ok_3pm_4)$ and $16A16,3$ with symbol $(al_3bf_3co_3di_4ep_3fb_3gk_3hm_4id_3jn_3kg_3la_4mh_3nj_3oc_3pe_4)$ admit tiles with the isometry group of order 4 generated by a rotation of order 2 and a reflection if suitable parameters are chosen (Fig. 6). For the Delone class $16A16,2$ with symbol $(ad_3bj_3ck_3da_4eh_3fn_3go_3he_4il_3jb_3kc_3li_4mp_3nf_3og_3pm_4)$ a tile admits the isometry group of order 8 generated by a rotation of order 4 and a reflection, with suitable choice of parameters (Fig. 6). For the 8 Delone classes $16A1$ with symbol $(ao_3bk_3cf_3dh_3ei_3fc_3gj_3hd_3ie_3jg_3kb_3ln_4mp_4nl_3oa_4pm_4)$, $16A2$ with the adjacency symbol $(ao_3be_3ch_3di_3eb_3fn_3gj_3hc_3id_3jg_3km_4lp_4mk_3nf_3oa_4pl_4)$, $16A3$ with symbol $(ao_3bf_3ch_3dm_3ei_3fb_3gn_3hc_3ie_3jl_4kp_4lj_3md_3ng_3oa_4pk_4)$, $16A4$ with symbol $(ao_3bf_3cl_3dh_3em_3fb_3gn_3hd_3ik_4jp_4ki_3lc_3me_3ng_3oa_4pj_4)$, $16A6$ with adjacency symbol $(aj_3be_3cg_3dh_3eb_3fi_3gc_3hd_3if_3ja_4kn_3lo_4mp_3nk_4ol_3pm_4)$, $16A10$ with symbol $(an_3be_3ci_3dj_3eb_3fm_4gp_3hk_3ic_3jd_3kh_3lo_4mf_3na_4ol_3pg_4)$, $16A12$ with symbol $(af_3bk_3co_3dh_3el_3fa_4gm_3hd_3in_4jp_3kb_3le_3mg_4ni_3oc_3pj_4)$, and $16A15$ with symbol $(am_3bg_3co_3di_3ek_4fp_3gb_3hl_3id_3jn_4ke_3lh_3ma_4nj_3oc_3pf_4)$ tiles admit the isometry group of order 2 generated by respective reflection, for suitable choice of parameters. For the Delone class $16A5$ which has the adjacency symbol $(ao_3be_3ck_3dl_3eb_3fn_3gi_4hp_4ig_3jm_3kc_3ld_3mj_3nf_3oa_4ph_4)$ a tile admits the isometry group of order 2 generated by a rotation of order 2, with suitable choice of parameters.

Among tilings with 18-gons, the 3 Delone classes $18,1$ with the adjacency symbol $(ad_3bf_3cg_3da_3eh_3fb_3gc_3he_3ir_3jm_3ko_3lp_3mj_3nq_3ok_3pl_3qn_3ri_3)$, $18,2$ with symbol $(ad_3bg_3ch_3da_3en_3fi_3gb_3hc_3if_3jm_3kp_3lq_3mj_3ne_3or_3pk_3ql_3ro_3)$ and $18,5$ with symbol $(ad_3bk_3cl_3da_3eh_3fo_3gp_3he_3ir_3jm_3kb_3lc_3mj_3nq_3of_3pg_3qn_3ri_3)$ admit tiles with the isometry group of order 4 generated by a rotation of order 2 and a reflection if

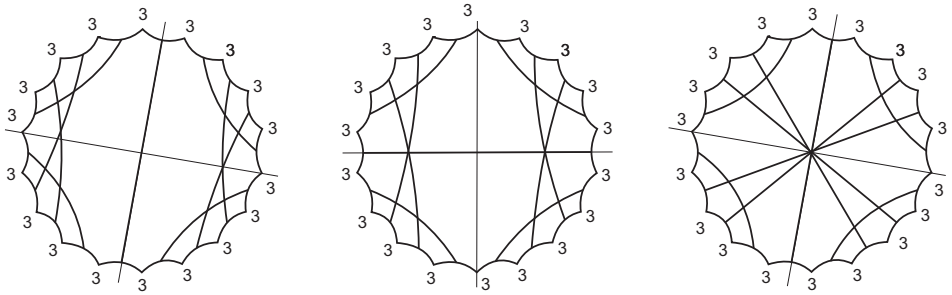


Figure 7. Adjacency diagrams for the Delone classes of isohedral tilings 18, 1, 18, 2 and 18, 5, with reflection axes

suitable parameters are chosen (Fig. 7). For the 3 Delone classes 18, 3 with symbol $(ad_3bh_3ci_3da_3em_3fp_3gj_3hb_3ic_3jg_3ko_3lq_3me_3nr_3ok_3pf_3ql_3rn_3)$, 18, 4 with symbol $(ad_3bi_3cj_3da_3em_3fo_3gq_3hk_3ib_3jc_3kh_3lp_3me_3nr_3of_3pl_3qg_3rn_3)$ and 18, 6 with symbol $(af_3bi_3cn_3dq_3ej_3fa_3gl_3ho_3ib_3je_3kp_3lg_3mr_3nc_3oh_3pk_3qd_3rm_3)$, a tile admits the isometry group of order 2 generated by respective reflection, with suitable choice of parameters.

Altogether there are 78 Delone classes of isohedral tilings of the hyperbolic plane with disks for translation group of genus two that admit tilings with additional isometries. The analysis of adjacency diagrams has yielded that 29 Delone classes admit tiles with the isometry group of order greater than 2, their adjacency diagrams are shown in Fig. 2–7. The adjacency diagrams are given for all the 78 Delone classes.

Acknowledgements. The work was partially supported by the SCSTD Project 15.817.02.26F.

References

[1] Macbeath A. M. *The classification of non-Euclidean plane crystallographic groups.* Canad. J. Math., 1967, **19**, 1192–1205.

[2] Conway J. H. *The orbifold notation for surface groups.* Groups, Combinatorics and Geometry. London Math. Soc. Lecture Note Series 165, Cambridge University Press, 1992, 438–447.

[3] Zamorzaeva E. *Isohedral tilings by 8-, 10- and 12-gons for hyperbolic translation group of genus two.* Bul. Acad. Ştiinţe Repub. Moldova. Mat., 2018, N 2, 74–84.

[4] Zamorzaeva E. *Isohedral tilings by 14-, 16- and 18-gons for hyperbolic translation group of genus two.* Bul. Acad. Ştiinţe Repub. Moldova. Mat., 2019, N 1, 91–102.

[5] Coxeter H. S. M., Moser W. O. J. *Generators and Relations for Discrete Groups,* 4th ed., Springer, Berlin–Heidelberg–New York, 1980.

[6] Delone B. N. *Theory of planigons.* Izvestiya Akad. Nauk SSSR. Ser. Matem., 1959, **23**, 365–386 (in Russian).

[7] Delone B. N., Dolbilin N. P., Shtogrin M. I. *Combinatorial and metric theory of planigons.* Tr. Mat. Inst. Steklov Akad. Nauk SSSR, 1978, **148**, 109–140 (in Russian). English translation: Proc. Steklov Inst. Math., 1980, **4**, 111–141.

[8] Grünbaum B. and Shephard G. C. *Tilings and Patterns,* Freeman, New York, 1987.

[9] Lučić Z. and Molnár E. *Combinatorial classification of fundamental domains of finite area for planar discontinuous isometry groups.* Archiv der Matematik, 1990, **54**, 511–520.

[10] Lučić Z. and Molnár E. *Fundamental domains for planar discontinuous groups and uniform tilings.* Geom. Dedicata, 1991, **40**, 125–143.

[11] Huson D. H. *The generation and classification of tile-k-transitive tilings of the Euclidean plane, the sphere and the hyperbolic plane.* Geom. Dedicata, 1993, **47**, 269–296.

[12] Balke L. and Huson D. H. *Two-dimensional groups, orbifolds and tilings.* Geom. Dedicata, 1996, **60**, 89–106.

[13] Zamorzaeva E. *On isohedral tilings of hyperbolic manifolds.* Analele Ştiinţifice ale Universităţii 'Al. I. Cuza' Iaşi, s. 1a, Matematică, 1997, **59**, fasc.1, 81–88.

[14] Zieschang H., Vogt E., Coldewey H. D., *Surfaces and Planar Discontinues Groups.* Lecture Notes in Mathematics, 835, Springer–Verlag, New York, 1980.

Elizaveta Zamorzaeva
'Vladimir Andrunachievici'
Institute of Mathematics and Computer Science,
5 Academiei str., Chişinău, MD-2028
Republic of Moldova
E-mail: *zamorzaeva@yahoo.com*

Received January 22, 2020

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 44–61
ISSN 1024–7696

Interior angle sums of geodesic triangles in $\mathbf{S}^2\times\mathbf{R}$ and $\mathbf{H}^2\times\mathbf{R}$ geometries

Jenő Szirmai

Abstract. In the present paper we study $\mathbf{S}^2\times\mathbf{R}$ and $\mathbf{H}^2\times\mathbf{R}$ geometries, which are homogeneous Thurston 3-geometries. We analyse the interior angle sums of geodesic triangles in both geometries and we prove that in $\mathbf{S}^2\times\mathbf{R}$ space it can be larger than or equal to π and in $\mathbf{H}^2\times\mathbf{R}$ space the angle sums can be less than or equal to π. This proof is a new direct approach to the issue and it is based on the projective model of $\mathbf{S}^2\times\mathbf{R}$ and $\mathbf{H}^2\times\mathbf{R}$ geometries described by E. Molnár in [7].

Mathematics subject classification: 53A20, 53A35, 52C35, 53B20.
Keywords and phrases: Thurston geometries, $\mathbf{S}^2\times\mathbf{R}$, $\mathbf{H}^2\times\mathbf{R}$ geometries, geodesic triangles, interior angle sum.

1 Introduction

A geodesic triangle in Riemannian geometry and more generally in metric geometry is a figure consisting of three different points together with the pairwise-connecting geodesic curves. The points are known as the vertices, while the geodesic curve segments are known as the sides of the triangle.

In the geometries of constant curvature $\mathbf{E}^3$, $\mathbf{H}^3$, $\mathbf{S}^3$ the well-known sums of the interior angles of geodesic triangles characterize the space. It is related to the Gauss-Bonnet theorem which states that the integral of the Gauss curvature on a compact 2-dimensional Riemannian manifold M is equal to $2\pi\chi(M)$ where $\chi(M)$ denotes the Euler characteristic of M. This theorem has a generalization to any compact even-dimensional Riemannian manifold (see e.g.[2, 5]).

Remark 1. In the Thurston spaces translation curves can be introduced in a natural way (see [7]). These curves are simpler than geodesics and differ from them in **Nil**, $\widetilde{\mathbf{SL}_2\mathbf{R}}$ and **Sol** geometries. In $\mathbf{E}^3$, $\mathbf{S}^3$, $\mathbf{H}^3$, $\mathbf{S}^2\times\mathbf{R}$ and $\mathbf{H}^2\times\mathbf{R}$ geometries the mentioned curves coincide with each other ([1, 4, 15, 21]).

In [4] we investigated the angle sums of translation and geodesic triangles in $\widetilde{\mathbf{SL}_2\mathbf{R}}$ geometry and proved that the possible sum of the interior angles in a translation triangle must be greater than or equal to π. However, in geodesic triangles this sum is less, greater or equal to π.

In [20] we considered the analogous problem for geodesic triangles in **Nil** geometry and proved that the sum of the interior angles of geodesic triangles in **Nil** space is larger than, less than or equal to π. In [1] K. Brodaczewska showed that sum of

the interior angles of translation triangles of the **Nil** space is larger than or equal to π.

In [21] we studied the interior angle sums of *translation triangles* in **Sol** geometry and proved that the possible sum of the interior angles in a translation triangle must be greater than or equal to π. Further interesting properties of translation triangles and tetrahedra are described in [15].

However, in $\mathbf{S}^2\times\mathbf{R}$, $\mathbf{H}^2\times\mathbf{R}$ and **Sol** Thurston geometries there are no results concerning the angle sums of *geodesic triangles*. Therefore, it is interesting to study this question in the above three geometries.

In the present paper, we are interested in *geodesic triangles* in $\mathbf{S}^2\times\mathbf{R}$ and $\mathbf{H}^2\times\mathbf{R}$ spaces [13, 22].

In Section 2 we describe the projective model and the isometry group of the considered geometries, moreover, we give an overview about its geodesic curves. *In Section 3 we study the* $\mathbf{S}^2\times\mathbf{R}$ *and* $\mathbf{H}^2\times\mathbf{R}$ *geodesic triangles and their properties. We analyse the interior angle sums of geodesic triangles in both geometries and we prove that in* $\mathbf{S}^2\times\mathbf{R}$ *space it can be larger than or equal to* π *and in* $\mathbf{H}^2\times\mathbf{R}$ *space the angle sums can be less than or equal to* π*. This is a consequence of comparison theorems in Riemannian geometry (Toponogov and Alexandrov's theorems, see [3]), since the sectional curvature of* $\mathbf{S}^2\times\mathbf{R}$ *is non-negative and the sectional curvature of* $\mathbf{H}^2\times\mathbf{R}$ *is non-positive.*

Our new proof gives a new direct approach to the issue and it is based on the projective model of $\mathbf{S}^2\times\mathbf{R}$ *and* $\mathbf{H}^2\times\mathbf{R}$ *geometries described by E. Molnár in [7].*

2 Projective models of $\mathbf{H}^2\times\mathbf{R}$ and $\mathbf{S}^2\times\mathbf{R}$ spaces

E. Molnár has shown in [7] that the homogeneous 3-spaces have a unified interpretation in the projective 3-sphere $\mathcal{PS}^3(\mathbf{V}^4, \boldsymbol{V}_4, \mathbf{R})$. In our work we shall use this projective model of $\mathbf{S}^2\times\mathbf{R}$ and $\mathbf{H}^2\times\mathbf{R}$ geometries. The Cartesian homogeneous coordinate simplex is $E_0(\mathbf{e}_0)$, $E_1^\infty(\mathbf{e}_1)$, $E_2^\infty(\mathbf{e}_2)$, $E_3^\infty(\mathbf{e}_3)$, ($\{\mathbf{e}_i\}\subset\mathbf{V}^4$ with the unit point $E(\mathbf{e}=\mathbf{e}_0+\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3)$), which is distinguished by an origin E_0 and by the ideal points of coordinate axes, respectively. Moreover, $\mathbf{y}=c\mathbf{x}$ with $0<c\in\mathbf{R}$ (or $c\in\mathbf{R}\setminus\{0\}$) defines a point $(\mathbf{x})=(\mathbf{y})$ of the projective 3-sphere $\mathcal{PS}^3$ (or that of the projective space $\mathcal{P}^3$ where opposite rays $(\mathbf{x})$ and $(-\mathbf{x})$ are identified). The dual system $\{(\boldsymbol{e}^i)\}\subset\boldsymbol{V}_4$ describes the simplex planes, especially the plane at infinity $(\boldsymbol{e}^0)=E_1^\infty E_2^\infty E_3^\infty$, and generally, $\boldsymbol{v}=\boldsymbol{u}\frac{1}{c}$ defines a plane $(\boldsymbol{u})=(\boldsymbol{v})$ of $\mathcal{PS}^3$ (or that of $\mathcal{P}^3$). Thus $0=\mathbf{x}\boldsymbol{u}=\mathbf{y}\boldsymbol{v}$ defines the incidence of point $(\mathbf{x})=(\mathbf{y})$ and plane $(\boldsymbol{u})=(\boldsymbol{v})$, as $(\mathbf{x})\mathrm{I}(\boldsymbol{u})$ also denotes it. Thus $\mathbf{S}^2\times\mathbf{R}$ can be visualized in the affine 3-space $\mathbf{A}^3$ (so in $\mathbf{E}^3$) as well.

2.1 Geodesic curves in $\mathbf{S}^2\times\mathbf{R}$ space

In this section we recall the important notions and results from the papers [7, 11, 14, 16, 17].

The well-known infinitesimal arc-length square at any point of $\mathbf{S}^2\times\mathbf{R}$ is as follows

$$(ds)^2=\frac{(dx)^2+(dy)^2+(dz)^2}{x^2+y^2+z^2}. \tag{2.1}$$

We shall apply the usual geographical coordiantes (ϕ,θ), $(-\pi<\phi\leq\pi,\ -\frac{\pi}{2}\leq\theta\leq\frac{\pi}{2})$ of the sphere with the fibre coordinate $t\in\mathbf{R}$. We describe points in the above coordinate system in our model by the following equations:

$$x^0=1,\quad x^1=e^t\cos\phi\cos\theta,\quad x^2=e^t\sin\phi\cos\theta,\quad x^3=e^t\sin\theta. \tag{2.2}$$

Then we have $x=\frac{x^1}{x^0}=x^1$, $y=\frac{x^2}{x^0}=x^2$, $z=\frac{x^3}{x^0}=x^3$, i.e. the usual Cartesian coordinates. We obtain by [7] that in this parametrization the infinitesimal arc-length square at any point of $\mathbf{S}^2\times\mathbf{R}$ is the following

$$(ds)^2=(dt)^2+(d\phi)^2\cos^2\theta+(d\theta)^2. \tag{2.3}$$

The geodesic curves of $\mathbf{S}^2\times\mathbf{R}$ are generally defined as having locally minimal arc length between their any two (near enough) points. The equation systems of the parametrized geodesic curves $\gamma(t(\tau),\phi(\tau),\theta(\tau))$ in our model can be determined by the general theory of Riemann geometry (see [5,17]).

Then by (2.2) we get with $c=\sin v$, $\omega=\cos v$ the equation systems of a geodesic curve, visualized in Fig. 3 in our Euclidean model:

$$\begin{gathered}x(\tau)=e^{\tau\sin v}\cos(\tau\cos v),\\ y(\tau)=e^{\tau\sin v}\sin(\tau\cos v)\cos u,\\ z(\tau)=e^{\tau\sin v}\sin(\tau\cos v)\sin u,\\ -\pi<u\leq\pi,\quad -\frac{\pi}{2}\leq v\leq\frac{\pi}{2}.\end{gathered} \tag{2.4}$$

Definition 1. The distance $d(P_1,P_2)$ between the points P_1 and P_2 is defined by the arc length of the shortest geodesic curve from P_1 to P_2.

2.2 Geodesic curves of $\mathbf{H}^2\times\mathbf{R}$ geometry

In this section we recall the important notions and results from the papers [7, 12,18].

The points of $\mathbf{H}^2\times\mathbf{R}$ space, forming an open cone solid in the projective space $\mathcal{P}^3$, are the following:

$$\mathbf{H}^2\times\mathbf{R}:=\left\{X(\mathbf{x}=x^i\mathbf{e}_i)\in\mathcal{P}^3:-(x^1)^2+(x^2)^2+(x^3)^2<0<x^0,\ x^1\right\}.$$

In this context E. Molnár [7] has derived the infinitesimal arc-length square at any point of $\mathbf{H}^2\times\mathbf{R}$ as follows

$$\begin{gathered}(ds)^2=\frac{1}{(-x^2+y^2+z^2)^2}\cdot[(x)^2+(y)^2+(z)^2](dx)^2+\\ +2dxdy(-2xy)+2dxdz(-2xz)+[(x)^2+(y)^2-(z)^2](dy)^2+\\ +2dydz(2yz)+[(x)^2-(y)^2+(z)^2](dz)^2.\end{gathered} \tag{2.5}$$

This becomes simpler in the following special (cylindrical) coordiantes (t, r, α), $(r \geq 0,\ -\pi < \alpha \leq \pi)$ with the fibre coordinate $t \in \mathbf{R}$. We describe points in our model by the following equations:

$$x^0 = 1, \quad x^1 = e^t \cosh r, \quad x^2 = e^t \sinh r \cos\alpha, \quad x^3 = e^t \sinh r \sin\alpha. \tag{2.6}$$

Then we have $x = \frac{x^1}{x^0} = x^1$, $y = \frac{x^2}{x^0} = x^2$, $z = \frac{x^3}{x^0} = x^3$, i.e. the usual Cartesian coordinates. We obtain by [7] that in this parametrization the infinitesimal arc-length square by (2.1) at any point of $\mathbf{H}^2 \times \mathbf{R}$ is the following

$$(ds)^2 = (dt)^2 + (dr)^2 + \sinh^2 r (d\alpha)^2. \tag{2.7}$$

The geodesic curves of $\mathbf{H}^2 \times \mathbf{R}$ are generally defined as having locally minimal arc length between their any two (near enough) points. The equation systems of the parametrized geodesic curves $\gamma(t(\tau), r(\tau), \alpha(\tau))$ in our model can be determined by the general theory of Riemann geometry:

By (2.5) the second order differential equation system of the $\mathbf{H}^2 \times \mathbf{R}$ geodesic curve is the following [18]:

$$\ddot{\alpha} + 2\coth(r)\ \dot{r}\dot{\alpha} = 0,\ \ddot{r} - \sinh(r)\cosh(r)\dot{\alpha}^2 = 0,\ \ddot{t} = 0, \tag{2.8}$$

from which we get first a line as "geodesic hyperbola" on our model of $\mathbf{H}^2$ times a component on $\mathbf{R}$ each running with constant velocity c and ω, respectively:

$$t = c \cdot \tau, \quad \alpha = 0, \quad r = \omega \cdot \tau, \quad c^2 + \omega^2 = 1. \tag{2.9}$$

We can assume that the starting point of a geodesic curve is $(1, 1, 0, 0)$, because we can transform a curve into an arbitrary starting point, moreover, unit velocity with "geographic" coordinates (u, v) can be assumed:

$$r(0) = \alpha(0) = t(0) = 0; \quad \dot{t}(0) = \sin v,\ \dot{r}(0) = \cos v \cos u, \dot{\alpha}(0) = \cos v \sin u;$$
$$-\pi < u \leq \pi,\ -\frac{\pi}{2} \leq v \leq \frac{\pi}{2}.$$

Then by (2.6) we get with $c = \sin v$, $\omega = \cos v$ the equation systems of a geodesic curve, visualized in Fig. 8 in our Euclidean model [18]:

$$\begin{aligned} x(\tau) &= e^{\tau \sin v} \cosh(\tau \cos v), \\ y(\tau) &= e^{\tau \sin v} \sinh(\tau \cos v) \cos u, \\ z(\tau) &= e^{\tau \sin v} \sinh(\tau \cos v) \sin u, \\ -\pi < u \leq \pi, &\quad -\frac{\pi}{2} \leq v \leq \frac{\pi}{2}. \end{aligned} \tag{2.10}$$

Definition 2. The distance $d(P_1, P_2)$ between the points P_1 and P_2 is defined by the arc length of the geodesic curve from P_1 to P_2.

Remark 2. $\mathbf{S}^2 \times \mathbf{R}$ and $\mathbf{H}^2 \times \mathbf{R}$ are affine metric spaces (affine-projective spaces – in the sense of the unified formulation of [7]). Therefore their linear, affine, unimodular, etc. transformations are defined as those of the embedding affine space.

3 Geodesic triangles

We consider 3 points A_1, A_2, A_3 in the projective model of X space (see Section 2) ($X \in \{\mathbf{S}^2\times\mathbf{R}, \mathbf{H}^2\times\mathbf{R}\}$. The *geodesic segments* a_k connecting the points A_i and A_j ($i<j,\ i,j,k \in \{1,2,3\}, k \neq i,j$) are called sides of the *geodesic triangle* with vertices A_1, A_2, A_3 (see Fig. 1, 2).

In Riemannian geometries the infinitesimal arc-lenght square (see (2.1) and (2.5)) is used to define the angle θ between two geodesic curves. If their tangent vectors at their common point are $\mathbf{u}$ and $\mathbf{v}$ and g_{ij} are the components of the metric tensor then

$$\cos(\theta) = \frac{u^i g_{ij} v^j}{\sqrt{u^i g_{ij} u^j \ v^i g_{ij} v^j}} \tag{3.1}$$

It is clear by the above definition of the angles and by the infinitesimal arc-lenght squares that the angles are the same as the Euclidean ones at the starting point of the geodesics.

Considering a geodesic triangle $A_1A_2A_3$ we can assume by the homogeneity of the considered geometries that one of its vertex coincides with the point $A_1 = (1,1,0,0)$ and the other two vertices are $A_2 = (1,x_2,y_2,z_2)$ and $A_3 = (1,x_3,y_3,z_3)$.

We will consider the *interior angles* of geodesic triangles that are denoted at the vertex A_i by ω_i ($i \in \{1,2,3\}$). We note here that the angle of two intersecting geodesic curves depends on the orientation of their tangent vectors.

