ARS MATHEMATICA CONTEMPORANEA

Volume 19, Number 1, Fall/Winter 2020, Pages 1–171

Covered by:
Mathematical Reviews
zbMATH (formerly Zentralblatt MATH)
COBISS
SCOPUS
Science Citation Index-Expanded (SCIE)
Web of Science
ISI Alerting Service
Current Contents/Physical, Chemical & Earth Sciences (CC/PC & ES)
dblp computer science bibliography

The University of Primorska

The Society of Mathematicians, Physicists and Astronomers of Slovenia
The Institute of Mathematics, Physics and Mechanics
The Slovenian Discrete and Applied Mathematics Society

The publication is partially supported by the Slovenian Research Agency from the Call for co-financing of scientific periodical publications.

DOI – Digital Object Identifier

It is clear that electronic publishing is here to stay, and that hard copies of mathematical journals are becoming less and less important for the reader. But it is not immediately obvious how to locate and identify items in the vast array of scientific papers in the electronic world.

We have followed the lead of many other scientific journals, and equipped the papers appearing in *AMC* with a DOI (a 'Digital Object Identifier') that helps uniquely identify the paper. DOIs were introduced in 2000 by the International DOI Foundation (IDF). From volume 15 (2018) onwards, we have assigned a DOI to the hard copy of each paper in *AMC*. In the same year, we also began the process of assigning a DOI to every electronic paper, from volume 1 onwards. This information is now displayed on the 'metapage' of each individual paper.

The DOI is a string of characters divided into two parts: a prefix and a suffix, separated by a slash. For example, the DOI of the first paper of this issue is 10.26493/1855-3974.2163.5df. The corresponding document can be accessed by visiting

https://doi.org/10.26493/1855-3974.2163.5df,

which forwards the visitor to the corresponding 'metapage' on the *AMC* website. The first part, namely 10.26493, is permanent and identifies the publisher, which in our case is the University of Primorska. The second part is variable, and depends on the specific paper. In our example above, 1855-3974 represents the ISSN of our journal. Also a journal issue has its own DOI; for example, the current issue (volume 19, issue 1) is identified by 10.26493/1855-3974.19_1.

Using DOIs proved to be especially helpful when we introduced the class of Accepted Manuscripts. Each article appearing in Accepted Manuscripts is assigned a DOI, and because a DOI is permanent identification, every article may be accurately identified when cited even before it is assigned to a particular issue, and when that happens, the DOI remains the same. Even later, if the platform of the journal moves to another website, the same DOI can be used to link to the new URL. (In our case, *AMC* began at http://amc.imfm.si/, but is now hosted at https://amc-journal.eu/.)

For more information about the DOI system, see the DOI Handbook, doi:10.1000/182.

Nino Bašić
Layout Editor

Klavdija Kutnar, Dragan Marušič and Tomaž Pisanski
Editors in Chief

ARS MATHEMATICA CONTEMPORANEA

Contents

Volume 19, Number 1, Fall/Winter 2020, Pages 1–171

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 1–15
https://doi.org/10.26493/1855-3974.2163.5df
(Also available at http://amc-journal.eu)

Hamilton cycles in primitive vertex-transitive graphs of order a product of two primes – the case $\mathrm{PSL}(2, q^2)$ acting on cosets of $\mathrm{PGL}(2, q)$*

Shaofei Du †

Capital Normal University, School of Mathematical Sciences, Bejing 100048, People's Republic of China

Klavdija Kutnar ‡

University of Primorska, UP FAMNIT, Glagoljaška 8, 6000 Koper, Slovenia, and *University of Primorska, UP IAM, Muzejski trg 2, 6000 Koper, Slovenia*

Dragan Marušič §

University of Primorska, UP FAMNIT, Glagoljaška 8, 6000 Koper, Slovenia, and *University of Primorska, UP IAM, Muzejski trg 2, 6000 Koper, Slovenia,* and *IMFM, Jadranska 19, 1000 Ljubljana, Slovenia*

Received 27 October 2019, accepted 4 May 2020, published online 26 October 2020

Abstract

A step forward is made in a long standing Lovász problem regarding hamiltonicity of vertex-transitive graphs by showing that every connected vertex-transitive graph of order a product of two primes arising from the group action of the projective special linear group $\mathrm{PSL}(2, q^2)$ on cosets of its subgroup isomorphic to the projective general linear group $\mathrm{PGL}(2, q)$ contains a Hamilton cycle.

Keywords: Vertex-transitive graph, Hamilton cycle, automorphism group, orbital graph.

Math. Subj. Class. (2020): 05C25, 05C45

*The authors wish to thank Kai Yuan for helpful conversations about the material of this paper and to the referees for their helpful comments and suggestions.

†This work is supported in part by the National Natural Science Foundation of China (11671276, 11971248).

‡Corresponding author. This work is supported in part by the Slovenian Research Agency (research program P1-0285 and research projects N1-0038, N1-0062, J1-6720, J1-6743, J1-7051, J1-9110, and J1-1695).

§This work is supported in part by the Slovenian Research Agency (I0-0035, research program P1-0285 and research projects N1-0038, N1-0062, J1-6720, J1-9108, and J1-1695), and in part by H2020 Teaming InnoRenew CoE (grant no. 739574).

E-mail addresses: dushf@mail.cnu.edu.cn (Shaofei Du), klavdija.kutnar@upr.si (Klavdija Kutnar), dragan.marusic@upr.si (Dragan Marušič)

1 Introduction

In 1969, Lovász [20] asked if there exists a finite, connected vertex-transitive graph without a Hamilton path, that is, a simple path going through all vertices of the graph. To this date no such graph is known to exist. Intriguingly, with the exception of K_2, only four connected vertex-transitive graphs that do not have a Hamilton cycle are known to exist. These four graphs are the Petersen graph, the Coxeter graph and the two graphs obtained from them by replacing each vertex by a triangle. The fact that none of these four graphs is a Cayley graph has led to a folklore conjecture that every Cayley graph is hamiltonian (see [1, 8, 9, 11, 12, 16, 22, 32, 37] and the survey paper [6] for the current status of this conjecture).

Coming back to the general class of vertex-transitive graphs, the existence of Hamilton paths, and in some cases also Hamilton cycles, in connected vertex-transitive graphs has been shown for graphs of particular orders, such as, kp, $k \leq 6$, p^j, $j \leq 5$ and $2p^2$ (see [5, 15, 17, 18, 23, 24, 26, 27, 38] and the survey paper [16]). (Throughout this paper p will always denote a prime number.) Further, some partial results have been obtained for graphs of order pq, $q < p$ a prime [2, 25]. The main obstacle to obtaining a complete solution lies in graphs with a primitive automorphism group having no imprimitive subgroup. It is the object of this paper to move a step closer to resolving Lovász question for vertex-transitive graphs of order a product of two primes by showing existence of Hamilton cycles in graphs arising from the action of $\mathrm{PSL}(2, q^2)$ on cosets of its subgroup isomorphic to $\mathrm{PGL}(2, q)$ (see Theorem 3.2). The strategy used in the proof is introduced in Section 3. In the next section we fix the terminology and notation, and gather same useful results and tools.

2 Terminology, notation and some useful results

2.1 Basic definitions and notation

Throughout this paper graphs are finite, simple and undirected, and groups are finite. Furthermore, a *multigraph* is a generalization of a graph in which we allow multiedges and loops. Given a graph X we let $V(X)$ and $E(X)$ be the vertex set and the edge set of X, respectively. For adjacent vertices $u, v \in V(X)$ we write $u \sim v$ and denote the corresponding edge by uv. Let U and W be disjoint subsets of $V(X)$. The subgraph of X induced by U will be denoted by $X\langle U\rangle$. Similarly, we let $X[U, W]$ denote the bipartite subgraph of X induced by the edges having one endvertex in U and the other endvertex in W.

Given a transitive group G acting on a set V, we say that a partition $\mathcal{B}$ of V is *G-invariant* if the elements of G permute the parts, the socalled *blocks* of $\mathcal{B}$, setwise. If the trivial partitions $\{V\}$ and $\{\{v\} : v \in V\}$ are the only G-invariant partitions of V, then G is *primitive*, and is *imprimitive* otherwise.

A graph X is *vertex-transitive* if its automorphism group, denoted by $\mathrm{Aut}\, X$, acts transitively on $V(X)$. A vertex-transitive graph is said to be *primitive* if every transitive subgroup of its automorphism group is primitive, and is said to be *imprimitive* otherwise.

A graph containing a Hamilton cycle will be sometimes referred as a *hamiltonian* graph.

2.2 Generalized orbital graphs

In this subsection we recall the orbital graph construction which is used throughout the rest of the paper. A permutation group G on a set V induces the action of G on $V \times V$. The corresponding orbits are called *orbitals*. An orbital is said to be *self-paired* if it simultaneously contains or does not contain ordered pairs (x, y) and (y, x), for $x, y \in V$. For an arbitrary union $\mathcal{O}$ of orbitals (having empty intersection with the diagonal $D = \{(x, x) : x \in V\}$), the *generalized orbital (di)graph* $X(V, \mathcal{O})$ of the action of G on V with respect to $\mathcal{O}$ is a simple (di)graph with vertex set V and edge set $\mathcal{O}$. (For simplicity reasons we will refer to any such (di)graph as an orbital (di)graph of G.) It is an (undirected) graph if and only if $\mathcal{O}$ coincides with its symmetric closure, that is, $\mathcal{O}$ has the property that $(x, y) \in \mathcal{O}$ implies $(y, x) \in \mathcal{O}$. Further, the generalized orbital graph $X(V, \mathcal{O})$ is said to be a *basic orbital graph* if $\mathcal{O}$ is a single orbital or a union of a single orbital and its symmetric closure. Note that the orbital graph $X(V, \mathcal{O})$ is vertex-transitive if and only if G is transitive on V, that the diagonal D is always an orbital provided G acts transitively on V, and that its complement, $V \times V - D$ is an orbital if and only if G is doubly transitive.

Every vertex-transitive (di)graph admitting a transitive group of automorphisms G with the corresponding vertex stabilizer H can be constructed as an orbital (di)graph of the action of the group G on the coset space G/H. The orbitals of the action of G on G/H are in 1-1 correspondence with the orbits of the action of H on G/H, called *suborbits* of G. A suborbit corresponding to a self-paired orbital is said to be *self-paired*. When presenting the (generalized) orbital (di)graph $X(G/H, \mathcal{O})$ with the corresponding (union) of suborbits $\mathcal{S}$ the (di)graph $X(G/H, \mathcal{O})$ is denoted by $X(G, H, \mathcal{S})$.

2.3 Semiregular automorphisms and quotient (multi)graphs

Let $m \geq 1$ and $n \geq 2$ be integers. An automorphism ρ of a graph X is called (m, n)-*semiregular* (in short, *semiregular*) if as a permutation on $V(X)$ it has a cycle decomposition consisting of m cycles of length n. The question whether all vertex-transitive graphs admit a semiregular automorphism is one of famous open problems in algebraic graph theory (see, for example, [3, 4, 7, 10, 21]). Let $\mathcal{P}$ be the set of orbits of ρ, that is, the orbits of the cyclic subgroup $\langle\rho\rangle$ generated by ρ. Let $A, B \in \mathcal{P}$. By $d(A)$ and $d(A, B)$ we denote the valency of $X\langle A\rangle$ and $X[A, B]$, respectively. (Note that the graph $X[A, B]$ is regular.) We let the *quotient graph corresponding to* $\mathcal{P}$ be the graph $X_{\mathcal{P}}$ whose vertex set equals $\mathcal{P}$ with $A, B \in \mathcal{P}$ adjacent if there exist vertices $a \in A$ and $b \in B$, such that $a \sim b$ in X. We let the *quotient multigraph corresponding to* ρ be the multigraph X_ρ whose vertex set is $\mathcal{P}$ and in which $A, B \in \mathcal{P}$ are joined by $d(A, B)$ edges. Note that the quotient graph $X_{\mathcal{P}}$ is precisely the underlying graph of X_ρ.

2.4 Useful number theory facts

For a prime power r a finite field of order r will be denoted by F_r, with the subscript r being omitted whenever the order of the field is clear from the context. As usual, set $F^* = F \setminus \{0\}$. Set $S^* = \{a^2 : a \in F^*\}$ and $N^* = F^* \setminus S^*$. The elements of S^* and N^* will be called *squares* and *non-squares*, respectively. The following basic number-theoretic results will be needed.

Proposition 2.1 ([35, Theorem 21.2]). *Let F be a finite field of odd prime order p. Then $-1 \in S^*$ if $p \equiv 1 \pmod 4$, and $-1 \in N^*$ if $p \equiv 3 \pmod 4$.*

Proposition 2.2 ([35, Theorem 21.4]). *Let F be a finite field of odd prime order p. Then $2 \in S^*$ if $p \equiv 1, 7 \pmod 8$, and $2 \in N^*$ if $p \equiv 3, 5 \pmod 8$.*

Proposition 2.3 ([29, p. 167]). *Let F be a finite field of odd prime order p. Then*

$$|(S^* + 1) \cap (-S^*)| = \begin{cases} (p-5)/4, & \text{if } p \equiv 1 \pmod 4, \\ (p+1)/4, & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

In particular, if $p \equiv 1 \pmod 4$ then $|S^* \cap (S^* + 1)| = (p-5)/4$, $|N^* \cap (N^* + 1)| = (p-1)/4$, and $|S^* \cap (N^* + 1)| = |S^* \cap (N^* - 1)| = (p-1)/4$.

Using Proposition 2.3 the following result may be easily deduced.

Proposition 2.4. *Let F be a finite field of odd prime order p. Then for any $k \in F^*$, the equation $x^2 + y^2 = k$ has $p - 1$ solutions if $p \equiv 1 \pmod 4$, and $p + 1$ solutions if $p \equiv 3 \pmod 4$.*

3 Vertex-transitive graphs of order *pq*

Vertex-transitive graphs whose order is a product of two different odd primes p and q, where $p > q$ can be conveniently split into three mutually disjoint classes. The first class consists of graphs admitting an imprimitive subgroup of automorphisms with blocks of size p – it coincides with (q, p)-metacirculants [2]. The second class consists of graphs admitting an imprimitive subgroup of automorphisms with blocks of size q but no imprimitive subgroup of automorphisms with blocks of size p – it coincides with the class of socalled Fermat graphs, which are certain q-fold covers of K_p where p is a Fermat prime [28]. The third class consists of vertex-transitive graphs with no imprimitive subgroup of automorphisms. Following [31, Theorem 2.1] the theorem below gives a complete classification of connected vertex-transitive graphs of order pq (see also [33, 34]). We would like to remark, however, that there is an additional family of primitive graphs of order $91 = 7 \cdot 13$ that was not covered neither in [31] nor in [34]. This is due to a missing case in Liebeck-Saxl's table [19] of primitive group actions of degree mp, $m < p$. This missing case consists of primitive groups of degree $91 = 7 \cdot 13$ with socle $\mathrm{PSL}(2, 13)$ acting on cosets of A_4. In the classification theorem below this missing case is included in Row 7 of Table 1.

Theorem 3.1 ([31, Theorem 2.1]). *A connected vertex-transitive graph of order pq, where p and q are odd primes and $p > q$, must be one of the following:*

(i) a metacirculant,

(ii) a Fermat graph,

(iii) a generalized orbital graph associated with one of the groups in Table 1.

The existence of Hamilton cycles in graphs given in Theorem 3.1(i) and (ii) was proved, respectively, in [2] and [25]. It is the aim of this paper to make the next step towards proving the existence of Hamilton cycles in every connected vertex-transitive of order a product of two primes with the exception of the Petersen graph, by showing existence of Hamilton cycles in graphs arising from Row 5 of Table 1.

Theorem 3.2. *Vertex-transitive graphs arising from the action of $\mathrm{PSL}(2, q^2)$ on $\mathrm{PGL}(2, q)$ given in Row 5 of Table 1 are hamiltonian.*

Table 1: Primitive groups of degree pq without imprimitive subgroups and with non-isomorphic generalized orbital graphs.

Row	$\mathrm{soc}\, G$	(p,q)	Action	Comment
1	$\mathrm{P}\Omega^{\epsilon}(2d,2)$	$(2^d-\epsilon, 2^{d-1}+\epsilon)$	singular 1-spaces	$\epsilon=+1 : d$ Fermat prime $\epsilon=-1 : d-1$ Mersenne prime
2	M_{22}	$(11,7)$	see Atlas	
3	A_7	$(7,5)$	triples	
4	$\mathrm{PSL}(2,61)$	$(61,31)$	cosets of A_5	
5	$\mathrm{PSL}(2,q^2)$	$(\frac{q^2+1}{2}, q)$	cosets of $\mathrm{PGL}(2,q)$	$q \geq 5$
6	$\mathrm{PSL}(2,p)$	$(p, \frac{p+1}{2})$	cosets of D_{p-1}	$p \equiv 1 \pmod 4$ $p \geq 13$
7	$\mathrm{PSL}(2,13)$	$(13,7)$	cosets of A_4	missing in [19]

The existence of Hamilton cycles needs to be proved for all connected generalized orbital graphs arising from these actions. Recall that a generalized orbital graph is a union of basic orbital graphs. Since the considered action is primitive and hence the corresponding basic orbital graphs are connected, it suffices to prove the existence of Hamilton cycles solely in basic orbital graphs of this action. This is done in Section 4. The method used is for the most part based on the socalled lifting cycle technique [1, 16, 22]. Lifts of Hamilton cycles from quotient graphs which themselves have a Hamilton cycle are always possible, for example, when the quotienting is done relative to a semiregular automorphism of prime order and when the corresponding quotient multigraph has two adjacent orbits joined by a double edge contained in a Hamilton cycle. This double edge gives us the possibility to conveniently "change direction" so as to get a walk in the quotient that lifts to a full cycle above. By [21, Theorem 3.4] a vertex-transitive graph of order pq, $q < p$ primes, contains a (q,p)-semiregular automorphism. The lifting cycle technique, however, can only be applied provided appropriate Hamilton cycles can be found in the corresponding quotients. It so happens that graphs arising from Row 5 of Table 1 also admit (p,q)-semiregular automorphisms, and it is with respect to these automorphisms that the lifting cycle technique is applied. In constructing Hamilton cycles, the corresponding quotients have proved to be easier to work with than the quotients obtained from (q,p)-semiregular automorphisms. Namely, as one would expect, it is precisely the existence of Hamilton cycles in the quotients that represents the hardest obstacle one needs to overcome in order to assure the existence of Hamilton cycles in the graphs in question. In this respect the well-known Jackson theorem will be useful.

Proposition 3.3 (Jackson Theorem [13, Theorem 6])**.** *Every 2-connected regular graph of order n and valency at least $n/3$ contains a Hamilton cycle.*

It will be useful to introduce the following terminology. Let X be a graph that admits an (m,n)-semiregular automorphism ρ. Let $\mathcal{P} = \{S_1, S_2, \dots, S_m\}$ be the set of orbits of ρ, and let $\pi\colon X \to X_{\mathcal{P}}$ be the corresponding projection of X to its quotient $X_{\mathcal{P}}$. For a (possibly closed) path $W = S_{i_1}S_{i_2}\cdots S_{i_k}$ in $X_{\mathcal{P}}$ we let the *lift* of W be the set of all paths in X that project to W. The proof of following lemma is straightforward and is just a reformulation of [26, Lemma 5].

Lemma 3.4. *Let X be a graph admitting an (m,p)-semiregular automorphism ρ, where p is a prime. Let C be a cycle of length k in the quotient graph $X_{\mathcal{P}}$, where $\mathcal{P}$ is the set of orbits of ρ. Then, the lift of C either contains a cycle of length kp or it consists of p disjoint k-cycles. In the latter case we have $d(S,S')=1$ for every edge SS' of C.*

4 Actions of $\mathrm{PSL}(2,q^2)$

The following group-theoretic result due to Manning will be needed in the proof of Theorem 3.2.

Proposition 4.1 ([36, Theorem 3.6']). *Let G be a transitive group on Ω and let $H = G_\alpha$ for some $\alpha \in \Omega$. Suppose that $K \le G$ and at least one G-conjugate of K is contained in H. Suppose further that the set of G-conjugates of K which are contained in H form t conjugacy classes under H with representatives $K_1, K_2, \dots, K_t$. Then K fixes $\sum_{i=1}^{t} |N_G(K_i) : N_H(K_i)|$ points of Ω.*

Let $F_{q^2} = F_q(\alpha)$, where $\alpha^2 = \theta$ for $F_q^* = \langle\theta\rangle$. Let $G = \mathrm{PSL}(2,q^2)$, where $q \ge 5$ is an odd prime. For simplicity reasons we refer to the elements of G as matrices; this should cause no confusion. Then G has two conjugacy classes of subgroups isomorphic to $\mathrm{PGL}(2,q)$, with the corresponding representatives H and H'. Since each element in $\mathrm{PGL}(2,q^2) \setminus \mathrm{PSL}(2,q^2)$ interchanges these two classes, it suffices to consider the action of G on the set $\mathcal{H}$ of right cosets of H in G. The degree of this action is pq, where $p = (q^2+1)/2$. Without loss of generality let

$$H = \left\{ \frac{1}{\sqrt{|A|}} A : A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, a,b,c,d \in F_q \right\} \le G,$$

and

$$H' = H^g \quad \text{where} \quad g = \begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix},$$

where $\beta \in F_{q^2}^* \setminus (F_{q^2}^*)^2$. Let

$$Q = \left\langle \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} \right\rangle \quad \text{and} \quad Q_1 = \left\langle \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\rangle.$$

Then $Q \le H'$ and $Q \cap H = 1$. Moreover, we have the following result.

Lemma 4.2. *The action of Q on $\mathcal{H}$ is semiregular. Furthermore, the action of its normalizer $N_G(Q)$ on $\mathcal{H}$ has $\frac{q+1}{2}$ orbits of length q and one orbit of length $\frac{q^2(q-1)}{2}$.*

Proof. We first prove that the action of Q on $\mathcal{H}$ is semiregular. Suppose on the contrary that there exists $g \in G$ such that $HgQ = Hg$. Then $HgQg^{-1} = H$, and so $gQg^{-1} \le H$. But this contradicts the choice of Q. Hence Q is semiregular on $\mathcal{H}$.

One can see that

$$N = N_G(Q) = \left\{ \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} : a \in \langle\alpha\rangle, b \in F_{q^2} \right\} \cong \mathbb{Z}_q^2 \rtimes \mathbb{Z}_{q-1}.$$

We now compute the orbits of N in its action on $\mathcal{H}$, by analyzing subgroups of N conjugate in G to subgroups of H. (Note that there is only one conjugacy class of subgroups in G isomorphic to N.) Observe that a subgroup of N is isomorphic to one of the following groups: $\mathbb{Z}_q^2$, $\mathbb{Z}_q^2 \rtimes \mathbb{Z}_{q-1}$, $\mathbb{Z}_q^2 \rtimes \mathbb{Z}_l$, where $2 \leq l < q-1$, $\mathbb{Z}_q$, $\mathbb{Z}_q \rtimes \mathbb{Z}_{q-1}$, $\mathbb{Z}_q \rtimes \mathbb{Z}_l$, where $2 \leq l < q-1$, and $\mathbb{Z}_l$, where l divides $q-1$. Since Q is semiregular on $\mathcal{H}$ no subgroup of N containing Q fixes a coset in $\mathcal{H}$ (that is, no subgroup of N containing Q is conjugate to a subgroup of H). Further, there exists unique subgroup of order q^2 in N, which clearly contains Q, and so this subgroup cannot fix a coset in $\mathcal{H}$ as well. Therefore, we only need to consider subgroups of N isomorphic to $\mathbb{Z}_q \rtimes \mathbb{Z}_l$ and $\mathbb{Z}_l$, where l divides $q-1$.

The group N contains $q+1$ conjugacy classes of maximal subgroups isomorphic to $\mathbb{Z}_q \rtimes \mathbb{Z}_{q-1}$, which are divided into two G-conjugate subsets of equal size, with the respective representatives:

$$K = \left\{ \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} : a \in \langle\alpha\rangle, b \in F_q \right\} \quad \text{and} \quad I = \left\{ \begin{bmatrix} a & b\beta \\ 0 & a^{-1} \end{bmatrix} : a \in \langle\alpha\rangle, b \in F_q \right\},$$

where K is contained in H and I is not. Let $K_i = K^g$ be a subgroup of N conjugate to K. Since H has only one conjugacy class of subgroups isomorphic to K, we have $t = 1$ (for the meaning of t, see Proposition 4.1). Since $N_G(K) = N_H(K) = K$, it therefore follows from Proposition 4.1 that K_i fixes only the coset Hg. In view of maximality of K_i in N, the N-orbit of Hg on $\mathcal{H}$ is of length $|N|/|K_i| = q$. Since the G-conjugates of K in N form $\frac{q+1}{2}$ different conjugacy classes inside N, we can conclude that N has $\frac{q+1}{2}$ orbits of length q.

Let K_0 be the subgroup of order q in K. Since $|N_G(K_0) : N_H(K_0)| = |N : K| = q$, any $K_0^g \leq N$ fixes q cosets, which form the N-orbit containing Hg (see the the previous paragraph). Let K_1 be a subgroup of K isomorphic to $\mathbb{Z}_q \rtimes \mathbb{Z}_l$, where $l \mid q-1$ and $l \notin \{1, q-1\}$. One may check that any $K_1^g \leq N$ has the same fixed cosets as K (and so it is a subgroup of a coset stabilizer in N). Consequently N does not have orbits of length $q \cdot \frac{q-1}{l}$ for $1 \leq l < q-1$. Further, for any subgroup $K_2 \leq K$ of H isomorphic to $\mathbb{Z}_l$, where l divides $q-1$ and $l \geq 3$, the fact that $|N_G(K_2) : N_H(K_2)| = |D_{q^2-1} : D_{2(q-1)}| = \frac{q+1}{2}$, implies that K_2 fixes $\frac{q+1}{2}$ cosets. These cosets are clearly contained in the above $\frac{q+1}{2}$ orbits of N of length q, and consequently N does not have orbits of length $\frac{q-1}{l}$.

We have therefore shown that the only other possible stabilizers are $\mathbb{Z}_2$ and $\mathbb{Z}_1$. Since $|\mathcal{H}| = q(q^2+1)/2$ and since the length of an orbit of N on $\mathcal{H}$ with coset stabilizer isomorphic to $\mathbb{Z}_2$ or to $\mathbb{Z}_1$ equals, respectively, $\frac{q^2(q-1)}{2}$ and $q^2(q-1)$, we have

$$\frac{q(q^2+1)}{2} = q\frac{q+1}{2} + a\frac{q^2(q-1)}{2} + bq^2(q-1), \tag{4.1}$$

where a is the number of orbits of N on $\mathcal{H}$ with coset stabilizer isomorphic to $\mathbb{Z}_2$ and b is the number of orbits of N on $\mathcal{H}$ on which N acts regularly. The equation (4.1) simplifies to $q^2 = q + aq(q-1) + 2bq(q-1)$, which clearly has $a = 1$ and $b = 0$ as the only possible solution. This completes the proof of Lemma 4.2. □

Lemma 4.2 will play an essential part in our construction of Hamilton cycles in basic orbital graphs arising from the action of $\mathrm{PSL}(2,q^2)$ on cosets of $\mathrm{PGL}(2,q)$ given in Row 5 of Table 1. The strategy goes as follows. Let X be such an orbital graph. By Lemma 4.2, the action of the normalizer $N = N_G(Q)$ on the quotient graph $X_{\mathcal{Q}}$ with respect to the orbits $\mathcal{Q}$ of a semiregular subgroup Q consists of one large orbit of length $q(q-1)/2$ and $(q+1)/2$ isolated fixed points. We will show the existence of a Hamilton cycle in X by first showing that the subgraph of $X_{\mathcal{Q}}$ induced on the large orbit has at most two connected components and that each component contains a Hamilton cycle with double edges in the corresponding quotient multigraph. If there is only one component then its Hamilton cycle is modified to a Hamilton cycle in $X_{\mathcal{Q}}$ by choosing in an arbitrary manner $(q+1)/2$ edges and replacing them by 2-paths having as central vertices the $(q+1)/2$ isolated fixed points of N in $X_{\mathcal{Q}}$. By Lemma 3.4, this cycle lifts to a Hamilton cycle in X. Such 2-paths indeed exist because every isolated fixed point has to be adjacent to every vertex in the large orbit (see Lemma 4.5). If the subgraph of $X_{\mathcal{Q}}$ induced on the large orbit has two components with corresponding Hamilton cycles C_0 and C_1, then a Hamilton cycle in X is constructed by first constructing a Hamilton cycle in $X_{\mathcal{Q}}$ in the following way. We use two isolated fixed points to modify these two cycles C_0 and C_1 into a cycle of length $q^2(q-1)/2+2$ by replacing an edge in C_0 and an edge in C_1 by two 2-paths each having one endvertex in C_0 and the other in C_1, whereas the central vertices are the above two isolated fixed points. In order to produce the desired Hamilton cycle in $X_{\mathcal{Q}}$ the remaining isolated fixed points are attached to this cycle in the same manner as in the case of one component. By Lemma 3.4, this cycle lifts to a Hamilton cycle in X. Formal proofs are given in Propositions 4.7 and 4.8.

It follows from the previous paragraph that we only need to prove that the subgraph of $X_{\mathcal{Q}}$ induced on the large orbit of N contains a Hamilton cycle with at least one double edge in the corresponding multigraph or two components each of which contains a Hamilton cycle with double edges in the corresponding multigraph. For this purpose we now proceed with the analysis of the structure of basic orbital graphs (and corresponding suborbits) arising from the action of $\mathrm{PSL}(2,q^2)$ on cosets of $\mathrm{PGL}(2,q)$ given in Row 5 of Table 1. We apply the approach taken in [34] where the computation of suborbits is done using the fact that $\mathrm{PSL}(2,q^2) \cong \mathrm{P}\Omega^-(4,q)$ and that the action of $\mathrm{PSL}(2,q^2)$ on the cosets of $\mathrm{PGL}(2,q)$ is equivalent to the induced action of $\mathrm{P}\Omega^-(4,q)$ on nonsingular 1-dimensional vector subspaces $\langle \mathbf{v} \rangle$ such that $\mathbf{Q}(\mathbf{v}) = 1$, where $\mathbf{Q}$ is the associated quadratic form. For the sake of completeness, we give a more detailed description of this action together with a short explanation of the isomorphism $\mathrm{PSL}(2,q^2) \cong \mathrm{P}\Omega^-(4,q)$ (see [14, p. 45] for details).

Let $\phi \in \mathrm{Aut}(F_{q^2})$ be the Frobenius automorphism of F_{q^2} defined by the rule $\phi(a) = a^q$, $a \in F_{q^2}$. (Note that ϕ is an involution.) Let $W = \langle \mathbf{w}_1, \mathbf{w}_2 \rangle = F_{q^2}^2$ be a natural $\mathrm{SL}(2,q^2)$-module. Then $\mathrm{SL}(2,q^2)$ acts on W in a natural way. In particular, the action of $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,q^2)$ on W is given by

$$\begin{aligned} \mathbf{w}_1 g &= a\mathbf{w}_1 + b\mathbf{w}_2, \\ \mathbf{w}_2 g &= c\mathbf{w}_1 + d\mathbf{w}_2. \end{aligned}$$

Let $\overline{W}$ be an $\mathrm{SL}(2,q^2)$-module with the underlying space W and the action of $\mathrm{SL}(2,q^2)$ defined by the rule $\mathbf{w} * g = \mathbf{w} g^{\phi}$, where $g = (a_{ij}) \in \mathrm{SL}(2,q^2)$ and $g^{\phi} = (\phi(a_{ij})_{ij}) = (a_{ij}^q)$. One can now see that the action $\cdot \colon W \otimes \overline{W} \times \mathrm{SL}(2,q^2) \to W \otimes \overline{W}$ defined by the rule

$$(\mathbf{w} \otimes \mathbf{w}') \cdot g = \mathbf{w}g \otimes \mathbf{w}' * g = \mathbf{w}g \otimes \mathbf{w}' g^{\phi}$$

is an action of $\mathrm{SL}(2, q^2)$ on the 4-dimensional space $W \otimes \overline{W}$ (that is, on a tensor product of W and $\overline{W}$). The kernel of this action equals $Z(\mathrm{SL}(2, q^2))$, and thus this is in fact a 4-dimensional representation of $G = \mathrm{PSL}(2, q^2)$ (an embedding of G into $\mathrm{GL}(4, q^2)$). Further, the set $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, where $\mathbf{v}_1 = \mathbf{w}_1 \otimes \mathbf{w}_1$, $\mathbf{v}_2 = \mathbf{w}_2 \otimes \mathbf{w}_2$, $\mathbf{v}_3 = \mathbf{w}_1 \otimes \mathbf{w}_2 + \mathbf{w}_2 \otimes \mathbf{w}_1$, $\mathbf{v}_4 = \alpha(\mathbf{w}_1 \otimes \mathbf{w}_2 - \mathbf{w}_2 \otimes \mathbf{w}_1)$, is a basis for $W \otimes \overline{W}$ over F_{q^2}.

Since G fixes the 4-dimensional space $V = \mathrm{span}_{F_q}(\mathcal{B})$ over F_q it can be viewed as a subgroup of $\mathrm{GL}(4, q)$. A non-degenerate symplectic form f of W and $\overline{W}$ defined by $f(\mathbf{w}_1, \mathbf{w}_2) = -f(\mathbf{w}_2, \mathbf{w}_1) = 1$ and $f(\mathbf{w}_1, \mathbf{w}_1) = f(\mathbf{w}_2, \mathbf{w}_2) = 0$ is fixed by $\mathrm{SL}(2, q^2)$. It follows that G fixes a non-degenerate symmetric bilinear form of $W \otimes \overline{W}$ defined by the rule

$$(\mathbf{w}_1' \otimes \mathbf{w}_2', \mathbf{w}_1'' \otimes \mathbf{w}_2'') = f(\mathbf{w}_1', \mathbf{w}_1'') f(\mathbf{w}_2', \mathbf{w}_2'').$$

Then we have

$$((\mathbf{v}_i, \mathbf{v}_j))_{4\times 4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 2\theta \end{pmatrix},$$

and so for $\mathbf{x} = \sum_{i=1}^4 x_i v_i \in V$ and $\mathbf{y} = \sum_{i=1}^4 y_i v_i \in V$ the symmetric form $(\mathbf{x}, \mathbf{y})$ and the associated quadratic form $\mathbf{Q}$ are given by the rules

$$(\mathbf{x}, \mathbf{y}) = x_2 y_1 + x_1 y_2 - 2x_3 y_3 + 2\theta x_4 y_4 \quad \text{and}$$
$$\mathbf{Q}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}, \mathbf{x}) = x_1 x_2 - x_3^2 + \theta x_4^2.$$

By computation it follows that $\mathbf{Q}$ has $q^2 + 1$ singular 1-dimensional subspaces of V. As for the remaining $q(q^2 + 1)$ nonsingular 1-dimensional subspaces, G has two orbits $\{\langle \mathbf{v} \rangle : \mathbf{Q}(\mathbf{v}) = 1, \mathbf{v} \in V\}$ and $\{\langle \mathbf{v} \rangle : \mathbf{Q}(\mathbf{v}) \in F_q^* \setminus S^*, \mathbf{v} \in V\}$. Since these two representations of G are equivalent, we set Ω to be the first of these two orbits. Then the action of G on $\mathcal{H}$ is equivalent to the action of G on Ω. By comparing their orders, we get $\mathrm{PSL}(2, q^2) \cong \mathrm{P}\Omega^-(4, q)$. The following result characterizing suborbits of the action of G on the cosets of $\mathrm{PGL}(2, q)$ in the context of the action of $\mathrm{P}\Omega^-(4, q)$ on Ω was proved in [34] (see also [30]).

Proposition 4.3 ([34, Lemma 4.1]). *For any $\langle \mathbf{v} \rangle \in \Omega$ where $Q(\mathbf{v}) = 1$, the nontrivial suborbits of the action of G on Ω (that is, the orbits of $G_{\langle \mathbf{v} \rangle}$) are the sets $\mathcal{S}_{\pm\lambda} = \{\langle \mathbf{x} \rangle \in \Omega : (\mathbf{x}, \mathbf{v}) = \pm 2\lambda, Q(\mathbf{x}) = 1\}$, where $\lambda \in F_q$, and*

(i) $|\mathcal{S}_0| = \frac{q(q\mp 1)}{2}$ *for* $q \equiv \pm 1 \pmod 4$;

(ii) $|\mathcal{S}_{\pm 1}| = q^2 - 1$;

(iii) $|\mathcal{S}_{\pm\lambda}| = q(q + 1)$ *for* $\lambda^2 - 1 \in N^*$;

(iv) $|\mathcal{S}_{\pm\lambda}| = q(q - 1)$ *for* $\lambda^2 - 1 \in S^*$.

Moreover, all the suborbits are self-paired.

Let $X = X(G, H, \mathcal{S}_\lambda)$ be the basic orbital graph associated with $\mathcal{S}_\lambda$, and take

$$\rho = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in G.$$

For $k \in F_q$ we have

$$\begin{aligned} \mathbf{v}_1\rho^k &= \mathbf{v}_1 + k^2\mathbf{v}_2 + k\mathbf{v}_3, \\ \mathbf{v}_2\rho^k &= \mathbf{v}_2, \\ \mathbf{v}_3\rho^k &= 2k\mathbf{v}_2 + \mathbf{v}_3, \\ \mathbf{v}_4\rho^k &= \mathbf{v}_4, \end{aligned}$$

and so ρ^k maps the vector $x = \sum_{i=1}^4 x_i\mathbf{v}_i \in V$ to

$$\mathbf{x}\rho^k = x_1\mathbf{v}_1 + (k^2x_1 + x_2 + 2kx_3)\mathbf{v}_2 + (kx_1 + x_3)\mathbf{v}_3 + x_4\mathbf{v}_4.$$

Identifying $\mathbf{x}$ with (x_1, x_2, x_3, x_4) we have $\mathbf{x}\rho^k = (x_1, k^2x_1 + x_2 + 2kx_3, kx_1 + x_3, x_4)$. One can check that for $k \neq 0$ we have $\langle \mathbf{x}\rho^k \rangle \neq \langle \mathbf{x} \rangle$, and thus ρ is (p, q)-semiregular. Let $Q = \langle \rho \rangle$, and let $\mathcal{Q}$ be the set of orbits of Q. These orbits will be referred to as blocks. The set Ω decomposes into two subsets each of which is a union of blocks from $\mathcal{Q}$:

$$\begin{aligned} \mathcal{I} = \langle (0, 0, x_3, x_4) \rangle Q &= \{\langle (0, 2kx_3, x_3, x_4) \rangle : k \in F_q\}, \\ &\qquad \text{where} \quad -x_3^2 + \theta x_4^2 = 1. \\ \mathcal{L} = \langle (x_1, x_2, 0, x_4) \rangle Q &= \{\langle (x_1, k^2x_1 + x_2, kx_1, x_4) \rangle : k \in F_q\}, \\ &\qquad \text{where} \quad x_1 \neq 0 \quad \text{and} \quad x_1x_2 + \theta x_4^2 = 1. \end{aligned}$$

Note that the subset $\mathcal{I}$ contains $\frac{q(q+1)}{2}$ vertices which form $\frac{q+1}{2}$ blocks, and the subset $\mathcal{L}$ contains $\frac{q^2(q-1)}{2}$ vertices which form $\frac{q(q-1)}{2}$ blocks. By $\mathcal{I}_Q$ and $\mathcal{L}_Q$, we denote, respectively, the set of blocks in $\mathcal{I}$ and $\mathcal{L}$; that is, $\mathcal{Q} = \mathcal{I}_Q \cup \mathcal{L}_Q$.

Remark 4.4. Recall that

$$N = N_G(Q) = \left\langle \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} : a \in \langle \alpha \rangle, b \in F_{q^2} \right\rangle.$$

One may check directly that $\mathcal{I}_Q$ consists precisely of the orbits of N of length q and that $\mathcal{L}$ is the orbit of N of length $\frac{q^2(q-1)}{2}$.

In the next lemma we observe that $X\langle\mathcal{L}\rangle$ and $X\langle\mathcal{L}\rangle_\mathcal{Q}$ are vertex-transitive and show that the bipartite subgraph of $X_\mathcal{Q}$ induced by $\mathcal{I}_Q$ and $\mathcal{L}_Q$ is a complete bipartite graph.

Lemma 4.5. *With the above notation, the following hold:*

(i) The induced subgraph $X\langle\mathcal{L}\rangle$ and the quotient graph $X\langle\mathcal{L}\rangle_\mathcal{Q}$ are both vertex-transitive.

(ii) For $\langle \mathbf{x} \rangle Q \in \mathcal{I}_Q$ and $\langle \mathbf{y} \rangle Q \in \mathcal{L}_Q$ we have

$$d(\langle \mathbf{x} \rangle Q, \langle \mathbf{y} \rangle Q) = \begin{cases} 1, & \textit{if } \lambda = 0, \\ 2, & \textit{if } \lambda \neq 0. \end{cases}$$

Proof. By Remark 4.4, N is transitive on $\mathcal{L}$, and so the induced subgraph $X\langle\mathcal{L}\rangle$ and the quotient graph $X\langle\mathcal{L}\rangle_Q$ are both vertex transitive, and thus (i) holds.

To prove (ii), take two arbitrary blocks $\langle\mathbf{x}\rangle Q \in \mathcal{I}_Q$ where $\mathbf{x} = (0,0,x_3,x_4)$ and $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$ where $\mathbf{y} = (y_1,y_2,0,y_4)$. Then $y_1 \neq 0$ and $x_3 \neq 0$, and $\langle\mathbf{x}\rangle \sim \langle\mathbf{y}\rho^k\rangle$ if and only if

$$(\mathbf{x},\mathbf{y}\rho^k) = ((0,0,x_3,x_4),(y_1,k^2y_1+y_2,ky_1,y_4)) = \pm 2\lambda,$$

that is, if and only if

$$-2x_3ky_1 + 2\theta x_4y_4 = \pm 2\lambda. \tag{4.2}$$

From (4.2) we get that $k = \frac{\theta x_4 y_4 \mp \lambda}{x_3 y_1}$ and so for given $\langle\mathbf{x}\rangle$ and $\langle\mathbf{y}\rangle$ we have a unique solution for k if $\lambda = 0$ and two solutions if $\lambda \neq 0$. It follows that for $\langle\mathbf{x}\rangle Q \in \mathcal{I}_Q$ and $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$ we have $d(\langle\mathbf{x}\rangle Q,\langle\mathbf{y}\rangle Q) = 1$ or 2, depending on whether $\lambda = 0$ or $\lambda \neq 0$, completing part (ii) of Lemma 4.5. □

In what follows, we divide the proof into two cases depending on whether $\lambda = 0$ or $\lambda \neq 0$.

4.1 Case $\mathcal{S}_0$

Let

$$\varepsilon = \begin{cases} 2, & \text{if } q \equiv 1,3 \pmod 8, \\ 0, & \text{if } q \equiv 5,7 \pmod 8). \end{cases}$$

The following lemma gives us the number of edges inside a block and between two blocks from $\mathcal{L}_Q$ for the orbital graph $X(G,H,\mathcal{S}_0)$.

Lemma 4.6. *Let $X = X(G,H,\mathcal{S}_0)$. Then for $\langle\mathbf{x}\rangle Q \in \mathcal{L}_Q$ the following hold:*

(i) $d(\langle\mathbf{x}\rangle Q) = \varepsilon$,

(ii) $d(\langle\mathbf{x}\rangle Q,\langle\mathbf{y}\rangle Q) = 1$ *for* $\frac{q+1}{2}$ *blocks* $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$,

(iii) $d(\langle\mathbf{x}\rangle Q,\langle\mathbf{y}\rangle Q) = 2$ *for* $\frac{1}{4}(q^2-3q-2(\varepsilon+1))$ *blocks* $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$ *if* $q \equiv 1 \pmod 4$, *and for* $\frac{1}{4}(q^2-q-2(\varepsilon+1))$ *blocks* $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$ *if* $q \equiv 3 \pmod 4$.

Proof. Fix a block $\langle\mathbf{x}\rangle Q \in \mathcal{L}_Q$ where $\mathbf{x} = (1,1,0,0)$. For any $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$, where $\mathbf{y} = (y_1,y_2,0,y_4)$ with $y_1 \neq 0$, we have $\langle\mathbf{x}\rangle \sim \langle\mathbf{y}\rangle\rho^k$ if and only if $(k^2+1)y_1 + y_2 = 0$, and therefore, since $y_1y_2 + \theta y_4^2 = 1$, if and only if

$$k^2 = -y_1^{-2} + \theta(y_1^{-1}y_4)^2 - 1. \tag{4.3}$$

It follows from (4.3) that $\langle\mathbf{x}\rangle$ is adjacent to one vertex in the block $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$ if $k = 0$ and to two vertices in this block if $k \neq 0$. Clearly, $k = 0$ if and only if

$$\theta y_4^2 = 1 + y_1^2. \tag{4.4}$$

Proposition 2.4 implies that (4.4) has $q+1$ solutions for (y_1,y_4), and therefore since $\langle\mathbf{y}\rangle = \langle-\mathbf{y}\rangle$ we have a total of $\frac{q+1}{2}$ choices for $\langle\mathbf{y}\rangle$. This implies that $d(\langle\mathbf{x}\rangle Q,\langle\mathbf{y}\rangle Q) = 1$ for $\frac{q+1}{2}$ blocks $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$, proving part (ii).

To prove part (i), take $\mathbf{y} = \pm\mathbf{x} = \pm(1, 1, 0, 0)$. Then, by (4.3), there are edges inside the block $\langle\mathbf{x}\rangle Q$ if and only if $k^2 = -2$. This equation has solutions if and only if $q \equiv 1, 3 \pmod 8$ (see Propositions 2.1 and 2.2), and thus the induced subgraph $X\langle\langle\mathbf{x}\rangle Q\rangle$ is a q-cycle for $q \equiv 1, 3 \pmod 8$ and a totally disconnected graph qK_1 if $q \equiv 5, 7 \pmod 8$.

Finally, to prove part (iii) let m be the number of blocks $\langle\mathbf{y}\rangle Q \in \mathcal{L}_Q$ for which $d(\langle\mathbf{x}\rangle Q, \langle\mathbf{y}\rangle Q) = 2$. Suppose first that $q \equiv 1 \pmod 4$. Then, combining together the facts that X is of valency $\frac{1}{2}q(q-1)$, that $d(\langle\mathbf{x}\rangle Q) = \varepsilon$ and that $\langle\mathbf{x}\rangle$ is adjacent to $\frac{1}{2}(q+1)$ vertices in the set $\mathcal{I}$ and to exactly one vertex from $\frac{q+1}{2}$ blocks in $\mathcal{L}_Q$, we have

$$m = \frac{1}{2}\left(\frac{1}{2}q(q-1) - \frac{q+1}{2} - \frac{q+1}{2} - \varepsilon\right) = \frac{1}{4}(q^2 - 3q - 2(1+\varepsilon)).$$

Suppose now that $q \equiv 3 \pmod 4$. Then, replacing the valency of X in the above computation with $\frac{1}{2}q(q+1)$ we obtain, as desired, that $m = \frac{1}{4}(q^2 - q - 2(1+\varepsilon))$. □

We are now ready to prove existence of a Hamilton cycle in $X(G, H, \mathcal{S}_0)$.

Proposition 4.7. *The graph $X = X(G, H, \mathcal{S}_0)$ is hamiltonian.*

Proof. Let $X\langle\mathcal{L}\rangle'$ be the graph obtained from $X\langle\mathcal{L}\rangle$ by deleting the edges between any two blocks $B_1, B_2 \in \mathcal{L}_Q$ for which $d(B_1, B_2) = 1$ (see Lemma 4.6(ii)). By Lemma 4.5, $X\langle\mathcal{L}\rangle_{\mathcal{Q}}$ is vertex-transitive, and consequently one can see that also $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ is vertex-transitive.

If $q \equiv 1 \pmod 4$ then Lemma 4.6(iii) implies that $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ is of valency $m = \frac{1}{4}(q^2 - 3q - 2(1+\varepsilon))$. If, however, $q \equiv 3 \pmod 4$ then Lemma 4.6(iii) implies that $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ is of valency $m = \frac{1}{4}(q^2 - q - 2(1+\varepsilon))$. If $q = 5$ then $\varepsilon = 0$ and $m = \frac{1}{4}(q^2 - 3q - 2(1+\varepsilon)) = 2$. If $q \geq 7$ then using the facts that $q^2 - 7q - 6(1+\varepsilon) \geq 0$ for $q \equiv 1 \pmod 4$ and that $q^2 - q - 6(1+\varepsilon) \geq 0$ for $q \equiv 3 \pmod 4$ one can see that

$$m = \frac{1}{4}(q^2 - (2 \pm 1)q - 2(1+\varepsilon)) \geq \frac{1}{3}\frac{q(q-1)}{2} = \frac{1}{3}|\mathcal{L}_Q|.$$

Suppose first that $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ is connected. If $q = 5$, then $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ is just a cycle C. For $q \geq 7$, by Proposition 3.3, $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ admits a Hamilton cycle, say C again. Clearly C is also a Hamilton cycle of $X\langle\mathcal{L}\rangle_{\mathcal{Q}}$. Form C a Hamilton cycle in $X_{\mathcal{Q}}$ can be constructed by choosing arbitrarily $(q+1)/2$ edges and replacing them by 2-paths having as central vertices the $(q+1)/2$ isolated fixed points of N in $X_{\mathcal{Q}}$. By Lemma 3.4, this lifts to a Hamilton cycle in X.

Next, suppose that $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ is disconnected. For $q = 5$, since $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ is a vertex transitive graph of order 10 and degree 2, it must be a union of two 5-cycles. For $q \geq 7$, since $m \geq \frac{1}{3}|\mathcal{L}_Q|$, it follows that $X\langle\mathcal{L}\rangle'_{\mathcal{Q}}$ has just two components. By Proposition 3.3, each component admits a Hamilton cycle. Take a respective Hamilton path for each component, say $\mathcal{U} = U_1U_2\cdots U_l$, and $\mathcal{U}' = U'_1U'_2\cdots U'_l$, where $l = \frac{q(q-1)}{4}$. Choose any two isolated fixed points W_1 and W_2 and construct the cycle $\mathcal{D} = W_1\mathcal{U}W_2\mathcal{U}'W_1$. Choose arbitrarily $(q+1)/2 - 2$ edges in $\mathcal{U} \cup \mathcal{U}'$ and replace them by 2-paths having as central vertices the remaining $(q+1)/2 - 2$ isolated fixed points. Then we get a Hamilton cycle in $X_{\mathcal{Q}}$, which, by Lemma 3.4, lifts to a Hamilton cycle in X. □

4.2 Case $\mathcal{S}_\lambda$ with $\lambda \neq 0$

Proposition 4.8. *The graph $X = X(G, H, \mathcal{S}_{\pm\lambda})$, where $\lambda \neq 0$, is hamiltonian.*

Proof. As in the proof of Lemma 4.6, fix a block $\langle \mathbf{x} \rangle Q \in \mathcal{L}_Q$ where $\mathbf{x} = (1, 1, 0, 0)$. For any $\langle \mathbf{y} \rangle Q \in \mathcal{L}_Q$ where $\mathbf{y} = (y_1, y_2, 0, y_4)$ with $y_1 \neq 0$, we have $\mathbf{y}\rho^k = (y_1, k^2 y_1 + y_2, k y_1, y_4)$, and so $\langle \mathbf{x} \rangle \sim \langle \mathbf{y}\rho^k \rangle$ if and only if $(k^2 + 1)y_1 + y_2 = \pm 2\lambda$, which implies, since $y_1 y_2 + \theta y_4^2 = 1$, that $k^2 = \pm 2\lambda y_1^{-1} - y_1^{-2} + \theta(y_1^{-1} y_4)^2 - 1$. It follows that there are at most four solutions for k. Hence each vertex in $\mathcal{L}$ is adjacent to at most four vertices in each block from $\mathcal{L}_Q$ (including the block containing this vertex).

Let m be the valency of $X\langle \mathcal{L} \rangle_Q$. Since, by Proposition 4.3, the valency of X is, respectively, $q^2 - 1$, $q^2 - q$ and $q^2 + q$, we get that $m \geq \frac{1}{3}|\mathcal{L}_P| = \frac{1}{3}\frac{q(q-1)}{2}$ provided

$$m \geq \frac{1}{4}((q^2 - j) - (q + 1) - 4) = \frac{1}{4}(q^2 - q - j - 5) \geq \frac{1}{3}\frac{q(q-1)}{2},$$

where $j \in \{1, q, -q\}$ for $q \geq 7$ and $j \in \{1, -q\}$ for $q = 5$. One can check that this inequality holds for all $q \geq 5$. We can therefore conclude that $X\langle \mathcal{L} \rangle_Q$, which is vertex-transitive by Lemma 4.5, has at most two connected components. The rest of the argument follows word by word from the argument given in the proof of Proposition 4.7, since, by Lemma 4.5, $d(\langle \mathbf{x} \rangle Q, \langle \mathbf{y} \rangle Q) = 2$, for any $\langle \mathbf{x} \rangle Q \in \mathcal{I}_Q$ and $\langle \mathbf{y} \rangle Q \in \mathcal{L}_Q$. □

5 Proof of Theorem 3.2

Proof of Theorem 3.2. Let X be a connected vertex-transitive graph of order pq, where q and $p = (q^2+1)/2$ are primes, arising the action given in Row 5 of Table 1. As explained in Section 3, we can assume that X is a basic orbital graph arising from a group action given in Row 5 of Table 1, and thus it admits a Hamilton cycle by Propositions 4.7 and 4.8. □

ORCID iDs

Shaofei Du https://orcid.org/0000-0001-6725-9293
Klavdija Kutnar https://orcid.org/0000-0002-9836-6398
Dragan Marušič https://orcid.org/0000-0002-8452-3057

References

[1] B. Alspach, Lifting Hamilton cycles of quotient graphs, *Discrete Math.* **78** (1989), 25–36, doi: 10.1016/0012-365x(89)90157-x.

[2] B. Alspach and T. D. Parsons, On Hamiltonian cycles in metacirculant graphs, in: E. Mendelsohn (ed.), *Algebraic and Geometric Combinatorics*, North-Holland, Amsterdam, volume 65 of *North-Holland Mathematics Studies*, pp. 1–7, 1982, doi:10.1016/s0304-0208(08)73249-3.

[3] P. J. Cameron, M. Giudici, G. A. Jones, W. M. Kantor, M. H. Klin, D. Marušič and L. A. Nowitz, Transitive permutation groups without semiregular subgroups, *J. London Math. Soc.* **66** (2002), 325–333, doi:10.1112/s0024610702003484.

[4] P. J. Cameron (ed.), Problems from the Fifteenth British Combinatorial Conference, *Discrete Math.* **167/168** (1997), 605–615, doi:10.1016/s0012-365x(96)00212-9.

[5] Y. Q. Chen, On Hamiltonicity of vertex-transitive graphs and digraphs of order p^4, *J. Comb. Theory Ser. B* **72** (1998), 110–121, doi:10.1006/jctb.1997.1796.

[6] S. J. Curran and J. A. Gallian, Hamiltonian cycles and paths in Cayley graphs and digraphs—a survey, *Discrete Math.* **156** (1996), 1–18, doi:10.1016/0012-365x(95)00072-5.

[7] E. Dobson, A. Malnič, D. Marušič and L. A. Nowitz, Semiregular automorphisms of vertex-transitive graphs of certain valencies, *J. Comb. Theory Ser. B* **97** (2007), 371–380, doi:10.1016/j.jctb.2006.06.004.

[8] E. Durnberger, Connected Cayley graphs of semidirect products of cyclic groups of prime order by abelian groups are Hamiltonian, *Discrete Math.* **46** (1983), 55–68, doi:10.1016/0012-365x(83)90270-4.

[9] E. Ghaderpour and D. W. Morris, Cayley graphs on nilpotent groups with cyclic commutator subgroup are Hamiltonian, *Ars Math. Contemp.* **7** (2014), 55–72, doi:10.26493/1855-3974.280.8d3.

[10] M. Giudici, Quasiprimitive groups with no fixed point free elements of prime order, *J. London Math. Soc.* **67** (2003), 73–84, doi:10.1112/s0024610702003812.

[11] H. Glover and D. Marušič, Hamiltonicity of cubic Cayley graphs, *J. Eur. Math. Soc.* **9** (2007), 775–787, doi:10.4171/jems/96.

[12] H. H. Glover, K. Kutnar, A. Malnič and D. Marušič, Hamilton cycles in $(2, \mathrm{odd}, 3)$-Cayley graphs, *Proc. Lond. Math. Soc.* **104** (2012), 1171–1197, doi:10.1112/plms/pdr042.

[13] B. Jackson, Hamilton cycles in regular 2-connected graphs, *J. Comb. Theory Ser. B* **29** (1980), 27–46, doi:10.1016/0095-8956(80)90042-8.

[14] P. Kleidman and M. Liebeck, *The Subgroup Structure of the Finite Classical Groups*, volume 129 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, 1990, doi:10.1017/cbo9780511629235.

[15] K. Kutnar and D. Marušič, Hamiltonicity of vertex-transitive graphs of order $4p$, *European J. Combin.* **29** (2008), 423–438, doi:10.1016/j.ejc.2007.02.002.

[16] K. Kutnar and D. Marušič, Hamilton cycles and paths in vertex-transitive graphs — Current directions, *Discrete Math.* **309** (2009), 5491–5500, doi:10.1016/j.disc.2009.02.017.

[17] K. Kutnar, D. Marušič and C. Zhang, Hamilton paths in vertex-transitive graphs of order $10p$, *European J. Combin.* **33** (2012), 1043–1077, doi:10.1016/j.ejc.2012.01.005.

[18] K. Kutnar and P. Šparl, Hamilton paths and cycles in vertex-transitive graphs of order $6p$, *Discrete Math.* **309** (2009), 5444–5460, doi:10.1016/j.disc.2008.12.005.

[19] M. W. Liebeck and J. Saxl, Primitive permutation groups containing an element of large prime order, *J. London Math. Soc.* **31** (1985), 237–249, doi:10.1112/jlms/s2-31.2.237.

[20] L. Lovász, The factorization of graphs, in: R. Guy, H. Hanani, N. Sauer and J. Schönheim (eds.), *Combinatorial Structures and Their Applications*, Gordon and Breach, New York, 1970 pp. 243–246, proceedings of the Calgary International Conference on Combinatorial Structures and Their Applications held at the University of Calgary, Calgary, Alberta, Canada, June 1969.

[21] D. Marušič, On vertex symmetric digraphs, *Discrete Math.* **36** (1981), 69–81, doi:10.1016/0012-365x(81)90174-6.

[22] D. Marušič, Hamiltonian circuits in Cayley graphs, *Discrete Math.* **46** (1983), 49–54, doi:10.1016/0012-365x(83)90269-8.

[23] D. Marušič, Vertex transitive graphs and digraphs of order p^k, in: B. R. Alspach and C. D. Godsil (eds.), *Cycles in Graphs*, North-Holland, Amsterdam, volume 115 of *North-Holland Mathematics Studies*, pp. 115–128, 1985, doi:10.1016/s0304-0208(08)73001-9, papers from the workshop held at Simon Fraser University, Burnaby, British Columbia, July 5 – August 20, 1982.

[24] D. Marušič, Hamiltonian cycles in vertex symmetric graphs of order $2p^2$, *Discrete Math.* **66** (1987), 169–174, doi:10.1016/0012-365x(87)90129-4.

[25] D. Marušič, Hamiltonicity of vertex-transitive pq-graphs, in: J. Nešetřil and M. Fiedler (eds.), *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity*, North-Holland, Amsterdam, volume 51 of *Annals of Discrete Mathematics*, pp. 209–212, 1992, doi:10.1016/s0167-5060(08)70630-7, proceedings of the symposium held in Prachatice, 1990.

[26] D. Marušič and T. D. Parsons, Hamiltonian paths in vertex-symmetric graphs of order $5p$, *Discrete Math.* **42** (1982), 227–242, doi:10.1016/0012-365x(82)90220-5.

[27] D. Marušič and T. D. Parsons, Hamiltonian paths in vertex-symmetric graphs of order $4p$, *Discrete Math.* **43** (1983), 91–96, doi:10.1016/0012-365x(83)90024-9.

[28] D. Marušič and R. Scapellato, Characterizing vertex-transitive pq-graphs with an imprimitive automorphism subgroup, *J. Graph Theory* **16** (1992), 375–387, doi:10.1002/jgt.3190160410.

[29] D. Marušič and R. Scapellato, A class of non-Cayley vertex-transitive graphs associated with $\mathrm{PSL}(2, p)$, *Discrete Math.* **109** (1992), 161–170, doi:10.1016/0012-365x(92)90287-p.

[30] D. Marušič and R. Scapellato, A class of graphs arising from the action of $\mathrm{PSL}(2, q^2)$ on cosets of $\mathrm{PGL}(2, q)$, *Discrete Math.* **134** (1994), 99–110, doi:10.1016/0012-365x(93)e0065-c.

[31] D. Marušič and R. Scapellato, Classifying vertex-transitive graphs whose order is a product of two primes, *Combinatorica* **14** (1994), 187–201, doi:10.1007/bf01215350.

[32] D. W. Morris, Cayley graphs on groups with commutator subgroup of order $2p$ are hamiltonian, *Art Discrete Appl. Math.* **1** (2018), #P1.04 (31 pages), doi:10.26493/2590-9770.1240.60e.

[33] C. E. Praeger, R. J. Wang and M. Y. Xu, Symmetric graphs of order a product of two distinct primes, *J. Comb. Theory Ser. B* **58** (1993), 299–318, doi:10.1006/jctb.1993.1046.

[34] C. E. Praeger and M. Y. Xu, Vertex-primitive graphs of order a product of two distinct primes, *J. Comb. Theory Ser. B* **59** (1993), 245–266, doi:10.1006/jctb.1993.1068.

[35] J. H. Silverman, *A Friendly Introduction to Number Theory*, Pearson Education, 4th edition, 2012, `https://www.math.brown.edu/~jhs/frint.html`.

[36] H. Wielandt, *Finite Permutation Groups*, Academic Press, New York-London, 1964.

[37] D. Witte Morris, Odd-order Cayley graphs with commutator subgroup of order pq are hamiltonian, *Ars Math. Contemp.* **8** (2015), 1–28, doi:10.26493/1855-3974.330.0e6.

[38] J.-Y. Zhang, Vertex-transitive digraphs of order p^5 are Hamiltonian, *Electron. J. Combin.* **22** (2015), #P1.76 (12 pages), doi:10.37236/4034.

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 17–23
https://doi.org/10.26493/1855-3974.1685.395
(Also available at http://amc-journal.eu)

Incidence structures near configurations of type (n_3)*

Peter J. Dukes †, Kaoruko Iwasaki
Mathematics and Statistics, University of Victoria, Victoria, Canada

Received 24 April 2018, accepted 28 June 2020, published online 26 October 2020

Abstract

An (n_3) configuration is an incidence structure equivalent to a linear hypergraph on n vertices which is both 3-regular and 3-uniform. We investigate a variant in which one constraint, say 3-regularity, is present, and we allow exactly one line to have size four, exactly one line to have size two, and all other lines to have size three. In particular, we study planar (Euclidean or projective) representations, settling the existence question and adapting Steinitz' theorem for this setting.

Keywords: Geometric configurations, incidence structures.

Math. Subj. Class. (2020): 52C30, 05C65

1 Introduction

A *geometric* (n_r, m_k) *configuration* is a set of n points and m lines in the Euclidean (or projective) plane such that every line contains exactly k points (uniformity) and every point is incident with exactly r lines (regularity). Grünbaum's book [3] is an excellent reference for the major results on this topic, including many variations.

Our study here is a variation in which uniformity (or, dually, regularity) is mildly relaxed; see also [1] and [4, §6.8] for prior investigations in this direction.

An *incidence structure* is a triple $(P, \mathcal{L}, \iota)$, where P is a set of *points*, $\mathcal{L}$ is a set of *lines*, and $\iota \subseteq P \times \mathcal{L}$ is a relation called *incidence*. Here, we assume P is finite, say $|P| = n$, and no two different lines are incident with the same set of points. In this case, we may identify $\mathcal{L}$ with a set system on P. Equivalently, the points and lines can be regarded as vertices and edges, respectively, of a (finite) hypergraph. A *point-line incidence structure* is an incidence structure in which any two different points are incident with at most one line.

*The authors would like to thank Gary MacGillivray and Vincent Pilaud for helpful discussions.
†Corresponding author. Research supported by NSERC grant number 312595–2017.
E-mail addresses: dukes@uvic.ca (Peter J. Dukes), mphysics.f@gmail.com (Kaoruko Iwasaki)

(Under this assumption, the associated hypergraph is called *linear*.) We make the standard assumption in this setting that each line is incident with at least two points.

A point-line incidence structure in which all lines have the same number of points and all points are incident with the same number of lines is called a (combinatorial) *configuration*. As before, the abbreviation (n_r, m_k) represents that there are n points, each incident with exactly r lines, and m lines, each incident with exactly k points. In alternate language, such an incidence structure is then an r-regular, k-uniform linear hypergraph on n vertices and m edges. Counting incidences in two ways, one has $nr = mk$. In either the geometric or combinatorial setting, notation in the case $n = m$ (hence $k = r$) is simplified to (n_k) *configuration*.

Examples 1.1. The complete graph K_4 is a $(4_3, 6_2)$ configuration, also known as a 'quadrangle'. The Desargues' configuration is a (10_3) configuration. The Fano plane is a (7_3) configuration. The quadrangle and Desargues' configuration are geometric, whereas the Fano plane is not.

A point-line incidence structure is *geometric* if it can be realized with points and lines in the plane, with the usual notion of incidence.

Consider a point-line incidence structure with n_i points of degree r_i, $i = 1, \dots, s$, and m_j lines of size k_j, $j = 1, \dots, t$. As before, counting incidences yields the relation

$$\sum_{i=1}^{s} n_i r_i = \sum_{j=1}^{t} m_j k_j. \tag{1.1}$$

The *signature* of such a point-line incidence structure is the pair of polynomials $(f(x), g(y))$, where $f(x) = \sum_{i=1}^{s} n_i x^{r_i}$ and $g(y) = \sum_{j=1}^{t} m_j y^{k_j}$. Note that (1.1) can be rewritten as $f'(1) = g'(1)$.