3.1 Interior angle sums in $\mathbf{S}^2\times\mathbf{R}$ geometry

In order to determine the interior angles of a geodesic triangle $A_1A_2A_3$ and its interior angle sum $\sum_{i=1}^{3}(\omega_i)$, we define *isometric transformations* $\mathbf{T}_{A_i}^{\mathbf{S}^2\times\mathbf{R}}$ ($i \in \{2,3\}$, as elements of the isometry group of $\mathbf{S}^2\times\mathbf{R}$ geometry that maps the A_i onto A_1). Let the isometry $\mathbf{T}_{A_2}^{\mathbf{S}^2\times\mathbf{R}}$ be given by the composition of some special types of $\mathbf{S}^2\times\mathbf{R}$ isometries which transforms a fixed $A_2 = (1,x_2,y_2,z_2)$ point of $\mathbf{S}^2\times\mathbf{R}$ into $(1,1,0,0)$ (up to a positive determinant factor):

$\mathcal{T} = (\mathbf{Id}., T)$ is a fibre translation,

$$\begin{gathered} A_2 = (1,x_2,y_2,z_2) \to A_2^{\mathcal{T}} = (1,x_2',y_2',z_2') = \\ = A_2^{\mathcal{T}} = \Big(1, \frac{x_2}{\sqrt{x_2^2+y_2^2+z_2^2}}, \frac{y_2}{\sqrt{x_2^2+y_2^2+z_2^2}}, \frac{z_2}{\sqrt{x_2^2+y_2^2+z_2^2}}\Big). \end{gathered} \tag{3.2}$$

($A_2^{\mathcal{T}}$ has 0 fibre coordinate). $\mathcal{R}_x = (\mathbf{R_x}, 0)$ is a special rotation about x axis with 0 fibre translation which moves the point $(1,x_2',y_2',z_2')$ into the $[x,y]$ plane.

$$\begin{gathered} A_2^{\mathcal{T}} = (1,x_2',y_2',z_2') \to A_2^{\mathcal{T}\mathcal{R}_x} = (1,x_2'',y_2'',0) = \\ = A_2^{\mathcal{T}\mathcal{R}_x} = (1,x_2',\sqrt{y_2'^2+z_2'^2},0). \end{gathered} \tag{3.3}$$

Similarly, $\mathcal{R}_z = (\mathbf{R_z}, 0)$ is a special rotation about z axis with 0 fibre translation which moves the point $(1,x_2'',y_2'',0)$ into the $(1,1,0,0)$ point.

$$A_2^{\mathcal{T}\mathcal{R}_x} = (1,x_2'',y_2'',0) \to A_2^{\mathcal{T}\mathcal{R}_x\mathcal{R}_z} = (1,1,0,0). \tag{3.4}$$

Finally we apply the inverse transformation $\mathcal{R}_x^{-1}$ of rotation $\mathcal{R}_x$ because the geodesic curve $g(A_1, A_2)$ between the points A_1 and A_2 and its image $g(A_1^2, A_1)$ under the transformation $\mathcal{T}\mathcal{R}_x\mathcal{R}_z\mathcal{R}_x^{-1}$ lie in the same plane in Euclidean sense. The matrix of the above transformation $\mathbf{T}_{A_2}^{\mathbf{S}^2\times\mathbf{R}} = \mathcal{T}\mathcal{R}_x\mathcal{R}_z\mathcal{R}_x^{-1}$ is the following:

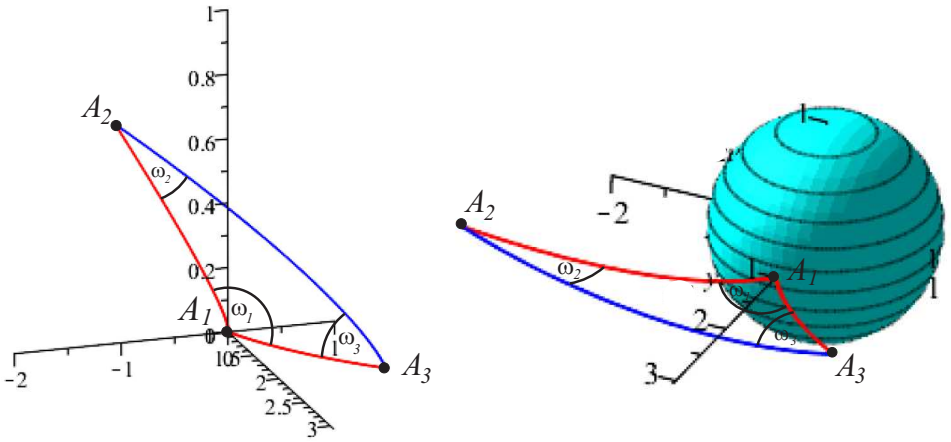


Figure 1. Geodesic triangle with vertices $A_1 = (1, 1, 0, 0)$, $A_2 = (1, 3, -2, 1)$, $A_3 = (1, 2, 1, 0)$ in $\mathbf{S}^2\times\mathbf{R}$ geometry.

$$\mathbf{T}_{A_2}^{\mathbf{S}^2\times\mathbf{R}} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{x_2}{(x_2)^2+(y_2)^2+(z_2)^2} & \frac{-y_2}{(x_2)^2+(y_2)^2+(z_2)^2} & \frac{-z_2}{(x_2)^2+(y_2)^2+(z_2)^2} \\ 0 & \frac{y_2}{(x_2)^2+(y_2)^2+(z_2)^2} & \frac{(y_2)^2x_2+(z_2)^2\sqrt{(x_2)^2+(y_2)^2+(z_2)^2}}{((x_2)^2+(y_2)^2+(z_2)^2)((y_2)^2+(z_2)^2)} & \frac{-y_2z_2(-x_2+\sqrt{(x_2)^2+(y_2)^2+(z_2)^2})}{((x_2)^2+(y_2)^2+(z_2)^2)((y_2)^2+(z_2)^2)} \\ 0 & \frac{z_2}{(x_2)^2+(y_2)^2+(z_2)^2} & \frac{(y_2)z_2(-x_2+\sqrt{(x_2)^2+(y_2)^2+(z_2)^2})}{((x_2)^2+(y_2)^2+(z_2)^2)((y_2)^2+(z_2)^2)} & \frac{(z_2)^2x_2+(y_2)^2\sqrt{(x_2)^2+(y_2)^2+(z_2)^2}}{((x_2)^2+(y_2)^2+(z_2)^2)((y_2)^2+(z_2)^2)} \end{pmatrix}, \tag{3.5}$$

and the images $\mathbf{T}_{A_2}^{\mathbf{S}^2\times\mathbf{R}}(A_i)$ of the vertices A_i $(i \in \{1, 2, 3\})$ are the following (see also Fig. 2):

$$\mathbf{T}_{A_2}^{\mathbf{S}^2\times\mathbf{R}}(A_1) = A_1^2 =$$

$$= \Big(1, \frac{x_2}{(x_2)^2+(y_2)^2+(z_2)^2}, \frac{-y_2}{(x_2)^2+(y_2)^2+(z_2)^2}, \frac{-z_2}{(x_2)^2+(y_2)^2+(z_2)^2}\Big),$$

$$\mathbf{T}_{A_2}^{\mathbf{S}^2\times\mathbf{R}}(A_2) = A_2^2 = (1, 1, 0, 0),$$

$$\mathbf{T}_{A_2}^{\mathbf{S}^2\times\mathbf{R}}(A_3) = A_3^2 = \Big(1, \frac{x_2x_3+y_2y_3}{(x_2)^2+(y_2)^2+(z_2)^2}, \tag{3.6}$$

$$\frac{y_3(z_2)^2\sqrt{(x_2)^2+(y_2)^2+(z_2)^2}+x_2(y_2)^2y_3-x_3(y_2)^3-x_3y_2(z_2)^2}{((y_2)^2+(z_2)^2)((x_2)^2+(y_2)^2+(z_2)^2)},$$

$$-\frac{z_2(y_3y_2(\sqrt{(x_2)^2+(y_2)^2+(z_2)^2}-x_2)+x_3(y_2)^2+x_3(z_2)^2)}{((y_2)^2+(z_2)^2)((x_2)^2+(y_2)^2+(z_2)^2)}\Big).$$

Remark 3. More information about the isometry group of $\mathbf{S}^2 \times \mathbf{R}$ and about its discrete subgroups can be found in [16] and [17].

Similarly to the above computation we get that the images $\mathbf{T}_{A_3}^{\mathbf{S}^2\times\mathbf{R}}(A_i)$ of the vertices A_i ($i \in \{1,2,3\}$) are the following (see also Fig. 2):

$$\begin{gathered}\mathbf{T}_{A_3}^{\mathbf{S}^2\times\mathbf{R}}(A_1) = A_1^3 = \Big(1, \frac{x_3}{(x_3)^2+(y_3)^2}, \frac{-y_3}{(x_3)^2+(y_3)^2}, 0\Big),\\ \mathbf{T}_{A_3}^{\mathbf{S}^2\times\mathbf{R}}(A_3) = A_3^3 = A_1 = (1,1,0,0),\\ \mathbf{T}_{A_3}^{\mathbf{S}^2\times\mathbf{R}}(A_2) = A_2^3 = \Big(1, \frac{x_2x_3+y_2y_3}{(x_3)^2+(y_3)^2}, \frac{x_3y_2-x_2y_3}{(x_3)^2+(y_3)^2}, \frac{z_2}{\sqrt{(x_3)^2+(y_3)^2}}\Big).\end{gathered} \tag{3.7}$$

Our aim is to determine angle sum $\sum_{i=1}^{3}(\omega_i)$ of the interior angles of geodesic

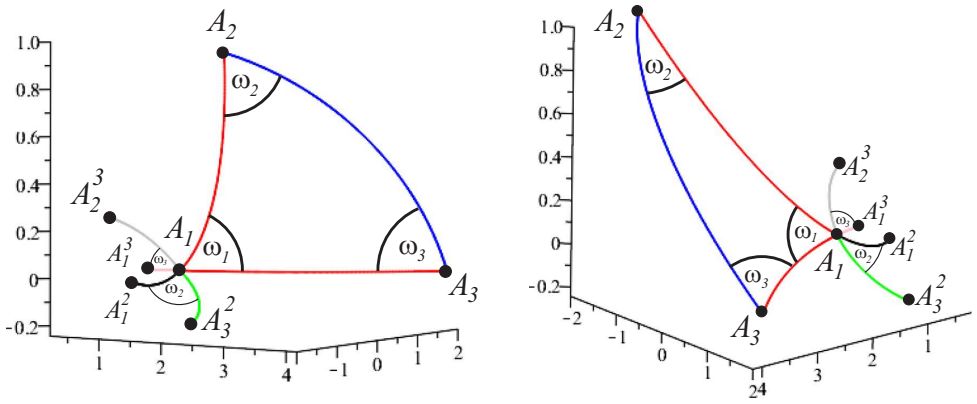


Figure 2. Geodesic triangle with vertices $A_1 = (1,1,0,0)$, $A_2 = (1,3,-2,1)$, $A_3 = (1,2,1,0)$ in $\mathbf{S}^2\times\mathbf{R}$ geometry, and transformed images of its geodesic side segments.

triangles $A_1A_2A_3$ (see Fig. 1, 2). We have seen that ω_1, the angle of geodesic curves with common point at the vertex A_1, is the same as the Euclidean one therefore it can be determined by usual Euclidean sense.

The $\mathbf{T}_{A_i}^{\mathbf{S}^2\times\mathbf{R}}$ ($i = 2,3$) are isometries in $\mathbf{S}^2 \times \mathbf{R}$ geometry thus ω_i is equal to the angle $(g(A_i^i, A_1^i), g(A_i^i, A_j^i))\angle$ ($i,j = 2,3$, $i \neq j$) (see Fig. 2) where $g(A_i^i, A_1^i)$, $g(A_i^i, A_j^i)$ are oriented geodesic curves ($A_1 = A_2^2 = A_3^3$) and ω_1 is equal to the angle $(g(A_1, A_2), g(A_1, A_3))\angle$ where $g(A_1, A_2)$, $g(A_1, A_3)$ are also oriented geodesic curves.

We denote the oriented unit tangent vectors of the oriented geodesic curves $g(A_1, A_i^j)$ with $\mathbf{t}_i^j$ where $(i,j) \in \{(1,3),(1,2),(2,3),(3,2),(3,0),(2,0)\}$ and $A_3^0 = A_3$, $A_2^0 = A_2$. The Euclidean coordinates of $\mathbf{t}_i^j$ (see Section 2.1) are :

$$\mathbf{t}_i^j = (\sin(v_i^j), \cos(v_i^j)\cos(u_i^j), \cos(v_i^j)\sin(u_i^j)). \tag{3.8}$$

In order to obtain the angle of two geodesic curves $g(A_1, A_i^j)$ and $g(A_1, A_k^l)$ ($(i,j) \neq (k,l)$; $(i,j),(k,l) \in \{(1,3),(1,2),(2,3),(3,2),(3,0),(2,0)\}$) intersected at the vertex A_1 we need to determine their tangent vectors $\mathbf{t}_s^r$ ($(s,r) \in \{(1,3),(1,2),$

$(2,3),(3,2),(3,0),(2,0)\}$) (see (3.8)) at their starting point A_1. From (3.8) it follows that a tangent vector at the origin is given by the parameters u and v of the corresponding geodesic curve (see (2.10)), which can be determined from the homogeneous coordinates of the endpoint of the geodesic curve as the following Lemma shows:

Lemma 1. *Let $(1,x,y,z)$ $(x,y,z\in\mathbf{R}, x^2+y^2+z^2\neq 0)$ be the homogeneous coordinates of the point $P\in\mathbf{S}^2\times\mathbf{R}$. The paramerters of the corresponding geodesic curve $g(A_1,P)$ are the following:*

1. *$y,z\in\mathbf{R}\setminus\{0\}$ and $x^2+y^2+z^2\neq 1$;*

$$v=\arctan\Big(\frac{\log\sqrt{x^2+y^2+z^2}}{\arccos\frac{x}{\sqrt{x^2+y^2+z^2}}}\Big),\ u=\arctan\Big(\frac{z}{y}\Big),$$
$$\tau=\frac{\log\sqrt{x^2+y^2+z^2}}{\sin v},\ \text{where}\ -\pi<u\leq\pi,\ -\pi/2\leq v\leq\pi/2,\ \tau\in\mathbf{R}^+. \tag{3.9}$$

2. *$y=0$, $z\neq 0$ and $x^2+z^2\neq 1$;*

$$u=\frac{\pi}{2},\ v=\arctan\Big(\frac{\log\sqrt{x^2+z^2}}{\arccos\frac{x}{\sqrt{x^2+z^2}}}\Big),$$
$$\tau=\frac{\log\sqrt{x^2+z^2}}{\sin v},\ \text{where}\ -\pi/2\leq v\leq\pi/2,\ \tau\in\mathbf{R}^+. \tag{3.10}$$

3. *$y=0$, $z\neq 0$ and $x^2+z^2=1$;*

$$u=\frac{\pi}{2},\ v=0,\ \tau=\arccos(x),\ \tau\in\mathbf{R}^+. \tag{3.11}$$

4. *$y,z=0$;*

$$u=0,\ v=\frac{\pi}{2},\ \tau=\log\sqrt{x^2+y^2+z^2},\ \tau\in\mathbf{R}^+. \tag{3.12}$$

5. *$x=0$, $y=0$ and $z\neq 1$;*

$$u=\frac{\pi}{2},\ v=\arctan\frac{2\log|z|}{\pi},\ \tau=\frac{\log|z|}{\sin v},$$
$$-\pi/2\leq v\leq\pi/2,\ \tau\in\mathbf{R}^+. \tag{3.13}$$

□

We obtain directly from the (2.4) equations of the geodesic curves the following

Lemma 2. *Let P be an arbitrary point and $g(A_1,P)$ ($A_1=(1,1,0,0)$) is a geodesic curve in the considered model of $\mathbf{S}^2\times\mathbf{R}$ geometry. The points of the geodesic curve $g(A_1,P)$ and the centre of the model E_0 lie in a plane in Euclidean sense (see Fig. 3).*
□

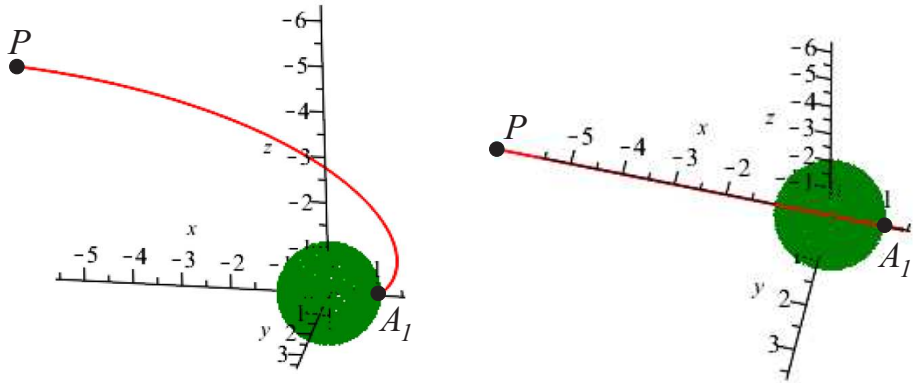


Figure 3. Geodesic curve $g(A_1, P)$ ($A_1 = (1, 1, 0, 0)$ and $P \in \mathbf{S}^2 \times \mathbf{R}$) with "base plane", the plane of a geodesic curve contains the origin $E_0 = (1, 0, 0, 0)$ of the model.

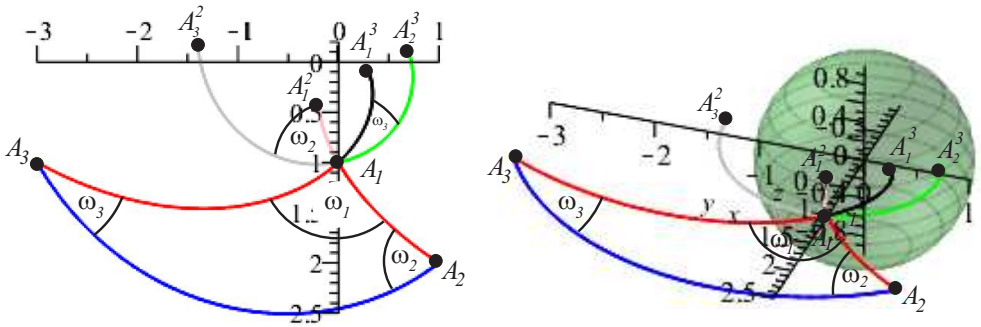


Figure 4. Geodesic triangle with vertices $A_1 = (1, 1, 0, 0)$, $A_2 = (1, 1, -3, 0)$, $A_3 = (1, 2, 1, 0)$ in $\mathbf{S}^2 \times \mathbf{R}$ geometry, and transformed images of its geodesic side segments. The geodesic curve segments $g(A_1, A_2)$, $g(A_2, A_3)$, $g(A_1, A_3)$ lie on the coordinate plane $[x, y]$ and the interior angle sum of this geodesic triangle is $\sum_{i=1}^{3}(\omega_i) = \pi$.

Theorem 1. *If the Euclidean plane of the vertices of an* $\mathbf{S}^2 \times \mathbf{R}$ *geodesic triangle* $A_1A_2A_3$ *contains the centre of model* E_0 *then its interior angle sum is equal to* π.

Proof: We can assume without loss of generality that the vertices A_1, A_2, A_3 of such a geodesic triangle lie in the $[x, y]$ plane of the model. Using Lemma 2 we get that the geodesic segments A_1A_2, A_1A_3 and A_2A_3 are contained by the [x,y] plane, too.

The $\mathbf{S}^2 \times \mathbf{R}$ transformations $\mathbf{T}_{A_2}^{\mathbf{S}^2 \times \mathbf{R}}$ and $\mathbf{T}_{A_3}^{\mathbf{S}^2 \times \mathbf{R}}$ are isometries in $\mathbf{S}^2 \times \mathbf{R}$ geometry, thus ω_2 is equal to the angle $(g(A_2^2, A_1^2), g(A_2^2, A_3^2))\angle$ (see Fig. 2, 4) of the oriented geodesic segments $g(A_2^2, A_1^2)$, $g(A_2^2, A_3^2)$ and ω_3 is equal to the angle $(g(A_3^3, A_1^3), g(A_3^3 A_2^3))\angle$ of the oriented geodesic segments $g(A_3^3, A_1^3)$ and $g(A_3^3, A_2^3)$ $(A_1 = A_2^2 = A_3^3)$.

Substituting the coordinates of the points A_i^j (see (3.5), (3.6) and (3.7)) $((i, j) \in$

$\{(1,3),(1,2),\ (2,3),(3,2),(3,0),(2,0)\}$) into the appropriate equations (3.8-12) of Lemma 1, it is easy to see that

$$\begin{aligned} v_2^0 = -v_1^2,\ u_2^0 - u_1^2 = \pm\pi \Rightarrow \mathbf{t}_2^0 = -\mathbf{t}_1^2, \\ v_3^0 = -v_1^3,\ u_3^0 - u_1^3 = \pm\pi \Rightarrow \mathbf{t}_3^0 = -\mathbf{t}_1^3, \\ v_3^2 = -v_2^3,\ u_3^2 - u_2^3 = \pm\pi \Rightarrow \mathbf{t}_3^2 = -\mathbf{t}_2^3. \end{aligned} \tag{3.13}$$

The endpoints T_i^j of the position vectors $\mathbf{t}_i^j = \overrightarrow{A_1 T_i^j}$ lie on the unit sphere centred at the origin. The measure of angle ω_i ($i \in \{1,2,3\}$) of the vectors $\mathbf{t}_i^j$ and $\mathbf{t}_r^s$ is equal to the spherical distance of the corresponding points T_i^j and T_r^s on the unit sphere (see Fig. 4). Moreover, a direct consequence of equations (3.13) is that each point pair (T_2, T_1^2), (T_3,T_1^3), (T_2^3,T_3^2) contains antipodal points related to the unit sphere with centre A_1.

Due to the antipodality $\omega_1 = T_2 A_1 T_3 \angle = T_1^2 A_1 T_1^3 \angle$, therefore their corresponding spherical distances are equal, as well (see Fig. 4). Now, the sum of the interior angles $\sum_{i=1}^{3}(\omega_i)$ can be considered as three consecutive spherical arcs ($T_3^2 T_1^2$), ($T_1^2 T_1^3$), $T_1^3 T_2^3$). Since the points T_2, T_1^2, T_3, T_1^3, T_2^3, T_3^2 lie in the $[x,y]$ plane (see Lemma 2) the sum of these arc lengths is equal to the half of the circumference of the main circle on the unit sphere,i.e. π. □

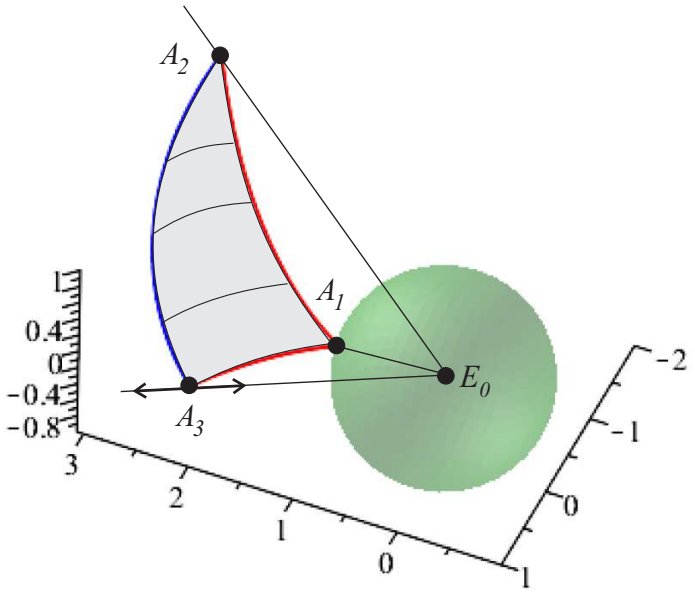


Figure 5. Geodesic triangle with vertices $A_1 = (1,1,0,0)$, $A_2 = (1,3,-2,1)$, $A_3 = (1,2,1,0)$ and the correspondig trihedron with base sphere of $\mathbf{S}^2 \times \mathbf{R}$ geometry.

We can determine the interior angle sum of arbitrary geodesic triangle. In the following table we summarize some numerical data of interior angles of given geodesic triangles:

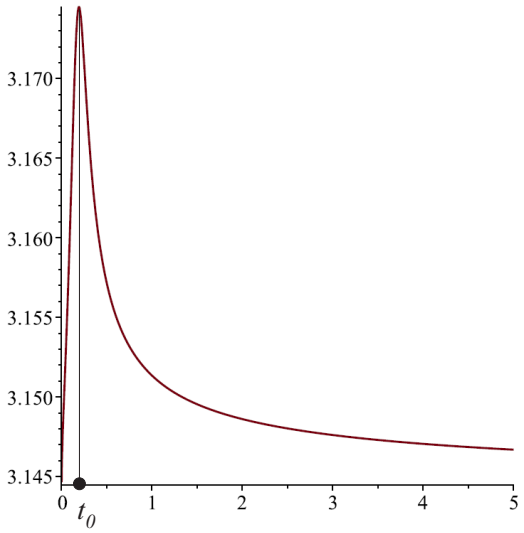


Figure 6. $S(\Delta(t))$ function related to parameters $x_2 = 3, y_2 = -2, z_2 = 1$ $x_3 = 2 \cdot t, y_3 = 1 \cdot t, z_3 = 0$.