In particular, an (n_r, m_k) configuration has signature (nx^r, my^k). As an example with mixed line sizes, deleting one point from the Fano plane results in a point-line incidence structure with signature $(6x^3, 3y^2 + 4y^3)$. For another example, placing an extra point on one line of the Desargues' configuration results in a point-line incidence structure with signature $(x + 10x^3, 4y + 9y^3)$.

We are interested here in point-line incidence structures which are 'approximately' configurations of type (n_3). As some motivation and context, Borowski and Pilaud [1] use point-line incidence structures with signature $(ax^3 + bx^4, cy^3 + dy^4)$ to produce the first known examples of (n_4) configurations for $n = 37$ and 43. In a different direction, Dumnicki et al. [2] study line arrangements in projective planes with maximum number of points of degree 3.

In this note, our focus is on structures with the specific signature

$$(nx^3, y^2 + (n-2)y^3 + y^4) = (nx^3, ny^3 + y^2(1-y)^2),$$

and in particular their geometric realizations. These are very nearly (n_3) configurations, but with a minimal change to the signature. Our main result is that point-line incidence structures with such signatures exist if and only if $n \geq 9$, and in fact there is one with a geometric representation in each case. This is shown in Section 2. On the other hand, not all structures with our signature are geometric. We observe that Steinitz' theorem on representing (n_3) configurations with at most one curved line assumes a stronger form in our setting, essentially characterizing which of our structures are realizable via their bipartite incidence graph. This is covered in Section 3. We conclude with a few extra remarks, including a connection to 'fuzzy' configurations, in which some points are replaced by intervals.

2 Existence

Our first result easily settles the existence question for the abstract combinatorial case.

Theorem 2.1. *There exists a* (*combinatorial*) *point-line incidence structure having signature* $(nx^3, ny^3 + y^2(1-y)^2)$ *if and only if* $n \geq 9$.

Proof. Suppose there exists such an incidence structure. Consider line $L = \{p_1, \dots, p_4\}$ of size four. Each point p_i is incident with exactly two other lines. These 8 additional lines are distinct, by linearity and the fact that p_i, p_j are already together on L for each $1 \leq i < j \leq 4$. It follows that the configuration has at least 9 lines, and so $n \geq 9$.

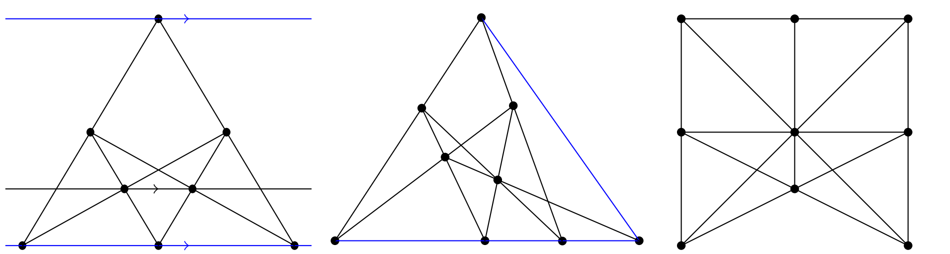
Conversely, suppose $n \geq 9$. Let $\mathcal{L}$ denote the set of translates of $\{0, 1, 3\}$ modulo $n-1$. Then $(\{0, 1, \dots, n-2\}, \mathcal{L})$ is an $((n-1)_3)$ configuration. Add a new point ∞ and take the adjusted family of lines

$$\mathcal{L}' := \mathcal{L} \setminus \{\{0,1,3\}, \{4,5,7\}\} \cup \{\{0,1,3,\infty\}, \{4,\infty\}, \{5,7,\infty\}\}.$$

It is simple to check that $(\{0, 1, \dots, n-2, \infty\}, \mathcal{L}')$ is a point-line incidence structure with the desired signature. □

We turn now to the geometric case. The line adjustments used in the proof of Theorem 2.1 cannot be applied to an arbitrary geometric $((n-1)_3)$ configuration. Moreover, the Mobiüs-Kantor (8_3) configuration does not admit a geometric realization. To this end, we start with an example on 9 points.

Example 2.2. Figure 1 shows a (geometric) point-line incidence structure with signature $(9x^3, y^2 + 7y^3 + y^4)$ given as (a) a projective realization with one point at infinity and (b) a Euclidean realization. The lines of size 2 and 4 are highlighted. The dual incidence structure, in which there is exactly one point of degree 2 and one point of degree 4, is drawn in (c).



(a) Projective (b) Euclidean (c) The dual

Figure 1: A point-line incidence structure with signature $(9x^3, y^2 + 7y^3 + y^4)$.

Here, it is helpful to cite a result which is used (very mildly) below, and (more crucially) for our structural considerations in Section 3. This result, due to Steinitz, roughly says that configurations of type (n_3) are 'nearly' geometric. However, the reader is encouraged to see Grünbaum's discussion [3, §2.6] of 'unwanted incidences', which in certain cases cannot be avoided.

Theorem 2.3 (Steinitz [5]; see also [3])**.** *For every combinatorial (n_3) configuration, there is a representation of all but at most one of its incidences by points and lines in the plane.*

This result, in combination with with Example 2.2 and the existence of combinatorial (n_3) configurations for $n \geq 7$, allows us to adapt the argument in Theorem 2.1 to a geometric one.

Theorem 2.4. *For each $n \geq 9$, there exists a point-line incidence structure having signature $(nx^3, ny^3 + y^2(1-y)^2)$ with incidences represented as points and lines in the plane.*

Proof. First, we consider values $n \geq 14$. Take disjoint configurations $(P, \mathcal{L})$ and $(Q, \mathcal{H})$ of types (7_3) and $((n-7)_3)$, respectively. Apply Steinitz' theorem to each. Assume that $\{p_1, p_2, p_3\} \in \mathcal{L}$, yet line p_1p_2 in the first Steinitz embedding does not contain p_3. Similarly, suppose $\{q_1, q_2, q_3\} \in \mathcal{H}$ where q_3 need not be placed on line q_1q_2. If it is not, we align the drawings in the plane such that line p_1p_2 coincides with line q_1q_2, creating a line of size four. We finish by including the additional line p_3q_3. If the latter configuration has a geometric realization (which we may assume for $n \geq 16$), we can simply place the drawings so that p_3 is on lines $q_1q_2q_3$, and include p_1p_2 as an additional line. In either case, we have a geometric point-line configuration with the desired signature.

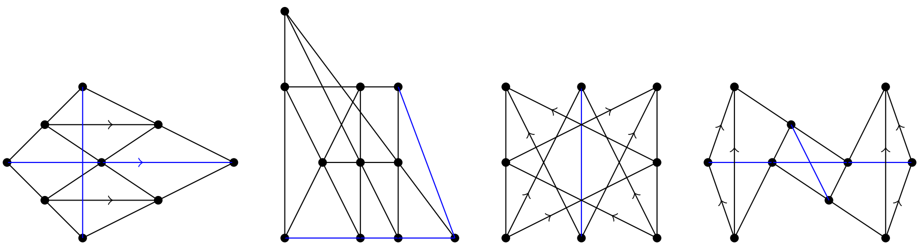


Figure 2: Examples for $n = 10, \ldots, 13$.

For the cases $9 \leq n \leq 13$, we use direct examples. The case $n = 9$ is shown in Example 2.2; the other values are shown in Figure 2. Note that projective points are used in some cases, as indicated with arrows on parallel (Eudlidean) lines. □

Example 2.5. As an illustration of the proof technique, Figure 3 shows how two Fano planes with missing incidences can be aligned to produce an example in the case $n = 14$.

3 Structure

In Theorem 2.4, we merely proved the existence of one incidence structure with signature $(nx^3, ny^3 + y^2(1-y)^2)$ having a geometric representation. It is easy to build (combinatorial) such incidence structures which are non-geometric using a similar method.

Proposition 3.1. *For each $n \geq 14$, there exists a point-line incidence structure having signature $(nx^3, ny^3 + y^2(1-y)^2)$ and with no planar representation.*

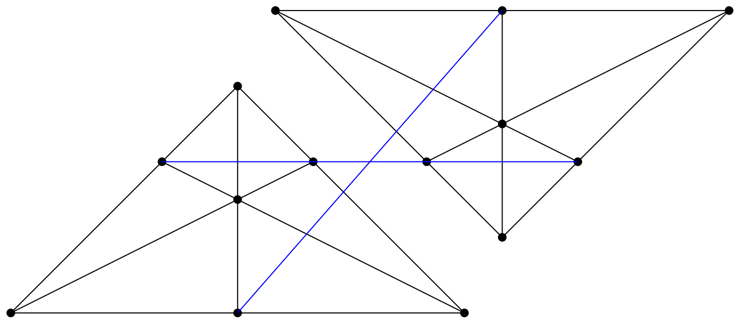


Figure 3: A compound example with $n = 14$ points.

Proof. Take disjoint configurations $(P, \mathcal{L})$ and $(Q, \mathcal{H})$ of types (7_3) and $((n-7)_3)$, respectively. Consider the configuration $(P \cup Q, \mathcal{K})$, where $\mathcal{K}$ is formed from the lines of $\mathcal{L} \cup \mathcal{H}$ by replacing lines $\{p_1, p_2, p_3\} \in \mathcal{L}$ and $\{q_1, q_2, q_3\} \in \mathcal{H}$ with $\{p_1, p_2, p_3, q_4\}$ and $\{q_2, q_3\}$. This incidence structure has the desired signature. It is non-geometric, since its 'restriction' to P is the Fano plane, which has no realization in the plane. □

Given an incidence structure $(P, \mathcal{L}, \iota)$, its *Levi graph* is the bipartite graph with vertex partition $(P, \mathcal{L})$ and edge set $\{\{p, L\} : (p, L) \in \iota\}$. It is easy to see that the Levi graph of a linear incidence structure has girth at least six. Steinitz' argument for obtaining geometric representations essentially works by iteratively removing vertices of lowest degree in the Levi graph, and then drawing the corresponding objects (as needed) in reverse.

We can use the Levi graph to test when a combinatorial point-line incidence structure of our signature has a geometric representation.

Theorem 3.2. *A point-line incidence structure with signature* $(nx^3, ny^3 + y^2(1-y)^2)$ *has a geometric realization if and only if its Levi graph contains no cut-edge whose removal leaves a* 3*-regular component whose corresponding configuration is non-geometric.*

Proof. Necessity of the condition is obvious.

For sufficiency, consider the Levi graph of a structure of the given type. Remove the unique vertex of degree 2, and iteratively remove a vertex of degree less than 3 until there are no more such vertices. Either we succeed in eliminating all vertices of the graph, or some 3-regular subgraph is left over. In the latter case, note that the last vertex removed corresponds to one of the points on the line of size four.

If all vertices are eliminated, simply reverse the list of deletions and follow Steinitz' argument. Since vertices get 'added back' with degree at most two, we obtain a sequence of instructions of one of the following types: placing a new point, drawing a new line, placing a point on an existing line, drawing a line through an existing point, placing a point on the intersection of two existing lines, and drawing a line through a pair of existing points. Each instruction can be carried out in the plane, and we have the desired representation.

Suppose a nonempty 3-regular graph remains after removing vertices. By our assumption, it is the Levi graph of an (n_3) configuration which admits a geometric representation. Carry out the instructions as above to reverse vertex deletions, ensuring to begin by placing a new point on the cut-vertex to create a line of size four. This again produces a representation of incidences in the plane. □

As a simple consequence, we may successfully carry out Steinitz' procedure in this setting provided the vertex of degree four in the Levi graph is incident with no cut-edge.

4 Discussion

A general method which we have found often works for the explicit construction of geometric 'approximate' configurations can be loosely described as 'moving incidences'. It is usually not possible to simply move a point onto an existing line. But, sometimes, a configuration can be perturbed slightly to achieve this. (In doing so, some care must be taken to avoid unwanted incidences.) Our hope is that such a process can be systematically described, producing a wide assortment of perturbations of known configurations. If done in sufficient generality, such a process might handle other types beyond our simple case study of $(nx^3, ny^3 + y^2(1 - y)^2)$.

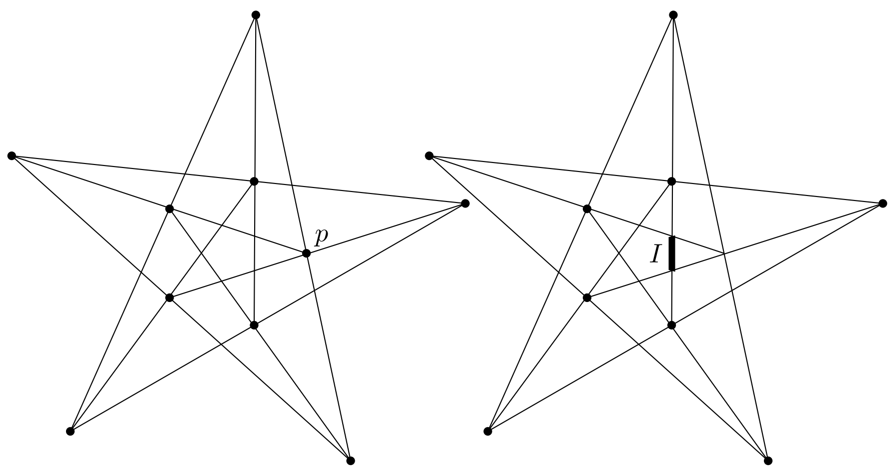


Figure 4: 'Fuzzy' modification of a (10_3) configuration.

One possible way this can succeed is through the intermediate step of a 'fuzzy' realization, which we loosely define as a collection of intervals (possibly points) and lines so that all incidences are represented in the usual way. In this setting, an interval is deemed incident with a line if it intersects that line. Consider the example (10_3) configuration shown on the left of Figure 4. To construct the incidence structure with signature $(10x^3, y^2 + 8y^3 + y^4)$ on the right, one point p is replaced by an abstract point represented geometrically by an interval I incident with two of the lines that went through p. If it is possible to shrink the interval I to a geometric point, maintaining other incidences, then the resulting configuration is geometric. This latter step is possible in the case of the fuzzy realization on the right of Figure 4, and in a few other cases we considered. We feel an interesting question is identifying sufficient conditions for when such interval-eliminating perturbation can be carried out in general.

We have not attempted any enumeration or classification work for structures with the signature $(nx^3, ny^3 + y^2(1 - y)^2)$. But such work may offer some insight into a partial correspondence with, say, geometric $((n - 1)_3)$ configurations.

ORCID iDs

Peter J. Dukes https://orcid.org/0000-0002-5617-083X

References

[1] J. Bokowski and V. Pilaud, Quasi-configurations: building blocks for point-line configurations, *Ars Math. Contemp.* **10** (2016), 99–112, doi:10.26493/1855-3974.642.bbb.

[2] M. Dumnicki, Ł. Farnik, A. Główka, M. Lampa-Baczyńska, G. Malara, T. Szemberg, J. Szpond and H. Tutaj-Gasińska, Line arrangements with the maximal number of triple points, *Geom. Dedicata* **180** (2016), 69–83, doi:10.1007/s10711-015-0091-7.

[3] B. Grünbaum, *Configurations of Points and Lines*, volume 103 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, Rhode Island, 2009, doi:10.1090/gsm/103.

[4] T. Pisanski and B. Servatius, *Configurations from a Graphical Viewpoint*, Birkhäuser Advanced Texts, Birkhäuser, New York, 2013, doi:10.1007/978-0-8176-8364-1.

[5] E. Steinitz, *Über die Konstruction der Configurationen* n_3, Ph.D. thesis, Universität Breslau, 1894.

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 25–35
https://doi.org/10.26493/1855-3974.2028.cb4
(Also available at http://amc-journal.eu)

Results on the domination number and the total domination number of Lucas cubes*

Zülfükar Saygı
Department of Mathematics, TOBB University of Economics and Technology, Turkey

Received 27 June 2019, accepted 2 July 2020, published online 28 October 2020

Abstract

Lucas cubes are special subgraphs of Fibonacci cubes. For small dimensions, their domination numbers are obtained by direct search or integer linear programming. For larger dimensions some bounds on these numbers are given. In this work, we present the exact values of total domination number of small dimensional Lucas cubes and present optimization problems obtained from the degree information of Lucas cubes, whose solutions give better lower bounds on the domination numbers and total domination numbers of Lucas cubes.

Keywords: Lucas cube, Fibonacci cube, domination number, total domination number, integer linear programming.

Math. Subj. Class. (2020): 05C69, 68R10, 11B39

1 Introduction

Fibonacci cubes and Lucas cubes are special subgraphs of the hypercube graph, which are introduced as an alternative model for interconnection networks [6, 11]. The structural and enumerative properties of these graphs are considered from various point of view in the literature [2, 6, 7, 8, 9, 10, 11, 15].

Let Q_n denote the hypercube of dimension $n \geq 1$. It is the graph with vertex set represented by all binary strings of length n and two vertices in Q_n are adjacent if they differ in one coordinate. For convenience $Q_0 = K_1$. *Fibonacci strings* of length n are defined as the binary strings $b_1 b_2 \dots b_n$ such that $b_i \cdot b_{i+1} = 0$ for all $i = 0, 1, \dots, n-1$, that is, binary strings of length n not containing two consecutive 1s. Using this representation n dimensional *Fibonacci cube* Γ_n is defined as the subgraph of Q_n induced by the vertices

*The author is grateful to the anonymous reviewers for their valuable comments and careful reading and would like to thank Ömer Eğecioğlu and Elif Saygı for useful discussions and suggestions.

E-mail address: zsaygi@etu.edu.tr (Zülfükar Saygı)

whose string representations are Fibonacci strings. *Lucas strings* of length n are defined as the Fibonacci strings $b_1 b_2 \dots b_n$ such that $b_1 \cdot b_n = 0$. Similar to the Fibonacci cubes n dimensional *Lucas cube* Λ_n is defined as the subgraph of Γ_n induced by the vertices whose string representations are Lucas strings.

Let $G = (V, E)$ be a graph with vertex set V and edge set E. $D \subseteq V$ is called a *dominating set* of G if every vertex in V either belongs to D or is adjacent to some vertex in D. Then the *domination number* $\gamma(G)$ of G is defined as the minimum cardinality of a dominating set of the graph G. Similarly, $D \subseteq V$ is called a *total dominating set* of a graph G without isolated vertex if every vertex in V is adjacent to some vertex in D and the *total domination number* $\gamma_t(G)$ of G is defined as the minimum cardinality of a total dominating set of G. The domination numbers of Γ_n and Λ_n are first considered in [2, 12]. Using integer linear programming, domination numbers of Γ_n and Λ_n are considered in [7] and total domination number of Γ_n is considered in [1]. Furthermore, upper bounds and lower bounds on $\gamma(\Gamma_n), \gamma_t(\Gamma_n), \gamma(\Lambda_n)$ are obtained in [1, 2, 13] and they are improved for Γ_n in [14].

In this work, we present optimization problems obtained from the degree information of Lucas cubes, whose solutions give better lower bounds on the domination numbers and total domination numbers of Lucas cubes. Our aim is to improve the known results on $\gamma(\Lambda_n)$ and present new results on $\gamma_t(\Lambda_n)$. Furthermore, we introduce the up-down degree polynomials for Λ_n containing the degree information of all vertices $V(\Lambda_n)$ in more detail. Using these polynomials we define optimization problems whose solutions give lower bound on $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$.

2 Preliminaries

For $n \geq 2$ we will use the fundamental decomposition of Γ_n (see, [8]):

$$\Gamma_n = 0\Gamma_{n-1} + 10\Gamma_{n-2}, \tag{2.1}$$

where $\Gamma_0 = Q_0$ and $\Gamma_1 = Q_1$. Here note that $0\Gamma_{n-1}$ is the subgraph of Γ_n induced by the vertices that start with 0 and Γ_{n-2} is the subgraph of Γ_n induced by the vertices that start with 10. Furthermore, $0\Gamma_{n-1}$ has a subgraph isomorphic to $00\Gamma_{n-2}$, and there exists a perfect matching between $00\Gamma_{n-2}$ and $10\Gamma_{n-2}$. Similar to this decomposition for $n \geq 3$ Lucas cubes can be written as

$$\Lambda_n = 0\Gamma_{n-1} + 10\Gamma_{n-3}0, \tag{2.2}$$

where $10\Gamma_{n-3}0$ is the subgraph of Λ_n induced by the vertices that start with 10 and end with 0. Here, there exists a perfect matching between $10\Gamma_{n-3}0$ and $00\Gamma_{n-3}0 \subset 0\Gamma_{n-1}$. By convention, $\Lambda_1 = \Gamma_0$ and $\Lambda_2 = \Gamma_2$.

The number of vertices of the Γ_n is f_{n+2}, where f_n are the Fibonacci numbers defined as $f_0 = 0$, $f_1 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$. Similarly, the number of vertices of the Λ_n is L_n, where L_n are the Lucas numbers defined as $L_0 = 2$, $L_1 = 1$ and $L_n = L_{n-1} + L_{n-2}$ for $n \geq 2$.

Let x, y be two binary strings of length n. Then the *Hamming distance* between x and y, $d_H(x, y)$ is the number of coordinates in which they differ. The *Hamming weight* of x, $w(x)$ is the number of nonzero coordinates in x. Note that Hamming distance is the usual graph distance in Q_n.

In Figure 1 we present small dimensional Lucas cubes and a minimal total dominating set with circled vertices for $2 \leq n \leq 5$.

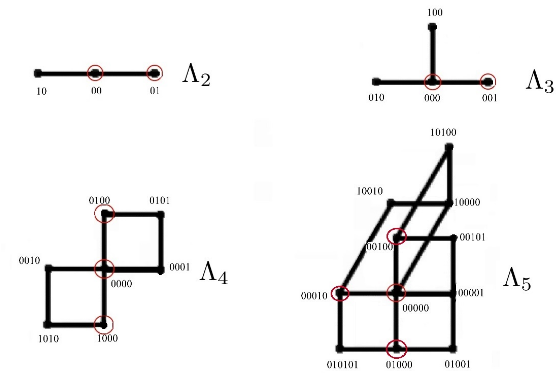


Figure 1: Lucas cubes and their minimal total dominating sets for $2 \leq n \leq 5$.

3 Integer linear programming for domination numbers

In this section, we describe a linear programming problem used in [7] for finding the domination number of Γ_n and Λ_n. A similar approach is used in [1] for finding the total domination number of Γ_n. The main difficulty for these methods are the number of variables and the number of constraints which are equal to the number of vertices in Γ_n and Λ_n. Using this approach we obtain the total domination number of Λ_n for $n \leq 12$.

Let $N(v)$ denote the set of vertices adjacent to v and $N[v] = N(v) \cup \{v\}$. Suppose each vertex $v \in V(\Lambda_n)$ is associated with a binary variable x_v. The problems of determining $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$ can be expressed as a problem of minimizing the objective function

$$\sum_{v \in V(\Lambda_n)} x_v \tag{3.1}$$

subject to the following constraints for every $v \in V(\Lambda_n)$:

$$\sum_{a \in N[v]} x_a \geq 1 \quad \text{(for domination number)},$$

$$\sum_{a \in N(v)} x_a \geq 1 \quad \text{(for total domination number)}.$$

The value of the objective function gives $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$ respectively. Note that this problem has L_n variables and L_n constraints. In [7] $\gamma(\Lambda_n)$ is obtained up to $n = 11$ and for larger values of n as the number of variables increases no results are presented.

We implemented the integer linear programming problem (3.1) using CPLEX in NEOS Server [3, 4, 5] for $n \leq 12$ and obtain the values of $\gamma_t(\Lambda_n)$ for $n \leq 12$ and obtain the estimates $49 \leq \gamma(\Lambda_{12}) \leq 54$ (takes approximately 2 hours). We collect the known values of $\gamma(\Lambda_n)$ for $n \leq 11$ (see, [7]) and the new values of $\gamma_t(\Lambda_n)$ that we obtained from (3.1) for $n \leq 12$ in Table 1.

The fundamental decompositions (2.1) and (2.2) of Γ_n and Λ_n are used to obtain the following relations between $\gamma(\Lambda_n)$ and $\gamma(\Gamma_n)$. The main idea in the proof is to partition the set of vertices into disjoint subsets.

Table 1: Values of $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$ for $n \leq 12$.

n	2	3	4	5	6	7	8	9	10	11	12
$\lvert V(\Lambda_n)\rvert$	3	4	7	11	18	29	47	76	123	199	322
$\gamma(\Lambda_n)$	1	1	3	4	5	7	11	16	23	35	
$\gamma_t(\Lambda_n)$	2	2	3	4	7	9	13	19	27	41	58

Proposition 3.1 ([2, Proposition 3.1])**.** *Let $n \geq 4$, then*

(i) $\gamma(\Lambda_n) \leq \gamma(\Gamma_{n-1}) + \gamma(\Gamma_{n-3})$,

(ii) $\gamma(\Lambda_n) \leq \gamma(\Gamma_n) \leq \gamma(\Lambda_n) + \gamma(\Gamma_{n-4})$.

Using a similar idea we obtain the following result.

Proposition 3.2. *Let $n \geq 4$, then*

(i) $\gamma_t(\Lambda_n) \leq \gamma_t(\Gamma_{n-1}) + \gamma_t(\Gamma_{n-3})$,

(ii) $\gamma_t(\Lambda_n) \leq \gamma_t(\Gamma_n) \leq \gamma_t(\Lambda_n) + \gamma_t(\Gamma_{n-4})$.

Proof. The proof mimics the proof of [2, Proposition 3.1].

(i): The vertices of Λ_n can be partitioned into vertices that start with 0 and vertices that start with 1. The subgraphs induced by these vertices are isomorphic to Γ_{n-1} and Γ_{n-3} respectively, hence we infer that $\gamma_t(\Lambda_n) \leq \gamma_t(\Gamma_{n-1}) + \gamma_t(\Gamma_{n-3})$.

(ii): Let D_T be a minimal total dominating set of Γ_n and set

$$D'_T = \{\alpha \mid \alpha \text{ is a Lucas string from } D_T\} \cup \{0\alpha 0 \mid 1\alpha 1 \in D_T\}.$$

Note that $|D'_T| \leq |D_T|$ and a vertex of the form $1\alpha 1$ dominates two Lucas vertices of the form $0\alpha 1$ and $1\alpha 0$. Since these two vertices are dominated by $0\alpha 0$, we say that D'_T is a dominating set of Λ_n. Then we need to show that it is also a total dominating set. We know that every vertex in $v \in V(\Lambda_n) \subseteq V(\Gamma_n)$ is adjacent to some vertex $\beta \in D_T$. Then if $\beta \in D'_T$ we are done. Otherwise, β must be of the form $1\alpha 1 \in D_T$. In this case $v \in V(\Lambda_n)$ must be of the form $1\alpha 0$ or $0\alpha 1$, which means that v is also adjacent to a vertex of the form $0\alpha 0 \in D'_T$. It follows that $\gamma_t(\Lambda_n) \leq \gamma_t(\Gamma_n)$. On the other hand, a total dominating set of Λ_n dominates all vertices of Γ_n but the vertices of the form $10\alpha 01$ where the subgraph induced by these vertices is isomorphic to Γ_{n-4}. Hence we have $\gamma_t(\Gamma_n) \leq \gamma_t(\Lambda_n) + \gamma_t(\Gamma_{n-4})$. □

Considering the vertices of high degrees a lower bound on $\gamma(\Lambda_n)$ is obtained in [2, Theorem 3.5] as $\gamma(\Lambda_n) \geq \lceil \frac{L_n - 2n}{n-3} \rceil$ where $n \geq 7$. Combining this result with the fact that $\gamma_t(\Lambda_n) \geq \gamma(\Lambda_n)$ we have the following lower bound on $\gamma_t(\Lambda_n)$.

Proposition 3.3. *For any $n \geq 7$, we have*

$$\gamma_t(\Lambda_n) \geq \gamma(\Lambda_n) \geq \left\lceil \frac{L_n - 2n}{n-3} \right\rceil.$$

4 Up-down degree enumerator polynomial

In this section we present the up-down degree enumerator polynomial for Λ_n similar to the one for Γ_n given in [14]. Using this polynomial we write optimization problems whose solutions give lower bounds on $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$.

By the definition of the edge set $E(\Lambda_n)$, $(v, v') \in E(\Lambda_n)$ if and only if the number of different coordinates of v and v' is 1, that is, the Hamming distance $d_H(v, v') = 1$. Here we have at most two kinds of neighbor v' for a vertex in $v \in V(\Lambda_n)$, whose weights can take the values $w(v) \pm 1$. If $w(v') = w(v) + 1$ we call v' is an *up neighbor* of v and if $w(v') = w(v) - 1$ we call v' is a *down neighbor* of v. We denote the number of up neighbors of v by u and the number of down neighbors of v by d which is equal to the $w(v)$ by the definition of Λ_n. Note that if the degree of v is k then we have $u = k - d$. For each fixed $v \in V(\Lambda_n)$ having degree $k = \deg(v)$, we write a monomial $x^u y^d$ where $d = w(v)$ is the Hamming weight of v and u is $k - d$. We call the polynomial

$$P_{\Lambda_n}(x, y) = \sum_{v \in V(\Lambda_n)} x^{\deg(v) - w(v)} y^{w(v)} = \sum_{v \in V(\Lambda_n)} x^u y^d$$

as the *up-down degree enumerator polynomial* of Λ_n.

We need the following useful result given in [10] to obtain the recursive structure of $P_{\Lambda_n}(x, y)$. Let $\ell_{n,k,w}$ be the number of vertices in Λ_n of degree k and weight w.

Theorem 4.1 ([10, Theorem 5.2])**.** *For all n, k, w such that $n \geq 2$, $1 \leq k \leq n$ and $0 \leq w \leq n$,*

$$\ell_{n,k,w} = \binom{w-1}{2w+k-n}\binom{n-2w}{k-w} + 2\binom{w}{2w+k-n}\binom{n-2w-1}{k-w}.$$

Let $\ell'_{n,u,d}$ be the number of vertices in Λ_n whose number of up neighbors are u and number of down neighbors are d. Setting $k = u + d$ and $w = d$ in Theorem 4.1 we have

$$\ell'_{n,u,d} = \binom{d-1}{3d+u-n}\binom{n-2d}{u} + 2\binom{d}{3d+u-n}\binom{n-2d-1}{u}. \tag{4.1}$$

Then using (4.1) we can write the up-down degree enumerator polynomial of Λ_n as

$$P_{\Lambda_n}(x, y) = \sum_{u,d} \ell'_{n,u,d}\, x^u y^d, \tag{4.2}$$

where $0 \leq u, d \leq n$. Furthermore, using (4.1) and (4.2) we obtain the following recursive relation which is very useful to calculate $P_{\Lambda_n}(x, y)$.

Theorem 4.2. *Let $P_{\Lambda_n}(x, y)$ be the up-down degree enumerator polynomial of Λ_n. Then for $n \geq 5$ we have*

$$P_{\Lambda_n}(x, y) = xP_{\Lambda_{n-1}}(x, y) + yP_{\Lambda_{n-2}}(x, y) + (y - xy)P_{\Lambda_{n-3}}(x, y), \tag{4.3}$$

where

$$P_{\Lambda_2}(x, y) = x^2 + 2y, \quad P_{\Lambda_3}(x, y) = x^3 + 3y \quad \textit{and} \quad P_{\Lambda_4}(x, y) = x^4 + 4xy + 2y^2.$$

Proof. The initial conditions are clear from the definition of Λ_n. For fixed integers $1 \leq u < n$ and $2 \leq d < \lfloor \frac{n}{2} \rfloor$, the coefficient of the monomial $x^u y^d$ in the right hand side of the equation (4.3) is the sum of $\ell'_{n-1,u-1,d}$ coming from $xP_{\Lambda_{n-1}}(x,y)$, $\ell'_{n-2,u,d-1}$ coming from $yP_{\Lambda_{n-2}}(x,y)$, $\ell'_{n-3,u,d-1}$ coming from $yP_{\Lambda_{n-3}}(x,y)$ and $-\ell'_{n-3,u-1,d-1}$ coming from $-xyP_{\Lambda_{n-3}}(x,y)$. Then we need to show that

$$\ell'_{n,u,d} = \ell'_{n-1,u-1,d} + \ell'_{n-2,u,d-1} + \ell'_{n-3,u,d-1} - \ell'_{n-3,u-1,d-1}.$$

By setting $X = 3d + u - n$ and $Y = n - 2d$ in (4.1) and using the binomial identities

$$\binom{m}{k} = \frac{m}{k}\binom{m-1}{k-1} = \frac{m+1-k}{k}\binom{m}{k-1} = \frac{m}{m-k}\binom{m-1}{k}$$

we have

$$\begin{aligned}
&\ell'_{n-1,u-1,d} + \ell'_{n-2,u,d-1} + \ell'_{n-3,u,d-1} - \ell'_{n-3,u-1,d-1} \\
&\quad = \binom{d-1}{X}\binom{Y-1}{u-1} + 2\binom{d}{X}\binom{Y-2}{u-1} \\
&\qquad + \binom{d-2}{X-1}\binom{Y}{u} + 2\binom{d-1}{X-1}\binom{Y-1}{u} \\
&\qquad + \binom{d-2}{X}\binom{Y-1}{u} + 2\binom{d-1}{X}\binom{Y-2}{u} \\
&\qquad - \binom{d-2}{X-1}\binom{Y-1}{u-1} + 2\binom{d-1}{X-1}\binom{Y-2}{u-1} \\
&\quad = \binom{d-1}{X}\binom{Y}{u}\left[\frac{u}{Y} + \frac{X}{d-1} + \frac{d-1-X}{d-1}\cdot\frac{Y-u}{Y} - \frac{X}{d-1}\cdot\frac{u}{Y}\right] \\
&\qquad + 2\binom{d}{X}\binom{Y-1}{u}\left[\frac{u}{Y-1} + \frac{X}{d} + \frac{d-X}{d}\cdot\frac{Y-1-u}{Y-1} - \frac{X}{d}\cdot\frac{u}{Y-1}\right] \\
&\quad = \binom{d-1}{X}\binom{Y}{u} + 2\binom{d}{X}\binom{Y-1}{u} \\
&\quad = \ell'_{n,u,d}.
\end{aligned}$$

In particular, the case $d = 0$ corresponds to the all 0 vertex in Λ_n and we have $\ell'_{n,n,0} = 1$, which means that the coefficient of the terms $x^n y^0$ in both sides of (4.3) are 1. Similarly, the case $u = 0$ corresponds to the vertices in Λ_n whose weights are $\lfloor \frac{n}{2} \rfloor$ and we have (see, Remark 5.1)

$$\ell'_{n,0,\lfloor \frac{n}{2} \rfloor} = \begin{cases} n & \text{if } n \text{ is odd,} \\ 2 & \text{if } n \text{ is even.} \end{cases}$$

Then one can easily see that the coefficient of the terms $x^0 y^{\lfloor \frac{n}{2} \rfloor}$ in both sides of (4.3) are equal to each other. The only remaining particular case is $d = 1$. For $P_{\Lambda_n}(x,y)$ this case corresponds to the vertices in Λ_n whose weights are 1. We know that there are n such vertices in Λ_n and their number of up neighbors are $n-3$. That is, the coefficient of the term $x^{n-3}y$ in $P_{\Lambda_n}(x,y)$ is n. On the other hand the coefficient of the term $x^{n-3}y$ is $n-1$ in $xP_{\Lambda_{n-1}}(x,y)$; 0 in $yP_{\Lambda_{n-2}}(x,y)$ and 1 in $(y-xy)P_{\Lambda_{n-3}}(x,y)$ respectively. Hence the coefficient of the terms $x^u y^d$ in both sides of (4.3) are equal to each other for all cases. □

Remark 4.3. The recursive relation for the up-down degree enumerator polynomial of Λ_n in Theorem 4.2 is the same with the recursive relation for the up-down degree enumerator polynomial of Γ_n, which is proved using the fundamental decomposition $\Gamma_n = 0\Gamma_{n-1} + 10\Gamma_{n-2}$. The only differences are the initial polynomials. For the proof we directly used the degree information of Λ_n obtained in [10], since Λ_n do not have a decomposition like $0\Lambda_{n-1} + 10\Lambda_{n-2}$.

5 Lower bounds on domination numbers using optimization problems

In this section, we present optimization problems giving lower bounds on $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$, whose number of variables and number of constraints are fewer than the general optimization problem described in Section 3.

We use the up-down degree enumerator polynomial $P_{\Lambda_n}(x, y)$ to construct an optimization problem, which is similar to the optimization problem given in [14]. Let D and D_T be a dominating set and a total dominating set of Λ_n respectively. Let $v_D \in D$ $\big(v_D \in D_T$ respectively$\big)$ and $x^u y^d$ be its corresponding monomial in $P_{\Lambda_n}(x, y)$. Then v_D dominates u distinct vertices $v \in V(\Lambda_n)$ having weight $w(v) = w(v_D) + 1$ and d distinct vertices $v \in V(\Lambda_n)$ having weight $w(v) = w(v_D) - 1$. Note that for all $v_D \in D$ $\big(v_D \in D_T$ respectively$\big)$ some of the vertices of Λ_n may be dominated more than one times. Note that for every vertex $v \in V(\Lambda_n)$ there must exist at least one vertex $v_D \in N[v] \cap D_T$ with $w(v_D) = w(v) \mp 1$ or $v_D = v$ for the dominating set D and $v_D \in N(v) \cap D_T$ with $w(v_D) = w(v) \mp 1$ for the total dominating set D_T.

Now we write the up-down degree enumerator polynomial of Λ_n (see, 4.2) as

$$P_{\Lambda_n}(x, y) = \sum_{u,d} c_d^u x^u y^d, \tag{5.1}$$

where $c_d^u = \ell'_{n,u,d}$. For each pair (u, d) in the monomials of the up-down degree enumerator polynomial $P_{\Lambda_n}(x, y)$ we associate an integer variable z_d^u which counts the number of vertices in D or D_T having d down neighbors and u up neighbors. For any fixed value of d, the number of vertices having weight d gives the bounds $0 \le z_d^u \le c_d^u$. Our aim is to minimize $|D|$ for domination number and to minimize $|D_T|$ for total domination number. Hence our objective function is to minimize

$$\sum_{u,d} z_d^u.$$

To dominate all the vertices having a fixed weight d such that $1 \le d \le \lfloor \frac{n}{2} \rfloor - 1$ we must have the following constraints r_d for domination number and r'_d for the total domination number.

$$r_d\colon \qquad \sum_u \big(u \cdot z_{d-1}^u + z_d^u + (d+1) \cdot z_{d+1}^u\big) \ge \sum_u c_d^u$$

$$r'_d\colon \qquad \sum_u \big(u \cdot z_{d-1}^u + (d+1) \cdot z_{d+1}^u\big) \ge \sum_u c_d^u$$

since any vertex corresponding to the monomial $x^u y^{d-1}$ can dominate u distinct vertices (u up neighbors) having weight d and any vertex corresponding to the monomial $x^{u'} y^{d+1}$

can dominate $d+1$ distinct vertices ($d+1$ down neighbors) having weight d. By the same argument, for $d=0$ we must have

$$r_0\colon \quad \sum_u z_0^u + z_1^u \geq \sum_u c_0^u = 1 \qquad \text{and} \qquad r'_0\colon \quad \sum_u z_1^u \geq \sum_u c_0^u$$

and for $d = \lfloor \frac{n}{2} \rfloor$ we must have

$$r_{\lfloor \frac{n}{2} \rfloor}\colon \qquad \sum_u u \cdot z^u_{\lfloor \frac{n}{2} \rfloor - 1} + z^u_{\lfloor \frac{n}{2} \rfloor} \geq \sum_u c^u_{\lfloor \frac{n}{2} \rfloor} = \begin{cases} n & \text{if } n \text{ is odd,} \\ 2 & \text{if } n \text{ is even.} \end{cases}$$

$$r'_{\lfloor \frac{n}{2} \rfloor}\colon \qquad \sum_u u \cdot z^u_{\lfloor \frac{n}{2} \rfloor - 1} \geq \sum_u c^u_{\lfloor \frac{n}{2} \rfloor} = \begin{cases} n & \text{if } n \text{ is odd,} \\ 2 & \text{if } n \text{ is even.} \end{cases}$$

Now subject to the above constraints $r_0, \dots, r_{\lfloor \frac{n}{2} \rfloor}$ (constraints $r'_0, \dots, r'_{\lfloor \frac{n}{2} \rfloor}$) the value of the objective function will be a lower bound on $\gamma(\Lambda_n)$ ($\gamma_t(\Lambda_n)$, respectively).

Remark 5.1. The number of vertices of Λ_n having weight d is equal to the right hand side of the above constraints r_d and r'_d. By setting $k = u + d$ and $w = d$ in [10, Corollary 5.3] we have

$$\sum_u c_d^u = \sum_{u=0}^{n-d} \ell'_{n,u,d} = \binom{n-d}{d} + \binom{n-d-1}{n-2d}.$$

Remark 5.2. The number of variables z_d^u in our optimization problem is equal to the number of monomials in $P_{\Lambda_n}(x, y)$. Assume that n is even. By the string representation of the vertices in Λ_n we have $n-3d \leq u \leq n-2d-1$. The bounds come from the maximum number of the sub-strings 010 and 10 in the representation of the vertices. That is, u can take $n-2d-1-(n-3d)+1 = d$ distinct values when d ranges from 1 up to $\lfloor \frac{n}{3} \rfloor$ and u can take $n-2d$ distinct values when $\frac{n}{3}+1 \leq d < \lfloor \frac{n}{2} \rfloor$. Furthermore, u can take only one values for $d=0$ and $d = \lfloor \frac{n}{2} \rfloor$. Therefore, the number of variables z_d^u becomes

$$2 + \sum_{d=1}^{\lfloor \frac{n}{3} \rfloor} d + \sum_{d=\lfloor \frac{n}{3} \rfloor + 1}^{\lfloor \frac{n}{2} \rfloor - 1} (n-2d)$$

which is equal to

$$2 + \frac{3}{2} \left\lfloor \frac{n}{3} \right\rfloor \left(\left\lfloor \frac{n}{3} \right\rfloor + 1 - \frac{2n}{3} \right) + \left\lfloor \frac{n}{2} \right\rfloor \left(n - \left\lfloor \frac{n}{2} \right\rfloor + 1 \right) - n. \tag{5.2}$$

For $n \geq 2$ this sequence starts as $2, 2, 3, 3, 5, 5, 7, 8, 10, 11, 14, 15, 18, 20, 23, 25, 29, \dots$ Note that in (3.1) the number of variables is L_n, which exhibit exponential growth. In our case, if we omit the floor functions in (5.2) then the number of variables z_d^u is approximately equals to $2 + \frac{n^2}{12}$.

For $n = 12$ we illustrate our optimization problem as follows. First we obtain $P_{\Lambda_{12}}(x, y)$

by using the recursion in Theorem 4.2 as

$$\begin{aligned} P_{\Lambda_{12}}(x,y) = {} & 2y^6+\\ & 12y^5x + 24y^5+\\ & 12y^4x^3 + 54y^4x^2 + 36y^4x + 3y^4+\\ & 12y^3x^5 + 60y^3x^4 + 40y^3x^3+\\ & 12y^2x^7 + 42y^2x^6+\\ & 12yx^9+\\ & x^{12} \end{aligned}$$

Then using $P_{\Lambda_{12}}(x,y)$ we have the corresponding optimization problem:

Objective function:

$$minimize: \; z_0^{12} + z_1^9 + z_2^7 + z_2^6 + z_3^5 + z_3^4 + z_3^3 + z_4^3 + z_4^2 + z_4^1 + z_4^0 + z_5^1 + z_5^0 + z_6^0;$$

Constraints for $\gamma(\Lambda_{12})$:

$$\begin{aligned} r_6: && z_5^1 + z_6^0 \geq 2;\\ r_5: && 3z_4^3 + 2z_4^2 + z_4^1 + z_5^1 + z_5^0 + 6z_6^0 \geq 36;\\ r_4: && 5z_3^5 + 4z_3^4 + 3z_3^3 + z_4^3 + z_4^2 + z_4^1 + z_4^0 + 5z_5^1 + 5z_5^0 \geq 105;\\ r_3: && 7z_2^7 + 6z_2^6 + z_3^5 + z_3^4 + z_3^2 + 4z_4^3 + 4z_4^2 + 4z_4^1 + 4z_4^0 \geq 112;\\ r_2: && 9z_1^9 + z_2^7 + z_2^6 + 3z_3^5 + 3z_3^4 + 3z_3^3 \geq 54;\\ r_1: && 12z_0^{12} + z_1^9 + 2z_2^7 + 2z_2^6 \geq 12;\\ r_0: && z_0^{12} + z_1^9 \geq 1; \end{aligned}$$

Constraints for $\gamma_t(\Lambda_{12})$:

$$\begin{aligned} r'_6: && z_5^1 \geq 2;\\ r'_5: && 3z_4^3 + 2z_4^2 + z_4^1 + 6z_6^0 \geq 36;\\ r'_4: && 5z_3^5 + 4z_3^4 + 3z_3^3 + 5z_5^1 + 5z_5^0 \geq 105;\\ r'_3: && 7z_2^7 + 6z_2^6 + 4z_4^3 + 4z_4^2 + 4z_4^1 + 4z_4^0 \geq 112;\\ r'_2: && 9z_1^9 + 3z_3^5 + 3z_3^4 + 3z_3^3 \geq 54;\\ r'_1: && 12z_0^{12} + 2z_2^7 + 2z_2^6 \geq 12;\\ r'_0: && z_1^9 \geq 1; \end{aligned}$$

Bounds:

$$\begin{array}{ccccccc} z_0^{12} \leq 1; & z_1^9 \leq 12; & z_2^7 \leq 12; & z_2^6 \leq 42; & z_3^5 \leq 12; & z_3^4 \leq 60; & z_3^3 \leq 40;\\ z_4^3 \leq 12; & z_4^2 \leq 54; & z_4^1 \leq 36; & z_4^0 \leq 3; & z_5^1 \leq 12; & z_5^0 \leq 24; & z_6^0 \leq 2. \end{array}$$

Depending on the constraints r_d and r'_d ($d = 0, 1, \ldots, 6$) the value of the objective function gives a lower bound on $\gamma(\Lambda_{12})$ and $\gamma_t(\Lambda_{12})$ respectively. The above problem has only 14 variables and 7 constraints (instead of having $L_{12} = 322$ variables and 322 constraints as in (3.1)). To find lower bounds on $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$ one can use the up-down degree enumerator polynomial $P_{\Lambda_n}(x, y)$ of Λ_n in Theorem 4.2 and one can write an optimization problem having fewer number of variables z_d^u (see Remark 5.2) and $\lfloor \frac{n}{2} \rfloor + 1$ constraints r_d or r'_d. The solutions of the optimization problems give lower bounds on $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$. It is easy to see that the number of variables and the number of constraints in our optimization problems are very smaller than the ones in the optimization problem (3.1).

For illustration we implemented the above integer linear programming problem using CPLEX in NEOS Server [3, 4, 5] for $12 \leq n \leq 26$ and immediately (less than 0.02 seconds) obtain the lower bounds on $\gamma(\Lambda_n)$ and $\gamma_t(\Lambda_n)$ presented in Table 2 and Table 3 (better than the ones in Proposition 3.3). Note that for $n = 26$, the number of variables in our optimization problem is 58 by Remark 5.2 and the number of constraints is 14, on the other hand, these numbers are equal to $L_{26} = 271443$ for the general optimization problem (3.1). In addition, the upper bounds in these tables are obtained by Proposition 3.1 and Proposition 3.2 by using the upper bounds on the values of $\gamma(\Gamma_n)$ and $\gamma_t(\Gamma_n)$ given in [14] for $n \geq 14$.

Table 2: Current best bounds on $\gamma(\Lambda_n)$, $12 \leq n \leq 26$.

n	$\gamma(\Lambda_n)$	n	$\gamma(\Lambda_n)$	n	$\gamma(\Lambda_n)$
12	49^*-54	17	$310-555$	22	$2686-6140$
13	61^*-86	18	$471-895$	23	$4184-9935$
14	$89-132$	19	$725-1450$	24	$6519-16075$
15	$134-215$	20	$1114-2345$	25	$10163-26010$
16	$203-340$	21	$1724-3795$	26	$15835-42085$

Table 3: Current best bounds on $\gamma_t(\Lambda_n)$, $12 \leq n \leq 26$.

n	$\gamma_t(\Lambda_n)$	n	$\gamma_t(\Lambda_n)$	n	$\gamma_t(\Lambda_n)$
12	$\mathbf{58}^*$	17	$340-567$	22	$2893-6140$
13	77^*-95	18	$514-909$	23	$4490-9935$
14	$101-145$	19	$787-1450$	24	$6974-16075$
15	$151-231$	20	$1205-2345$	25	$10839-26010$
16	$225-362$	21	$1862-3795$	26	$16838-42085$

Remark 5.3. It is shown in [1, 7] that $\gamma(\Gamma_9) = 17$, $\gamma(\Gamma_{10}) = 25$, $54 \leq \gamma(\Gamma_{12}) \leq 61$ and $78 \leq \gamma(\Gamma_{13}) \leq 93$ (shown in [14]). Substituting these results in Proposition 3.1 we obtain the bounds for $n = 13$ in Table 2.

Similarly, it is shown in [1, 7] that $\gamma_t(\Gamma_9) = 20$, $\gamma_t(\Gamma_{10}) = 30$, $\gamma_t(\Gamma_{12}) = 65$ and $97 \leq \gamma_t(\Gamma_{13}) \leq 101$. Substituting these results in Proposition 3.2 we obtain the bounds for $n = 13$ in Table 3.

Note that our optimization problems obtained from up-down degree enumerator polynomial give $\gamma(\Lambda_{12}) \geq 39$, $\gamma_t(\Lambda_{12}) \geq 45$ and $\gamma(\Lambda_{13}) \geq 59$, $\gamma_t(\Lambda_{13}) \geq 68$. Furthermore, using (3.1) we obtain that $49 \leq \gamma(\Lambda_{12}) \leq 54$ and $\gamma_t(\Lambda_{12}) = 58$. For these reasons we put a * to the lower bounds for the cases $n = 12$ and $n = 13$ in Table 2 and Table 3.

ORCID iD

Zülfükar Saygı https://orcid.org/0000-0002-7575-3272

References

[1] J. Azarija, S. Klavžar, Y. Rho and S. Sim, On domination-type invariants of Fibonacci cubes and hypercubes, *Ars Math. Contemp.* **14** (2018), 387–395, doi:10.26493/1855-3974.1172.bae.

[2] A. Castro, S. Klavžar, M. Mollard and Y. Rho, On the domination number and the 2-packing number of Fibonacci cubes and Lucas cubes, *Comput. Math. Appl.* **61** (2011), 2655–2660, doi:10.1016/j.camwa.2011.03.012.

[3] J. Czyzyk, M. P. Mesnier and J. J. Moré, The NEOS server, *IEEE J. Comput. Sci. Eng.* **5** (1998), 68–75, doi:10.1109/99.714603.

[4] E. D. Dolan, *The NEOS Server 4.0 Administrative Guide*, Technical Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory, 2001.

[5] W. Gropp and J. J. Moré, Optimization environments and the NEOS server, in: M. D. Buhmann and A. Iserles (eds.), *Approximation Theory and Optimization*, Cambridge University Press, Cambridge, pp. 167–182, 1997, selected papers from the Conference on Numerical Mathematics, in honor of M. J. D. Powell on the occasion of his 60th birthday, held in Cambridge, July 27 – 30, 1996.

[6] W.-J. Hsu, Fibonacci cubes—A new interconnection topology, *IEEE Trans. Parallel Distrib. Syst.* **4** (1993), 3–12, doi:10.1109/71.205649.

[7] A. Ilić and M. Milošević, The parameters of Fibonacci and Lucas cubes, *Ars Math. Contemp.* **12** (2017), 25–29, doi:10.26493/1855-3974.915.f48.

[8] S. Klavžar, Structure of Fibonacci cubes: a survey, *J. Comb. Optim.* **25** (2013), 505–522, doi:10.1007/s10878-011-9433-z.

[9] S. Klavžar and M. Mollard, Cube polynomial of Fibonacci and Lucas cubes, *Acta Appl. Math.* **117** (2012), 93–105, doi:10.1007/s10440-011-9652-4.

[10] S. Klavžar, M. Mollard and M. Petkovšek, The degree sequence of Fibonacci and Lucas cubes, *Discrete Math.* **311** (2011), 1310–1322, doi:10.1016/j.disc.2011.03.019.

[11] E. Munarini, C. Perelli Cippo and N. Zagaglia Salvi, On the Lucas cubes, *Fibonacci Quart.* **39** (2001), 12–21, `https://www.fq.math.ca/Scanned/39-1/munarini.pdf`.

[12] D. A. Pike and Y. Zou, The domination number of Fibonacci cubes, *J. Combin. Math. Combin. Comput.* **80** (2012), 433–444.

[13] E. Saygı, Upper bounds on the domination and total domination number of Fibonacci cubes, *SDU J. Nat. Appl. Sci.* **21** (2017), 782–785, doi:10.19113/sdufbed.05851.

[14] E. Saygı, On the domination number and the total domination number of Fibonacci cubes, *Ars Math. Contemp.* **16** (2019), 245–255, doi:10.26493/1855-3974.1591.92e.

[15] E. Saygı and Ö. Eğecioğlu, q-counting hypercubes in Lucas cubes, *Turkish J. Math.* **42** (2018), 190–203, doi:10.3906/mat-1605-2.

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 37–49
https://doi.org/10.26493/1855-3974.2025.5d9
(Also available at http://amc-journal.eu)

Classification of virtual string links up to cobordism*

Robin Gaudreau

Department of Mathematics, University of Toronto, Toronto, Canada

Received 20 June 2019, accepted 26 February 2020, published online 29 October 2020

Abstract

Cobordism of virtual string links on n strands is a combinatorial generalization of link cobordism. There exists a bijection between virtual string links on n strands up to cobordisms and elements of the direct product of $n(n-1)$ copies of the integers. This paper also shows that virtual string links up to unwelded equivalence are classified by those groups. Finally, the related theory of welded string link cobordism is defined herein and shown to be trivial for string links with one component.

Keywords: Virtual links, string links, cobordism, concordance, welded knots.

Math. Subj. Class. (2020): 57K12, 57Q60

1 Introduction

Virtual knot theory, as understood from [9], is a combinatorial extension of classical knot theory. When picturing oriented knots as diagrams in the plane, crossings are vertices of a planar, oriented, tetra-valent graph with a cyclic orientation of the edges and a distinguished over-crossing pair. Removing the planarity requirement on such graphs yields virtual knot diagrams, whose equivalence classes up to the appropriate (generalized) Reidemeister moves are called virtual knots. Similarly, by understanding a classical string link as an equivalence class of diagrams, one defines a virtual string link. The goal of this paper is to relate three generalizations of concepts from classical knot theory through the following result:

*This work would not have been possible without the feedback and encouragement from Dror Bar-Natan and Hans Boden. The result in this paper was first presented at the Knots in Washington, and I am grateful to the organizers for creating such a motivating and creative atmosphere. I would also like to thank the reviewer for providing positive and useful feedback. This research was supported by funding from the National Science Research Council of Canada.

E-mail address: gaudreau@tutamail.com (Robin Gaudreau)

Theorem 1.1. *Let L_1 and L_2 be virtual string link diagrams. Then, the following are equivalent:*

(1) L_1 is cobordant to L_2;

(2) L_1 is unwelded equivalent to L_2;

(3) The pairwise virtual linking number of the components of L_1 equal those of L_2.

The first classification, string link cobordism, is a generalization of the notion of virtual link cobordism introduced by Carter, Kamada, and Saito in [5]. It relies on an interpretation of virtually knotted objects as curves in thickened surfaces, but yields the same theory as the one that is exposed below.

The second classification has been studied under many other names, notably as *fused isotopy* and the equivalence between statements (2) and (3) is a generalization of Theorem 2 of [7] and of Theorem 8 in [12]. This result appears as Theorem 4 in [10] (where the word *unrestricted* is synonymous to unwelded) and as Proposition 3.6 in [2]. Both of these papers offer view on unwelded equivalence which unfortunately lacks an intrinsic topological interpretation.

My new contribution is relating cobordance with the pairwise linking number. This paper is structured as follows: relevant definitions are given in Section 2, results and topological notions needed to prove the main theorem appear in Section 3, followed by its proof in Section 4. Finally, Section 5 is contains partial results on welded string links up to concordance.

2 Vocabulary

From now on, fix $n \geq 1$ to be an integer, and let $I = [0, 1]$ denote the closed unit interval.

2.1 Virtual string links

Classical string links were defined in [8] as a *self-concordance* of n points in $D^2 \times I$, where D^2 is the closed unit disk in the plane. This abstract and succinct definition contains all the details needed to understand these objects, but it does not allow a straightforward generalization to virtual string links.

Following the approach to virtual knot theory from [9], let a virtual string link diagram be a diagram consisting of n smooth curves, oriented from $(\frac{i}{n+1}, 0)$ to $(\frac{i}{n+1}, 1)$, with $i = 1, 2, \ldots, n$, such that singularities are at most a finite number of transverse double points, decorated in one of the ways depicted in Figure 1. The classical Reidemeister moves, as shown in Figure 2 can be applied to string link diagrams and generate the expected equivalence classes.

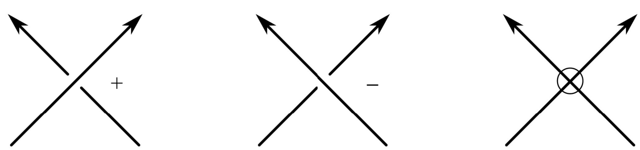


Figure 1: Positive, negative, and virtual crossings in planar diagrams.

Therefore, virtual string links, to be generalizations of virtual pure braids need to be defined combinatorially. While one could do this process using any knot presentation, the following will only use planar and Gauss diagrams. *Virtual string links* are then the equivalence class generated by such a diagram, up to the extended Reidemeister moves from [9] and planar isotopies.

Given such a planar diagram, one can create its associated Gauss diagram by drawing the n intervals and connecting the pre-images of a classical crossing by an arrow oriented from the overcrossing component to the undercrossing one, decorated by signs using the convention shown in Figure 1. The writhe function of a crossing c, $w(c)$, takes value $+1$ or -1 if c is positive or negative respectively. The writhe is not defined for virtual crossings.

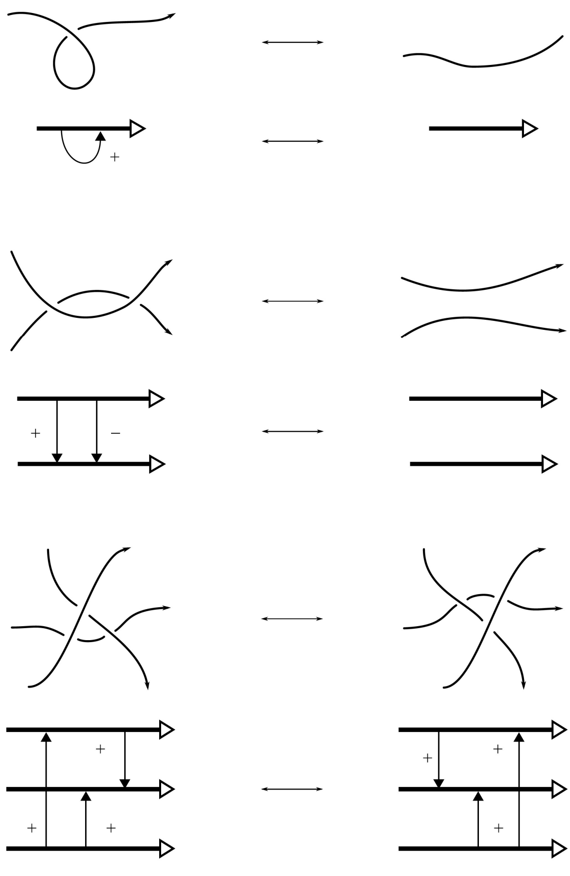


Figure 2: Reidemeister moves on planar and Gauss diagrams.

Alternatively, a virtual string link Gauss diagram can be constructed abstractly, by drawing a finite number of signed oriented chords with distinct endpoints on the interior of n intervals. As with virtual knots and links, it is immediate that any such Gauss diagram can be realized as a virtual string link planar diagram. Gauss diagrams admit their own version of Reidemeister moves, which are the same for classical and virtual SL_1 since Reidemeister moves which involve virtual crossings leave the Gauss diagram unchanged.

In this paper, “Reidemeister moves” is used to mean simultaneously the classical moves on planar diagrams, their extended version, and the analogous moves on Gauss diagrams. Figure 2 shows the equivalence between the two approaches, and therefore a virtual string link can be defined strictly from the Gauss diagrams.

2.2 Notation

Following [3], the set of classical string links on n strands is denoted uSL_n. Its virtual extension is vSL_n, while the welded version, $vSL_n/(f1)$ is wSL_n. Finally, $vSL_n/(f1, f2) =: uwSL_n$ are unwelded string links on n strands. The moves $f1$ and $f2$ are discussed in more details in Section 2.4.

Each of those sets is closed under an ordered, binary operation, concatenation. It is written using the $\#$ operator, following the notation for connected sum in uSL_1, which corresponds to classical long knots, and to carry on the analogy, diagrams are drawn such that the strands connect vertical intervals from left to right. Given two string link diagrams on n strands, D_1 and D_2, the string link $D_1 \# D_2$ is represented a diagram obtained by connected the end of the ith strand of D_1 to the beginning of the ith strand of D_2. For Gauss diagrams, the concatenation, as seen in Figure 11, is represented by joining the pre-images of the strands together.

Because moves can be applied to each part of $D_1 \# D_2$ independently, the result of concatenation is independent of the choice of diagrams. Moreover, the operation is associative, thus makes the sets of string links into monoids.

2.3 Virtual cobordims

A cobordism between two virtual knot diagrams K_0 and K_1 is a finite sequence of Reidemeister moves, births and deaths of closed unknotted components, and oriented saddle moves, as pictured in Figure 3. By *closed* unknotted component, we mean a closed circle that is an unknot as understood in the context of classical link theory. This restriction to the kind of components that can be created (during a birth) or removed (during a death) will be preserved for cobordisms of string link. Diagrammatic cobordism are generalized to string links from [5], with the added restriction that the abstract surface with corners described by the cobordism of an n-component virtual string link must have precisely n connected components. This corresponds to the topological restriction that is imposed on link cobordisms. As with other cases of cobordism, the genus can be computed by using the formula

$$(s_i - b_i + d_i)/2,$$

where s_i is the number of saddle moves in the cobordism that involve the ith component, b_i the number of birthed unknots that get saddled to it and d_i the number of deaths related to the component.

Given a virtual string link diagram D on n strands, the cobordism class it generates is $\mathcal{B}(D)$, and the set of all such classes is $vSL_n\mathcal{B}$. Similarly, the restriction to classical diagrams is $uSL_n\mathcal{B}$.

For classical knots, a cobordism between $K_0, K_1 \subset \mathbb{R}^3$ is called a concordance if it is realized by an annulus $S^1 \times [0, 1] \subset \mathbb{R}^3 \times I$ where its boundary component $S^1 \times \{i\}$ represents K_i. For long knots, the cobording surface is $\mathbb{R} \times [0, 1]$, and a simple truncated example is in Figure 4. The set of diagrams that are concordant to some classical string link D is $\mathcal{C}(D)$, an element $uSL_n\mathcal{C}$, the n-strand classical string link concordance group

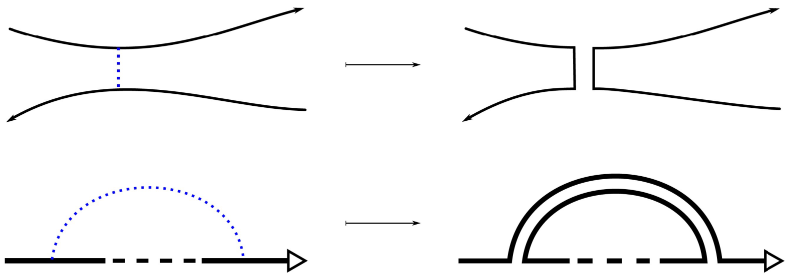


Figure 3: The saddle move on a planar and on a Gauss diagram.

(with the inverse of an planar diagram being its vertical mirror image). Using the abstract definition of genus above, a *concordance* between two virtual string link diagrams on n strands consists of a series of extended Reidemeister moves, births, deaths and saddles, such that the genus of the cobordism on each component is 0. The quotient by concordance of vSL_n is denoted $vSL_n\mathcal{C}$, and called the n-strand virtual string link concordance group (see Proposition 5.2). As with cobordism, all quotients of vSL_n can be factored by concordance equivalence, and there are many questions about the maps between those groups. It is known from [4] that $uSL_1\mathcal{C}$ embeds in $vSL_1\mathcal{C}$, but it is an open problem whether this continues to hold for $n > 1$ i.e. is the natural map $uSL_n\mathcal{C} \to vSL_n\mathcal{C}$ one-to-one?

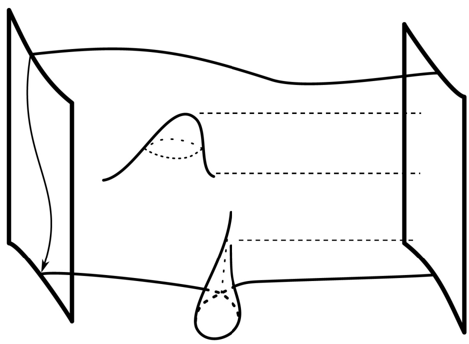


Figure 4: Concordance between a standard long unknot and one with a kink.

While round classical knots up to concordance form a group, round virtual knots do not have a well-defined concatenation, hence the appropriate virtual concordance group uses long virtual knots and agrees with $vSL_1\mathcal{C}$. This motivates the study of the problem above.

2.4 Forbidden moves

On planar diagrams, forbidden moves are the tempting operations that appear similar to a third Reidemeister move and would allow a strand to slide either over ($f1$) or under ($f2$) a virtual crossing. On Gauss diagrams, the difference between those operations and the other moves is more evident as the forbidden moves allow certain arrow endpoints to commute without compensating for it elsewhere in the link, as depicted in Figure 5.

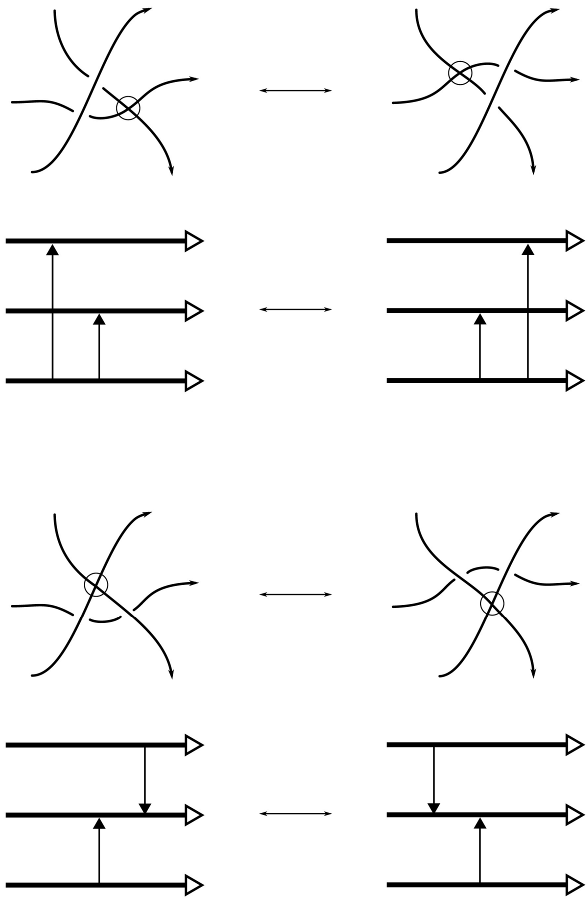


Figure 5: Forbidden moves of planar and Gauss diagrams.

The first forbidden move has appeared, and been allowed, in the literature long before virtual knot theory was ever popular. Keeping with the notation introduced in [6], the objects defined up to Reidemeister moves and the first forbidden move are called *welded*. That paper is focused on welded braids, and proves that the welded pure braid groups are not trivial, and distinct from the classical pure braid groups. Further allowing the second forbidden move yields *unwelded* objects. In particular, all knots are trivial as unwelded knots, as shown in [11] and references therein. The main theorem of this paper is proved in Section 4.2 following Nelson's approach.

Let $vSL_n\mathcal{B}$ denote the monoid whose elements are equivalence classes of virtual string link diagrams on n components up to cobordisms and whose operation is concatenation. The monoids $wSL_n\mathcal{B}$ and $uwSL_n\mathcal{B}$ are defined similarly by allowing one and both forbidden moves respectively.

3 Fundamental results

The classical linking between two components of a classical link was first defined as an integral over the paths of a representative of the link and it admits combinatorial formulas that compute it from a planar or Gauss diagram. Using the un-normalization version,

$$\mathrm{ulk}(L_{(1)}, L_{(2)}) = \sum_{c \in L_{(1)} \cap L_{(2)}} \mathrm{w}(c),$$

where $L_{(i)}$ are components of a link L, c a crossing, an $\mathrm{w}(c)$ the sign of c.

If L is a classical knot, then this "usual" linking number is even, and often normalized by multiplication by a factor of $\frac{1}{2}$. For virtual link, the symmetry that this relies on needs not hold and the *ordered* linking numbers are different. Let

$$\mathrm{lk}(L_{(1)}, L_{(2)}) = \sum_{c:\ L_{(1)} \to L_{(2)}} \mathrm{w}(c)$$

be the linking number of $L_{(1)}$ over $L_{(2)}$, that is, the sum of the writhes of the crossings where $L_{(1)}$ goes over $L_{(2)}$. The notation $L_{(1)} \to L_{(2)}$ reflects that the arrows that are counted in the Gauss diagram point from $L_{(1)}$ to $L_{(2)}$. Then, $\mathrm{ulk}(L_{(1)}, L_{(2)}) = \mathrm{ulk}(L_{(2)}, L_{(1)}) = \mathrm{lk}(L_{(1)}, L_{(2)}) + \mathrm{lk}(L_{(2)}, L_{(1)})$.

These definitions can be used verbatim for components of virtual string links.

Lemma 3.1. *The linking numbers between components of a virtual link or virtual string link are invariant under the forbidden moves and cobordisms.*

Proof. Let L and L' differ by a single forbidden move. Let $L_{(1)}$ and $L_{(2)}$ be components of L, and $L'_{(i)}$, $i = 1, 2$, be the corresponding components in L'. Since forbidden moves change neither the number nor the sign of arrows between any two components, $\mathrm{lk}(L_{(1)}, L_{(2)}) = \mathrm{lk}(L'_{(1)}, L'_{(2)})$. For cobordisms, first notice that the corresponding claim also holds for the first and third Reidemeister moves. For the second Reidemeister move, assume that L' is obtained from L by canceling a pair of arrows from $L_{(1)}$ to $L_{(2)}$. Then, those arrows contribute $+1$ and -1 respectively to $\mathrm{lk}(L_{(1)}, L_{(2)})$ and therefore $\mathrm{lk}(L_{(1)}, L_{(2)}) = \mathrm{lk}(L'_{(1)}, L'_{(2)})$. Finally, the restrictions on the death, birth, and saddle moves make it so that the order of the endpoints of arrows can be changed, but the component on which they lie is preserved. Therefore, pairwise linking numbers are invariants of $uwSL_n$ and $vSL_n\mathcal{B}$. □

Proposition 4.9 in [5] shows that cobordism classes of virtual links are completely classified by pairwise virtual linking numbers. Any virtual string link can be mapped to an oriented virtual link with the same number of components by connecting the endpoints of each strand together without creating new crossings. This operation is called the *closure*. It immediately follows from the main theorem that the closure on $vSL_n\mathcal{B}$ is an injection onto cobordism classes of virtual links with n components.

Lemma 3.2. *Forbidden moves can be realized by cobordisms.*

Proof. Consider chords c and d in a Gauss diagram such that they have adjacent endpoints. Then, by using two saddle moves, these endpoints can be first exiled to some small closed component, and then reintroduced to their original location with the opposite order. Such a cobordism realizes both the first and the second forbidden move, and any other move which commutes arrows. □

Additivity of the linking numbers under concatenation is immediate from its definition. This is where the third classification from Theorem 1.1 comes from. Let $\mathrm{LK}_n\colon uwSL_n \to \mathbb{Z}^{n(n-1)}$ be the map that takes an unwelded string link L to the ordered list of its linking numbers, $(\mathrm{lk}(L_{(1)}, L_{(2)}), \mathrm{lk}(L_{(1)}, L_{(3)}), \dots, \mathrm{lk}(L_{(n-1)}, L_{(n)})$. In particular, there is standard form for any unwelded string link which displays exactly the crossings which contribute to $\mathrm{lk}(L_{(i)}, L_{(j)})$ in lexicographic order on ij. This is illustrated in Figure 6, with the sign of the crossing replaced by the signed number of parallel arrows with that sign. Conversely, given any list w of $n(n-1)$ integers, there is a unique unwelded string W such that $LK_n(W) = w$, using that standard form. Combining Lemmas 3.1 and 3.2, this discussion also holds for $\mathrm{LK}_n\colon vSL_n\mathcal{B} \to \mathbb{Z}^{n(n-1)}$.

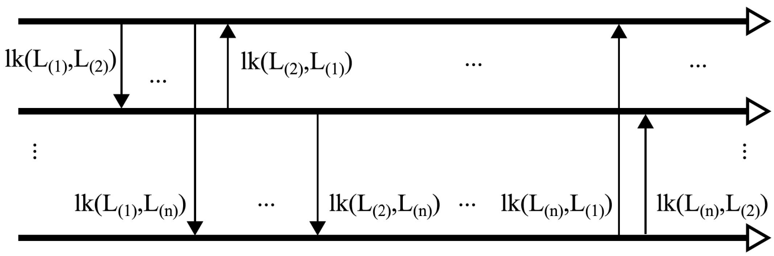


Figure 6: The standard form of an unwelded string link.

Proposition 3.3. *The monoids $vSL_n\mathcal{B}$, $wSL_n\mathcal{B}$, and $uwSL_n\mathcal{B}$ are isomorphic for all $n \geq 1$, and $vSL_1\mathcal{B}$ has exactly one element.*

Proof. First notice that $vSL_n\mathcal{B} \twoheadrightarrow wSL_n\mathcal{B} \twoheadrightarrow uwSL_n\mathcal{B}$, since each monoid is obtained from the previous one by allowing one more move. By Lemma 3.2, $uwSL_n\mathcal{B} \twoheadrightarrow vSL_n\mathcal{B}$, and thus all those maps are isomorphisms.

For the second part of the statement, it suffices to show that any arrow in a one-component string link Gauss diagram D can be erased using a cobordism and classical Reidemeister moves.

This is done by creating a saddle parallel to the crossing such that its head and foot are adjacent. Then, it can be removed using a single $RM1$ and the saddled off component can be reconnected to its original component of the link with a saddle more. Since D has only one component, every arrow can be canceled that way, and D is cobordant to the empty Gauss diagram on one long component. Thus, $vSL_1\mathcal{B}$ is isomorphic to $\{1\}$. □

4 Proof of the theorem

The main step is to show that the endpoints of any two adjacent chords on the Gauss diagram can be commuted using unwelded equivalences or using cobordisms. The proof could equivalently be illustrated with planar diagrams, but the simplicity of the standard form can be lost in the sea of virtual crossings that is required to realize it.

4.1 Standard form with cobordisms

Let D be a virtual string link diagram. By Lemma 3.2 the endpoints on each strand commute with each other. Thus, self-crossings can be isolated and removed, while the rest of the diagram can be organized to be in standard form, by canceling parallel arrows with opposite signs as needed.

4.2 Standard form with forbidden moves

This argument is a generalization of the proof that forbidden moves unknot virtual knots as it appears in [11]. The first and second forbidden moves on planar diagrams admit many orientations which give all possible choices of signs to the pairs of chords depicted in Figure 7. Thus, any two adjacent arrowheads or arrowfeet on a component of a string link Gauss diagram can commute.

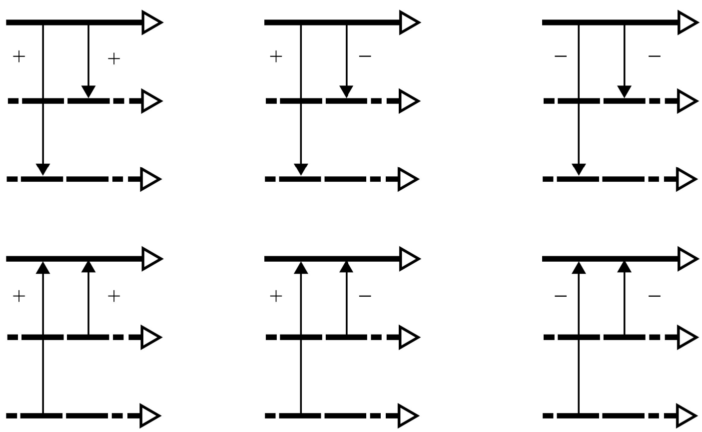


Figure 7: Pairs of crossings which commute with forbidden moves.

There are four different choices of signs that can occur in this situation. Two of them are depicted in Figure 8. The other cases can be obtained from these by applying various symmetries to the diagrams and changes of orientation of the strings.

It follows that the order of arrows on each component is irrelevant to the unwelded string link represented by a Gauss diagram. As with cobordisms, placing any self-crossing as an isolated crossing allows them to be canceled and the rest of the link can be put

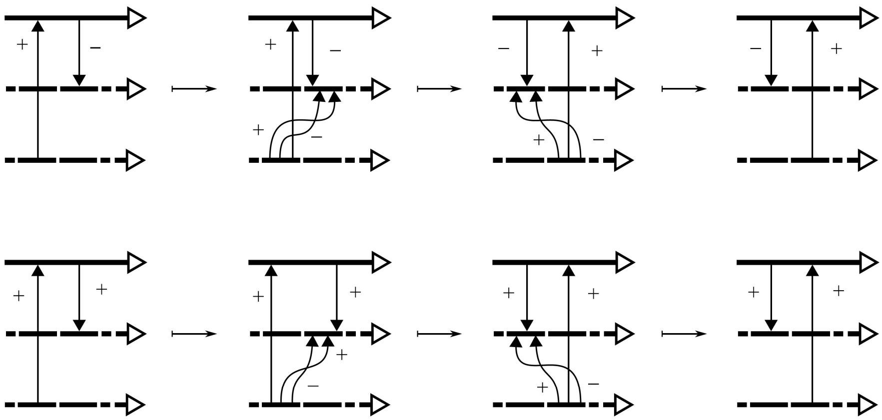


Figure 8: Commuting crossings using both forbidden moves.

in standard form, which is uniquely determined by the $n(n-1)$ pairwise virtual linking numbers.

5 Welded knot concordance

As an attempt to reach a midpoint between unwelded equivalence and cobordism, say that two virtual string link diagrams are *welded concordant* if one can be obtained from the other by a sequence of generalized Reidemeister moves, first forbidden moves, and genus 0 cobordisms. The welded moves are allowed to happen at any point of the cobordism.

The Tube map was defined on virtual knot diagrams by Satoh in [13], and gives a topological setting in which to interpret the first forbidden move, which is then more accurately called the *overcrossings commute* move, by mapping a planar knot diagram to a ribbon knotted torus in four dimensional space. Consider the Tube of each diagram appearing in a concordance movie between welded knots. The birth and death of unknotted components correspond respectively to creating and filling a ribbon (un)knotted torus. A proposed geometric realization of the saddle move is seen in Figure 9. The Tube map of welded string links is defined in [1] while a concordance theory for ribbon knotted surfaces which agrees with welded concordance has yet to be studied.

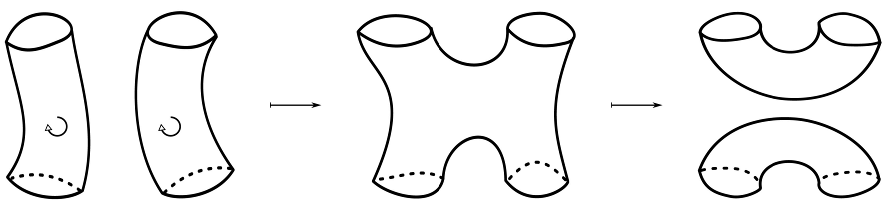


Figure 9: The saddle move realized on a ribbon knotted surface.

In the case of one component string links, Theorem 5.1 shows that allowing the first forbidden move trivializes the virtual knot concordance group, which is surprising, considering the well-known results that classical knots inject into welded knots.

Theorem 5.1. *Any long welded knot is concordant to the unknot.*

Proof. Again, this is shown by putting an arbitrary Gauss diagram D_0 with k crossings into the standard form using the allowed moves.

Let D_1 be the diagram obtained by adding an isolated crossing of opposite sign, pointing near the arrowhead of every crossing of D_0. This is shown in Figure 10. Then, using saddle moves, each pair of arrowheads can be isolated to its own closed component. The resulting diagram is D_2. Since the long component of D_2 contains only overcrossings and this is a welded link, they can be commuted such that the pairs of crossings have adjacent feet. This is diagram D_3 of Figure 10.

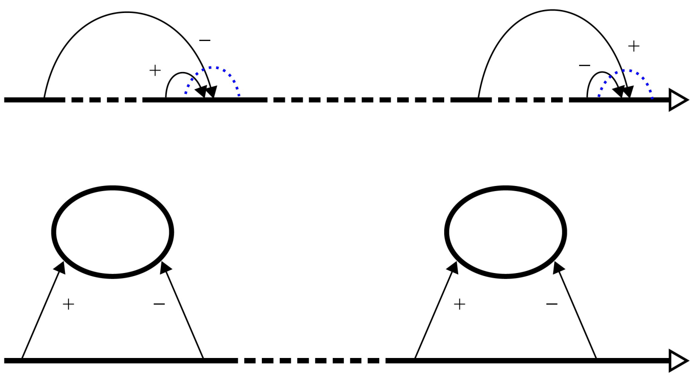


Figure 10: Gauss diagram where the long component contains only undercrossings, and the n closed components each have one positive and one negative overcrossing.

The second Reidemeister move allows to cancel each of the pairs, and finally, deaths delete the closed components. Since there were k saddle moves and the same number of deaths, this is a genus 0 cobordism to the unknot. □

For completeness, let's mention that as for classical knots, the concordance inverse of a welded string link is obtained by taking mirror images. Proposition 5.2 should be self-evident and it is presented here using Gauss diagram language.

Proposition 5.2. *Let S be a string link Gauss diagram with n components. Let $-S$ denote the diagram obtained by changing the direction of each core component and the sign of each arrow. Then, $S \# -S = -S \# S = U_n \in wSL_n\mathcal{C}$.*

Proof. Since $-(-S) = S$, it suffices to prove the second equality. Enumerate the chords of S as $c_1, c_2, \dots, c_k$ such that the first crossing in the diagram is from c_1, the one after that from c_2, and so on. Denote by $-c_i$ the mirror image of c_i in $-S$. Then, the innermost pair of crossings is $(-c_1, c_1)$. Using a saddle move which connects the arcs on the outside of the far endpoints of $\pm c_1$ to each other, that pair of crossings can be canceled using a second Reidemeister move. Figure 11 shows how crossings pair up with their inverses in the Gauss diagram.

Repeat this as needed (at most $k-1$ times) creating round components, and removing them with deaths as needed until the diagram is empty. □

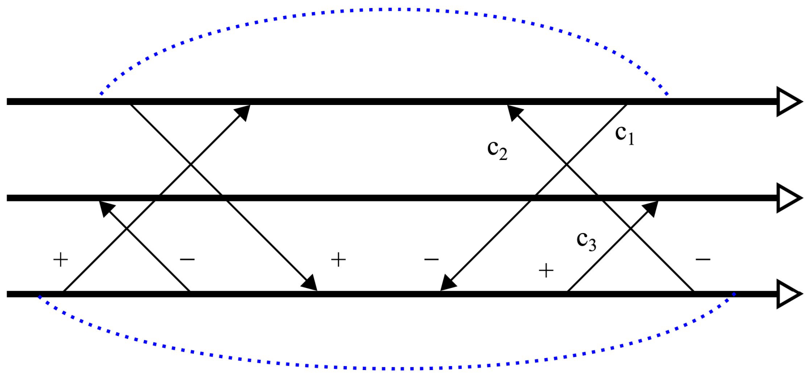


Figure 11: A virtual string link diagram concatenated to its concordance inverse. Dashed curves show the saddle moves needed to trivialize it.

As a corollary, $wSL_n\mathcal{C}$ is also a group for any n.

ORCID iD

Robin Gaudreau https://orcid.org/ 0000-0003-1186-9679

References

[1] B. Audoux, P. Bellingeri, J.-B. Meilhan and E. Wagner, Homotopy classification of ribbon tubes and welded string links, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **17** (2017), 713–761, doi: 10.2422/2036-2145.201507_003.

[2] B. Audoux, P. Bellingeri, J.-B. Meilhan and E. Wagner, Extensions of some classical local moves on knot diagrams, *Michigan Math. J.* **67** (2018), 647–672, doi:10.1307/mmj/1531447373.

[3] V. G. Bardakov, P. Bellingeri and C. Damiani, Unrestricted virtual braids, fused links and other quotients of virtual braid groups, *J. Knot Theory Ramifications* **24** (2015), 1550063 (23 pages), doi:10.1142/s0218216515500637.

[4] H. U. Boden and M. Nagel, Concordance group of virtual knots, *Proc. Amer. Math. Soc.* **145** (2017), 5451–5461, doi:10.1090/proc/13667.

[5] J. S. Carter, S. Kamada and M. Saito, Stable equivalence of knots on surfaces and virtual knot cobordisms, *J. Knot Theory Ramifications* **11** (2002), 311–322, doi:10.1142/s0218216502001639.

[6] R. Fenn, R. Rimányi and C. Rourke, The braid-permutation group, *Topology* **36** (1997), 123–135, doi:10.1016/0040-9383(95)00072-0.

[7] A. Fish and E. Keyman, Classifying links under fused isotopy, *J. Knot Theory Ramifications* **25** (2016), 1650042 (8 pages), doi:10.1142/s0218216516500425.

[8] N. Habegger and X.-S. Lin, The classification of links up to link-homotopy, *J. Amer. Math. Soc.* **3** (1990), 389–419, doi:10.2307/1990959.

[9] L. H. Kauffman, Virtual knot theory, *European J. Combin.* **20** (1999), 663–690, doi:10.1006/eujc.1999.0314.

[10] T. Nasybullov, Classification of fused links, *J. Knot Theory Ramifications* **25** (2016), 1650076 (21 pages), doi:10.1142/s0218216516500760.

[11] S. Nelson, Unknotting virtual knots with Gauss diagram forbidden moves, *J. Knot Theory Ramifications* **10** (2001), 931–935, doi:10.1142/s0218216501001244.

[12] T. Okabayashi, Forbidden moves for virtual links, *Kobe J. Math.* **22** (2005), 49–63, doi:10.1142/s0218216501000731.

[13] S. Satoh, Virtual knot presentation of ribbon torus-knots, *J. Knot Theory Ramifications* **9** (2000), 531–542, doi:10.1142/s0218216500000293.

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 51–60
https://doi.org/10.26493/1855-3974.2034.dad
(Also available at http://amc-journal.eu)

Association schemes with a certain type of p-subschemes*

Wasim Abbas † , Mitsugu Hirasaka

Department of Mathematics, College of Sciences, Pusan National University, 63 Beon-gil 2, Busandaehag-ro, Geumjung-gu, Busan 609-735, Korea

Received 4 July 2019, accepted 16 July 2020, published online 2 November 2020

Abstract

In this article, we focus on association schemes with some properties derived from the orbitals of a transitive permutation group G with a one-point stabilizer H satisfying $H < N_G(H) < N_G(N_G(H)) \trianglelefteq G$ and $|N_G(N_G(H))| = p^3$ where p is a prime. By a corollary of our main result we obtain some inequality which corresponds to the fact $|G : N_G(N_G(H))| \le p + 1$.

Keywords: Association schemes, p-schemes.

Math. Subj. Class. (2020): 05E15, 05E30

1 Introduction

Let G be a finite group with a subgroup H which satisfies

$$H < N_G(H) < N_G(N_G(H)) \trianglelefteq G \quad \text{and} \quad |N_G(N_G(H))| = p^3 \tag{1.1}$$

where p is a prime. In this article we focus on association schemes axiom-zing some properties derived from the orbitals of the action of G on G/H.

We shall recall some terminologies to show that the definition of coherent configurations is derived from properties of the binary relations obtained from a permutation group. Let G be a permutation group of a finite set Ω. Then G acts on $\Omega \times \Omega$ by its entry-wise action, i.e.,

$$(\alpha, \beta)^x := (\alpha^x, \beta^x) \quad \text{for } \alpha, \beta \in \Omega \text{ and } x \in G.$$

We denote the set of orbits of the action of G on $\Omega \times \Omega$ by $\mathrm{Inv}(G)$, which satisfies the following conditions:

*This work was supported by BK21 Dynamic Math Center, Department of Mathematics at Pusan National University.

†Corresponding author.

E-mail addresses: wasimabbas@pusan.ac.kr (Wasim Abbas), hirasaka@pusan.ac.kr (Mitsugu Hirasaka)

(i) The diagonal relation 1_Ω is a union of elements of $\mathrm{Inv}(G)$;

(ii) For each $s \in \mathrm{Inv}(G)$ we have $s^* \in \mathrm{Inv}(G)$ where $s^* := \{(\alpha, \beta) \mid (\beta, \alpha) \in s\}$;

(iii) For all $s, t, u \in \mathrm{Inv}(G)$ we have $\sigma_s \sigma_t = \sum_{u \in S} c_{st}^u \sigma_u$ for $c_{st}^u \in \mathbb{N}$ uniquely determined by s, t, u where σ_u is the adjacency matrix of u, i.e., $(\sigma_u)_{\alpha,\beta} = 1$ if $(\alpha, \beta) \in u$ and $(\sigma_u)_{\alpha,\beta} = 0$ if $(\alpha, \beta) \notin u$.

A *coherent configuration* is a pair (Ω, S) of a finite set Ω and a partition S of $\Omega \times \Omega$ which satisfies the conditions obtained from the above by replacing $\mathrm{Inv}(G)$ by S. We say that a coherent configuration (Ω, S) is *schurian* if $S = \mathrm{Inv}(G)$ for some permutation group G of Ω, and it is *homogeneous* or an *association scheme* if $1_\Omega \in S$ (see [2] and [3] for its background).

Suppose that G has a subgroup H which satisfies (1.1). Then $|H| = p$, $|N_G(H)| = p^2$ and for each $g \in G$ we have the following:

(i) $|HgH|/|H| \in \{1, p\}$ and $|N_G(H)gN_G(H)|/|N_G(H)| \in \{1, p\}$;

(ii) $|HgH|/|H| = 1$ if and only if $g \in N_G(H)$;

(iii) $|N_G(H)gN_G(H)|/|N_G(H)| = 1$ if and only if $g \in N_G(N_G(H))$;

(iv) $N_G(N_G(H))$ is the smallest normal subgroup of G containing H.

Since G acts faithfully and transitively on the set of right cosets of H in G by its right multiplication, it induces a schurian association scheme (Ω, S) where $\Omega = \{Hx \mid x \in G\}$ and $S = \mathrm{Inv}(G)$ such that, for each $s \in S$ we have the following:

(i) $n_s \in \{1, p\}$ where $n_s := c_{ss^*}^{1_\Omega}$;

(ii) $\mathbf{O}_\theta(S)$ forms a group of order p where $\mathbf{O}_\theta(S) := \{s \in S \mid n_s = 1\}$;

(iii) $\mathbf{O}^\theta(S) = \{s \in S \mid ss^*s = s\}$ where $\mathbf{O}^\theta(S)$ is the thin residue of S (see Section 2, [9] or [10] for its definition).

The following is our main result:

Theorem 1.1. *Let (Ω, S) be an association scheme with $\mathbf{O}_\theta(S) < \mathbf{O}^\theta(S)$ such that $n_s \in \{1, p\}$ for each $s \in S$ and $n_{\mathbf{O}^\theta(S)} = p^2$ where p is a prime. Then $|\Omega| \leq p^2(p+1)$.*

In [4] they give a criterion on association schemes whose thin residue $\mathbf{O}^\theta(S)$ induces the subschemes isomorphic to either

$$C_{p^2}, \quad C_p \times C_p \quad \text{or} \quad C_p \wr C_p.$$

Here we denote $(G, \mathrm{Inv}(G))$ by G when G acts on itself by its right multiplication and we denote the wreath product of one scheme (Δ, U) by another scheme (Γ, V) by $(\Delta, U) \wr (\Gamma, V)$, i.e.,

$$(\Delta, U) \wr (\Gamma, V) := (\Delta \times \Gamma, \{1_\Gamma \otimes u \mid u \in U\} \cup \{v \otimes U \mid v \in V \setminus \{1_\Gamma\}\})$$

where

$$1_\Gamma \otimes u := \{((\delta_1, \gamma), (\delta_2, \gamma)) \mid (\delta_1, \delta_2) \in u, \gamma \in \Gamma\} \quad \text{and}$$
$$v \otimes U := \{((\delta_1, \gamma_1), (\delta_2, \gamma_2)) \mid \delta_1, \delta_2 \in \Delta, (\gamma_1, \gamma_2) \in v\}.$$

For the case of $\mathbf{O}^\theta(S) \simeq C_{p^2}$ we can apply the main result in [7] to conclude that (Ω, S) is schurian. For the case of $\mathbf{O}^\theta(S) \simeq C_p \times C_p$ we can say that $|\Omega| \le p^2(p^2+p+1)$ under the assumption that $n_s = p$ for each $s \in S \setminus \mathbf{O}^\theta(S)$. For the case of $\mathbf{O}^\theta(S) \simeq C_p \wr C_p$ we had no progression for the last five years.

In [6] all association schemes of degree 27 are classified by computational enumeration, and there are three pairs of non-isomorphic association schemes with $\mathbf{O}^\theta(S) \simeq C_3 \wr C_3$ which are algebraic isomorphic. These examples had given an impression that we need some complicated combinatorial argument to enumerate p-schemes (Ω, S) with $\mathbf{O}^\theta(S) \simeq C_p \wr C_p$ and $\{n_s \mid s \in S \setminus \mathbf{O}^\theta(S)\} = \{p\}$. The following reduces our argument to the p-schemes of degree p^3 where an association scheme (Ω, S) is called a p-scheme if $|s|$ is a power of p for each $s \in S$:

Corollary 1.2. *For each p-scheme (Ω, S) with $\mathbf{O}^\theta(S) \simeq C_p \wr C_p$, if $n_s = p$ for each $s \in S \setminus \mathbf{O}^\theta(S)$, then $|\Omega| = p^3$.*

In the proof of Theorem 1.1 the theory of coherent configurations plays an important role through the thin residue extension which is a way of construction of coherent configurations from association schemes (see [5, Theorem 2.1] or [8]) . The following is the kernel of our paper:

Theorem 1.3. *For each coherent configuration (Ω, S) whose fibers are isomorphic to $C_p \wr C_p$, if $|s| = p^3$ for each $s \in S$ with $\sigma_s\sigma_s = 0$, then either $|\Omega| \le p^2(p+1)$ or $ss^*s = s$ for each $s \in S$.*

In Section 2 we prepare necessary terminologies on coherent configurations. In Section 3 we prove our main results.

2 Preliminaries

Throughout this section, we assume that (Ω, S) is a coherent configuration. An element of Ω and an element of S are called a *point* and a *basis relation*, respectively. Furthermore, $|\Omega|$ and $|S|$ are called the *degree* and *rank* of (Ω, S), respectively. For all $\alpha, \beta \in \Omega$ the unique element in S containing (α, β) is denoted by $r(\alpha, \beta)$. For $s \in S$ and $\alpha \in \Omega$ we set

$$\alpha s := \{\beta \in \Omega \mid (\alpha, \beta) \in s\}.$$

A subset Δ of Ω is called a *fiber* of (Ω, S) if $1_\Delta \in S$. For each $s \in S$, there exists a unique pair (Δ, Γ) of fibers such that $s \subseteq \Delta \times \Gamma$. For fibers Δ, Γ of (Ω, S) we denote the set of $s \in S$ with $s \subseteq \Delta \times \Gamma$ by $S_{\Delta,\Gamma}$, and we set $S_\Delta := S_{\Delta,\Delta}$. It is easily verified that (Δ, S_Δ) is a homogeneous coherent configuration. Now we define the *complex product* on the power set of S as follows: For all subsets T and U of S we set

$$TU := \{s \in S \mid c_{tu}^s > 0 \text{ for some } t \in T \text{ and } u \in U\}$$

where the singleton $\{t\}$ in the complex product is written without its parenthesis.

The following equations are frequently used without any mention:

Lemma 2.1. *Let (Ω, S) be a coherent configuration. Then we have the following:*

(i) *For all $r, s \in S$, if $rs \neq \emptyset$, then $n_r n_s = \sum_{t\in S} c_{rs}^t n_t$;*

(ii) *For all $r, s, t \in S$ we have $|t|c_{rs}^{t^*} = |r|c_{st}^{r^*} = |s|c_{tr}^{s^*}$;*

(iii) *For all* $r, s \in S$ *we have* $|\{t \in S \mid t \in rs\}| \leq \gcd(n_r, n_s)$.

For $T \subseteq S_{\Delta,\Gamma}$ we set

$$n_T := \sum_{t \in T} n_t.$$

Here we mention closed subsets, their subschemes and factor scheme according to the terminologies given in [10]. Let (Ω, S) be an association scheme and $T \subseteq S$. We say that a non-empty subset T of S is *closed* if $TT^* \subseteq T$ where

$$T^* := \{t^* \mid t \in T\},$$

equivalently $\bigcup_{t \in T} t$ is an equivalence relation on Ω whose equivalence classes are

$$\{\alpha T \mid \alpha \in \Omega\}$$

where $\alpha T := \{\beta \in \Omega \mid (\alpha, \beta) \in t \text{ for some } t \in T\}$. Let T be a closed subset of S and $\alpha \in \Omega$. It is well-known (see [9]) that

$$(\Omega, S)_{\alpha T} := (\alpha T, \{t \cap (\alpha T \times \alpha T) \mid t \in T\})$$

is an association scheme, called the *subscheme* of (Ω, S) induced by αT, and that

$$(\Omega, S)^T := (\Omega/T, S/\!\!/T)$$

is also an association scheme where

$$\Omega/T := \{\alpha T \mid \alpha \in \Omega\}, \quad S/\!\!/T = \{s^T \mid s \in S\} \quad \text{and}$$
$$s^T := \{(\alpha T, \beta T) \mid (\gamma, \delta) \in s \text{ for some } (\gamma, \delta) \in \alpha T \times \beta T\},$$

which is called the *factor scheme* of (Ω, S) over T.

We say that a closed subset T is *thin* if $n_t = 1$ for each $t \in T$, and $\mathbf{O}_\theta(S)$ is called the *thin radical* of S, and the smallest closed subset T such that $S/\!\!/T$ is thin is called the *thin residue* of S, which is denoted by $\mathbf{O}^\theta(S)$.

3 Proof of the main theorem

Let (Ω, S) be a coherent configuration whose distinct fibers are $\Omega_1, \Omega_2, \ldots, \Omega_m$. For all integers i, j with $1 \leq i, j \leq m$ we set

$$S_{ij} := S_{\Omega_i, \Omega_j} \quad \text{and} \quad S_i := S_{ii}.$$

Throughout this section we assume that $(\Omega_i, S_i) \simeq C_p \wr C_p$ for $i = 1, 2, \ldots, m$ where p is a prime and $C_p \wr C_p$ is a unique non-thin p-scheme of degree p^2 up to isomorphism.

For $s \in S$ we say that s is *regular* if $ss^*s = \{s\}$ and we denote by R the set of regular elements in S.

Lemma 3.1. *For each regular element* $s \in S_{ij}$ *with* $n_s = p$ *we have*

$$\sigma_s \sigma_{s^*} = p(\textstyle\sum_{t \in \mathbf{O}_\theta(S_i)} \sigma_t) \quad \textit{and} \quad \sigma_{s^*} \sigma_s = p(\textstyle\sum_{t \in \mathbf{O}_\theta(S_j)} \sigma_t).$$

In particular, $ss^* = \mathbf{O}_\theta(S_i)$ *and* $s^*s = \mathbf{O}_\theta(S_j)$.

Proof. Notice that $\{1_{\Omega_i}\} \subsetneq ss^* \subset S_i$ and $ts = \{s\}$ for each $t \in ss^*$. Since $\{t \in S_i \mid ts = \{s\}\}$ is a closed subset of valency at most n_s, it follows from $(\Omega_i, S_i) \simeq C_p \wr C_p$ that $ss^* = \mathbf{O}_\theta(S_i)$, and hence for each $t \in ss^*$

$$c_{ss^*}^t = c_{st}^s n_{s^*}/n_{t^*} = p.$$

This implies that $\sigma_s\sigma_{s^*} = p(\sum_{t\in\mathbf{O}_\theta(S_i)} \sigma_t)$. By the symmetric argument we have $\sigma_{s^*}\sigma_s = p(\sum_{t\in\mathbf{O}_\theta(S_j)} \sigma_t)$. □

Lemma 3.2. *For each non-regular element $s \in S_{ij}$ with $n_s = p$ we have*

$$\sigma_s\sigma_{s^*} = p\sigma_{1_{\Omega_i}} + \sum_{u\in S_i\setminus\mathbf{O}_\theta(S_i)} \sigma_u \quad \textit{and} \quad \sigma_{s^*}\sigma_s = p\sigma_{1_{\Omega_j}} + \sum_{u\in S_j\setminus\mathbf{O}_\theta(S_j)} \sigma_u.$$

Proof. Notice that $\{t \in S_i \mid ts = \{s\}\} = \{1_{\Omega_i}\}$, otherwise, s is regular or $n_s = p^2$, a contradiction. This implies that the singletons ts with $t \in \mathbf{O}_\theta(S_i)$ are distinct elements of valency p. Since

$$p^2 = |\Omega_j| = \sum_{s\in S_{ij}} n_s \geq \sum_{t\in\mathbf{O}_\theta(S_i)} n_{ts} = p + p + \cdots + p = p^2,$$

it follows that $\mathbf{O}_\theta(S_i)s = S_{ij}$.

We claim that $S_i \setminus \mathbf{O}_\theta(S_i) \subseteq ss^*$. Let $u \in S_i \setminus \mathbf{O}_\theta(S_i)$. Then there exists $t \in \mathbf{O}_\theta(S_i)$ such that $u \in tss^*$ since $u \in S_{ij}s^* = \mathbf{O}_\theta(S_i)ss^*$. This implies that $u = t^*u \subseteq t^*(tss^*) = ss^*$.

By the claim with $p^2 = n_s n_{s^*} = \sum_{t\in S_i} c_{ss^*t} n_t$ and $c_{ss^*1_{\Omega_i}} = n_s = p$ we have the first statement, and the second statement is obtained by the symmetric argument. □

For the remainder of this section we assume that $n_s = p$ for each $s \in \bigcup_{i\neq j} S_{ij}$.

Lemma 3.3. *The set $\bigcup_{s\in R} s$ is an equivalence relation on Ω.*

Proof. Since $1_{\Omega_i} \in S_i \subseteq R$ for $i = 1, 2, \ldots, m$, $\bigcup_{s\in R} s$ is reflexive. Since $ss^*s = \{s\}$ is equivalent to $s^*ss^* = \{s^*\}$, $\bigcup_{s\in R} s$ is symmetric.

Let $\alpha \in \Omega_i$, $\beta \in \Omega_j$ and $\gamma \in \Omega_k$ with $r(\alpha,\beta), r(\beta,\gamma) \in R$. Then we have

$$r(\alpha,\gamma)r(\alpha,\gamma)^* \subseteq r(\alpha,\beta)r(\beta,\gamma)r(\beta,\gamma)^*r(\alpha,\beta)^*.$$

If one of $r(\alpha,\beta)$, $r(\beta,\gamma)$ is thin, then $(\alpha,\gamma)r(\alpha,\gamma)^*$, and hence $r(\alpha,\gamma) \in R$. Now we assume that both of them are non-thin. Since $r(\beta,\gamma)r(\beta,\gamma)^* = \mathbf{O}_\theta(S_j) = r(\alpha,\beta)^*r(\alpha,\beta)$, it follows that

$$r(\alpha,\gamma)r(\alpha,\gamma)^* \subseteq r(\alpha,\beta)r(\alpha,\beta)^* = \mathbf{O}_\theta(S_i).$$

Applying Lemma 3.1 and 3.2 we obtain that $r(\alpha,\gamma)$ is regular, and hence $\bigcup_{s\in R} s$ is transitive. □

Lemma 3.4. *The set $\bigcup_{s\in N} s$ is an equivalence relation on Ω where $N := \bigcup_{i=1}^m S_i \cup (S \setminus R)$.*

Proof. Since $1_{\Omega_i} \in S_i \subseteq N$ for $i = 1, 2, \dots, m$, $\bigcup_{s \in N} s$ is reflexive. By Lemma 3.3, $\bigcup_{s \in R}$ is symmetric, so that $\bigcup_{s \in N} s$ is symmetric.

Let $\alpha \in \Omega_i$, $\beta \in \Omega_j$ and $\gamma \in \Omega_k$ with $r(\alpha, \beta), r(\beta, \gamma) \in N$. Since $\bigcup_{i=1}^m S_i \subseteq R$, it follows from Lemma 3.3 that it suffices to show that

$$r(\alpha, \gamma) \in S \setminus R$$

under the assumption that

$$r(\alpha, \beta), r(\beta, \gamma) \in S \setminus R \quad \text{with } i \neq k.$$

Suppose the contrary, i.e., $r(\alpha, \gamma) \in R$. Then, by Lemma 3.3, $S_{ik} \subseteq R$. Since

$$r(\alpha, \beta) r(\beta, \gamma) \subseteq S_{ik} \subseteq R,$$

it follows that

$$\mathbf{O}_\theta(S_i) r(\alpha, \beta) r(\beta, \gamma) = r(\alpha, \beta) r(\beta, \gamma).$$

On the other hand, we have

$$\mathbf{O}_\theta(S_i) r(\alpha, \beta) r(\beta, \gamma) = S_{ij} r(\beta, \gamma) = S_{ik}.$$

Thus, $r(\alpha, \beta) r(\beta, \gamma) = S_{ik}$. Since $i \neq k$, each element of S_{ik} has valency p, and hence,

$$\sigma_{s_1} \sigma_{s_2} = \sum_{u \in S_{ik}} \sigma_u$$

where $s_1 := r(\alpha, \beta)$ and $s_2 := r(\beta, \gamma)$. By Lemma 3.2,

$$p^2 = \langle \sigma_{s_1} \sigma_{s_2}, \sigma_{s_1} \sigma_{s_2} \rangle = \langle \sigma_{s_1}^* \sigma_{s_1}, \sigma_{s_2} \sigma_{s_2}^* \rangle = p^2 + p(p-1),$$

a contradiction where $\langle\ ,\ \rangle$ is the inner product defined by

$$\langle A, B \rangle := 1/p^2 \operatorname{tr}(AB^*) \quad \text{for all } A, B \in M_\Omega(\mathbb{C}).$$

Therefore, $\bigcup_{s \in N} s$ is transitive. □

Lemma 3.5. *We have either $R = S$ or $N = S$.*

Proof. Suppose $R \neq S$. Let $\alpha, \beta \in \Omega$ with $r(\alpha, \beta) \in R$. Since $R \neq S$, there exists $\gamma \in \Omega$ with $r(\alpha, \gamma) \in N$. Notice that $r(\beta, \gamma) \in R \cup N$. By Lemma 3.3, $r(\beta, \gamma) \in N$, and hence, by Lemma 3.4,

$$r(\alpha, \beta) \in R \cap N = \bigcup_{i=1}^m S_i.$$

Since $\alpha, \beta \in \Omega$ are arbitrarily taken, it follows that

$$R = \bigcup_{i=1}^m S_i \quad \text{and} \quad N = S.$$ □

Lemma 3.6. *Suppose that $S = N$ and $s_1 \in S_{ij}, s_2 \in S_{jk}$ and $s_3 \in S_{ik}$ with distinct i, j, k. Then $\sigma_{s_1} \sigma_{s_2} = \sigma_{s_3} (\sum_{t \in \mathbf{O}_\theta(S_k)} a_t \sigma_t)$ for some non-negative integers a_t with $\sum_{t \in \mathbf{O}_\theta(S_k)} a_t = p$, $\sum_{t \in \mathbf{O}_\theta(S_k)} a_t^2 = 2p - 1$ and for each $u \in \mathbf{O}_\theta(S_k) \setminus \{1_{\Omega_k}\}$, $\sum_{t \in \mathbf{O}_\theta(S_k)} a_t a_{tu} = p - 1$.*

Proof. Since $s_1s_2 \subseteq S_{ij} = s_3\mathbf{O}_\theta(S_k)$, $\sigma_{s_1}\sigma_{s_2} = \sum_{t\in\mathbf{O}_\theta(S_k)} a_t\sigma_{s_3t}$ for some non-negative integers a_t. Since $\sigma_{s_3t} = \sigma_{s_3}\sigma_t$ and

$$p^2 = n_{s_1}n_{s_2} = \sum_{t\in\mathbf{O}_\theta(S_k)} a_t n_{s_3t} = p\sum_{t\mathbf{O}_\theta(S_j)} a_t,$$

it remains to show the last two equalities on a_t with $t \in \mathbf{O}_\theta(S_j)$. Expanding $\sigma^*_{s_2}\sigma^*_{s_1}\sigma_{s_1}\sigma_{s_2}$ by two ways we obtain from Lemma 3.2 that

$$(2p^2-p)\sigma_{1_{\Omega_j}} + (p^2-p)\sum_{t\in\mathbf{O}_\theta(S_j)\setminus\{1_{\Omega_j}\}}\sigma_t + (p^2-2p)\sum_{u\in S_j\setminus\mathbf{O}_\theta(S_j)}\sigma_u$$
$$= \sum_{t\in\mathbf{O}_\theta(S_k)} a_t\sigma^*_t\sigma^*_{s_3}\sigma_{s_3}\sum_{t\in\mathbf{O}_\theta(S_k)} a_t\sigma_t.$$

Therefore, we conclude from Lemma 3.2 that

$$p\sum_{t\in\mathbf{O}_\theta(S_k)} a_t^2 = 2p^2-p \quad \text{and} \quad p\sum_{t\in\mathbf{O}_\theta(S_k)} a_ta_{tu} = p^2-p$$

for each $u \in \mathbf{O}_\theta(S_k)$ with $u \neq 1_{\Omega_k}$. □

For the remainder of this section we assume that

$$S = N.$$

For $i = 1, 2, \ldots, m$ we take $\alpha_i \in \Omega_i$ and we define $t_i \in S_i$ such that $t_1 \in \mathbf{O}_\theta(S_1)\setminus\{1_{\Omega_1}\}$, and for $i = 2, 3, \ldots, m$, t_i is a unique element in $\mathbf{O}_\theta(S_i)$ with $r(\alpha_1t_1, \alpha_it_i) = r(\alpha_1, \alpha_i)$. Then C_p acts semi-regularly on Ω such that

$$\Omega\times C_p \to \Omega, \quad (\beta_i, t^j) \mapsto \beta_i t_i^j,$$

where $C_p = \langle t\rangle$ and β_i is an arbitrary element in Ω_i.

Lemma 3.7. *The above action acts semi-regularly on* Ω *as an automorphism of* (Ω, S).

Proof. Since C_p acts regularly on each of geometric coset of $\mathbf{O}_\theta(S_i)$ for $i = 1, 2, \ldots, m$, the action is semi-regular on Ω. By the definition of $\{t_i\}$, it is straightforward to show that $r(\alpha_1, \alpha_i)$ is fixed by the action on $\Omega\times\Omega$, and hence each element of $\bigcup_{j=2}^m S_{1j}\cup S_{j1}$ is also fixed since $S_{1j} = \mathbf{O}_\theta(S_1)r(\alpha_1, \alpha_j)$. Let $s \in S_{ij}$with $2 \leq i, j$. Notice that $r(\alpha_i, \alpha_1)r(\alpha_1, \alpha_j)$ is a proper subset of S_{ij} by Lemma 3.6. This implies that s is obtained as the intersection of some of $t_i^k r(\alpha_i, \alpha_1)r(\alpha_1, \alpha_j)$ with $0 \leq k \leq p-1$, and hence s is fixed. □

For each $i = 1, 2, \ldots, m$ we take $\{\alpha_{ik} \mid k = 1, 2, \ldots, m\}$ to be a complete set of representatives with respect to the equivalence relation $\bigcup_{t\in\mathbf{O}_\theta(S_i)} t$ on Ω_i.

Lemma 3.8. *For each* $s \in S_{ij}$ *with* $i \neq j$ *and all* $k, l = 1, 2, \ldots, p$ *there exists a unique* $h(s)_{kl} \in \mathbb{Z}_p$ *such that* $r(\alpha_{ik}, \alpha_{jl}t^{h(s)_{kl}}) = s$. *Moreover, if* $s_1 \in S_{ij}$ *and* $t^a \in \mathbf{O}_\theta(S_k)$ *with* $s_1 = st^a$, *then* $h(s_1)_{kl} = h(s)_{kl} + a$ *for all* $k, l = 1, 2, \ldots, m$.

Proof. Since $\mathbf{O}_\theta(S_j)$ acts regularly on S_{ij} by its right multiplication, the first statement holds. The second statement is obtained by a direct computation. □

Lemma 3.9. *For each $s \in S_{ij}$ with $i \neq j$ and all $k, l = 1, 2, \ldots, p$ we have*

$$s \cap (\alpha_{ik}\mathbf{O}_\theta(S_i) \times \alpha_{jl}\mathbf{O}_\theta(S_j)) = \{(\alpha_{ik}t_i^a, \alpha_{jl}t_j^b) \mid b - a = h(s)_{kl}\}.$$

Proof. Notice that

$$r(\alpha_{ik}t_i^a, \alpha_{jl}t_j^b) = (t_i^a)^* r(\alpha_{ik}, \alpha_{jl})t_j^b = r(\alpha_{ik}, \alpha_{jl})t_j^{b-a}.$$

Since $r(\alpha_{ik}, \alpha_{jl}t^{h(s)_{kl}}) = s$ by Lemma 3.8, it follows that $r(\alpha_{ik}t_i^a, \alpha_{jl}t_j^b) = s$ if and only if $b - a = h(s)_{kl}$. □

Proposition 3.10. *For each $s \in S_{ij}$ with $i \neq j$ the matrix $(h(s)_{kl}) \in M_{p\times p}(\mathbb{Z}_p)$ satisfies that, for all distinct $k_1, k_2 \in \{1, 2, \ldots, p\}$,*

$$\{h(s)_{k_1,l} - h(s)_{k_2,l} \mid l = 1, 2, \ldots, p\} = \mathbb{Z}_p.$$

In other word the matrix is a generalized Hadamard matrix of degree p over $\mathbb{Z}_p$, equivalently, the matrix $(\xi^{h(s)_{kl}}) \in M_{p\times p}(\mathbb{C})$ is a complex Hadamard matrix of Butson type (p, p) where ξ is a primitive p-th root of unity.

Proof. Notice that, for all distinct k, l, by Lemma 3.9,

$$\{\gamma \in \Omega \mid r(\alpha_{ik}t_i^a, \gamma) = r(\alpha_{il}t_i^b, \gamma) = s\}$$

equals

$$\bigcup_{r=1}^{p} \{\alpha_{jr}t_j^c \mid c - a = h(s)_{kr}, c - b = h(s)_{lr}\}.$$

Since the upper one is a singleton by Lemma 3.2, there exists a unique $r \in \{1, 2, \ldots, p\}$ such that $b - a = h(s)_{kr} - h(s)_{lr}$. Since a and b are arbitrarily taken, the first statement holds.

The second statement holds since $\sum_{i=0}^{p-1} x^i$ is the minimal polynomial of ξ over $\mathbb{Q}$. □

We shall write the matrix $(\xi^{h(s)_{kl}})$ as $H(s)$. For $s \in S_{ij}$ with $i \neq j$, the restriction of σ_s to $\Omega_i \times \Omega_j$ can be viewed as a $(p \times p)$-matrix whose (k, l)-entry is the matrix $P_i^{h(s)_{kl}}$ where P_i is the permutation matrix corresponding to the mapping $\beta_i \mapsto \beta_i t_i$ where we may assume that $P_i = P_j$, say P, for all $i, j = 1, 2, \ldots, m$ by Lemma 3.7. Notice that $H(s)$ is obtained from $(P^{h(s)_{kl}})$ by sending $P^{h(s)_{kl}}$ to $\xi^{h(s)_{kl}}$.

Proposition 3.11. *For all $s_1 \in S_{ij}$, $s_2 \in S_{jk}$ and $s_3 \in S_{ik}$ with distinct i, j, k we have $H(s_1)H(s_2) = \alpha H(s_3)$ for some $\alpha \in \mathbb{C}$ with $|\alpha| = \sqrt{p}$.*

Proof. By Lemma 3.6, $H(s_1)H(s_2) = H(s_3)(\sum_{i=0}^{p-1} a_i\xi^i)$ for some $a_i \in \mathbb{Z}$ where $a_i = c_{t_k^i}$. Thus, it suffices to show that $|(\sum_{i=0}^{p-1} a_i\xi^i)|^2 = p$. By Lemma 3.6, the left hand side equals

$$\sum_{i=0}^{p-1}\sum_{j=0}^{p-1} a_i a_j \xi^{i-j} = \sum_{i=0}^{p-1} a_i^2 + \sum_{i=1}^{p-1}\sum_{j=0}^{p-1} a_j a_{i+j}\xi^i = (2p-1) + (p-1)(-1) = p. \quad \square$$

Corollary 3.12. *Let $s_i := r(\alpha_1, \alpha_i)$ for $i = 2, 3, \ldots, m$ and $\mathbf{B}_i$ denote the basis consisting of the rows of $H(s_i)$, $i = 2, 3, \ldots, m$, and $\mathbf{B}_1$ be the standard basis. Then $\{\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_m\}$ is a mutually unbiased bases for $\mathbb{C}^p$, and $m \leq p + 1$.*

Proof. The first statement is an immediate consequence of Proposition 3.10, and the second statement follows from a well-known fact that the number of mutually unbiased bases for $\mathbb{C}^n$ is at most $n + 1$ (see [1]). □

Proof of Theorem 1.3. Suppose that $R \neq S$. Then $N = S$ and the theorem follows from Corollary 3.12. □

Proof of Theorem 1.1. Since $n_{\mathbf{O}^\theta(S)} = p^2$ and $\mathbf{O}_\theta(S) < \mathbf{O}^\theta(S)$, it follows from [5, Theorem 2.1] (or see [8]) that the thin residue extension of (Ω, S) is a coherent configuration with all fibers isomorphic to $C_p \wr C_p$ such that each basic relation out of the fibers has valency p.

We claim that $S = N$. Otherwise, $S = R$, which implies that $\langle ss^* \mid s \in S\rangle$ has valency p. Since $\mathbf{O}^\theta(S) = \langle ss^* \mid s \in S\rangle$ (see [9]), it contradicts that $\mathbf{O}^\theta(S)$ has valency p^2.

By the claim, $S = N$. Since the number of fibers of the thin residue extension of (Ω, S) equals $|\Omega/\mathbf{O}^\theta(S)|$, the theorem follows from Theorem 1.3. □

Proof of Corollary 1.2. Since (Ω, S) is a p-scheme and $\mathbf{O}^\theta(S) \simeq C_p \times C_p$, $|\Omega|$ is a power of p greater than p^2. By Theorem 1.1, $|\Omega| \leq (p + 1)p^2$, and hence, $|\Omega| = p^3$. □

ORCID iDs

Wasim Abbas https://orcid.org/0000-0002-1706-1462

References

[1] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury and F. Vatan, A new proof for the existence of mutually unbiased bases, *Algorithmica* **34** (2002), 512–528, doi:10.1007/s00453-002-0980-7.

[2] E. Bannai and T. Ito, *Algebraic Combinatorics I: Association Schemes*, The Benjamin/Cummings Publishing, Menlo Park, CA, 1984.

[3] A. E. Brouwer, A. M. Cohen and A. Neumaier, *Distance-Regular Graphs*, volume 18 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Springer-Verlag, Berlin, 1989, doi:10.1007/978-3-642-74341-2.

[4] J. R. Cho, M. Hirasaka and K. Kim, On p-schemes of order p^3, *J. Algebra* **369** (2012), 369–380, doi:10.1016/j.jalgebra.2012.06.026.

[5] S. A. Evdokimov and I. N. Ponomarenko, Schemes of relations of the finite projective plane, and their extensions, *Algebra i Analiz* **21** (2009), 90–132, `http://mi.mathnet.ru/aa996`, *St. Petersburg Math. J.* **21** (2010), 65–93, doi:10.1090/s1061-0022-09-01086-3.

[6] A. Hanaki and I. Miyamoto, Classification of association schemes of small order, *Discrete Math.* **264** (2003), 75–80, doi:10.1016/s0012-365x(02)00551-4.

[7] M. Hirasaka and P.-H. Zieschang, Sufficient conditions for a scheme to originate from a group, *J. Comb. Theory Ser. A* **104** (2003), 17–27, doi:10.1016/s0097-3165(03)00104-3.

[8] M. Muzychuk and I. Ponomarenko, On quasi-thin association schemes, *J. Algebra* **351** (2012), 467–489, doi:10.1016/j.jalgebra.2011.11.012.

[9] P.-H. Zieschang, *An Algebraic Approach to Association Schemes*, volume 1628 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 1996, doi:10.1007/bfb0097032.

[10] P.-H. Zieschang, *Theory of Association Schemes*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005, doi:10.1007/3-540-30593-9.

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 61–76
https://doi.org/10.26493/1855-3974.2106.423
(Also available at http://amc-journal.eu)

Dominating sets in finite generalized quadrangles*

Tamás Héger †

MTA–ELTE Geometric and Algebraic Combinatorics Research Group, ELTE Eötvös Loránd University, Pázmány Péter sétány 1/C, Budapest, Hungary, H-1117

Lisa Hernandez Lucas

Department of Mathematics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Received 30 August 2019, accepted 25 March 2020, published online 2 November 2020

Abstract

A dominating set in a graph is a set of vertices such that each vertex not in the set has a neighbor in the set. The domination number is the smallest size of a dominating set. We consider this problem in the incidence graph of a generalized quadrangle. We show that the domination number of a generalized quadrangle with parameters s and t is at most $2st + 1$, and we prove that this bound is sharp if $s = t$ or if $s = q - 1$ and $t = q + 1$. Moreover, we give a complete classification of smallest dominating sets in generalized quadrangles where $s = t$, and give some general results for small dominating sets in the general case.

Keywords: Dominating set, finite generalized quadrangle.

Math. Subj. Class. (2020): 05B25, 05C69, 51E12

1 Preliminaries

Dominating sets in graphs have already been studied in 1958, but there was a boost of interest after the publishing of a survey paper in the '70s by Cockayne and Hedetniemi [3], in which the authors show that the domination problem is related to the well-known problem of colorings of graphs. In [8] a dominating set is defined as follows:

*The authors gratefully acknowledge the support of the FWO–HAS mobility grant 'Substructures in finite projective spaces: algebraic and extremal questions'. The authors also wish to thank Jan De Beule and Leo Storme for the discussions during this work, and the anonymous referee for a clarifying comment.

†Supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by NKFIH OTKA Grant No. K124950.

E-mail addresses: heger@caesar.elte.hu (Tamás Héger), lihernan@vub.be (Lisa Hernandez Lucas)

Definition 1.1. Let $G = (V, E)$ be a graph. The *closed neighborhood* $N[S]$ of a set of vertices S is defined as the set of vertices adjacent to any vertex in S, joint with the vertices of S itself. A set $D \subseteq V$ is a *dominating set* if $N[D] = V$.

It is desirable to find the smallest dominating sets in a graph. The number of vertices in the smallest dominating set in a graph G is the *domination number of* G, and a common notation for it is $\Gamma(G)$.

The problem of domination has been studied before in incidence graphs of geometric structures, see for instance [6] and [11]. Also, perfect dominating sets of the incidence graphs of finite generalized quadrangles were considered in [4] (see also [9]), and for the particular quadrangle $Q(4, q)$, they were studied in detail in [2]; see Section 5 for further information. In this paper, we will consider dominating sets in the incidence graph of finite generalized quadrangles.

Generalized quadrangles were first introduced by Tits [14]. In [13], Payne and Thas give the following definition of finite generalized quadrangles:

Definition 1.2. A finite generalized quadrangle $\mathrm{GQ}(s, t)$ with parameters s and t, where $s, t \geq 1$, is a point-line incidence structure $(\mathcal{P}, \mathcal{B}, \mathrm{I})$, in which $\mathcal{P}$ is the set of points, $\mathcal{B}$ is the set of lines and I is a symmetric point-line incidence relation, satisfying the following axioms:

- Each point is incident with $t + 1$ lines and two distinct points are incident with at most one line.
- Each line is incident with $s+1$ points and two distinct lines are incident with at most one point.
- If x is a point and L is a line not incident with x, then there is a unique pair $(y, M) \in \mathcal{P} \times \mathcal{B}$ for which $x \mathrel{\mathrm{I}} M \mathrel{\mathrm{I}} y \mathrel{\mathrm{I}} L$.

We will refer to this third property as the *projection property*. A generalized quadrangle $\mathrm{GQ}(s, t)$ with parameters s and t is said to have *order* (s, t). It is well known that the number of points in a generalized quadrangle of order (s, t) is $(s + 1)(st + 1)$, and the number of lines is $(t + 1)(st + 1)$. For two points P and Q, we will write $P \sim Q$ if there exists a line incident with both (that is, they are collinear), and we will use this notation dually for lines as well.

For readability, we will refer to a dominating set in the incidence graph of $\mathrm{GQ}(s, t)$ and the domination number of the incidence graph of $\mathrm{GQ}(s, t)$ as *a dominating set in* $\mathrm{GQ}(s, t)$ and *the domination number of* $\mathrm{GQ}(s, t)$.

When viewed from a geometric perspective, a dominating set in $\mathrm{GQ}(s, t)$ becomes the union of a set of points and a set of lines such that each point which is not in the set is incident with a line from the set and such that each line which is not in the set is incident with a point from the set. This is closely related to the concept of *blocking sets* and the dual concept of *covers*.

Definition 1.3. A *blocking set* in $\mathrm{GQ}(s, t)$ is a set of points such that each line is incident with at least one of these points. A *cover* in $\mathrm{GQ}(s, t)$ is a set of lines such that each point is incident with at least one of these lines.

A blocking set $\mathcal{O}$ such that no two points from $\mathcal{O}$ are collinear is called an *ovoid*, a cover $\mathcal{S}$ such that no two lines from $\mathcal{S}$ are concurrent is called a *spread*. An arbitrary set of points $\mathcal{O}$ such that no two points from the set are collinear is called a *partial ovoid*, an arbitrary set of lines $\mathcal{S}$ such that no two lines from $\mathcal{S}$ are concurrent is called a *partial spread*. Let us note that an ovoid or a spread in $\mathrm{GQ}(s,t)$ contains exactly $st+1$ elements, and this is the smallest possibe size for a blocking set or a covering set.

In Section 2 we will show that there exist dominating sets in $\mathrm{GQ}(s,t)$ of size $2st+1$ and that the union of a blocking set and a cover exceeds this size. This motivates us to call a dominating set in $\mathrm{GQ}(s,t)$ *small* when it has size *at most* $2st+1$. This section also lists some properties of small dominating sets in $\mathrm{GQ}(s,t)$. In Section 3 we show that all small dominating sets in $\mathrm{GQ}(s,t)$, where $|s-t|\le 3$ must have size $2st+1$, which shows that the domination number in this case is $2st+1$. In particular, this gives us the domination number of $\mathrm{GQ}(q,q)$ and $\mathrm{GQ}(q-1,q+1)$. In Section 4 we give a classification of small dominating sets in $\mathrm{GQ}(q,q)$. In Section 5 we give a summary of the main results, and add some open problems.

2 Examples and properties of small dominating sets in $\mathrm{GQ}(s,t)$

Consider a generalized quadrangle $\mathrm{GQ}(s,t)$. We will construct a dominating set D of size $2st+1$ as follows. Let P be a point in $\mathrm{GQ}(s,t)$. Number the lines through P as $\ell_1,\ell_2,\dots,\ell_{t+1}$. Now define $\mathcal{P}$ as the set of all points which are incident with one of the first t lines $\ell_1,\dots,\ell_t$ through P, including P itself. Then $|\mathcal{P}|=st+1$. Now define $\mathcal{L}$ as the set of lines which intersect the last line ℓ_{t+1} in a point different from P. Then $|\mathcal{L}|=st$. Define $D=\mathcal{P}\cup\mathcal{L}$. The construction is also shown in Figure 1.

The size of D is $2st+1$. Now take an arbitrary point Q in $\mathrm{GQ}(s,t)$. If Q is incident with ℓ_{t+1}, then it is either contained in the dominating set (if $Q=P$), or covered by t different lines from the dominating set. So assume Q is not incident with ℓ_{t+1}. Then by the projection property of generalized quadrangles, there exists a unique point-line pair (R,m) such that R is incident with ℓ_{t+1} and m is incident with both Q and R. If m is one of the lines $\ell_1,\dots,\ell_t$, then Q is a point of the dominating set. Otherwise, m must be a line intersecting ℓ_{t+1} in a point different from P. Then m is in the dominating set and Q is covered.

Now take an arbitrary line ℓ. Assume ℓ is not incident with P. Then again by the projection property of generalized quadrangles, there exists a unique point-line pair (R,m) such that R is incident with ℓ and m is incident with both P and R. If R is incident with ℓ_{t+1}, then ℓ intersects ℓ_{t+1} in a point different from P, so it must be one of the lines in the dominating set. Otherwise, R is incident with one of the lines $\ell_1,\dots,\ell_t$, hence it is in the dominating set. In this case ℓ is blocked.

So D is indeed a dominating set of $\mathrm{GQ}(s,t)$. We now have Theorem 2.1.

Theorem 2.1. *For any finite generalized quadrangle* $\mathrm{GQ}(s,t)$ *there exists a dominating set of size* $2st+1$.

Note that the construction in Theorem 2.1 can be dualized, giving us a second example of a dominating set of size $2st+1$. We can also get this dual structure by omitting the point P from the dominating set and adding the line ℓ_{t+1} to it. See also Figure 1.

From a graph-theoretical point of view, we can get the dominating set from Theorem 2.1 or its dual as follows. Fix one edge $\{P,\ell\}$ in the incidence graph, then all points with distance two from $\{P,\ell\}$ together with $\{P\}$ (resp. together with $\{\ell\}$) form the dominating

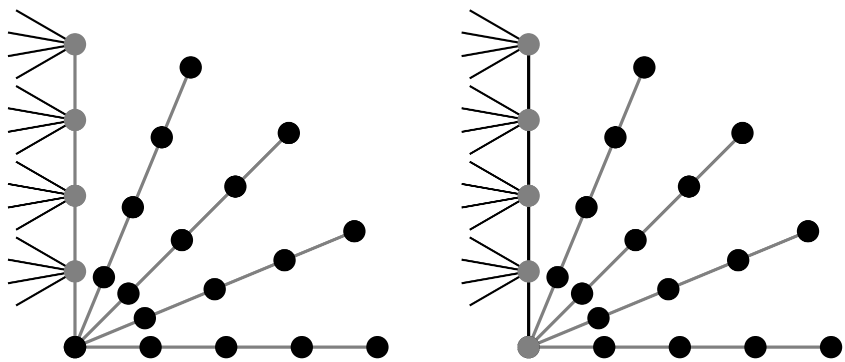


Figure 1: The dominating set from Theorem 2.1 (left) and its dual (right).

set from Theorem 2.1 (resp. its dual)[1].

Let us note that this dominating set is a maximal independent set in the incidence graph. Maximal independent sets are clearly dominating sets; however, the converse is not always true. Even when the dominating set has the smallest size possible, it is not necessarily a independent set. Families of graphs for which the smallest dominating sets are independent also form a subject of study.

In the case of non-thick generalized quadrangles; that is, when $s = 1$ or $t = 1$, we immediately have the following result.

Theorem 2.2. *The domination number of* $\mathrm{GQ}(q, 1)$ *and* $\mathrm{GQ}(1, q)$ *is* $2q + 1$*; furthermore, dominating sets of size* $2q + 1$ *are independent.*

Proof. By duality, it is sufficient to show this result for $\mathrm{GQ}(q, 1)$. The points and lines of $\mathrm{GQ}(q, 1)$ may naturally be viewed as the $(q + 1)^2$ points together with the $q + 1$ horizontal and $q + 1$ vertical lines of a $(q + 1) \times (q + 1)$ grid.