Table 1: $A_1 = (1,0,0,0)$, $A_2 = (1,3,-2,1)$				
A_3/ω_i	ω_1	ω_2	ω_3	$\sum_{i=1}^{3}(\omega_i)$
$(1, 2/\sqrt{5}, 1/\sqrt{5}, 0)$	1.97206	0.26028	0.92635	3.15869
$(1, 2, 1, 0)$	0.94654	0.68775	1.51707	3.15135
$(1, 4, 2, 0)$	0.73193	1.29546	1.12123	3.14862
$(1, 12, 6, 0)$	0.61470	1.99926	0.53246	3.14643
$(1, 2000, 1000, 0)$	0.50628	2.52677	0.11050	3.14355

By the above experiences and computations we obtain the following

Theorem 2. *If the Euclidean plane of the vertices of a* $\mathbf{S}^2 \times \mathbf{R}$ *geodesic triangle* $A_1A_2A_3$ *does not contain the centre of model* E_0 *then its interior angle sum is greater than* π.

Proof: We can assume without loss of generality that the vertices A_1, A_2 of such a geodesic triangle lie in the $[x, y]$ plane of the model. Using Lemma 2 we get that the geodesic segment A_iA_j $((i, j) \in \{(1, 2), (1, 3), 2, 3)\})$ is contained in the $A_iA_jE_0$ plane, therefore the sides of triangle $A_1A_2A_3$ lie on the boundary of trihedron given by the points E_0, A_1, A_2, A_3 (see Fig. 2 and 5). It is clear that all types of geodesic triangles can be described by such a triangle. Therefore, it is sufficient to investigate the interior angle sums of geodesic triangles where we fix two of the vertices, e.g. A_1 and A_2, and move the third vertex A_3 on the half straight line E_0A_3 with starting point $E_0 \neq A_3(t)$.

Remark 4. It is well known that if the vertices A_1, A_2, A_3 lie on a sphere of radius $R \in \mathbf{R}^+$ centred at E_0 then the interior angle sum of spherical triangle $A_1A_2A_3$ is greater than π.

Let $\Delta^{\mathbf{S}^2\times\mathbf{R}}(t)$ ($t\in\mathbf{R}^+$) denote the above geodesic triangle with *interior angles* $\omega_i(t)$ at the vertex A_i ($i\in\{1,2,3\}$).

The interior angle sum function $S(\Delta(t))=\sum_{i=1}^{3}(\omega_i(t))$ can be determined relative to the parameters $x_2,y_2,z_2,x_3,y_3\in\mathbf{R}$ by the formulas (2.4), (3.6), (3.7) and by Lemma 1. Analyzing the above complicated continuous functions of single real variable t we get that its maximum is achieved at a point $t_0\in(0,\infty)$ depending on given parameters. Moreover, $S(\Delta^{\mathbf{S}^2\times\mathbf{R}}(t))$ is stricly increasing on the interval $(0,t_0)$, stricly decreasing on the interval (t_0,∞) and

$$\lim_{t\to 0}S(\Delta^{\mathbf{S}^2\times\mathbf{R}}(t))=\pi,\qquad \lim_{t\to\infty}S(\Delta^{\mathbf{S}^2\times\mathbf{R}}(t))=\pi.$$

In Fig. 6 we described the $S(\Delta^{\mathbf{S}^2\times\mathbf{R}}(t))$ function related to geodesic triangle $\Delta^{\mathbf{S}^2\times\mathbf{R}}(t)$ ($t\in(0,5)$) with vertices $A_1=(1,1,0,0)$, $A_2=(1,3,-2,1)$, $A_3=(1,2\cdot t,1\cdot t,0)$. Its maximum is achieved at $t_0\approx 0.19316$ where $S(\Delta^{\mathbf{S}^2\times\mathbf{R}}(t_0))\approx 3.17450$. □

Finally we get the following

Theorem 3. *The sum of the interior angles of a geodesic triangle of* $\mathbf{S}^2\times\mathbf{R}$ *space is greater than or equal to* π. □

3.2 Interior angle sums in $\mathbf{H}^2\times\mathbf{R}$ geometry

Similarly to the $\mathbf{S}^2\times\mathbf{R}$ space we investigate the interior angles of a geodesic triangle $A_1A_2A_3$ and its interior angle sum $\sum_{i=1}^{3}(\omega_i)$ in the $\mathbf{H}^2\times\mathbf{R}$ space. Therefore we define *isometric transformations* $\mathbf{T}_{A_i}^{\mathbf{H}^2\times\mathbf{R}}$ ($i\in\{2,3\}$) as elements of the isometry group of $\mathbf{H}^2\times\mathbf{R}$ geometry that maps the A_i onto the vertex A_1. Let the isometry $\mathbf{T}_{A_2}^{\mathbf{H}^2\times\mathbf{R}}$ be given by the composition of some special types of $\mathbf{H}^2\times\mathbf{R}$ isometries, which transforms a fixed $A_2=(1,x_2,y_2,z_2)$ point of $\mathbf{H}^2\times\mathbf{R}$ into $A_1=(1,1,0,0)$ (up to a positive determinant factor). The methods, the considered transformations and the determinations of their matrices are similar to the $\mathbf{S}^2\times\mathbf{R}$ case and therefore are not detailed here. The images $\mathbf{T}_{A_2}^{\mathbf{H}^2\times\mathbf{R}}(A_i)$ of the vertices A_i ($i\in\{1,2,3\}$) are the following (see also Fig. 7, 9):

$$\mathbf{T}_{A_2}^{\mathbf{H}^2\times\mathbf{R}}(A_1)=A_1^2=$$
$$=\Big(1,\frac{x_2}{(x_2)^2-(y_2)^2-(z_2)^2},\frac{-y_2}{(x_2)^2-(y_2)^2-(z_2)^2},\frac{-z_2}{(x_2)^2-(y_2)^2-(z_2)^2}\Big),$$
$$\mathbf{T}_{A_2}^{\mathbf{H}^2\times\mathbf{R}}(A_2)=A_2^2=(1,1,0,0),$$
$$\mathbf{T}_{A_2}^{\mathbf{H}^2\times\mathbf{R}}(A_3)=A_3^2=\Big(1,\frac{x_2x_3-y_2y_3}{(x_2)^2-(y_2)^2-(z_2)^2},\tag{3.14}$$
$$\frac{y_3(z_2)^2\sqrt{(x_2)^2-(y_2)^2-(z_2)^2}+x_2(y_2)^2y_3-x_3(y_2)^3-x_3y_2(z_2)^2}{((y_2)^2+(z_2)^2)((x_2)^2-(y_2)^2-(z_2)^2)},$$
$$-\frac{z_2(y_3y_2(\sqrt{(x_2)^2-(y_2)^2-(z_2)^2}-x_2)+x_3(y_2)^2+x_3(z_2)^2)}{((y_2)^2+(z_2)^2)((x_2)^2-(y_2)^2-(z_2)^2)}\Big).$$

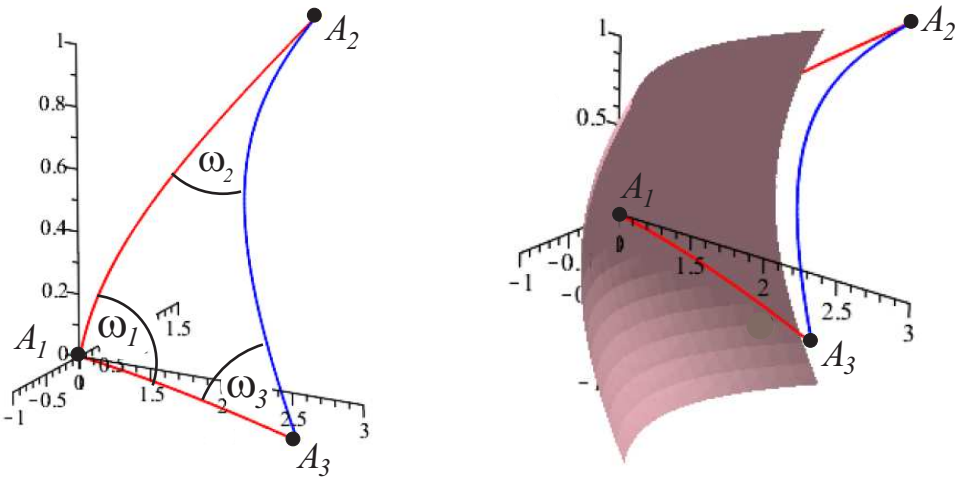


Figure 7. Geodesic triangle with vertices $A_1 = (1, 1, 0, 0)$, $A_2 = (1, 2, 3/2, 1)$, $A_3 = (1, 3, -1, 0)$ in $\mathbf{H}^2 \times \mathbf{R}$ geometry.

Remark 5. More information about the isometry group of $\mathbf{H}^2 \times \mathbf{R}$ and about its discrete subgroups can be found in [18].

Similarly to the above computation we get that the images $\mathbf{T}_{A_3}^{\mathbf{H}^2\times\mathbf{R}}(A_i)$ of the vertices A_i ($i \in \{1, 2, 3\}$) are the following (see also Fig. 7, 9):

$$\begin{aligned} \mathbf{T}_{A_3}^{\mathbf{H}^2\times\mathbf{R}}(A_1) = A_1^3 &= \Big(1, \frac{x_3}{(x_3)^2 - (y_3)^2}, \frac{-y_3}{(x_3)^2 - (y_3)^2}, 0\Big), \\ \mathbf{T}_{A_3}^{\mathbf{H}^2\times\mathbf{R}}(A_3) = A_3^3 &= A_1 = (1, 1, 0, 0), \\ \mathbf{T}_{A_3}^{\mathbf{H}^2\times\mathbf{R}}(A_2) = A_2^3 &= \Big(1, \frac{x_2x_3 - y_2y_3}{(x_3)^2 - (y_3)^2}, \frac{x_3y_2 - x_2y_3}{(x_3)^2 - (y_3)^2}, \frac{z_2}{\sqrt{(x_3)^2 - (y_3)^2}}\Big). \end{aligned} \tag{3.15}$$

The method is the same as that used for $\mathbf{S}^2\times\mathbf{R}$ case to determine angle sum $\sum_{i=1}^{3}(\omega_i)$ of the interior angles of geodesic triangles $A_1A_2A_3$ (see Fig. 7, 9). We have seen that ω_1, the angle of geodesic curves with common point at the vertex A_1, is the same as the Euclidean one therefore it can be determined in usual Euclidean sense.

ω_i is equal to the angle $(g(A_i^i, A_1^i), g(A_i^i, A_j^i))\angle$ $(i, j = 2, 3,\ i \neq j)$ (see Fig. 7, 9) where $g(A_i^i, A_1^i)$, $g(A_i^i, A_j^i)$ are oriented geodesic curves ($A_1 = A_2^2 = A_3^3$) and ω_1 is equal to the angle $(g(A_1, A_2), g(A_1, A_3))\angle$ where $g(A_1, A_2)$, $g(A_1, A_3)$ are also oriented geodesic curves. We denote the oriented unit tangent vectors of the oriented geodesic curves $g(A_1, A_i^j)$ with $\mathbf{t}_i^j$ where $(i, j) \in \{(1, 3), (1, 2), (2, 3), (3, 2), (3, 0), (2, 0)\}$ and $A_3^0 = A_3$, $A_2^0 = A_2$. The Euclidean coordinates of $\mathbf{t}_i^j$ coincide with the coordinates in (3.8) (see Section 2.2). In order to obtain the angle of two geodesic curves $g(A_1, A_i^j)$ and $g(A_1, A_k^l)$ $((i, j) \neq (k, l);\ (i, j), (k, l) \in \{(1, 3), (1, 2), (2, 3), (3, 2), (3, 0), (2, 0)\})$ intersected at the vertex A_1 we need to determine their tangent vectors $\mathbf{t}_s^r$ $((s, r) \in \{(1, 3), (1, 2), (2, 3), (3, 2), (3, 0), (2, 0)\})$

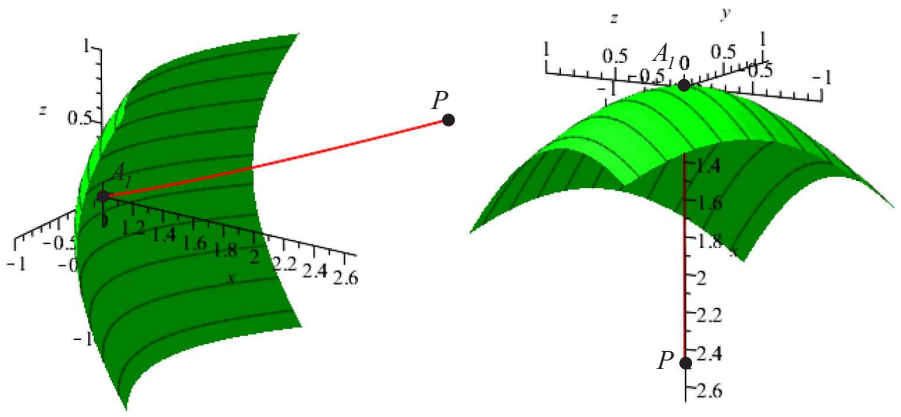


Figure 8. Geodesic curve $g(A_1, P)$ ($A_1 = (1, 1, 0, 0)$ and $P \in \mathbf{H}^2 \times \mathbf{R}$) with "base plane" (the "upper" sheet of the two-sheeted hyperboloid), the plane of a geodesic curve contains the origin $E_0 = (1, 0, 0, 0)$ of the model.

(see (2.10) and (3.8)) at their starting point A_1. From (3.8) it follows that a tangent vector at the origin is given by the parameters u and v of the corresponding geodesic curve (see (2.10)), which can be determined from the homogeneous coordinates of the endpoint of the geodesic curve as the following Lemma shows:

Lemma 3. *Let $(1, x, y, z)$ $(x, y, z \in \mathbf{R}, x^2 - y^2 - z^2 \geq 0,\ x \geq 0)$ be the homogeneous coordinates of the point $P \in \mathbf{H}^2 \times \mathbf{R}$. The paramerters of the corresponding geodesic curve $g(A_1, P)$ are the following:*

1. *$y, z \in \mathbf{R} \setminus \{0\}$ and $x^2 - y^2 - z^2 \neq 1$;*

$$v = \arctan\Big(\frac{\log\sqrt{x^2-y^2-z^2}}{\operatorname{arccosh}\frac{x}{\sqrt{x^2-y^2-z^2}}}\Big),\ u = \arctan\Big(\frac{z}{y}\Big),$$
$$\tau = \frac{\log\sqrt{x^2-y^2-z^2}}{\sin v},\ \textit{where}\ -\pi < u \leq \pi,\ -\pi/2 \leq v \leq \pi/2,\ \tau \in \mathbf{R}^+. \tag{3.16}$$

2. *$y = 0$, $z \neq 0$ and $x^2 - z^2 \neq 1$;*

$$u = \frac{\pi}{2},\ v = \arctan\Big(\frac{\log\sqrt{x^2-z^2}}{\operatorname{arccosh}\frac{x}{\sqrt{x^2-z^2}}}\Big),$$
$$\tau = \frac{\log\sqrt{x^2-z^2}}{\sin v},\ \textit{where}\ -\pi/2 \leq v \leq \pi/2,\ \tau \in \mathbf{R}^+. \tag{3.17}$$

3. *$y = 0$, $z \neq 0$ and $x^2 - z^2 = 1$;*

$$u = \frac{\pi}{2},\ v = 0,\ \tau = \operatorname{arccosh}(x),\ \tau \in \mathbf{R}^+. \tag{3.18}$$

4. $y, z = 0$*;*

$$u = 0,\ v = \frac{\pi}{2},\ \tau = \log(x),\ \tau \in \mathbf{R}^+. \tag{3.19}$$

□

We obtain directly from the (2.10) equations of the geodesic curves the following

Lemma 4. *Let P be an arbitrary point and $g(A_1, P)$ ($A_1 = (1, 1, 0, 0)$) is a geodesic curve in the considered model of $\mathbf{H}^2 \times \mathbf{R}$ geometry. The points of the geodesic curve $g(A_1, P)$ and the centre of the model E_0 lie in a plane in Euclidean sense (see Fig. 8).* □

The proof of the next theorem essentially is the same as the proof of Theorem 1.

Theorem 4. *If the Euclidean plane of the vertices of a $\mathbf{H}^2 \times \mathbf{R}$ geodesic triangle $A_1A_2A_3$ contains the centre of model E_0 then its interior angle sum is equal to π (see Fig. 9).* □

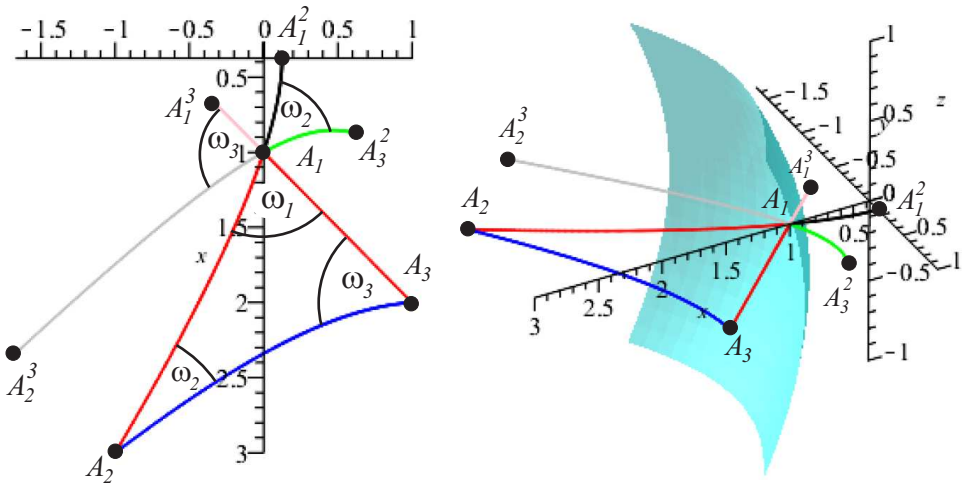


Figure 9. Geodesic triangle with vertices $A_1 = (1, 1, 0, 0)$, $A_2 = (1, 2, 3/2, 1)$, $A_3 = (1, 3, -1, 0)$ in $\mathbf{H}^2 \times \mathbf{R}$ geometry, and transformed images of its geodesic side segments. The geodesic curve segments $g(A_1, A_2)$, $g(A_2, A_3)$, $g(A_3, A_1)$ lie on the coordinate plane $[x, y]$ and the interior angle sum of this geodesic triangle is $\sum_{i=1}^{3}(\omega_i) = \pi$.

We can determine the interior angle sum of arbitrary $\mathbf{H}^2 \times \mathbf{R}$ geodesic triangle. In the following table we summarize some numerical data of interior angles of given geodesic triangles:

Table 2: $A_1 = (1,0,0,0)$, $A_2 = (1,2,3/2,1)$				
A_3	ω_1	ω_2	ω_3	$\sum_{i=1}^{3}(\omega_i)$
$(1,3/\sqrt{8},-1/\sqrt{8},0)$	2.54659	0.06953	0.41780	3.03392
$(1,3,-1,0)$	1.93230	0.49280	0.69816	3.12325
$(1,6,-2,0)$	1.83102	0.71611	0.58348	3.13061
$(1,9,-3,0)$	1.80083	0.81224	0.51964	3.13270
$(1,3000,-1000,0)$	1.70394	1.25735	0.17793	3.13922

By the above experiences and computations we obtain the following

Theorem 5. *If the Euclidean plane of the vertices of a* $\mathbf{H}^2\times\mathbf{R}$ *geodesic triangle* $A_1A_2A_3$ *does not contain the centre of model* E_0 *then its interior angle sum is less than* π*.*

Proof: The proof is similar to the $\mathbf{S}^2\times\mathbf{R}$ case.

We can assume without loss of generality that the vertices A_1, A_2 of such a geodesic triangle lie in the $[x,y]$ plane of the model. Using Lemma 4 we get that the geodesic segment A_iA_j $((i,j)\in\{(1,2),(1,3),2,3)\})$ is contained in the $A_iA_jE_0$ plane, therefore the sides of triangle $A_1A_2A_3$ lie on the boundary of trihedron given by the points E_0, A_1, A_2, A_3. It is clear that all types of geodesic triangles can be described by such a triangle. Therefore, it is sufficient to investigate the interior angle sums of geodesic triangles where we fix two of the vertices, e.g. A_1 and A_2, and move the third vertex A_3 on the half straight line E_0A_3 with starting point $E_0 \neq A_3(t)$.

Remark 6. It is well known that if the vertices A_1, A_2, A_3 lie in an "upper" sheet of the two-sheeted hyperboloid (in the hyperboloid model of the hyperbolic plane geometry where the straight lines of hyperbolic 2-space are modeled by geodesics on the hyperboloid) centred at E_0 then the interior angle sum of hyperbolic triangle $A_1A_2A_3$ is less than π.

Let $\Delta(t)$ $(t\in\mathbf{R}^+)$ denote the above geodesic triangle with *interior angles* $\omega_i(t)$ at the vertex A_i $(i\in\{1,2,3\})$.

The interior angle sum function $S(\Delta^{\mathbf{H}^2\times\mathbf{R}}(t)) = \sum_{i=1}^{3}(\omega_i(t))$ can be determined relative to the parameters $x_2, y_2, z_2, x_3, y_3 \in \mathbf{R}$ by the formulas (2.10), (3.14), (3.15) and by Lemma 3. Analyzing the above complicated continuous functions of single real variable t we get that its maximum is achieved at a point $t_0 \in (0,\infty)$ depending on given parameters. Moreover, $S(\Delta^{\mathbf{H}^2\times\mathbf{R}}(t))$ is stricly increasing on the interval $(0,t_0)$, stricly decreasing on the interval (t_0,∞) and

$$\lim_{t\to 0} S(\Delta^{\mathbf{H}^2\times\mathbf{R}}(t)) = \pi, \qquad \lim_{t\to\infty} S(\Delta^{\mathbf{H}^2\times\mathbf{R}}(t)) = \pi.$$

In Fig. 10 we described the $S(\Delta(t))$ function related to geodesic triangle $\Delta(t)$ $(t\in(0,5))$ with vertices $A_1 = (1,1,0,0)$, $A_2 = (1,2,3/2,1)$, $A_3 = (1,3\cdot t,-1\cdot t,0)$. Its minimum is achieved at $t_0 \approx 0.36392$ where $S(\Delta^{\mathbf{H}^2\times\mathbf{R}}(t_0)) \approx 3.03236$. □

Finally we obtain the following

Theorem 6. *The sum of the interior angles of a geodesic triangle of* $\mathbf{H}^2\times\mathbf{R}$ *space is less than or equal to* π*.* □

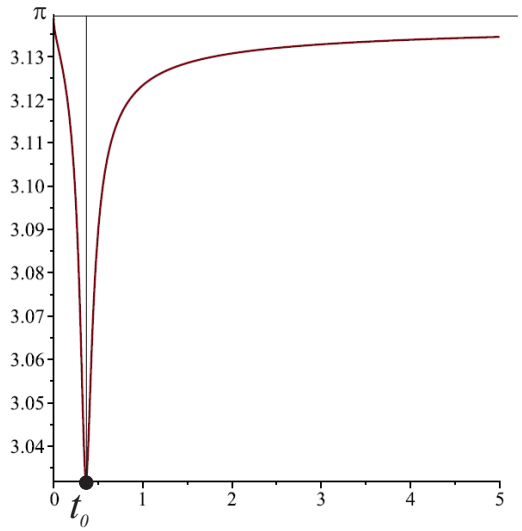


Figure 10. $S(\Delta^{\mathbf{H}^2\times\mathbf{R}}(t))$ function related to parameters $x_2 = 2, y_2 = 3/1, z_2 = 1$ $x_3 = 3 \cdot t, y_3 = -1 \cdot t, z_3 = 0$.

References

[1] Brodaczewska, K. Elementargeometrie in **Nil**. Dissertation (Dr. rer. nat.) Fakultät Mathematik und Naturwissenschaften der Technischen Universität Dresden (2014).

[2] Chavel, I. *Riemannian Geometry: A Modern Introduction.* Cambridge Studies in Advances Mathematics (2006).