Let D be a dominating set of size $|D| \leq 2q+1$, and let l_v and l_h stand for the number of vertical and horizontal lines in D, respectively. If $l_h = q + 1$, then for each vertical line we must have either the line or a point on the line in D, hence $|D| \geq 2q + 2$, a contradiction. Thus $l_h \leq q$, and similarly $l_v \leq q$.

If $l_h = q$, then for all the $q + 1$ points on the horizontal line not in D, D must contain either the point or the vertical line through it, whence $|D| \geq 2q + 1$. Since we assumed $|D| \leq 2q + 1$, we now have $|D| = 2q + 1$. Moreover, note that D is independent in this case. A similar argument works for $l_v = q$.

Suppose now $l_h \leq q - 1$ and $l_v \leq q - 1$. The lines of D leave exactly $(q + 1 - l_h)(q + 1 - l_v)$ points not covered, which all must be in D, whence

$$(q + 1)^2 - (l_h + l_v)(q + 1) + l_h l_v + (l_h + l_v) \leq |D| \leq 2q + 1.$$

As $l_h \leq q-1$ and $l_v \leq q-1$, we have $l_v + l_h \leq 2q-2$ and $l_h l_v \geq (q-1)(l_h + l_v - (q-1))$, hence

$$2q + 1 \geq (q + 1)^2 - (l_h + l_v)q + l_h l_v \geq (q + 1)^2 - (l_h + l_v) - (q - 1)^2 \geq 4q - (2q - 2),$$

[1]The authors wish to thank Sam Mattheus (VUB) for this remark.

a contradiction.

We can conclude that the size of a dominating set D is at least $2q+1$, and that D is independent in this case. Theorem 2.1 assures that such a dominating set indeed exists. □

From now on, we will assume that our generalized quadrangle is thick; that is, $s \geq 2$ and $t \geq 2$.

The most trivial dominating sets are the union of a blocking set and a cover. If an ovoid and a spread exist, then their union forms a dominating set of size $2st+2$. So by Theorem 2.1, the union of an ovoid and a spread is definitely not the smallest dominating set. Moreover, we can prove that any dominating set containing a blocking set or a cover exceeds the size $2st+1$.

Lemma 2.3. *Let $D = \mathcal{P}_D \cup \mathcal{L}_D$ be a dominating set in $\mathrm{GQ}(s,t)$, $s,t \geq 2$, with $\mathcal{P}_D$ the points and $\mathcal{L}_D$ the lines of D. If $\mathcal{P}_D$ is a blocking set or $\mathcal{L}_D$ is a cover, then $|D| \geq 2st+2$. Moreover, if equality holds, then $\mathcal{P}_D$ is an ovoid and $\mathcal{L}_D$ is a spread.*

Proof. Assume without loss of generality that $\mathcal{P}_D$ is a blocking set. The case where $\mathcal{L}_D$ is a cover can be showed analogously. Assume $|D| \leq 2st+2$, then $|\mathcal{P}_D| \leq 2st+2-|\mathcal{L}_D|$. Any point that is not in $\mathcal{P}_D$ needs to be covered at least once by a line of $\mathcal{L}_D$. But since $\mathcal{P}_D$ is a blocking set, each line of $\mathcal{L}_D$ contains at least one point of $\mathcal{P}_D$, meaning it can only cover at most s points not in $\mathcal{P}_D$. This gives us the following inequality:

$$(s+1)(st+1) \leq |\mathcal{P}_D| + s|\mathcal{L}_D| \leq 2st+2-|\mathcal{L}_D| + s|\mathcal{L}_D|,$$

whence $(s-1)(st+1) \leq (s-1)|\mathcal{L}_D|$, and thus $st+1 \leq |\mathcal{L}_D|$ follows. Since $\mathcal{P}_D$ is a blocking set, we have that $|\mathcal{P}_D| \geq st+1$. But then $|D| = |\mathcal{P}| + |\mathcal{L}| \geq 2st+2$, so the lower bound on $|D|$ is proved.

Assume now that equality holds. Then $|\mathcal{P}_D| = |\mathcal{L}_D| = st+1$ and $\mathcal{L}_D$ covers each point not in D exactly once. We want to show that $\mathcal{L}_D$ covers each point of $\mathcal{P}_D$ exactly once as well. As $\mathcal{P}_D$ is a blocking set, its size implies it being an ovoid, so each line of $\mathcal{L}_D$ covers at most one point of $\mathcal{P}_D$. Suppose that there exists a point $P \in \mathcal{P}_D$ not covered by $\mathcal{L}_D$. Then each of the $(t+1)s$ points collinear with P, which are not in $\mathcal{P}_D$, must be covered by a line of $\mathcal{L}_D$ which, due to the projection property, are pairwise distinct, implying $|\mathcal{L}_D| \geq st+s > st+1$, a contradiction. Hence $\mathcal{L}_D$ is a cover of size $st+1$, that is, a spread. □

Lemma 2.3, together with Theorem 2.1, motivates the following definition:

Definition 2.4. Let D be a dominating set in $\mathrm{GQ}(s,t)$. Then D is a *small dominating set* if $|D| \leq 2st+1$.

The following two lemmas also provide us with some information regarding the size of a dominating set or, more precisely, regarding the size of the set of points and the set of lines contained in a dominating set.

Lemma 2.5. *Let $\mathcal{P}$ be an arbitrary point set in $\mathrm{GQ}(s,t)$. Assume there exists a number Δ such that*

$$\forall P \in \mathcal{P}\colon |\{Q \in \mathcal{P} \mid Q \neq P, Q \sim P\}| \geq \Delta.$$

Then the number of lines in $\mathrm{GQ}(s,t)$ not blocked by $\mathcal{P}$ is at least

$$(t+1)(st+1-|\mathcal{P}|) + \frac{|\mathcal{P}|\Delta}{s+1}.$$

Proof. For any line ℓ in $\mathrm{GQ}(s,t)$, define the degree $\mathrm{d}(\ell)$ of ℓ as the number of points in $\mathcal{P}$ which are incident with ℓ. Let $\mathcal{L}_b$ be the set of lines blocked by $\mathcal{P}$. Then

$$|\mathcal{L}_b| = |\mathcal{P}|(t+1) - \sum_{\ell\in\mathcal{L}_b}(\mathrm{d}(\ell)-1). \tag{2.1}$$

Now let $X := \{(P,Q) \mid P,Q \in \mathcal{P}, P \neq Q, P \sim Q\}$. Then we have the following inequalities:

$$|\mathcal{P}|\Delta \leq |X| = \sum_{\ell\in\mathcal{L}_b}\mathrm{d}(\ell)(\mathrm{d}(\ell)-1) \leq (s+1)\sum_{\ell\in\mathcal{L}_b}(\mathrm{d}(\ell)-1),$$

implying that $\frac{|\mathcal{P}|\Delta}{s+1} \leq \sum_{\ell\in\mathcal{L}_b}(\mathrm{d}(\ell)-1)$. Together with (2.1) this yields:

$$|\mathcal{L}_b| \leq |\mathcal{P}|(t+1) - \frac{|\mathcal{P}|\Delta}{s+1}.$$

Since the total number of lines in $\mathrm{GQ}(s,t)$ is $(t+1)(st+1)$ it follows that the number of lines not blocked by $\mathcal{P}$ is at least $(t+1)(st+1-|\mathcal{P}|) + \frac{|\mathcal{P}|\Delta}{s+1}$. □

If $\mathcal{P}$ is the point set of a dominating set, then Lemma 2.5 gives a lower bound on the number of lines contained in this dominating set. By dualizing this lemma we find:

Lemma 2.6. *Let $\mathcal{L}$ be an arbitrary line set in $\mathrm{GQ}(s,t)$. Assume there exists a number Δ such that*

$$\forall \ell \in \mathcal{L}\colon |\{m \in \mathcal{L} \mid m \neq \ell, m \sim \ell\}| \geq \Delta.$$

Then the number of points in $\mathrm{GQ}(s,t)$ not covered by the line set $\mathcal{L}$ is at least

$$(s+1)(st+1-|\mathcal{L}|) + \frac{|\mathcal{L}|\Delta}{t+1}.$$

Notation 2.7. Let D be a dominating set in $\mathrm{GQ}(s,t)$. Let $\mathcal{P}_D$ and $\mathcal{L}_D$ denote the point set and the line set of D, resp. Define $\mathcal{P}'$ and $\mathcal{L}'$ as the set of points and the set of lines resp., that are not covered by $\mathcal{L}_D$ and not blocked by $\mathcal{P}_D$ resp. We will use this notation in the sequel implicitly. Note that by the definition of a dominating set, $\mathcal{P}' \subseteq \mathcal{P}_D$ and $\mathcal{L}' \subseteq \mathcal{L}_D$.

The following Lemma allows us to apply Lemma 2.5.

Lemma 2.8. *Let D be a dominating set in $\mathrm{GQ}(s,t)$. For any point in $\mathcal{P}'$, the number of points in $\mathcal{P}'$ collinear with it is at least $\Delta_{\mathcal{P}} := st+s-|\mathcal{L}_D|$. For any line in $\mathcal{L}'$ the number of lines in $\mathcal{L}'$ concurrent with it is at least $\Delta_{\mathcal{L}} := st+t-|\mathcal{P}_D|$.*

Proof. Let P be an arbitrary point in $\mathcal{P}'$. Then each line of D covers at most one point collinear with P. Hence, there are at least $\Delta_{\mathcal{P}} := (t+1)s - |\mathcal{L}_D|$ points collinear with P which are not covered by a line of D. These points must be in $\mathcal{P}'$. So each point of $\mathcal{P}'$ is collinear with at least $\Delta_{\mathcal{P}}$ other points of $\mathcal{P}'$. Dually, we have that each line of $\mathcal{L}'$ is concurrent with at least $\Delta_{\mathcal{L}} := st+t-|\mathcal{P}_D|$ other lines of $\mathcal{L}'$. □

From the next lemma follows that $\Delta_{\mathcal{P}}$ and $\Delta_{\mathcal{L}}$ are non-negative.

Lemma 2.9. *Let D be a dominating set in* $\mathrm{GQ}(s,t)$*. If* $|D| \leq 2st+1$*, then*

$$st - s + 1 \leq |\mathcal{P}_D| \leq st + t, \quad (2.2)$$

$$st - t + 1 \leq |\mathcal{L}_D| \leq st + s. \quad (2.3)$$

Proof. Assume $|\mathcal{P}_D| = st + 1 - \epsilon$, then at least $\epsilon(t+1)$ lines are not blocked by $\mathcal{P}_D$ and have to be in $\mathcal{L}_D$. This implies that

$$2st + 1 \geq |D| = |\mathcal{P}_D| + |\mathcal{L}_D| \geq st + 1 - \epsilon + \epsilon(t+1),$$

from which follows that $\epsilon \leq s$. Hence $|\mathcal{P}_D| \geq st - s + 1$. From this we obtain

$$2st + 1 \geq |D| = |\mathcal{P}_D| + |\mathcal{L}_D| \geq st - s + 1 + |\mathcal{L}_D|,$$

hence $|\mathcal{L}_D| \leq st + s$. The other two inequalities follow similarly. □

3 The domination number of $\mathrm{GQ}(s,t)$, $|s-t|$ small

Theorem 3.1. *The domination number of* $\mathrm{GQ}(s,t)$*, where* $|s-t| \leq 3$*, is* $2st+1$*.*

Proof. By Theorem 2.1, it is enough to show $\Gamma(\mathrm{GQ}(s,t)) \geq 2st+1$. Assume a dominating set D exists with size smaller than $2st+1$. When lines or points are added to a dominating set, it still remains a dominating set, so without loss of generality we may assume that D has size $|D| = 2st$.

Let $l = |\mathcal{L}_D|$ and $p = |\mathcal{P}_D|$, and let $\Delta_\mathcal{P}$ and $\Delta_\mathcal{L}$ be as in Lemma 2.8. By Lemma 2.5 we have that the number of lines not blocked by $\mathcal{P}'$ is at least $(st+1-|\mathcal{P}'|)(t+1) + \frac{|\mathcal{P}'|\Delta_\mathcal{P}}{s+1}$. Since each point of D can block at most $t+1$ lines, the number of lines $|\mathcal{L}'|$ not blocked by D is at least

$$\begin{aligned} |\mathcal{L}'| &\geq (st+1-|\mathcal{P}'|)(t+1) + \frac{|\mathcal{P}'|\Delta_\mathcal{P}}{s+1} - (p - |\mathcal{P}'|)(t+1) \\ &= (st+1-p)(t+1) + \frac{|\mathcal{P}'|\Delta_\mathcal{P}}{s+1}. \end{aligned}$$

Dually, by Lemma 2.6, we find that

$$|\mathcal{P}'| \geq (st+1-l)(s+1) + \frac{|\mathcal{L}'|\Delta_\mathcal{L}}{t+1}.$$

Suppose, say, $l \leq p$ (we may consider the dual quadrangle otherwise). Let $0 \leq \epsilon \leq t$ be such that $p = st+\epsilon, l = st-\epsilon$ (cf. Lemma 2.9). Filling in these and $\Delta_\mathcal{P} = st+s-l = s+\varepsilon$ and $\Delta_\mathcal{L} = st+t-p = t-\varepsilon$, multiplying by $s+1$ and $t+1$ resp., and rearranging we get

$$(s+1)|\mathcal{L}'| - (s+\varepsilon)|\mathcal{P}'| \geq (1-\epsilon)(t+1)(s+1), \quad (3.1)$$

$$(t+1)|\mathcal{P}'| - (t-\varepsilon)|\mathcal{L}'| \geq (1+\epsilon)(t+1)(s+1). \quad (3.2)$$

Suppose $|\mathcal{P}'| \leq |\mathcal{L}'|$. Then (3.2) yields

$$(1+\epsilon)(t+1)(s+1) \leq (t+1)|\mathcal{P}'| - (t-\varepsilon)|\mathcal{L}'| \leq (1+\epsilon)|\mathcal{L}'| \leq (1+\epsilon)|\mathcal{L}_D| \leq (1+\epsilon)st,$$

a contradiction. Hence $|\mathcal{P}'| > |\mathcal{L}'|$. Then (3.1) gives

$$(1-\epsilon)(t+1)(s+1) \leq (s+1)|\mathcal{L}'| - (s+\varepsilon)|\mathcal{P}'| < (1-\epsilon)|\mathcal{P}'| \leq (1-\epsilon)(st+t),$$

a contradiction if $\epsilon \leq 1$; thus $\epsilon \geq 2$. Adding up (3.1) and (3.2) we find

$$2(t+1)(s+1) \leq (t-s-\epsilon+1)|\mathcal{P}'|-(t-s-\epsilon-1)|\mathcal{L}'| = (t-s-\epsilon-1)(|\mathcal{P}'|-|\mathcal{L}'|)+2|\mathcal{P}'|.$$

As $t - s \leq 3$, $\epsilon + 1 \geq 3$, $|\mathcal{P}'| - |\mathcal{L}'| > 0$ and $|\mathcal{P}'| \leq |\mathcal{P}_D| \leq st + t$, this is a contradiction. Consequently, $|D| > 2st$. □

Corollary 3.2. *The domination number of* $\mathrm{GQ}(q, q)$ *is* $2q^2+1$*, and the domination number of* $\mathrm{GQ}(q + 1, q - 1)$ *and* $\mathrm{GQ}(q - 1, q + 1)$ *is* $2q^2 - 1$*.*

This corollary applies to the well-known quadrangles $W(q)$, $Q(4, q)$, $T_2(O)$ (these have order (q, q)), $T_2^*(O)$ (of order $(q - 1, q + 1)$, q even), $\mathrm{AS}(q)$ (of order $(q - 1, q + 1)$, q odd) and their duals (of order $(q + 1, q - 1)$). Let us note that these quadrangles yield isomorphic incidence graphs in many cases. Clearly, the incidence graphs of a GQ and its dual are isomorphic. Let us now fix q. It is known that $W(q)$ is isomorphic to the dual of $Q(4, q)$, and that $T_2(O)$ is isomorphic to $Q(4, q)$ if and only if the oval O is a conic [13, Section 3.2], which is certainly the case when q is odd by B. Segre's celebrated result. However, when q is even, O may be an oval that is not a conic, in which case the construction $T_2(O)$ gives new instances of GQs of order (q, q) and corresponding incidence graphs. In case of order $(q - 1, q + 1)$, q even, there are also examples of GQs other than $T_2^*(O)$ [13].

4 Classification of the smallest dominating sets in $\mathrm{GQ}(q, q)$

Corollary 3.2 shows that all small dominating sets in $\mathrm{GQ}(q, q)$ have size $2q^2 + 1$. Moreover, we already have two constructions of small dominating sets, namely the construction from Theorem 2.1 and its dual. In this section we show that these are the only two small dominating sets.

First we need a few lemmas regarding the structure of small dominating sets in $\mathrm{GQ}(q, q)$.

Lemma 4.1. *Let* $D = \mathcal{P}_D \cup \mathcal{L}_D$ *be a dominating set in* $\mathrm{GQ}(q, q)$ *of size* $2q^2 + 1$*. Then* $\mathcal{P}'$ *is not a partial ovoid and* $\mathcal{L}'$ *is not a partial spread.*

Proof. It is sufficient to show that $\mathcal{P}'$ cannot be a partial ovoid. It then follows by duality that $\mathcal{L}'$ cannot be a spread. So assume to the contrary that $\mathcal{P}'$ is a partial ovoid, this will lead to a contradiction.

Take a point $P \in \mathcal{P}'$. Since $\mathcal{P}'$ is a partial ovoid, all points collinear with P are not in $\mathcal{P}'$. So they need to be covered by at least $q^2 + q$ different lines from $\mathcal{L}_D$. By Lemma 2.9 we now have that $|\mathcal{L}_D| = q^2 + q$ and $|\mathcal{P}_D| = q^2 - q + 1 \geq |\mathcal{P}'|$.

Now let η be the number of lines that are blocked by points of $\mathcal{P}_D \setminus \mathcal{P}'$, but are not in $\mathcal{L}_D$ and are not blocked by a point of $\mathcal{P}'$. Then we can count the total number of lines in $\mathrm{GQ}(q, q)$:

$$q^3 + q^2 + q + 1 = |\mathcal{L}_D| + (q + 1)|\mathcal{P}'| + \eta.$$

Note that $\eta \leq (q^2 - q + 1 - |\mathcal{P}'|)q$, since each point of $\mathcal{P}_D \setminus \mathcal{P}'$ is covered at least once by $\mathcal{L}_D$, so it contributes at most q lines to η. Remember that $|\mathcal{L}_D| = q^2 + q$. We now find that:

$$q^3 + q^2 + q + 1 \leq q^2 + q + (q + 1)|\mathcal{P}'| + (q^2 - q + 1 - |\mathcal{P}'|)q,$$

which implies that $|\mathcal{P}'| \geq q^2 - q + 1$. This means that $\mathcal{P}' = \mathcal{P}_D$. So all points of the dominating set are uncovered.

Since $|\mathcal{L}_D| > q^2 + 1$, the set of lines $\mathcal{L}_D$ cannot form a partial spread, meaning some of these lines must intersect. Assume there exist three lines $\ell_1, \ell_2, \ell_3 \in \mathcal{L}_D$ such that ℓ_1 intersects both ℓ_2 and ℓ_3, in different points. Each point $Q \in \mathcal{P}_D$ can be projected onto ℓ_1. The projection point needs to be different from the points where ℓ_2 and ℓ_3 intersect ℓ_1. Otherwise, there would not be enough lines in $\mathcal{L}_D$ to cover all points collinear to Q. Different points from $\mathcal{P}_D$ will have different projection lines, since no two points from $\mathcal{P}_D$ are collinear. This implies that $|\mathcal{P}_D| \leq (q-1)q = q^2 - q$, which is a contradiction. From this we can conclude that each line of $\mathcal{L}_D$ must cover at least q points which are not covered by any other line of $\mathcal{L}_D$.

We can now start counting again:

$$q^3 + q^2 + q + 1 > |\mathcal{P}_D| + q|\mathcal{L}_D| = q^2 - q + 1 + q(q^2 + q) = q^3 + 2q^2 - q + 1,$$

from which $q < 2$ follows, yielding an obvious contradiction. Hence, $\mathcal{P}'$ cannot be a partial ovoid. □

Note that since $\mathcal{P}' \subseteq \mathcal{P}_D$, this lemma implies that $\mathcal{P}_D$ cannot be a partial ovoid either and, dually, $\mathcal{L}_D$ is not a partial spread.

The following two theorems give a characterization for the dominating set constructed in Theorem 2.1, and its dual.

Theorem 4.2. *Let $D = \mathcal{P}_D \cup \mathcal{L}_D$ be a dominating set in $\mathrm{GQ}(q,q)$ of size $2q^2+1$. Let the degree $\mathrm{d}(\ell)$ of a line ℓ be the number of points of $\mathcal{P}'$ that are incident with ℓ. If all lines in $\mathrm{GQ}(q,q)$ have $\mathrm{d}(\ell) \in \{0, 1, q+1\}$, then D is the dominating set from Theorem 2.1.*

Proof. Suppose that every line ℓ of the $\mathrm{GQ}(q,q)$ admits $\mathrm{d}(\ell) \in \{0,1,q+1\}$. If there is no line with $\mathrm{d}(\ell) = q+1$, then $\mathcal{P}'$ is a partial ovoid, which is not possible according to Lemma 4.1. So there is at least one line with degree $q+1$. Then every point of $\mathcal{P}'$ must be contained in a line of degree at least two since either it is contained in a line of degree $q+1$ or it can be projected to one such line, and then the projection line has degree at least two. Since as soon as a line has degree at least two, it is completely contained in $\mathcal{P}'$ as a point set, this yields that $\mathcal{P}'$ can be obtained as the union of some lines.

Assume there are two non-intersecting lines ℓ and m contained (as a set of points) in $\mathcal{P}'$. Then each point of ℓ can be projected onto m. All these projection lines have degree at least two, so they are contained in $\mathcal{P}'$ as well (as point sets). But then $|\mathcal{P}_D| \geq (q+1)^2 > q^2+q$, which contradicts Lemma 2.9.

Hence, $\mathcal{P}'$ is a set of k lines through a point P. Note that $|\mathcal{P}'| = kq+1$, so by Lemma 2.9, $1 \leq k \leq q$. There are $q+1-k$ lines through P that, aside from P itself, do not contain points of $\mathcal{P}'$. So all points on these lines, except for P, must be covered. This leads to $(q+1-k)q$ lines from $\mathcal{L}_D$. These lines of $\mathcal{L}$ cover altogether at most $(q+1-k)q(q+1)$ points. The other lines of $\mathcal{L}_D$ can cover at most q points that are not covered yet by these first lines. This leads to the following inequality:

$$\begin{aligned} q^3 + q^2 + q + 1 &\leq kq + 1 + (q+1-k)q(q+1) + (|\mathcal{L}_D| - (q+1-k)q)\, q \\ &= kq + 1 + (q+1-k)q^2 + q^2 + q - kq + |\mathcal{L}_D|q - (q+1-k)q^2 \\ &= q^2 + q + 1 + |\mathcal{L}_D|q, \end{aligned}$$

hence $|\mathcal{L}_D| \geq q^2$.

The number of lines blocked by the elements of $\mathcal{P}'$ is $kq^2 + q + 1$. Consider a point Q in $\mathcal{P}_D \setminus \mathcal{P}'$. This point needs to be covered at least once by a line of $\mathcal{L}_D$. By projecting this point on one of the k lines through P, we see that there is at least one line through Q that is already blocked by a point from $\mathcal{P}'$. Hence, each point of $\mathcal{P}_D \setminus \mathcal{P}'$ can block at most $q-1$ lines that are not in $\mathcal{L}_D$ and are not blocked by a point of $\mathcal{P}'$. So for the size of this set we obtain

$$|\mathcal{P}_D \setminus \mathcal{P}'| \geq \frac{q^3+q^2+q+1-|\mathcal{L}_D|-(kq^2+q+1)}{q-1} = \frac{q^3+q^2-kq^2-|\mathcal{L}_D|}{q-1}.$$

Using this inequality and the observation that $|\mathcal{L}_D| \geq q^2$, we find for the size of the dominating set $D = (\mathcal{P}_D \setminus \mathcal{P}') \cup \mathcal{P}' \cup \mathcal{L}_D$ that

$$\begin{aligned}
|D| &\geq \frac{q^3+q^2-kq^2-|\mathcal{L}_D|}{q-1} + kq + 1 + |\mathcal{L}_D| \\
&= \frac{q^3+q^2-kq^2}{q-1} + kq + 1 + \left(1 - \frac{1}{q-1}\right)|\mathcal{L}_D| \geq \frac{q^3-kq^2}{q-1} + kq + 1 + q^2 \\
&= q^2 + (1-k)q + (1-k) + \frac{1-k}{q-1} + kq + 1 + q^2 \\
&= 2q^2 + q + 2 - k - \frac{k-1}{q-1}.
\end{aligned}$$

Now assuming $k < q$, we find that $|D| > 2q^2 + 1$, which is a contradiction. So the only possibility left is $k = q$. In this case $\mathcal{P}'$ consists of the points on q lines through P, and $|\mathcal{P}'| = q^2 + 1$. The number of lines blocked by these points is $q^3 + q + 1$. Since $|\mathcal{L}_D| \geq q^2$, all lines not blocked by the points of $\mathcal{P}'$ must be in $\mathcal{L}_D$. So D is the dominating set constructed in Theorem 2.1. □

Dualizing this theorem gives us a characterization for the dual of the construction in Theorem 2.1.

Theorem 4.3. *Let $D = \mathcal{P}_D \cup \mathcal{L}_D$ be a dominating set in $\mathrm{GQ}(q,q)$ of size $2q^2+1$. Let the degree $\mathrm{d}(P)$ of a point P be the number of lines of $\mathcal{L}'$ that are incident with P. If all points in $\mathrm{GQ}(q,q)$ have $\mathrm{d}(P) \in \{0, 1, q+1\}$, then D is the dual dominating set from Theorem 2.1.*

We will need the following lemma, which is actually a variation on Lemma 2.5.

Lemma 4.4. *Let $D = \mathcal{P}_D \cup \mathcal{L}_D$ be a dominating set in $\mathrm{GQ}(q,q)$ of size $2q^2+1$. Let $p := |\mathcal{P}_D|$, $l := |\mathcal{L}_D|$, and define $\Delta_{\mathcal{P}} = q^2 + q - l$ and $\Delta_{\mathcal{L}} = q^2 + q - p$ as in Lemma 2.8. Define the degree $\mathrm{d}(\ell)$ of a line ℓ as the number of points from $\mathcal{P}_D$ incident with ℓ, and the degree $\mathrm{d}(P)$ of a point P as the number of lines from $\mathcal{L}_D$ incident with P. Finally, we introduce*

$$c(D) = \sum_{Q \notin \mathcal{P}_D} (q+1-\mathrm{d}(Q))(\mathrm{d}(Q)-1) + \sum_{\ell \notin \mathcal{L}_D} (q+1-\mathrm{d}(\ell))(\mathrm{d}(\ell)-1) + \sum_{P \in \mathcal{P}_D} \mathrm{d}(P) + \sum_{\ell \in \mathcal{L}_D} \mathrm{d}(\ell).$$

Then

$$p \geq (q^2+1-l)(q+1) + \frac{1}{q+1}\left(l\Delta_{\mathcal{L}} + c(D)\right),$$
$$l \geq (q^2+1-p)(q+1) + \frac{1}{q+1}\left(p\Delta_{\mathcal{P}} + c(D)\right).$$

Proof. Let $p' := |\mathcal{P}'|$ and $l' := |\mathcal{L}'|$. Note that for a line ℓ, $\mathrm{d}(\ell) \geq 1$ iff $\ell \notin \mathcal{L}'$. With $p(q+1)$, we count each line ℓ blocked by $\mathcal{P}_D$ exactly $\mathrm{d}(\ell)$ times, hence the number of lines blocked by $\mathcal{P}_D$ is

$$\begin{aligned} &p(q+1) - \sum_{\ell\notin\mathcal{L}_D}(\mathrm{d}(\ell)-1) - \sum_{\ell\in\mathcal{L}_D\setminus\mathcal{L}'}(\mathrm{d}(\ell)-1) \\ &= p(q+1) - \sum_{\ell\notin\mathcal{L}_D}(\mathrm{d}(\ell)-1) - \left(\sum_{\ell\in\mathcal{L}_D}\mathrm{d}(\ell) - l + l'\right). \end{aligned} \tag{4.1}$$

Recall that l' equals the number of lines not blocked by $\mathcal{P}_D$, hence $l' = (q^2+1)(q+1) -$ (4.1). From this it follows that

$$l = (q^2+1-p)(q+1) + \sum_{\ell\notin\mathcal{L}_D}(\mathrm{d}(\ell)-1) + \sum_{\ell\in\mathcal{L}_D}\mathrm{d}(\ell). \tag{4.2}$$

We will estimate the middle term using

$$(q+1)\sum_{\ell\notin\mathcal{L}_D}(\mathrm{d}(\ell)-1) = \sum_{\ell\notin\mathcal{L}_D}\mathrm{d}(\ell)(\mathrm{d}(\ell)-1) + \sum_{\ell\notin\mathcal{L}_D}(q+1-\mathrm{d}(\ell))(\mathrm{d}(\ell)-1). \tag{4.3}$$

Note that the second sum on the right-hand side is a part of $c(D)$. For the sum $\sum_{\ell\notin\mathcal{L}_D}\mathrm{d}(\ell)(\mathrm{d}(\ell)-1)$ we can find a lower bound as follows.

For $P \in \mathcal{P}_D$, let $N'(P)$ denote the number of points of $\mathcal{P}'$ collinear with P. Then

$$\begin{aligned} \sum_{\ell\notin\mathcal{L}_D}\mathrm{d}(\ell)(\mathrm{d}(\ell)-1) &= |\{(P,Q,\ell)\colon \ell\notin\mathcal{L}_D, P\in\mathcal{P}_D, Q\in\mathcal{P}_D, P\sim Q, PQ=\ell\}| \\ &\geq |\{(P,Q,\ell)\colon \ell\notin\mathcal{L}_D, P\in\mathcal{P}_D, Q\in\mathcal{P}', P\sim Q, PQ=\ell\}| \\ &= |\{(P,Q)\colon P\in\mathcal{P}_D, Q\in\mathcal{P}', P\sim Q\}| = \sum_{P\in\mathcal{P}_D} N'(P). \end{aligned}$$

Let $P \in \mathcal{P}_D$. Then we have

$$\begin{aligned} l &\geq \mathrm{d}(P) + \sum_{\substack{Q\sim P\\ PQ\notin\mathcal{L}_D}}\mathrm{d}(Q) = \mathrm{d}(P) + \sum_{\substack{Q\sim P\\ Q\notin\mathcal{P}'\\ PQ\notin\mathcal{L}_D}}\mathrm{d}(Q) \\ &= \mathrm{d}(P) + (q+1-\mathrm{d}(P))q - N'(P) + \sum_{\substack{Q\sim P\\ Q\notin\mathcal{P}'\\ PQ\notin\mathcal{L}_D}}(\mathrm{d}(Q)-1), \end{aligned}$$

whence

$$N'(P) \geq \Delta_{\mathcal{P}} - (q-1)\mathrm{d}(P) + \sum_{\substack{Q\sim P\\ Q\notin\mathcal{P}'\\ PQ\notin\mathcal{L}_D}}(\mathrm{d}(Q)-1).$$

With this we find

$$\begin{aligned}\sum_{P\in\mathcal{P}_D} N'(P) &\geq p\Delta_{\mathcal{P}} - (q-1)\sum_{P\in\mathcal{P}_D} \mathrm{d}(P) + \sum_{P\in\mathcal{P}_D}\sum_{\substack{Q\sim P\\ Q\notin\mathcal{P}'\\ PQ\notin\mathcal{L}_D}} (\mathrm{d}(Q)-1)\\ &\geq p\Delta_{\mathcal{P}} - (q-1)\sum_{P\in\mathcal{P}_D} \mathrm{d}(P) + \sum_{P\in\mathcal{P}_D}\sum_{\substack{Q\sim P\\ Q\notin\mathcal{P}_D\\ PQ\notin\mathcal{L}_D}} (\mathrm{d}(Q)-1)\\ &= p\Delta_{\mathcal{P}} - (q-1)\sum_{P\in\mathcal{P}_D} \mathrm{d}(P) + \sum_{Q\notin\mathcal{P}_D}\sum_{\substack{P\in\mathcal{P}_D\\ Q\sim P\\ PQ\notin\mathcal{L}_D}} (d(Q)-1).\end{aligned}$$

As for each point $Q \notin \mathcal{P}_D$, there are $q+1-\mathrm{d}(Q)$ lines on Q that are not in $\mathcal{L}_D$ and each of these must be incident with a point of $\mathcal{P}_D$, we find that

$$\sum_{P\in\mathcal{P}_D} N'(P) \geq p\Delta_{\mathcal{P}} - (q-1)\sum_{P\in\mathcal{P}_D} \mathrm{d}(P) + \sum_{Q\notin\mathcal{P}_D} (q+1-\mathrm{d}(Q))(\mathrm{d}(Q)-1).$$

As $\sum_{\ell\in\mathcal{L}_D} \mathrm{d}(\ell) = \sum_{P\in\mathcal{P}_D} \mathrm{d}(P)$, we conclude

$$\sum_{\ell\notin\mathcal{L}_D} \mathrm{d}(\ell)(\mathrm{d}(\ell)-1) \geq p\Delta_{\mathcal{P}} - (q-1)\sum_{\ell\in\mathcal{L}_D} \mathrm{d}(\ell) + \sum_{Q\notin\mathcal{P}_D} (q+1-\mathrm{d}(Q))(\mathrm{d}(Q)-1).$$

Together with (4.2) and (4.3), this gives the second desired inequality. The other inequality is showed analogously. □

Note that $\mathcal{P}' = \mathcal{P}_D$ and $\mathcal{L}' = \mathcal{L}_D$ are equivalent. Also note that if this is the case, then $\sum_{\ell\in\mathcal{L}_D} \mathrm{d}(\ell) = 0$. We can now prove the following Theorem, giving a classification of the small dominating sets in $\mathrm{GQ}(q,q)$.

Theorem 4.5. *Let $D = \mathcal{P}_D \cup \mathcal{L}_D$ be a dominating set in $\mathrm{GQ}(q,q)$ with size $|D| = 2q^2+1$. Then D is the dominating set from Theorem 2.1 or its dual.*

Proof. Define $p = |\mathcal{P}_D|$, $l = |\mathcal{L}_D|$ and $p' = |\mathcal{P}'|$; note that $p + l = 2q^2 + 1$. By duality, we may assume that $p > l + 1$ or $p = q^2$ (and $l = q^2 + 1$). Define the degree $\mathrm{d}(\ell)$ of a line ℓ as the number of points from $\mathcal{P}_D$ incident with ℓ, and the degree $\mathrm{d}(P)$ of a point P as the number of lines from $\mathcal{L}_D$ incident with P. We will find lower bounds on the sums from Lemma 4.4; let $c(D)$ be defined as therein.

Define $\Delta := \Delta_{\mathcal{P}} = q^2 + q - l$ as in Lemma 2.8. For any point $P \in \mathcal{P}_D$, define the number of neighbors $N(P) := |\{Q \mid Q \sim P, Q \in \mathcal{P}_D\}|$. We immediately have that $N(P) \geq \Delta$ if $P \in \mathcal{P}'$. If $l = q^2+1$, then $\Delta = q-1$. If $p > l+1$, then $q^2-q+1 \leq l < q^2$, by Lemma 2.9. In both cases, we have that $\Delta \not\equiv 0 \pmod q$. We now consider two types of points in $\mathcal{P}'$ and their contributions to $c(D)$.

- Type 1: P is incident with at least one line e with $2 \leq \mathrm{d}(e) \leq q$.

  Since P is a point from $\mathcal{P}'$, the line e is not in $\mathcal{L}_D$. So this line e contributes at least $q-1$ to $\sum_{\ell\notin\mathcal{L}_D} (q+1-\mathrm{d}(\ell))\,(\mathrm{d}(\ell)-1)$. Note that on this line there are at most

q points of Type 1. Assume there are k points of Type 1, then we find the following lower bound:

$$\sum_{\ell \notin \mathcal{L}_D} (q+1-\mathrm{d}(\ell))\,(\mathrm{d}(\ell)-1) \geq k\frac{q-1}{q}. \tag{4.4}$$

- Type 2: All lines through P have degree 1 or $q+1$.

  If $\Delta = N(P)$, then there are exactly Δ points in $\mathcal{P}_D$ collinear with P. But $\Delta \not\equiv 0 \pmod{q}$, so then there must be at least one line ℓ through P with degree $2 \leq \mathrm{d}(\ell) \leq q$. Then P would be a point of Type 1. So we have $N(P) > \Delta$.

  Denote by $x_i(P) := |\{R \mid R \notin \mathcal{P}_D, R \sim P, \mathrm{d}(R) = i\}|$, for $i = 1, \dots, q$. Note that, as $P \in \mathcal{P}'$ there are no points collinear with P with degree $q+1$.

  Each point $R \notin \mathcal{P}_D$, with degree $1 \leq i \leq q$ contributes $(q+1-i)(i-1)$ to the sum $\sum_{Q \notin \mathcal{P}_D} (q+1-\mathrm{d}(Q))\,(\mathrm{d}(Q)-1)$. Such a point is collinear with at most $q+1-i$ points of Type 2, since a line through a point of Type 2 is not in $\mathcal{L}_D$, and either contains no other points of $\mathcal{P}_D$ or contains only points of $\mathcal{P}_D$. So the contribution of P to $\sum_{Q \notin \mathcal{P}_D} (q+1-\mathrm{d}(Q))\,(\mathrm{d}(Q)-1)$ is at least

$$\sum_{i=1}^{q} x_i \frac{(q+1-i)(i-1)}{q+1-i} = \sum_{i=1}^{q} x_i(i-1).$$

  Now we show that the contribution is strictly positive for each point of Type 2. So assume this is not the case for a point $P \in \mathcal{P}'$ of Type 2, so each point $Q \notin \mathcal{P}_D$ collinear with P has degree $\mathrm{d}(Q) = 1$. Since each line through P has either degree 1 or $q+1$, there must be λ lines through P which contain all points collinear with it from $\mathcal{P}$, for some $1 \leq \lambda \leq q+1$ ($\lambda > 0$ as $N(P) > 0$). So we already have $\lambda q + 1$ points in the dominating set.

  On each of the other $q+1-\lambda$ lines through P there are q points with degree 1. This gives us $q(q+1-\lambda)$ lines in the dominating set. Through each of these points $\notin \mathcal{P}_D$, collinear with P there are $q-1$ lines which are not in the dominating set and are not blocked yet. Say there are x points in D, which are not collinear with P and different from P itself. Each of these points can block at most $q+1-\lambda$ lines which are not yet blocked. From this follows $(q-1)q(q+1-\lambda) \leq x(q+1-\lambda)$, hence $x \geq q^2 - q$. Since $2q^2 + 1 = |D| \geq \lambda q + 1 + x + (q+1-\lambda)q = q^2 + q + 1 + x$, we have $x = q^2 - q$. As $q^2 + q \geq |\mathcal{P}_D| = x + \lambda q + 1 = q^2 + (\lambda - 1)q + 1$, we have $\lambda = 1$ and $|\mathcal{P}_D| = q^2 + 1$, contrary to our assumptions.

  From this follows that we may assume that for each point P of Type 2 we have $\sum_{i=1}^{q} x_i(i-1) \geq 1$. Note that there are $p' - k$ points of Type 2. This gives us the following inequality:

$$\sum_{Q \notin \mathcal{P}_D} (q+1-\mathrm{d}(Q))\,(\mathrm{d}(Q)-1) \geq p' - k. \tag{4.5}$$

Note that $\mathrm{d}(P) \geq 1$ for each $P \in \mathcal{P}_D \setminus \mathcal{P}'$, so $\sum_{\ell \in \mathcal{L}_D} \mathrm{d}(\ell) = \sum_{P \in \mathcal{P}_D} \mathrm{d}(P) \geq p - p'$.

Combining this with (4.4) and (4.5), and with $k \leq p' \leq p$ and $p \geq q^2$, we get the following:

$$\begin{aligned} c(D) &= \sum_{\ell \notin \mathcal{L}_D} (q+1-\mathrm{d}(\ell))\,(\mathrm{d}(\ell)-1) \\ &+ \sum_{Q \notin \mathcal{P}_D} (q+1-\mathrm{d}(Q))\,(\mathrm{d}(Q)-1) + 2\sum_{P \in \mathcal{P}_D} \mathrm{d}(P) \\ &\geq k\frac{q-1}{q} + p' - k + 2(p-p') \geq k\frac{q-1}{q} + p - k \\ &\geq p - \frac{k}{q} \geq p\frac{q-1}{q} \geq q^2 - q. \end{aligned}$$

Thus, according to Lemma 4.4, we find

$$\begin{aligned} p &\geq (q^2+1-l)(q+1) + \frac{1}{q+1}\left(l\Delta_l + q^2 - q\right), \\ l &\geq (q^2+1-p)(q+1) + \frac{1}{q+1}\left(p\Delta_p + q^2 - q\right). \end{aligned}$$

Now using $2q^2+1 = |D| = p + l$, $\Delta_{\mathcal{P}} = q^2+q-l$, $\Delta_{\mathcal{L}} = q^2+q-p$, and that $pl \leq q^2(q^2+1)$, we calculate the sum of these two inequalities:

$$\begin{aligned} 2q^2+1 &\geq q+1+\frac{p(q^2+q-l)+l(q^2+q-p)+2(q^2-q)}{q+1} \\ &= q+1+\frac{(p+l)q(q+1)-2pl+2(q^2-q)}{q+1} \\ &= q+1+q(2q^2+1)+\frac{-2pl+2(q^2-q)}{q+1} \\ &\geq 2q^3+2q+1+\frac{-2q^4-2q^2+2q^2-2q}{q+1} \\ &\geq 2q^3+2q+1-2q\frac{q^3+1}{q+1} \\ &\geq 2q^3+2q+1-2q(q^2-q+1) = 2q^2+1. \end{aligned}$$

So we see that we actually reach equality. This means that all the estimates we used during our countings were exact, hence we have $k = p' = p$. As $\mathcal{P}_D = \mathcal{P}'$, every point $P \in \mathcal{P}_D$ has $\mathrm{d}(P) = 0$. Equality with zero in (4.5) yields that for all $Q \notin \mathcal{P}_D$ we have $\mathrm{d}(Q) = 1$ or $\mathrm{d}(Q) = q+1$. By Theorem 4.3, D is the dual of the dominating set from Theorem 2.1. □

5 Conclusion, remarks and open problems

The main results of this paper can be summarized as follows.

Theorem 5.1.

- *The domination number of* $\mathrm{GQ}(s,t)$ *is at most* $2st+1$.
- *The domination number of* $\mathrm{GQ}(q,q)$ *equals* $2q^2+1$.
- *A dominating set of* $\mathrm{GQ}(q,q)$ *of size* $2q^2+1$ *is one of the two examples seen in Figure 1.*

- *The domination number of* $\mathrm{GQ}(q-1, q+1)$ *and* $\mathrm{GQ}(q+1, q-1)$ *equals* $2q^2 - 1$.

Let us outline some possible areas of further investigation in this topic. For the other classical parameters, the calculations in Lemma 3.1 fail in general. Is the bound $2st + 1$ for $\mathrm{GQ}(s, t)$ still sharp for general s and t? It would be also interesting to answer this question for the classical generalized quadrangles. We have checked by computer, using a simple linear integer programming model and Gurobi [7], for all classical GQs of order $(3, 9)$, $(4, 8)$, $(4, 16)$ as given on Moorhouse's webpage [12], and we have found the bound sharp.

In $W(q)$, or $Q(4, q)$, if q is even, there exists an ovoid and a spread as well, giving rise to a dominating set of size $2q^2 + 2$. This implies that there is no general stability phenomenon for smallest dominating sets in GQs, unlike in the case of generalized triangles (projective planes; see [11]); that is, the size of minimal examples (with respect to containment) may be arbitrarily close. However, the structure of the mentioned dominating sets are immensely dissimilar. Is it true that minimal dominating sets of size $2q^2 + 2$ of a $\mathrm{GQ}(q, q)$ (or, more specifically, $W(q)$) are the union of an ovoid and a spread? What size does the next smallest minimal example for a dominating set in $\mathrm{GQ}(q, q)$ have?

Dominating sets of a graph $G = (V, E)$ may be also viewed as a set D of vertices such that $V \setminus D$ induces a subgraph of G such that every vertex has degree at least one smaller than originally. If equality holds for every vertex of $V \setminus D$, then D is called a perfect dominating set. More generally, if we replace 'at least one smaller' by 'at least t smaller', we talk about t-fold dominating sets and perfect t-fold dominating sets. Perfect t fold dominating sets of the incidence graphs of projective planes, generalized quadrangles and generalized hexagons have already been studied in order to produce upper bounds on the order of some particular cage graphs (see [9] for an overview). In [5], perfect t-fold dominating sets of the incidence graph of the desarguesian projective plane $\mathrm{PG}(2, q)$ are completely described for small enough t, while the characterization of small dominating sets of projective planes can be found in [11]. In the case $t = 1$, describing smallest dominating and perfect dominating sets is quite easy, unlike in the here discussed case of generalized quadrangles; see also [2] for results on perfect (1-fold) dominating sets of $Q(4, q)$. It would be also interesting to study t-fold (ordinary and perfect) dominating sets of generalized quadrangles, $t \geq 2$. Also, as a counterpart of t-fold dominating sets, finding a (preferably large) subset D of the incidence graph such that each vertex not in D has at most t neighbors in D would be also interesting, as its complement induces a subgraph of high minimum degree. Such subsets, asides being interesting in themselves, also may find their applications in different topics as they do when the host graph is the incidence graph of a projective plane; see [1, 10].

ORCID iDs

Tamás Héger https://orcid.org/0000-0002-9586-966X
Lisa Hernandez Lucas https://orcid.org/0000-0002-1356-5970

References

[1] G. Araujo-Pardo and C. Balbuena, Constructions of small regular bipartite graphs of girth 6, *Networks* **57** (2011), 121–127, doi:10.1002/net.20392.

[2] L. Beukemann and K. Metsch, Regular graphs constructed from the classical generalized quadrangle $Q(4, q)$, *J. Combin. Des.* **19** (2011), 70–83, doi:10.1002/jcd.20266.

[3] E. J. Cockayne and S. T. Hedetniemi, Towards a theory of domination in graphs, *Networks* **7** (1977), 247–261, doi:10.1002/net.3230070305.

[4] A. Gács and T. Héger, On geometric constructions of (k, g)-graphs, *Contrib. Discrete Math.* **3** (2008), 63–80, doi:10.11575/cdm.v3i1.61998.

[5] A. Gács, T. Héger and Zs. Weiner, On regular graphs of girth six arising from projective planes, *European J. Combin.* **34** (2013), 285–296, doi:10.1016/j.ejc.2012.07.005.

[6] F. Goldberg, D. Rajendraprasad and R. Mathew, Domination in designs, arXiv:1405.3436 [math.CO].

[7] Gurobi Optimization, LLC, Gurobi Optimizer (Version 8.1), 2020, https://www.gurobi.com.

[8] S. T. Hedetniemi and R. C. Laskar, Bibliography on domination in graphs and some basic definitions of domination parameters, *Discrete Math.* **86** (1990), 257–277, doi:10.1016/0012-365x(90)90365-o.

[9] T. Héger, *Some graph theoretic aspects of finite geometries*, Ph.D. thesis, ELTE Eötvös Loránd University, 2013.

[10] T. Héger and S. Mattheus, Bipartite Ramsey numbers for the four-cycle versus stars emerging from projective planes, 2020, manuscript.

[11] T. Héger and Z. L. Nagy, Dominating sets in projective planes, *J. Combin. Des.* **25** (2017), 293–309, doi:10.1002/jcd.21527.

[12] E. Moorhouse, Generalised Polygons of Small Order, 2003, http://ericmoorhouse.org/pub/genpoly/.

[13] S. E. Payne and J. A. Thas, *Finite Generalized Quadrangles*, EMS Series of Lectures in Mathematics, European Mathematical Society (EMS), Zürich, 2nd edition, 2009, doi:10.4171/066.

[14] J. Tits, Sur la trialité et certains groupes qui s'en déduisent, *Inst. Hautes Études Sci. Publ. Math.* **2** (1959), 13–60, http://www.numdam.org/item?id=PMIHES_1959__2__13_0.

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 77–82
https://doi.org/10.26493/1855-3974.2301.63f
(Also available at http://amc-journal.eu)

Counterexamples to "A conjecture on induced subgraphs of Cayley graphs"*

Florian Lehner †

Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria

Gabriel Verret ‡

Department of Mathematics, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

Received 3 April 2020, accepted 28 May 2020, published online 7 November 2020

Abstract

Recently, Huang gave a very elegant proof of the Sensitivity Conjecture by proving that hypercube graphs have the following property: every induced subgraph on a set of more than half the vertices has maximum degree at least $\sqrt{d}$, where d is the valency of the hypercube. This was generalised by Alon and Zheng who proved that every Cayley graph on an elementary abelian 2-group has the same property. Very recently, Potechin and Tsang proved an analogous results for Cayley graphs on abelian groups. They also conjectured that all Cayley graphs have the analogous property. We disprove this conjecture by constructing various counterexamples, including an infinite family of Cayley graphs of unbounded valency which admit an induced subgraph of maximum valency 1 on a set of more than half the vertices.

Keywords: Cayley graphs, vertex-transitive graphs, sensitivity conjecture.

Math. Subj. Class. (2020): 05E18

1 Introduction

All graphs and groups in this paper are finite. Recently, Huang [5] gave a very elegant proof of the Sensitivity Conjecture [6] by proving that hypercube graphs have the following

*The authors would like to thank the anonymous referees for a number of helpful suggestions.

†Supported by the Austrian Science Fund (FWF), grants J 3850-N32 and P 31889-N35.

‡Supported by the N.Z. Marsden Fund, grant UOA1824.

E-mail addresses: mail@florian-lehner.net (Florian Lehner), g.verret@auckland.ac.nz (Gabriel Verret)

property: every induced subgraph on a set of more than half the vertices has maximum degree at least $\sqrt{d}$, where d is the valency of the hypercube. This is best possible, as shown by Chung, Füredi, Graham and Seymour [2]. This was generalised by Alon and Zheng who proved that every Cayley graph on an elementary abelian 2-group has the same property [1]. In their concluding remarks, they point out that this result cannot generalise directly to other groups, but that it would be interesting to investigate possible analogs for Cayley graphs on other groups.

Very recently, Potechin and Tsang proved such an analogous result for Cayley graphs on all abelian groups [7], by replacing the $\sqrt{d}$ bound by $\sqrt{d/2}$. (More precisely, they prove their result with the bound $\sqrt{x + x'/2}$, where x is the number of involutions in the connection set of the Cayley graph, and x' the number of non-involutions.) They also conjectured that their result should hold for all Cayley graphs [7, Conjecture 1] and asked whether even all vertex-transitive graphs might have this property.

In this short note, we give three infinite families of vertex-transitive graph such that every graph in these families admits an induced subgraph of maximum valency 1 on a set of more than half the vertices. First is the well-known family of odd graphs. Note that this family has unbounded valency and so these graphs fail to have the required property in a very strong sense. On the other hand, they are not Cayley. The second family consists of some 3-regular Cayley graphs on dihedral groups. The last family consists of an infinite family of graphs of unbounded valency which are Cayley on groups defined via iterated wreath products. Both the latter families are thus counterexamples to [7, Conjecture 1]. (We also note that, in our first two families, the induced subgraph in question is a matching, that is, each vertex has valency 1.)

2 Odd graphs

For $n \geq 1$, the *odd graph* O_n has vertex-set the n-subsets of a $(2n+1)$-set Ω, with two vertices adjacent if and only if the corresponding subsets are disjoint. For example $\mathrm{O}_1 \cong \mathrm{K}_3$, while O_2 is isomorphic to the Petersen graph. Note that there is an obvious action of S_{2n+1} as a vertex-transitive group of automorphism of O_n, so these graphs are vertex-transitive. On the other hand, these graphs are not Cayley graphs for $n \geq 2$ [4]. It is easy to check that O_n is $(n+1)$-regular. Let $\omega \in \Omega$ and let U be the set of n-subsets of Ω that do not contain ω. Note that

$$\frac{|U|}{|\mathrm{V}(\mathrm{O}_n)|} = \frac{\binom{2n}{n}}{\binom{2n+1}{n}} = \frac{n+1}{2n+1} > \frac{1}{2},$$

but the induced subgraph on U is 1-regular.

3 3-valent Cayley graphs on dihedral groups

For a group G and an inverse-closed and identity-free subset S of G, the *Cayley graph* $\mathrm{Cay}(G, S)$ is the graph with vertex-set G and two vertices g and h adjacent if and only $g^{-1}h \in S$. For $n \geq 1$, we denote by D_{2n} the dihedral group $\langle a, b \mid a^n = b^2 = (ab)^2 = 1\rangle$ of order $2n$.

Let $\Gamma_{18} = \mathrm{Cay}(\mathrm{D}_{18}, \{b, ab, a^3b\})$. There is a set U_{18} of 10 vertices of Γ_{18} such that the induced subgraph on U_{18} is 1-regular, as can be seen on the picture below, where the elements of U_{18} are coloured gray.

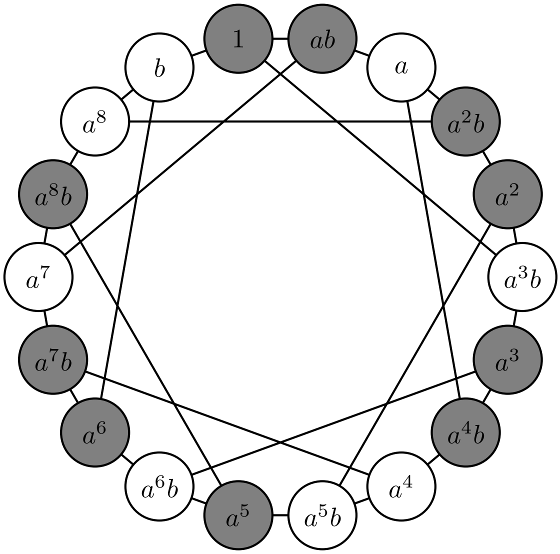


Figure 1: $\Gamma_{18} = \mathrm{Cay}(\mathrm{D}_{18}, \{b, ab, a^3b\})$.

Starting from Γ_{18}, it is easy to get an infinite family of further examples using covers: for $m \geq 1$, let $\Gamma_{18m} = \mathrm{Cay}(\mathrm{D}_{18m}, \{b, ab, a^3b\})$ and let N be the subgroup of D_{18m} generated by a^9. Note that N is a normal subgroup of D_{18m} of order m and $\mathrm{D}_{18m}/N \cong \mathrm{D}_{18}$. Let $\varphi\colon \mathrm{D}_{18m} \to \mathrm{D}_{18}$ be the natural projection and let U_{18m} be the preimage of U_{18}. Now, $|U_{18m}| = 10m$ and, since Γ_{18m} is a (normal) cover of Γ_{18}, the induced subgraph on U_{18m} is 1-regular.

4 Cayley graphs on iterated wreath products

Let $\mathbb{Z}_2 = \{0, 1\}$ denote the cyclic group of order 2, let G be a group with identity element 1_G and let $(\mathbb{Z}_2)^G$ denote the set of functions from G to $\mathbb{Z}_2$. Note that $(\mathbb{Z}_2)^G$ forms a group under pointwise addition. Let $\mathbf{0}$ be the identity element of $(\mathbb{Z}_2)^G$ (that is, the function mapping every element of G to 0). Note also that there is natural action of G on $(\mathbb{Z}_2)^G$. Written in exponential notation, we have that if $\mathbf{a} \in (\mathbb{Z}_2)^G$ and $g \in G$, then $\mathbf{a}^g$ is the element of $(\mathbb{Z}_2)^G$ defined by $\mathbf{a}^g(x) = \mathbf{a}(g^{-1}x)$ for every $x \in G$.

The *wreath product* $\mathbb{Z}_2 \wr G$ is the group consisting of all pairs $(\mathbf{a}, g)$ where $\mathbf{a} \in (\mathbb{Z}_2)^G$ and $g \in G$, with the group operation given by

$$(\mathbf{a}, g)(\mathbf{b}, h) = (\mathbf{a} + \mathbf{b}^g, gh).$$

For $g \in G$, let $\mathbf{a}_g \in (\mathbb{Z}_2)^G$ be the function mapping g to 1 and every other element of G to 0. If S is a generating set for G, then

$$\{(\mathbf{a}_1, 1_G)\} \cup \{(\mathbf{0}, s) \mid s \in S\}$$

is a generating set for $\mathbb{Z}_2 \wr G$ which we call the *canonical* generating set for $\mathbb{Z}_2 \wr G$ (with respect to S).

Example 4.1. Let $G = \mathbb{Z}_2 = \{0, 1\}$, let $S = \{1\} \subseteq G$, let $\hat{G} = \mathbb{Z}_2 \wr G$ and let $a = (\mathbf{a}_0, 0), c = (\mathbf{0}, 1) \in \hat{G}$. The canonical generating set for $\hat{G}$ with respect to S is $\hat{S} = \{a, c\}$. Writing $b = a^c = (\mathbf{a}_0, 0)$, $\mathrm{Cay}(\hat{G}, \hat{S})$ is illustrated in Figure 2. (The colours of the vertices will be explained in Example 4.3.)

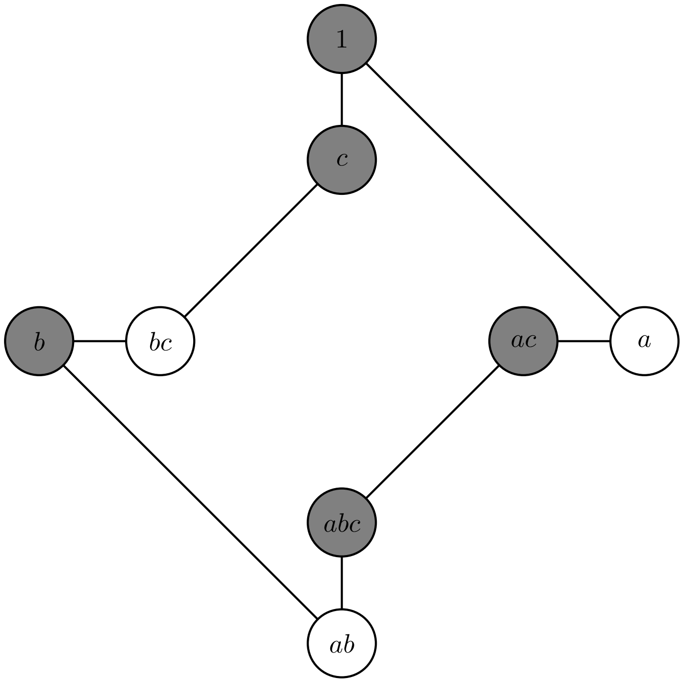


Figure 2: $\mathrm{Cay}(\mathbb{Z}_2 \wr \mathbb{Z}_2, \{a, c\})$.

Lemma 4.2. *Let S be a generating set for a group G, let $\hat{G} = \mathbb{Z}_2 \wr G$ and let $\hat{S}$ be the canonical generating set for $\hat{G}$ with respect to S. If $\mathrm{Cay}(G, S)$ is bipartite and has an induced subgraph of maximum degree 1 on more than half the vertices, then the same is true for $\mathrm{Cay}(\hat{G}, \hat{S})$.*

Proof. We first show that $\mathrm{Cay}(\hat{G}, \hat{S})$ is bipartite. Call an element of G *even* if it lies in the same part of the bipartition of $\mathrm{Cay}(G, S)$ as 1_G, and *odd* otherwise. Call an element $\mathbf{a}$ of $(\mathbb{Z}_2)^G$ *even* if $\mathbf{a}$ maps an even number of elements of G to 1, and call $\mathbf{a}$ *odd* otherwise. Finally, call an element $\hat{g} = (\mathbf{a}, g) \in \hat{G}$ *even* if $\mathbf{a}$ and g are either both even or both odd, and call $\hat{g}$ *odd* otherwise. It is straightforward to check that if $\hat{s} \in \hat{S}$, then $\hat{g}$ is even if and only if $\hat{g}\hat{s}$ is odd. Thus the partition of $\hat{G}$ into even and odd elements is a bipartition of $\mathrm{Cay}(\hat{G}, \hat{S})$.

Let $H \subseteq G$ be such that $|H| > |G|/2$ and the subgraph of $\mathrm{Cay}(G, S)$ induced by H has maximum degree 1. Denote by G_{even} and G_{odd} the set of even and odd elements of G, respectively. For each $\mathbf{a} \in (\mathbb{Z}_2)^G$, let $[\mathbf{a}] = \{(\mathbf{a}, g) \mid g \in G\} \subseteq \hat{G}$ and define a subset $H_{\mathbf{a}}$

of G as follows:

$$H_{\mathbf{a}} = \begin{cases} H & \text{if } \mathbf{a} = \mathbf{0}, \\ G_{\text{odd}} & \text{if } \mathbf{a} = \mathbf{a}_g \text{ for some } g \in G_{\text{even}}, \\ G_{\text{even}} & \text{otherwise.} \end{cases}$$

Let $\hat{H} = \{(\mathbf{a}, h) \mid \mathbf{a} \in (\mathbb{Z}_2)^G, h \in H_{\mathbf{a}}\} \subseteq \hat{G}$. Clearly, $|\hat{H} \cap [\mathbf{0}]| = |H| > |G|/2$ while, for $\mathbf{a} \neq \mathbf{0}$, we have $|\hat{H} \cap [\mathbf{a}]| = |G|/2$. It follows that $|\hat{H}| > |\hat{G}|/2$. Let $(\mathbf{a}, h) \in \hat{H}$. We show that $(\mathbf{a}, h)$ has at most one neighbour in $\hat{H}$. By definition, $h \in H_{\mathbf{a}}$.

If $\mathbf{a} = \mathbf{0}$, then $h \in H$. Note that, if $g \in G_{\text{even}}$, then $H_{\mathbf{a}_g} = G_{\text{odd}}$, whereas if $g \in G_{\text{odd}}$, then $H_{\mathbf{a}_g} = G_{\text{even}}$. In particular, for every $g \in G$, $(\mathbf{a}_g, g) \notin \hat{H}$. It follows that $(\mathbf{0}, h)(\mathbf{a}_1, 1_G) = (\mathbf{a}_1^h, h) = (\mathbf{a}_h, h) \notin \hat{H}$. On the other hand, for $s \in S$, $(\mathbf{0}, h)(\mathbf{0}, s) = (\mathbf{0}, hs) \in \hat{H}$ if and only if $hs \in H$. This shows that the number of neighbours of $(\mathbf{0}, h)$ in $\hat{H}$ is the same as the number of neighbours of h in H and therefore at most 1 (with equality reached for some $h \in H$).

Assume now $\mathbf{a} \neq \mathbf{0}$. Since the partition of $\hat{G}$ into even and odd elements is a bipartition of $\mathrm{Cay}(\hat{G}, \hat{S})$, it follows from the definition of $H_{\mathbf{a}}$ that there are no two adjacent elements in $\hat{H} \cap [\mathbf{a}]$. Since $(\mathbf{a}, h)$ has exactly one neighbour outside of $[\mathbf{a}]$ (namely $(\mathbf{a}, h)(\mathbf{a}_1, 1_G)$), it follows that $(\mathbf{a}, h)$ has at most one neighbour in $\hat{H}$. This concludes the proof that the subgraph induced by $\hat{H}$ has maximum degree 1. □

By starting with, say $(G, S) = (\mathbb{Z}_2, \{1\})$ and applying Lemma 4.2 repeatedly, one obtains an infinite family of Cayley graphs of unbounded valency such that every graph in the family admits an induced subgraph of maximum degree 1 on a set of more than half the vertices, as claimed.

Example 4.3. To illustrate Lemma 4.2, we pick off where Example 4.1 left off. We have $G_{\text{even}} = \{0\}$ and $G_{\text{odd}} = \{1\}$. Let $H = G$. Following Lemma 4.2, we have

$$\begin{aligned} H_{\mathbf{0}} &= \{0, 1\}, \\ H_{\mathbf{a}_0} &= \{1\}, \\ H_{\mathbf{a}_1} &= \{0\}, \\ H_{\mathbf{a}_0 + \mathbf{a}_1} &= \{0\}. \end{aligned}$$

This gives $\hat{H} = \{1, c, ac, b, abc\}$, which is the set of vertices coloured in gray in Figure 2. One can observe that this is more than half the vertices of the graph and that the induced subgraph on $\hat{H}$ has maximum valency 1.

5 Concluding remarks

1. Recall that a *covering map* f from a graph $\hat{\Gamma}$ to a graph Γ is a surjective map that is a local isomorphism. If such a map exists, then $\hat{\Gamma}$ is a *covering graph* of Γ. It is easy to see that if Γ is a d-regular graph admitting an induced subgraph of maximum degree 1 on a set of more than half the vertices, then $\hat{\Gamma}$ has the same property. (It is well known that all the vertex-fibers have the same cardinality, so one can simply take the preimage of the set of more than half the vertices of Γ.) Starting from one example, one can thus construct infinitely many having the same valency, vertex-transitivity, Cayleyness, etc.

2. As a consequence of the above remark, we can construct infinite families of d-regular Cayley graphs of order n admitting induced subgraphs of maximum degree 1 on $\frac{1+\varepsilon(d)}{2}n$ vertices.
3. It is an easy exercise that if Γ is a d-regular graph of order n admitting a subset X of vertices such that the induced graph on X has maximum valency at most 1, then $|X|/n \leq \frac{d}{2d-1}$. Note that the Odd graph O_{d-1} attains this bound and so is extremal from this perspective. It would be interesting to know if this bound can be achieved by Cayley graphs. (Such examples were later found by García-Marco and Knauer [3].)
4. In light of the results from [1, 7] on Cayley graphs of abelian groups, it seems natural ask if there is a natural family of nonabelian groups having the same property. In Sections 3 and 4, we give examples of Cayley graphs on some dihedral groups and some 2-groups that do not have this property. Given that both these families of groups are in some sense close to being abelian (dihedral groups have a cyclic subgroup of index 2, while 2-groups are nilpotent), the question of determining whether any natural family of nonabelian groups has this property seems even more interesting.

ORCID iDs

Florian Lehner https://orcid.org/0000-0002-0599-2390
Gabriel Verret https://orcid.org/0000-0003-1766-4834

References

[1] N. Alon and K. Zheng, Unitary signings and induced subgraphs of Cayley graphs of $\mathbb{Z}_2^n$, arXiv:2003.04926 [math.CO].

[2] F. R. K. Chung, Z. Füredi, R. L. Graham and P. Seymour, On induced subgraphs of the cube, *J. Comb. Theory Ser. A* **49** (1988), 180–187, doi:10.1016/0097-3165(88)90034-9.

[3] I. García-Marco and K. Knauer, On sensitivity in bipartite Cayley graphs, arXiv:2009.00554 [math.CO].

[4] C. D. Godsil, More odd graph theory, *Discrete Math.* **32** (1980), 205–207, doi:10.1016/0012-365x(80)90055-2.

[5] H. Huang, Induced subgraphs of hypercubes and a proof of the sensitivity conjecture, *Ann. of Math.* **190** (2019), 949–955, doi:10.4007/annals.2019.190.3.6.

[6] N. Nisan and M. Szegedy, On the degree of Boolean functions as real polynomials, *Comput. Complex.* **4** (1994), 301–313, doi:10.1007/bf01263419.

[7] A. Potechin and H. Y. Tsang, A conjecture on induced subgraphs of Cayley graphs, arXiv:2003.13166 [math.CO].

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 83–93
https://doi.org/10.26493/1855-3974.2161.f55
(Also available at http://amc-journal.eu)

On sign-symmetric signed graphs*

Ebrahim Ghorbani

Department of Mathematics, K. N. Toosi University of Technology, P.O. Box 16765-3381, Tehran, Iran

Willem H. Haemers

Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands

Hamid Reza Maimani, Leila Parsaei Majd †

Mathematics Section, Department of Basic Sciences, Shahid Rajaee Teacher Training University, P.O. Box 16785-163, Tehran, Iran

Received 24 October 2019, accepted 27 March 2020, published online 10 November 2020

Abstract

A signed graph is said to be sign-symmetric if it is switching isomorphic to its negation. Bipartite signed graphs are trivially sign-symmetric. We give new constructions of non-bipartite sign-symmetric signed graphs. Sign-symmetric signed graphs have a symmetric spectrum but not the other way around. We present constructions of signed graphs with symmetric spectra which are not sign-symmetric. This, in particular answers a problem posed by Belardo, Cioabă, Koolen, and Wang in 2018.

Keywords: Signed graph, spectrum.

Math. Subj. Class. (2020): 05C22, 05C50

1 Introduction

Let G be a graph with vertex set V and edge set E. All graphs considered in this paper are undirected, finite, and simple (without loops or multiple edges).

A *signed graph* is a graph in which every edge has been declared positive or negative. In fact, a signed graph Γ is a pair (G, σ), where $G = (V, E)$ is a graph, called the underlying

*The authors would like to thank the anonymous referees for their helpful comments and suggestions.

†Corresponding author.

E-mail addresses: ghorbani@kntu.ac.ir (Ebrahim Ghorbani), haemers@uvt.nl (Willem H. Haemers), maimani@ipm.ir (Hamid Reza Maimani), leila.parsaei84@yahoo.com (Leila Parsaei Majd)

graph, and $\sigma\colon E \to \{-1,+1\}$ is the sign function or signature. Often, we write $\Gamma = (G,\sigma)$ to mean that the underlying graph is G. The signed graph $(G,-\sigma) = -\Gamma$ is called the *negation* of Γ. Note that if we consider a signed graph with all edges positive, we obtain an unsigned graph.

Let v be a vertex of a signed graph Γ. *Switching* at v is changing the signature of each edge incident with v to the opposite one. Let $X \subseteq V$. Switching a vertex set X means reversing the signs of all edges between X and its complement. Switching a set X has the same effect as switching all the vertices in X, one after another.

Two signed graphs $\Gamma = (G,\sigma)$ and $\Gamma' = (G,\sigma')$ are said to be *switching equivalent* if there is a series of switching that transforms Γ into Γ'. If Γ' is isomorphic to a switching of Γ, we say that Γ and Γ' are *switching isomorphic* and we write $\Gamma \simeq \Gamma'$. The signed graph $-\Gamma$ is obtained from Γ by reversing the sign of all edges. A signed graph $\Gamma = (G,\sigma)$ is said to be *sign-symmetric* if Γ is switching isomorphic to $(G,-\sigma)$, that is: $\Gamma \simeq -\Gamma$.

For a signed graph $\Gamma = (G,\sigma)$, the adjacency matrix $A = A(\Gamma) = (a_{ij})$ is an $n\times n$ matrix in which $a_{ij} = \sigma(v_iv_j)$ if v_i and v_j are adjacent, and 0 if they are not. Thus A is a symmetric matrix with entries $0, \pm 1$ and zero diagonal, and conversely, any such matrix is the adjacency matrix of a signed graph. The spectrum of Γ is the list of eigenvalues of its adjacency matrix with their multiplicities. We say that Γ has a *symmetric spectrum* (with respect to the origin) if for each eigenvalue λ of Γ, $-\lambda$ is also an eigenvalues of Γ with the same multiplicity.

Recall that (see [4]), the *Seidel adjacency matrix* of a graph G with the adjacency matrix A is the matrix S defined by

$$S_{uv} = \begin{cases} 0 & \text{if } u = v \\ -1 & \text{if } u \sim v \\ 1 & \text{if } u \nsim v \end{cases}$$

so that $S = J - I - 2A$. The Seidel adjacency spectrum of a graph is the spectrum of its Seidel adjacency matrix. If G is a graph of order n, then the Seidel matrix of G is the adjacency matrix of a signed complete graph Γ of order n where the edges of G are precisely the negative edges of Γ.

Proposition 1.1. *Suppose S is a Seidel adjacency matrix of order n. If n is even, then S is nonsingular, and if n is odd,* $\operatorname{rank}(S) \geq n-1$. *In particular, if n is odd, and S has a symmetric spectrum, then S has an eigenvalue* 0 *of multiplicity* 1.

Proof. We have $\det(S) \equiv \det(I-J) \pmod 2$, and $\det(I-J) = 1-n$. Hence, if n is even, $\det(S)$ is odd. So, S is nonsingular. Now, if n is odd, any principal submatrix of order $n-1$ is nonsingular. Therefore, $\operatorname{rank}(S) \geq n-1$. □

The goal of this paper is to study sign-symmetric signed graphs as well as signed graphs with symmetric spectra. It is known that bipartite signed graphs are sign-symmetric. We give new constructions of non-bipartite sign-symmetric graphs. It is obvious that sign-symmetric graphs have a symmetric spectrum but not the other way around (see Remark 4.1 below). We present constructions of graphs with symmetric spectra which are not sign-symmetric. This, in particular answers a problem posed in [2].

2 Constructions of sign-symmetric graphs

We note that the property that two signed graphs Γ and Γ' are switching isomorphic is equivalent to the existence of a 'signed' permutation matrix P such that $PA(\Gamma)P^{-1} = A(\Gamma')$. If Γ is a bipartite signed graph, then we may write its adjacency matrix as

$$A = \begin{bmatrix} O & B \\ B^\top & O \end{bmatrix}.$$

It follows that $PAP^{-1} = -A$ for

$$P = \begin{bmatrix} -I & O \\ O & I \end{bmatrix},$$

which means that bipartite graphs are 'trivially' sign-symmetric. So it is natural to look for non-bipartite sign-symmetric graphs. The first construction was given in [1] as follows.

Theorem 2.1. *Let n be an even positive integer and V_1 and V_2 be two disjoint sets of size $n/2$. Let G be an arbitrary graph with the vertex set V_1. Construct the complement of G, that is G^c, with the vertex set V_2. Assume that $\Gamma = (K_n, \sigma)$ is a signed complete graph in which $E(G) \cup E(G^c)$ is the set of negative edges. Then Γ is sign-symmetric.*

2.1 Constructions for general signed graphs

Let $\mathcal{M}_{r,s}$ denote the set of $r \times s$ matrices with entries from $\{-1, 0, 1\}$. We give another construction generalizing the one given in Theorem 2.1:

Theorem 2.2. *Let $B, C \in \mathcal{M}_{k,k}$ be symmetric matrices where B has a zero diagonal. Then the signed graph with the adjacency matrices*

$$A = \begin{bmatrix} B & C \\ C & -B \end{bmatrix}$$

is sign-symmetric on $2k$ vertices.

Proof.

$$\begin{bmatrix} O & -I \\ I & O \end{bmatrix} \begin{bmatrix} B & C \\ C & -B \end{bmatrix} \begin{bmatrix} O & I \\ -I & O \end{bmatrix} = \begin{bmatrix} -B & -C \\ -C & B \end{bmatrix} = -A \qquad \square$$

Note that Theorem 2.2 shows that there exists a sign-symmetric graph for every even order.

We define the family $\mathcal{F}$ of signed graphs as those which have an adjacency matrix satisfying the conditions given in Theorem 2.2. To get an impression on what the role of $\mathcal{F}$ is in the family of sign-symmetric graphs, we investigate small complete signed graphs. All but one complete signed graphs with symmetric spectra of orders $4, 6, 8$ are illustrated in Figure 1 (we show one signed graph in the switching class of the signed complete graphs induced by the negative edges). There is only one sign-symmetric complete signed graph of order 4. There are four complete signed graphs with symmetric spectrum of order 6, all of which are sign-symmetric, and twenty-one complete signed graphs with symmetric spectrum of order 8, all except the last one are sign-symmetric, and together with the negation of the last signed graph, Figure 1 gives all complete signed graphs with symmetric spectrum of order $4, 6$ and 8. Interestingly, all of the above sign-symmetric signed graphs belong to $\mathcal{F}$.

The following proposition shows that $\mathcal{F}$ is closed under switching.

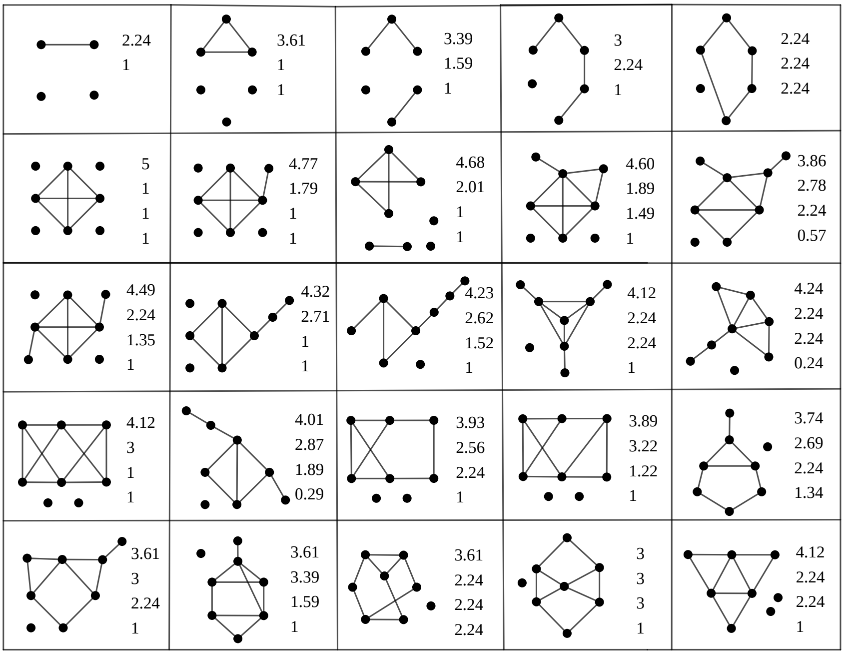


Figure 1: Complete signed graphs (up to switching isomorphism and negation) of order $4, 6, 8$ having symmetric spectrum. The numbers next to the graphs are the non-negative eigenvalues. Only the last graph on the right is not sign-symmetric.

Proposition 2.3. *If* $\Gamma \in \mathcal{F}$ *and* Γ' *is obtained from* Γ *by switching, then* $\Gamma' \in \mathcal{F}$.

Proof. Let $\Gamma \in \mathcal{F}$. It is enough to show that if Γ' is obtained from Γ by switching with respect to its first vertex, then $\Gamma' \in \mathcal{F}$. We may write the adjacency matrix of Γ as follows:

$$A = \left[\begin{array}{c|c|c|c} 0 & \mathbf{b}^\top & c & \mathbf{c}^\top \\ \hline \mathbf{b} & B' & \mathbf{c} & C' \\ \hline c & \mathbf{c}^\top & 0 & -\mathbf{b}^\top \\ \hline \mathbf{c} & C' & -\mathbf{b} & -B' \end{array}\right].$$

After switching with respect to the first vertex of Γ, the adjacency matrix of the resulting signed graph is

$$\left[\begin{array}{c|c|c|c} 0 & -\mathbf{b}^\top & -c & -\mathbf{c}^\top \\ \hline -\mathbf{b} & B' & \mathbf{c} & C' \\ \hline -c & \mathbf{c}^\top & 0 & -\mathbf{b}^\top \\ \hline -\mathbf{c} & C' & -\mathbf{b} & -B' \end{array}\right].$$

Now by interchange the 1st and $(k+1)$-th rows and columns we obtain

$$\left[\begin{array}{c|c|c|c} 0 & \mathbf{c}^\top & -c & -\mathbf{b}^\top \\ \hline \mathbf{c} & B' & -\mathbf{b} & C' \\ \hline -c & -\mathbf{b}^\top & 0 & -\mathbf{c}^\top \\ \hline -\mathbf{b} & C' & -\mathbf{c} & -B' \end{array}\right]$$

which is a matrix of the form given in Theorem 2.2 and thus Γ' is isomorphic with a signed graph in $\mathcal{F}$. □

In the following we present two constructions for complete sign-symmetric signed graphs using self-complementary graphs.

2.2 Constructions for complete signed graphs

In the following, the meaning of a self-complementary graph is the same as defined for unsigned graphs. Let G be a self-complementary graph so that there is a permutation matrix P such that $PA(G)P^{-1} = A(\overline{G})$ and $PA(\overline{G})P^{-1} = A(G)$. It follows that if Γ is a complete signed graph with $E(G)$ being its negative edges, then $A(\Gamma) = A(\overline{G}) - A(G)$ (in other words, $A(\Gamma)$ is the Seidel matrix of G). It follows that $PA(\Gamma)P^{-1} = -A(\Gamma)$. So we obtain the following:

Observation 2.4. If Γ is a complete signed graph whose negative edges induce a self-complementary graph, then Γ is sign-symmetric.

We give one more construction of sign-symmetric signed graphs based on self-complementary graphs as a corollary to Observation 2.4. We remark that a self-complementary graph of order n exists whenever $n \equiv 0$ or $1 \pmod 4$.

Proposition 2.5. *Let G, H be two self-complementary graphs, and let Γ be a complete signed graph whose negative edges induce the join of G and H (or the disjoint union of G and H). Then Γ is sign symmetric. In particular, if G has n vertices, and if H is a singleton, then the complete signed graph Γ of order $n+1$ with negative edges equal to $E(G)$ is sign-symmetric.*

In the following remark we present a sign-symmetric construction for non-complete signed graphs.

Remark 2.6. Let Γ', Γ'' be two signed graphs which are isomorphic to $-\Gamma', -\Gamma''$, respectively. Consider the signed graph Γ obtained from joining Γ' and Γ'' whose negative edges are the union of negative edges in Γ' and Γ''. Then, Γ is sign-symmetric.

Remark 2.7. By Proposition 2.5, we have a construction of sign-symmetric complete signed graphs of order $n \equiv 0, 1$ or $2 \pmod 4$. All complete sign-symmetric signed graphs of order 5 and 9 (depicted in Figure 2) can be obtained in this way. There is just one sign-symmetric signed graph of order 5 which is obtained by joining a vertex to a complete signed graph of order 4 whose negative edges form a path of length 3 (which is self-complementary). Moreover, there exist sixteen complete signed graphs of order 9 with symmetric spectrum of which ten are sign-symmetric; the first three are not sign-symmetric, and when we include their negations we get them all. All of these ten complete sign-symmetric signed graphs can be obtained by joining a vertex to a complete signed graph of order 8 whose negative edges induce a self-complementary graph. Note that there are exactly ten self-complementary graphs of order 8.

Theorem 2.8. *There exists a complete sign-symmetric signed graph of order n if and only if $n \equiv 0, 1$ or $2 \pmod 4$.*

Proof. Using the previous results obviously one can construct a sign-symmetric signed graph of order n whenever $n \equiv 0, 1$ or $2 \pmod 4$. Now, suppose that there is a complete sign-symmetric signed graph Γ of order n with $n \equiv 3 \pmod 4$. By [6, Corollary 3.6], the determinant of the Seidel matrix of Γ is congruent to $1 - n \pmod 4$. Since $n \equiv 3 \pmod 4$, the determinant of the Seidel matrix (obtained from the negative edges of Γ) is not zero. Hence, we can conclude that all eigenvalues of Γ are non-zero. Therefore, Γ cannot have a symmetric spectrum, and also it cannot be sign-symmetric. □

In [7] all switching classes of Seidel matrices of order at most seven are given. There is a error in the spectrum of one of the graphs on six vertices in [7, Table 4.1] (2.37 should be 2.24), except for that, the results in [7] coincide with ours.

3 Positive and negative cycles

A graph whose connected components are K_2 or cycles is called an *elementary graph*. Like unsigned graphs, the coefficients of the characteristic polynomial of the adjacency matrix of a signed graph Γ can be described in terms of elementary subgraphs of Γ.

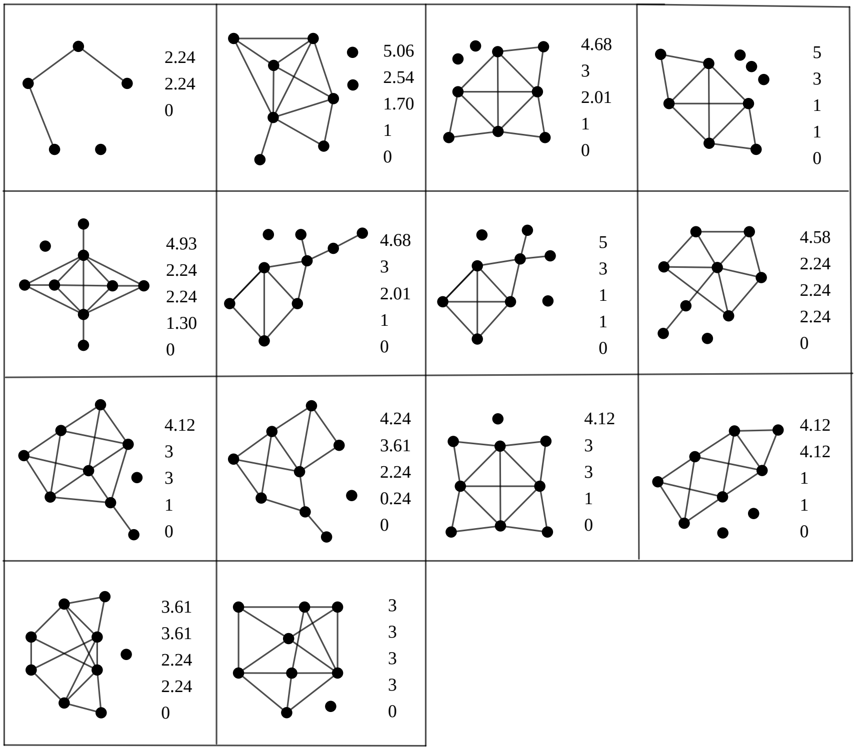


Figure 2: Complete signed graphs (up to switching isomorphism and negation) of order $5, 9$ having symmetric spectrum. The numbers next to the graphs are the non-negative eigenvalues.The first three signed graphs of order 9 are not sign-symmetric.

Theorem 3.1 ([3, Theorem 2.3])**.** *Let* $\Gamma = (G, \sigma)$ *be a signed graph and*

$$P_\Gamma(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n \tag{3.1}$$

be the characteristic polynomial of the adjacency matrix of Γ*. Then*

$$a_i = \sum_{B \in \mathcal{B}_i} (-1)^{p(B)} 2^{|c(B)|} \sigma(B),$$

where $\mathcal{B}_i$ *is the set of elementary subgraphs of* G *on* i *vertices,* $p(B)$ *is the number of components of* B*,* $c(B)$ *the set of cycles in* B*, and* $\sigma(B) = \prod_{C \in c(B)} \sigma(C)$*.*

Remark 3.2. It is clear that Γ has a symmetric spectrum if and only if in its characteristic polynomial (3.1), we have $a_{2k+1} = 0$, for $k = 1, 2, \ldots$.

In a signed graph, a cycle is called *positive* or *negative* if the product of the signs of its edges is positive or negative, respectively. We denote the number of positive and negative ℓ-cycles by c_ℓ^+ and c_ℓ^-, respectively.

Observation 3.3. For sign-symmetric signed graph, we have

$$c^+_{2k+1} = c^-_{2k+1} \quad \text{for } k = 1, 2, \ldots.$$

Remark 3.4. If in a signed graph Γ, $c^+_{2k+1} = c^-_{2k+1}$ for all $k = 1, 2, \ldots$, then it is not necessary that Γ is sign-symmetric. See the complete signed graph given in Figure 5. For this complete signed graph we have $c^+_{2k+1} = c^-_{2k+1}$ for all $k = 1, 2, \ldots$, but it is not sign-symmetric. Moreover, one can find other examples among complete and non-complete signed graphs. For example, the signed graph given in Figure 4 is a non-complete signed graph with the property that $c^+_{2k+1} = c^-_{2k+1}$ for all $k = 1, 2, \ldots$, but it is not sign-symmetric.

By Theorem 3.1, we have that $a_3 = 2(c^-_3 - c^+_3)$. By Theorem 3.1 and Remark 3.2 for signed graphs having symmetric spectrum, we have $c^+_3 = c^-_3$. Further, for each complete signed graph with a symmetric spectrum, it can be seen that $c^+_5 = c^-_5$. However, the equality $c^+_{2k+1} = c^-_{2k+1}$ does not necessarily hold for $k \geq 3$. The complete signed graph in Figure 3 has a symmetric spectrum for which $c^+_7 \neq c^-_7$.

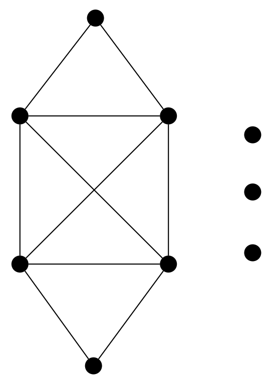


Figure 3: The graph induced by negative edges of a complete signed graph on 9 vertices with a symmetric spectrum but $c^+_7 \neq c^-_7$.

Remark 3.5. There are some examples showing that for a non-complete signed graph we have $c^+_{2k+1} = c^-_{2k+1}$ for all $k = 1, 2, \ldots$, but their spectra are not symmetric. As an example see Figure 4 (dashed edges are negative; solid edges are positive).

Now, we may ask a weaker version of the result mentioned in Remark 3.4 as follows.

Question 3.6. Is it true that if in a complete signed graph Γ, $c^+_{2k+1} = c^-_{2k+1}$ for all $k = 1, 2, \ldots$, then Γ has a symmetric spectrum?

4 Sign-symmetric vs. symmetric spectrum

Remark 4.1. Consider the complete signed graph whose negative edges induces the graph of Figure 5. This graph has a symmetric spectrum, but it is not sign-symmetric. Note that this complete signed graph has the minimum order with this property. Moreover, for this complete signed graph we have the equalities $c^+_{2k+1} = c^-_{2k+1}$ for $k = 1, 2, 3$.

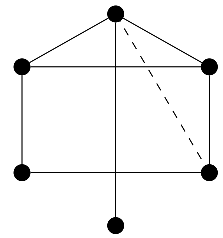


Figure 4: A signed graph with $c^+_{2k+1} = c^-_{2k+1}$ for $k = 1, 2, \ldots$, but its spectrum is not symmetric.

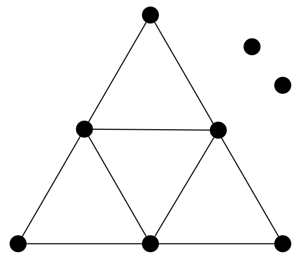


Figure 5: The graph induced by negative edges of a complete signed graph on 8 vertices with a symmetric spectrum but not sign-symmetric.

Remark 4.2. A *conference matrix* C of order n is an $n \times n$ matrix with zero diagonal and all off-diagonal entries ± 1, which satisfies $CC^\top = (n-1)I$. If C is symmetric, then C has eigenvalues $\pm\sqrt{n-1}$. Hence, its spectrum is symmetric. Conference matrices are well-studied; see for example [4, Section 10.4]. An important example of a symmetric conference matrix is the Seidel matrix of the Paley graph extended with an isolated vertex, where the *Paley graph* is defined on the elements of a finite field $\mathbf{F}_q$, with $q \equiv 1 \pmod 4$, where two elements are adjacent whenever the difference is a nonzero square in $\mathbf{F}_q$. The Paley graph is self-complementary. Therefore, by Proposition 2.5, C is the adjacency matrix of a sign-symmetric complete signed graph. However, there exist many more symmetric conference matrices, including several that are not sign-symmetric (see [5]).

In [2], the authors posed the following problem on the existence of the non-complete signed graphs which are not sign-symmetric but have symmetric spectrum.

Problem 4.3 ([2])**.** Are there non-complete connected signed graphs whose spectrum is symmetric with respect to the origin but they are not sign-symmetric?

We answer this problem by showing that there exists such a graph for any order $n \geq 6$. For $s \geq 0$, define the signed graph Γ_s to be the graph illustrated in Figure 6.

Theorem 4.4. *For $s \geq 0$, the graph Γ_s has a symmetric spectrum, but it is not sign-symmetric.*

Proof. Let S be the set of s vertices adjacent to both 1 and 5. The positive 5-cycles of Γ_s are 123461 together with $u1645u$ for any $u \in S$, and the negative 5-cycles are $u1465u$ for

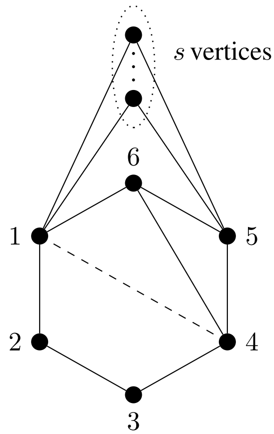


Figure 6: The graph Γ_s.

any $u \in S$. Hence, $c_5^+ = s+1$ and $c_5^- = s$. In view of Observation 3.3, this shows that Γ_s is not sign-symmetric.

Next, we show that Γ_s has a symmetric spectrum. It suffices to verify that $a_{2k+1} = 0$ for $k = 1, 2, \ldots$.

The graph Γ_s contains a unique positive cycle of length 3: 4564 and a unique negative cycle of length 3: 1461. It follows that $a_3 = 0$.

As discussed above, we have $c_5^+ = s+1$ and $c_5^- = s$. We count the number of positive and negative copies of $K_2 \cup C_3$. For the negative triangle 1461, there are $s+1$ non-incident edges, namely 23 and $5u$ for any $u \in S$ and for the positive triangle 4564, there are $s+2$ non-incident edges, namely 12, 23 and $1u$ for any $u \in S$. It follows that

$$a_5 = -2((s+1)-s) + 2((s+2)-(s+1) = 0.$$

Now, we count the number of positive and negative elementary subgraphs on 7 vertices:

C_7: s positive: $u123465u$ for any $u \in S$, and no negative;

$K_2 \cup C_5$: $2s$ positive: $u5 \cup 123461$, and $23 \cup u1645u$ for any $u \in S$, and s negative: $23 \cup u1465u$ for any $u \in S$;

$2K_2 \cup C_3$: $s+1$ positive: $u1 \cup 23 \cup 4564$ for any $u \in S$, and $s+1$ negative: $u5 \cup 23 \cup 1461$ for any $u \in S$;

$C_4 \cup C_3$: none.

Therefore,

$$a_7 = -2(s-0) + 2(2s-s) - 2((s+1)-(s+1)) = 0.$$

The graph Γ_s contains no elementary subgraph on 8 vertices or more. The result now follows. □

More families of non-complete signed graphs with a symmetric spectrum but not sign-symmetric can be found. Consider the signed graphs $\Gamma_{s,t}$ depicted in Figure 7, in which the number of upper repeated pair of vertices is $s \geq 0$ and the number of upper repeated pair of vertices is $t \geq 1$. In a similar fashion as in the proof of Theorem 4.4 it can be verified that $\Gamma_{s,t}$ has a symmetric spectrum, but it is not sign-symmetric.

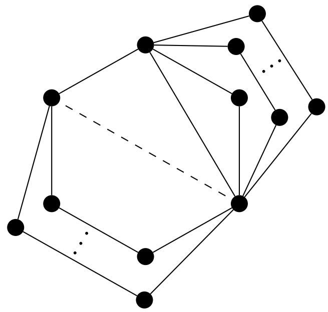


Figure 7: The family of signed graphs $\Gamma_{s,t}$.

ORCID iDs

Willem H. Haemers https://orcid.org/0000-0001-7308-8355
Hamid Reza Maimani https://orcid.org/0000-0001-9020-0871
Leila Parsaei Majd https://orcid.org/0000-0003-0775-3147

References

[1] S. Akbari, H. R. Maimani and L. Parsaei Majd, On the spectrum of some signed complete and complete bipartite graphs, *Filomat* **32** (2018), 5817–5826, doi:10.2298/fil1817817a.

[2] F. Belardo, S. M. Cioabă, J. Koolen and J. Wang, Open problems in the spectral theory of signed graphs, *Art Discrete Appl. Math.* **1** (2018), #P2.10 (23 pages), doi:10.26493/2590-9770.1286.d7b.

[3] F. Belardo and S. K. Simić, On the Laplacian coefficients of signed graphs, *Linear Algebra Appl.* **475** (2015), 94–113, doi:10.1016/j.laa.2015.02.007.

[4] A. E. Brouwer and W. H. Haemers, *Spectra of Graphs*, Universitext, Springer, New York, 2012, doi:10.1007/978-1-4614-1939-6.

[5] F. C. Bussemaker, R. A. Mathon and J. J. Seidel, Tables of two-graphs, in: S. B. Rao (ed.), *Combinatorics and Graph Theory*, Springer, Berlin-New York, volume 885 of *Lecture Notes in Mathematics*, 1981 pp. 70–112, proceedings of the Second Symposium held at the Indian Statistical Institute, Calcutta, February 25 – 29, 1980.

[6] G. Greaves, J. H. Koolen, A. Munemasa and F. Szöllősi, Equiangular lines in Euclidean spaces, *J. Comb. Theory Ser. A* **138** (2016), 208–235, doi:10.1016/j.jcta.2015.09.008.

[7] J. H. van Lint and J. J. Seidel, Equilateral point sets in elliptic geometry, *Nederl. Akad. Wetensch. Proc. Ser. A* **69** (1966), 335–348.

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 95–124
https://doi.org/10.26493/1855-3974.2120.085
(Also available at http://amc-journal.eu)

Balancing polyhedra*

Gábor Domokos

MTA-BME Morphodynamics Research Group and Department of Mechanics, Materials and Structures, Budapest University of Technology, Műegyetem rakpart 1-3., Budapest, Hungary, 1111

Flórián Kovács

MTA-BME Morphodynamics Research Group and Department of Structural Mechanics, Budapest University of Technology, Műegyetem rakpart 1-3., Budapest, Hungary, 1111

Zsolt Lángi †

MTA-BME Morphodynamics Research Group and Department of Geometry, Budapest University of Technology, Egry József utca 1., Budapest, Hungary, 1111

Krisztina Regős, Péter T. Varga

MTA-BME Morphodynamics Research Group, Műegyetem rakpart 1-3., Budapest, Hungary, 1111

Dedicated to the memory of John Horton Conway.

Received 18 September 2019, accepted 20 May 2020, published online 10 November 2020

Abstract

We define the mechanical complexity $C(P)$ of a 3-dimensional convex polyhedron P, interpreted as a homogeneous solid, as the difference between the total number of its faces, edges and vertices and of its static equilibria; and the mechanical complexity $C(S,U)$ of primary equilibrium classes $(S,U)^E$ with S stable and U unstable equilibria as the infimum of the mechanical complexity of all polyhedra in that class. We prove that the mechanical complexity of a class $(S,U)^E$ with $S,U > 1$ is the minimum of $2(f+v-$

*The authors acknowledge the support of the BME Water Sciences & Disaster Prevention TKP2020 IE grant of NKFIH Hungary (BME IE-VIZ TKP2020) and the NKFIH grant K119245. The authors thank Mr. Otto Albrecht for backing the prize for the complexity of the Gömböc-class. Any solution should be sent to the corresponding author as an accepted publication in a mathematics journal of worldwide reputation and it must also have general acceptance in the mathematics community two years after. The authors are indebted to Dr. Norbert Krisztián Kovács for his invaluable advice and help in printing the 9 tetrahedra and 7 pentahedra, and the two referees for their valuable comments and one of them for posing Problems 5.2 and 5.4.

†Corresponding author. The author has been supported by the Bolyai Fellowship of the Hungarian Academy of Sciences and partially supported by the UNKP-20-5 New National Excellence Program of the Ministry of Innovation and Technology.

$S - U$) over all polyhedral pairs (f, v), where a pair of integers is called a polyhedral pair if there is a convex polyhedron with f faces and v vertices. In particular, we prove that the mechanical complexity of a class $(S, U)^E$ is zero if and only if there exists a convex polyhedron with S faces and U vertices. We also give asymptotically sharp bounds for the mechanical complexity of the monostatic classes $(1, U)^E$ and $(S, 1)^E$, and offer a complexity-dependent prize for the complexity of the Gömböc-class $(1, 1)^E$.

Keywords: Polyhedron, static equilibrium, monostatic polyhedron, f-vector.

Math. Subj. Class. (2020): 52B10, 70C20, 52A38

1 Introduction

1.1 Basic concepts and the main result

Polyhedra may be regarded as purely geometric objects, however, they are also often intuitively identified with solids. Among the most obvious sources of such intuition are dice which appear in various polyhedral shapes: while classical, cubic dice have 6 faces, a large diversity of other dice exists as well: dice with $2, 3, 4, 6, 8, 10, 12, 16, 20, 24, 30$ and 100 faces appear in various games [37]. The key idea behind throwing dice is that each of the aforementioned faces is associated with a stable mechanical equilibrium point where dice may be at rest on a horizontal plane. Dice are called *fair* if the probabilities to rest on any face (after a random throw) are equal [10], otherwise they are called *loaded* [9]. The concept of mechanical equilibrium may also be defined in purely geometric terms:

Definition 1.1. Let P be a 3-dimensional convex polyhedron, let $\operatorname{int} P$ and $\operatorname{bd} P$ denote its interior and boundary, respectively and let $c \in \operatorname{int} P$. We say that $q \in \operatorname{bd} P$ is an *equilibrium point* of P with respect to c if the plane H through q and perpendicular to $[c, q]$ supports P at q. In this case q is *nondegenerate*, if $H \cap P$ is the (unique) vertex, edge, or face of P, respectively, that contains q in its relative interior. A nondegenerate equilibrium point q is called *stable, saddle-type* or *unstable*, if $\dim(H \cap P) = 2, 1$ or 0, respectively.

Throughout this paper we deal only with equilibrium points with respect to the center of mass of polyhedra, assuming uniform density. A *support plane* is a generalization of the tangent plane for non-smooth objects. While it is a central concept of convex geometry its name may be related to the mechanical concept of equilibrium. If c coincides with the center of mass of P, then equilibrium points gain intuitive interpretation as locations on $\operatorname{bd} P$ where P may be balanced if it is supported on a horizontal surface (identical to the support plane) without friction in the presence of uniform gravity. Equilibrium points may belong to three stability types: faces may carry stable equilibria, vertices may carry unstable equilibria and edges may carry saddle-type equilibria. Denoting their respective numbers by S, U, H, by the Poincaré-Hopf formula [25] for a convex polyhedron one obtains the following relation for them:

$$S + U - H = 2, \tag{1.1}$$

E-mail addresses: domokos@iit.bme.hu (Gábor Domokos), kovacs.florian@epito.bme.hu (Flórián Kovács), zlangi@math.bme.hu (Zsolt Lángi), regoskriszti@gmail.com (Krisztina Regős), petercobbler@gmail.com (Péter T. Varga)

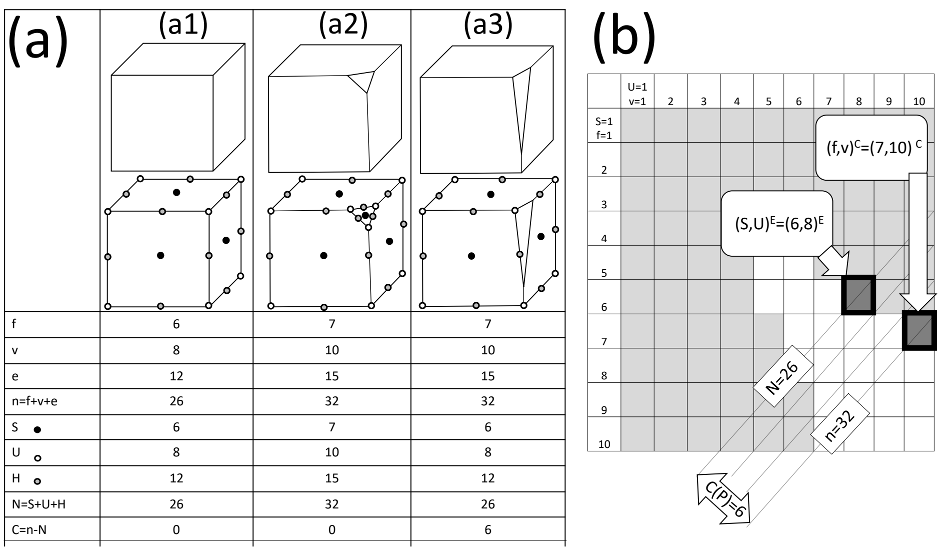
which is strongly reminiscent of the well-known Euler formula

$$f + v - e = 2, \tag{1.2}$$

relating the respective numbers f, v and e of the faces, vertices and edges of a convex polyhedron. In the case of regular, homogeneous, cubic dice the formulae (1.1) and (1.2) appear to express the same fact, however, in case of irregular polyhedra the connection is much less apparent. While the striking similarity between (1.1) and (1.2) can only be fully explained via deep topological and analytic ideas [25], our goal in this paper is to demonstrate an interesting connection at an elementary, geometric level. To this end, we define

$$N = S + U + H,$$
$$n = f + v + e.$$

Figure 1 shows three polyhedra where the values for all these quantities can be compared.



	(a1)	(a2)	(a3)
f	6	7	7
v	8	10	10
e	12	15	15
n=f+v+e	26	32	32
S ●	6	7	6
U ○	8	10	8
H ○	12	15	12
N=S+U+H	26	32	26
C=n-N	0	0	6

Figure 1: (a): Three polyhedra interpreted as homogeneous solids with given numbers for faces (f), vertices (v), edges (e), stable equilibria (S), unstable equilibria (U) and saddle-type equilibria (H), their respective sums $n = f + v + e$, $N = S + U + H$ and mechanical complexity $C = n - N$ (given in Definition 1.2). (b): Polyhedron in column (a3) shown on the overlay of the (S, U) and (f, v) grids, complexity obtained from distance between corresponding diagonals.

The numbers S, U, H may serve, from the mechanical point of view, as a first-order characterization of P and via (1.1) the triplet (S, U, H) may be uniquely represented by the pair (S, U), which is called *primary equilibrium class* of P [35]. Based on this, we denote by $(S, U)^E$ the family of all convex polyhedra having S stable and U unstable equilibrium points with respect to their centers of mass. In an analogous manner, the numbers (v, e, f) (also called the f-vector of P) serve as a first-order combinatorial characterization of P, and via (1.2) they may be uniquely represented by the pair (f, v). Here, we call the family of all convex polyhedra having v vertices and f faces the *primary combinatorial class* of P, and denote it by $(f, v)^C$. The face structure of a convex polyhedron P permits a finer combinatorial description of P. In the literature, the family of convex polyhedra having the same face lattice is called a combinatorial class; here we call it a *secondary combinatorial class*, and discuss it in Section 5. In an entirely analogous manner, one can define also secondary equilibrium classes of convex bodies, for more details the interested reader is referred to [16]. While it is immediately clear that for any polyhedron P we have

$$f \geq S, \qquad v \geq U, \tag{1.3}$$

inverse type relationships (e.g. defining the minimal number of faces and vertices for given numbers of equilibria) are much less obvious.

A trivial necessary condition for any die to be fair can be stated as $f = S$ and it is relatively easy to construct a polyhedron with this property. The opposite extreme case (when a polyhedron is stable only on one of its faces) appears to be far more complicated. John H. Conway was first to notice this curious fact [5] and ever since, his idea has been expanded in various ways [2, 28]. In broader terms, it appears that, as the number of equilibria in a given equilibrium class gets smaller, it is getting increasingly difficult to identify the corresponding geometry. In other words, the difference $(n - N)$ between the topological and mechanical characteristics of the polyhedron appears to indicate some kind of complexity of the geometry. Motivated by this intuition we define the *mechanical complexity* of polyhedra.

Definition 1.2. Let P be a convex polyhedron and let $N(P), n(P)$ denote the total number of its equilibria and the total number of its k*-faces* (i.e., faces of k dimensions) for all values $k = 0, 1, 2$, respectively. Then $C(P) = n(P) - N(P)$ is called the mechanical complexity of P.

We remark that the term *mechanical complexity* has been used in various contexts, ranging from robotics [1] to cell biology [21], to describe phenomena where the observed complexity is rooted in the mechanical properties of the investigated subject. In our case we witness the same phenomenon: the apparent complexity of some polyhedral shapes arises from the mechanical constraint that the number of static equilibria is kept, compared to the number of vertices, edges and faces, very low.

Mechanical complexity may not only be associated with individual polyhedra but also with primary equilibrium classes.

Definition 1.3. If $(S, U)^E$ is a primary equilibrium class, then the quantity

$$C(S, U) = \min\{C(P) : P \in (S, U)^E\}$$

is called the mechanical complexity of $(S, U)^E$.

Our goal is to find the values of $C(S, U)$ for all primary equilibrium classes. For $S, U > 1$ we will achieve this goal while for $S = 1$ or $U = 1$ we provide some partial results. To formulate our main results, we introduce the following concept:

Definition 1.4. Let x, y be positive integers. We say that (x, y) is a polyhedral pair if and only if $x \geq 4$ and $\frac{x}{2} + 2 \leq y \leq 2x - 4$.

The combinatorial classification of convex polyhedra was established by Steinitz [30, 31] (for a proof, see also [20]), who proved, in particular, the following.

Theorem 1.5. *For any positive integers f, v, there is a convex polyhedron P with f faces and v vertices if and only if (f, v) is a polyhedral pair.*

Based on this theorem, we call a primary equilibrium class $(S, U)^E$ a polyhedral (primary equilibrium) class if (S, U) is a polyhedral pair, and the remaining primary equilibrium classes non-polyhedral classes.

Definition 1.6. For any primary equilibrium class $(S, U)^E$ with $S, U \geq 1$, let

$$R(S, U) = \min\{f + v - S - U : (f, v) \text{ is a polyhedral pair and } f, v \text{ satisfy } (1.3)\}.$$

The geometric interpretation of $R(S, U)$ is given in the left panel of Figure 2. Since (1.3) holds for any polyhedron $P \in (S, U)^E$, we immediately have the trivial lower bound for mechanical complexity:

$$C(S, U) \geq 2R(S, U). \tag{1.4}$$

Remark 1.7. Based on Definition 1.4, the function $R(S, U)$ can be expressed as

$$R(S, U) = \begin{cases} \lceil \frac{S}{2} \rceil - U + 2, & \text{if } S > 4 \text{ and } S > 2U - 4, \\ \lceil \frac{U}{2} \rceil - S + 2, & \text{if } U > 4 \text{ and } U > 2S - 4, \\ 8 - S - U, & \text{if } S, U \leq 4, \\ 0, & \text{otherwise.} \end{cases} \tag{1.5}$$

Our main result is Theorem 1.8, stating that this bound is sharp if $S, U > 1$:

Theorem 1.8. *Let $S, U \geq 2$ be positive integers. Then $C(S, U) = 2R(S, U)$.*

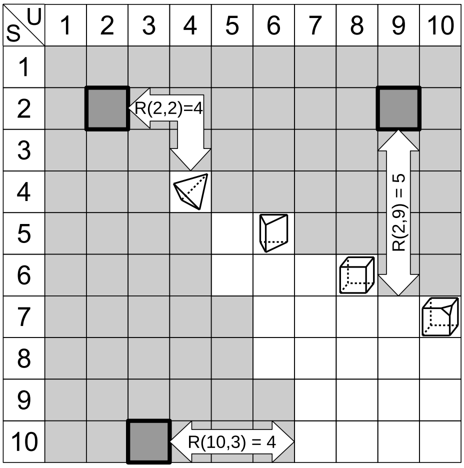
We remark that, as a consequence of Theorem 1.8, $C(S, U) = 0$ if and only if (S, U) is a polyhedral pair. For monostatic equilibrium classes ($S = 1$ or $U = 1$) we cannot provide a sharp value for their mechanical complexity. However, we will provide an upper bound for their complexity, which differs from $2R(S, U)$ only by a constant:

Theorem 1.9. *If $S \geq 4$ then $C(S, 1) \leq 59 + (-1)^S + 2R(S, 1)$; if $U \geq 4$ then $C(1, U) \leq 90 + 2R(1, U)$.*

We also improve the lower bound (1.4) in some of these classes by generalizing a theorem of Conway [7] about the non-existence of a homogeneous tetrahedron with only one stable equilibrium point. We state our result in the following form:

Theorem 1.10. *Any homogeneous tetrahedron has $S \geq 2$ stable and $U \geq 2$ unstable equilibrium points.*

We summarize all results (including those about monostatic classes) in Figure 2.



S \ U	1	2	3	4	5	6	7	8	9	10
1	[12]	[10, 70]	[8, 64]	[6, 96]	[8, 98]	[8, 98]	[10, 100]	[10, 100]	[12, 102]	[12, 102]
2	[10, 66]	8	6	4	6	6	8	8	10	10
3	[8, 64]	6	4	2	4	4	6	6	8	8
4	[6, 66]	4	2	0	2	2	4	4	6	6
5	[8, 66]	6	4	2	0	0	2	2	4	4
6	[8, 68]	6	4	2	0	0	0	0	2	2
7	[10, 68]	8	6	4	2	0	0	0	0	0
8	[10, 70]	8	6	4	2	0	0	0	0	0
9	[12, 70]	10	8	6	4	2	0	0	0	0
10	[12, 72]	10	8	6	4	2	0	0	0	0

Figure 2: Summary of results for $S, U \leq 10$. Left panel: the (S, U) grid with some selected polyhedra as examples. Polyhedral pairs on the (S, U) grid have white background. The function $R(S, U)$ illustrated for classes $(2, 2)^E$, $(2, 9)^E$, $(10, 3)^E$. Right panel: Mechanical complexity of equilibrium classes $(S, U)^E$. Polyhedral pairs on the (S, U) grid have white background. Sharp values for mechanical complexity $C(S, U)$ are given as integers without brackets. In column $U = 1$ and row $S = 1$ we give bounds. If two integers are given in square brackets then they are the lower and upper bounds for $C(S, U)$, if only one integer is given in square brackets then it is the lower bound (and no upper bound is available).

1.2 Sketch of the proof

The main idea of the proofs of Theorems 1.8 and 1.9 is to provide explicit constructions for at least one polyhedron P in each class $(S, U)^E$, $S, U > 1$ with mechanical complexity $C(P) = 2R(S, U)$, in class $(S, 1)^E$, $S \geq 4$ with $C(P) = 59 + (-1)^S + 2R(S, 1)$, and in class $(1, U)^E, U \geq 4$ with $C(P) = 90 + 2R(1, U)$. By Definition 1.3, such a construction establishes an upper bound for $C(S, U)$. In case of $S > 1$ and $U > 1$, by Definition 1.6, this coincides with the lower bound while for $S = 1$ or $U = 1$ the bounds remain separate.

Our proof consists of five parts:

(a) for classes $(S, S)^E$ with $S \geq 4$, suitably chosen pyramids have zero mechanical complexity (Section 3);

(b) for classes $1 < S \leq 5$ and $1 < U \leq 5$, $(S, U)^E \neq (4, 4)^E, (5, 5)^E$, we provide examples found by computer search (Subsection 3.2, Tables 1 and 2);

(c) for polyhedral classes with $S \neq U$, we construct examples by recursive, local manipulations of the pyramids mentioned in (a) (Subsection 3.1);

(d) for non-polyhedral classes with $U > S \geq 6$, we construct examples by recursive, local manipulations starting with polyhedral classes containing simple polyhedra (Subsection 3.2);

(e) for non-polyhedral classes with $6 \leq U < S$ we provide examples by using the polyhedra obtained in (d) and the properties of polarity proved in Section 2. We also show how to modify the construction in (d) for this case (Subsection 3.2);

(f) for monostatic classes with $S = 1$ or $U = 1$ we provide examples using Conway's polyhedron P_C in class $(1, 4)^E$, we also construct a polyhedron P_3 in class $(3, 1)$ and subsequently we apply recursive, local truncations (Section 4).

In Section 2, we prove a number of lemmas which help us keep track of the change of the center of mass of a convex polyhedron under local deformations and establish a connection between equilibrium points of a convex polyhedron and its polar. The local manipulations in our proof may be regarded as generalizations of the algorithm of Steinitz [20]. Figure 3 and Figure 4 summarize the steps outlined above.

CONSTRUCTION SYMBOLS FOR EQUILIBRIUM CLASSES

S \ U	1	2	3	4	5	6	7	8	9	10
1	-	P_R	P_B	P_C	L4	L3	L3	L3	L3	L3
2	P_2	P_4	P_4	P_4	P_5	L3	L4	L3	L4	L3
3	P_3	P_4	P_4	P_4	P_5	L3	L4	L3	L4	L3
4	L6	P_4	P_4	P_4(h)	P_5	L3	L4	L3	L4	L3
5	L5	P_5	P_5	P_5	P_5(h)	L1	L4	L3	L4	L3
6	L5	R / L5	R / L5	R / L5	L2	P_6(h)	L1	L1	L4	L3
7	L5	R / L6	R / L6	R / L6	R / L6	L2	P_7(h)	L1	L1	L1
8	L5	R / L5	R / L5	R / L5	R / L5	L2	L2	P_8(h)	L1	L1
9	L5	R / L6	R / L6	R / L6	R / L6	R / L6	L2	L2	P_9(h)	L1
10	L5	R / L5	R / L5	R / L5	R / L5	R / L5	L2	L2	L2	P_{10}(h)

CONSTRUCTION		SYMBOL	DESCRIPTION	SECTION, SUBSECTION
DIRECT CONSTRUCTION	ANALYTICAL	P_n(h)	PYRAMIDS WITH REGULAR n-GONAL BASE AND SUiTABLY CHOSEN HEIGHT	SS 3.1
		P_C , P_B	MONOSTABLE POLYHEDRA BY CONWAY & GUY, BEZDEK	SS 4.1
		P_2 , P_3	MONO-UNSTABLE POLYHEDRA CONSTRUCTED USING THE IDEA OF CONWAY & GUY	
	COMPUTER SEARCH	P_R	MONOSTABLE POLYHEDRON FOUND BY RESHETOV	
		P_4	TETRAHEDRA	S 1
		P_5	PENTAHEDRA	SS 3.2
POLARITY		R	Construction of polyhedra via polarity, resulting in $(S,U) \to (U,S)$	S 2
LOCAL MANIPULATION	POLYHEDRAL CLASSES	L1	Construction of minimal polyhedra via steps $(S,U) \to (S+1, U+2)$	SS 3.1
		L2	Construction of minimal polyhedra via steps $(S,U) \to (S+2, U+1)$	
	NON-POLYHEDRAL CLASSES	L3	Construction of polyhedra via steps $(S,U) \to (S, U+2)$	SS 3.2, S 4
		L4	Construction of polyhedra via steps $(S,U) \to (S, U+1)$	
		L5	Construction of polyhedra via steps $(S,U) \to (S+2, U)$	
		L6	Construction of polyhedra via steps $(S,U) \to (S+1, U)$	

Figure 3: Summary of the proof. Left panel: Symbols on the (S, U) grid indicate how polyhedra in the given equilibrium class $(S, U)^E$ have been constructed. Dark background corresponds to classes where polyhedra have been identified by computer search. Light grey background corresponds to polyhedral pairs. Symbols are explained in the right panel. For $S, U > 1$ the indicated constructions provide minimal complexity and thus the complexity of the class itself. Hyphen indicates that no polyhedron is known in that class. Right panel: Symbols in the left panel explained briefly with reference to sections, subsections and sub-subsections of the paper.

2 Preliminaries

Before we prove some lemmas that we need for Theorem 1.8, we make a general remark about small perturbations:

Remark 2.1. Observe that

(i) a nondegenerate (stable) equilibrium point s_F on face F of a convex polyhedron P exists if and only if the orthogonal projection s_F of $c(P)$ (the center of mass of P) onto F is in the relative interior of F;

(ii) a vertex q is a nondegenerate (unstable) equilibrium point of P if and only if the plane perpendicular to $q - c(P)$ and containing q contains no other point of P;

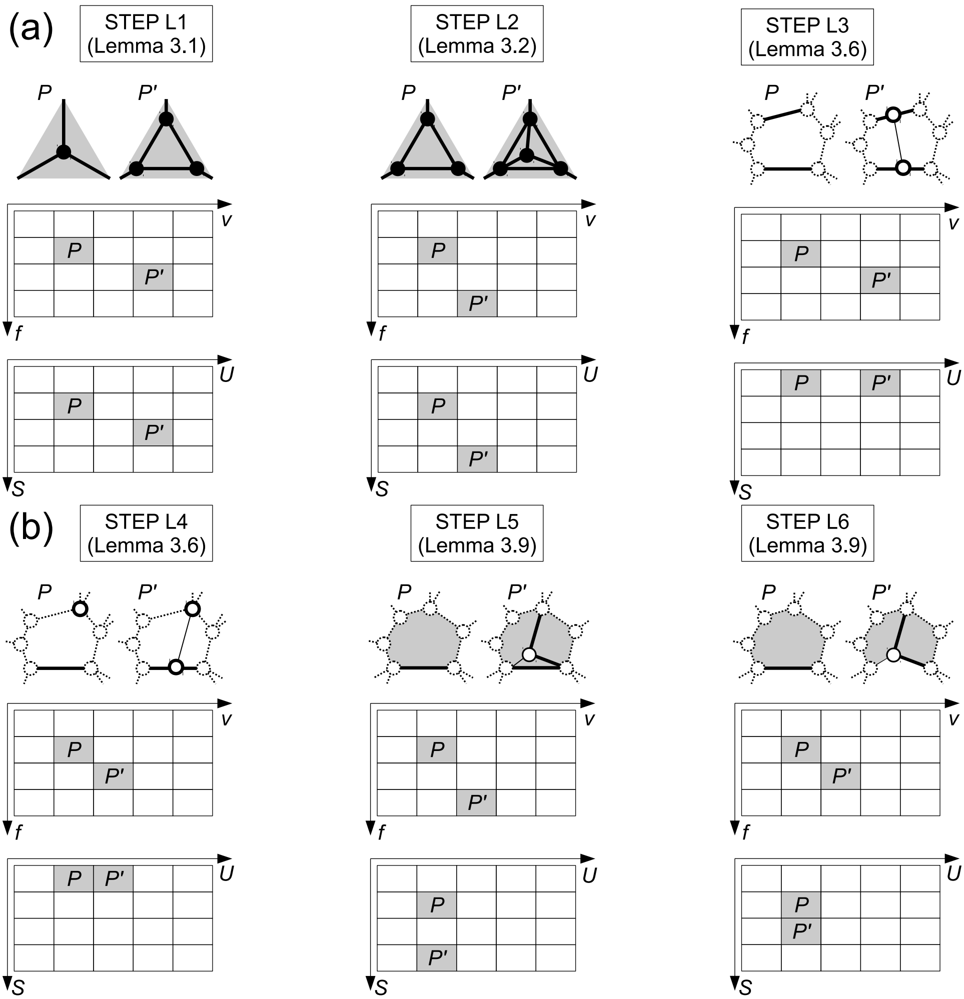


Figure 4: Summary of the proof. (a) – (b): Upper row: schematic picture of local manipulations L1 – L6, showing local face structure and equilibria on original and manipulated polyhedra P and P', respectively. Lower rows: Original and manipulated polyhedra P and P' shown on the (f, v) and (S, U) grids.

(iii) a nondegenerate equilibrium point s_E on an edge E of P exists if and only if the orthogonal projection s_E of $c(P)$ onto E is in the relative interior of E, and the angle between $c(P) - s_E$ and any of the two faces of P containing E is acute.

The subject of our investigation is the family of 3-dimensional convex polyhedra which have only nondegenerate equilibria, and all polyhedra appearing in the paper satisfy this property. Then the following observation is used many times in the paper for some 3-dimensional convex polyhedron P:

If all equilibria are nondegenerate then we will find the same number of equilibria

(a) after applying any sufficiently small perturbation of vertices which leaves the combinatorial structure unchanged on all vertices, edges and faces,

(b) after applying a truncation to the polyhedron with sufficiently small volume, on vertices, edges or faces left unchanged by the truncation and

(c) after applying an augmentation (inverse of truncation) to the polyhedron with sufficiently small volume, on vertices, edges or faces left unchanged by the augmentation.

In the following, $\operatorname{conv} X$, $\operatorname{aff} X$, $\operatorname{int} X$ and $\operatorname{cl} X$ denote the convex hull, the affine hull, the interior and the closure of the set $X \subset \mathbb{R}^d$, respectively. The origin is denoted by o. For any convex polytope P in $\mathbb{R}^d$, we denote by $V(P)$ the set of vertices of P, and the volume and the center of mass of P by $w(P)$ and $c(P)$, respectively. The polar of the set X is denoted by X°.

The first three lemmas investigate the behavior of the center of mass of a convex polyhedron under local deformations.

Lemma 2.2. *Let P be a convex polyhedron and let q be a vertex of P. Let P_ε be a convex polyhedron such that $P_\varepsilon \subset P$, and every point of $P \setminus P_\varepsilon$ is contained in the ε-neighborhood of q. Let $c = c(P)$ and $c_\varepsilon = C(P_\varepsilon)$. Then there is a constant $\gamma > 0$, independent of ε, such that $|c_\varepsilon - c| \leq \gamma\varepsilon^3$ holds for every polyhedron P_ε satisfying the above conditions.*

Proof. Without loss of generality, let $c = o$, $\bar{c}_\varepsilon = c(\operatorname{cl}(P \setminus P_\varepsilon))$, $w = w(P)$ and $w_\varepsilon = w(P_\varepsilon)$. Then $o = w_\varepsilon c_\varepsilon + (w - w_\varepsilon)\bar{c}_\varepsilon$, implying that $c_\varepsilon = -\frac{w-w_\varepsilon}{w_\varepsilon}\bar{c}_\varepsilon$. Note that for some $\gamma' > 0$ independent of ε, we have $0 \leq \frac{w-w_\varepsilon}{w_\varepsilon} < 2\frac{w-w_\varepsilon}{w} \leq \gamma'\varepsilon^3$. Furthermore, for some $\gamma'' > 0$, $|q - \bar{c}_\varepsilon| \leq \gamma''\varepsilon$, which yields that $|\bar{c}_\varepsilon|$ is bounded. Thus, the assertion readily follows. □

Lemma 2.3. *Let F be a triangular face of the convex polyhedron P, and assume that each vertex of P lying in F has degree 3. Let q_1, q_2 and q_3 be the vertices of P on F, and for $i = 1, 2, 3$, let L_i denote the line containing the edge of P through q_i that is not contained in F. For $i = 1, 2, 3$ and $\tau \in \mathbb{R}$, let $q_i(\tau)$ denote the point of L_i at the signed distance τ from q_i, where we orient each L_i in such a way that $q_i(\tau)$ is a point of P for any sufficiently small negative value of τ. Let U be a neighborhood of o, and for any $t = (\tau_1, \tau_2, \tau_3) \in U$, let $W(t) = w(P(t))$ and $C(t) = c(P(t))$, where $P(t) = \operatorname{conv}\left((V(P) \setminus \{q_1, q_2, q_3\}) \cup \{q_1(\tau_1), q_2(\tau_2), q_3(\tau_3)\}\right)$. Then the Jacobian of the function $W(t)C(t)$ is nondegenerate at $t = o$.*

Proof. It is sufficient to show that the partial derivatives of the examined function span $\mathbb{R}^3$. Without loss of generality, we may assume that q_1, q_2 and q_3 are linearly independent.

Consider the polyhedron $P(\tau_1, 0, 0)$ for some $\tau_1 > 0$, and let $\bar{W}(\tau_1) = w(T(\tau_1))$, $\bar{C}(\tau_1) = c(T(\tau_1))$ and $T(\tau_1) = \operatorname{conv}\{q_1, q_2, q_3, q_1(\tau_1)\}$. Let A be the area of the triangle $\operatorname{conv}\{q_1, q_2, q_3\}$. If $\tau_1 > 0$ is sufficiently small, then

$$\left.\frac{\partial}{\partial\tau_1} W(t)C(t)\right|_{t=(0,0,0)} = \frac{\sin\alpha_1 A}{12}(2q_1 + q_2 + q_3),$$

$$W(\tau_1, 0, 0)C(\tau_1, 0, 0) = w(P)c(P) + \bar{W}(\tau_1)\bar{C}(\tau_1).$$

Since $\bar{C}(\tau_1) = \frac{1}{4}(q_1 + q_2 + q_3 + q_1(\tau_1))$, it follows that

$$\left.\frac{\partial}{\partial\tau_1} W(t)C(t)\right|_{t=(0,0,0)} = \frac{\sin\alpha_1 A}{12}(2q_1 + q_2 + q_3),$$

where α_i denotes the angle between L_i and the plane through q_1, q_2, q_3.

Using a similar consideration, we obtain the same formula if $\tau_1 < 0$, and similar formulas, where q_2 or q_3 plays the role of q_1, in the partial derivatives with respect to τ_2 or τ_3, respectively. Note that $0 < \alpha_1, \alpha_2, \alpha_3 \leq \frac{\pi}{2}$. Thus, to show that the three partial derivatives are linearly independent, it suffices to show that the vectors $2q_1 + q_2 + q_3$, $q_1 + 2q_2 + q_3$ and $q_1 + q_2 + 2q_3$ are linearly independent. To show it under the assumption that q_1, q_2, q_3 are linearly independent can be done using elementary computations, which we leave to the reader. □

Remark 2.4. We remark that Lemma 2.3 can be 'dualized' in the following form: Assume that q is a 3-valent vertex of P, and each face of P that q lies on is a triangle. Furthermore, let Y be a neighborhood of q, and for any $x \in Y$, let

$$W(x) = w\left(\operatorname{conv}\left((V(P) \setminus \{q\}) \cup \{x\}\right)\right),$$

and $C(x) = c\left(\operatorname{conv}\left((V(P) \setminus \{q\}) \cup \{x\}\right)\right)$. Then the Jacobian matrix of the function $W(\cdot)C(\cdot) \colon Y \to \mathbb{R}^3$ is nondegenerate at q.

Remark 2.5. If the Jacobian of a smooth vector-valued function in $\mathbb{R}^3$ is nondegenerate, by the Inverse Function Theorem it follows that the function is surjective. Thus, a geometric interpretation of Lemma 2.3 and Remark 2.4 is that under the given conditions, by slight modifications of a vertex or a face of P the function $w(P)c(P)$ moves everywhere within a small neighborhood of its original position.

In the forthcoming two lemmas we investigate the connection between polarity and equilibrium points.

Lemma 2.6. *Let S be a d-dimensional simplex in the Euclidean space $\mathbb{R}^d$ such that $o \in \operatorname{int} S$. Then $o = c(S^\circ)$ if and only if $o = c(S)$.*

Proof. Let the vertices of S be denoted by $p_1, p_2, \ldots, p_{d+1}$. For $i = 1, 2, \ldots, d+1$, let n_i denote the orthogonal projection of o onto the facet hyperplane H_i of S not containing p_i, and let H_i' be the hyperplane through o and parallel to H_i. We remark that since $o \in \operatorname{int} S$, none of the p_is and the n_is is zero. Finally, let α_i denote the angle between p_i and n_i.

Assume that $o = c(S)$. Then for all values of i, we have $\operatorname{dist}(p_i, H_i') = d \operatorname{dist}(H_i', H_i)$, where $\operatorname{dist}(A, B) = \inf\{|a - b| : a \in A, b \in B\}$ is the distance of the sets A and B. This implies that the projection of p_i onto the line through o and n_i is $-dn_i$ for all values of i, or in other words,

$$\cos \alpha_i |p_i| = -d|n_i| \tag{2.1}$$

for all values of i. On the other hand, it is easy to see that if (2.1) holds for all values of i, then $o = c(S)$.

The vertices of S° are the points $p_i^\star = \frac{n_i}{|n_i|^2}$, where $i = 1, 2, \ldots, d+1$, and the projection of o onto the facet hyperplane of P° not containing $p_i^\star$ is $n_i^\star = \frac{p_i}{|p_i|^2}$. Hence, the angle between $p_i^\star$ and $n_i^\star$ is α_i. Similarly like in the previous paragraph, $o = c(S^\circ)$ if and only if

$$\cos \alpha_i |p_i^\star| = -d|n_i^\star| \tag{2.2}$$

holds for all values of i. On the other hand, if $\cos \alpha_i |p_i| = -d|n_i|$ for some value of i, then $\cos \alpha_i |p_i^\star| = \frac{\cos \alpha_i}{|n_i|} = -\frac{d}{|p_i|} = -d|n_i^\star|$, and vice versa. Thus, (2.1) and (2.2) are equivalent, implying Lemma 2.6. □

Lemma 2.7. *Let P be a convex d-polytope in the Euclidean space $\mathbb{R}^d$ such that $o \in \operatorname{int} P$, and let P° be its polar. Let F be a k-face of P, where $0 \leq k \leq d-1$, and let $F^\star$ denote the corresponding $(d-k-1)$-face of P°. Then F contains a nondegenerate equilibrium point of P with respect to o if and only if $F^\star$ contains a nondegenerate equilibrium point of P° with respect to o.*

Proof. Let $F = \operatorname{conv}\{p_i : i \in I\}$, where I is the set of the indices of the vertices of P such that p_i is contained in F, and let p be the orthogonal projection of o onto $\operatorname{aff} F$. Let $L = \operatorname{aff}(F \cup \{o\})$, and let L^c denote the orthogonal complement of L passing through o. For any facet hyperplane of P containing F, let n_j, $j \in J$ denote the projection of o onto this hyperplane. Let H_j^+ be the closed half space $\{q \in \mathbb{R}^d : \langle q, n_j \rangle \leq \langle n_j, n_j \rangle\}$. Let $\bar{H}_i^+ = H_i^+ \cap H$ for any $i \notin I$. Finally, let $\bar{n}_i$ be the component of n_i parallel to H.