[3] Cheeger, J., Ebin, D.G. *Comparison Theorems in Riemannian Geometry.* American Mathematical Society (2006).

[4] Csima, G., Szirmai, J. *Interior angle sum of translation and geodesic triangles in* $\widetilde{\mathbf{SL}_2\mathbf{R}}$ *space.* Filomat, **32/14**, (2018) 5023–5036.

[5] Kobayashi, S. – Nomizu, K. *Foundation of differential geometry, I.* Interscience, Wiley, New York (1963).

[6] Milnor, J. *Curvatures of left Invariant metrics on Lie groups.* Advances in Math., **21** (1976), 293–329.

[7] Molnár, E. *The projective interpretation of the eight 3-dimensional homogeneous geometries.* Beitr. Algebra Geom., **38** No. 2 (1997), 261–288, .

[8] Molnár, E., Szirmai, J. *Symmetries in the 8 homogeneous 3-geometries.* Symmetry Cult. Sci., **21/1-3** (2010), 87-117.

[9] Molnár, E., Szirmai, J. *Classification of* **Sol** *lattices.* Geom. Dedicata, **161/1** (2012), 251-275.

[10] Molnár, E., Szirmai, J., Vesnin, A. *Projective metric realizations of cone-manifolds with singularities along 2-bridge knots and links.* J. Geom., **95** (2009), 91-133.

[11] Pallagi, J., Schultz, B., Szirmai, J. *Visualization of geodesic curves, spheres and equidistant surfaces in* $\mathbf{S}^2\times\mathbf{R}$ *space.* KoG, **14** (2010), 35–40.

[12] Pallagi, J., Schultz, B., Szirmai, J. *Equidistant surfaces in* $\mathbf{H}^2\times\mathbf{R}$ *space.* KoG, **15** (2011), 3–6.

[13] Scott, P. *The geometries of 3-manifolds.* Bull. London Math. Soc., **15** (1983), 401–487.

[14] Szirmai, J. *A candidate to the densest packing with equal balls in the Thurston geometries.* Beitr. Algebra Geom., **55**(2) (2014), 441–452.

[15] Szirmai, J. *Bisector surfaces and circumscribed spheres of tetrahedra derived by translation curves in* **Sol** *geometry.* New York J. Math., **25** (2019), 107–122.

[16] Szirmai, J. *Simply transitive geodesic ball packings to* $\mathbf{S^2} \times \mathbf{R}$ *space groups generated by glide reflections.* Ann. Mat. Pur. Appl., **193/4** (2014), 1201–1211.

[17] Szirmai, J. *Geodesic ball packings in* $\mathbf{S}^2 \times \mathbf{R}$ *space for generalized Coxeter space groups.* Beitr. Algebra Geom., **52** (2011), 413–430.

[18] Szirmai, J. *Geodesic ball packings in* $\mathbf{H}^2 \times \mathbf{R}$ *space for generalized Coxeter space groups.* Math. Commun., **17/1** (2012), 151–170.

[19] Szirmai, J. The densest translation ball packing by fundamental lattices in **Sol** space. *Beitr. Algebra Geom.,* **51**(2) (2010), 353–373.

[20] Szirmai, J. **Nil** *geodesic triangles and their interior angle sums. Bull. Braz. Math. Soc. (N.S.),* **49** (2018), 761–773.

[21] Szirmai, J. *Triangle angle sums related to translation curves in* **Sol** *geometry.* Stud. Univ. Babes-Bolyai Math. (to appear), (2020).

[22] Thurston, W. P. (and Levy, S. editor) *Three-Dimensional Geometry and Topology.* Princeton University Press, Princeton, New Jersey, vol. **1** (1997).

Jenő Szirmai
Budapest University of Technology and
Economics Institute of Mathematics,
Department of Geometry,
Budapest, P. O. Box: 91, H-1521
E-mail: *szirmai@math.bme.hu*

Received January 25, 2020

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 62–67
ISSN 1024–7696

Signature Schemes on Algebras, Satisfying Enhanced Criterion of Post-quantum Security

N. A. Moldovyan

Abstract. The paper introduces an enhanced criterion of the post-quantum security for designing post-quantum digital signature schemes based on the hidden discrete logarithm problem. The proposed criterion requires that it is computationally impossible to construct a periodic function containing a period whose length depends on the value of a discrete logarithm in a hidden cyclic group when using public parameters of the signature scheme. A practical post-quantum signature scheme which satisfies the criterion is proposed.

Mathematics subject classification: 94A60, 16Z05, 14G50, 11T71, 16S50.
Keywords and phrases: finite associative algebra, non-commutative algebra, discrete logarithm problem, hidden logarithm problem, post-quantum cryptography, digital signature.

1 Introduction

Currently the development of practical signature schemes is one of the challenges in the field of cryptography [1, 2]. A signature scheme is called post-quantum if it resists attacks that use hypothetic quantum computers. Signature schemes based on the computational difficulty of the discrete logarithm problem (DLP) and the factorization problem (FP), which are widely used at the present time, are not post-quantum because both the DLP and the FP can be solved in polynomial time with quantum computer [3, 4]. The quantum algorithms for solving each of these problems are based on the extremely high efficiency of a quantum computer to perform discrete Fourier transform of periodic functions that take on values in some fixed finite group [5, 6]. The algorithms for solving the DLP (the FP) use the reducibility of each of these two problems to the problem of finding length of the period depending on the value of the discrete logarithm (divisor of the integer to be factorized) [3, 7]. A post-quantum signature scheme is to be based on a problem that is different from the DLP and the FP, which has superpolynomial computational complexity when solving it with quantum computer.

The hidden DLP (HDLP) was proposed as the base primitive of the post-quantum public key-agreement protocols [8, 9] and post-quantum digital signature schemes [10, 11]. The HDLP-based signature algorithms introduced earlier satisfy the following criterion of the post-quantum security, which requires that periodic functions, constructed on the basis of public parameters of the signature scheme,

take on values that lie in a sufficiently large number of different groups contained in the finite algebra used as the algebraic carrier of the cryptoscheme. However, in the future new quantum algorithms are assumed to be developed for finding the period length of a periodic function whose values are not limited to a single finite group or a sufficiently small number of different finite groups.

We can propose the following enhanced criterion of post-quantum security for designing the HDLP-based signature algorithms: construction of the periodic functions containing a period depending on the value of the discrete logarithm should be a computationally intractable problem, when using the public parameters of the signature scheme.

This paper presents the developed signature scheme that implements the introduced enhanced criterion of the post-quantum security, which is of interest for application as a practical post-quantum signature scheme having high performance and comparatively small size (1550 bits).

2 Notion of the HDLP and its algebraic supports

Usual DLP is defined in a finite cyclic group as finding the value x in the equation $Y = G^x$, where the group elements Y and G are known; G is the generator of the group. For example, in the Schnorr signature algorithm [12] the value G having prime order q of sufficiently large size (≥ 160 bits) is a common parameter, Y is a public key, and x ($x < q$) is the private key of the owner of the public key Y. The quantum algorithm for finding the value x uses the periodic function $f(i,j) = Y^i G^j$ that contains a period of the length $(-1, x)$: $f(i-1, j+x) = Y^{i-1} G^{j+x} = f(i,j)$, where the function $f(i,j)$ takes on the values in the said group.

The HDLP is set in finite non-commutative associative algebras (FNAAs) [10,11] in frame of which one can set sufficintly large number of different cyclic groups. A hidden cyclic group of large prime order q is selected, in it the basic exponentiation operation is performed $Y = G^x$. Then the masking operations ψ_1 and ψ_2 (each of them is mutually commutative with the exponentiation operation) are performed over the values Y and G: $W = \psi_1(Y)$ and $Z = \psi_2(G)$. The values W and Z are elements of the public key in the signature schemes introduced in [10, 11]. In some other signature algorithms [13, 14] the public key includes the third element T that is a matching element needed to provide correctness of the signature scheme. The value T is defined by the selected private operations ψ_1 and ψ_2. Using the public parameters of the known HDLP-based signature schemes one can easily compose the periodic function $f'(i,j) = W^i \circ G^j$ or $f''(i,j) = W^i \circ T \circ G^j$ (where $\circ$ denotes the multiplication operation in the FNAA), which also contains a period of the length $(-1, x)$, however each of the functions $f'(i,j)$ and $f''(i,j)$ takes on the values related to sufficiently large number of different finite cyclic groups contained in the FNAA used as the algebraic support of the signature scheme, therefore, the known quantum algorithms can not be applied for finding the value x.

In the signature scheme introduced in the next section, which satisfies the enhanced criterion of the post-quantum security, it is assumed to use algebraic supports

Table 1. The BVMT setting the quaternion-like FNAA with the unit $(0,1,0,0)$.

$\circ$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_0$	$\lambda\mathbf{e}_1$	$\mathbf{e}_0$	$-\mathbf{e}_3$	$-\lambda\mathbf{e}_2$
$\mathbf{e}_1$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_2$	$-\mathbf{e}_1$	$-\mathbf{e}_0$
$\mathbf{e}_3$	$\lambda\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_0$	$\lambda\mathbf{e}_1$

Table 2. The BVMT of the quaternion-like FNAA with the unit $(0,0,1,0)$.

$\circ$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_0$	$\lambda\mathbf{e}_2$	$-\mathbf{e}_3$	$\mathbf{e}_0$	$-\lambda\mathbf{e}_1$
$\mathbf{e}_1$	$\mathbf{e}_3$	$-\mathbf{e}_2$	$\mathbf{e}_1$	$-\mathbf{e}_0$
$\mathbf{e}_2$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_3$	$\lambda\mathbf{e}_1$	$\mathbf{e}_0$	$\mathbf{e}_3$	$\lambda\mathbf{e}_2$

being 4-dimensional FNAAs containing global two-sided unit, which are defined over the ground finite field $GF(p)$ with the characteristic equal to the prime $p = 2q - 1$, where q is a 256-bit prime. For example, one can use the FNAAs described in [10,14] or quaternion-like FNAAs with the multiplication operation defined with the basis vector multiplication tables (BVMTs) shown as Tables 1, 2, and 3, where $\lambda \neq 0$. (Description of the procedure for performing the multiplication operation in FNAAs is given, for example, in [10]).

Usually the multiplication operation of two vectors A and $B = \sum_{i=0}^{m-1} b_i\mathbf{e}_i$ is defined with the formula $A \circ B = \sum_{j=0}^{m-1}\sum_{i=0}^{m-1} a_i b_j(\mathbf{e}_i \circ \mathbf{e}_j)$, in which products of different pairs of basis vectors $\mathbf{e}_i \circ \mathbf{e}_j$ are to be substituted by a single-component vector indicated in the so-called basis vector multiplication table (BVMT), namely, at the intersection of the ith row and jth column.

In every of the FNAAs defined with Tables 1, 2, and 3 the set of all invertible 4-dimensional vectors forms a finite non-commutative group with the group operation

Table 3. The BVMT of the quaternion-like FNAA with the unit $(0,0,0,1)$.

$\circ$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_0$	$-\lambda\mathbf{e}_3$	$\mathbf{e}_2$	$-\lambda\mathbf{e}_1$	$\mathbf{e}_0$
$\mathbf{e}_1$	$-\mathbf{e}_2$	$\mathbf{e}_3$	$-\mathbf{e}_0$	$\mathbf{e}_1$
$\mathbf{e}_2$	$\lambda\mathbf{e}_1$	$\mathbf{e}_0$	$\lambda\mathbf{e}_3$	$\mathbf{e}_2$
$\mathbf{e}_3$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$

$\circ$, the order Ω of which is described by the following formula:

$$\Omega = p(p-1)\left(p^2-1\right).$$

Due to the used structure of the prime p, the prime value q divides the value Ω, therefore the group contains elements of the order q. In the above group the maximum order of the group elements is equal to p^2-1, like in the case of the finite quaternion algebra defined over the field $GF(p)$ [8].

3 The proposed post-quantum signature scheme

The procedure for generating the public key includes the following steps:

1. Select at random an invertible vector U that is a generator of certain finite cyclic group with the order p^2-1 and compute the vector $G = U^{\frac{p^2-1}{q}}$ that has order equal to the prime q.

2. Select at random invertible vectors X and D with the order p^2-1, which satisfy the conditions $X \circ D \neq D \circ X$, $X \circ G \neq G \circ X$, and $D \circ G \neq G \circ D$.

3. Generate two random natural numbers $x < q$ and $t < q$.

4. Compute the vectors $Z_1 = \psi_D(G \circ U) = D \circ G \circ U \circ D^{-1}$ and $Z_2 = \psi_X\left(G^t \circ U\right) = X \circ G^t \circ U \circ X^{-1}$.

5. Compute the vectors $W_1 = \psi_X(G^x) = X \circ G^x \circ X^{-1}$ and $W_2 = \psi_D\left(G^{tx}\right) = D \circ G^{tx} \circ D^{-1}$.

The public key constitutes two pairs of the vectors (Z_1, W_1) and (Z_2, W_2). All other values used in the public-key generation procedure, except the integers q and p, are secret. The set of secret values that are needed to compute a signature (i. e., the vectors X, D, G, U, and the integers x and t) represent the private key. Computing the private key from the public one is the proposed version of the HDLP that is used as the base primitive of the developed signature scheme described as follows.

Procedure for generation of the signature (h, s, S) to the electronic document M:

1. Select two random integers $w < q$ and $u < q$ and compute the vector $K = G^w \circ U^u$.

2. Generate a random integer $k < q$ and compute the vectors $V_1 = X \circ G^k \circ K \circ D^{-1}$ and $V_2 = X \circ G^{tk} \circ K \circ D^{-1}$.

3. Using some specified 256-bit hash-function f_h compute the hash value h from the document M to which the vectors V_1 and V_2 are concatenated: $h = f_h(M, V_1, V_2)$. The value h is the first signature element.

4. Then compute the second signature element s: $s = k - xh \bmod q$.

5. Compute the third signature element in the form of the vector $S = X \circ G^w \circ U^{u-s} \circ D^{-1}$.

Signature verification procedure is executed as follows:

1. Compute the vector $V_1' = W_1^h \circ S \circ Z_1^s$.

2. Compute the vector $V_2' = Z_2^s \circ S \circ W_2^h$.

3. Compute the value $h' = f_h\left(M, V_1', V_2'\right)$.

4. If $h' = h$, then the signature is accepted as genuine. Otherwise it is rejected.

The masking operations ψ_X and ψ_D define two different automomorphism maps of the FNAA used as algebraic support of the developed signature scheme, therefore each of the above two operations is mutually commutative with the exponentiation operation and the signature scheme performs correctly.

Correctness proof of the signature scheme is as follows:

$$
\begin{aligned}
V_1' &= \left(X \circ G^x \circ X^{-1}\right)^h \circ \left(X \circ G^w \circ U^{u-s} \circ D^{-1}\right) \circ \left(D \circ G \circ U \circ D^{-1}\right)^s = \\
&= X \circ G^{xh} \circ G^w \circ U^{u-s} \circ G^s \circ U^s \circ D^{-1} = X \circ G^{xh} \circ G^{k-xh} \circ G^w \circ U^u \circ D^{-1} = \\
&= X \circ G^k \circ G^w \circ U^u \circ D^{-1} = X \circ G^k \circ K \circ D^{-1} = V_1; \\
V_2' &= \left(X \circ G^t \circ U \circ X^{-1}\right)^s \circ \left(X \circ G^w \circ U^{u-s} \circ D^{-1}\right) \circ \left(D \circ G^t \circ D^{-1}\right)^h = \\
&= X \circ G^{ts} \circ U^s \circ G^w \circ U^{u-s} \circ G^{txh} \circ D^{-1} = \\
&= X \circ G^{t(k-xh)} \circ G^{txh} \circ U^u \circ G^w \circ D^{-1} = X \circ G^{tk} \circ K \circ D^{-1} = V_2; \\
&\left\{V_1' = V_1;\ V_2' = V_2\right\} \Rightarrow f_h\left(M, V_1', V_2'\right) = f_h\left(M, V_1, V_2\right) \Rightarrow h' = h.
\end{aligned}
$$

Thus, the correctly computed signature (h, s, S) passes the verification procedure as genuine signature.

4 Discussion and conclusion

To define computaional complexity of constructing a periodic function containing period depending on the value x, the value U has been imbedded in the public key elements Z_1 and Z_2, which masks well the potential periodicity connected with x. However, when performing the signature verification you need to eliminate the influence of the vector U, which depends on a random value s, so the third element of the signature is used in the form of the vector S calculated depending on the value U^s. To prevent the possibility of using the third element of the signature (which is included as a multiplier of the first degree in the signature verification equation) for signature forgery, the proposed signature scheme uses a double verification equation as compared with the signature schemes [11, 13] used as prototype.

The proposed post-quantum signature scheme is of practical interest, since it has sufficiently high performance and low size of the public key (about 4100 bits) and of the signature (about 1550 bits) in comparison with the candidates for post-quantum signature standard which were proposed in the framework of the world competition for the development of post-quantum two-key cryptosystems [2, 16].

The introduced signature scheme uses computations in cyclic hidden group. A more efficient masking of the periodicity of the periodic functions, which depends on the private value x, is supposed to be provided when using the commutative hidden group with 2-dimensional cyclicity (a group generated by the generator system containing two elements G and U of the same order). However, some particular FNAA are to be used as algebraic carries to implement this idea. To set new specific FNAAs one can use unified methods [10,17] for defining FNAAs of an arbitrary even dimension.

References

[1] *Post-Quantum Cryptography. 9th International Conference, PQCrypto 2018 Proceedings*. Fort Lauderdale, FL, USA, April 9-11, 2018. Springer Verlag LNCS, 2018, **10786**.

[2] *Post-Quantum Cryptography. Proceedings of the 10th International Conference, PQCrypto 2019*. Chongqing, China, May 8−10, 2019. Springer Verlag LNCS, 2019, **11505**.

[3] Shor P. W. *Polynomial-time algorithms for prime factorization and discrete logarithms on quantum computer*. SIAM Journal of Computing, 1997, **26**, 1484–1509.

[4] Smolin J. A., Smith G., Vargo A. *Oversimplifying quantum factoring*, Nature. **499** (2013), No. 7457, 163–165.

[5] A. Ekert, R. Jozsa, *Quantum computation and Shorś factoring algorithm*. Rev. Mod. Phys. **68** (1996), 733.

[6] Jozsa R. *Quantum algorithms and the Fourier transform*. Proc. Roy. Soc. London Ser A, **454** (1998), 323–337.

[7] Yan S. Y. *Quantum Attacks on Public-Key Cryptosystems*, Springer, 2014. 207 p.

[8] Moldovyan D. N. *Non-Commutative Finite Groups as Primitive of Public-Key Cryptoschemes*. Quasigroups and Related Systems, 2010, **18**, No. 2, 165–176.

[9] Kuzmin A. S., Markov V. T., Mikhalev A. A., Mikhalev A. V., Nechaev A. A. *Cryptographic Algorithms on Groups and Algebras*. Journal of Mathematical Sciences, 2017, **223**, No. 5, 629–641.

[10] Moldovyan N. A., Moldovyan A. A. *Finite Non-commutative Associative Algebras as Carriers of Hidden Discrete Logarithm Problem*. Bulletin of the South Ural State University. Ser. Mathematical Modelling, Programming & Computer Software, 2019, **12**, No. 1, 66–81.

[11] Moldovyan N. A. *Finite Non-commutative Associative Algebras for Setting the Hidden Discrete Logarithm Problem and Post-quantum Cryptoschemes on Its Base*. Buletinul Academiei de Stiinte a Republicii Moldova. Matematica, 2019, No. 1 (89), 71 − 78.

[12] Schnorr C.P. *Efficient signature generation by smart cards*. J. Cryptology, 1991, **4**, 161–174.

[13] Moldovyan A. A., Moldovyan N. A. *Post-quantum signature algorithms based on the hidden discrete logarithm problem*. Computer Science J. of Moldova, 2018, **26**, No. 3(78), 301–313.

[14] Moldovyan N. A., Abrosimov I. K. *Post-quantum electronic digital signature scheme based on the enhanced form of the hidden discrete logarithm problem*. Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes, 2019, **15**, No. 2, 212–220

[15] D. N. Moldovyan *New Form of the Hidden Logarithm Problem and Its Algebraic Support*. Buletinul Academiei de Stiinte a Republicii Moldova. Matematica, 2020, No. 2 (93), 3–10.

[16] First NIST standardization conference – April 11–13, 2018. http://prometheuscrypt.gforge.inria.fr/2018-04-18.pqc2018.html

[17] Moldovyan N. A. *Unified Method for Defining Finite Associative Algebras of Arbitrary Even Dimensions*, Quasigroups and Related Systems, 2018, **26**, No. 2. P. 263-270.

Nikolay Moldovyan
St. Petersburg Institute for Informatics and Automation of
Russian Academy of Sciences
14-th line 39, 199178, St. Petersburg
Russia
E-mail: *nmold@mail.ru*

Received February 6, 2020

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 68–75
ISSN 1024–7696

Quasi-Kählerian manifolds and quasi-Sasakian hypersurfaces axiom

Ahmad Abu-Saleem, Mihail B. Banaru,
Galina A. Banaru, Lidia V. Stepanova

Abstract. It is proved that if a quasi-Kählerian manifold satisfies the quasi-Sasakian hypersurfaces axiom, then it is an almost Kahlerian manifold

Mathematics subject classification: 53B35; 53B50.
Keywords and phrases: Almost contact metric structure, quasi-Sasakian structure, hypersurface, almost Hermitian manifold, quasi-Kählerian manifold, almost Kählerian manifold.

1 Introduction

The geometry of almost Hermitian manifolds and geometry of almost contact metric manifolds belong to the most intensively developing areas of modern mathematics. We mark out their profound inner content as well as their diverse applications in many domains of mathematics and theoretical physics [12]. It is known that almost contact metric structures are induced on oriented hypersurfaces of almost Hermitian manifolds. Almost contact metric structures on hypersurfaces of almost Hermitian manifolds were studied since 1960s by such outstanding mathematicians as S. Sasaki [13], S. Goldberg [9] and H. Yanamoto [17]. In the present paper, we consider the case when the almost Hermitian manifold is quasi-Kählerian (i.e. it belongs to the class $W_1 \oplus W_2$ in Gray–Hervella notation [10]). We remark that the class of quasi-Kählerian manifolds contains all Kählerian, nearly Kählerian and almost Kählerian manifolds that are the best studied types of almost Hermitian manifolds. The main result of our note is the following:

Theorem 1. *If a quasi-Kählerian manifold satisfies the quasi-Sasakian hypersurfaces axiom, then it is an almost Kählerian manifold.*

This short article is a continuation of the authors' researches in the area of interconnection of almost Hermitian and almost contact metric structures (see [1, 3, 4, 6, 7, 14–16] and others).

2 Preliminaries

An almost Hermitian manifold is an even-dimensional manifold M^{2n} with a Riemannian metric $g = \langle \cdot, \cdot \rangle$ and an almost complex structure J if the following

condition holds

$$\langle JX,\ JY\rangle = \langle X,\ Y\rangle\,, \quad X,Y \in \aleph(M^{2n}),$$

where $\aleph(M^{2n})$ is the module of smooth vector fields on M^{2n}[10]. The specification of an almost Hermitian structure on a manifold is equivalent to the setting of a G-structure, where G is the unitary group $U(n)$ [7, 12]. Its elements are the frames adapted to the structure (A-frames) that look as follows:

$$(p,\ \varepsilon_1,\ \ldots, \varepsilon_n,\ \varepsilon_{\hat{1}},\ \ldots,\ \varepsilon_{\hat{n}}\,),$$

where ε_a are the eigenvectors corresponded to the eigenvalue $i = \sqrt{-1}$, and $\varepsilon_{\hat{a}}$ are the eigenvectors corresponded to the eigenvalue $-i$. Here the index a ranges from 1 to n, and we state $\hat{a} = a + n$. Therefore, the matrixes of the operator of the almost complex structure and of the Riemannian metric written in an A-frame look as follows, respectively:

$$\left(J_j^k\right) = \left(\begin{array}{c|c} iI_n & 0 \\ \hline 0 & -iI_n \end{array}\right);\ \ (g_{kj}) = \left(\begin{array}{c|c} 0 & I_n \\ \hline I_n & 0 \end{array}\right),$$

where I_n is the identity matrix; $k, j = 1, \ldots, 2n$.