Before proving the lemma, we observe that for any given vectors $n_1, n_2, \ldots, n_k$ spanning $\mathbb{R}^d$, the following are equivalent:

(a) o is an interior point of a polytope Q in $\mathbb{R}^d$ with outer facet normals $n_1, n_2, \ldots, n_k$.

(b) There are some $\lambda_1, \lambda_2, \ldots, \lambda_k > 0$ such that $o \in \operatorname{int} Q'$, where

$$Q' = \operatorname{conv}\{\lambda_1 n_1, \lambda_2 n_2, \ldots, \lambda_k n_k\}.$$

(c) We have $o \in \operatorname{int} \operatorname{conv}\{\lambda_1 n_1, \lambda_2 n_2, \ldots, \lambda_k n_k\}$ for any $\lambda_1, \lambda_2, \ldots, \lambda_k > 0$.

We note that if a polytope Q satisfies the conditions in (a), then its polar $Q' = Q^\circ$ satisfies the conditions in (b), and vice versa. Finally, observe that if F contains an equilibrium point, then by exclusion it is p.

We show that p is a nondegenerate equilibrium point of F if and only if it is contained in the relative interiors of the conic hulls of the p_is as well as those of the n_js. First, let p be a nondegenerate equilibrium point. Then $p \in \operatorname{relint} F$, that is, it is in the relative interior of the conic hull (in particular, the convex hull) of the p_is. Observe that since the projection of o onto $\operatorname{aff} F$ is p, for any $j \in J$, the projection of n_j onto $\operatorname{aff} F$ is p. In other words, $n_j \in L' = \operatorname{aff}(L^c \cup \{p\})$ for all $j \in J$. Since p is a vertex of the polytope $P \cap L'$, the vectors n_j, $j \in J$ span this linear subspace, or equivalently, the vectors $\bar{n}_j$ span L^c. Observe that the intersection of P with the affine subspace $(1-\varepsilon)p + L^c$, for sufficiently small values of $\varepsilon > 0$, is a $(d-k-1)$-polytope, with outer facet normals $\bar{n}_j$, $j \in J$, which contains $(1-\varepsilon)p$ in its relative interior. By the observation in the previous paragraph, it follows that o is contained in the relative interior of the convex hull of the $\bar{n}_j$s, which implies that p is contained in the relative interior of the conic hull of the n_js. On the other hand, if p is contained in the relative interior of the conic hull of the p_is, then the fact that $p \in \operatorname{aff} F$ implies that $p \in \operatorname{relint} F$. Furthermore, if p is contained in the relative interior of the conic hull of the n_js, then o is contained in the relative interior of the convex hull of the $\bar{n}_j$s. Thus, the only solution for $q \in L^c$ of the system of linear inequalities $\langle q, \bar{n}_j \rangle \leq 0$, where $j \in J$, is $q = p$, which implies that the only point of P in $p + L^c$ is p. This means that p is a nondegenerate equilibrium point of P.

Finally, observe that the vertices of $F^\star$ are the points $\frac{n_j}{|n_j|^2}$, and the projections of o onto the facet hyperplanes of P° containing $F^\star$ are the points $\frac{p_i}{|p_i|^2}$. Furthermore, $\operatorname{aff} F^\star = \frac{p}{|p|^2} + L^c$, which yields that the projection of o onto $\operatorname{aff} F^\star$ is $\frac{p}{|p|^2}$. Combining it with the consideration in the previous paragraph, this yields the assertion. □

The next corollary is an immediate consequence of Lemmas 2.6 and 2.7 and, together with the result of Conway [7], implies Theorem 1.10.

Corollary 2.8. *Every homogeneous tetrahedron has at least two vertices which are equilibrium points. Furthermore, there are inhomogeneous tetrahedra with exactly one vertex which is an equilibrium point.*

3 Polyhedra with many stable or unstable equilibria: proof of Theorem 1.8

3.1 Proof of Theorem 1.8 for polyhedral pairs

We need to show that if $(S, U)^E$ is a polyhedral class (see the remark after Theorem 1.5), then there is a polyhedron with S faces and U vertices. For brevity, we call such a polyhedron a *minimal polyhedron* in class $(S, U)^E$. We do the construction separately in several cases.

3.1.1 Case 1: $S = U \geq 4$

Let $S \geq 4$, and consider a regular $(S-1)$-gon R_S in the (x, y)-plane, centered at o and with unit inradius. Let $P_v(h)$ be the pyramid with base R_v and apex $(0, 0, h)$. By its symmetry properties, $P_S(h)$ is a minimal polyhedron in the class $(S, S)^E$ for all $h > 0$.

3.1.2 Case 2: $S > 4$ and $S < U \leq 2S - 4$

In this case the proof is based on Lemma 3.1.

Lemma 3.1. *Assume that P is a minimal polyhedron in class $(S, U)^E$ having a vertex of degree 3. Then there is a minimal polyhedron in class $(S + 1, U + 2)^E$ having a vertex of degree 3.*

Proof. Let P be a minimal polyhedron in class $(S, U)^E$ with a vertex q of degree 3. For sufficiently small $\varepsilon > 0$, let $P_\varepsilon \subset P$ be the intersection of P with the closed half space with inner normal vector $c - q$, at the distance ε from q. We show that if ε is sufficiently small, then P_ε satisfies the conditions in the lemma.

If ε is sufficiently small, the boundary of this half space intersects only those edges of P that start at q. Thus, P_ε has one new triangular face F, and three new vertices q_1, q_2, q_3 on F. Since q is not a vertex of P_ε, P_ε has $S + 1$ faces and $U + 2$ vertices. Furthermore, q_1, q_2 and q_3 have degree 3, which means that we need only to show that P_ε is a minimal polyhedron. To do it, we set $c = c(P)$ and $c_\varepsilon = c(P_\varepsilon)$.

Note that by (1.2) and (1.1), every edge of a minimal polyhedron contains an equilibrium point. Thus, by Remark 2.1, if ε is sufficiently small, then every edge of P_ε, apart from those having at least one common point with F, contains an equilibrium point with respect to c_ε. For a sufficiently small ε, clearly, $|c - c_\varepsilon|$ is also small enough such that the edges starting at (but not contained by) F still contain equilibrium points with respect to c_ε. We intend to show that if ε is sufficiently small, then the edges of P_ε in F also contain equilibrium points with respect to c_ε, which, by (1.2) and (1.1) clearly implies that P_ε is a minimal polyhedron.

Consider, e.g. the edge $E = [q_1, q_2]$, and let F_3 be the face of P_ε different from F and containing E. Let h and s be the equilibrium point on E and on F_3, respectively, with respect to c. Let α and β denote the dihedral angles between the planes $\operatorname{aff}(E \cup \{c\})$ and $\operatorname{aff} F$, and the planes $\operatorname{aff}(E \cup \{c\})$ and $\operatorname{aff} F_3$, respectively. The fact that h is an equilibrium point with respect to c is equivalent to saying that the orthogonal projection of c onto the line of E is h, and that $0 < \alpha, \beta < \frac{\pi}{2}$.

Since h is contained in the plane $\operatorname{aff}\{q, c, s\}$ for all values of ε, and it is easy to see that there is some constant $\gamma' > 0$ independent of ε such that $|q_1 - h|, |q_2 - h| \geq \gamma'\varepsilon$. Similarly, an elementary computation shows that for some constant $\gamma'' > 0$ independent of ε, we have $0 < \alpha, \beta \leq \frac{\pi}{2} - \gamma''\varepsilon$. Thus, Lemma 2.2 implies that for small values of ε, E contains an equilibrium point with respect to c_ε, implying that P_ε is a minimal polyhedron. □

Now, consider some class $(S, U)^E$ with $S > 4$ and $S < U \leq 2S - 4$. Then, if we set $k = U - S$ and $S_0 = S - k$, we have $0 < k \leq S - 4$ and $4 \leq S_0$. In other words, $(S, U)^E = (S_0 + k, S_0 + 2k)^E$ for some $S_0 \geq 4$ and $k > 0$. Now, by the proof in Case 1, the class $(S_0, S_0)^E$ contains a minimal polyhedron, e.g. a right pyramid $P_{S_0}(h)$ with a regular $(S_0 - 1)$-gon as its base, where $h > 0$ is arbitrary. Note that the degree of every vertex of $P_{S_0}(h)$ on its base is 3, and thus, applying Lemma 3.1 yields a minimal polyhedron in class $(S_0 + 1, S_0 + 2)^E$ having a vertex of degree 3. Repeating this argument $(k - 1)$ times, we obtain a minimal polyhedron in class $(S, U)^E$.

3.1.3 Case 3: $S > 4$ and $\frac{S}{2} + 2 \leq U < S$

Note that these inequalities are equivalent to $U > 4$ and $U < S \leq 2U - 4$. For the proof in this case we need Lemma 3.2.

Lemma 3.2. *Assume that there is a minimal polyhedron P in class $(S, U)^E$ having a triangular face. Then there is a minimal polyhedron P' in class $(S + 2, U + 1)^E$ having a triangular face F'.*

Proof. Let $c = c(P)$, and let c_F be its orthogonal projection on the plane of F. Since P is a minimal polyhedron, c_F is a relative interior point of F, and an equilibrium point with respect to c (see also Figure 5 for illustration). Let $\bar{c}$ be the centroid of F and define the vector u as $\bar{c} - c_F$. Let v be the outer unit normal vector of F, and for any $0 < \varepsilon$ and $0 \leq \alpha \leq 1$, let $T_{\varepsilon\alpha}$ denote the tetrahedron with base F and apex $q = c_F + \varepsilon v + \alpha u$ such that $T_{\varepsilon\alpha} \cap P = F$. Let $P_{\varepsilon\alpha} = T_{\varepsilon\alpha} \cup P$, $c' = c(P_{\varepsilon\alpha})$, and c'_F be the orthogonal projection of c' on the plane of F. By Remark 2.1, for a sufficiently small ε, equilibrium points on all vertices of $P_{\varepsilon\alpha}$ except q, as well as on all edges and faces of $P_{\varepsilon\alpha}$ not containing q will be preserved.

It is also easy to see from simple geometric considerations that for small values of ε, every face and vertex of $P_{\varepsilon\alpha}$ contains an equilibrium point with respect to c' if q, c'_F and c' are collinear. In the special case of $u = 0$, those points are obviously collinear. In any other case, it is also straightforward to see that $c'_F \in \operatorname{relint} \operatorname{conv}\{c_F, c_F/4 + 3\bar{c}/4\}$. Let us define $d(\alpha) = (c_F + \alpha u - c'_F) \cdot u$. Since d continuously varies with α and $d(0) < 0, d(1) > 0$, for any small $\varepsilon > 0$ there is an α_0 such that apex q and all edges and faces it is contained in have equilibrium points. □

Remark 3.3. Note that the argument also yields a polyhedron P' such that there is an equilibrium point on each face and at every vertex of P' *with respect to the original reference point*: in this case we may choose the value of α in the proof simply as $\alpha = 0$.

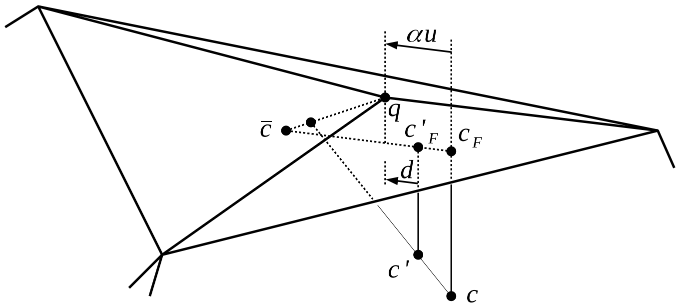


Figure 5: Building a tetrahedron on a triangular face of a convex polyhedron.

Now we prove Theorem 1.8 for Case 3. Like in Case 2, if we set $k = S - U$ and $S_0 = U - k$, then $S_0 \geq 4$, $k > 0$, and $(S, U)^E = (S_0 + 2k, S_0 + k)^E$. Consider the right pyramid $P_{S_0}(h)$ in Case 1. This pyramid has S_0 faces consisting of $S_0 - 1$ triangles and one regular $(S_0 - 1)$-gon shaped face. Thus, applying the construction in Lemma 3.2 k times subsequently yields the desired polyhedron.

3.2 Proof of Theorem 1.8 for non-polyhedral pairs

3.2.1 Case 1: $2 \leq S \leq 4$ and $2 \leq U \leq 4$

Lemma 3.4. *Let $S, U \in \{2, 3, 4\}$. Then $C(S, U) = 2R(S, U)$.*

Proof. Table 1 contains an example for a tetrahedron in each of the 9 classes (illustrated in Figure 6) and for the tetrahedron we have $n = f + v + e = 14$, consequently an upper bound for complexity can be computed as $C(S, U) \leq 14 - S - U - H = 16 - 2S - 2U$. Since from (1.5) we have the same for the lower bound we proved the claim. □

3.2.2 Case 2: $2 \leq S \leq 4, U = 5$ or $2 \leq U \leq 4, S = 5$

This case follows from Lemma 3.5.

Lemma 3.5. *Let $2 \leq S \leq 4, U = 5$ or $2 \leq U \leq 4, S = 5$. Then $C(S, U) = 2R(S, U)$.*

Proof. Table 2 contains an example for a pentahedron in each of the 6 classes (illustrated in Figure 6) and for the pentahedron we have $n = f + v + e = 18$, consequently an upper bound for complexity can be computed as $C(S, U) \leq 18 - S - U - H = 20 - 2S - 2U$. From (1.5) we obtain the same lower bound for all 6 classes so we proved the claim. □

3.2.3 Case 3: $S \geq 5$ and $U > 2S - 4$, or $2 \leq S \leq 4$ and $U \geq 6$

First, we prove the following lemma.

Lemma 3.6. *Let $P \in (S, U)^E$ be a convex polyhedron with f faces and v vertices. Let q_i, $i = 1, \ldots, j$, be successive vertices of an m-gonal ($m \geq, j \geq 3$) face F of P such that*

(i) *the lines $\mathrm{aff}(\{q_1, q_2\})$ and $\mathrm{aff}(\{q_{j-1}, q_j\})$ intersect at some point q with the property $|q - q_1| > |q - q_2|$;*

Class	Non-constant vertex coordinates					Equilibria on													
						faces				vertices				edges					
	C_x	C_y	D_x	D_y	D_z	ABC	ABD	ACD	BCD	A	B	C	D	AB	AC	AD	BC	BD	CD
$(2,2)$	3.2	1.9	−2.2	0.3	1.8	0	0	1	1	0	0	1	1	1	0	0	0	0	1
$(2,3)$	1.9	5.3	1.9	−0.9	5.2	0	0	1	1	1	0	1	1	0	1	1	0	0	1
$(2,4)$	−0.9	5.3	1.9	0.9	5.2	0	0	1	1	1	1	1	1	1	1	0	0	1	1
$(3,2)$	1.0	2.7	−0.9	−4.1	3.4	0	1	1	1	0	0	1	1	0	0	1	0	1	1
$(3,3)$	1.0	5.7	0.5	−0.5	1.3	1	0	1	1	1	0	1	1	0	1	1	1	0	1
$(3,4)$	0.5	2.8	0.5	−0.7	1.2	1	0	1	1	1	1	1	1	0	1	1	1	1	1
$(4,2)$	3.2	3.8	−2.2	−2.9	2.5	1	1	1	1	0	0	1	1	1	1	0	0	1	1
$(4,3)$	1.9	5.3	1.9	5.0	1.8	1	1	1	1	1	0	1	1	0	1	1	1	1	1

Table 1: Examples for tetrahedra in equilibrium classes $(S,U)^E$, $S,U \in \{2,3,4\}$, $(S,U) \neq 4,4$. Constant vertex coordinates for all tetrahedra are $A_x = A_y = A_z = B_y = C_z = 0$, $B_x = 1$.

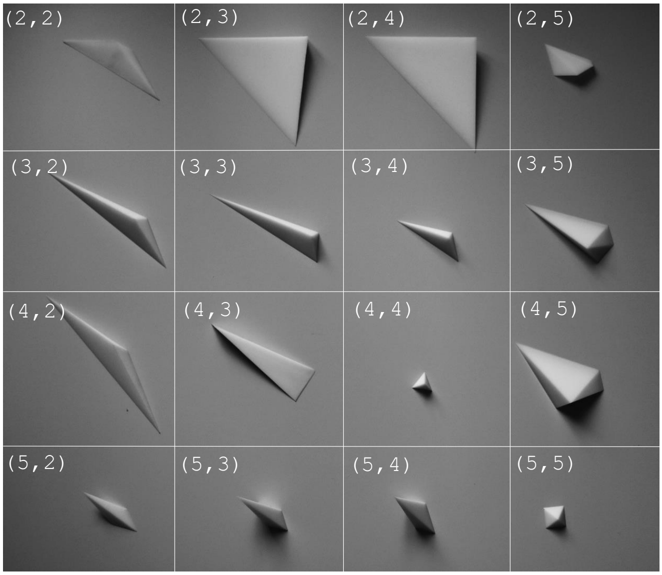


Figure 6: The 8 tetrahedra in Table 1 and the 6 pentahedra in Table 2, the regular tetrahedron and the symmetrical pyramid in equilibrium classes $(S,U)^E$, $S,U \in \{2,3,4,5\}$ produced by 3D printing.

Class	Non-constant vertex coordinates						
	C_x	C_y	D_x	D_y	E_x	E_y	E_z
$(2,5)$	1.0	1.7	0.5	-0.3	2.1	1.2	1.2
$(3,5)$	1.0	1.7	3.8	-2.2	1.6	0.9	0.9
$(4,5)$	2.5	1.4	3.8	-2.2	2.0	1.2	1.2
$(5,2)$	1.0	1.7	0.9	0.5	-0.6	-1.1	-1.1
$(5,3)$	1.0	1.7	0.9	0.5	1.5	2.6	2.6
$(5,4)$	1.0	1.7	1.3	0.8	1.5	2.6	2.6

Table 2: Examples for pentahedra in equilibrium classes $(i,5)$ and $(5,i)$, $i \in \{2,3,4\}$. Constant vertex coordinates for all pentahedra are $A_x = A_y = A_z = B_x = C_z = D_z = 0$, $B_y = 1$.

(ii) *both edges* $E_a = [q_1, q_2]$ *and* $E_b = [q_{j-1}, q_j]$ *contain saddle points;*

(iii) *the vertices* q_i, $i = 2, \dots, j-1$, *are trivalent.*

Then there is convex polyhedron $P' \in (S, U+2)^E$ *with* $f+1$ *faces and* $v+2$ *vertices.*

Proof. Let the saddle points on E_a and E_b be denoted by x_a and x_b. In the proof, based on Remark 2.1, we show that there is an arbitrarily small truncation of P by a plane that intersects F in a line close to x_a and x_b that results in two new unstable vertices u_a and u_b.

We choose a suitable truncation from a 2-parameter family of truncations defined as follows (see also Figure 7 for explanation): For any $t \in [0,1]$, set $y_a(t) = tq_2 + (1-t)q_1$ and $y_b(t) = tq_{j-1} + (1-t)q_j$. Let $G(s,t)$ be the plane that intersects $[q_1, q_2]$ at $y_a(s)$

and $[q_{j-1}, q_j]$ at $y_b(t)$, whose angle with the plane of F is a sufficiently small value $\varepsilon > 0$ (the term 'sufficiently small' is explained in the next paragraph) and truncates the vertices $q_2, q_3, \ldots, q_{j-1}$. For $i = 2, 3, \ldots, j-1$, let $q_i(s,t)$ be the intersection of $G(s,t)$ with the edge of P starting at q_i and not contained in F. Finally, let $P(s,t)$ be the truncation of P by $G(s,t)$, that is, $P(s,t) = \mathrm{cl}(P \setminus \mathrm{conv}\{y_a(s), y_b(t), q_2, \ldots, q_{j-1}, q_2(s,t), \ldots, q_{j-1}(s,t)\}$. We denote the center of mass of $P(s,t)$ by $c(s,t)$, and the projection of c and $c(s,t)$ onto the plane of F by c_F and $c_F(s,t)$, respectively. Furthermore, we denote the new edge of $P(s,t)$ starting at $y_a(s)$ and different from $[y_a(s), y_b(t)]$ by $Y_a(s,t)$, and define $Y_b(s,t)$ similarly.

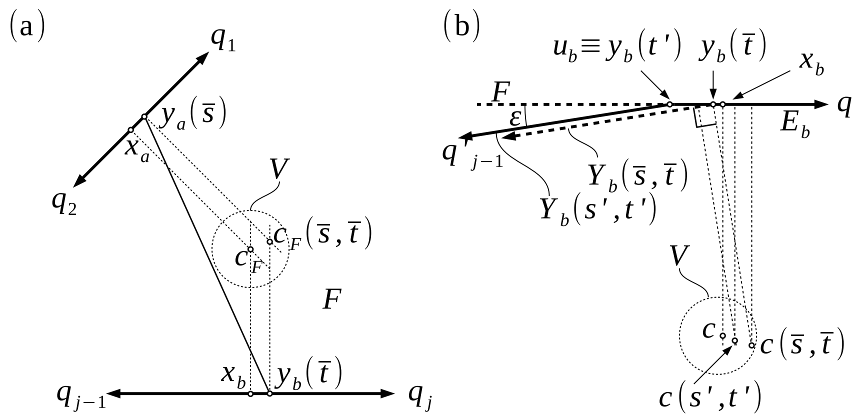


Figure 7: Increasing the number of unstable equilibria by two. Views perpendicular to the plane F (a) and edge $[u_a, u_b]$ (b).

We choose some $\varepsilon > 0$ to satisfy the following conditions: with respect to any point $c' \in V$, the *original polyhedron* P has equilibrium points on the same faces and edges, and at the same vertices, as with respect to the center of mass c of P, where V is the locus of the centers of mass of all truncations of P by the plane $G(s,t)$, $s, t \in [0,1]$ (for a sufficiently small ε, clearly, $|c - c'|$ and the volume removed by truncation are also small enough for the number of original equilibrium points on vertices, edges and faces to be preserved as well; new face and edges included in $G(s,t)$ have no equilibrium points). Furthermore, we assume also that $G(s,t)$ truncates no vertex or equilibrium point of P other than those on F, and that there is some arbitrarily small, fixed value $\delta > 0$ (independent of (s,t)) such that $c(s,t)$ is a Lipschitz function at every (s,t) with Lipschitz constant δ, i.e. $|c(s + \Delta s, t + \Delta t) - c(s,t)| \leq \delta\sqrt{(\Delta s)^2 + (\Delta t)^2}$ for all $s, t \in [0,1]$.

First, we show that for some suitable choice of s and t, the orthogonal projections of $c(s,t)$ onto the lines of E_a and E_b are $y_a(s)$ and $y_b(t)$, respectively. To do this, we use a consequence of Brouwer's fixed point theorem, the so-called Cube Separation Theorem from [27], which states the following: Let the pairs of opposite facets of a d-dimensional cube K be denoted by F_i' and F_i'', $i = 1, 2, \ldots, d$, and let C_i, $i = 1, 2, \ldots, d$, be compact sets such that C_i 'separates' F_i' and F_i'', or in other words, $K \setminus C_i$ is the disjoint union of two open sets Q_i', Q_i'' such that $F_i' \subset Q_i'$, and $F_i'' \subset Q_i''$. Then $\bigcap_{i=1}^{d} C_i \neq \emptyset$.

To apply this theorem, we set $K = \{(s,t) : 0 \leq s, t \leq 1\}$, and define Q_1', C_1 and Q_1'' as the set of pairs (s,t) such that the orthogonal projection of $c(s,t)$ onto the line of

E_a is a relative interior point of $[y_a(s), q_1]$, coincides with $y_a(s)$, or does not belong to $[y_a(s), q_1]$, respectively. We define Q_2', C_2 and Q_2'' similarly. Then these sets satisfy the conditions of theorem, and we obtain a pair $(\bar{s}, \bar{t})$ with the desired property. Note that by the choice of $\varepsilon > 0$, it holds that in a neighborhood of $(\bar{s}, \bar{t})$, the orthogonal projection of $c(s,t)$ onto the line of $Y_a(s,t)$ is in the relative interior of $Y_a(s,t)$, and the same holds also for the projection onto the line of $Y_b(s,t)$. Now we choose some (s', t') sufficiently close to $(\bar{s}, \bar{t})$ such that the intersections of $G(s', t')$ and $G(\bar{s}, \bar{t})$ with F are parallel, and that of $G(s', t')$ is closer to q_2 and q_{j-1} than that of $G(\bar{s}, \bar{t})$. By the Lipschitz property of $c(s,t)$, we have that the distance of the two intersection lines is greater than $|c(s', t') - c(\bar{s}, \bar{t})|$, and hence, the projections of $c(s', t')$ onto the lines of E_a and E_b lie in the relative interior of the segments $[y_a(s'), q_1]$ and $[y_b(t'), q_j]$, respectively. From this it readily follows that both these edges of $P' = P(s', t')$ and also $Y_a(s', t')$ and $Y_b(s', t')$ contain saddle points with respect to $c(s', t')$. This implies also that $y_a(s')$ and $y_b(t')$ are vertices of P' carrying unstable equilibrium points, and the assertion follows. □

Corollary 3.7. *Let conditions (i) and (iii) of Lemma 3.6 hold and (ii) be modified as follows:*

(ii) q_1 contains an unstable and $E_b = [q_{j-1}, q_j]$ contains a saddle-type equilibrium point.

Then there exists a polyhedron $P'' \in (S, U+1)^E$ with $f+1$ faces and $v+1$ vertices.

Remark 3.8. A simplified version of the proof of Lemma 3.6 can be used to prove the same statement for a *fixed* reference point c.

To prove Theorem 1.8 in Case 3, we construct a simple polyhedron with U vertices that has S stable and U unstable points. Since any polyhedron in class $(S,U)^E$ has at least U vertices, and among polyhedra with U vertices those with a minimum number of faces are the simple ones, such a polyhedron clearly has minimal mechanical complexity in class $(S,U)^E$.

First, consider the case that $S \geq 5$ and $U > 2S-4$. Let $U_0 = 2S-4$. By the construction in Subsection 3.1, class $(S, U_0)^E$ contains a simple polyhedron P_0 with U_0 vertices and S faces. Remember that to construct P_0 we started with a tetrahedron T in class $(4,4)^E$, and in each step we truncated a vertex of the polyhedron sufficiently close to this vertex. Throughout the process, the vertex can be chosen as one of those created during the previous step. Since in this case the conditions of Lemma 3.6 are satisfied for any face of P_0, applying Lemma 3.6 to it we obtain a polyhedron P_1 with two more vertices, one more face, two more unstable and the same number of stable points. By subsequently applying the same procedure, we can construct a convex polyhedron in class $(S,U)^E$ for every even value of U. To obtain a polyhedron in class $(S,U)^E$ where U is odd, we can modify a polyhedron in class $(S, U-1)^E$ according to Corollary 3.10.

Now, consider the case that $2 \leq S \leq 4$, and $U \geq 6$. Then, starting with a tetrahedron in class $(S,4)^E$ (based on the data of Table 1, all three tetrahedra meet the conditions of Lemma 3.6) we can repeat the argument in the previous paragraph.

3.2.4 Case 4: $U \geq 5$ and $S > 2U - 4$, or $2 \leq U \leq 4$ and $S \geq 6$

Theorem 1.8 in Case 4 can be deduced from Case 3 using direct geometric properties of polarity. Nevertheless, also the proof in Case 3 via Lemma 3.6 can be dualized as well. In

Lemma 3.9 and Corollary 3.10 we prove dual versions of Lemma 3.6 and Corollary 3.7, respectively, which we are going to use also in Section 4, in our investigation of monostatic polyhedra. Since Theorem 1.8 follows from Lemma 3.9 and Corollary 3.10 similarly like in the proof of Case 3, we leave it to the reader.

We start with the proof using polarity. Considering a tetrahedron T centered at o, a straightforward modification of the construction in Lemma 3.1 and by Remark 3.8, we may construct a simple polyhedron P with U vertices that has S stable and U unstable equilibrium points with respect to o. Using small truncations, we may assume that P is arbitrarily close to T measured in Hausdorff distance. Furthermore, without loss of generality, we may assume that a face of T, and all vertices of this face have degree 3 in P. Let this face of T be denoted by F.

Recall that P° denotes the polar of P. By Lemma 2.6, $c(T^\circ) = o$, and by the continuity of polar and the center of mass, $c(P^\circ)$ is 'close' to o. On the other hand, since the vertex q of P° corresponding to F has degree 3, and each face containing q is a triangle, Lemma 2.3 implies that by a slight modification of q we obtain a polyhedron Q such that $c(Q) = o$, and a face/edge/vertex of Q contains an equilibrium point with respect to o if and only if the corresponding vertex/edge/face of P contains an equilibrium point with respect to o. Thus, Q satisfies the required properties.

As we mentioned, an alternative way to prove Theorem 1.8 in Case 4 is using Lemma 3.9 and Corollary 3.10.

Lemma 3.9. *Let $P \in (S, U)^E$ be a convex polyhedron with f faces and v vertices. Let q_i, $i = 1, \ldots, j-1, j, \ldots, m$ ($j \geq 3$), be successive vertices of an m-gonal ($m \geq 3$) face F of P such that*

(i) *P has a stable equilibrium point c_F on F, which is contained in the relative interior of the triangle $T = \operatorname{conv}\{q_1, q_{j-1}, q_j\}$;*

(ii) *the edge $E = [q_{j-1} q_j]$ contains a saddle-type equilibrium point c_E;*

(iii) *the vertices $q_i, i = 2, \ldots, j-1$ and $i = j+1, \ldots, m$, are trivalent;*

(iv) *q_1, c_F and c_E are not collinear.*

Then there exists a polyhedron $P' \in (S+2, U)^E$ with $f+2$ faces and $v+1$ vertices.

Proof. In the proof, we show that for a sufficiently small pyramid erected over the triangle $T = \operatorname{conv}\{q_1, q_{j-1}, q_j\}$ (which is contained by F and carries a stable equilibrium point) followed by a truncation of P by the plane of the three new faces of the pyramid, results in three new faces instead of F all carrying stable equilibrium and two new edges both carrying saddle-type equilibrium, see Figure 8 (subfigure (a)).

Let the intersection point of the line through q_1 and c_F with E be denoted by x. We choose the apex q of the pyramid from a fixed, sufficiently small neighborhood V of x. Let U be the set of the centers of mass of the modified convex polyhedra, which we denote by $P(q)$. We choose V in such a way that, apart from the three new faces and edges, and the new vertex, $P(q)$ and P have equilibrium points on the same faces and edges, and at the same vertices. Furthermore, we choose V such that for all $q \in V$, the face structure of the resulting polyhedron $P(q)$ is the one described in the previous paragraph, and for any $y \in U$, the Euclidean distance function from y on $[q_{j-1}, q] \cup [q, q_j]$ has a unique

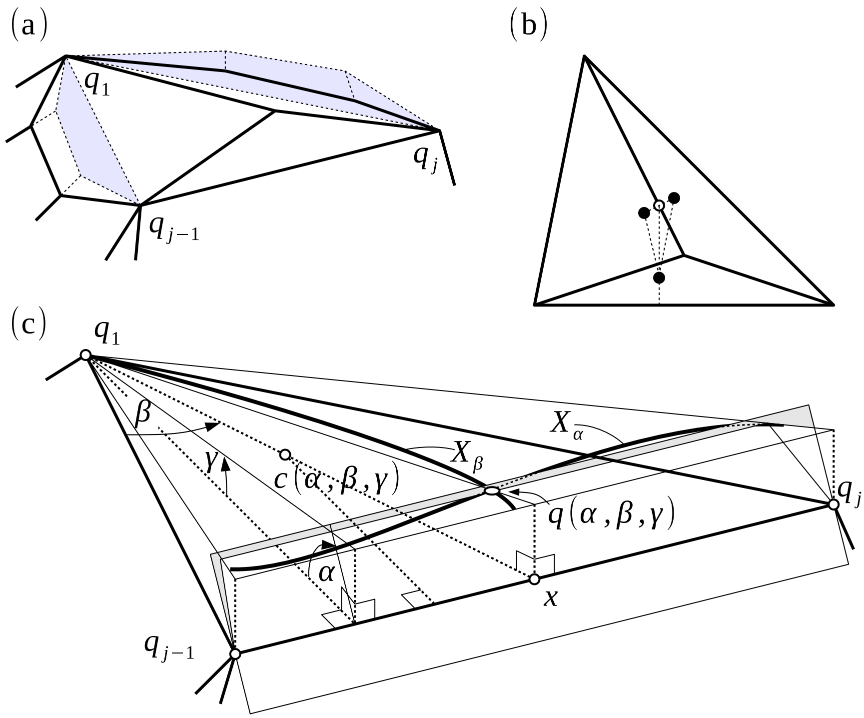


Figure 8: Increasing the number of stable equilibria by two. Schematic view of the pyramid with three light faces instead of the original dark one denoted as F (a); view perpendicular to face F: full circles mean stable equilibrium points, the empty circle is the projection of $c(\alpha, \beta, \gamma)$ onto F (b); illustration for the application of the Cube Separation Theorem for compact sets X_α and X_β (c).

local minimum, and this point is different from q, for all $q \in V$. Note that the latter condition implies that the new vertex q is *not* an unstable equilibrium point. Thus, we need to prove only that, with a suitable choice of q, all the three new faces contain a new stable equilibrium point.

We parametrize q using the following parameters:

- the angle α of the plane of $\operatorname{conv}\{q_{j-1}, q_j, q\}$ and the plane of F. Here we assume that $0 \leq \alpha \leq \alpha_0$, where the sum of α_0 and the dihedral angle of P at E is π;
- the angle β between two rays, both starting at q_1, and containing q_{j-1} and the orthogonal projection q_F of q onto the plane of F, respectively. Here we set $\beta_1 \leq \beta \leq \beta_2$, where $[\beta_1, \beta_2]$ is a sufficiently small interval containing the angle $\angle q_{j-1} q_1 c_F$;
- the angle γ between the ray starting at q_1 and containing q, and the plane of F. Here we assume that $0 < \gamma < \gamma_0$ for some small, fixed value γ_0.

We choose the values of $\beta_1, \beta_2, \gamma_0$ such that in the permitted range of the parameters, $q \in V$. For brevity, we may refer to $P(q(\alpha, \beta, \gamma))$ as $P(\alpha, \beta, \gamma)$, $c(P(\alpha, \beta, \gamma))$ as $c(\alpha, \beta, \gamma)$ and observe that these three quantities determine q.

Note that, using the idea of the proof of Lemma 2.2, we have that $|c(P(q)) - c(P)| = O(\gamma)$, and for some constant $C > 0$ independent of α, β, γ, if $|\alpha' - \alpha| \leq \gamma$, then $|c(\alpha', \beta, \gamma) - c(\alpha, \beta, \gamma)| \leq C\gamma^2$.

Fix some $\gamma > 0$, and let X_α be the set of pairs $(\alpha, \beta) \in [0, \alpha_0] \times [\beta_1, \beta_2]$ such that the planes through E, and containing $c(\alpha, \beta, \gamma)$ and $q(\alpha, \beta, \gamma)$, respectively, are perpendicular. Furthermore, let X_β be the set of pairs $(\alpha, \beta) \in [0, \alpha_0] \times [\beta_1, \beta_2]$ such that q_1, and the projections of $c(\alpha, \beta, \gamma)$ and $q(\alpha, \beta, \gamma)$ onto the plane of F are collinear. If $\gamma > 0$ is sufficiently small, the property $|c(P(q)) - c(P)| = O(\gamma)$ implies that X_α strictly separates the sets $\{(0, \beta) : \beta \in [\beta_1, \beta_2]\}$ and $\{(\alpha_0, \beta) : \beta \in [\beta_1, \beta_2]\}$, and X_β strictly separates the sets $\{(\alpha, \beta_1) : \alpha \in [0, \alpha_0]\}$ and $\{(\alpha, \beta_2) : \alpha \in [0, \alpha_0]\}$. Since X_α and X_β are compact, we may apply the Cube Separation Theorem [27] as in the proof of Lemma 3.6. From this, it follows that there is some $(\alpha_\gamma, \beta_\gamma) \in X_\alpha \cap X_\beta$.

It is easy to see that $(\alpha_\gamma, \beta_\gamma) \in X_\alpha$ implies that for sufficiently small values γ, the polyhedron $P(\alpha_\gamma, \beta_\gamma, \gamma)$ has stable equilibrium points on both faces containing the new edge $[q_1, q(\alpha_\gamma, \beta_\gamma, \gamma)]$. Furthermore, the orthogonal projection of $c(\alpha_\gamma, \beta_\gamma, \gamma)$ onto the plane containing E and $q = q(\alpha_\gamma, \beta_\gamma, \gamma)$ lies on E. Now, let us replace α_γ by $\alpha' = \alpha_\gamma - \gamma$. Then, since in this case $|c(\alpha', \beta_\gamma, \gamma) - c(\alpha_\gamma, \beta_\gamma, \gamma)| \leq C\gamma^2$, we have that if γ is sufficiently small, then the orthogonal projection of $c(\alpha', \beta_\gamma, \gamma)$ onto the face $\mathrm{conv}\{q_{j-1}, q_j, q(\alpha', \beta_\gamma, \gamma)\}$ lies inside the face; that is, P has a stable equilibrium point on this face. This yields the assertion. □

Corollary 3.10. *If all conditions (i)–(iv) of Lemma 3.9 hold, then there is a polyhedron* $P'' \in (S+1, U)^E$ *with* $f+2$ *faces and* $v+1$ *vertices.*

4 Monostatic polyhedra: proof of Theorem 1.9

Our theory of mechanical complexity highlights the special role of polyhedra in the first row and first column of the (S, U) grid. These objects have either only one stable equilibrium point (first row) or just one unstable equilibrium point (first column) and therefore they are called collectively *monostatic*. In particular, the first row is sometimes referred to as *mono-stable* and the first column as *mono-unstable*. Our theory provided only a rough lower bound for their mechanical complexity. While no general upper bound is known, individual constructions provide upper bounds for some particular classes; based on these values one might think that the mechanical complexity of these classes, in particular when both S and U are relatively low, is very high. Monostatic objects have peculiar properties, apparently the overall shape in these equilibrium classes is constrained. In [35] the thinness T and the flatness F of convex bodies is defined ($1 \leq T, F \leq \infty$) and it is shown that, for nondegenerate convex bodies, $T = 1$ if and only if $U = 1$, and $F = 1$ if and only if $S = 1$. This constrained overall geometry may partly account for the high mechanical complexity of monostatic polyhedra.

4.1 Known examples

The first (and probably best) known such object is the monostatic polyhedron P_C constructed by Conway and Guy in 1969 [19] (cf. Figure 9) having mechanical complexity $C(P_C) = 96$. Recently, there have been two additions: the polyhedron P_B by Bezdek [2] (cf. Figure 10) and the polyhedron P_R by Reshetov [28] with respective mechanical complexities $C(P_B) = 64$ and $C(P_R) = 70$. It is apparent that all of these authors were primarily interested in minimizing the number of faces on the condition that there is only one stable equilibrium, so, if one seeks minimal complexity in any of these classes it is possible that these constructions could be improved. Also, as we show below, the same ideas can be used to construct examples of mono-unstable polyhedra. The construction in [19] relies on a delicate calculation for a certain discretized planar spiral, defining a planar polygon P, serving as the basis of a prism which is truncated in an oblique manner (cf. Figure 9). The spiral consists of $2m$ similar right triangles, each having an angle $\beta = \pi/m$ at the point o. The cathetus of the smallest pair of triangles has length r_0, and this will be the vertical height of o when the solid stands in stable equilibrium.

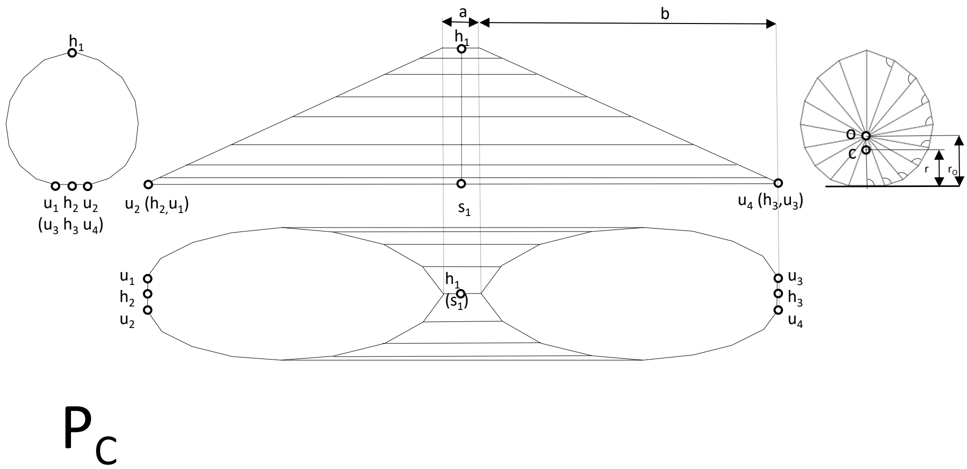


Figure 9: Schematic view of the monostatic polyhedron $P_C \in (1,4)^E, (19,34)^C$ constructed by Conway and Guy in 1969 [19]. Stable, unstable and saddle-type equilibria are marked with s_i, u_j, h_k, $i = 1$, $j = 1,2,3,4$, $k = 1,2,3$, respectively. Complexity can be computed as $C(P_C) = 2(19 + 34 - 1 - 4) = 96$.

We denote the height of the center of mass c by r in the same configuration. It is evident from the construction that if P is a homogeneous planar disc then we have $r > r_0$ since such a disc cannot be monostatic [17]. However, it is also clear that for a non-uniform mass distribution resulting in $r < r_0$, P would be monostatic (cf. Figure 9). In the construction of Conway and Guy we can regard r as a function $r(a,b)$ of the geometric parameters a, b (cf. Figure 9). Apparently, $r(0,b) = r_1$ and $r(a,0) = r_2$ are constants. If P is the aforementioned homogeneous disc then we have $r = r_2 > r_0$. Next we state a corollary to the main result of [19]:

Corollary 4.1. *If* $m \geq 9$ *then* $r_1 < r_0$.

4.2 Examples in $(3,1)^E$ and $(2,1)^E$

Consider a Conway construction with $b = 0$ and denote its vertical centroidal coordinate by r_3: it equals the centroidal coordinate of a plane polygon depicted on the right of Figure 9. Now erect a mirror-symmetric pyramid over the polygon with its apex close to the bottom edge: the vertical coordinate of the body centre of the pyramid will then be close to $3r_3/4$.

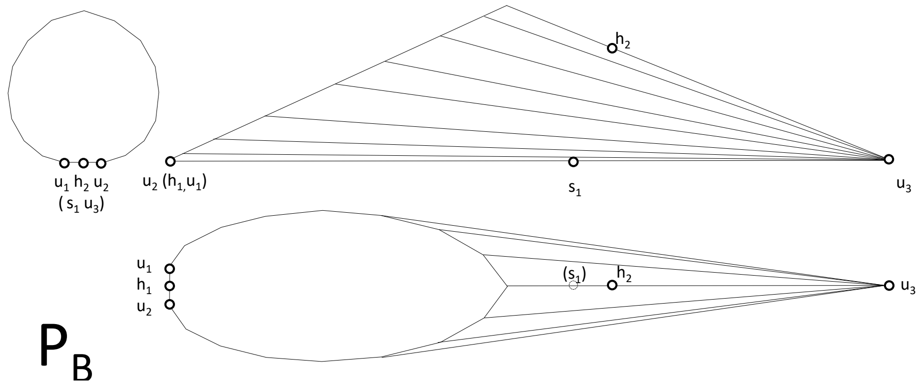


Figure 10: Schematic view of the monostatic polyhedron $P_B \in (1,3)^E, (18,18)^C$ constructed by Bezdek in 2011 [2]. Stable, unstable and saddle-type equilibria are marked with s_i, u_j, h_k, $i = 1$, $j = 1,2,3$, $k = 1,2$, respectively. Complexity can be computed as $C(P_B) = 2(18+18-1-3) = 64$.

It can be shown that for a sufficiently flat pyramid (we call it P_3) will be in classes $(3,1)^E$ and $(18,18)^C$. Introducing a small asymmetry to P_3 by moving the apex off the symmetry plane, a polyhedron P_2 is obtained which belongs to classes $(2,1)^E$ and $(18,18)^C$.

These 'mono-unstable' polyhedra are illustrated in Figure 11. An overview of the discussed monostatic polyhedra is shown in Figure 12 on an overlay of the $(f,v)^C$ and $(S,U)^E$ grids.

4.3 Proof of Theorem 1.9

Proof. Consider the case $C(1,U)$ first. The polyhedron P_C has a narrow rectangular face with a stable point and two saddle points on opposite short edges of the same face. They do not satisfiy condition (i) of Lemma 3.6 because of being collinear, but both 17-gonal faces of P_C can slightly be rotated to get P'_C according to Remark 2.1 in a way that no equilibrium points appear or disappear but the two edges with saddle points become nonparallel, and thus Lemma 3.6 turns to be applicable.

Since the same face of P_C contains four unstable points as well (and none of them is collinear with the stable and any saddle point), Corollary 3.7 can directly be applied to get P_D with $C(P_D) = C(P_C) + 2 = 98$. It means that $C(1,4) \leq 2R(1,4) + 90$ and $C(1,5) \leq 2R(1,5) + 90$. Applying now Lemma 3.6 on both P_C and P_D successively, the assertion readily follows. Note that P_B could not be used as departure instead of P_C, since its saddle points are not on edges of the same face.

A similar path is taken for the case $C(S,1)$. Depart now P_3 with $C(P_3) = 64$: that polyhedron has a 17-gonal face with a stable equilibrium and there is a vertex and an edge on its perimeter having an unstable and a saddle point, respectively. Now it is possible again to slightly rotate the plane of the symmetric triangular face about an axis which is perpendicular to the 17-gon and runs through the apex of the pyramid, making the stable (s_3) and saddle (h_1) point to move off the symmetry axis of the 17-gon, so that they become non-collinear with u_1 (Remark 2.1 guarantees that it can always be done without changing the number of equilibrium points of any kind). Applying or not Corollary 3.10 first then

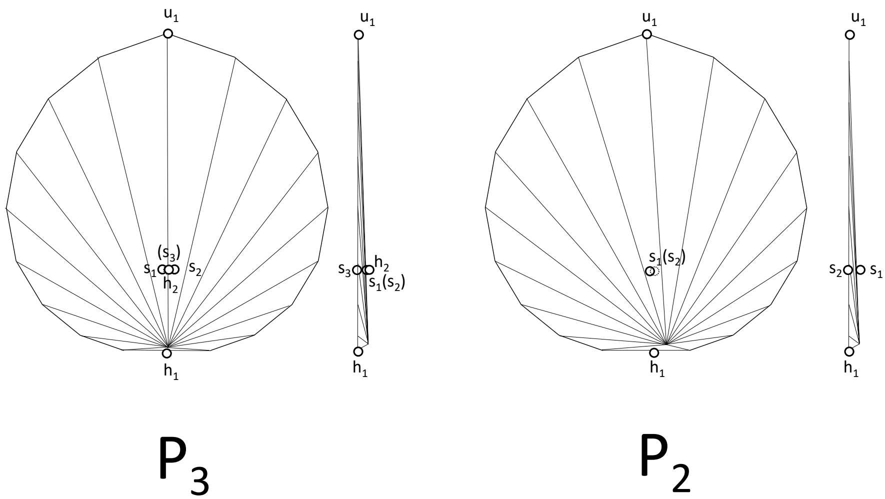


Figure 11: Schematic view of two polyhedra $P_3 \in (3,1)^E, (18,18)^C$ and $P_2 \in (2,1)^E, (18,18)^C$, obtained by using the ideas of the Conway and Bezdek constructions. Stable, unstable and saddle-type equilibria are marked with s_i, u_j, h_k. In case of P_3 we have $i = 1,2,3$, $j = 1$, $k = 1,2$ and in case of P_2 we have $i = 1,2$, $j = 1$, $k = 1$. Complexity can be computed as $C(P_3) = 2(18+18-3-1) = 64, C(P_2) = 2(18+18-2-1) = 66$.

Lemma 3.9 successively gives $C(S,1) \leq S + 61$ and $C(S,1) \leq S + 62$ for odd and even S, respectively, which is equivalent to the second statement of the theorem. □

4.4 Gömböcedron prize

While the construction of monostatic polyhedra with less than 34 edges appears to be challenging (cf. Figure 12), the only case which has been excluded is the tetrahedron with $e = 6$ edges.

It also appears to be very likely that Gömböc-like polyhedra in class $(1,1)^E$ do exist, however, based on this chart and the previous results, one would expect polyhedra with high mechanical complexity. To further motivate this research we offer a prize for establishing the mechanical complexity $C(1,1)$, the amount p of the prize is given in US dollars as

$$p = \frac{10^6}{C(1,1)}.$$

5 Generalizations and applications

5.1 Complexity of secondary equilibrium classes

A special case of Theorem 1.8 states that for any polyhedral pair (f, v) one can construct a homogeneous polyhedron P with f faces and v vertices in such a manner that $C(P) = 0$. In other words, in any primary combinatorial class there exist polyhedra with zero complexity. A natural generalization of this statement is to ask whether this is also true for any *secondary* combinatorial class of convex polyhedra. While we do not have this result,

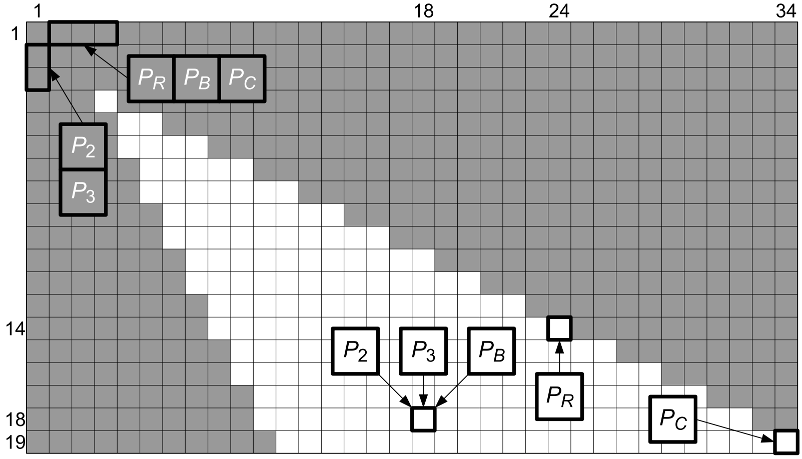


Figure 12: Polyhedra with a single stable or unstable equilibrium point. The grid shown is an overlay of the (f, v) and the (S, U) grids. White squares correspond to polyhedral pairs. Location of monostatic polyhedra is shown with black capital letters on the (f, v) grid and white capital letters on the (S, U) grid. Abbreviations: P_C: Conway and Guy, 1969 [19], P_B: Bezdek, 2011 [2], P_R: Reshetov, 2014 [28]. P_2, P_3: current paper, Figure 11. Complexity for these polyhedra can be readily computed as $C(P_C) = 96, C(P_B) = 64, C(P_R) = 70, C(P_3) = 64, C(P_2) = 66$.

we present an affirmative statement for the inhomogeneous case:

Proposition 5.1. *Let P be a Koebe polyhedron, i.e. a convex polyhedron midscribed (edge-circumscribed) about the unit sphere $\mathbb{S}^2$ with center o. Then every face, edge and vertex of P carries an equilibrium point with respect to o.*

Proof. By (1.1) and (1.2) it is sufficient to show that every edge of P contains an equilibrium point with respect to o.

Let E be an edge of P that touches $\mathbb{S}^2$ at a point q, and let H be the plane touching $\mathbb{S}^2$ at q. Clearly, H is orthogonal to q, and since every face of P intersects the interior of the sphere, we have $H \cap P = E$. Thus, q is an equilibrium point of P with respect to o. □

Here it might be worth noting that for any convex body K and any point $p \in \operatorname{int} K$ there is a density function $\rho\colon K \to [0, \infty)$ such that the center of mass of K with respect to this density function is p, implying that Proposition 5.1 can indeed be reformulated in terms of Koebe polyhedra with inhomogeneous densities. Thus, since a variant of the Circle Packing Theorem [4] states that every combinatorial class contains a Koebe polyhedron, it follows that every combinatorial class contains an inhomogeneous polyhedron with zero mechanical complexity. To find a homogeneous representative appears to be a challenge.

In [24], the author strengthened the result in [4] by showing the existence of a Koebe polyhedron P in each combinatorial class such that the center of mass of the k-dimensional skeleton of P, where $k = 0, 1$ or 2, coincides with o. This result and Proposition 5.1 imply that replacing $c(P)$ by the center of mass of the k-skeleton of a polyhedron with $0 \le k \le 2$, every combinatorial class contains a polyhedron with zero mechanical complexity.

5.2 Inverse type questions

The basic goal of this paper is to explore the nontrivial links between the combinatorial $(f,v)^C$ and the mechanical $(S,U)^E$ classification of convex polyhedra. The concept of mechanical complexity (Definition 1.2) helps to explore the $(S,U)^E \to (f,v)^C$ direction of this link. Inverse type questions may be equally useful to understand this relationship: for example, a natural question to ask is the following: Is it true that any equilibrium class $(S,U)^E$ intersects all but at most finitely many combinatorial classes $(f,v)^C$? Here it is worth noting that it is easy to carry out local deformations on a polyhedron that increase the number of faces and vertices, but not the number of equilibria. Alternatively, one may ask to provide the list of all $(S,U)^E$ classes represented by homogeneous polyhedra in a given combinatorial class $(f,v)^C$. A similar question may be asked for a secondary combinatorial class of polyhedra. In general, we know little about the answers, however we certainly know that (1.3) holds and we also know that $S = f, U = v$ is a part of this list. The minimal values for S and U are less clear. In particular, based on our previous results it appears that the values $S = 1$ and $U = 1$ can be only achieved for sufficiently high values of f, v. On the other hand, Theorem 1.10 and Lemma 3.4 resolve this problem at least for the $(4,4)^C$ class. The latter is based on a global numerical search and this could be done at least for some polyhedral classes, although the computational time grows with exponent $(f+v)$.

5.3 Inhomogeneity and higher dimensions

While here we described only 3D shapes, the generalization of Definitions 1.2 and 1.3 to arbitrary dimensions is straightforward. While the actual values of mechanical complexity are trivial in the planar case (class $(2)^E$ has mechanical complexity 2 and every other equilibrium class has mechanical complexity zero), the $d > 3$ dimensional case appears an interesting question in the light of the results of Dawson et al. on monostatic simplices in higher dimensions [6, 7, 8]. We formulated all our results for homogeneous polyhedra, nevertheless, some remain valid in the inhomogeneous case which also offers interesting open questions. In particular, the universal lower bound (1.3) is independent of the material density distribution so it remains valid for inhomogeneous polyhedra and as a consequence, so does Theorem 1.8. However, our other results (in particular the bounds for monostatic equilibrium classes) are only valid for the homogeneous case. In the latter context it is interesting to note that Conway proved the existence of inhomogeneous, monostatic tetrahedra [7].

5.4 Classification of centric minimal polyhedra

Recall from Subsection 3.1 that a 3-dimensional convex polyhedron is called a minimal polyhedron if its every vertex, edge and face contains an equilibrium point. To further specify these polyhedra, we define a centric minimal polyhedron P as a minimal polyhedron with the additional property that the orthogonal projection of the center of mass of P onto the affine hull of every face/edge of P coincides with the center of mass of the corresponding face/edge of P. It is worth noting that the notions of both minimal polyhedra and centric minimal polyhedra can be defined for d-dimensional convex polytopes in an analogous manner for any $d \geq 2$. Then, for any centric minimal d-polytope P and k-face F, F is a centric minimal k-polytope.

Problem 5.2. Characterize the family of centric minimal polyhedra in $\mathbb{R}^3$.

In Proposition 5.3 we collect some elementary properties of centric minimal polyhedra.

Proposition 5.3. *Let $P \subset \mathbb{R}^d$ with $d \geq 2$ be a d-dimensional convex polytope.*

(i) If $d = 2$, then P is a centric minimal polygon if and only if every vertex of P is at the same distance from the center of mass of P.

(ii) If $d \geq 3$ and P is a centric minimal polytope, then every vertex of P is at the same distance from the center of mass of P.

Proof. Let o be the center of mass of P. Then the orthogonal projection of o onto the line through any edge of P is the midpoint of the edge, which yields that the two endpoints of the edge are at the same distance from o. Thus, both (i) and (ii) follow from the connectedness of the edge graph of P. □

By relaxing the definition of centric minimal polyhedra, we may define a weakly centric d-dimensional minimal polytope P as a minimal polytope such that the orthogonal projection of the center of mass of P onto the affine hull of every $(d-1)$-face of P coincides with the center of mass of the corresponding $(d-1)$-face of P.

Problem 5.4. Characterize the family of weakly centric minimal polyhedra in $\mathbb{R}^3$.

5.5 Applications

Here we describe some problems in mineralogy, geomorphology and industry where the concept of mechanical complexity could potentially contribute to the efficient description and the better understanding of the main phenomena.

5.5.1 Crystal shapes

Crystal shapes are probably the best known examples of polyhedra appearing in Nature and the literature on their morphological, combinatorial and topological classification is substantial [22]. However, as crystals are not just geometric objects but also (nearly homogeneous) 3D solids, their equilibrium classification appears to be relevant. The number of static balance points has been recognized as a meaningful geophysical shape descriptor [11, 18, 33] and it has also been investigated in the context of crystal shapes [32]. The theory outlined in our paper may help to add new aspects to their understanding. While the study of a broader class of crystal shapes is beyond the scope of this paper, we can illustrate this idea in Figure 13 by two examples of quartz crystals with identical number of faces displaying a large difference in mechanical complexity. The length a of the middle, prismatic part of the hexagonal crystal shape (appearing on the left side of Figure 13) is not fixed in the crystal. As we can observe, for sufficiently small values of the length a the crystal will be still in the same combinatorial class $(18, 14)^C$, however, its mechanical complexity will be reduced to zero.

5.5.2 Random polytopes, chipping models and natural fragments

There is substantial literature on the shape of random polytopes [29] which are obtained by successive intersections of planes at random positions. Under rather general assumptions

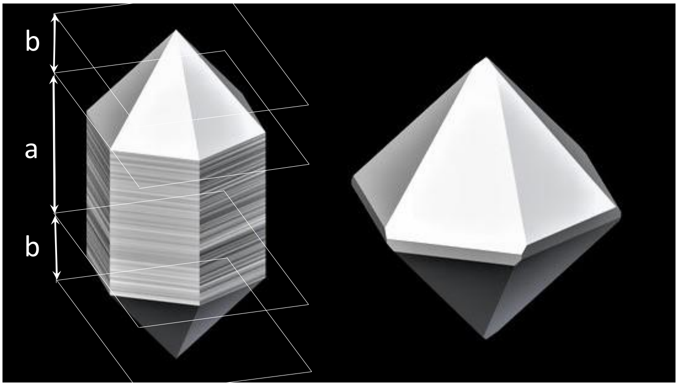


Figure 13: Quartz crystals. Left: Hexagonal habit in classes $(18,14)^C$ and $(6,2)^E$, $C(P) = 48$. Right: Cumberland habit [36] in classes$(18,32)^C$ and $(12,8)^E$, $C(P) = 60$. Picture source [26].

on the distribution of the intersecting planes it can be shown that the *expected* primary combinatorial class of such a random polytope is $(6,8)^C$ (see Theorem 10.3.1 in [29]), however, there are no results on the mechanical complexity. A very special limit of random polytopes can be created if we use a *chipping model* [13, 23] where one polytope is truncated with planes in such a manner that the truncated pieces are small compared to the polytope. Although not much is known about the combinatorial properties of these polytopes, it can be shown [15] that under a sufficiently small truncation the mechanical complexity either remains constant or it increases (this is illustrated in Figure 1). Apparently, random polytopes can be used to approximate natural fragments [12, 14]. There is data available on the number and type of static equilibria of the latter, so any result on the mechanical complexity of random polytopes could be readily tested and also used to identify fragmentation processes.

5.5.3 Assembly processes

In industrial assembly processes parts are processed by a feeder and often these parts can be approximated by polyhedra. These polyhedra arrive in a random orientation on a horizontal surface (tray) and end up ultimately on one of their faces carrying a stable equilibrium. Based on the relative frequency of this position, one can derive *face statistics* and the throughput of a part feeder is heavily influenced by the face statistics of the parts processed by the feeder. Design algorithms for feeders are often investigated from this perspective [3, 34]. It is apparent that one key factor determining the entropy of the face statistics is the mechanical complexity of the polyhedron, in particular, higher mechanical complexity leads to better predictability of the assembly process so this concept may add a useful aspect to the description of this industrial problem.

5.6 Concluding remarks

We showed an elementary connection between the Euler and Poincaré-Hopf formulae (1.1) and (1.2): the mechanical complexity of a polyhedron is determined jointly by its equilibrium class $(S, U)^E$ and combinatorial class $(f, v)^C$. Mechanical complexity appears to be a good tool to highlight the special properties of monostatic polyhedra and offers a new approach to the classification of crystal shapes. We defined polyhedral pairs (x, y) of integers (cf. Definition 1.4) and showed that they play a central role in both classifications: they define all possible combinatorial classes $(f, v)^C$ while in the mechanical classification they correspond to classes with zero complexity.

ORCID iDs

Flórián Kovács https://orcid.org/0000-0002-8374-8035
Zsolt Lángi https://orcid.org/0000-0002-5999-5343

References

[1] J. E. Auerbach and J. C. Bongard, On the relationship between environmental and mechanical complexity in evolved robots, *Artif. Life* **13** (2012), 309–316, doi:10.7551/978-0-262-31050-5-ch041.

[2] A. Bezdek, On stability of polyhedra, 2011, a lecture given at the Fields Institute, Canada, on September 14, 2011 as part of the Workshop on Discrete Geometry (September 13 – 16, 2011), `https://www.fields.utoronto.ca/audio/11-12/wksp_geometry/bezdek/`.

[3] K. F. Bohringer, B. R. Donald, L. E. Kavraki and F. Lamiraux, Part orientation with one or two stable equilibria using programmable vector fields, *IEEE Trans. Robot. Automat.* **16** (2000), 157–170, doi:10.1109/70.843172.

[4] G. R. Brightwell and E. R. Scheinerman, Representations of planar graphs, *SIAM J. Discrete Math.* **6** (1993), 214–229, doi:10.1137/0406017.

[5] J. H. Conway and R. K. Guy, Stability of polyhedra, *SIAM Rev.* **8** (1966), 381–381, doi:10.1137/1008075.

[6] R. Dawson, W. Finbow and P. Mak, Monostatic simplexes II, *Geom. Dedicata* **70** (1998), 209–219, doi:10.1023/a:1004941706441.

[7] R. J. MacG. Dawson, Monostatic simplexes, *Amer. Math. Monthly* **92** (1985), 541–546, doi:10.2307/2323158.

[8] R. J. MacG. Dawson and W. Finbow, Monostatic simplexes III, *Geom. Dedicata* **84** (2001), 101–113, doi:10.1023/a:1010339220243.

[9] R. J. MacG. Dawson and W. A. Finbow, What shape is a loaded die?, *Math. Intelligencer* **21** (1999), 32–37, doi:10.1007/bf03024844.

[10] P. Diaconis and J. B. Keller, Fair dice, *Amer. Math. Monthly* **96** (1989), 337–339, doi:10.2307/2324089.

[11] G. Domokos, Natural numbers, natural shapes, *Axiomathes* (2018), doi:10.1007/s10516-018-9411-5.

[12] G. Domokos, D. J. Jerolmack, F. Kun and J. Török, Plato's cube and the natural geometry of fragmentation, *Proc. Natl. Acad. Sci. USA* **117** (2020), 18178–18185, doi:10.1073/pnas.2001037117.

[13] G. Domokos, D. J. Jerolmack, A. A. Sipos and A. Török, How river rocks round: Resolving the shape-size paradox, *PLoS One* **9** (2014), 1–7, doi:10.1371/journal.pone.0088657.

[14] G. Domokos, F. Kun, A. A. Sipos and T. Szabó, Universality of fragment shapes, *Sci. Rep.* **5** (2015), Article 9147, doi:10.1038/srep09147.

[15] G. Domokos and Z. Lángi, The robustness of equilibria on convex solids, *Mathematika* **60**

(2014), 237–256, doi:10.1112/s0025579313000181.

[16] G. Domokos, Z. Lángi and T. Szabó, A topological classification of convex bodies, *Geom. Dedicata* **182** (2016), 95–116, doi:10.1007/s10711-015-0130-4.

[17] G. Domokos, J. Papadopulos and A. Ruina, Static equilibria of planar, rigid bodies: is there anything new?, *J. Elasticity* **36** (1994), 59–66, doi:10.1007/bf00042491.

[18] G. Domokos, A. A. Sipos, T. Szabó and P. L. Várkonyi, Pebbles, shapes and equilibria, *Math. Geosci.* **42** (2010), 29–47, doi:doi.org/10.1007/s11004-009-9250-4.

[19] M. Goldberg and R. K. Guy, Problem 66-12, Stability of polyhedra (by J. H. Conway and R. K. Guy), *SIAM Rev.* **11** (1969), 78–82, doi:10.1137/1011014.

[20] B. Grünbaum, *Convex Polytopes*, volume 221 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, 2nd edition, 2003, doi:10.1007/978-1-4613-0019-9.

[21] S. Huang and D. Ingber, The structural and mechanical complexity of cell-growth control, *Nat. Cell Biol.* **1** (1999), E131–E138, doi:10.1038/13043.

[22] C. Klein, *Minerals and Rocks: Exercises in Crystal and Mineral Chemistry, Crystallography, X-ray Powder Diffraction, Mineral and Rock Identification, and Ore Msineralogy*, Wiley, 3rd edition, 2007.

[23] P. L. Krapivsky and S. Redner, Smoothing a rock by chipping, *Phys. Rev. E* **75** (2007), 031119 (7 pages), doi:10.1103/physreve.75.031119.