We recall that the fundamental form (or Kählerian form [10]) of an almost Hermitian manifold is determined by the relation

$$F(X,\ Y) = \langle X,\ JY\rangle\,, \quad X,Y \in \aleph(M^{2n}).$$

By direct computing it is easy to obtain that in an A-frame the fundamental form matrix looks as follows:

$$(F_{kj}) = \left(\begin{array}{c|c} 0 & iI_n \\ \hline -iI_n & 0 \end{array}\right).$$

An almost Hermitian manifold is quasi-Kählerian if the following identity holds [10, 12]:

$$\nabla_X(F)(Y,Z) + \nabla_{JX}(F)(JY,Z) = 0,$$

where $X, Y, Z \in \aleph(M^{2n})$. We also remind that the following three identities determine almost Kählerian, nearly Kählerian and Kählerian manifolds, respectively:

$$\nabla F = 0, \quad \nabla_X(F)(X,Y) = 0, \quad dF = 0.$$

We recall also that an almost contact metric structure on an odd-dimensional manifold N is defined by the system of tensor fields $\{\Phi,\ \xi,\ \eta,\ g\}$ on this manifold, where ξ is a vector field, η is a covector field, Φ is a tensor of the type (1, 1) and $g = \langle\cdot,\ \cdot\rangle$ is the Riemannian metric [12]. Moreover, the following conditions are fulfilled:

$$\eta(\xi) = 1,\ \Phi(\xi) = 0,\ \eta \circ \Phi = 0,\ \Phi^2 = -id + \xi \otimes \eta,$$
$$\langle \Phi X, \Phi Y\rangle = \langle X, Y\rangle - \eta\,(X)\,\eta\,(Y)\,,\ X,Y \in \aleph(N),$$

where $\aleph(N)$ is the module of smooth vector fields on N. As the most important examples of almost contact metric structures we can mark out the cosymplectic

structure, the nearly cosymplectic structure, the Sasakian structure and the Kenmotsu structure.

As it was mentioned above, the almost contact metric structures are closely connected to the almost Hermitian structures. For instance, if $(N, \{\Phi, \xi, \eta, g\})$ is an almost contact metric manifold, then an almost Hermitian structure is induced on the product $N \times R$ [12]. If this almost Hermitian structure is integrable, then the input almost contact metric structure is called normal. As it is known, a normal contact metric structure is called Sasakian. On the other hand, we can characterize the Sasakian structure by the following condition [7, 12]:

$$\nabla_X(\Phi)Y = \langle X, Y\rangle\, \xi - \eta(Y)X, \quad X, Y \in \aleph(N).$$

For example, Sasakian structures are induced on totally umbilical hypersurfaces in a Kählerian manifold [13]. As it is well known, the Sasakian structures have many remarkable properties and play a fundamental role in contact geometry. A natural generalization of the Sasakian structure is the quasi-Sasakian structure [8]. An almost contact metric structure $\{\Phi, \xi, \eta, g\}$ is called quasi-Sasakian (qS-) if its fundamental form $\Omega(X, Y) = \langle X, \Phi Y\rangle$ is closed and the following condition holds:

$$N_\Phi + \frac{1}{2}\, d\eta \otimes \xi = 0,$$

where N_Φ is the Nijenhuis tensor of Φ . The theory of quasi-Sasakian structures was created by the outstanding American geometer D. E. Blair [8]. He has established some sufficient conditions under which a qS-manifold is a product of a Sasakian and a Kählerian manifolds.

We remind that an almost Hermitian manifold M^{2n} satisfies the quasi-Sasakian hypersurfaces axiom if a qS-hypersurface passes through every point of this manifold. This terminology was introduced by V. F. Kirichenko [11].

At the end of this section, note that all considered manifolds, tensor fields and similar objects are assumed to be smooth of the class C^∞ .

3 Proof of the theorem

Let us consider a quasi-Kählerian manifold M^{2n} , let N^{2n-1} be its oriented hypersurface. The first group of the Cartan structural equations of a quasi-Kählerian structure written in an A-frame looks as follows [5, 14, 15]:

$$d\,\omega^a = \omega^a_b \wedge \omega^b \;+\; B^{abc}\,\omega_b \wedge \omega_c;$$

$$d\,\omega_a = -\omega^b_a \wedge \omega_b \;+\; B_{abc}\,\omega^b \wedge \omega^c,$$

where

$$B^{abc} = \frac{i}{2}\, J^a_{[\hat{b},\hat{c}]};\; B_{abc} = -\frac{i}{2}\, J^{\hat{a}}_{[b,c]}.$$

The systems of functions $\{\ B^{abc}\ \}$, $\{\ B_{abc}\ \}$ are components of the Kirichenko tensors of the almost Hermitian manifold M^{2n} [2]; $\{J^j_{k,m}\}$ are components of ∇J;

here and further $a, b, c = 1, ..., n;$, $\hat{a} = a + n$. The similar tensors were introduced by L. V. Stepanova [14]:

$$\tilde{B}^{abc} \;=\; -\frac{i}{2}\, J^{a}_{\hat{b},\,\hat{c}};\quad \tilde{B}_{abc} \;=\; \frac{i}{2}\, J^{\hat{a}}_{b,\,c}.$$

Let us consider the Cartan structural equations of the almost contact metric structure on an oriented hypersurface N^{2n-1} of a quasi-Kählerian manifold M^{2n} [7, 14]:

$$\begin{gathered}
d\,\omega^{\alpha} \;=\; \omega^{\alpha}_{\beta} \wedge \omega^{\beta} \;+\; B^{\alpha\beta\gamma}\,\omega_{\beta} \wedge \omega_{\gamma} + \\
+\; i\sigma^{\alpha}_{\beta}\,\omega^{\beta} \wedge \omega \;+\; \left(-\sqrt{2}\,\tilde{B}^{n\alpha\beta} \;-\; \frac{1}{\sqrt{2}}\tilde{B}^{\alpha\beta n} \;+\; i\,\sigma^{\alpha\beta}\right)\,\omega_{\beta} \wedge \omega; \\
d\,\omega_{\alpha} \;=\; -\omega^{\beta}_{\alpha} \wedge \omega_{\beta} \;+\; B_{\alpha\beta\gamma}\,\omega^{\beta} \wedge \omega^{\gamma} - \\
-\; i\sigma^{\beta}_{\alpha}\,\omega_{\beta} \wedge \omega \;+\; \left(-\sqrt{2}\,\tilde{B}_{n\alpha\beta} - \frac{1}{\sqrt{2}}\tilde{B}_{\alpha\beta n} \;-\; i\,\sigma_{a\beta}\right)\,\omega^{\beta} \wedge \omega; \qquad (1) \\
d\,\omega \;=\; \sqrt{2}\,B_{n\alpha\beta}\,\omega^{\alpha} \wedge \omega^{\beta} \;+\; \sqrt{2}\,B^{n\alpha\beta}\,\omega_{\alpha} \wedge \omega_{\beta} - \\
-2i\,\sigma^{\alpha}_{\beta}\,\omega^{\beta} \;\wedge\; \omega_{\alpha} \;+\; \left(\tilde{B}_{n\beta n} \;+\; i\,\sigma_{n\beta}\right)\,\omega \wedge \omega^{\beta} \;+\; \left(\tilde{B}^{n\beta n} \;-\; i\,\sigma^{\beta}_{n}\right)\,\omega \wedge \omega_{\beta},
\end{gathered}$$

where σ is the second fundamental form of the immersion of the hypersurface N^{2n-1} into M^{2n}; here and further $\alpha, \beta, \gamma = 1, ..., n-1$.

Comparing the equations (1) with the first group of Cartran structural equations of a qS-structure [12, 14, 16]

$$d\omega^{\alpha} = \omega^{\alpha}_{\beta} \wedge \omega^{\beta} + B^{\alpha}_{\beta}\omega \wedge \omega^{\beta};$$

$$d\omega_{\alpha} = -\omega^{\beta}_{\alpha} \wedge \omega_{\beta} - B^{\beta}_{\alpha}\omega \wedge \omega_{\beta};$$

$$d\omega = 2\,B^{\alpha}_{\beta}\,\omega^{\beta} \wedge \omega_{\alpha},$$

we obtain the conditions that are necessary and sufficient for an almost contact structure on N^{2n-1} to be quasi-Sasakian:

$$1)\; B^{\alpha\beta\gamma} \;=\; 0 \;\; ;2)\; \sigma^{\alpha}_{\beta} \;=\; i\,B^{\alpha}_{\beta} \;\; ;3)\; B^{n\alpha\beta} = 0; \qquad (2)$$

$$4)\; -\sqrt{2}\,\tilde{B}^{n\alpha\beta} \;-\; \frac{1}{\sqrt{2}}\tilde{B}^{\alpha\beta n} + \,i\,\sigma^{\alpha\beta} = 0; 5)\; \tilde{B}^{n\beta n} \;-\, i\,\sigma^{\beta}_{n} = 0$$

and the formulae obtained by complex conjugation (no need to write them explicitly). From $(2)_3$ we obtain:

$$B^{n\alpha\beta} \;=\; 0 \quad \Rightarrow \tilde{B}^{n[\alpha\beta]} \;=\; 0 \quad \Rightarrow \tilde{B}^{n\alpha\beta} = \tilde{B}^{n\beta\alpha}.$$

By alternating $(2)_4$ we get:

$$0 = \sigma^{[\alpha\beta]} \;=\; -i\sqrt{2}\tilde{B}^{n[\alpha\beta]} \;-\frac{i}{\sqrt{2}}\;\tilde{B}^{[\alpha\beta]n} =$$

$$= -\frac{i}{2}\left(\sqrt{2}\tilde{B}^{n\alpha\beta} - \sqrt{2}\tilde{B}^{n\beta\alpha} + \frac{1}{\sqrt{2}}\tilde{B}^{\alpha\beta n} - \frac{1}{\sqrt{2}}\tilde{B}^{\beta\alpha n}\right) = -i\sqrt{2}\tilde{B}^{\alpha\beta n}.$$

We have $\tilde{B}^{\alpha\beta n} = 0$, from this equality we obtain $\sigma^{\alpha\beta} = -i\sqrt{2}\,\tilde{B}^{n\alpha\beta}$. That is why we can rewrite (2) as follows:

$$1)\ B^{\alpha\beta\gamma} = 0;\ \ 2)\ B^{n\alpha\beta} = 0;\ \ 3)\ \tilde{B}^{\alpha\beta n} = 0;\ \ 4)\ \sigma^{\alpha}_{\beta} = i\,B^{\alpha}_{\beta}; \tag{3}$$

$$5)\ \sigma^{\alpha\beta} = -i\sqrt{2}\,\tilde{B}^{n\alpha\beta};\ \ 6)\ \sigma^{\beta}_{n} = i\,\tilde{B}^{n\beta n}$$

and the formulae obtained by complex conjugation.

Now, let us fix a point $p \in M^{2n}$. If the hypersurface N^{2n-1} passes through this point, then the conditions (3) are fulfilled at this point. For $a, b, c = 1, ..., n$ we have:

$$2\,B^{abc} = \tilde{B}^{acb} - \tilde{B}^{abc};$$

$$2\,B^{bca} = \tilde{B}^{bac} - \tilde{B}^{bca};$$

$$2\,B^{cab} = \tilde{B}^{cba} - \tilde{B}^{cab}.$$

Adding the first and second equalities and subtracting the third equality, and also taking into account that the tensor $\tilde{B}^{abc}$ is skew-symmetric with respect to the indices a and b, we obtain

$$\tilde{B}^{abc} = -B^{abc} - B^{bca} + B^{cab}.$$

That is why we get:

$$-\tilde{B}^{\alpha\beta n} = B^{\alpha\beta n} + B^{\beta n\alpha} - B^{n\alpha\beta} = 0;$$

$$B^{nn\beta} + B^{n\beta n} + B^{\beta nn} = B^{nn\beta} - B^{nn\beta} = 0;$$

$$B^{\alpha\beta\gamma} + B^{\beta\gamma\alpha} + B^{\gamma\alpha\beta} = 0.$$

So, we have

$$B^{abc} + B^{bca} + B^{cab} = 0, \quad a, b, c = 1, ..., n. \tag{4}$$

It is not difficult to show that the condition (4) is equivalent to the closure of the fundamental form F , i.e. $dF = 0$, or

$$(dF)_p = 0; \quad p \in M^{2n}.$$

Indeed, it is known [7] that $F = -2i\,\omega^a \wedge \omega_a$. Therefore,

$$d\,F = -2i\,d\omega^a \wedge \omega_a + 2i\,\omega^a \wedge d\,\omega_a;$$

$$d\,F = -2i\,(\omega^a_b \wedge \omega^b \wedge \omega_a + B^{abc}\,\omega_b \wedge \omega_c \wedge \omega_a) +$$

$$+2\,i\,(\omega^a \wedge (-\omega^b_a \wedge \omega_b) + B_{abc}\,\omega^a \wedge \omega^b \wedge \omega^c);$$

$$d\,F = -2i\,(\omega^a_b \wedge \omega^b \wedge \omega_a + B^{[abc]}\,\omega_b \wedge \omega_c \wedge \omega_a) +$$

$$+2\,i\,(\,\omega_a^b \wedge \omega^a \wedge \omega_b) \;\; + \;\; B_{[abc]}\,\omega^a \wedge \omega^b \,\wedge \omega^c);$$

$$d\,F = -2i\,(B^{[abc]}\,\omega_b \wedge \omega_c \,\wedge \omega_a - B_{[abc]}\,\omega^a \wedge \omega^b \,\wedge \omega^c\,).$$

So,

$$d\,F = 0 \quad \Leftrightarrow \quad B^{[abc]} = B_{[abc]} = 0.$$

Taking into account that

$$B^{[abc]} = 0 \quad \Leftrightarrow \quad B^{abc} + B^{bca} + B^{cab} = 0,$$

we conclude

$$d\,F = 0 \quad \Leftrightarrow \quad B^{abc} \; + \; B^{bca} + \; B^{cab} = 0.$$

But the condition $B^{abc} \; + \; B^{bca} + \; B^{cab} = 0$ is the well-known [5, 12] criterion in terms of Kirichenko tensor for an arbitrary quasi-Kählerian manifold to be almost Kählerian (or to be a manifold of class W_2 in Gray–Hervella notation [10]).

We obtain that if a quasi-Sasakian hypersurface N^{2n-1} passes through an arbitrary point of a quasi-Kählerian manifold M^{2n}, then the condition $d\,F = 0$ holds at this point. That is why the manifold M^{2n} is almost Kählerian, Q.E.D.

4 Some comments

Using the well-known facts that the class of quasi-Kählerian manifolds contains the classes of nearly Kählerian (NK-) and almost Kählerian (AK-) manifolds [10], and class of Kählerian manifolds is the intersection of these classes:

$$K = NK \cap AK,$$

we can state the following consequence.

Corollary. If a nearly Kählerian manifold satisfys the quasi-Sasakian hypersurfaces axiom, then it is a Kählerian manifold.

We remark that this result was established by L. V. Stepanova [14] in a different way. Namely, this result was obtained from the fact that the class of Kählerian manifolds is also the intersection of the classes of nearly Kählerian and Hermitian manifolds:

$$K = NK \cap H.$$

We note that the theory of almost contact metric hypersurfaces of Hermitian manifolds (i.e. of almost Hermitian manifolds with integrable almost complex structure) is studied much better than the theory of almost contact metric hypersurfaces of quasi-Kählerian manifolds. The survey [7] contains a large list of papers on this subject.

Acknowledgment

The authors sincerely thank Professor Aligadzhi R. Rustanov for his useful comments on the subject of this paper.

References

[1] Abu-Saleem A., Banaru M. *Two theorems on Kenmotsu hypersurfaces in a W_3-manifold*, Stud. Univ. Babeş-Bolyai, Math., 2005, **51**, N3, 3–11.

[2] Abu-Saleem A., Banaru M. *Some applications of Kirichenko tensors*, Analele Univ. Oradea, Fasc. Mat., 2010, **17**, N2, 201–208.

[3] Abu-Saleem A., Banaru M. *On almost contact metric hypersurfaces of nearly Kählerian 6-sphere*, Malaysian Journal of Mathematical Sciences, 2014, **8**, N1, 35–46.

[4] Abu-Saleem A., Banaru M., Banaru G. *A note on 2-hypersurfaces of the nearly Kählerian six-sphere*, Buletinul Academiei de Ştiinţe a Republicii Moldova. Matematica, 2017, N3(**85**), 107–114.

[5] Banaru M. *On the Gray-Hervella classes of AH-structures on six-dimensional submanifolds of Cayley algebra*, Annuaire de l'Université de Sofia St. Kliment Ohridski. Faculté de Mathématiques et Informatique, 2004, **95**, 125–131.

[6] Banaru M. *Almost contact metric hypersurfaces with type number 0 or 1 in nearly-Kählerian manifolds*, Moscow University Mathematics Bulletin, 2014, **69**, N3, 132–134.

[7] Banaru M., Kirichenko V. *Almost contact metric structures on the hypersurface of almost Hermitian manifolds*, Journal of Mathematical Sciences (New York), 2015, **207**, N4, 513–537.

[8] Blair D. E. *The theory of quasi-Sasakian structures*, J. Diff. Geom., 1967, **1**, 331–345.

[9] Goldberg S. *Totally geodesic hypersurfaces of Kaehler manifolds*, Pacif. J. Math., 1968, **27**. N2, 275–281.

[10] Gray A., Hervella L. M. *The sixteen classes of almost Hermitian manifolds and their linear invariants*, Ann. Mat. Pura Appl., **123**, N4, 1980, 35–58.

[11] Kirichenko V. *The axiom of holomorphic planes in generalized Hermitian geometry*, Sov. Math. Dokl., 1981, **24**, 336–341.

[12] Kirichenko V. *Differential-geometric structures on manifolds*, Pechatnyi Dom, Odessa, 2013 (in Russian).

[13] Sasaki S. *On differentiable manifolds with certain structures which are closely related to almost contact structures*. I, Tohoku Math. J., 1960, **12**, N3, 459–476.

[14] Stepanova L. *Contact geometry of hypersurfaces of quasi-Kählerian manifolds*, Moscow State Pedagogical University V.I. Lenin (PhD thesis), 1995 (in Russian).

[15] Stepanova L., Banaru M. *On hypersurfaces of quasi-Kählerian manifolds*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi, Ser. Nouă, Mat., 2001, **47**, N1, 65–70.

[16] Stepanova L., Banaru M., Banaru G. *On geometry of QS-hypersurfaces of Kählerian manifolds*, Siberian Electronic Mathematical Reports, 2018, **15**, 815–822.

[17] Yanamoto H. *Quasi-Sasakian hypersurfaces in almost Hermitian manifolds*, Res. Rep. Nagaoka Tech. Coll. 1969, **5**, N2, 148–158.

Ahmad Abu-Saleem
Department of Mathematics Al al-Bayt University
Marfaq JORDAN
E-mail: *dr_ahmad57@yahoo.com*

Mihail B. Banaru, Galina A. Banaru
Smolensk State University
4, Przhevalsky Street, Smolensk – 214 000 RUSSIA
E-mail: *mihail.banaru@yahoo.com*

Lidia V. Stepanova
Smolensk Physical and Mathematical Lyceum under MIFI
5, Pamfilov Street, Smolensk – 214 018 RUSSIA
E-mail: *lide@yandex.ru*

Received April 5, 2020

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 76–87
ISSN 1024–7696

Laurent-Padé approximation for locating singularities of meromorphic functions with values given on simple closed contours

Capcelea Maria, Capcelea Titu

Abstract. In the present paper the Padé approximation with Laurent polynomials is examined for a meromorphic function on a finite domain of the complex plane. Values of the function are given at the points of a simple closed contour from this domain. Based on this approximation, an efficient numerical algorithm for locating singular points of the function is proposed.

Mathematics subject classification: 65E05, 41A21.
Keywords and phrases: Padé approximation, meromorphic function, simple closed contour, localization of singular points .

1 Introduction and problem formulation

Let consider a meromorphic function $f(z)$ defined on a finite domain $\Omega \subset \mathbb{C}$ that contains a simple closed contour Γ. The domain within the contour Γ is denoted by Ω^+, whilst the complementary domain to $\Omega^+ \cup \Gamma$ is denoted by $\Omega^- := \bar{\mathbb{C}} \setminus \{\Omega^+ \cup \Gamma\}$, where $\bar{\mathbb{C}}$ is the complex plane extended with the point at infinity. We consider that the point $z = 0 \in \Omega^+$.

According to the Riemann mapping theorem there exists a conformal map $z = \psi(w)$ of the domain $D^- := \{w \in \mathbb{C} : |w| > 1\}$ onto Ω^- such that $\psi(\infty) = \infty$, $\psi'(\infty) > 0$. The function $\psi(w)$ transforms the unit circle $\Gamma_0 := \{w \in \mathbb{C} : |w| = 1\}$ onto the contour Γ. Next, we consider that the points of the contour Γ are defined by means of the Riemann function $\psi(w)$.

The function $f(z)$ admits a finite number of singular points of polar type on the domain Ω, but their number and locations are not known. Also, on the contour Γ the function $f(z)$ can have both poles and jump discontinuity points (see Figure 1). Considering that the finite values $f_j := f(t_j)$ of the function $f(z)$ are known at the points $t_j \in \Gamma$ and that these values form a dense set on Γ, we aim to determine the locations of the singular points on Ω (including those on the contour Γ) for the function $f(z)$.

The approach applied here for singular points determination on Ω is based on the fact that the Padé approximation of a meromorphic function $f(z)$, defined on the domain Ω, allows to locate the poles of $f(z)$ on Ω. According to Montessus de Ballore's theorem [1,2], the poles of the sequence of Padé approximations to $f(z)$ form convergent sequences to the corresponding poles of the function $f(z)$ on Ω. But

it is necessary to keep in mind that although the mentioned theorem theoretically ensures pointwise convergence, numerically this convergence does not take place due to the influence of rounding errors.

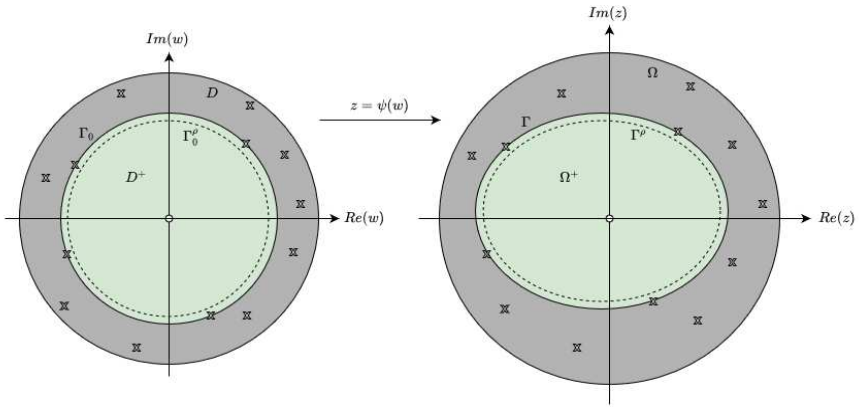


Figure 1: Data representation for the considered problem

Since jump discontinuities can be considered as apparent singularities (or "zero" order poles), they can be detected according to the same approach of searching for the poles of $f(z)$.

In Section 2 we set out the theoretical basis for the algorithm of locating the singular points of the function, based on the Laurent-Padé approximation of $f(z)$. In Section 3 we examine a formula for numerical approximation of the Laurent coefficients from the Padé approximation, and in Section 4 an algorithm for evaluating the number of poles of f on $\Omega^+ \cup \Gamma$ and $\Omega^- \cup \Gamma$. In Section 5 we consider some numerical examples that confirm the efficiency of the proposed algorithm and simultaneously show some of its numerical difficulties.

2 Detection of singularities based on Laurent-Padé approximation

Let consider that the function $f(z)$ is analytic in the annulus $A := \{z \in \mathbb{C} : r < |z| < R\}$, $r > 0$, $R < \infty$. Then the Laurent series of the function $f(z)$, $f(z) = \sum_{k=-\infty}^{\infty} c_k z^k$, can be represented in the form

$$f(z) = f^+(z) + f^-(z), \ z \in A, \tag{1}$$

where the function $f^+(z) = c_0/2 + \sum_{k=1}^{\infty} c_k z^k$ is analytic on the set $\{z \in \mathbb{C} : |z| < R\}$ and $f^-(z) = c_0/2 + \sum_{k=-\infty}^{-1} c_k z^k$ is analytic on $\{z \in \mathbb{C} : |z| > r\}$.

Next, for $M, N \in \mathrm{N}$ such that $N \geq M$, we consider the Laurent-Padé approximation of order (N, M) for $f(z)$ [1], that is defined as the sum of two classical Padé approximations [2] for $f^+(z)$ at $z = 0$ and, respectively, for $f^-(z)$ at $z = \infty$. Thus, two series are approximated: one is a power series in z and the other is a power series in $1/z$. Then the obtained approximations are summed.

Let $R^+_{(N_1,M_1)}(z) = P^+_{N_1}(z) \Big/ Q^+_{M_1}(z)$ be the Padé approximation of order (N_1, M_1), $N_1 \geq M_1$ for the function $f^+(z)$ at $z = 0$ that satisfies the following condition

$$Q^+_{M_1}(z) f^+(z) - P^+_{N_1}(z) = \mathcal{O}\left(z^{M_1+N_1+1}\right). \tag{2}$$

The polynomials $P^+_{N_1}(z)$ and $Q^+_{M_1}(z)$ are of the form $P^+_{N_1}(z) = \sum_{k=0}^{N_1} p^+_k z^k$, $Q^+_{M_1}(z) = \sum_{j=0}^{M_1} q^+_j z^j$ and we consider that they do not have common zeros. In order to avoid division by zero in the relation for $R^+_{(N_1,M_1)}(z)$, the normalization condition $q^+_0 = 1$ is imposed.

Analogously we define the (N_2, M_2) - order Padé approximation for the function $f^-(z)$ at $z = \infty$, $R^-_{(N_2,M_2)}(z) = P^-_{N_2}(1/z) \Big/ Q^-_{M_2}(1/z)$, where $P^-_{N_2}(1/z) = \sum_{k=0}^{N_2} p^-_k (1/z)^k$, $Q^-_{M_2}(1/z) = \sum_{j=0}^{M_2} q^-_j (1/z)^j$, $q^-_0 = 1$, and the coefficients of the approximation are determined from the following condition

$$Q^-_{M_2}(1/z) f^-(z) - P^-_{N_2}(1/z) = \mathcal{O}\left(z^{-(M_2+N_2+1)}\right). \tag{3}$$

The Laurent-Padé approximation of order (N, M) for $f(z)$ is defined as follows:

$$R_{(N,M)}(z) := R^+_{(N_1,M_1)}(z) + R^-_{(N_2,M_2)}(z) = P_{N_1,N_2}(z) / Q_{M_1,M_2}(z),$$

where $P_{N_1,N_2}(z) := P^+_{N_1}(z) Q^-_{M_2}(1/z) + P^-_{N_2}(1/z) Q^+_{M_1}(z)$ and $Q_{M_1,M_2}(z) := Q^+_{M_1}(z) Q^-_{M_2}(1/z)$ are Laurent polynomials of the form $\sum_{j=-N_2}^{N_1} p_j z^j$ and, respectively, $\sum_{j=-M_2}^{M_1} q_j z^j$, and $N = \max(N_1, N_2)$, $M = \max(M_1, M_2)$. Taking into account the relation for $R_{(N,M)}(z)$ we can deduce that the zeros of $Q^+_{M_1}(z)$ and $Q^-_{M_2}(1/z)$ are approximations of the singular points of the function $f(z)$.

Let the function $f(z)$ allow a meromorphic extension on the annulus $\tilde{A} := \left\{z \in \mathbb{C} : \ \tilde{r} < |z| < \tilde{R}\right\}$ so that the poles of $f(z)$ belong to the set $\tilde{A}$, more precisely to the annulus $A_1 := \left\{z \in \mathbb{C} : \ R \leq |z| < \tilde{R}\right\}$ and $A_2 := \{z \in \mathbb{C} : \ \tilde{r} < |z| \leq r\}$ (see Figure 2). Taken into account the relation (1), we can conclude that the function $f^+(z)$ admits a meromorphic extension on the set A_1 because $f^-(z)$ is analytic on $\{z \in \mathbb{C} : \ |z| > r\}$. Analogously it can be shown that the function $f^-(z)$ admits a meromorphic extension on the set A_2.

Based on the Montessus de Ballore theorem it can be shown that the condition $\lim\limits_{N_1 \to \infty} R^+_{(N_1,M_1)}(z) = f^+(z)$ is satisfied, that means that the Padé approximation $R^+_{(N_1,M_1)}(z)$ converges to $f^+(z)$ for a sufficiently large and fixed value M_1 and $N_1 \to \infty$. Convergence is locally uniform over the set $\left\{z \in \mathbb{C} : \ |z| < \tilde{R}\right\} \setminus \left\{z^+_j\right\}_{j=1}^{k^+}$, where z^+_j, $j = 1, ..., k^+$, are the poles of the function $f^+(z)$ of multiplicities l^+_j correspondingly. Moreover, the condition $\lim\limits_{N_1 \to \infty} Q^+_{M_1}(z) = \prod_{j=1}^{k^+} \left(1 - z \Big/ z^+_j\right)^{l^+_j}$ is satisfied. Analogously it can be shown that $\lim\limits_{N_2 \to \infty} R^-_{(N_2,M_2)}(z) = f^-(z)$ and

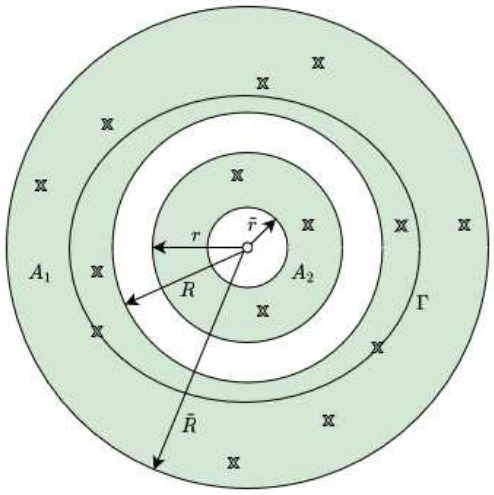


Figure 2: Regions of meromorphic extension for the function f

$\lim_{N_2\to\infty} Q^-_{M_2}(1/z) = \prod_{j=1}^{k^-}\left(1 - z_j^- \Big/ z\right)^{l_j^-}$, where z_j^-, $j = 1, ..., k^-$, are the poles of the function $f^-(z)$ of multiplicities l_j^- correspondingly. The zeros of polynomials $Q^+_{M_1}(z)$ and $Q^-_{M_2}(1/z)$ are approximations of the poles of the function $f(z)$. Thus, the zeros of the polynomial $Q^+_{M_1}(z)$ approximate the poles from the domain $\Omega^- \cup \Gamma$ and the zeros of $Q^-_{M_2}(1/z)$ - the poles from $\Omega^+ \cup \Gamma$.

The relations (2) and (3) can be used to determine the coefficients p_k^+, q_j^+, $k = 0, 1, ..., N_1$, $j = 0, 1, ..., M_1$, and p_r^-, q_m^-, $r = 0, 1, ..., N_2$, $m = 0, 1, ..., M_2$, of the polynomials $P^+_{N_1}(z), Q^+_{M_1}(z)$ and $P^-_{N_2}(1/z), Q^-_{M_2}(1/z)$, correspondingly. Based on the relation (2) we obtain the system of linear algebraic equations

$$\sum_{j=0}^{k-1} c_{k-j} q_j^+ + 0.5 c_0 q_k^+ - p_k^+ = 0,\ k = 0, 1, ..., M_1,$$

$$\sum_{j=0}^{M_1} c_{k-j} q_j^+ - p_k^+ = 0,\ k = M_1 + 1, ..., N_1,$$

$$\sum_{j=0}^{M_1} c_{k-j} q_j^+ = 0,\ k = N_1 + 1, ..., N_1 + M_1.$$

The coefficients q_j^+, $j = 1, ..., M_1$, of the polynomial $Q^+_{M_1}(z)$ are determined using only the relations $\sum_{j=0}^{M_1} c_{k-j} q_j^+ = 0$, $k = N_1 + 1, ..., N_1 + M_1$, that define a system of M_1 linear algebraic equations with $M_1 + 1$ unknowns and a Toeplitz matrix as a coefficient matrix of the system. Taking into account the normalization condition

$q_0^+ = 1$, the relation that defines the system is written as

$$\sum_{j=1}^{M_1} c_{k-j} q_j^+ = -c_k,\ k = N_1 + 1, ..., N_1 + M_1. \tag{4}$$

The system (4) has a unique solution.

Analogously from condition (3) we obtain the system of linear algebraic equations

$$\sum_{j=1}^{M_2} c_{j-k} q_j^- = -c_{-k},\ k = N_2 + 1, ..., N_2 + M_2. \tag{5}$$

The components q_j^-, $j = 1, ..., M_2$, of the solution of the system (5) (together with $q_0^- = 1$) are the coefficients of the polynomial $Q_{M_2}^-(1/z)$. The zeros of $Q_{M_2}^-(1/z)$ are obtained by inverting the zeros of the polynomial $Q_{M_2}^-(t) = \sum_{j=0}^{M_2} q_j^- t^j$.

In order to find the solutions of the systems (4) and (5) it is necessary firstly to determine the Laurent coefficients $c_{N_1-M_1+1}, ..., c_{N_1+M_1}$ and $c_{-(N_2+M_2)}, ..., c_{-(N_2-M_2+1)}$.

3 Numerical approximation of Laurent coefficients

In this section, we examine a formula for numerical approximation of Laurent coefficients $c_k = \frac{1}{2\pi i} \int_\Gamma f(z) z^{-k-1} dz$, where Γ is a simple closed contour that satisfies the conditions mentioned in Section 1. The values f_j of the function $f(z)$ on the contour Γ are used for the calculation of the Laurent coefficients that define the coefficient matrices of the systems (4) and (5).

Since the function $f(z)$ can have a finite number of poles on the contour Γ, we insignificantly disturb the contour and consider that the values f_j of the continuous function $f(z)$ on the disturbed contour Γ^ρ are known. Since the contour Γ is defined by the Riemann function $\psi(w)$, that transforms the unit circle $\Gamma_0 = \{w \in \mathbb{C} : |w| = 1\}$ onto the contour Γ, we define the perturbed contour Γ^ρ as follows:

$$\Gamma^\rho := \{t \in \mathbb{C} :\ t = \psi(w),\ w \in \Gamma_0^\rho\},$$

where $\Gamma_0^\rho := \{w \in \mathbb{C} :\ |w| = \rho,\ \rho = 1 \pm \varepsilon,\ \varepsilon > 0\}$. We consider small values for ε, for example, $\varepsilon = 0.001$.

The Laurent coefficient c_k is approximated as follows:

$$c_k \approx \frac{1}{2\pi i} \int_{\Gamma^\rho} f(z) z^{-k-1} dz =$$

$$= \frac{1}{2\pi i} \int_{\Gamma_0^\rho} (\psi(w))^{-k-1} f(\psi(w)) \psi'(w) dw = \frac{1}{2\pi i} \int_{\Gamma_0^\rho} g(w) dw,$$

where $g(w) := (\psi(w))^{-k-1} f(\psi(w)) \psi'(w),\ w \in \Gamma_0^\rho$.

We approximate the function $g(w)$ with Lagrange interpolation polynomial

$$(L_n g)(w) = \sum_{s=-n}^{n} \Lambda_s w^s,\ \Lambda_s = \frac{1}{2n+1} \sum_{j=0}^{2n} g(w_j)\, w_j^{-s},\ s = -n, -n+1, ..., n,$$

defined on the uniform set of nodes on Γ_0^ρ, $w_j = \rho e^{(2\pi j/(2n+1))i}$, $i^2 = -1$, $j = 0, 1, ..., 2n$. Then, taking into account the relation $\frac{1}{2\pi i}\int_{\Gamma_0^\rho} w^s dw = \begin{cases} 0, & for\ s \neq -1 \\ 1, & for\ s = -1 \end{cases}$, we can write the approximation for the coefficient c_k as follows:

$$c_k \approx \frac{1}{2\pi i}\int_{\Gamma_0^\rho} g(w)\, dw \approx c_k^{(n)} := \frac{1}{2\pi i}\int_{\Gamma_0^\rho} \sum_{s=-n}^{n} \Lambda_s w^s dw =$$

$$= \sum_{s=-n}^{n} \Lambda_s \frac{1}{2\pi i}\int_{\Gamma_0^\rho} w^s dw = \Lambda_{-1} = \frac{1}{2n+1} \sum_{j=0}^{2n} g(w_j)\, w_j.$$

If we apply the approach presented in [3] for estimating the error of approximation of the coefficient c_k by $c_k^{(n)}$, then it can be shown that when $n \to \infty$ the sequence of approximations $c_k^{(n)}$ converges to c_k with rate of a geometric progression.

4 Evaluating the number of poles

The solutions q_j^+, $j = 0, 1, ..., M_1$, and q_j^-, $j = 0, 1, ..., M_2$, of the systems of equations (4) and, respectively, (5), are the coefficients of the polynomials $Q_{M_1}^+(z)$ and, respectively, $Q_{M_2}^-(1/z)$, whose zeros are considered as approximations of the poles of the function $f(z)$. Thus, the algorithm based on the Laurent-Padé approximation generates $M_1 + M_2$ approximations of the poles of the function $f(z)$.

The zeros of the polynomial $Q_{M_1}^+(z)$ are approximations of the poles that belong to the domain $\Omega^+ \cup \Gamma$ and the zeros of $Q_{M_2}^-(1/z)$ are approximations of the poles that belong to $\Omega^- \cup \Gamma$. In order the Laurent-Padé approximation algorithm to be applied efficiently, it is necessary to know in advance the number of poles of $f(z)$ belonging to $\Omega^+ \cup \Gamma$ and $\Omega^- \cup \Gamma$, respectively. We denote these values by m_1 and m_2, respectively. Next, we present an approach that allows evaluation of values m_1 and m_2.

Consider the matrices of the systems of equations (4) and (5), that have the orders $M_1 \times M_1$ and $M_2 \times M_2$, respectively, i.e.

$$C_1 = \begin{pmatrix} c_{N_1} & c_{N_1-1} & \cdots & c_{N_1-(M_1-1)} \\ c_{N_1+1} & c_{N_1} & \cdots & c_{N_1-(M_1-2)} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N_1+M_1-1} & c_{N_1+M_1-2} & \cdots & c_{N_1} \end{pmatrix}$$

and

$$C_2 = \begin{pmatrix} c_{-N_2} & c_{-(N_2-1)} & \cdots & c_{-(N_2-(M_2-1))} \\ c_{-(N_2+1)} & c_{-N_2} & \cdots & c_{-(N_2-(M_2-2))} \\ \vdots & \vdots & \ddots & \vdots \\ c_{-(N_2+M_2-1)} & c_{-(N_2+M_2-2)} & \cdots & c_{-N_2} \end{pmatrix}.$$

The idea on which the mentioned approach is built is that the determinants of the matrices C_1 and C_2 are non-zero when $M_1 = m_1$ and $M_2 = m_2$, respectively, and the determinants of higher order matrices (which are obtained by adding rows and columns) have close to zero values.

Thus, we define a parameter $m > 0$ that means an estimate for the maximum value of the number of poles for $f(z)$ on $\Omega^+ \cup \Gamma$, $m_1 \leq m$ (respectively, $m_2 \leq m$ on $\Omega^- \cup \Gamma$). If for the initial value m we have nonzero determinant, then we consider a higher value for m. For values $m, m-1, m-2, ...$ we calculate the determinants of the Toeplitz matrices C_1 (respectively, C_2) of order $m \times m$, until we have $\det C_1 \neq 0$ (respectively, $\det C_2 \neq 0$). Obviously, a sufficiently small parameter $\delta > 0$ must be used to evaluate the last condition, for example, $\delta = 0.001$. The first value of the parameter m for which $\det C_1 \neq 0$ (respectively, $\det C_2 \neq 0$), represents an evaluation of the number of poles on $\Omega^+ \cup \Gamma$ (respectively, on $\Omega^- \cup \Gamma$).

Numerical experiments show that the accuracy of the result depends on the value of the parameters $N_1 \geq m$ $(N_2 \geq m)$ and δ as well as on the presence of multiple poles and discontinuity points.

5 Numerical examples

In this section we analyze some numerical examples that confirm the correctness of the examined localization algorithm and provide us with necessary data to analyze its efficiency. Below we examine four test functions with singularities that highlight the possibilities of the algorithm:

1. the case when $f(z)$ has only simple poles on Ω;
2. the case when $f(z)$ has multiple poles on Ω;
3. the case when $f(z)$ has points of jump discontinuity on Ω;
4. the case when $f(z)$ has both poles and discontinuity points on Ω.

In our examples we consider that the Riemann function $z = \psi(w)$ that performs the conformal transformation of the set $\{ w \in \mathbb{C} \mid |w| > 1 \}$ on the domain Ω^- from the outside of the contour Γ is $\psi(w) = w + 1/\left(3w^3\right)$. Thus, $\psi(w)$ transforms the circle Γ_0 onto the astroid Γ.

We consider that the values f_j of the examined function $f(z)$ are given at the points

$$z_j = \psi\left(\rho e^{i\theta_j}\right) \in \Gamma^\rho,\ \theta_j = 2\pi j/m,\ m \in \mathbb{N},\ j = 0, 1, ..., m,$$

where $\rho = 1$.

The following notations are used in the below graphical representations:
$\square$ is location of the pole;
$\diamond$ $(\times)$ is the approximation of the pole determined by the Padé approximation of $f^+(z)$ $(f^-(z))$;
$\circ$ is jump discontinuity point on the contour Γ.

The following parameter values are considered in the algorithm that evaluates the number of singular points $N_1 = m+1,\ N_2 = m+1,\ \delta = 10^{-3}$, where m changes starting with the value $m^* = 10$, i.e. $m = 10, 9, 8, ...$

Example 1. The function of a complex variable $f(z)$ is defined as follows:

$$f(z) = \frac{\cos\left(z^2\right)}{(z - zi_1)(z - zi_2)(z - ze_1)(z - ze_2)(z - zc_1)(z - zc_2)(z - zc_3)},$$

where

$$zc_1 = \psi\left(e^{\pi i/4}\right),\ zc_2 = \psi\left(e^{6\pi i/7}\right),\ zc_3 = \psi\left(e^{9\pi i/7}\right),$$

$$zi_1 = 0.3 + 0.1i,\ zi_2 = -0.4 - 0.2i,\ ze_1 = \psi(1.3 + 0.7i),\ ze_2 = \psi(-1.5 - 0.75i),$$

there are seven simple poles (two on Ω^-, three on the contour Γ and two on Ω^+) of $f(z)$, $i^2 = -1$.

The algorithm for evaluating the number of singular points of $f(z)$ returns $m_1 = 5$ for the number of poles on $\Omega^+ \cup \Gamma$ and $m_2 = 6$ for the number of poles on $\Omega^- \cup \Gamma$.

For the function $f^+(z)$ we apply the Padé approximation of order $(N_1, M_1) = (6, 5)$ at $z = 0$, and for $f^-(z)$ - the Padé approximation of order $(N_2, M_2) = (5, 5)$ at $z = \infty$. The approximations for the poles obtained by examined algorithm are presented in Table 1 and Figure 3.

Poles of the function $f(z)$		*Approximations of the poles of $f^+(z)$*	*Approximations of the poles of $f^-(z)$*
Poles from Ω^-	1.3092 + 0.5968i	1.3100 + 0.5907i	—
	-1.5126 - 0.6805i	-1.5164 - 0.6818i	—
Poles on Γ	0.4714 + 0.4714i	0.4714 + 0.4714i	0.4714 + 0.4714i
	-0.9751 + 0.1089i	-0.9753 + 0.1091i	-0.9751 + 0.1089i
	-0.3232 - 0.6372i	-0.3232 - 0.6372i	-0.3232 - 0.6372i
Poles from Ω^+	0.3000 + 0.1000i	—	0.3000 + 0.1000i
	-0.4000 - 0.2000i	—	-0.4000 - 0.2000i

Table 1: The approximations obtained in Example 1

We can see that if the function $f(z)$ has only simple poles on Ω, then the localization algorithm converges rapidly, requiring values for N_1, M_1, N_2, M_2 close to the number of poles of $f(z)$.

Example 2. The function of a complex variable $f(z)$ is defined as follows:

$$f(z) = \frac{\cos\left(z^2\right)}{(z - zi_1)^3 (z - zi_2)(z - ze_1)^2 (z - ze_2)(z - zc_1)(z - zc_2)^2 (z - zc_3)},$$

where

$$zc_1 = \psi\left(e^{\pi i/4}\right),\ zc_2 = \psi\left(e^{6\pi i/7}\right),\ zc_3 = \psi\left(e^{9\pi i/7}\right)$$

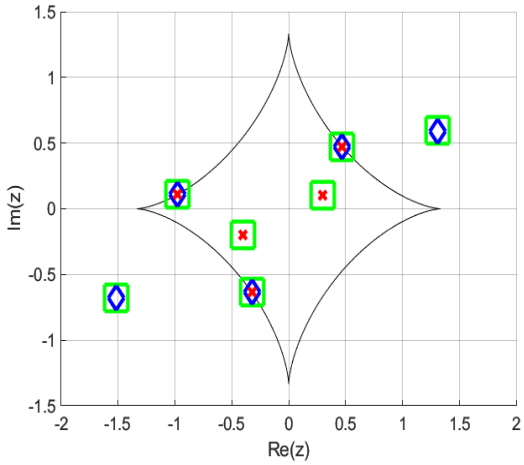


Figure 3: Approximations of the poles in Example 1

$$zi_1 = 0.3 + 0.1i,\ zi_2 = -0.4 - 0.2i,\ ze_1 = \psi\left(1.3 + 0.7i\right),\ ze_2 = \psi\left(-1.5 - 0.75i\right),$$

there are seven poles (two on Ω^-, three on the contour Γ and two on Ω^+) of $f\left(z\right)$, $i^2 = -1$. The poles ze_1 and zc_2 are of the second order, zi_1 is of the third order, and the other poles are simple.

The algorithm for evaluating the number of singular points of $f\left(z\right)$ returns $m_1 = 7$ for the number of poles on $\Omega^+ \cup \Gamma$ and $m_2 = 7$ for the number of poles on $\Omega^- \cup \Gamma$. Thus, the values m_1 and m_2 returned by the algorithm also include the multiplicities of the poles.

For the function $f^+\left(z\right)$ we apply the Padé approximation of order $(N_1, M_1) = (8, 7)$ at $z = 0$, and for $f^-\left(z\right)$ – the Padé approximation of order $(N_2, M_2) = (8, 8)$ at $z = \infty$. The approximations for the poles obtained by examined algorithm are presented in Table 2 and Figure 4.

Poles of the function $f\left(z\right)$		*Approximations of the poles of $f^+\left(z\right)$*	*Approximations of the poles of $f^-\left(z\right)$*
Poles from Ω^-	1.3092 + 0.5968i	1.3008 + 0.6365i	—
		1.3044 + 0.5583i	—
	-1.5126 - 0.6805i	-1.5029 - 0.6807i	—
Poles on Γ	0.4714 + 0.4714i	0.4714 + 0.4714i	0.4714 + 0.4714i
	-0.9751 + 0.1089i	-0.9751 + 0.1090i	-0.9751 + 0.1089i
		-0.9314 + 0.1182i	-0.9751 + 0.1089i
	-0.3232 - 0.6372i	-0.3232 - 0.6372i	-0.3232 - 0.6372i
Poles from Ω^+	0.3000 + 0.1000i	—	0.2997 + 0.0999i
		—	0.3000 + 0.1003i
		—	0.3002 + 0.0998i
	-0.4000 - 0.2000i	—	-0.4000 - 0.2000i

Table 2: The approximations obtained in Example 2

For each multiple pole a number of approximations is generated corresponding to its order and this requires the Padé approximations to have higher orders.

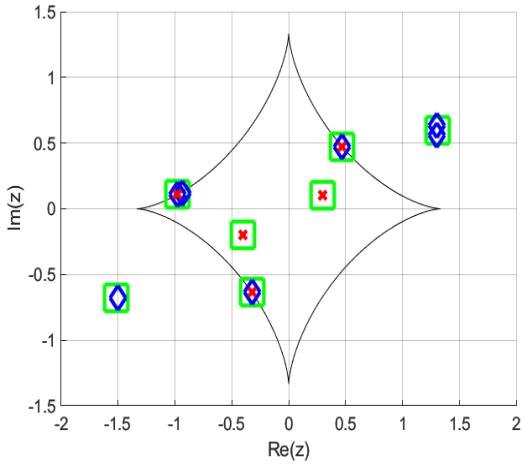


Figure 4: Approximations of the poles in Example 2

Example 3. The function of a complex variable $f(z)$ is defined as follows:

$$f(z) = \begin{cases} z^2, & for\ \ \theta \in [\zeta_2, \zeta_1] \\ \cos\left(z^2\right), & for\ \ \theta \in (\zeta_1, 2\pi + \zeta_2) \end{cases},$$

where $\zeta_1 = 7\pi/4, \zeta_2 = 7\pi/10$. We can see that $f(z)$ has two points of jump discontinuity on Γ (see Figure 5).