[24] Z. Lángi, Centering Koebe polyhedra via Möbius transformations, *Groups Geom. Dyn.* (2020), in press.

[25] J. W. Milnor, *Topology from the Differentiable Viewpoint*, The University Press of Virginia, Charlottesville, Virginia, 1965.

[26] Mindat.org, Quartz, About Quartz, `https://www.mindat.org/min-3337.html`.

[27] M. M. Postnikov, *Lectures in Geometry, Semester III, Smooth Manifolds*, MIR Publishers, Moscow, 1989, translated from Russian by V. V. Shokurov.

[28] A. Reshetov, A unistable polyhedron with 14 faces, *Internat. J. Comput. Geom. Appl.* **24** (2014), 39–59, doi:10.1142/s0218195914500022.

[29] R. Schneider and W. Weil, *Stochastic and Integral Geometry*, Probability and its Applications (New York), Springer-Verlag, Berlin, 2008, doi:10.1007/978-3-540-78859-1.

[30] E. Steinitz, Über die Eulersche Polyederrelationen, *Arch. Math. Phys.* **11** (1906), 86–88.

[31] E. Steinitz, Polyeder und Raumeinteilungen, in: W. F. Meyer and H. Mohrmann (eds.), *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen, Band III.1.2*, Teubner, Leipzig, pp. 1–139, 1922.

[32] T. Szabó and G. Domokos, A new classification system for pebble and crystal shapes based on static equilibrium points, *Cent. Eur. Geol.* **53** (2010), 1–19, doi:10.1556/ceugeol.53.2010.1.1.

[33] T. Szabó, S. Fityus and G. Domokos, Abrasion model of downstream changes in grain shape and size along the Williams River, Australia, *J. Geophys. Res. Earth Surf.* **118** (2013), 2059–2071, doi:10.1002/jgrf.20142.

[34] P. L. Várkonyi, Estimating part pose statistics with application to industrial parts feeding and shape design: New metrics, algorithms, simulation experiments and datasets, *IEEE Trans. Autom. Sci. Eng.* **11** (2014), 658–667, doi:10.1109/tase.2014.2318831.

[35] P. L. Várkonyi and G. Domokos, Static equilibria of rigid bodies: Dice, pebbles and the Poincaré-Hopf theorem, *J. Nonlinear Sci.* **16** (2006), 255–281, doi:10.1007/s00332-005-0691-8.

[36] J. S. White, Let's get it right: The cumberland habit, *Rocks Miner.* **78** (2003), 196–197, doi:10.1080/00357529.2003.9926720.

[37] Wikipedia contributors, Dice — Wikipedia, The Free Encyclopedia, `http://en.wikipedia.org/wiki/Dice`.

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 125–145
https://doi.org/10.26493/1855-3974.2137.7fa
(Also available at http://amc-journal.eu)

A new family of maximum scattered linear sets in $\mathrm{PG}(1, q^6)$*

Daniele Bartoli

Dipartimento di Matematica e Informatica, Università degli Studi di Perugia, Perugia, Italy

Corrado Zanella

Dipartimento di Tecnica e Gestione dei Sistemi Industriali, Università degli Studi di Padova,Vicenza, Italy

Ferdinando Zullo †

Dipartimento di Matematica e Fisica, Università degli Studi della Campania "Luigi Vanvitelli", Caserta, Italy

Received 5 October 2019, accepted 11 July 2020, published online 12 November 2020

Abstract

We generalize the example of linear set presented by the last two authors in "Vertex properties of maximum scattered linear sets of $\mathrm{PG}(1, q^n)$" (2019) to a more general family, proving that such linear sets are maximum scattered when q is odd and, apart from a special case, they are new. This solves an open problem posed in "Vertex properties of maximum scattered linear sets of $\mathrm{PG}(1, q^n)$" (2019). As a consequence of Sheekey's results in "A new family of linear maximum rank distance codes" (2016), this family yields to new MRD-codes with parameters $(6, 6, q; 5)$.

Keywords: Scattered linear set, MRD-code, linearized polynomial.

Math. Subj. Class. (2020): 51E20, 05B25, 51E22

*The research of all the authors was supported by the Italian National Group for Algebraic and Geometric Structures and their Applications (GNSAGA – INdAM).

†This research of the third author was supported by the project "VALERE: VAnviteLli pEr la RicErca" of the University of Campania "Luigi Vanvitelli" and was partially funded by a fellowship from the Department of Management and Engineering (DTG) of the Padua University.

E-mail addresses: daniele.bartoli@unipg.it (Daniele Bartoli), corrado.zanella@unipd.it (Corrado Zanella), ferdinando.zullo@unicampania.it (Ferdinando Zullo)

1 Introduction

Let $\Lambda = \mathrm{PG}(V, \mathbb{F}_{q^n}) = \mathrm{PG}(1, q^n)$, where V is a vector space of dimension 2 over $\mathbb{F}_{q^n}$. If U is a k-dimensional $\mathbb{F}_q$-subspace of V, then the $\mathbb{F}_q$-*linear set* L_U is defined as

$$L_U = \{\langle \mathbf{u} \rangle_{\mathbb{F}_{q^n}} : \mathbf{u} \in U \setminus \{\mathbf{0}\}\},$$

and we say that L_U has *rank* k. Two linear sets L_U and L_W of $\mathrm{PG}(1, q^n)$ are said to be $\mathrm{P\Gamma L}$-*equivalent* if there is an element ϕ in $\mathrm{P\Gamma L}(2, q^n)$ such that $L_U^{\phi} = L_W$. It may happen that two $\mathbb{F}_q$-linear sets L_U and L_W of $\mathrm{PG}(1, q^n)$ are $\mathrm{P\Gamma L}$-equivalent even if the $\mathbb{F}_q$-vector subspaces U and W are not in the same orbit of $\Gamma\mathrm{L}(2, q^n)$ (see [5, 12] for further details). In this paper we focus on *maximum scattered* $\mathbb{F}_q$-linear sets of $\mathrm{PG}(1, q^n)$, that is, $\mathbb{F}_q$-linear sets of rank n in $\mathrm{PG}(1, q^n)$ of size $(q^n - 1)/(q - 1)$.

If $\langle (0, 1) \rangle_{\mathbb{F}_{q^n}}$ is not contained in the linear set L_U of rank n of $\mathrm{PG}(1, q^n)$ (which we can always assume after a suitable projectivity), then $U = U_f := \{(x, f(x)) : x \in \mathbb{F}_{q^n}\}$ for some linearized polynomial (or *q-polynomial*) $f(x) = \sum_{i=0}^{n-1} a_i x^{q^i} \in \mathbb{F}_{q^n}[x]$. In this case we will denote the associated linear set by L_f. If L_f is scattered, then $f(x)$ is called a *scattered* q-polynomial; see [24].

The first examples of scattered linear sets were found by Blokhuis and Lavrauw in [3] and by Lunardon and Polverino in [18] (recently generalized by Sheekey in [24]). Apart from these, very few examples are known, see Section 3.

In [24, Section 5], Sheekey established a connection between maximum scattered linear sets of $\mathrm{PG}(1, q^n)$ and MRD-codes, which are interesting because of their applications to random linear network coding and cryptography. We point out his construction in the last section. By the results of [1] and [2], it seems that examples of maximum scattered linear sets are rare.

In this paper we will prove that any

$$f_h(x) = h^{q-1} x^q - h^{q^2-1} x^{q^2} + x^{q^4} + x^{q^5}, \quad h \in \mathbb{F}_{q^6}, \quad h^{q^3+1} = -1, \quad q \text{ odd} \tag{1.1}$$

is a scattered q-polynomial. This will be done by considering two cases:

Case 1: $h \in \mathbb{F}_q$, that is, $f_h(x) = x^q - x^{q^2} + x^{q^4} + x^{q^5}$; the condition $h^{q^3+1} = -1$ implies $q \equiv 1 \pmod 4$.

Case 2: $h \notin \mathbb{F}_q$. In this case $h \neq \pm\sqrt{-1}$, otherwise $h \in \mathbb{F}_{q^2}$ and then we have $h^{q+1} = 1$, a contradiction to $h^{q^3+1} = -1$.

Note that in Case 1, this example coincides with the one introduced in [27], where it has been proved that f_h is scattered for $q \equiv 1 \pmod 4$ and $q \leq 29$. In Corollary 3.11 we will prove that the linear set $\mathcal{L}_h$ associated with $f_h(x)$ is new, apart from the case of q a power of 5 and $h \in \mathbb{F}_q$. This solves an open problem posed in [27].

Finally, in Section 4 we prove that the $\mathbb{F}_q$-linear MRD-codes with parameters $(6, 6, q; 5)$ arising from linear sets $\mathcal{L}_h$ are not equivalent to any previously known MRD-code, apart from the case $h \in \mathbb{F}_q$ and q a power of 5; see Theorem 4.1.

2 $\mathcal{L}_h$ is scattered

A *q-polynomial* (or *linearized polynomial*) over $\mathbb{F}_{q^n}$ is a polynomial of the form

$$f(x) = \sum_{i=0}^{t} a_i x^{q^i},$$

where $a_i \in \mathbb{F}_{q^n}$ and t is a positive integer. We will work with linearized polynomials of degree less than or equal to q^{n-1}. For such a kind of polynomial, the *Dickson matrix*[1] $M(f)$ is defined as

$$M(f) := \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1}^q & a_0^q & \cdots & a_{n-2}^q \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{q^{n-1}} & a_2^{q^{n-1}} & \cdots & a_0^{q^{n-1}} \end{pmatrix} \in \mathbb{F}_{q^n}^{n\times n},$$

where $a_i = 0$ for $i > t$.

Recently, different results regarding the number of roots of linearized polynomials have been presented, see [4, 9, 22, 23, 26]. In order to prove that a certain polynomial is scattered, we make use of the following result; see [4, Corollary 3.5].

Theorem 2.1. *Consider the q-polynomial* $f(x) = \sum_{i=0}^{n-1} a_i x^{q^i}$ *over* $\mathbb{F}_{q^n}$ *and, with* m *as a variable, consider the matrix*

$$M(m) := \begin{pmatrix} m & a_1 & \cdots & a_{n-1} \\ a_{n-1}^q & m^q & \cdots & a_{n-2}^q \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{q^{n-1}} & a_2^{q^{n-1}} & \cdots & m^{q^{n-1}} \end{pmatrix}.$$

The determinant of the $(n-i)\times(n-i)$ *matrix obtained by* $M(m)$ *after removing the first* i *columns and the last* i *rows of* $M(m)$ *is a polynomial* $M_{n-i}(m) \in \mathbb{F}_{q^n}[m]$*. Then the polynomial* $f(x)$ *is scattered if and only if* $M_0(m)$ *and* $M_1(m)$ *have no common roots.*

2.1 Case 1

Let

$$f(x) = x^q - x^{q^2} + x^{q^4} + x^{q^5} \in \mathbb{F}_{q^6}[x].$$

By Theorem 2.1, $f(x)$ is scattered if and only if for each $m \in \mathbb{F}_{q^6}$ the determinants of the following two matrices do not vanish at the same time

$$M_5(m) = \begin{pmatrix} 1 & -1 & 0 & 1 & 1 \\ m^q & 1 & -1 & 0 & 1 \\ 1 & m^{q^2} & 1 & -1 & 0 \\ 1 & 1 & m^{q^3} & 1 & -1 \\ 0 & 1 & 1 & m^{q^4} & 1 \end{pmatrix},$$

$$M_6(m) = \begin{pmatrix} m & 1 & -1 & 0 & 1 & 1 \\ 1 & m^q & 1 & -1 & 0 & 1 \\ 1 & 1 & m^{q^2} & 1 & -1 & 0 \\ 0 & 1 & 1 & m^{q^3} & 1 & -1 \\ -1 & 0 & 1 & 1 & m^{q^4} & 1 \\ 1 & -1 & 0 & 1 & 1 & m^{q^5} \end{pmatrix}.$$

[1] This is sometimes called *autocirculant matrix*.

Theorem 2.2. *The polynomial $f(x)$ is scattered if and only if $q \equiv 1 \pmod 4$.*

Proof. If q is even, then for $m = 0$ the matrix $M_6(0)$ has rank two and $f(x)$ is not scattered.

Suppose now $q \equiv 3 \pmod 4$. Then let $\overline{m} \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$ such that $\overline{m}^2 = -4$. So $\overline{m} = \overline{m}^{q^2} = \overline{m}^{q^4} = -\overline{m}^q = -\overline{m}^{q^3} = -\overline{m}^{q^5}$ and, by direct checking,

$$\det(M_5(\overline{m})) = (\overline{m}^2 + 4)^2 = 0, \quad \det(M_6(\overline{m})) = -(\overline{m}^2 + 4)^3 = 0$$

and $f(x)$ is not scattered.

Assume $q \equiv 1 \pmod 4$ and suppose that $f(x)$ is not scattered. Then there exists $m_0 \in \mathbb{F}_{q^6}$ such that

$$(\det(M_5(m_0)))^{q^s} = 0, \quad (\det(M_6(m_0)))^{q^t} = 0, \quad s, t = 0, 1, 2, 3, 4, 5. \tag{2.1}$$

Consider

$$P_1 = \det\begin{pmatrix} 1 & -1 & 0 & 1 & 1 \\ Y & 1 & -1 & 0 & 1 \\ 1 & Z & 1 & -1 & 0 \\ 1 & 1 & U & 1 & -1 \\ 0 & 1 & 1 & V & 1 \end{pmatrix}, \quad P_2 = \det\begin{pmatrix} X & 1 & -1 & 0 & 1 & 1 \\ 1 & Y & 1 & -1 & 0 & 1 \\ 1 & 1 & Z & 1 & -1 & 0 \\ 0 & 1 & 1 & U & 1 & -1 \\ -1 & 0 & 1 & 1 & V & 1 \\ 1 & -1 & 0 & 1 & 1 & W \end{pmatrix}. \tag{2.2}$$

Therefore,

$$X = m_0, \ Y = m_0^q, \ \ldots, \ W = m_0^{q^5} \tag{2.3}$$

is a root of $P_1 =: P_1^{(0)}$, $P_2 =: P_2^{(0)}$ and of the polynomials inductively defined by

$$P_i^{(j)}(X, Y, Z, U, V, W) = P_i^{(j-1)}(Y, Z, U, V, W, X), \quad j = 1, 2, 3, 4, 5, \quad i = 1, 2,$$

which arise from Equation 2.1. These polynomials satisfy

$$\left(P_i^{(j-1)}(m_0, m_0^q, m_0^{q^2}, m_0^{q^3}, m_0^{q^4}, m^{q^5})\right)^q = P_i^{(j)}(m_0, m_0^q, m_0^{q^2}, m_0^{q^3}, m_0^{q^4}, m^{q^5}).$$

One obtains a set S of twelve equations in X, Y, Z, U, V, W having a nonempty zero set. The following arguments are based on the fact that taking the resultant R of two polynomials in S with respect to any variable, the equations $S \cup \{R\}$ admit the same solutions.

We have

$$P_1 = YZUV - YZU - 2YZ + 2YU + 4Y - ZUV + 2ZV - 2UV + 4V + 16 = 0. \tag{2.4}$$

Consider the following resultants:

$$\begin{aligned} Q_1 := Res_V(P_1^{(3)}, P_1) = 2(&XY^2ZU - XY^2ZW + XY^2UW + 2XY^2W \\ &- 2XYZU + 2XYZW - 2XYUW + 8XYW + 8XY - 8XW + 16X \\ &- Y^2ZUW - 2Y^2ZU + 2YZUW - 8YZU - 8YZ + 8YU - 8YW \\ &+ 8ZU - 16Z + 16U - 16W), \end{aligned}$$

$$\begin{aligned} Q_2 := Res_V(P_1^{(4)}, P_1) = &XYZW - XYZ - XYW + 2XZ \\ &- 2XW - 2YZ + 2YW + 4Z + 4W + 16, \end{aligned}$$

$$\begin{aligned} Q_3 := Res_V(P_1^{(5)}, P_1) = &XYZU - XYZ - 2XY + 2XZ \\ &+ 4X - YZU + 2YU - 2ZU + 4U + 16. \end{aligned}$$

They all must be zero, as well as

$$Res_W(Res_U(Q_1, Q_3), Q_2) = 8(YZ - 4)(Y^2 + 4)(X - Z)(XZ + 4)(XY - 4). \quad (2.5)$$

We distinguish a number of cases.

1. Suppose that $Y^2 = -4$. Since $q \equiv 1 \pmod 4$, $X = Y = Z = U = V = W$. So

   $$P_1 = X^4 - 2X^3 + 8X + 16$$

   and the resultant between $X^2 + 4$ and P_1 with respect to X is $2^{27} \neq 0$ and then (2.3) is not a root of P_1, a contradiction.
2. Condition $YZ = 4$ is clearly equivalent to $XY = 4$. This means that $Y = U = W = 4/X$, $Z = V = X$. Therefore, by (2.4) we get $X^2 + 4 = 0$ and we proceed as above.
3. Case $XZ = -4$. In this case $Z = -4/X$, $U = -4/Y$, $V = -4/Z = X$, $W = Y$, $X = Z$ and therefore $X^2 = -4$ and we can proceed as above.
4. Condition $X = Z$ implies $X \in \mathbb{F}_{q^2}$ and so $X = Z = V$ and $Y = U = W$. By substituting in P_1 and P_2,

   $$X^3Y^3 + 3X^3Y - 6X^2Y^2 - 12X^2 + 3XY^3 + 24XY - 12Y^2 - 64 = 0,$$
   $$X^2Y^2 - X^2Y + 2X^2 - XY^2 - 4XY + 4X + 2Y^2 + 4Y + 16 = 0.$$

   Eliminating Y from these two equations one gets

   $$8(X^2 + 4)^6 = 0,$$

   and so $X^2 + 4 = 0$. We proceed as in the previous cases.

This proves that such $m_0 \in \mathbb{F}_{q^6}$ does not exist and the assertion follows. □

2.2 Case 2

We apply the same methods as in Section 2.1. In the following preparatory lemmas (and in the rest of the paper) q is a power of an arbitrary prime p.

Lemma 2.3. *Let $h \in \mathbb{F}_{q^6}$ be such that $h^{q^3+1} = -1$, $h^4 \neq 1$. Then*

1. *$h^q \neq -h$;*
2. *$h^{q^2+1} \neq 1$;*
3. *$h^{q^2+1} \neq \pm h^q$, if q is odd;*
4. *$h^{4q^2+4} + 14h^{2q^2+2q+2} + h^{4q} = 0$ implies $p = 2$ and $h^{q^2-q+1} = 1$ or $q = 3^{2s}$, $s \in \mathbb{N}^*$, $h^{q^2-q+1} = \pm\sqrt{-1}$.*

Proof. The first three are easy computations. Consider now

$$h^{4q^2+4} + 14h^{2q^2+2q+2} + h^{4q} = 0.$$

For $p = 2$ the equation above implies $h^{q^2-q+1} = 1$.

Assume now $p \neq 2$. Since $h \neq 0$, it is equivalent to

$$(h^{q^2-q+1})^4 + 14(h^{q^2-q+1})^2 + 1 = 0,$$

that is $(h^{q^2-q+1})^2 = -7 \pm 4\sqrt{3} = (\sqrt{-3} \pm 2\sqrt{-1})^2$. Let $z = -7 \pm 4\sqrt{3}$. Note that $h^{q^2-q+1} = \pm\sqrt{z}$ belongs to $\mathbb{F}_{q^2}$. We distinguish two cases.

- $\sqrt{z} \in \mathbb{F}_q$. Then

$$-1 = h^{q^3+1} = (h^{q^2-q+1})^{q+1} = \left(\pm\sqrt{z}\right)^{q+1} = z = -7 \pm 4\sqrt{3},$$

  a contradiction if $p \neq 3$. Also, $z = -1$, q is an even power of 3, and $h^{q^2-q+1} = \pm\sqrt{-1}$.

- $\sqrt{z} \notin \mathbb{F}_q$. Then

$$-1 = h^{q^3+1} = (h^{q^2-q+1})^{q+1} = \left(\pm\sqrt{z}\right)^{q+1} = -z = 7 \mp 4\sqrt{3},$$

  a contradiction if $p \neq 2$. □

Lemma 2.4. *Let $h \in \mathbb{F}_{q^6}$ be such that $h^{q^3+1} = -1$, $h^4 \neq 1$. If a root σ of the polynomial*

$$h^{q+1}T^{q+1} + (h^{q^2+q+2} + h^{2q^2+2})T^q + (h^{2q^2+2} - h^{q^2+1})T + h^{q^2+2q+1} + h^{2q^2+q+1} - h^{2q} - h^{q^2+q} \in \mathbb{F}_{q^6}[T]$$

belongs to $\mathbb{F}_{q^6}$, then one of the following cases occurs:

- *$p = 2$, $h^{q^2-q+1} = 1$; or*
- *$q = 3^{2s}$, $s > 0$, $h^{q^2-q+1} = \pm\sqrt{-1}$; or*
- *$\sigma = \pm(h^{q^2} + h^q)$; or*
- *$h \in \mathbb{F}_q$.*

Proof. First, note that $\sigma = 0$ would imply $h^q(h^q+h)^q(h^{q^2+1}-1) = 0$ which is impossible by Lemma 2.3. Therefore $\sigma \neq 0$ and $\sigma^{q^i} = \frac{\ell_i(X)}{m_i(X)}$, where

$$\begin{aligned}
\ell_1(X) &= -(h^{q^2+1} - 1)(h^{q^2+1}X + h^{2q} + h^{q^2+q}) \\
m_1(X) &= h(h^qX + h^{q^2+q+1} + h^{2q^2+1}) \\
\ell_2(X) &= -(h^q + h)(2h^{q^2+q+1}X + h^{2q^2+q+2} + h^{3q^2+2} + h^{3q} + h^{q^2+2q}) \\
m_2(X) &= h^{q+1}(h^{2q^2+2}X + h^{2q}X + 2h^{q^2+2q+1} + 2h^{2q^2+q+1}) \\
\ell_3(X) &= (h^q + h)^q(3h^{2q^2+q+2}X + h^{3q}X + h^{3q^2+q+3} + h^{4q^2+3} + 3h^{q^2+3q+1} \\
&\quad + 3h^{2q^2+2q+1}) \\
m_3(X) &= h^{q^2+q}(h^{3q^2+3}X + 3h^{q^2+2q+1}X + 3h^{2q^2+2q+2} + 3h^{3q^2+q+2} + h^{4q} + h^{q^2+3q})
\end{aligned}$$

$$
\begin{aligned}
\ell_4(X) &= (h^{q^2+1}-1)(h^{4q^2+4}X+6h^{2q^2+2q+2}X+h^{4q}X+4h^{3q^2+2q+3}+4h^{4q^2+q+3}\\
&\quad+4h^{q^2+4q+1}+4h^{2q^2+3q+1})\\
m_4(X) &= h^{q^2}(4h^{3q^2+q+3}X+4h^{q^2+3q+1}X+h^{4q^2+q+4}+h^{5q^2+4}+6h^{2q^2+3q+2}\\
&\quad+6h^{3q^2+2q+2}+h^{5q}+h^{q^2+4q})\\
\ell_5(X) &= -(h^q+h)(h^{5q^2+5}X+10h^{3q^2+2q+3}X+5h^{q^2+4q+1}X+5h^{4q^2+2q+4}\\
&\quad+5h^{5q^2+q+4}+10h^{2q^2+4q+2}+10h^{3q^2+3q+2}+h^{6q}+h^{q^2+5q})\\
m_5(X) &= 5h^{4q^2+q+4}X+10h^{2q^2+3q+2}X+h^{5q}X+h^{5q^2+q+5}+h^{6q^2+5}\\
&\quad+10h^{3q^2+3q+3}+10h^{4q^2+2q+3}+5h^{q^2+5q+1}+5h^{2q^2+4q+1}\\
\ell_6(X) &= (h^q+h)^q(6h^{5q^2+q+5}X+20h^{q^3+3q+3}X+6Xh^{q^2+5q+1}+h^{6q^2+q+6}\\
&\quad+h^{7q^2+6}+15h^{4q^2+3q+4}+15h^{5q^2+2q+4}+15h^{2q^2+5q+2}\\
&\quad+15h^{3q^2+4q+2}+h^{7q}+h^{q^2+6q})\\
m_6(X) &= h^{6q^2+6}X+15h^{4q^2+2q+4}X+15h^{2q^2+4q+2}X+h^{q^6}X+6h^{5q^2+2q+5}\\
&\quad+6h^{6q^2+q+5}+20h^{3q^2+4q+3}+20h^{4q^2+3q+3}+6h^{q^2+6q+1}+6h^{2q^2+5q+1}.
\end{aligned}
$$

Since $\sigma^{q^6}=\sigma$, in particular

$$(h^{2q^2+2}+h^{2q})(h^{4q^2+4}+14h^{2q^2+2q+2}+h^{4q})(h^{q^2}-h^q)(\sigma+h^q+h^{q^2})(\sigma-h^q-h^{q^2})=0.$$

The claim follows from Lemma 2.3. □

Lemma 2.5. *Let* $h\in\mathbb{F}_{q^6}$ *be such that* $h^{q^3+1}=-1$, $h^4=1$. *If a root* σ *of the polynomial*

$$h^{q+1}T^{q^2+1}+(h^q+h)^{q+1}\in\mathbb{F}_{q^6}[T]$$

belongs to $\mathbb{F}_{q^6}$*, then*

$$\sigma=\pm(h^{q^2}+h^q).$$

Proof. If $\sigma=0$, then $h^q+h=0$, a contradiction to Lemma 2.3. So we can suppose $\sigma\neq 0$. Then

$$
\begin{aligned}
\sigma^{q^2} &= -\frac{(h^{q-1}+1)^{q+1}}{\sigma}\\
\sigma^{q^4} &= (h^{q-1}+1)^{q^3+q^2-q-1}\sigma\\
\sigma^{q^6} &= -\frac{(h^{q-1}+1)^{q^5+q^4-q^3-q^2+q+1}}{\sigma}=\frac{(h^q+h)^{2q}}{\sigma}.
\end{aligned}
$$

So, $\sigma=\pm(h^{q^2}+h^q)$. □

Let $h\in\mathbb{F}_{q^6}$ be such that $h^{q^3+1}=-1$, $h^4\neq 1$. By Theorem 2.1 the polynomial

$$f_h(x)=h^{q-1}x^q-(h^{q^2-1})x^{q^2}+x^{q^4}+x^{q^5}$$

is scattered if and only if for each $m \in \mathbb{F}_{q^6}$ the determinant of the following two matrices do not vanish at the same time

$$M_6(m) = \begin{pmatrix} m & h^{q-1} & -h^{q^2-1} & 0 & 1 & 1 \\ 1 & m^q & h^{q^2-q} & h^{-q-1} & 0 & 1 \\ 1 & 1 & m^{q^2} & -h^{-q^2-1} & h^{-q^2-q} & 0 \\ 0 & 1 & 1 & m^{q^3} & h^{1-q} & -h^{1-q^2} \\ h^{q+1} & 0 & 1 & 1 & m^{q^4} & h^{q-q^2} \\ -h^{q^2+1} & h^{q^2+q} & 0 & 1 & 1 & m^{q^5} \end{pmatrix}, \tag{2.6}$$

$$M_5(m) = \begin{pmatrix} h^{q-1} & -h^{q^2-1} & 0 & 1 & 1 \\ m^q & h^{q^2-q} & h^{-q-1} & 0 & 1 \\ 1 & m^{q^2} & -h^{-q^2-1} & h^{-q^2-q} & 0 \\ 1 & 1 & m^{q^3} & h^{1-q} & -h^{1-q^2} \\ 0 & 1 & 1 & m^{q^4} & h^{q-q^2} \end{pmatrix}. \tag{2.7}$$

Theorem 2.6. *Let $h \in \mathbb{F}_{q^6}$, $q = 2^s$, be such that $h^{q^3+1} = 1$. Then the polynomial $f_h(x) = h^{q-1}x^q - (h^{q^2-1})x^{q^2} + x^{q^4} + x^{q^5}$ is not scattered.*

Proof. Consider $\overline{m} = h^{q^2} + h^q$. So,

$$\overline{m}^q = \frac{1}{h} + h^{q^2}, \qquad \overline{m}^{q^2} = \frac{1}{h^q} + \frac{1}{h}, \qquad \overline{m}^{q^3} = \frac{1}{h^{q^2}} + \frac{1}{h^q},$$
$$\overline{m}^{q^4} = h + \frac{1}{h^{q^2}}, \qquad \overline{m}^{q^5} = h^q + h.$$

By direct checking, in this case, both $\det(M_6(\overline{m})) = \det(M_5(\overline{m})) = 0$ and therefore $f_h(x)$ is not scattered. □

Theorem 2.7. *Let $h \in \mathbb{F}_{q^6}$, $q = p^s$, $p > 2$, be such that $h^{q^3+1} = -1$ and $h \notin \mathbb{F}_q$. Then the polynomial $f_h(x) = h^{q-1}x^q - (h^{q^2-1})x^{q^2} + x^{q^4} + x^{q^5}$ is scattered.*

Proof. First we note that $h^4 \neq 1$ since q is odd, $h \notin \mathbb{F}_q$, and $h^{q^3+1} = -1$. Suppose that $f(x)$ is not scattered. Then $\det(M_6(m_0)) = \det(M_5(m_0)) = 0$ for some $m_0 \in \mathbb{F}_{q^6}$. Consider

$$X = m_0, \quad Y = m_0^q, \quad Z = m_0^{q^2}, \quad U = m_0^{q^3}, \quad V = m_0^{q^4}, \quad W = m_0^{q^5}.$$

With a procedure similar to the one in the proof of Theorem 2.2, we will compute resultants starting from the polynomials associated with $\det(M_6(m_0))$, $\det(M_5(m_0))^{q^3}$, and $\det(M_5(m_0))^{q^5}$.

Eliminating W using $\det(M_5(m_0))^{q^3} = 0$ and U using $\det(M_5(m_0))^{q^5} = 0$, one gets from $\det(M_6(m_0)) = 0$

$$h^{q^2+2q+1}\varphi_1(X,Y)\varphi_2(X,Y,Z,V)\varphi_3(X,Y,Z,V) = 0,$$

where

$$\begin{aligned}
\varphi_1(X,Y) &= h^{q+1}XY + h^{2q^2+2}X - h^{q^2+1}X + h^{q^2+q+2}Y + h^{2q^2+2}Y\\
&\quad + h^{q^2+2q+1} + h^{2q^2+q+1} - h^{2q} - h^{q^2+q};\\
\varphi_2(X,Y,Z,V) &= h^{q^2+q+2}XYZV - h^{q^2+q+2}XYZ - h^2XY - h^{q+1}XY\\
&\quad - h^{2q^2+q+1}XZV - h^{2q^2+2}XV - h^{2q^2+q+1}XV - h^{q^2+2q+3}YZ\\
&\quad - h^{2q^2+q+3}YZ - h^{q^2+q+2}Y - h^{2q^2+2}Y - h^{q^2+2q+1}Y\\
&\quad - h^{2q^2+q+1}Y - h^{q^2+2q+1}ZV - h^{2q^2+q+1}ZV - h^{2q^2+q+1}V\\
&\quad - h^{3q^2+1}V - h^{2q^2+2q}V - h^{3q^2+q}V + h^{2q^2+q+3} + h^{3q^2+3}\\
&\quad + h^{2q^2+2q+2} + h^{3q^2+q+2} - 2h^{q^2+q+2} - 2h^{2q^2+2} - 2h^{q^2+2q+1}\\
&\quad - 2h^{2q^2+q+1} + h^{q+1} + h^{q^2+1} + h^{2q} + h^{q^2+q};\\
\varphi_3(X,Y,Z,V) &= h^{q^2+q+2}XYZV + h^{q^2+q+2}XYZ - h^2XY - h^{q+1}XY\\
&\quad + h^{2q^2+q+1}XZV - h^{2q^2+2}XV - h^{2q^2+q+1}XV - h^{q^2+2q+3}YZ\\
&\quad - h^{2q^2+q+3}YZ + h^{q^2+q+2}Y + h^{2q^2+2}Y + h^{q^2+2q+1}Y\\
&\quad + h^{2q^2+q+1}Y - h^{q^2+2q+1}ZV - h^{2q^2+q+1}ZV + h^{2q^2+q+1}V\\
&\quad + h^{3q^2+1}V + h^{2q^2+2q}V + h^{3q^2+q}V + h^{2q^2+q+3} + h^{3q^2+3}\\
&\quad + h^{2q^2+2q+2} + h^{3q^2+q+2} - 2h^{q^2+q+2} - 2h^{2q^2+2} - 2h^{q^2+2q+1}\\
&\quad - 2h^{2q^2+q+1} + h^{q+1} + h^{q^2+1} + h^{2q} + h^{q^2+q}.
\end{aligned}$$

- If $\varphi_1(X,Y) = 0$, then by Lemma 2.4 either $q = 3^{2s}$ and $h^{q^2-q+1} = \pm\sqrt{-1}$, or $X = \pm(h^{q^2} + h^q)$.

  In this last case,

$$\begin{aligned}
&Y = \pm(-h^{-1} + h^{q^2}), \quad Z = \pm(-h^{-q} - h^{-1}), \quad U = \pm\left(-h^{-q^2} - h^{-q}\right)\\
&V = \pm(h - h^{-q^2}), \quad W = \pm(h^q + h).
\end{aligned} \tag{2.8}$$

  By substituting in $\det(M_5(m_0))$ one obtains

$$4(h+h^q)^{q+1}(h^{q^2+1}-1)(h^{q^2+1}-h^q) = 0$$

  and

$$4(h+h^q)^{q+1}(h^{q^2+1}-1)(h^{q^2+1}+h^q) = 0,$$

  respectively. Both are not possible due to Lemma 2.3.

  Consider now the case $q = 3^{2s}$, $h^{q^2-q+1} = \pm\sqrt{-1}$ and $X \neq \pm(h^{q^2} + h^q)$. So, using $\varphi_1(X,Y) = 0$ and $h^{q^2-q+1} = \pm\sqrt{-1}$,

$$\begin{aligned}
&\det(M_5(m_0)) = 0 \Longrightarrow\\
&\quad h^{q^2+2q+1}(h^{q^2}+h^q)(h^q+h)(h^{q^2+1}-1)(h^{q^2+q}+h^q)^3(h^{q^2+q}-h^q)^3\cdot\\
&\quad\cdot(h^{2q^2+2} - h^{q^2+1} + h^{2q})(X + h^q + h^{q^2})^2(X - h^q - h^{q^2})^2 = 0.
\end{aligned}$$

  By Lemma 2.3 we get

$$h^{2q^2+2} - h^{q^2+1} + h^{2q} = 0,$$

  which yields to a contradiction.

- If $\varphi_2(X,Y,Z,V)=0$ and $\varphi_1(X,Y)\neq 0$, eliminating V in $\det(M_5(m_0))=0$ one gets

$$2h^{3q^2+2q+1}(h^{q+2}YZ-h^{q^2+2}-h^{q^2+q+1}+h^q+h)\cdot$$
$$\cdot(hXY+h^{q^2+q+1}+h^{2q^2+1}-h^{q^2}-h^q)\cdot$$
$$\cdot(h^{q+1}XZ+h^{q+1}+h^{q^2+1}+h^{2q}+h^{q^2+q})\cdot$$
$$\cdot(h^{q+2}YZ+hY+h^qY-h^{q^2+q+1}Z+h^qZ-h^{q^2+2}-h^{q^2+q+1}+h^q+h)=0.$$

  - If $h^{q+2}YZ-h^{q^2+2}-h^{q^2+q+1}+h^q+h=0$ then, from

$$Z=\frac{h^{q^2+2}+h^{q^2+q+1}-h^q-h}{h^{q+2}Y},$$

    $\det(M_5)=0$ gives

$$(h^q+h)^{q+1}(hY-h^{q^2+1}+1)(hY+h^{q^2+1}-1)=0.$$

    So, (2.8) holds and as in the case $\varphi_1(X,Y)=0$ a contradiction arises.
  - If $hXY+h^{q^2+q+1}+h^{2q^2+1}-h^{q^2}-h^q=0$ then, from

$$Y=\frac{-h^{q^2+q+1}-h^{2q^2+1}+h^{q^2}+h^q}{hX},$$

    the equation $\det(M_5(m_0))=0$ yields

$$(h^q+h)(h^{q^2+1}-1)(X-h^{q^2}-h^q)(X+h^{q^2}+h^q)=0.$$

    So, (2.8) holds and as in the case $\varphi_1(X,Y)=0$, a contradiction.
  - If $h^{q+1}XZ+h^{q+1}+h^{q^2+1}+h^{2q}+h^{q^2+q}=0$ then by Lemma 2.5

$$(X-h^{q^2}-h^q)(X+h^{q^2}+h^q)=0,$$

    again a contradiction as before.
  - If $h^{q+2}YZ+hY+h^qY-h^{q^2+q+1}Z+h^qZ-h^{q^2+2}-h^{q^2+q+1}+h^q+h=0$ then

$$Z=-\frac{(h^q+h)Y-h^{q^2+2}-h^{q^2+q+1}+h^q+h}{h^{q+2}Y-h^{q^2+q+1}+h^q}.$$

    So, substituting $U=Z^q, V=Z^{q^2}, W=Z^{q^3}, X=Z^{q^4}$ in $\det(M_5(m_0))=0$ we get

$$(h-1)^{q+1}(h+1)^{q+1}(h^q+h)^{q+1}(h^{q^2+1}-1)\cdot$$
$$\cdot(hY-h^{q^2+1}+1)^2(hY+h^{q^2+1}-1)^2=0.$$

    By Lemma 2.3, $(hY-h^{q^2+1}+1)(hY+h^{q^2+1}-1)=0$. Since $Y=\pm(h^{q^2}-1/h)$ then (2.8) holds and a contradiction arises as in the case $\varphi_1(X,Y)=0$.

- If $\varphi_3(X,Y,Z,V)=0$ and $\varphi_1(X,Y)\neq 0$, eliminating U from $\det(M_5(m_0))=0=\det(M_5(m_0))^{q^5}$ and then eliminating V using $\varphi_3(X,Y,Z,V)=0$ one gets

$$2h^{3q^2+q+1}(h^q+h)^q(h^{q+2}YZ-h^{q^2+2}-h^{q^2+q+1}+h^q+h)^2\cdot \\ \cdot(hXY+h^{q^2+q+1}+h^{2q^2+1}-h^{q^2}-h^q)\cdot \\ \cdot(h^{q+1}XZ+h^{q+1}+h^{q^2+1}+h^{2q}+h^{q^2+q})=0.$$

A contradiction follows as in the case $\varphi_2(X,Y,Z,V)=0$ and $\varphi_1(X,Y)\neq 0$. □

3 The equivalence issue

We will deal with the linear sets $\mathcal{L}_h = L_{f_h}$ associated with the polynomials defined in (1.1). Note that when $h\in\mathbb{F}_q$, such a linear set coincide with the one introduced in [27, Section 5].

3.1 Preliminary results

We start by listing the non-equivalent (under the action of $\Gamma\mathrm{L}(2,q^6)$) maximum scattered subspaces of $\mathbb{F}_{q^6}^2$, i.e. subspaces defining maximum scattered linear sets.

Example 3.1.

1. $U^1 := \{(x,x^q) : x\in\mathbb{F}_{q^6}\}$, defining the linear set of pseudoregulus type, see [3, 11];
2. $U^2_\delta := \{(x,\delta x^q + x^{q^5}) : x\in\mathbb{F}_{q^6}\}$, $\mathrm{N}_{q^6/q}(\delta)\notin\{0,1\}$, defining the linear set of LP-type, see [16, 18, 20, 24];
3. $U^3_\delta := \{(x,x^q+\delta x^{q^4}) : x\in\mathbb{F}_{q^6}\}$, $\mathrm{N}_{q^6/q^3}(\delta)\notin\{0,1\}$, satisfying further conditions on δ and q, see [6, Theorems 7.1 and 7.2] and [23] [2];
4. $U^4_\delta := \{(x,x^q+x^{q^3}+\delta x^{q^5}) : x\in\mathbb{F}_{q^6}\}$, q odd and $\delta^2+\delta=1$, see [10, 21].

In order to simplify the notation, we will denote by L^1 and L^i_δ the $\mathbb{F}_q$-linear set defined by U^1 and U^i_δ, respectively. We will also use the following notation:

$$\mathcal{U}_h := U_{h^{q-1}x^q - h^{q^2-1}x^{q^2}+x^{q^4}+x^{q^5}}.$$

Remark 3.2. Consider the non-degenerate symmetric bilinear form of $\mathbb{F}_{q^6}$ over $\mathbb{F}_q$ defined by

$$\langle x,y\rangle = \mathrm{Tr}_{q^6/q}(xy),$$

for each $x,y\in\mathbb{F}_{q^6}$. Then the *adjoint* $\hat f$ of the linearized polynomial $f(x)=\sum_{i=0}^5 a_i x^{q^i}\in\tilde{\mathcal{L}}_{6,q}$ with respect to the bilinear form $\langle\ ,\ \rangle$ is

$$\hat f(x)=\sum_{i=0}^{5} a_i^{q^{6-i}} x^{q^{6-i}},$$

i.e.

$$\mathrm{Tr}_{q^6/q}(xf(y))=\mathrm{Tr}_{q^6/q}(y\hat f(x)),$$

for any $x,y\in\mathbb{F}_{q^6}$.

[2] Here $q>2$, otherwise it is not scattered.

In [10, Propositions 3.1, 4.1 and 5.5] the following result has been proved.

Lemma 3.3. *Let L_f be one of the maximum scattered of $\mathrm{PG}(1,q^6)$ listed before. Then a linear set L_U of $\mathrm{PG}(1,q^6)$ is $\mathrm{P\Gamma L}$-equivalent to L_f if and only if U is $\Gamma\mathrm{L}$-equivalent either to U_f or to $U_{\hat{f}}$ Furthermore, L_U is $\mathrm{P\Gamma L}$-equivalent to L^3_δ if and only if U is $\Gamma\mathrm{L}$-equivalent to U^3_δ.*

We will work in the following framework. Let $x_0,\dots,x_5$ be the homogeneous coordinates of $\mathrm{PG}(5,q^6)$ and let

$$\Sigma=\{\langle(x,x^q,\dots,x^{q^5})\rangle_{\mathbb{F}_{q^6}} : x\in\mathbb{F}_{q^6}\}$$

be a fixed canonical subgeometry of $\mathrm{PG}(5,q^6)$. The collineation $\hat{\sigma}$ of $\mathrm{PG}(5,q^6)$ defined by $\langle(x_0,\dots,x_5)\rangle^{\hat{\sigma}}_{\mathbb{F}_{q^6}}=\langle(x_5^q,x_0^q,\dots,x_4^q)\rangle_{\mathbb{F}_{q^6}}$ fixes precisely the points of Σ. Note that if σ is a collineation of $\mathrm{PG}(5,q^6)$ such that $\mathrm{Fix}(\sigma)=\Sigma$, then $\sigma=\hat{\sigma}^s$, with $s\in\{1,5\}$.

Let Γ be a subspace of $\mathrm{PG}(5,q^6)$ of dimension $k\geq 0$ such that $\Gamma\cap\Sigma=\emptyset$, and $\dim(\Gamma\cap\Gamma^\sigma)\geq k-2$. Let r be the least positive integer satisfying the condition

$$\dim(\Gamma\cap\Gamma^\sigma\cap\Gamma^{\sigma^2}\cap\dots\cap\Gamma^{\sigma^r})>k-2r. \tag{3.1}$$

Then we will call the integer r the *intersection number of* Γ w.r.t. σ and we will denote it by $\mathrm{intn}_\sigma(\Gamma)$; see [27].

Note that if $\hat{\sigma}$ is as above, then $\mathrm{intn}_{\hat{\sigma}}(\Gamma)=\mathrm{intn}_{\hat{\sigma}^5}(\Gamma)$ for any Γ.

As a consequence of the results of [11, 27] we have the following result.

Result 3.4. *Let L be a scattered linear set of $\Lambda=\mathrm{PG}(1,q^6)$ which can be realized in $\mathrm{PG}(5,q^6)$ as the projection of $\Sigma=\mathrm{Fix}(\sigma)$ from $\Gamma\simeq\mathrm{PG}(3,q^6)$ over Λ. If $\mathrm{intn}_\sigma(\Gamma)\neq 1,2$, then L is not equivalent to any linear set neither of pseudoregulus type nor of LP-type.*

3.2 $\mathcal{L}_h$ is new in most of the cases

The linear set $\mathcal{L}_h$ can be obtained by projecting the canonical subgeometry

$$\Sigma=\{\langle(x,x^q,x^{q^2},x^{q^3},x^{q^4},x^{q^5})\rangle_{\mathbb{F}_{q^6}} : x\in\mathbb{F}^*_{q^6}\}$$

from

$$\Gamma\colon\begin{cases}x_0=0\\ h^{q-1}x_1-h^{q^2-1}x_2+x_4+x_5=0\end{cases}$$

to

$$\Lambda\colon\begin{cases}x_1=0\\ x_2=0\\ x_3=0\\ x_4=0.\end{cases}$$

Then

$$\Gamma^{\hat{\sigma}}\colon\begin{cases}x_1=0\\ h^{q^2-q}x_2+h^{-q-1}x_3+x_5+x_0=0\end{cases}$$

and

$$\Gamma^{\hat{\sigma}^2}\colon \begin{cases} x_2 = 0 \\ -h^{-1-q^2}x_3 + h^{-q^2-q}x_4 + x_0 + x_1 = 0. \end{cases}$$

Therefore,

$$\Gamma \cap \Gamma^{\hat{\sigma}}\colon \begin{cases} x_0 = 0 \\ x_1 = 0 \\ -h^{q^2-1}x_2 + x_4 + x_5 = 0 \\ h^{q^2-q}x_2 + h^{-q-1}x_3 + x_5 = 0 \end{cases}$$

and

$$\Gamma \cap \Gamma^{\hat{\sigma}} \cap \Gamma^{\hat{\sigma}^2}\colon \begin{cases} x_0 = 0 \\ x_1 = 0 \\ x_2 = 0 \\ x_4 + x_5 = 0 \\ h^{-q-1}x_3 + x_5 = 0 \\ -h^{-q^2-1}x_3 + h^{-q^2-q}x_4 = 0. \end{cases}$$

Hence, $\dim_{\mathbb{F}_{q^6}}(\Gamma\cap\Gamma^{\hat{\sigma}}) = 1$ and $\dim_{\mathbb{F}_{q^6}}(\Gamma\cap\Gamma^{\hat{\sigma}}\cap\Gamma^{\hat{\sigma}^2}) = -1$, since q is odd and $h^{q^3+1} \neq 1$. So, $\mathrm{intn}_\sigma(\Gamma) = 3$ and hence, by Result 3.4 it follows that $\mathcal{L}_h$ is not equivalent neither to L^1 nor to L^2_δ.

Generalizing [27, Propositions 5.4 and 5.5] we have the following two propositions.

Proposition 3.5. *The linear set $\mathcal{L}_h$ is not $\mathrm{P\Gamma L}$-equivalent to L^3_δ.*

Proof. By Lemma 3.3, we have to check whether $\mathcal{U}_h$ and U^3_δ are $\Gamma\mathrm{L}$-equivalent, with $\mathrm{N}_{q^6/q^3}(\delta) \notin \{0,1\}$. Suppose that there exist $\rho \in \mathrm{Aut}(\mathbb{F}_{q^6})$ and an invertible matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ such that for each $x \in \mathbb{F}_{q^6}$ there exists $z \in \mathbb{F}_{q^6}$ satisfying

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x^\rho \\ h^{\rho(q-1)}x^{\rho q} - h^{\rho(q^2-1)}x^{\rho q^2} + x^{\rho q^4} + x^{\rho q^5} \end{pmatrix} = \begin{pmatrix} z \\ z^q + \delta z^{q^4} \end{pmatrix}.$$

Equivalently, for each $x \in \mathbb{F}_{q^6}$ we have[3]

$$\begin{aligned} cx^\rho + d(h^{q-1}x^{\rho q} - h^{q^2-1}x^{\rho q^2} + x^{\rho q^4} + x^{\rho q^5}) = \\ a^q x^{\rho q} + b^q(h^{q^2-q}x^{\rho q^2} + h^{-q-1}x^{\rho q^3} + x^{\rho q^5} + x^\rho) \\ + \delta[a^{q^4}x^{\rho q^4} + b^{q^4}(h^{-q^2+q}x^{\rho q^5} - h^{q+1}x^\rho + x^{\rho q^2} + x^{\rho q^3})]. \end{aligned}$$

This is a polynomial identity in x^ρ and hence we have the following relations:

$$\begin{cases} c = b^q + \delta h^{q+1}b^{q^4} \\ dh^{q-1} = a^q \\ -dh^{q^2-1} = h^{q^2-q}b^q + \delta b^{q^4} \\ 0 = h^{-1-q}b^q + \delta b^{q^4} \\ d = \delta a^{q^4} \\ d = b^q + \delta h^{q-q^2}b^{q^4}. \end{cases} \tag{3.2}$$

[3]We may replace h^ρ by h, since $h^{q^3+1} = -1$ if and only if $(h^\rho)^{q^3+1} = -1$.

From the second and the fifth equations, if $a \neq 0$ then $\delta h^{q-1} = a^{q-q^4}$ and $\mathrm{N}_{q^6/q^3}(\delta) = 1$, which is not possible and so $a = d = 0$ and $b, c \neq 0$. By the last equation, we would get $\mathrm{N}_{q^6/q^3}(\delta) = 1$, a contradiction. □

Proposition 3.6. *The linear set $\mathcal{L}_h$ is* $\mathrm{P\Gamma L}$*-equivalent to L^4_δ (with $\delta^2 + \delta = 1$) if and only if there exist $a, b, c, d \in \mathbb{F}_{q^6}$ and $\rho \in \mathrm{Aut}(\mathbb{F}_{q^6})$ such that $ad - bc \neq 0$ and either*

$$\begin{cases} c = b^q - \delta k^{q^2+1} b^{q^5} \\ a = -k^{q+1} b^{q^4} - \delta^q b^{q^2} \\ d = k^{-q+1} b^{q^3} + \delta b^{q^5} \\ b^{q^3} + (k^{q-1} + \delta k^{q+q^2}) b^{q^5} = 0 \\ k^{q^2-q} b^q + (1 + k^{q^2-q}) b^{q^3} + \delta k^{q^2-1} b^{q^5} = 0 \\ -\delta b^q + (k^{-q+1} + \delta^2 k^{1-q^2}) b^{q^3} + \delta b^{q^5} = 0 \end{cases} \tag{3.3}$$

or

$$\begin{cases} c = \delta b^q - k^{q^2+1} b^{q^5} \\ a = -\delta^q k^{q+1} b^{q^4} - b^{q^2} \\ d = k^{-q+1} b^{q^3} + b^{q^5} \\ \delta b^{q^3} + (k^{q-1} - \delta k^{q^2+q}) b^{q^5} = 0 \\ \delta k^{q^2-q} b^q + (k^{q^2-q} + 1) b^{q^3} + k^{q^2-1} b^{q^5} = 0 \\ \delta^2 b^q + (k^{-q+1} + \delta^2 k^{-q^2+1}) b^{q^3} + b^{q^5} = 0, \end{cases} \tag{3.4}$$

where $k = h^\rho$.

Proof. By Lemma 3.3 we have to check whether $\mathcal{U}_h$ is equivalent either to U^4_δ or to $(U^4_\delta)^\perp$. Suppose that there exist $\rho \in \mathrm{Aut}(\mathbb{F}_{q^6})$ and an invertible matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ such that for each $x \in \mathbb{F}_{q^6}$ there exists $z \in \mathbb{F}_{q^6}$ satisfying

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x^\rho \\ h^{\rho(q-1)} x^{\rho q} - h^{\rho(q^2-1)} x^{\rho q^2} + x^{\rho q^4} + x^{\rho q^5} \end{pmatrix} = \begin{pmatrix} z \\ z^q + z^{q^3} + \delta z^{q^5} \end{pmatrix}.$$

Equivalently, for each $x \in \mathbb{F}_{q^6}$ we have

$$\begin{aligned} cx^\rho + d(k^{q-1} x^{\rho q} - k^{q^2-1} x^{\rho q^2} + x^{\rho q^4} + x^{\rho q^5}) = \\ a^q x^{\rho q} + b^q (k^{q^2-q} x^{\rho q^2} + k^{-1-q} x^{\rho q^3} + x^{\rho q^5} + x^\rho) \\ + a^{q^3} x^{\rho q^3} + b^{q^3} (k^{-q+1} x^{\rho q^4} - k^{-q^2+1} x^{\rho q^5} + x^{\rho q} + x^{\rho q^2}) \\ + \delta [a^{q^5} x^{\rho q^5} + b^{q^5} (-k^{1+q^2} x^\rho + k^{q^2+q} x^{\rho q} + x^{\rho q^3} + x^{\rho q^4})]. \end{aligned}$$

This is a polynomial identity in x^ρ which yields to the following equations

$$\begin{cases} c = b^q - \delta k^{q^2+1} b^{q^5} \\ dk^{q-1} = a^q + b^{q^3} + \delta k^{q+q^2} b^{q^5} \\ -dk^{q^2-1} = k^{q^2-q} b^q + b^{q^3} \\ 0 = k^{-q-1} b^q + a^{q^3} + \delta b^{q^5} \\ d = k^{-q+1} b^{q^3} + \delta b^{q^5} \\ d = b^q - k^{-q^2+1} b^{q^3} + \delta a^{q^5} \end{cases}$$

which can be written as (3.3).

Now, suppose that there exist $\rho \in \mathrm{Aut}(\mathbb{F}_{q^6})$ and an invertible matrix $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$ such that for each $x \in \mathbb{F}_{q^6}$ there exists $z \in \mathbb{F}_{q^6}$ satisfying

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x^\rho \\ h^{\rho(q-1)}x^{\rho q} - h^{\rho(q^2-1)}x^{\rho q^2} + x^{\rho q^4} + x^{\rho q^5} \end{pmatrix} = \begin{pmatrix} z \\ \delta z^q + z^{q^3} + z^{q^5} \end{pmatrix}.$$

Equivalently, for each $x \in \mathbb{F}_{q^6}$ we have

$$\begin{aligned} cx^\rho + d(k^{q-1}x^{\rho q} - k^{q^2-1}x^{\rho q^2} + x^{\rho q^4} + x^{\rho q^5}) = \\ \delta[a^q x^{\rho q} + b^q(k^{q^2-q}x^{\rho q^2} + k^{-1-q}x^{\rho q^3} + x^{\rho q^5} + x^\rho)] \\ + a^{q^3}x^{\rho q^3} + b^{q^3}(k^{-q+1}x^{\rho q^4} - k^{-q^2+1}x^{\rho q^5} + x^{\rho q} + x^{\rho q^2}) \\ + a^{q^5}x^{\rho q^5} + b^{q^5}(-k^{1+q^2}x^\rho + k^{q^2+q}x^{\rho q} + x^{\rho q^3} + x^{\rho q^4}). \end{aligned}$$

This is a polynomial identity in x^ρ which yields to the following equations

$$\begin{cases} c = \delta b^q - k^{q^2+1}b^{q^5} \\ dk^{q-1} = \delta a^q + b^{q^3} + k^{q+q^2}b^{q^5} \\ -dk^{q^2-1} = \delta k^{q^2-q}b^q + b^{q^3} \\ 0 = \delta k^{-q-1}b^q + a^{q^3} + b^{q^5} \\ d = k^{-q+1}b^{q^3} + b^{q^5} \\ d = \delta b^q - k^{-q^2+1}b^{q^3} + a^{q^5} \end{cases}$$

which can be written as (3.4). □

We are now ready to prove that when $h \notin \mathbb{F}_{q^2}$, $\mathcal{L}_h$ is new.

Proposition 3.7. *If $h \notin \mathbb{F}_{q^2}$, then $\mathcal{L}_h$ is not* $\mathrm{P\Gamma L}$*-equivalent to L_δ^4 (with $\delta^2 + \delta = 1$).*

Proof. By Proposition 3.6 we have to show that there are no a, b, c and d in $\mathbb{F}_{q^6}$ such that $ad - bc \neq 0$ and (3.3) or (3.4) are satisfied. Note that $b = 0$ in (3.3) and (3.4) yields $a = c = d = 0$, a contradiction. So, suppose $b \neq 0$. Since $h \notin \mathbb{F}_{q^2}$ then $k \notin \mathbb{F}_{q^2}$. We start by proving that the last three equations of (3.3), i.e.

$$\begin{cases} \mathrm{Eq}_1\colon b^{q^3} + (k^{q-1} + \delta k^{q+q^2})b^{q^5} = 0 \\ \mathrm{Eq}_2\colon k^{q^2-q}b^q + (1 + k^{q^2-q})b^{q^3} + \delta k^{q^2-1}b^{q^5} = 0 \\ \mathrm{Eq}_3\colon -\delta b^q + (k^{-q+1} + \delta^2 k^{1-q^2})b^{q^3} + \delta b^{q^5} = 0, \end{cases}$$

yield a contradiction. As in the above section, we will consider the q-th powers of Eq_1, Eq_2 and Eq_3 replacing b^{q^i}, k^{q^j}, and δ^{q^ℓ} (respectively) by X_i, Y_j, and Z_ℓ with $i, j \in \{0,1,2,3,4,5\}$ and $\ell \in \{0,1\}$. Consider the set S of polynomials in the variables X_i, Y_j, and Z_ℓ

$$S := \{\mathrm{Eq}_1^{q^\alpha}, \mathrm{Eq}_2^{q^\beta}, \mathrm{Eq}_3^{q^\gamma} : \alpha, \beta, \gamma \in \{0,1,2,3,4,5\}\}.$$

By eliminating from S the variables X_5, X_4, X_3, and X_2 using Eq_1, Eq_1^q, $\mathrm{Eq}_1^{q^4}$, and $\mathrm{Eq}_1^{q^3}$ respectively we obtain

$$X_0Y_1(Z_1Y_0^2Y_2 - Z_1Y_0Y_2^2 - Z_1Y_0 + Z_1Y_2 - Z_0^2Z_2 - Z_2) = 0.$$

By the conditions on b and k, $X_0Y_1 \neq 0$ and therefore

$$P := Z_1Y_0^2Y_2 - Z_1Y_0Y_2^2 - Z_1Y_0 + Z_1Y_2 - Z_0^2Z_2 - Z_2 = 0.$$

We eliminate Z_1 in S using P, obtaining, w.r.t. b, k, and δ,

$$bk^{q^2+1}(k - k^q)(k + k^q)(k^{q^2+1} - 1)(k^{q^2+1} + 1) = 0,$$

a contradiction to $k \notin \mathbb{F}_{q^2}$.

Consider now the last three equations of (3.4), i.e.

$$\begin{cases} \mathrm{Eq}_1 \colon \delta b^{q^3} + (k^{q-1} - \delta k^{q^2+q})b^{q^5} = 0 \\ \mathrm{Eq}_2 \colon \delta k^{q^2-q}b^q + (k^{q^2-q} + 1)b^{q^3} + k^{q^2-1}b^{q^5} = 0 \\ \mathrm{Eq}_3 \colon \delta^2 b^q + (k^{-q+1} + \delta^2 k^{-q^2+1})b^{q^3} + b^{q^5} = 0. \end{cases}$$

As before, we will consider the q-th powers of Eq_1, Eq_2, and Eq_3 replacing b^{q^i}, k^{q^j}, and δ^{q^ℓ} (respectively) by X_i, Y_j, and Z_ℓ with $i, j \in \{0, 1, 2, 3, 4, 5\}$ and $\ell \in \{0, 1\}$. Consider the set S of polynomials in the variables X_i, Y_j and Z_ℓ

$$S := \{\mathrm{Eq}_1^{q^\alpha}, \mathrm{Eq}_2^{q^\beta}, \mathrm{Eq}_3^{q^\gamma} : \alpha, \beta, \gamma \in \{0, 1, 2, 3, 4, 5\}\}.$$

We eliminate in S the variables X_5, X_4, X_3, and X_2 using Eq_1, Eq_1^q, $\mathrm{Eq}_1^{q^4}$, and $\mathrm{Eq}_1^{q^3}$ respectively, and we get

$$Y_0X_0(Z_1Y_0^2Y_2^2 + 2Z_1Y_0Y_1^2Y_2 + 2Z_1Y_0Y_2 + Z_1Y_1^2 - Y_0^2Y_2^2 - Y_0Y_1^2Y_2 - Y_0Y_2 - Y_1^2) = 0.$$

Since $b \neq 0$ and $k \notin \mathbb{F}_{q^2}$, $X_0Y_0 \neq 0$ and therefore

$$P := Z_1Y_0^2Y_2^2 + 2Z_1Y_0Y_1^2Y_2 + 2Z_1Y_0Y_2 + Z_1Y_1^2 - Y_0^2Y_2^2 - Y_0Y_1^2Y_2 - Y_0Y_2 - Y_1^2 = 0.$$

Once again we consider the resultants of the polynomials in S and P w.r.t. Z_1 and we obtain

$$bk^{q^2+2q}(k - k^q)(k + k^q)(k^{q^2+1} - 1)(k^{q^2+1} + 1) = 0,$$

a contradiction to $k \notin \mathbb{F}_{q^2}$. □

As a consequence of the above considerations and Propositions 3.5 and 3.7, we have the following.

Corollary 3.8. *If $h \notin \mathbb{F}_{q^2}$, then $\mathcal{L}_h$ is not* $\mathrm{P\Gamma L}$*-equivalent to any known scattered linear set in* $\mathrm{PG}(1, q^6)$.

3.3 $\mathcal{L}_h$ may be defined by a trinomial

Suppose that $h \in \mathbb{F}_{q^2}$, then the condition on h becomes $h^{q+1} = -1$. For such h we can prove that the linear set $\mathcal{L}_h$ can be defined by the q-polynomial $(h^{-1} - 1)x^q + x^{q^3} + (h - 1)x^{q^5}$.

Proposition 3.9. *If $h \in \mathbb{F}_{q^2}$, then the linear set $\mathcal{L}_h$ is* $\mathrm{P\Gamma L}$*-equivalent to*

$$L_{\mathrm{tri}} := \{\langle (x, (h^{-1} - 1)x^q + x^{q^3} + (h - 1)x^{q^5})\rangle_{\mathbb{F}_{q^6}} : x \in \mathbb{F}_{q^6}^*\}.$$

Proof. Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}(2,q^6)$ with $a = -h + h^{-1}, b = 1, c = h^{-1} - 1 - h^3 + h^2$ and $d = h - h^2 - 1$. Straightforward computations show that the subspaces $\mathcal{U}_h$ and $U_{(h^{-1}-1)x^q + x^{q^3} + (h-1)x^{q^5}}$ are $\Gamma\mathrm{L}(2,q^6)$-equivalent under the action of the matrix A. Hence, the linear sets $\mathcal{L}_h$ and L_{tri} are PΓL-equivalent. □

The fact that $\mathcal{L}_h$ can also be defined by a trinomial will help us to completely close the equivalence issue for $\mathcal{L}_h$ when $h \in \mathbb{F}_{q^2}$. Indeed, we can prove the following:

Proposition 3.10. *If $h \in \mathbb{F}_{q^2}$, then the linear set $\mathcal{L}_h$ is* PΓL*-equivalent to some L^4_δ ($\delta^2 + \delta = 1$) if and only if $h \in \mathbb{F}_q$ and q is a power of* 5*.*

Proof. Recall that by [27, Proposition 5.5] if $h \in \mathbb{F}_q$ and q is a power of 5, then $\mathcal{L}_h$ is PΓL-equivalent to some L^4_δ. As in the proof of Proposition 3.6, by Lemma 3.3 we have to check whether $U_{(h^{-1}-1)x^q + x^{q^3} + (h-1)x^{q^5}}$ is ΓL-equivalent either to U^4_δ or to $(U^4_\delta)^\perp$. Suppose that there exist $\rho \in \mathrm{Aut}(\mathbb{F}_{q^6})$ and an invertible matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ such that for each $x \in \mathbb{F}_{q^6}$ there exists $z \in \mathbb{F}_{q^6}$ satisfying

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x^\rho \\ (h^{-\rho}-1)x^{\rho q} + x^{\rho q^3} + (h^\rho - 1)x^{\rho q^5} \end{pmatrix} = \begin{pmatrix} z \\ z^q + z^{q^3} + \delta z^{q^5} \end{pmatrix}.$$

Let $k = h^\rho$, for which $k^{q+1} = -1$. As in Proposition 3.5, we obtain a polynomial identity, whence

$$\begin{cases} c = b^q(k^q - 1) + b^{q^3} + \delta b^{q^5}(k^{-q} - 1) \\ d(k^{-1} - 1) = a^q \\ 0 = b^q(k^{-q} - 1) + b^{q^3}(k^q - 1) + b^{q^5}\delta \\ d = a^{q^3} \\ 0 = b^q + b^{q^3}(k^{-q} - 1) + b^{q^5}(k^q - 1)\delta \\ d(k-1) = \delta a^{q^5}. \end{cases} \tag{3.5}$$

By subtracting the fifth equation from the third equation raised to q^2, we get

$$b^q = b^{q^5}(k^q - 1),$$

i.e. either $b = 0$ or $k^q - 1 = (b^q)^{q^4 - 1}$, whence we get either $b = 0$ or $\mathrm{N}_{q^6/q^2}(k^q - 1) = 1$.

If $b \neq 0$, since $k - 1 \in \mathbb{F}_{q^2}$ and $\mathrm{N}_{q^6/q^2}(k-1) = (k-1)^3 = 1$, then

$$k^3 - 3k^2 + 3k - 2 = 0$$

and, since $\mathrm{N}_{q^6/q^2}(k^q - 1) = 1$ and $k^q = -1/k$,

$$2k^3 + 3k^2 + 3k + 1 = 0,$$

from which we get

$$9k^2 - 3k + 5 = 0. \tag{3.6}$$

- If $k \notin \mathbb{F}_q$ then k and k^q are the solutions of (3.6) and

$$-1 = k^{q+1} = \frac{5}{9},$$

  which holds if and only if q is a power of 7. By (3.6) it follows that $k \in \mathbb{F}_q$, a contradiction.

- If $k \in \mathbb{F}_q$, then $k^2 = -1$ and by (3.6) we have $k = -4/3$, which is possible if and only if q is a power of 5.

Hence, if either $k \notin \mathbb{F}_q$ or $k \in \mathbb{F}_q$ with q not a power of 5, we have that $b = 0$ and hence $c = 0$, $a \neq 0$ and $d \neq 0$.