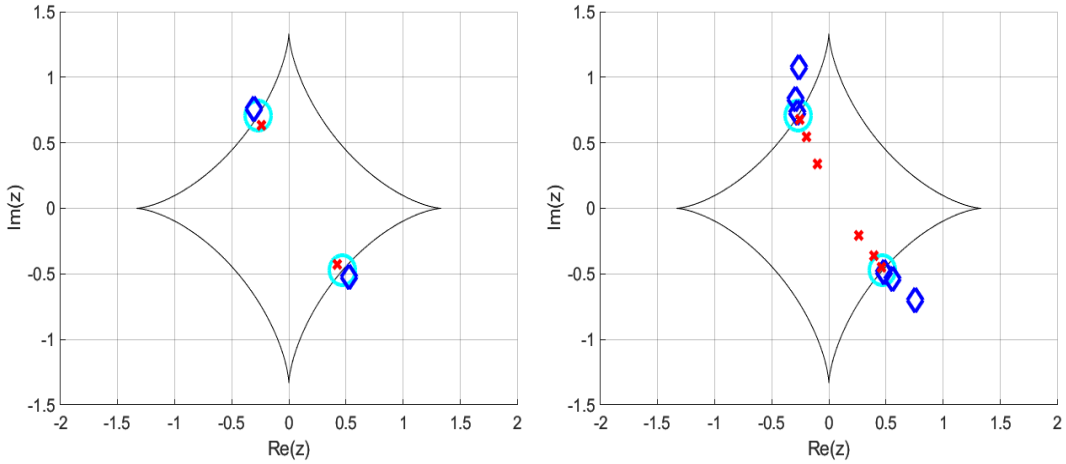


Figure 5: Approximations of the discontinuity points in Example 3

The algorithm for evaluating the number of singular points of $f(z)$ returns $m_1 = 2$ for the number of poles on $\Omega^+ \cup \Gamma$ and $m_2 = 5$ for the number of poles on $\Omega^- \cup \Gamma$. So, we observe that the algorithm gives wrong estimates in the case when the function has discontinuity points.

We apply Padé approximation of order $(N_1, M_1) = (10, 2)$ to the function $f^+(z)$ and Padé approximation of order $(N_2, M_2) = (8, 2)$ to $f^-(z)$. The approximations

obtained for the discontinuity points of the function $f(z)$ on Γ are presented in Figure 5 (left).

The approximations for $(N_1, M_1) = (16, 6)$ and $(N_2, M_2) = (8, 6)$ are presented in Figure 5 (right). It can be seen that in the neighborhood of each discontinuity point a convergent sequence of approximations is generated.

Example 4. The function of a complex variable $f(z)$ is defined as follows:

$$f(z) = \begin{cases} \frac{z^2}{(z-zi_2)(z-ze_2)(z-zc_2)(z-zc_3)}, & for\ \ \theta \in [\zeta_2, \zeta_1] \\ \frac{\cos(z^2)}{(z-zi_1)(z-zc_1)(z-ze_1)}, & for\ \ \theta \in (\zeta_1, 2\pi + \zeta_2) \end{cases},$$

where

$$\zeta_1 = 7\pi/4, \zeta_2 = 7\pi/10,\ zc_1 = \psi\left(e^{\pi i/4}\right),\ zc_2 = \psi\left(e^{6\pi i/7}\right),\ zc_3 = \psi\left(e^{9\pi i/7}\right),$$

$$zi_1 = 0.3 + 0.1i,\ zi_2 = -0.4 - 0.2i,\ ze_1 = \psi(1.3 + 0.7i),\ ze_2 = \psi(-1.5 - 0.75i).$$

The function $f(z)$ has two poles inside Γ, two poles outside of Γ and three poles on the contour Γ (all simple), as well as two jump discontinuity points on Γ (see Figure 6).

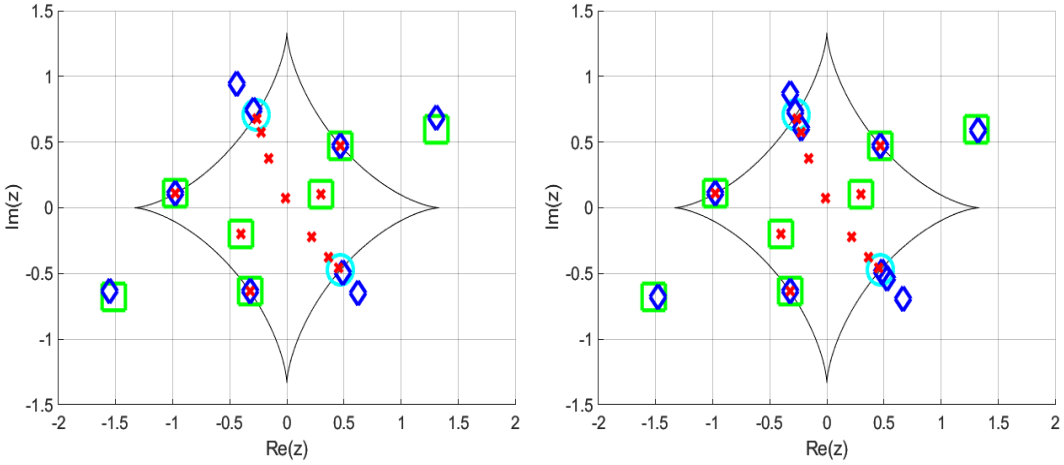


Figure 6: Approximations of the poles and discontinuity points in Example 4

The algorithm for evaluating the number of singular points of $f(z)$ returns $m_1 = 5$ for the number of poles on $\Omega^+ \cup \Gamma$ and $m_2 = 10$ for the number of poles on $\Omega^- \cup \Gamma$. The obtained numerical results confirm the idea that the examined algorithm gives a correct evaluation for the number of poles only in the case when the function has no discontinuity points.

We apply Padé approximation of order $(N_1, M_1) = (11, 9)$ to $f^+(z)$ and Padé approximation of order $(N_2, M_2) = (12, 12)$ to $f^-(z)$. The approximations obtained for the poles and discontinuity points of the function $f(z)$ are presented in Figure 6 (left).

It can be seen that in the neighborhood of each discontinuity point a convergent sequence of approximations is generated and the number of elements of the sequence increases simultaneously with the order of approximations. The approximations for $(N_1, M_1) = (17, 11)$ and $(N_2, M_2) = (12, 12)$ are presented in Figure 6 (right).

It should be noted that the presence of discontinuity points essentially reduces the convergence speed of the examined localization algorithm. This is explained by the fact that the discontinuity points "attract" the generated approximations and do not allow them to locate quickly certain poles.

As the problem of constructing the Padé approximation is generally poorly conditioned, then amplifying the order of the Padé approximations for the functions $f^+(z)$ and $f^-(z)$ usually leads to the appearance of spurious poles. Most of the spurious poles can be removed by applying the residual analysis [4].

References

[1] Bultheel A. *Laurent series and their Padé approximations*. Birkhäuser Verlag, Basel-Boston, 1987.

[2] Baker G., Graves-Morris P. *Padé approximants*. Second ed., Cambridge University Press, 1996.

[3] Lyness J. N. *Differentiation formulas for analytic functions*. Math. Comp., 1968, **22**, no. 102, 352–362.

[4] Gonnet P., Guttel S., Trefethen L. N. *Robust Padé approximation via SVD*. SIAM Review, 2013, **55**, no. 1, 101–117.

Capcelea Maria, Capcelea Titu
Moldova State University
E-mail: *mariacapcelea@yahoo.com; titu.capcelea@gmail.com*

Received May 22, 2020

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(93), 2020, Pages 88–101
ISSN 1024–7696

Asymmetric Separation of Convex Sets

Valeriu Soltan

Abstract. Based on various types of asymmetric hyperplane separation of a given pair of convex sets K_1 and K_2 in the n-dimensional Euclidean space, we derive a uniform description of existing types of separation. Our argument uses properties of the polar cone $(K_1 - K_2)^\circ$. Also, we consider asymmetric separation of convex cones with a common apex.

Mathematics subject classification: 52A20.
Keywords and phrases: Separation, hyperplane, convex, cone.

1 Introduction

Support and separation properties of convex sets are among the most important topics in convexity. Studied by Minkowski [7, 8] on the turn of 20th century, they became useful tools in many mathematical disciplines, especially in convex geometry, linear analysis, convex analysis, and optimization.

We recall that nonempty convex sets K_1 and K_2 in the n-dimensional Euclidean space $\mathbb{R}^n$ are *separated* by a hyperplane $H \subset \mathbb{R}^n$ provided K_1 and K_2 lie in the opposite closed halfspaces determined by H. The concept of separation was gradually refined in the literature, and various types of separation of convex sets are known nowadays. The existing classification and the respective terminology in this regard is mainly due to Klee [5] and Rockafellar [9]. For instance, K_1 and K_2 are called *properly* separated provided $K_1 \cup K_2 \not\subset H$, and they are called *strongly* separated if suitable open ρ-neighborhoods $U_\rho(K_1)$ and $U_\rho(K_2)$ of these sets are separated by H. These two types of separation are the most popular in the literature due to the existence of simple criteria (see [9], § 11), and many other types of separation are often viewed as their derivatives.

In this paper, we deal with an alternative approach to the classification of separating hyperplanes. Namely, we start with various types of asymmetric hyperplane separation of convex sets, and then derive from them existing types of separation. Also, unlike many existing results, which provide conditions for the existence of at least one separating hyperplane, we tend to describe all such hyperplanes.

2 Preliminaries

This section contains necessary definitions, notation, and results on convex sets in $\mathbb{R}^n$ (see, e. g., [9] and [11] for details). The elements of $\mathbb{R}^n$ are called vectors, or

points. We denote by $[u, v]$ and (u, v) the closed and open segments with endpoints $u, v \in \mathbb{R}^n$. Also, $u \cdot v$ will mean the dot product of u and v. The zero vector of $\mathbb{R}^n$ is denoted o. A set $L \subset \mathbb{R}^n$ is called an r-dimensional plane if it is translate of a suitable r-dimensional subspace S of $\mathbb{R}^n$: $L = c + S$, where $c \in \mathbb{R}^n$.

In what follows, K stands for a nonempty convex set in $\mathbb{R}^n$. The open ρ-neighborhood of K, denoted $U_\rho(K)$, is the union of all open balls $U_\rho(x)$ of radius $\rho > 0$ centered at $x \in K$. Convex sets K_1 and K_2 are called strongly disjoint provided $U_\rho(K_1) \cap U_\rho(K_2) = \varnothing$ for a suitable $\rho > 0$; the latter occurs if and only if the *inf*-distance $\delta(K_1, K_2)$, defined by

$$\delta(K_1, K_2) = \inf\{\|x_1 - x_2\| : x_1 \in K_1, x_2 \in K_2\},$$

is positive. The notations $\operatorname{cl} K$, $\operatorname{int} K$, $\operatorname{rint} K$, $\operatorname{rbd} K$, and $K^\perp$ stand, respectively, for the closure, interior, relative interior, relative boundary, and the orthogonal complement of K. The linear span of K, denoted $\operatorname{span} K$, is the smallest subspace containing K, affine span of K, denoted $\operatorname{aff} K$, is the intersection of all planes containing K, and $\dim K$ is defined as the dimension of $\operatorname{aff} K$. Also, the direction space and the orthospace of K are defined by $\operatorname{dir} K = \operatorname{span}(K - K)$ and $\operatorname{ort} K = (\operatorname{dir} K)^\perp$, respectively.

A hyperplane in $\mathbb{R}^n$ is a plane which can be described as

$$H = \{x \in \mathbb{R}^n : x \cdot e = \gamma\}, \quad e \neq o, \quad \gamma \in \mathbb{R}. \tag{1}$$

Consequently, a hyperplane of the form (1) is a translate of the hypersubspace

$$S = \{x \in \mathbb{R}^n : x \cdot e = 0\}, \quad e \neq o. \tag{2}$$

Every hyperplane of the form (1) determines a pair of opposite closed halfspaces

$$V_1 = \{x \in \mathbb{R}^n : x \cdot e \leq \gamma\} \quad \text{and} \quad V_2 = \{x \in \mathbb{R}^n : x \cdot e \geq \gamma\} \tag{3}$$

and a pair of opposite open halfspaces

$$W_1 = \{x \in \mathbb{R}^n : x \cdot e < \gamma\} \quad \text{and} \quad W_2 = \{x \in \mathbb{R}^n : x \cdot e > \gamma\}. \tag{4}$$

The (negative) polar cone of a convex set $K \subset \mathbb{R}^n$ is the set

$$K^\circ = \{e \in \mathbb{R}^n : x \cdot e \leq 0 \ \text{ for all } \ x \in K\}.$$

The recession cone of K is defined by

$$\operatorname{rec} K = \{e \in \mathbb{R}^n : x + \lambda e \in K \text{ whenever } x \in K \text{ and } \lambda \geq 0\},$$

and the lineality space of K is the subspace given by $\operatorname{lin} K = \operatorname{rec} K \cap (-\operatorname{rec} K)$.

3 Asymmetric Separation

Definition 1. Let convex sets K_1 and K_2 in $\mathbb{R}^n$ be separated by a hyperplane $H \subset \mathbb{R}^n$. We will say that

1) H *nontrivially* separates K_1 *from* K_2 if $K_1 \not\subset H$,

2) H *strictly* separates K_1 *from* K_2 if $K_1 \cap H = \varnothing$,

3) H *strongly* separates K_1 *from* K_2 if there is an open ρ-neighborhood $U_\rho(K_1)$ of K_1 satisfying the condition $U_\rho(K_1) \cap H = \varnothing$.

Remark 1. The term *nontrivially separates from* is, probably, new. The expression *strictly separates from* is equivalent to that *openly separates from* used by Klee [4]; our choice here is due to the direct relation with the well established term *strict separation* (see Definition 2 below). Finally, the expression *strongly separates from* is introduced by Klee [4].

The following obvious lemma reformulates Definition 1 in analytic terms.

Lemma 1. *For convex sets K_1 and K_2 in $\mathbb{R}^n$ and a hyperplane H of the form* (1)*, the assertions below hold.*

1) *H separates K_1 and K_2 if and only if e and γ can be chosen such that*

$$\sup\{x_1 \cdot e : x_1 \in K_1\} \le \gamma \le \inf\{x_2 \cdot e : x_2 \in K_2\}. \tag{5}$$

2) *H nontrivially separates K_1 from K_2 if and only if e and γ can be chosen to satisfy either of the conditions* (6) *and* (7) *below:*

$$\sup\{x_1 \cdot e : x_1 \in K_1\} < \gamma \le \inf\{x_2 \cdot e : x_2 \in K_2\}, \tag{6}$$

$$\begin{aligned} &\inf\{x_1 \cdot e : x_1 \in K_1\} < \sup\{x_1 \cdot e : x_1 \in K_1\} \\ &\quad = \gamma = \inf\{x_2 \cdot e : x_2 \in K_2\}. \end{aligned} \tag{7}$$

3) *H strictly separates K_1 from K_2 if and only if e and γ can be chosen such that*

$$x_1 \cdot e < \gamma \le \inf\{x_2 \cdot e : x_2 \in K_2\} \quad \forall\, x_1 \in K_1. \tag{8}$$

4) *H strongly separates K_1 from K_2 if and only if e and γ can be chosen to satisfy the inequalities* (6). □

The next proposition describes known results on hyperplane separation related to Definition 1. Clearly, these results provide the existence of at least one such hyperplane; they do not describe all possible separating hyperplanes of a given type.

Proposition 1. Given convex sets K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.

1. If K_2 is a polyhedron, then there is a hyperplane nontrivially separating K_1 from K_2 if and only if $\operatorname{rint} K_1 \cap K_2 = \varnothing$ (Rockafellar [9], Theorem 20.2).

2. If K_1 and K_2 are disjoint and K_1 contains no halfline in its boundary, then K_1 is strictly separated from K_2 by a hyperplane (Klee [3]).

3. Let K_1 and K_2 be disjoint and closed, at least one of them being compact. Let $z_1 \in K_1$ and $z_2 \in K_2$ be points for which $\delta(K_1, K_2) = \|z_1 - z_2\|$. Then any hyperplane perpendicular to the closed segment $[z_1, z_2]$ and passing through any point of the semi-open segment $(z_1, z_2]$ strongly separates K_1 from K_2 (Minkowki [8], p. 141, for the case when both K_1 and K_2 are compact). □

We observe that the first assertion in Proposition 1 cannot be extended to the case of arbitrary convex sets. Namely, the following example shows that the condition $\operatorname{rint} K_1 \cap \operatorname{cl} K_2 = \varnothing$ is not sufficient for nontrivial separation of K_1 from K_2.

Example 1. Let K_1 and K_2 be planar circular disks in $\mathbb{R}^3$, given by

$$K_1 = \{(x, y, 0) : x^2 + (y-1)^2 \le 1\},$$
$$K_2 = \{(0, y, z) : y^2 + (z-1)^2 \le 1\}.$$

Both K_1 and K_2 are closed convex sets and $\operatorname{rint} K_1 \cap K_2 = \varnothing$. It is easy to see that the coordinate xy-plane (which contains K_1) is the only plane separating K_1 and K_2. Hence K_1 is not nontrivially separated from K_2.

The next two corollaries, which immediately follow from [12], describe all hyperplanes separating a pair of convex bodies or strongly separating a given convex set from another one.

Corollary 1. *For convex sets K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

1) *There is a hyperplane separating K_1 and K_2 if and only if any of the following two conditions is satisfied:*

   (*a*) $o \notin \operatorname{int}(K_1 - K_2)$,

   (*b*) $(K_1 - K_2)^\circ \ne \{o\}$.

2) *There is a translate of a hypersubspace* (2) *separating K_1 and K_2 if and only if one of the vectors e and $-e$ belongs to $(K_1 - K_2)^\circ \setminus \{o\}$.*

3) *For any vector $e \in (K_1 - K_2)^\circ \setminus \{o\}$, one has*

$$\sup\{x_1 \cdot e : x_1 \in K_1\} \le \inf\{x_2 \cdot e : x_2 \in K_2\}. \tag{9}$$

Consequently, any scalar γ satisfying the inequalities (5) *can be used in the description* (1) *of a hyperplane which separates K_1 and K_2.* □

Corollary 2. *For convex sets K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

1) *There is a hyperplane strongly separating K_1 from K_2 if and only if $o \notin \operatorname{cl}(K_1 - K_2)$.*

2) *If* $o \notin \operatorname{cl}(K_1 - K_2)$ *and* e *is a vector in* $\mathbb{R}^n$ *such that one of the vectors* e *and* $-e$ *belongs to* $\operatorname{rint}(K_1 - K_2)^\circ$, *then a suitable translate of the hypersubspace* (2) *strongly separates* K_1 *from* K_2.

3) *If there is a translate of a hypersubspace* (2) *strongly separating* K_1 *from* K_2, *then one of the vectors* e *and* $-e$ *belongs to* $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$. *If, additionally, both* K_1 *and* K_2 *are bounded sets, then one of these vectors belongs to* $\operatorname{rint}(K_1 - K_2)^\circ$.

4) *If* $o \notin \operatorname{cl}(K_1 - K_2)$, *then for any vector* $e \in \operatorname{rint}(K_1 - K_2)^\circ$, *the inequality*

$$\sup\{x_1 \cdot e : x_1 \in K_1\} < \inf\{x_2 \cdot e : x_2 \in K_2\} \tag{10}$$

holds. Consequently, any scalar γ *satisfying the inequalities* (6) *can be used in the description* (1) *of a hyperplane which strongly separates* K_1 *from* K_2. □

The next theorem describes hyperplanes nontrivially separating a given convex set from another one.

Theorem 1. *For convex sets* K_1 *and* K_2 *in* $\mathbb{R}^n$, *the assertions below hold.*

1) *There is a hyperplane nontrivially separating* K_1 *from* K_2 *if and only if any of the following two conditions is satisfied:*

   (*a*) $o \notin \operatorname{cl}(K_1 - K_2)$,

   (*b*) $o \in \operatorname{cl}(K_1 - K_2)$ *and* $(K_1 - K_2)^\circ \setminus \operatorname{ort} K_1 \neq \varnothing$.

2) *There is a translate of a hypersubspace* (2) *nontrivially separating* K_1 *from* K_2 *if and only if any of the following two conditions is satisfied:*

   (*c*) $o \notin \operatorname{cl}(K_1 - K_2)$ *and one of the vectors* e *and* $-e$ *belongs to the set*

$$\operatorname{rint}(K_1 - K_2)^\circ \cup (\operatorname{rbd}(K_1 - K_2)^\circ \setminus \operatorname{ort} K_1), \tag{11}$$

   (*d*) $o \in \operatorname{cl}(K_1 - K_2)$ *and one of the vectors* e *and* $-e$ *belongs to the set*

$$(K_1 - K_2)^\circ \setminus \operatorname{ort} K_1.$$

3) *If* $o \notin \operatorname{cl}(K_1 - K_2)$ *and* e *belongs to the set* (11), *then one of the relations* (10) *and*

$$\inf\{x_1 \cdot e : x_1 \in K_1\} < \sup\{x_1 \cdot e : x_1 \in K_1\} = \inf\{x_2 \cdot e : x_2 \in K_2\} \tag{12}$$

holds and any scalar γ *satisfying the respective conditions* (6) *and* (7) *can be used in the description* (1) *of a hyperplane which nontrivially separates* K_1 *from* K_2.

4) *If* $o \in \operatorname{cl}(K_1 - K_2)$ *and* $e \in (K_1 - K_2)^\circ \setminus \operatorname{ort} K_1$, *then the conditions* (12) *hold and any scalar* γ *satisfying* (7) *can be used in the description* (1) *of a hyperplane which nontrivially separates* K_1 *from* K_2.

Proof. 1) Let a hyperplane $H \subset \mathbb{R}^n$ nontrivially separate K_1 from K_2. We will assume that H is described by (1) and that K_1 and K_2 are contained, respectively, in the closed halfspaces V_1 and V_2 given by (3). By Lemma 1, one of the relations (6) and (7) holds. The obvious equality

$$\sup\{x\cdot e : x \in K_1 - K_2\} = \sup\{x_1\cdot e : x_1 \in K_1\} - \inf\{x_2\cdot e : x_2 \in K_2\}, \tag{13}$$

combined with (6) and (7), gives

$$\sup\{x\cdot e : x \in K_1 - K_2\} \le 0. \tag{14}$$

Hence the set $K_1 - K_2$ is contained in the homogeneous closed halfspace

$$V = \{x \in \mathbb{R}^n : x\cdot e \le 0\}.$$

Consequently, $e \in V^\circ \subset (K_1 - K_2)^\circ$.

Suppose that $o \in \mathrm{cl}\,(K_1 - K_2)$. Then (14) implies

$$\sup\{x\cdot e : x \in K_1 - K_2\} = o\cdot e = 0,$$

and (13) gives

$$\sup\{x_1\cdot e : x_1 \in K_1\} = \inf\{x_2\cdot e : x_2 \in K_2\}.$$

The latter equality shows that γ should satisfy the conditions (7).

Under the assumption $o \in \mathrm{cl}\,(K_1 - K_2)$, suppose that $e \in \mathrm{ort}\,K_1$. Then

$$\dim K_1 = (\mathrm{ort}\,K_1)^\perp \subset \{e\}^\perp := S = \{x \in \mathbb{R}^n : x\cdot e = 0\}.$$

As a translate of $\dim K_1$, the plane $\mathrm{aff}\,K_1$ is expressible in the form $\mathrm{aff}\,K_1 = z + \dim K_1$ for a suitable vector $z \in \mathbb{R}^n$. Consequently,

$$\begin{aligned} &K_1 \subset \mathrm{aff}\,K_1 = z + \dim K_1 \subset z + S = z + \{x \in \mathbb{R}^n : x\cdot e = 0\} \\ &= \{x \in \mathbb{R}^n : x\cdot e = \mu\},\ \text{where } \mu = z\cdot e. \end{aligned} \tag{15}$$

Comparing (1) and (15), we conclude that $\gamma = \mu$ and $H = z + S$. Hence $K_1 \subset H$, contrary to the hypothesis that H nontrivially separates K_1 from K_2. Thus $e \notin \mathrm{ort}\,K_1$. The latter exclusion shows that $(K_1 - K_2)^\circ \setminus \mathrm{ort}\,K_1 \ne \varnothing$.

Conversely, assume that any of the conditions (a) and (b) is satisfied. If (a) is satisfied, then, by Corollary 2, there is a hyperplane $H \subset \mathbb{R}^n$ strongly separating K_1 from K_2, and thus nontrivially separating K_1 from K_2. Suppose now that the condition (b) is satisfied. Choose a vector $e \in (K_1 - K_2)^\circ \setminus \mathrm{ort}\,K_1$. Repeating the above argument in the converse order, we conclude that the inclusion $e \in (K_1 - K_2)^\circ$ implies the existence of a hyperplane H of the form (1) separating K_1 and K_2, while the exclusion $e \notin \mathrm{ort}\,K_1$ guarantees that H does not contain K_1. Summing up, K_1 is nontrivially separated from K_2.

2) Let a translate of a hypersubspace (2) nontrivially separate K_1 from K_2. By Corollary 1, one of the vectors e and $-e$, say e, belongs to $(K_1 - K_2)^\circ$. If

$o \in \operatorname{cl}(K_1 - K_2)$, then, repeating the argument from part 1) above, we obtain the inclusion $e \in (K_1 - K_2)^\circ \setminus \operatorname{ort} K_1$. Suppose that $o \notin \operatorname{cl}(K_1 - K_2)$. If $e \in \operatorname{rint}(K_1 - K_2)^\circ$, then, by Corollary 2, the inequality (10) holds, and for any scalar γ satisfying the condition (6), the hyperplane (1) strongly separates K_1 from K_2. Finally, if $e \notin \operatorname{rint}(K_1 - K_2)^\circ$, then $e \in \operatorname{rbd}(K_1 - K_2)^\circ$, and, as above, e should not be in $\operatorname{ort} K_1$. Summing up, $e \in \operatorname{rbd}(K_1 - K_2)^\circ \setminus \operatorname{ort} K_1$.

Conversely, repeating the above argument in the converse order, we obtain that each of the conditions (c) and (d) implies the existence of a translate of a hypersubspace (2) nontrivially separating K_1 from K_2.

Assertions 3) and 4) follow from Lemma 1 and the above parts 1) and 2). □

Problem 1. Describe, in the spirit of Corollaries 1 and 2 and Theorem 1, all hyperplanes which *strictly separate* a given convex set from another one.

4 Weak Asymmetric Separation

In this section, we consider *weak types* of asymmetric separation of convex sets. Namely, we discuss the conditions under which (at least) one of the convex sets K_1 and K_2 in $\mathbb{R}^n$ is separated from the other.

Remark 2. The following terminology on weak types of asymmetric separation of convex sets K_1 and K_2 by a hyperplane $H \subset \mathbb{R}^n$ is known in the literature:

1. *H properly separates* K_1 and K_2 if one of the sets is nontrivially separated by H from the other (Rockafellar [9, p. 95]),

2. *H nicely separates* K_1 and K_2 if one of the sets is strictly separated by H from the other (Klee [5]).

The corollary below immediately follows from [12].

Corollary 3. *For convex sets K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

1) *There is a hyperplane nontrivially separating one of the sets K_1 and K_2 from the other if and only if any of the following three conditions is satisfied:*

   (a) $\operatorname{rint} K_1 \cap \operatorname{rint} K_2 = \varnothing$ (see [9, Theorem 11.3]).

   (b) $o \notin \operatorname{rint}(K_1 - K_2)$,

   (c) $(K_1 - K_2)^\circ$ *is not a subspace,*

   (d) $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ \neq \varnothing$.

2) *There is a translate of a hypersubspace* (2) *nontrivially separating one of the sets K_1 and K_2 from the other if and only if one of the vectors e and $-e$ belongs to $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$.*

3) *For any vector* $e \in (K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$, *both inequalities* (9) *and*

$$\inf\{x_1 \cdot e : x_1 \in K_1\} < \sup\{x_1 \cdot e : x_1 \in K_1\}$$

hold. Consequently, any scalar γ *satisfying the inequalities* (5) *can be used in the description* (1) *of a hyperplane which nontrivially separates one of the sets* K_1 *and* K_2 *from the other.*

For the case of strict separation, we consider, following Brøndsted [2], the "polarity" operation K^Δ on a convex set $K \subset \mathbb{R}^n$ defined by

$$K^\Delta = \{e \in \mathbb{R}^n : x \cdot e < 0 \ \ \forall\, x \in K \setminus \{o\}\}.$$

We observe that, generally, $K^\Delta \neq \operatorname{rint} K^\circ$.

Theorem 2. *For disjoint convex sets* K_1 *and* K_2 *in* $\mathbb{R}^n$, *the assertions below hold.*

1) *There is a hyperplane strictly separating one of the sets* K_1 *and* K_2 *from the other if and only if* $(K_1 - K_2)^\Delta \not\subset \{o\}$.

2) *There is a translate of a hypersubspace* (2) *strictly separating one of the sets* K_1 *and* K_2 *from the other if and only if one of the vectors* e *and* $-e$ *belongs to* $(K_1 - K_2)^\Delta \setminus \{o\}$.

3) *For any vector* $e \in (K_1 - K_2)^\Delta \setminus \{o\}$, *one of the conditions below is satisfied:*

$$x_1 \cdot e < \inf\{x_2 \cdot e : x_2 \in K_2\} \quad \forall\, x_1 \in K_1, \tag{16}$$

$$\sup\{x_1 \cdot e : x_1 \in K_1\} < x_2 \cdot e \quad \forall\, x_2 \in K_2. \tag{17}$$

Consequently, any scalar γ *satisfying the respective conditions*

$$x_1 \cdot e < \gamma \leq \inf\{x_2 \cdot e : x_2 \in K_2\} \quad \forall\, x_1 \in K_1,$$
$$\sup\{x_1 \cdot e : x_1 \in K_1\} \leq \gamma < x_2 \cdot e \quad \forall\, x_2 \in K_2$$

can be used in the description (1) *of a hyperplane strictly separating one of the sets* K_1 *and* K_2 *from the other.*

Proof. 1) Let a hyperplane $H \subset \mathbb{R}^n$ strictly separate one of the sets K_1 and K_2 from the other. We will assume that H is described by (1) and that K_1 and K_2 are contained, respectively, in the complementary halfspaces

$$W_1 = \{x \in \mathbb{R}^n : x \cdot e < \gamma\} \quad \text{and} \quad V_2 = \{x \in \mathbb{R}^n : x \cdot e \geq \gamma\}.$$

(The case when K_1 is contained is the closed halfspace V_1 and K_2 is in the complementary open halfspace W_2 is similar.) For any points $x_1 \in K_1$ and $x_2 \in K_2$, one has

$$(x_1 - x_2) \cdot e = x_1 \cdot e - x_2 \cdot e < \gamma - \gamma = 0.$$

Consequently, $K_1 - K_2 = (K_1 - K_2) \setminus \{o\}$ is contained in the homogeneous open halfspace

$$W = \{x \in \mathbb{R}^n : x \cdot e < 0\}, \tag{18}$$

which gives the inclusion $e \in (K_1 - K_2)^{\Delta}$. Hence $(K_1 - K_2)^{\Delta} \not\subset \{o\}$.

Conversely, assume that $(K_1 - K_2)^{\Delta} \not\subset \{o\}$ and choose a vector $e \in (K_1 - K_2)^{\Delta} \setminus \{o\}$. Then $K_1 - K_2$ is contained in the open halfspace (18). So, $(x_1 - x_2) \cdot e < 0$ whenever $x_1 \in K_1$ and $x_2 \in K_2$. Equivalently, $x_1 \cdot e < x_2 \cdot e$ for all $x_1 \in K_1$ and $x_2 \in K_2$. Let

$$\gamma_1 = \sup\{x_1 \cdot e : x_1 = K_1\} \quad \text{and} \quad \gamma_2 = \inf\{x_2 \cdot e : x_2 = K_2\}.$$

Then $\gamma_1 \leq \gamma_2$ due to the inclusion $K_1 - K_2 \subset W$ and the inequality

$$\begin{aligned} \gamma_1 - \gamma_2 &= \sup\{x_1 \cdot e : x_1 \in K_1\} - \inf\{x_2 \cdot e : x_2 \in K_2\} \\ &= \sup\{x \cdot e : x \in K_1 - K_2\} \leq 0. \end{aligned}$$

If $\gamma_1 < \gamma_2$, then, by Corollary 2, the hyperplane $H' = \{x \in \mathbb{R}^n : x \cdot e = \gamma'\}$ strongly separates K_1 from K_2 for any choice of $\gamma' \in (\gamma_1, \gamma_2]$. Suppose that $\gamma_1 = \gamma_2$ and put $\gamma' = \gamma_1 = \gamma_2$. If there is a point $x_2 \in K_2$ such that $x_2 \cdot e = \gamma'$, then $x_1 \cdot e < \gamma'$ for all $x_1 \in K_1$, implying that H' strictly separates K_1 from K_2. Similarly, if there is a point $x_1 \in K_1$ such that $x_1 \cdot e = \gamma'$, then $\gamma' < x_2 \cdot e$ for all $x_2 \in K_2$, implying that H' strictly separates K_2 from K_1.

2) Let a translate, say H, of a hypersubspace (2) strictly separate one of the sets K_1 and K_2 from the other. Then H is described by (1). By the argument of part 1), one of the vectors e and $-e$ should belong to $(K_1 - K_2)^{\Delta} \setminus \{o\}$. In a similar way, the converse assertion holds.

Assertion 3) follows from Lemma 1 and the above parts 1) and 2). □

Remark 3. A description of hyperplanes strongly separating one of the convex sets K_1 and K_2 from the other repeats Corollary 2 with one variation: in part 4), the scalar γ should be chosen to satisfy (6) or the symmetric conditions

$$\sup\{x_1 \cdot e : x_1 \in K_1\} \leq \gamma < \inf\{x_2 \cdot e : x_2 \in K_2\}.$$

5 Symmetric Separation

In this section, we consider *symmetric separation* of convex sets. Namely, we describe the conditions under which each of the convex sets K_1 and K_2 in $\mathbb{R}^n$ is separated from the other.

Definition 2. Let convex sets K_1 and K_2 in $\mathbb{R}^n$ be separated by a hyperplane $H \subset \mathbb{R}^n$. We will say that

1) H *nontrivially* separates K_1 *and* K_2 if $K_1 \not\subset H$ and $K_2 \not\subset H$,

2) H *strictly* separates K_1 *and* K_2 if $K_1 \cap H = K_2 \cap H = \varnothing$,

3) *H strongly* separates K_1 *and* K_2 if there is a scalar $\rho > 0$ such that

$$U_\rho(K_1) \cap H = U_\rho(K_2) \cap H = \varnothing.$$

Remark 4. Nontrivial separation is called *real* separation by Bair and Jongmans [1] and *definite* separation in [11, Definition 10.1]. The terms *strict* and *strong* separation are used in the survey of Klee [5] and in numerous publications afterwards.

The next proposition, proved in [11, Theorem 10.6], relates Definitions 1 and 2.

Proposition 2. Let K_1 and K_2 be convex sets and H_1 and H_2 be hyperplanes in $\mathbb{R}^n$ such that H_i separates (nontrivially, strictly, or strongly) K_i from K_{3-i}, $i = 1, 2$. Then there is a hyperplane containing the set $H_1 \cap H_2$ and separating (nontrivially, strictly, or strongly) K_1 and K_2. □

The corollary below immediately follows from Theorem 1 and Proposition 2.

Corollary 4. *For convex sets K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

1) *There is a hyperplane nontrivially separating K_1 and K_2 if and only if one of the following two conditions is satisfied:*

   (*a*) $o \notin \operatorname{cl}(K_1 - K_2)$,

   (*b*) $o \in \operatorname{cl}(K_1 - K_2)$ *and* $(K_1 - K_2)^\circ \setminus (\operatorname{ort} K_1 \cup \operatorname{ort} K_2) \neq \varnothing$.

2) *There is a translate of a hypersubspace* (2) *nontrivially separating K_1 and K_2 if and only if one of the following conditions is satisfied:*

   (*c*) $o \notin \operatorname{cl}(K_1 - K_2)$ *and one of the vectors e and $-e$ belongs to the set*

$$\operatorname{rint}(K_1 - K_2)^\circ \cup (\operatorname{rbd}(K_1 - K_2)^\circ \setminus (\operatorname{ort} K_1 \cup \operatorname{ort} K_2)), \tag{19}$$

   (*d*) $o \in \operatorname{cl}(K_1 - K_2)$ *and one of the vectors e and $-e$ belongs to the set*

$$(K_1 - K_2)^\circ \setminus (\operatorname{ort} K_1 \cup \operatorname{ort} K_2).$$

3) *If $o \notin \operatorname{cl}(K_1 - K_2)$ and e belongs to the set* (19), *then one of the inequalities* (10), (12), *and*

$$\sup\{x_1 \cdot e : x_1 \in K_1\} = \inf\{x_2 \cdot e : x_2 \in K_2\} < \sup\{x_2 \cdot e : x_2 \in K_2\}$$

   holds and the respective value of γ satisfying (5) *can be used in the description* (1) *of a hyperplane which nontrivially separates K_1 and K_2.*

4) *If $o \in \operatorname{cl}(K_1 - K_2)$ and $e \in (K_1 - K_2)^\circ \setminus (\operatorname{ort} K_1 \cup \operatorname{ort} K_2)$, then*

$$\begin{aligned} &\inf\{x_1 \cdot e : x_1 \in K_1\} < \sup\{x_1 \cdot e : x_1 \in K_1\} \\ &= \inf\{x_2 \cdot e : x_2 \in K_2\} < \sup\{x_2 \cdot e : x_2 \in K_2\} \end{aligned}$$

   and any scalar γ satisfying (5) *can be used in the description* (1) *of a hyperplane which nontrivially separates K_1 and K_2.* □

Problem 2. Describe all hyperplanes which *strictly separate* a given pair of convex sets K_1 and K_2 in $\mathbb{R}^n$.

Remark 5. A description of hyperplanes strongly separating one of the convex sets K_1 and K_2 from the other repeats Corollary 2 with one variation: in part 4), the scalar γ in the description of separating hyperplane (1) can be chosen to satisfy the conditions

$$\sup\{x_1 \cdot e : x_1 \in K_1\} < \gamma < \inf\{x_2 \cdot e : x_2 \in K_2\}.$$

6 Separation of Convex Cones

This section deals with various types of separation of convex cones which have a common apex. We recall that a convex set $C \subset \mathbb{R}^n$ is called a *cone* with apex $a \in \mathbb{R}^n$ provided $a + \lambda(x - a) \in C$ whenever $x \in C$ and $\lambda \geq 0$. This definition implies (letting $\lambda = 0$) that C contains its apex a, although a stronger condition $\lambda > 0$ can be beneficial; see, e. g., [6]. The set $\operatorname{ap} C = C \cap (2a - C)$ is called the *apex set* of C. It is known that $\operatorname{ap} C$ is the largest plane through a contained in C (see [6] and [11, Theorem 5.17]). Obviously, $C \neq \operatorname{ap} C$ if and only if C is not a plane.

The next corollary follows from [12] and Corollary 4.

Corollary 5. *Let C_1 and C_2 be convex cones with a common apex a, and let $D_1 = C_1 - a$ and $D_2 = C_2 - a$. The assertions below hold.*

1) *There is a hyperplane nontrivially separating one of the cones C_1 and C_2 from the other if and only if $(C_1 - C_2)^\circ$ is not a subspace.*

2) *A hyperplane $H \subset \mathbb{R}^n$ of the form*

$$H = \{x \in \mathbb{R}^n : x \cdot e = a \cdot e\}, \quad e \neq o, \tag{20}$$

nontrivially separates one of the cones C_1 and C_2 from the other if and only if one of the vectors e and $-e$ belongs to $(C_1 - C_2)^\circ \setminus \operatorname{lin}(C_1 - C_2)$.

3) *There is a hyperplane nontrivially separating C_1 from C_2 if and only if*

$$(C_1 - C_2)^\circ \setminus \operatorname{ort} C_1 \neq \varnothing.$$

4) *A hyperplane $H \subset \mathbb{R}^n$ of the form* (20) *nontrivially separates C_1 from C_2 if and only if one of the vectors e and $-e$ belongs to $(C_1 - C_2)^\circ \setminus \operatorname{ort} C_1$.*

5) *There is a hyperplane nontrivially separating C_1 and C_2 if and only if*

$$(C_1 - C_2)^\circ \setminus (\operatorname{ort} C_1 \cup \operatorname{ort} C_2) \neq \varnothing. \tag{21}$$

6) *A hyperplane $H \subset \mathbb{R}^n$ of the form* (20) *nontrivially separates C_1 and C_2 if and only if one of the vectors e and $-e$ belongs to the set* (21). □

If closed convex cones C_1 and C_2 with a common apex are separated by a hyperplane $H \subset \mathbb{R}^n$, then H supports both C_1 and C_2. Consequently, $\operatorname{ap} C_1 \cup \operatorname{ap} C_2 \subset H$ (see, e. g., [11], Theorem 9.43). In this regard, we recall the definition from [10]: a hyperplane H *sharply separates* C_1 and C_2 provided H separates them and

$$C_1 \cap H = \operatorname{ap} C_1 \quad \text{and} \quad C_2 \cap H = \operatorname{ap} C_2. \tag{22}$$

An asymmetric version of this definition is formulated as follows.

Definition 3. Let C_1 and C_2 be closed convex cones in $\mathbb{R}^n$, with a common apex, and let $H \subset \mathbb{R}^n$ be a hyperplane separating C_1 and C_2. We will say that H *sharply* separates C_1 *from* C_2 if $H \cap C_1 = \operatorname{ap} C_1$.

The next theorem gives a criterion for sharp separation of a convex cone from another one in terms of their polar cones.

Theorem 3. *If C_1 and C_2 are closed convex cones in $\mathbb{R}^n$ with a common apex a, then the following conditions are equivalent.*

1) *C_1 is sharply separated from C_2.*

2) *The set $E = \operatorname{rint}(C_1 - a)^\circ \cap (a - C_2)^\circ$ has positive dimension.*

Proof. Put $F_1 = C_1 - a$ and $F_2 = C_2 - a$. Then both F_1 and F_2 are closed convex cones with common apex o. Furthermore, $\operatorname{ap} F_i = \operatorname{ap} C_i - a$, $i = 1, 2$, and the set E from the condition 2) can be described as

$$E = \operatorname{rint} F_1^\circ \cap (-F_2)^\circ = \operatorname{rint} F_1^\circ \cap (-F_2^\circ). \tag{23}$$

1) $\Rightarrow$ 2) Let C_1 be sharply separated from C_2 by a hyperplane of the form

$$H = \{x \in \mathbb{R}^n : x \cdot e = \gamma\}, \quad e \neq o.$$

Clearly, F_1 and F_2 are separated by the $(n-1)$-dimensional subspace

$$S = H - a = \{x \in \mathbb{R}^n : x \cdot e = o\}.$$

Furthermore, S sharply separates F_1 from F_2 due to

$$S \cap F_1 = (H - a) \cap (C_1 - a) = H \cap C_1 - a = \operatorname{ap} C_1 - a = \operatorname{ap} F_1,$$

Without loss of generality, we may assume that

$$F_1 \subset V_1 = \{x \in \mathbb{R}^n : x \cdot e \leq 0\} \quad \text{and} \quad F_2 \subset V_2 = \{x \in \mathbb{R}^n : x \cdot e \geq 0\}.$$

We assert that $e \in E$. To show the inclusion $e \in \operatorname{rint} F_1^\circ$, we will consider separately the cases when F_1 is or is not a subspace.

Assume first that F_1 is a subspace. Then $F_1 = \operatorname{ap} F_1 \subset S$, which gives the inclusion $e \in S^\perp \subset F_1^\perp$. Since $F_1^\perp$ is a subspace, we obtain $e \in F_1^\perp = F_1^\circ = \operatorname{rint} F_1^\circ$.

Suppose now that F_1 is not a subspace. Then $F_1 \neq \operatorname{ap} F_1$, and the condition $S \cap F_1 = \operatorname{ap} F_1$ implies the inclusion $F_1 \setminus \operatorname{ap} F_1 \subset \operatorname{int} V_1$. In this case, Theorem 8.6 from [11] shows that $e \in \operatorname{rint} F_1^\circ$.

For the inclusion $e \in (-F_2^\circ)$, we observe first that V_2 can be expressed as

$$V_2 = \{x \in \mathbb{R}^n : x \cdot (-e) \leq 0\}.$$

Consequently, the inclusion $F_2 \subset V_2$ gives $-e \in V_2^\circ \subset F_2^\circ$, or $e \in -F_2^\circ$.

Summing up, $e \in \operatorname{rint} F_1^\circ \cap (-F_2^\circ) = E$, implying that $\dim E > 0$.

2) $\Rightarrow$ 1) Suppose that $\dim E > 0$, and choose a nonzero vector $e \in E$. By the above argument, $F_1 \subset V_1$ and $F_2 \subset V_2$ such that $F_1 \setminus \operatorname{ap} F_1 \subset \operatorname{int} V_1$ if F_1 is not a plane, and $F_1 \setminus \operatorname{ap} F_1 = \varnothing$ if F_1 is a plane. Hence $S \cap F_1 = \operatorname{ap} F_1$, implying that S sharply separates F_1 from F_2. Consequently, H sharply separates C_1 from C_2. □

Analysis of the proof of Theorem 3 reveals the following corollary.

Corollary 6. *Let C_1 and C_2 be closed convex cones in $\mathbb{R}^n$, with a common apex. If C_1 is not a plane and is sharply separated from C_2, then C_1 is nontrivially separated from C_2.* □

Remark 6. The converse to Corollary 6 assertion is not true. For instance, in $\mathbb{R}^2$, the cone $C_1 = \{(x, 0) : x \in \mathbb{R}\}$ is separated sharply but not properly from the cone $C_2 = \{(x, y) : 0 \leq x, 0 \leq y \leq x\}$, while C_2 is separated properly but not sharply from C_1.

Theorem 4. *Let C_1 and C_2 be closed convex cones in $\mathbb{R}^n$, with common apex a. The following conditions are equivalent.*

1) *C_1 and C_2 are sharply separated.*

2) *Each of the cones C_1 and C_2 is sharply separated from the other.*

3) *The set $D = \operatorname{rint}(C_1 - a)^\circ \cap \operatorname{rint}(a - C_2)^\circ$ has positive dimension.*

Proof. The equivalence of conditions 1) and 3) is proved in [11, Theorem 10.16] (initially given in [10] for the case when neither C_1 nor C_2 is a plane). Since 1) obviously implies 2), it suffices to show that 2) $\Rightarrow$ 3).

So, assume that each of the cones C_1 and C_2 is sharply separated from the other. By Theorem 3, there are nonzero vectors

$$e_1 \in \operatorname{rint}(C_1 - a)^\circ \cap (a - C_2)^\circ \quad \text{and} \quad e_2 \in \operatorname{rint}(C_2 - a)^\circ \cap (a - C_1)^\circ.$$

Obviously, the second inclusion can be rewritten as

$$-e_2 \in (C_1 - a)^\circ \cap \operatorname{rint}(a - C_2)^\circ.$$

If $e_1 = -e_2$, then $e_1 \in \operatorname{rint}(C_1 - a)^\circ \cap \operatorname{rint}(a - C_2)^\circ = D$. Because D is a convex cone with improper apex o, one has $(o, e_1] \subset D$, which implies the inequality $\dim D > 0$.

Let $e_1 \neq -e_2$. Then the open segment $I = (e_1, -e_2)$ is one-dimensional. Because $(C_1 - a)^\circ$ is a closed convex cone, the inclusions $e_1 \in \operatorname{rint}(C_1 - a)^\circ$ and $-e_2 \in (C_1 - a)^\circ$ imply that $I \subset \operatorname{rint}(C_1 - a)^\circ$ (see [9, Theorem 6.1]). By a similar argument, $I \subset \operatorname{rint}(a - C_2)^\circ$. So,

$$I \subset \operatorname{rint}(C_1 - a)^\circ \cap \operatorname{rint}(a - C_2)^\circ = D,$$

again resulting in the inequality $\dim D > 0$. □

References

[1] J. Bair, F. Jongmans, *La séparation vraie dans un espace vectoriel*. Bull. Soc. Roy. Sci. Liège 41 (1972), 163–170.

[2] A. Brøndsted, *The inner aperture of a convex set*. Pacific J. Math. 72 (1977), 335–340.

[3] V. L. Klee, *Strict separation of convex sets*. Proc. Amer. Math. Soc. 7 (1956), 735–737.

[4] V. L. Klee, *Maximal separation theorems for convex sets*, Trans. Amer. Math. Soc. 134 (1968), 133–147.

[5] V. L. Klee, *Separation and support properties of convex sets – a survey*, Control Theory and the Calculus of Variations (University of California, LA, 1968), pp. 235–303, Academic Press, New York, 1969.

[6] J. Lawrence, V. Soltan, *On unions and intersections of nested families of cones*, Beitr. Algebra Geom. 57 (2016), 655–665.

[7] H. Minkowski, *Geometrie der Zahlen*. I, Teubner, Leipzig, 1896; II. Teubner, Leipzig, 1910.

[8] H. Minkowski, *Gesammelte Abhandlungen*. Bd 2, Teubner, Leipzig, 1911.

[9] R. T. Rockafellar, *Convex Analysis*. Princeton Universty Press, Princeton, NJ, 1970.

[10] V. Soltan, *Polarity and separation of cones*, Linear Algebra Appl. 538 (2018), 212–224.

[11] V. Soltan, *Lectures on convex sets. Second edition*, World Scientific, Hackensack, NJ, 2020.

[12] V. Soltan, *Separating hyperplanes of convex sets*, J. Convex Anal. 28 (2021), accepted.

Valeriu Soltan
George Mason University
Fairfax, Virginia 22030, USA
E-mail: *vsoltan@gmu.edu*

Received September 1, 2020