By combining the second and the fourth equation of (3.5), we get $\mathrm{N}_{q^6/q^2}(k^{-1} - 1) = 1$ and, since $k^q = -1/k$, $\mathrm{N}_{q^6/q^2}(k^q + 1) = -1$. Arguing as above, we get a contradiction whenever $k \notin \mathbb{F}_q$ or $k \in \mathbb{F}_q$ with q not a power of 5.

Now, suppose that there exist $\rho \in \mathrm{Aut}(\mathbb{F}_{q^6})$ and an invertible matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ such that for each $x \in \mathbb{F}_{q^6}$ there exists $z \in \mathbb{F}_{q^6}$ satisfying

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x^\rho \\ (h^{-\rho} - 1)x^{\rho q} + x^{\rho q^3} + (h^\rho - 1)x^{\rho q^5} \end{pmatrix} = \begin{pmatrix} z \\ \delta z^q + z^{q^3} + z^{q^5} \end{pmatrix}.$$

Let $k = h^\rho$. As before, we get the following equations

$$\begin{cases} c = \delta b^q(k^q - 1) + b^{q^3} + b^{q^5}(k^{-q} - 1) \\ d(k^{-1} - 1) = \delta a^q \\ 0 = \delta b^q(k^{-q} - 1) + b^{q^3}(k^q - 1) + b^{q^5} \\ d = a^{q^3} \\ 0 = \delta b^q + b^{q^3}(k^{-q} - 1) + b^{q^5}(k^q - 1) \\ d(k - 1) = a^{q^5}. \end{cases} \tag{3.7}$$

By subtracting the fifth equation from the third raised to q^2 of the above system we get

$$b^q = b^{q^3}(k^{-q} - 1).$$

If $b \neq 0$, then $\mathrm{N}_{q^6/q^2}(k^{-q} - 1) = 1$. Hence, arguing as above, we get that $b = 0$ and hence $c = 0$, $a, d \neq 0$. By combining the fourth equation with the second and the fifth equation of (3.7) we get $\mathrm{N}_{q^6/q^2}(k - 1) = 1$, which yields again to a contradiction when $k \notin \mathbb{F}_q$ or $k \in \mathbb{F}_q$ with q not a power of 5. □

So, as a consequence of Corollary 3.8 and of the above proposition, we have the following result.

Corollary 3.11. *Apart from the case $h \in \mathbb{F}_q$ and q a power of 5, the linear set $\mathcal{L}_h$ is not $\mathrm{P\Gamma L}$-equivalent to any known scattered linear set in $\mathrm{PG}(1, q^6)$.*

By Proposition 3.9, when $h \in \mathbb{F}_{q^2}$, $\mathcal{L}_h$ is a linear set of the family presented in [23, Section 7]. Also, we get an extension of [21, Table 1], where it is shown examples of scattered linear sets which could generalize the family presented in [10]. We do not know whether the linear set $\mathcal{L}_h$, for each $h \in \mathbb{F}_{q^6} \setminus \mathbb{F}_{q^2}$ with $h^{q^3+1} = -1$, may be defined by a trinomial or not.

4 New MRD-codes

Delsarte in [13] (see also [14]) introduced in 1978 rank metric codes as follows. A *rank metric code* (or *RM*-code for short) $\mathcal{C}$ is a subset of the set of $m \times n$ matrices $\mathbb{F}_q^{m\times n}$ over $\mathbb{F}_q$ equipped with the distance function

$$d(A, B) = \mathrm{rk}\,(A - B)$$

for $A, B \in \mathbb{F}_q^{m\times n}$. The *minimum distance* of $\mathcal{C}$ is

$$d = \min\{d(A,B) : A, B \in \mathcal{C},\ A \neq B\}.$$

We will say that a rank metric code of $\mathbb{F}_q^{m\times n}$ with minimum distance d has parameters $(m,n,q;d)$. When $\mathcal{C}$ is an $\mathbb{F}_q$-subspace of $\mathbb{F}_q^{m\times n}$, we say that $\mathcal{C}$ is $\mathbb{F}_q$-linear. In the same paper, Delsarte also showed that the parameters of these codes fulfill a Singleton-like bound, i.e.

$$|\mathcal{C}| \leq q^{\max\{m,n\}(\min\{m,n\}-d+1)}.$$

When the equality holds, we call $\mathcal{C}$ a *maximum rank distance* (*MRD* for short) code. We will consider only the case $m = n$ and we will use the following equivalence definition for codes of $\mathbb{F}_q^{m\times m}$. Two $\mathbb{F}_q$-linear RM-codes $\mathcal{C}$ and $\mathcal{C}'$ are *equivalent* if and only if there exist two invertible matrices $A, B \in \mathbb{F}_q^{m\times m}$ and a field automorphism σ such that $\{AC^\sigma B : C \in \mathcal{C}\} = \mathcal{C}'$, or $\{AC^{T\sigma} B : C \in \mathcal{C}\} = \mathcal{C}'$, where T denotes transposition. Also, the *left* and *right idealisers* of $\mathcal{C}$ are $L(\mathcal{C}) = \{A \in \mathrm{GL}(m,q) : A\mathcal{C} \subseteq \mathcal{C}\}$ and $R(\mathcal{C}) = \{B \in \mathrm{GL}(m,q) : \mathcal{C}B \subseteq \mathcal{C}\}$ [17, 19]. They are important invariants for linear rank metric codes, see also [15] for further invariants.

In [24, Section 5] Sheekey showed that scattered $\mathbb{F}_q$-linear sets of $\mathrm{PG}(1,q^n)$ of rank n yield $\mathbb{F}_q$-linear MRD-codes with parameters $(n,n,q;n-1)$ with left idealiser isomorphic to $\mathbb{F}_{q^n}$; see [7, 8, 25] for further details on such kind of connections. We briefly recall here the construction from [24]. Let $U_f = \{(x, f(x)) : x \in \mathbb{F}_{q^n}\}$ for some scattered q-polynomial $f(x)$. After fixing an $\mathbb{F}_q$-basis for $\mathbb{F}_{q^n}$ we can define an isomorphism between the rings $\mathrm{End}(\mathbb{F}_{q^n}, \mathbb{F}_q)$ and $\mathbb{F}_q^{n\times n}$. In this way the set

$$\mathcal{C}_f := \{x \mapsto af(x) + bx : a, b \in \mathbb{F}_{q^n}\}$$

corresponds to a set of $n \times n$ matrices over $\mathbb{F}_q$ forming an $\mathbb{F}_q$-linear MRD-code with parameters $(n,n,q;n-1)$. Also, since $\mathcal{C}_f$ is an $\mathbb{F}_{q^n}$-subspace of $\mathrm{End}(\mathbb{F}_{q^n}, \mathbb{F}_q)$ its left idealiser $L(\mathcal{C}_f)$ is isomorphic to $\mathbb{F}_{q^n}$. For further details see [6, Section 6].

Let $\mathcal{C}_f$ and $\mathcal{C}_h$ be two MRD-codes arising from maximum scattered subspaces U_f and U_h of $\mathbb{F}_{q^n} \times \mathbb{F}_{q^n}$. In [24, Theorem 8] the author showed that there exist invertible matrices A, B and $\sigma \in \mathrm{Aut}(\mathbb{F}_q)$ such that $A\mathcal{C}_f^\sigma B = \mathcal{C}_h$ if and only if U_f and U_h are $\Gamma\mathrm{L}(2,q^n)$-equivalent

Therefore, we have the following.

Theorem 4.1. *The $\mathbb{F}_q$-linear MRD-code $\mathcal{C}_{f_h}$ arising from the $\mathbb{F}_q$-subspace $\mathcal{U}_h$ has parameters $(6,6,q;5)$ and left idealiser isomorphic to $\mathbb{F}_{q^6}$, and is not equivalent to any previously known MRD-code, apart from the case $h \in \mathbb{F}_q$ and q a power of 5.*

Proof. From [6, Section 6], the previously known $\mathbb{F}_q$-linear MRD-codes with parameters $(6,6,q;5)$ and with left idealiser isomorphic to $\mathbb{F}_{q^6}$ arise, up to equivalence, from one of the maximum scattered subspaces of $\mathbb{F}_{q^6} \times \mathbb{F}_{q^6}$ described in Section 3. From Corollaries 3.8 and 3.11 the result then follows. □

ORCID iDs

Daniele Bartoli https://orcid.org/0000-0002-5767-1679
Corrado Zanella https://orcid.org/0000-0002-5031-1961
Ferdinando Zullo https://orcid.org/0000-0002-5087-2363

References

[1] D. Bartoli and M. Montanucci, Towards the full classification of exceptional scattered polynomials, *J. Comb. Theory Ser. A*, in press, `arXiv:1905.11390 [math.CO]`.

[2] D. Bartoli and Y. Zhou, Exceptional scattered polynomials, *J. Algebra* **509** (2018), 507–534, doi:10.1016/j.jalgebra.2018.03.010.

[3] A. Blokhuis and M. Lavrauw, Scattered spaces with respect to a spread in $\mathrm{PG}(n, q)$, *Geom. Dedicata* **81** (2000), 231–243, doi:10.1023/a:1005283806897.

[4] B. Csajbók, Scalar q-subresultants and Dickson matrices, *J. Algebra* **547** (2020), 116–128, doi:10.1016/j.jalgebra.2019.10.056.

[5] B. Csajbók, G. Marino and O. Polverino, Classes and equivalence of linear sets in $\mathrm{PG}(1, q^n)$, *J. Comb. Theory Ser. A* **157** (2018), 402–426, doi:10.1016/j.jcta.2018.03.007.

[6] B. Csajbók, G. Marino, O. Polverino and C. Zanella, A new family of MRD-codes, *Linear Algebra Appl.* **548** (2018), 203–220, doi:10.1016/j.laa.2018.02.027.

[7] B. Csajbók, G. Marino, O. Polverino and F. Zullo, Generalising the scattered property of subspaces, *Combinatorica*, in press, `arXiv:1906.10590 [math.CO]`.

[8] B. Csajbók, G. Marino, O. Polverino and F. Zullo, Maximum scattered linear sets and MRD-codes, *J. Algebraic Combin.* **46** (2017), 517–531, doi:10.1007/s10801-017-0762-6.

[9] B. Csajbók, G. Marino, O. Polverino and F. Zullo, A characterization of linearized polynomials with maximum kernel, *Finite Fields Appl.* **56** (2019), 109–130, doi:10.1016/j.ffa.2018.11.009.

[10] B. Csajbók, G. Marino and F. Zullo, New maximum scattered linear sets of the projective line, *Finite Fields Appl.* **54** (2018), 133–150, doi:10.1016/j.ffa.2018.08.001.

[11] B. Csajbók and C. Zanella, On scattered linear sets of pseudoregulus type in $\mathrm{PG}(1, q^t)$, *Finite Fields Appl.* **41** (2016), 34–54, doi:10.1016/j.ffa.2016.04.006.

[12] B. Csajbók and C. Zanella, On the equivalence of linear sets, *Des. Codes Cryptogr.* **81** (2016), 269–281, doi:10.1007/s10623-015-0141-z.

[13] P. Delsarte, Bilinear forms over a finite field, with applications to coding theory, *J. Comb. Theory Ser. A* **25** (1978), 226–241, doi:10.1016/0097-3165(78)90015-8.

[14] E. M. Gabidulin, Theory of codes with maximum rank distance, *Problemy Peredachi Informatsii* **21** (1985), 3–16, `http://mi.mathnet.ru/eng/ppi967`.

[15] L. Giuzzi and F. Zullo, Identifiers for MRD-codes, *Linear Algebra Appl.* **575** (2019), 66–86, doi:10.1016/j.laa.2019.03.030.

[16] M. Lavrauw, G. Marino, O. Polverino and R. Trombetti, Solution to an isotopism question concerning rank 2 semifields, *J. Combin. Des.* **23** (2015), 60–77, doi:10.1002/jcd.21382.

[17] D. Liebhold and G. Nebe, Automorphism groups of Gabidulin-like codes, *Arch. Math. (Basel)* **107** (2016), 355–366, doi:10.1007/s00013-016-0949-4.

[18] G. Lunardon and O. Polverino, Blocking sets and derivable partial spreads, *J. Algebraic Combin.* **14** (2001), 49–56, doi:10.1023/a:1011265919847.

[19] G. Lunardon, R. Trombetti and Y. Zhou, On kernels and nuclei of rank metric codes, *J. Algebraic Combin.* **46** (2017), 313–340, doi:10.1007/s10801-017-0755-5.

[20] G. Lunardon, R. Trombetti and Y. Zhou, Generalized twisted Gabidulin codes, *J. Comb. Theory Ser. A* **159** (2018), 79–106, doi:10.1016/j.jcta.2018.05.004.

[21] G. Marino, M. Montanucci and F. Zullo, MRD-codes arising from the trinomial $x^q + x^{q^3} + cx^{q^5} \in \mathbb{F}_{q^6}[x]$, *Linear Algebra Appl.* **591** (2020), 99–114, doi:10.1016/j.laa.2020.01.004.

[22] G. McGuire and J. Sheekey, A characterization of the number of roots of linearized and projective polynomials in the field of coefficients, *Finite Fields Appl.* **57** (2019), 68–91, doi: 10.1016/j.ffa.2019.02.003.

[23] O. Polverino and F. Zullo, On the number of roots of some linearized polynomials, *Linear Algebra Appl.* **601** (2020), 189–218, doi:10.1016/j.laa.2020.05.009.

[24] J. Sheekey, A new family of linear maximum rank distance codes, *Adv. Math. Commun.* **10** (2016), 475–488, doi:10.3934/amc.2016019.

[25] J. Sheekey and G. Van de Voorde, Rank-metric codes, linear sets, and their duality, *Des. Codes Cryptogr.* **88** (2020), 655–675, doi:10.1007/s10623-019-00703-z.

[26] C. Zanella, A condition for scattered linearized polynomials involving Dickson matrices, *J. Geom.* **110** (2019), Paper no. 50 (9 pages), doi:10.1007/s00022-019-0505-z.

[27] C. Zanella and F. Zullo, Vertex properties of maximum scattered linear sets of $\mathrm{PG}(1, q^n)$, *Discrete Math.* **343** (2020), 111800 (14 pages), doi:10.1016/j.disc.2019.111800.

ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 147–154
https://doi.org/10.26493/1855-3974.2219.554
(Also available at http://amc-journal.eu)

Noncommutative frames revisited

Karin Cvetko-Vah

Department of Mathematics, Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 21, SI-1000 Ljubljana, Slovenia

Jens Hemelaer *

Department of Mathematics, University of Antwerp, Middelheimlaan 1, B-2020 Antwerp, Belgium

Jonathan Leech

Department of Mathematics, Westmont College, Santa Barbara, CA 93108, USA

Received 10 January 2020, accepted 22 June 2020, published online 12 November 2020

Abstract

In this note we revisit noncommutative frames. Special attention is devoted to the study of join completeness and related properties in skew lattices.

Keywords: Noncommutative frame, skew lattice, completeness, lattice section.

Math. Subj. Class. (2020): 06D75, 06D22

1 Introduction

In [5], the first author introduced noncommutative frames, motivated by a noncommutative topology constructed by Le Bruyn [7] on the points of the Connes-Consani Arithmetic Site [2, 3]. The definition of noncommutative frame fits in the general theory of skew lattices, a theory that goes back to Pascual Jordan [6] and is an active research topic starting with a series of papers of the third author [8, 9, 10]. For an overview of the primary results on skew lattices, we refer the reader to [12] or the earlier systematic survey [11].

Recall that a frame is a complete lattice which satisfies the infinite distributive laws. Noncommutative frames are noncommutative generalizations of frames, the precise definition is given in Section 1. Loosely speaking, a noncommutative frame is a frame of certain

*Corresponding author. Jens Hemelaer was supported in part by a PhD fellowship of the Research Foundation (Flanders) and in part by the University of Antwerp (BOF).

E-mail addresses: karin.cvetko@fmf.uni-lj.si (Karin Cvetko-Vah), jens.hemelaer@uantwerpen.be (Jens Hemelaer), leech@westmont.edu (Jonathan Leech)

congruence classes, $\mathcal{D}$-classes. A noncommutative frame containing both the top and the bottom elements would necessarily be commutative. There are thus two natural ways of generalizing frames to the noncommutative setting:

1. We keep the bottom element, but replace the top element with a top $\mathcal{D}$-class. This approach is carried out in the present paper.

2. We keep the top element, but replace the bottom element with a bottom $\mathcal{D}$-class. This approach was carried out in [4].

Note that the two approaches are essentially different as they do not dualize one another.

The notion of completeness for noncommutative lattices is much more complex than for lattices. For example, join completeness and meet completeness turn out to be non-equivalent properties. The main purpose of this note is to study aspects of [join, meet] completeness for certain types of skew lattices as well as certain related properties, which we define and explore in Section 3. In Section 4 we study join completeness in terms of $\mathcal{D}$-classes. In Section 5 we state and prove a correction of Theorem 4.4 of [5], where the assumption of join completeness was erroneously omitted. Theorem 5.1 states that if S is a join complete, strongly distributive skew lattice with 0, then S is a noncommutative frame if and only if its commutative shadow $S/\mathcal{D}$ is a frame. Examples 3.2 and 3.4 show that the assumption of join completeness is indeed necessary.

2 Preliminaries

A *skew lattice* is a set A endowed with a pair of idempotent, associative operations $\wedge$ and $\vee$ which satisfy the absorption laws:

$$x \wedge (x \vee y) = x = x \vee (x \wedge y) \quad \text{and} \quad (x \wedge y) \vee y = y = (x \vee y) \wedge y.$$

The terms meet and join are still used for $\wedge$ and $\vee$, but without assuming commutativity. Given skew lattices A and B, a *homomorphism* of skew lattices is a map $f\colon A \to B$ that preserves finite meets and joins, i.e. it satisfies the following pair of axioms:

- $f(a \wedge b) = f(a) \wedge f(b)$, for all $a, b \in A$;

- $f(a \vee b) = f(a) \vee f(b)$, for all $a, b \in A$.

A *natural partial order* is defined on any skew lattice A by: $a \leq b$ iff $a \wedge b = b \wedge a = a$, or equivalently, $a \vee b = b = b \vee a$. The *Green's equivalence relation* $\mathcal{D}$ is defined on A by: $a \,\mathcal{D}\, b$ iff $a \wedge b \wedge a = a$ and $b \wedge a \wedge b = b$, or equivalently, $a \vee b \vee a = a$ and $b \vee a \vee b = b$. By Leech's First Decomposition Theorem [8], relation $\mathcal{D}$ is a congruence on a skew lattice A and $A/\mathcal{D}$ is a maximal lattice image of A, also referred to as the *commutative shadow* of A.

Skew lattices are always *regular* in that they satisfy the identities:

$$a \wedge x \wedge a \wedge y \wedge a = a \wedge x \wedge y \wedge a \quad \text{and} \quad a \vee x \vee a \vee y \vee a = a \vee x \vee y \vee a.$$

The following result is an easy consequence of regularity.

Lemma 2.1. *Let a, b, u, v be elements of a skew lattice A such that $\mathcal{D}_u \leq \mathcal{D}_a$, $\mathcal{D}_u \leq \mathcal{D}_b$, $\mathcal{D}_a \leq \mathcal{D}_v$ and $\mathcal{D}_b \leq \mathcal{D}_v$. Then:*

1. $a \wedge v \wedge b = a \wedge b$,
2. $a \vee u \vee b = a \vee b$.

A skew lattice is *strongly distributive* if it satisfies the identities:

$$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z) \quad \text{and} \quad x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

By a result of Leech [10], a skew lattice is strongly distributive if and only if it is symmetric, distributive and normal, where a skew lattice A is called:

- *symmetric* if for any $x, y \in A$, $x \vee y = y \vee x$ iff $x \wedge y = y \wedge x$;
- *distributive* if it satisfies the identities:

$$x \wedge (y \vee z) \wedge x = (x \wedge y \wedge x) \vee (x \wedge z \wedge x)$$
$$x \vee (y \wedge z) \vee x = (x \vee y \vee x) \wedge (x \vee z \vee x);$$

- *normal* if it satisfies the identity $x \wedge y \wedge z \wedge x = x \wedge z \wedge y \wedge x$.

Further, it is shown in [10] that a skew lattice A is normal if and only if given any $a \in A$ the set

$$a{\downarrow} = \{u \in A \mid u \leq a\}$$

is a lattice. For this reason, normal skew lattices are sometimes called *local lattices*. Given any comparable $\mathcal{D}$-classes $D < C$ in a normal skew lattice A and any $c \in C$ there exist a unique $d \in D$ such that $d < c$ with respect to the natural partial order.

Finally, a *skew lattice with* 0 is a skew lattice with a distinguished element 0 satisfying $x \vee 0 = x = 0 \vee x$, or equivalently, $x \wedge 0 = 0 = 0 \wedge x$.

Example 2.2. Let A, B be non-empty sets and denote by $\mathcal{P}(A, B)$ the set of all partial functions from A to B. We define the following operations on $\mathcal{P}(A, B)$:

$$f \wedge g = f|_{\mathrm{dom}(f) \cap \mathrm{dom}(g)}$$
$$f \vee g = g \cup f|_{\mathrm{dom}(f) \setminus \mathrm{dom}(g)}.$$

Leech [10] proved that $(\mathcal{P}(A, B); \wedge, \vee)$ is a strongly distributive skew lattice with 0. Moreover, given $f, g \in (\mathcal{P}(A, B); \wedge, \vee)$ the following hold:

- $f \; \mathcal{D} \; g$ iff $\mathrm{dom}(f) = \mathrm{dom}(g)$;
- $f \leq g$ iff $f = g|_{\mathrm{dom}(f) \cap \mathrm{dom}(g)}$;
- $\mathcal{P}(A, B)/\mathcal{D} \cong \mathcal{P}(A)$, the Boolean algebra of subsets of A;
- $\mathcal{P}(A, B)$ is left-handed in that $x \wedge y \wedge x = x \wedge y$ and dually, $x \vee y \vee x = y \vee x$ hold.

A *commuting subset* of a skew lattice A is a nonempty subset $\{x_i \mid i \in I\} \subseteq A$ such that $x_i \wedge x_j = x_j \wedge x_i$ and $x_i \vee x_j = x_j \vee x_i$ hold for all $i, j \in I$. The following result is a direct consequence of the definitions.

Lemma 2.3. *Let A and B be skew lattices, $f\colon A \to B$ be a homomorphism of skew lattices, and $\{x_i \mid i \in I\} \subseteq A$ be a commuting subset of A. Then $\{f(x_i) \mid i \in I\}$ is a commuting subset of B.*

A skew lattice is said to be *join [meet] complete* if all commuting subsets have suprema [infima] with respect to the natural partial ordering. By a result of Leech [9], the choice axiom implies that any join complete symmetric skew lattice has a top $\mathcal{D}$-class. If it occurs, we denote the top $\mathcal{D}$-class of a skew lattice A by T (or T_A). Dually, if A is a meet complete symmetric skew lattice, then it always has a bottom $\mathcal{D}$-class, denoted by B (or B_A).

A *frame* is a lattice that has all joins (finite and infinite), and satisfies the infinite distributive law:

$$x \wedge \bigvee_i y_i = \bigvee_i (x \wedge y_i).$$

A *noncommutative frame* is a strongly distributive, join complete skew lattice A with 0 that satisfies the infinite distributive laws:

$$(\bigvee_i x_i) \wedge y = \bigvee_i (x_i \wedge y) \qquad \text{and} \qquad x \wedge (\bigvee_i y_i) = \bigvee_i (x \wedge y_i) \tag{2.1}$$

for all $x, y \in A$ and all commuting subsets $\{x_i \mid i \in I\}, \{y_i \mid i \in I\} \subseteq A$.

By a result of Bignall and Leech [1], any join complete, normal skew lattice A with 0 (for instance, any noncommutative frame) satisfies the following:

- A is meet complete, with the meet of a commuting subset C denoted by $\bigwedge C$;
- any nonempty subset $C \subseteq A$ has an infimum with respect to the natural partial order, to be denoted by $\bigcap C$ (or by $x \cap y$ in the case $C = \{x, y\}$);
- if C is a nonempty commuting subset of A, then $\bigwedge C = \bigcap C$.

We call the $\bigcap C$ the *intersection* of C.

A *lattice section* L of a skew lattice S is a subalgebra that is a lattice (i.e. both $\wedge$ and $\vee$ are commutative on L) and that intersects each $\mathcal{D}$-class in exactly one element. When it exists, a lattice section is a maximal commuting subset and it is isomorphic to the maximal lattice image, as shown by Leech in [8]. If a normal skew lattice S has a top $\mathcal{D}$-class T then given $t \in T$, $t{\downarrow} = \{x \in S \mid x \leq t\}$ is a lattice section of S; moreover, all lattice sections are of the form $t{\downarrow}$ for some $t \in T$. Further, it is shown in [8] that any symmetric skew lattice S such that $S/\mathcal{D}$ is countable has a lattice section.

We say that a commuting subset C in a symmetric skew lattice S *extends to a lattice section* if there exists a lattice section L of C such that $C \subseteq L$.

3 Comparison of completeness properties

Let S be a normal, symmetric skew lattice. We will consider the following four properties that S might have:

(JC) S is join complete;

(BA) S is *bounded from above*, i.e. for every commuting subset C there is an element $s \in S$ such that $c \leq s$ for all $c \in C$;

(EX) every commuting subset extends to a lattice section;

(LS) there exists a lattice section.

Note that the last two properties are trivially satisfied if S is commutative.

Proposition 3.1. *For normal, symmetric skew lattices, the following implications hold:*

$$\textbf{(JC)} \Rightarrow \textbf{(BA)} \Rightarrow \textbf{(EX)} \Rightarrow \textbf{(LS)}.$$

Proof. We only prove **(BA)** $\Rightarrow$ **(EX)**, the other two implications are trivial. Take a normal, symmetric skew lattice S, such that every commutative subset has a join. Let $C \subseteq S$ be a commuting subset. We have to prove that C extends to a lattice section. For every chain $C_0 \subseteq C_1 \subseteq \cdots$ of commuting subsets, the union $\bigcup_{i=0}^{\infty} C_i$ is again a commuting subset. So by Zorn's Lemma, C is contained in a maximal commuting subset C'. Take an element $s \in S$ such that $s \geq c$ for all $c \in C'$. Then $s{\downarrow}$ contains C' and it is a commuting subset because S is normal. By maximality, $C' = s{\downarrow}$. Again by maximality, s is a maximal element for the natural partial order on S. This also means that s is in the top $\mathcal{D}$-class (if $y \in S$ has a $\mathcal{D}$-class with $[y] \not\leq [s]$, then $s \vee y \vee s > s$, a contradiction). So C' is a lattice section. □

We claim that the converse implications do not hold in general. We will give a counterexample to all three of them. In each case, the counterexamples are strongly distributive skew lattices with 0.

Example 3.2 (**(BA)** $\not\Rightarrow$ **(JC)**)**.** Consider the set $S = \mathbb{N} \cup \{\infty_a, \infty_b\}$ and turn S into a skew lattice by setting

$$x \wedge y = \min(x, y) \qquad x \vee y = \max(x, y)$$

whenever x or y is in $\mathbb{N}$ (∞_a and ∞_b are both greater than every natural number), and

$$\infty_a \wedge \infty_b = \infty_a = \infty_b \vee \infty_a$$
$$\infty_b \wedge \infty_a = \infty_b = \infty_a \vee \infty_b.$$

Then S is a left-handed strongly distributive skew lattice with 0. The commuting subsets of S are precisely the subsets that do not contain both ∞_a and ∞_b. Clearly, S is bounded from above (as well as meet complete). However, the commuting subset $\mathbb{N} \subseteq S$ does not have a join.

Note that there are commutative examples as well, for example the real interval $[0, 1]$ with join and meet given by respectively maximum and minimum. The element 1 is an upper bound for every subset, but the lattice is not join complete. However, we preferred an example where the commutative shadow $S/\mathcal{D}$ is join complete.

Example 3.3 (**(EX)** $\not\Rightarrow$ **(BA)**)**.** Here we give a commutative example. Take $S = \mathbb{N}$ with the meet and join given by respectively the minimum and maximum of two elements. Then **(EX)** is satisfied, but **(BA)** does not hold.

If S satisfies **(EX)** and $S/\mathcal{D}$ is bounded from above, then for any commuting subset $C \subseteq S$ we can find a lattice section $L \supseteq C$ and an element $y \in L$ such that $[y] \geq [c]$ for all $c \in C$. It follows that $y \geq c$ for all $c \in C$, so S is bounded from above. So any example as the one above essentially reduces to a commutative example.

Example 3.4 (**(LS)** $\not\Rightarrow$ **(EX)**)**.** Consider the subalgebra S of $\mathcal{P}(\mathbb{N}, \mathbb{N})$ consisting of all partial functions with finite image sets in $\mathbb{N}$. Note that $S/\mathcal{D} = \mathcal{P}(\mathbb{N})$. The skew lattice S has lattice sections, for example the subalgebra of all functions in $\mathcal{P}(\mathbb{N}, \mathbb{N})$ whose image set is $\{1\}$. The set of 1-point functions $\{n \mapsto n \mid n \in \mathbb{N}\}$ is clearly a commuting subset, but it cannot be extended to an entire lattice section.

Even the weakest property **(LS)**, the existence a lattice section, does not always hold for strongly distributive skew lattices.

Example 3.5 (**(LS)** does not hold)**.** Let S be the subalgebra of $\mathcal{P}(\mathbb{R}, \mathbb{N})$ consisting of all partial functions f such that $f^{-1}(n)$ is finite for all $n \in \mathbb{N}$. In particular, if $f \in S$, then the domain of f is at most countable. Conversely, for any at most countable subset $U \subseteq \mathbb{R}$ we can construct an element $f \in S$ with domain U. Suppose now that $Q \subseteq S$ is a lattice section. Then there is an entire function $q\colon \mathbb{R} \to \mathbb{N}$ such that every $f \in Q$ can be written as a restriction $f = q|_U$ with $U \subseteq \mathbb{R}$ at most countable. Take $n \in \mathbb{N}$ such that $q^{-1}(n)$ is infinite, and take a countably infinite subset $V \subseteq q^{-1}(n)$. Then $q|_V \notin S$, by definition. But this shows that there is no element $f \in Q$ with domain V, which contradicts that Q is a lattice section.

By [8], any symmetric skew lattice S with $S/\mathcal{D}$ at most countable has a lattice section. This shows that in the above example it is necessary that the commutative shadow $S/\mathcal{D}$ is uncountable.

4 Join completeness in terms of $\mathcal{D}$-classes

Let S be a normal, symmetric skew lattice. Recall that for an element $a \in S$, we write its $\mathcal{D}$-class as $[a]$. For a $\mathcal{D}$-class $u \leq [a]$, the unique element b with $b \leq a$ and $[b] = u$ will be called the *restriction of* a *to* u. We will denote the restriction of a to u by $a|_u$. For $u, v \leq [a]$ two $\mathcal{D}$-classes, we calculate that

$$(a|_u)|_v = a|_v \qquad \text{if } v \leq u,$$

and in particular

$$a|_u \leq a|_v \quad \Leftrightarrow \quad u \leq v.$$

Proposition 4.1. *Let S be a normal, symmetric skew lattice and take a commuting subset $\{a_i \mid i \in I\} \subseteq S$. Then the following are equivalent:*

(1) the join $\bigvee_{i \in I} a_i$ exists;

(2) the join $\bigvee_{i \in I}[a_i]$ exists and there is a unique $a \in S$ with $[a] = \bigvee_{i \in I}[a_i]$ and $a_i \leq a$ for all $i \in I$.

In this case, $a = \bigvee_{i \in I} a_i$. In particular, $\left[\bigvee_{i \in I} a_i\right] = \bigvee_{i \in I}[a_i]$.

Proof. (1) $\Rightarrow$ (2): We claim that $\left[\bigvee_{i \in I} a_i\right]$ is the join of the $\mathcal{D}$-classes $[a_i]$. Because taking $\mathcal{D}$-classes preserves the natural partial order, $[a_i] \leq \left[\bigvee_{i \in I} a_i\right]$ for all $i \in I$. If $\left[\bigvee_{i \in I} a_i\right]$ is not the join of the $[a_i]$'s, then we can find a $\mathcal{D}$-class $u < \left[\bigvee_{i \in I} a_i\right]$ such that $[a_i] \leq u$ for all $i \in I$. But then

$$a_i \leq \left.\left(\bigvee_{i \in I} a_i\right)\right|_u < \bigvee_{i \in I} a_i$$

for all $i \in I$, a contradiction. So $\bigvee_{i\in I}[a_i]$ exists and is equal to $\left[\bigvee_{i\in I} a_i\right]$. For the remaining part of the statement, it is a straightforward calculation to show that $a = \bigvee_{i\in I} a_i$ is the unique element with the given properties.

(2) ⇒ (1): Write $u = \bigvee_{i\in I}[a_i]$. Let $b \in S$ be an element such that $a_i \leq b$ for all $i \in I$. Then $u \leq [b]$ and $a_i \leq b|_u$ for all $i \in I$. It follows that $a = b|_u$, in particular $a \leq b$. So a is the join of the a_i's. □

Corollary 4.2. *Let S be a normal, symmetric skew lattice. Suppose that S is bounded from above and that $S/\mathcal{D}$ is join complete. If every two elements $a, b \in S$ have an infimum $a \cap b$ for the natural partial order, then S is join complete.*

Proof. Let $\{a_i \mid i \in I\} \subseteq S$ be a commuting subset. Because S is bounded from above, we can take an element $s \in S$ such that $a_i \leq s$ for all $i \in I$. Set $u = \bigvee_{i\in I}[a_i]$. By Proposition 4.1 it is enough to show that there is a unique $a \in S$ with $[a] = u$ and $a_i \leq a$ for all $i \in I$. Existence follows by taking the restriction $s|_u$. To show uniqueness, take two elements a and a' with $[a] = [a'] = u$ and $a_i \leq a$, $a_i \leq a'$ for all $i \in I$. It follows that $[a \cap a'] = \bigvee_{i\in I}[a_i] = u$. But this shows that $a = a \cap a' = a'$. □

In Example 3.2, the two elements ∞_a and ∞_b do not have an infimum.

5 Noncommutative frames

The following is a correction of a result in [5], where the assumption of being join complete was erroneously omitted.

Theorem 5.1. *Let S be a join complete, strongly distributive skew lattice with 0. Then S is a noncommutative frame if and only if $S/\mathcal{D}$ is a frame.*

Proof. Suppose that $S/\mathcal{D}$ is a frame. We prove the infinite distributivity laws (2.1). Take $x \in S$ and let $\{y_i \mid i \in I\} \subseteq S$ be a commuting subset. It is enough to show that

$$x \wedge \bigvee_{i\in I} y_i = \bigvee_{i\in I} x \wedge y_i$$

(the proof for the other infinite distributivity law is analogous). Using that S is strongly distributive, it is easy to compute that $y \leq z$ implies $x \wedge y \leq x \wedge z$. In particular, $x \wedge y_i \leq x \wedge \bigvee_i y_i$ for all $i \in I$. This shows:

$$\bigvee_{i\in I} x \wedge y_i \leq x \wedge \bigvee_{i\in I} y_i. \tag{5.1}$$

Further, we can use Proposition 4.1 to compute

$$\left[\bigvee_{i\in I} x \wedge y_i\right] = \bigvee_{i\in I}[x] \wedge [y_i] = [x] \wedge \bigvee_{i\in I}[y_i] = \left[x \wedge \bigvee_{i\in I} y_i\right],$$

where for the middle equality we use that $S/\mathcal{D}$ is a frame. Since left- and right-hand side in (5.1) are in the same $\mathcal{D}$-class, the inequality must be an equality, so that S is seen to be a noncommutative frame. Conversely, suppose that S is a noncommutative frame. Then S has a maximal $\mathcal{D}$-class, T_S. Let t be in T_S. Then $t{\downarrow}$ is a copy of $S/\mathcal{D}$. □

The extra assumption that S is join complete is necessary: the strongly distributive skew lattices from Examples 3.2 and 3.4 have a frame as commutative shadow, but they are not noncommutative frames, since they are not join complete.

ORCID iDs

Karin Cvetko-Vah https://orcid.org/0000-0003-3664-0731
Jens Hemelaer https://orcid.org/0000-0001-7228-272X

References

[1] R. Bignall and J. Leech, Skew Boolean algebras and discriminator varieties, *Algebra Universalis* **33** (1995), 387–398, doi:10.1007/bf01190707.

[2] A. Connes and C. Consani, The arithmetic site, *C. R. Math. Acad. Sci. Paris* **352** (2014), 971–975, doi:10.1016/j.crma.2014.07.009.

[3] A. Connes and C. Consani, Geometry of the arithmetic site, *Adv. Math.* **291** (2016), 274–329, doi:10.1016/j.aim.2015.11.045.

[4] K. Cvetko-Vah, On skew Heyting algebras, *Ars Math. Contemp.* **12** (2017), 37–50, doi:10.26493/1855-3974.757.7ec.

[5] K. Cvetko-Vah, Noncommutative frames, *J. Algebra Appl.* **18** (2019), 1950011 (13 pages), doi:10.1142/s0219498819500117.

[6] P. Jordan, Über nichtkommutative Verbände, *Arch. Math. (Basel)* **2** (1949), 56–59, doi:10.1007/bf02036754.

[7] L. Le Bruyn, Covers of the arithmetic site, `arXiv:1602.01627` [math.RA].

[8] J. Leech, Skew lattices in rings, *Algebra Universalis* **26** (1989), 48–72, doi:10.1007/bf01243872.

[9] J. Leech, Skew Boolean algebras, *Algebra Universalis* **27** (1990), 497–506, doi:10.1007/bf01188995.

[10] J. Leech, Normal skew lattices, *Semigroup Forum* **44** (1992), 1–8, doi:10.1007/bf02574320.

[11] J. Leech, Recent developments in the theory of skew lattices, *Semigroup Forum* **52** (1996), 7–24, doi:10.1007/bf02574077.

[12] J. Leech, My journey into noncommutative lattices and their theory, *Art Discrete Appl. Math.* **2** (2019), #P2.01 (19 pages), doi:10.26493/2590-9770.1282.e7a.

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 19 (2020) 155–171
https://doi.org/10.26493/1855-3974.2286.ece
(Also available at http://amc-journal.eu)

Oriented area as a Morse function on polygon spaces*

Daniil Mamaev †

Chebyshev Laboratory, St. Petersburg State University,
14th Line V.O., 29, Saint Petersburg 199178 Russia

Received 20 March 2020, accepted 16 July 2020, published online 13 November 2020

Abstract

We study polygon spaces arising from planar configurations of necklaces with some of the beads fixed and some of the beads sliding freely. These spaces include configuration spaces of flexible polygons and some other natural polygon spaces. We characterise critical points of the oriented area function in geometric terms and give a formula for the Morse indices. Thus we obtain a generalisation of isoperimetric theorems for polygons in the plane.

Keywords: Flexible polygons, configuration spaces, Morse index, critical points.

Math. Subj. Class. (2020): 58K05, 52B60

1 Preliminaries: necklaces, configuration spaces, and the oriented area function

Suppose one has a closed string with a number of labelled beads, a *necklace*. Some of the beads are fixed and some can slide freely (although the beads never pass through one another). Having the necklace in hand, one can try to put it on the plane in such a way that the string is strained between every two consecutive beads. We will call this *a (strained planar) configuration of the necklace*. The space of all configurations (up to rotations and translations) of a given necklace, called *the configuration space of the necklace*, together with the oriented area function on it is the main object of the present paper.

Let us now be precise. Given a tuple $(n_1, \dots, n_k)$ of positive integers and a tuple $(L_1, \dots, L_k)$ of positive reals, we define *a necklace* $\mathbf{N}$ to be a tuple $\big((n_1, L_1), \dots, (n_k, L_k)\big)$ interpreted as follows:

*I am deeply indebted to Gaiane Panina for posing the problem and supervising my research. I am also thankful to Joseph Gordon and Alena Zhukova for fruitful discussions and to Nathan Blacher for his valuable comments on the linguistic quality of the paper.

†The research is supported by «Native towns», a social investment program of PJSC «Gazprom Neft».

E-mail address: dan.mamaev@gmail.com (Daniil Mamaev)

- the necklace has the total of $n = n(\mathbf{N}) = n_1 + \cdots + n_k$ *beads* on it;
- $k = k(\mathbf{N})$ of the beads are *fixed* and numbered by the index $j = 1, \dots, k$ in counter-clockwise order, the index j is considered to be cyclic (that is, $j = 6k+5$ is the same as $j = 5$);
- there are $(n_j - 1)$ *freely sliding* beads between the j-th and the $(j+1)$-th fixed bead;
- the *total length* of the string is $L = L(\mathbf{N}) = L_1 + \cdots + L_k$;
- the length of the string between the j-th and the $(j+1)$-th fixed bead is equal to L_j.

We fix some notation concerning polygons.

- *A planar n-gon* is a collection of n (labelled) points (called *vertices*) $(p_1, \dots, p_n)$ in the Euclidean plane $\mathbb{R}^2$. Note that all kinds of degenerations, including self-intersection and collision of vertices, are allowed.
- *The space of all planar n-gons* Poly_n is thus just $\left(\mathbb{R}^2\right)^n$.
- *The edges* of a polygon $P = (p_1, \dots, p_n)$ are the segments $p_i p_{i+1}$ for $i = 1, \dots, n$, the *length of the i-th edge* is $l_i = l_i(P) = |p_i p_{i+1}|$. Note that the index $i = 1, \dots, n$ is cyclic (that is, $i = 10n + 3$ is the same as $i = 3$).

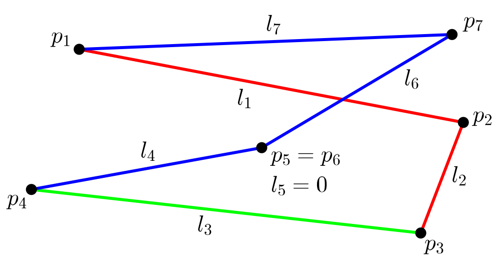
To avoid messy indices, we introduce some additional notation associated with a necklace $\mathbf{N} = \big((n_1, L_1), \dots, (n_k, L_k)\big)$ (see Figure 1 for an example). For index $j = 1, \dots, k$

- denote by j^* the set of *indices corresponding to the j-th piece of* $\mathbf{N}$:

$$j^* = \{n_1 + \cdots + n_{j-1} + 1, \dots, n_1 + \cdots + n_j\}; \tag{1.1}$$

- define a function $\mathcal{L}_j \colon \operatorname{Poly}_n \to \mathbb{R}$, *the total length of the edges of a polygon corresponding to the j-th piece of* $\mathbf{N}$, that is

$$\mathcal{L}_j(P) = \sum_{i \in j^*} l_i(P).$$



$$1^* = \{1,2\},\ 2^* = \{3\},\ 3^* = \{4,5,6,7\};$$
$$s(1) = 1,\ s(2) = 3,\ s(3) = 4;$$
$$\mathcal{L}_1(P) = l_1 + l_2 = L_1;$$
$$\mathcal{L}_2(P) = l_3 = L_2;$$
$$\mathcal{L}_3(P) = l_4 + l_6 + l_7 = L_3.$$

Figure 1: A configuration $P = (p_1, \dots, p_7)$ of $\mathbf{N} = \big((2, L_1), (1, L_2), (4, L_3)\big)$.

Definition 1.1. A (strained planar) configuration of a necklace $\mathbf{N} = \big((n_1, L_1), \dots, (n_k, L_k)\big)$ is a polygon $P \in \operatorname{Poly}_n$ with $\mathcal{L}_j(P) = L_j$ for all $j = 1, \dots, k$.

All configurations of a necklace $\mathbf{N} = \big((n_1, L_1), \ldots, (n_k, L_k)\big)$ modulo translations and rotations form a space $\mathcal{M}(\mathbf{N}) = \mathcal{M}\big((n_1, L_1), \ldots, (n_k, L_k)\big)$ called *configuration space of the necklace* $\mathbf{N}$. More formally,

- Consider all the strained planar configurations of $\mathbf{N} = \big((n_1, L_1), \ldots, (n_k, L_k)\big)$:

$$\widetilde{\mathcal{M}}(\mathbf{N}) = \{P \in \operatorname{Poly}_n : \mathcal{L}_j(P) = L_j \text{ for } j = 1, \ldots, k\}.$$

- The group $\operatorname{Iso}_+\left(\mathbb{R}^2\right)$ of orientation-preserving isometries of the Euclidean plane $\mathbb{R}^2$ acts diagonally on the space of all planar n-gons $\operatorname{Poly}_n = \left(\mathbb{R}^2\right)^n$.
- *The configuration space of the necklace* $\mathbf{N}$ is the space of orbits:

$$\mathcal{M}(\mathbf{N}) = \widetilde{\mathcal{M}}(\mathbf{N})\Big/ \operatorname{Iso}_+\left(\mathbb{R}^2\right).$$

Definition 1.2. The oriented area $\mathcal{A}$ of an n-gon $P = ((x_1, y_1), \ldots, (x_n, y_n)) \in \left(\mathbb{R}^2\right)^n$ is defined to be

$$\mathcal{A}(P) = \frac{1}{2}\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \frac{1}{2}\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \cdots + \frac{1}{2}\begin{vmatrix} x_n & x_1 \\ y_n & y_1 \end{vmatrix}. \tag{1.2}$$

The oriented area is preserved by the action of $\operatorname{Iso}_+\left(\mathbb{R}^2\right)$ and thus gives rise to a well-defined continuous function on $\mathcal{M}(\mathbf{N})$ for all the necklaces $\mathbf{N}$. We will denote these functions by the same letter $\mathcal{A}$. The study of critical points (i.e. the solutions of $d\mathcal{A}(P) = 0$) of $\mathcal{A}\colon \mathcal{M}(\mathbf{N}) \to \mathbb{R}$ is the subject of the present paper.

The paper is organised as follows. In Section 2 we review previously studied extreme cases: polygonal linkages (the 'all beads are fixed' case) and polygons with fixed perimeter (the 'one bead is fixed' case, which is clearly the same as 'none of the beads are fixed' case). In Section 3 we discuss the regularity properties of configuration spaces of necklaces. In the subsequent sections we study the non-singular part of the configuration space. In Section 4 we give a geometric description of critical points of the oriented area in the general case (Theorem 4.1) and deduce a formula for their Morse indices (Theorem 4.2). In Section 5 the auxiliary Lemmata 4.3 and 4.4 concerning orthogonality of certain spaces with respect to the Hessian form of the oriented area function are proven. In Section 6 we discuss the 'two consecutive beads are fixed' case and give a proof of Lemma 4.5.

2 An overview of existing results

2.1 Configuration spaces of polygonal linkages

In the notation of the present paper these are the spaces $\mathcal{M}\big((1, l_1), \ldots, (1, l_n)\big)$, i.e. the spaces $\mathcal{M}(\mathbf{N})$ for necklaces $\mathbf{N}$ with all beads being fixed. These spaces are studied in many aspects (see e.g. [1] or [2] for a thorough survey). On the side of studying the oriented area on these spaces, the first general fact about its critical points was noticed by Thomas Banchoff (unpublished), reproved by Khimshiashvili and Panina [5] (their technique required some non-degeneracy assumptions) and then proved again by Leger [8] in full generality.

Theorem 2.1 (*Critical configurations in the 'all beads are fixed' case*; Bunchoff, Khimshiashvili, Leger, and Panina)**.** *Let* $\mathbf{N}$ *be a necklace with all the beads fixed. Then a polygon* $P \in \mathcal{M}_{sm}(\mathbf{N})$ *is a critical point of* $\mathcal{A}$ *if and only if it is cyclic (i.e. inscribed in a circle).*

After describing critical points, the following natural question arises: are these critical points Morse (i.e. whether $\operatorname{Hess}_P \mathcal{A}$, the Hessian of $\mathcal{A}$ at P, is a non-degenerate bilinear form on $T_P\mathcal{M}(\mathbf{N})$) and if they are, what is the Morse index (the maximal dimension of a subspace on which $\operatorname{Hess}_P \mathcal{A}$ is negative definite). The state-of-the-art answer to this question requires some more notation (see Figure 2 for an example).

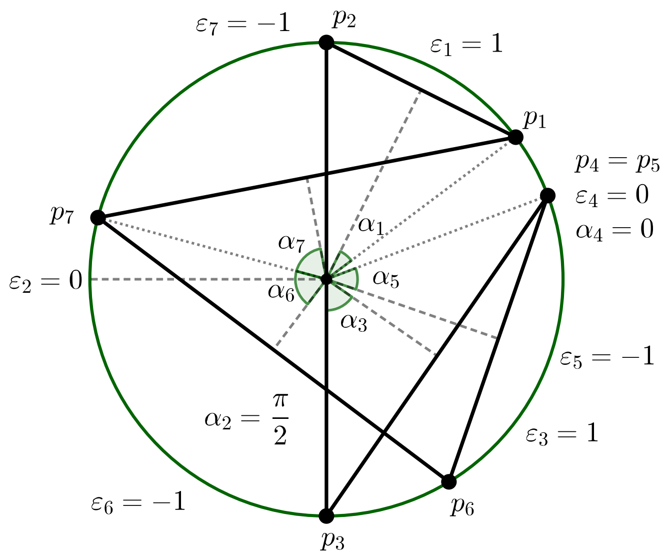


Figure 2: Notation for a cyclic polygon.

Definition 2.2. Let P be a cyclic n-gon, o be the circumcentre of P, and $i \in \{1, \dots, n\}$.

- The central half-angle of the i-th edge of P is

$$\alpha_i(P) = \frac{|\angle p_i o p_{i+1}|}{2} \in [0, \pi/2].$$

- The orientation of the i-th edge of P is

$$\varepsilon_i(P) = \begin{cases} 1, & \text{if } \angle p_i o p_{i+1} \in (0, \pi); \\ 0, & \text{if } \angle p_i o p_{i+1} \in \{0, \pi\}; \\ -1, & \text{if } \angle p_i o p_{i+1} \in (-\pi, 0). \end{cases} \tag{2.1}$$

We will denote by $\mathcal{C}_n$ *the configuration space of cyclic n-gons with at least three vertices*. More precisely,

$$\mathcal{C}_n = \left\{ P \in \left(\mathbb{R}^2\right)^n : \begin{array}{c} P \text{ is a cyclic polygon;} \\ \operatorname{AffineHull}(P) = \mathbb{R}^2 \end{array} \right\} \Big/ \operatorname{Iso}_+\left(\mathbb{R}^2\right). \tag{2.2}$$

For $P \in \mathcal{C}_n$ denote by Ω_P its circumscribed circle, by o_P its circumcentre, and by R_P the radius of Ω_P.

Definition 2.3. Let P be a cyclic polygon with at least three distinct vertices. It is called admissible if no edge of P passes through its circumcentre o. In this case its winding number $w_P = w(P, o)$ with respect to o is well-defined.

Theorem 2.4 (*Morse indices in the 'all beads are fixed' case*; Gordon, Khimshiashvili, Panina, Teplitskaya, and Zhukova)**.** *Let* $\mathbf{N} = \big((1, l_1), \dots, (1, l_n)\big)$ *be a necklace without freely moving beads, and let* $P \in \mathcal{M}_{sm}(\mathbf{N})$ *be an admissible cyclic polygon. Then* P *is a Morse point of* $\mathcal{A}$ *if and only if* $\sum_{i=1}^n \varepsilon_i \tan\alpha_i \neq 0$ *and in this case its Morse index is*

$$\mu_P(\mathcal{A}) = \#\big\{i \in \{1, \dots, n\} : \varepsilon_i > 0\big\} - 1 - 2w_P - \begin{cases} 0, & \textit{if } \sum_{i=1}^n \varepsilon_i \tan\alpha_i > 0; \\ 1, & \textit{otherwise}. \end{cases}$$

The formula more or less explicitly appeared in [6, 9], and [11], but in this form, with the precise condition of being Morse, the theorem was proved only in [3]. The following definition was first given in [3].

Definition 2.5. An admissible cyclic polygon P is called bifurcating if $\sum_{i=1}^n \varepsilon_i \tan\alpha_i = 0$.

2.2 Configuration space of n-gons with fixed perimeter

This is the space $\mathcal{M}\big((n, L)\big) = \mathcal{M}(n, L)$ (for different L these spaces are isomorphic, so usually L is set to 1). It is no secret since antiquity, that, with perimeter fixed, convex regular polygons maximise the area. All the critical points of the oriented area together with their indices were determined only in a recent paper [7] by Khimshiashvili, Panina and Siersma.

Definition 2.6. A regular star is a cyclic polygon P such that all its edges are equal and have the same orientation (see (2.1)). A complete fold is a regular star P with $p_i = p_{i+2}$ for all $i = 1, \dots, n$. It exists for even n only.

Theorem 2.7 (*Critical configurations and Morse indices in the 'one bead is fixed' case*; Khimshiashvili, Panina, and Siersma)**.**

(1) $\mathcal{M}(n, L)$ *is homeomorphic to* $\mathbb{C}\mathrm{P}^{n-2}$.

(2) A polygon $P \in \mathcal{M}_{sm}(n, L)$ *is a critical point of* $\mathcal{A}$ *if and only if it is a regular star.*

(3) The stars with maximal winding numbers are Morse critical points of $\mathcal{A}$.

(4) Under assumption that all regular stars are Morse critical points, the Morse indices are:

$$\mu_P^n(\mathcal{A}) = \begin{cases} 2w_P - 2, & \textit{if } w_P < 0; \\ 2n - 2w_P - 2, & \textit{if } w_P > 0; \\ n - 2, & \textit{if } P \textit{ is a complete fold}. \end{cases}$$

Remark 2.8. The super-index in $\mu_P(\mathcal{A})$ allows one to identify the domain of $\mathcal{A}$. For example, $\mu_P^n(\mathcal{A})$ is the Morse index of $\mathcal{A}\colon \mathcal{M}\big(n, \mathcal{L}_1(P)\big) \to \mathbb{R}$ at point P (as in (4) of Theorem 2.7), and $\mu_P^{1,\dots,1}(\mathcal{A})$ is the Morse index of $\mathcal{A}\colon \mathcal{M}\big(\big(1, \mathcal{L}_1(P)\big), \dots, \big(1, \mathcal{L}_n(P)\big)\big) \to \mathbb{R}$ at point P (as in Theorem 2.4). This notation will be of much use in the proof of Theorem 4.2.

Let us also mention an auxiliary statement proven in [7].

Lemma 2.9 (Khimshiashvili, Panina, and Siersma)**.** *Let P be a regular star which is not a complete fold with $w_P > 0$. Then P is a non-degenerate local maximum on $\mathcal{C}_n$.*

In fact, this lemma together with Theorem 2.4 and Lemma 4.3 allows one to omit the assumption in (4) of Theorem 2.7. All the critical points of the oriented area on $\mathcal{C}_n$ were described in a recent preprint [10] by Siersma.

3 Singular locus of the configuration space

Definition 3.1. Let P be a configuration of a necklace $\mathbf{N} = \big((n_1, L_1), \ldots, (n_k, L_k)\big)$. It is called non-singular if $\mathcal{L} = (\mathcal{L}_1, \ldots, \mathcal{L}_k)$ is a smooth submersion at P (i.e. $\mathcal{L}$ is differentiable at P and its differential $D_P\mathcal{L}\colon T_P \operatorname{Poly}_n \to T_P\mathbb{R}^k$ is a surjective linear map), otherwise it is called singular.

First we give a geometric characterisation of singular configurations. Consider a polygon $P = (p_1, \ldots, p_n)$, with $p_i = (x_i, y_i) \in \mathbb{R}^2$ and $l_i = |p_{i+1} - p_i| \neq 0$ for all $i = 1, \ldots, n$. Define β_i to be the oriented angle between vectors $(1, 0)$ and $(p_{i+1} - p_i)$. Denote by $s(j) = n_1 + \cdots + n_{j-1} + 1$ the index of the j-th fixed bead. Then every $\mathcal{L}_j$ is differentiable at P and the derivative of $\mathcal{L}_j$ with respect to x_i and y_i is as follows:

$$\frac{\partial \mathcal{L}_j}{\partial x_i}(P) = \begin{cases} -\cos\beta_i, & \text{if } i = s(j); \\ \cos\beta_{i-1} - \cos\beta_i, & \text{if } i \in j^* \setminus \{s(j)\}; \\ \cos\beta_{i-1}, & \text{if } i = s(j+1); \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

$$\frac{\partial \mathcal{L}_j}{\partial y_i}(P) = \begin{cases} -\sin\beta_i, & \text{if } i = s(j); \\ \sin\beta_{i-1} - \sin\beta_i, & \text{if } i \in j^* \setminus \{s(j)\}; \\ \sin\beta_{i-1}, & \text{if } i = s(j+1); \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

Definition 3.2. Let $\mathbf{N} = \big((n_1, L_1), \ldots, (n_k, L_k)\big)$ be a necklace. An index $i \in \{1, \ldots, n(\mathbf{N})\}$ is called boundary if it is equal to $s(j)$ for some $j \in \{1, \ldots, k\}$, otherwise it is called inner.

In Figure 1 the indices $1, 3, 4$ are boundary and the indices $2, 5, 6, 7$ are inner.

Lemma 3.3. *A configuration $P \in \operatorname{Poly}_n$ of the necklace $\mathbf{N} = \big((n_1, L_1), \ldots, (n_k, L_k)\big)$ is singular if and only if one of the following holds:*

(1) $l_i = 0$ for some $i \in \{1, \ldots, n(\mathbf{N})\}$;

(2) P fits in a straight line in such a way that $\beta_i = \beta_{i-1}$ for all inner indices i.

Proof. The first condition is equivalent to $\mathcal{L}$ being differentiable at P. Therefore, what is left to prove, is that for $P \in \widetilde{\mathcal{M}}(\mathbf{N})$ with no vanishing edges, the second condition holds if and only if the gradients $\operatorname{grad}_P \mathcal{L}_1, \ldots, \operatorname{grad}_P \mathcal{L}_k$ are linearly dependent.

Suppose that $\lambda_1 \operatorname{grad}_P \mathcal{L}_1 + \cdots + \lambda_k \operatorname{grad}_P \mathcal{L}_k = 0$ is a non-trivial vanishing linear combination. If $\lambda_j \neq 0$, then, using formulae (3.1) and (3.2) for boundary index $s(j)$, we get $-\lambda_j \cos\beta_{s(j)} + \lambda_{j-1}\cos\beta_{s(j)-1} = 0$ and $-\lambda_j \sin\beta_{s(j)} + \lambda_{j-1}\sin\beta_{s(j)-1} = 0$. It means that points $\lambda_j(\cos\beta_{s(j)}, \sin\beta_{s(j)}) \neq (0,0)$ and $\lambda_{j-1}(\cos\beta_{s(j)-1}, \sin\beta_{s(j)-1})$ coincide, which implies that $2(\beta_{s(j)} - \beta_{s(j)-1}) = 0$ and $\lambda_{j-1} = \cos(\beta_{s(j)} - \beta_{s(j)-1})\lambda_j \neq 0$. It follows then that $\lambda_j \neq 0$ for all $j = 1, \dots, k$, consequently, (we now use (3.1) and (3.2) for inner indices) $\beta_i = \beta_{i-1}$ for all inner indices i, meaning that P is composed of straight segments of lengths $L_1, \dots, L_k$. Taking into account previously deduced formula $2(\beta_i - \beta_{i-1}) = 0$ for boundary i, we conclude that P does satisfy condition (2). Reversing the above argument, we get the reverse implication. □

Definition 3.4. A singular configuration P of a necklace $\mathbf{N}$ is called strongly singular, if it satisfies (2) in Lemma 3.3. Otherwise it is called weakly singular.

Remark 3.5. Let $\mathbf{N} = \big((n_1, L_1), \dots, (n_k, L_k)\big)$ be a necklace. Then

- weakly singular configurations of $\mathbf{N}$ are, in a sense, inessential (for instance, $\mathcal{M}(\mathbf{N})$ is a topological manifold around them);
- if $(L_1, \dots, L_n)$ is such that $\pm L_1 \pm \cdots \pm L_n \neq 0$ for any choice of $\pm$ (such tuples are called *generic* in [2]), then there are no strongly singular configurations of $\mathbf{N}$.

Together, these two facts allow to deduce some information about topology of $\mathcal{M}(\mathbf{N})$ for generic $\mathbf{N}$ from Theorems 4.1 and 4.2, but this is not the subject of the present paper.

Now let $\widetilde{\mathcal{M}}_{sm}(\mathbf{N})$ be the set of non-singular configurations of necklace $\mathbf{N}$ and $\mathcal{M}_{sm}(\mathbf{N})$ be the non-singular part of $\mathcal{M}(\mathbf{N})$:

$$\mathcal{M}_{sm}(\mathbf{N}) = \frac{\widetilde{\mathcal{M}}_{sm}(\mathbf{N})}{\mathrm{Iso}_+\left(\mathbb{R}^2\right)} = \left\{ P \in \mathrm{Poly}_n : \begin{matrix} P \text{ is a non-singular} \\ \text{configuration of } \mathbf{N} \end{matrix} \right\} \Big/ \, \mathrm{Iso}_+\left(\mathbb{R}^2\right)$$

If these spaces are non-empty, they are smooth manifolds. This statement generalises previous results on smoothness of configuration spaces of polygonal linkages by Kapovich–Millson [4] and Farber [2].

Definition 3.6. A necklace $\mathbf{N} = \big((n_1, L_1), \dots, (n_k, L_k)\big)$ is called realisable if for all $j = 1, \dots, k$, such that $n_j = 1$, the inequality $2L_j < L_1 + \cdots + L_k$ holds.

Proposition 3.7. *Let* $\mathbf{N}$ *be a realisable necklace. Then*

(1) $\widetilde{\mathcal{M}}_{sm}(\mathbf{N})$ *is a smooth* $(2n-k)$*-dimensional submanifold of* $\mathrm{Poly}_n = \mathbb{R}^{2n}$*;*

(2) $\mathcal{M}_{sm}(\mathbf{N})$ *is a topological manifold of dimension* $2n-k-3$ *with a unique smooth structure making the quotient map* $\widetilde{\mathcal{M}}_{sm}(\mathbf{N}) \to \mathcal{M}_{sm}(\mathbf{N})$ *a smooth submersion;*

(3) the oriented area function $\mathcal{A}$ *is a smooth function on* $\mathcal{M}_{sm}(\mathbf{N})$*.*

Proof. It follows from Lemma 3.3, that the inequalities $2L_j < L_1 + \cdots + L_k$ are necessary and sufficient for $\widetilde{\mathcal{M}}_{sm}(\mathbf{N})$ to be non-empty.

The first claim is clear since $\widetilde{\mathcal{M}}_{sm}(\mathbf{N})$ is locally a level of a smooth submersion $\mathcal{L} = (\mathcal{L}_1, \dots, \mathcal{L}_k)\colon \left(\mathbb{R}^2\right)^n \to \mathbb{R}^k$.

To establish the second claim, we first note that $\mathcal{M}_{sm}(\mathbf{N})$ is an orbit space of the action of 3-dimensional Lie group $\mathrm{Iso}_+\left(\mathbb{R}^2\right)$ on the smooth manifold $\widetilde{\mathcal{M}}_{sm}(\mathbf{N})$. Thus, it suffices to observe that the action is free and proper, which is indeed the case.

The third claim is obvious since the smooth structure on $\mathcal{M}_{sm}(\mathbf{N})$ is induced from Poly_n, and the oriented area $\mathcal{A}$ is a smooth function (cf. (1.2)) on Poly_n preserved by the action of $\mathrm{Iso}_+\left(\mathbb{R}^2\right)$. □

4 Main results: critical configurations and their Morse indices in the general case

The following theorem describes critical points of the oriented area on configuration spaces of necklaces. It generalises Theorem 2.1 and (2) in Theorem 2.7.

Theorem 4.1 (Critical configurations in the general case)**.** *A polygon $P \in \mathcal{M}_{sm}\big((n_1, L_1), \dots, (n_k, L_k)\big)$ is a critical point of $\mathcal{A}$ if and only if all of the following conditions hold:*

(1) P is cyclic;

(2) $l_i(P) = L_j/n_j$ for all $i \in j^*$;

(3) $\varepsilon_{i_1}(P) = \varepsilon_{i_2}(P)$ for all $i_1, i_2 \in j^*$,

where j^* is the set of indices corresponding to the j-th piece of a necklace (see (1.1)) and $\varepsilon_i(P)$ is the orientation of the i-th edge of a cyclic polygon P (see (2.1)).

The proof essentially is a reformulation of geometric arguments into the language of Lagrange multipliers, so we first write partial derivatives of $\mathcal{A}$ with respect to x_i and y_i:

$$2 \cdot \frac{\partial \mathcal{A}}{\partial x_i}(P) = l_{i-1} \sin \beta_{i-1} + l_i \sin \beta_i \tag{4.1}$$

$$2 \cdot \frac{\partial \mathcal{A}}{\partial y_i}(P) = -l_{i-1} \cos \beta_{i-1} - l_i \cos \beta_i. \tag{4.2}$$

We follow the convention $0 \cdot \text{undefined} = 0$ hence both sides are defined for all $P \in \mathrm{Poly}_n$.

Proof of Theorem 4.1. Let P be a non-singular configuration of a necklace $\mathbf{N} = \big((n_1, L_1), \dots, (n_k, L_k)\big)$. Then P is a critical point of $\mathcal{A}$ if and only if there exist $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, such that $2 \cdot \mathrm{grad}_P\, \mathcal{A} = \lambda_1 \,\mathrm{grad}_P\, \mathcal{L}_1 + \dots + \lambda_k \,\mathrm{grad}_P\, \mathcal{L}_k$.

Assume that $2 \cdot \mathrm{grad}_P\, \mathcal{A} = \lambda_1 \,\mathrm{grad}_P\, \mathcal{L}_1 + \dots + \lambda_k \,\mathrm{grad}_P\, \mathcal{L}_k$. Applying formulae (3.1), (3.2), (4.1), (4.2) to an inner index $i \in j^*$, one deduces

$$\begin{aligned} l_{i-1} \sin \beta_{i-1} + l_i \sin \beta_i &= \lambda_j \left(\cos \beta_{i-1} - \cos \beta_i\right); \\ -l_{i-1} \cos \beta_{i-1} - l_i \cos \beta_i &= \lambda_j \left(\sin \beta_{i-1} - \sin \beta_i\right). \end{aligned}$$

If $\beta_i = \beta_{i-1}$, then $l_i = l_{i-1} = 0$, but P is non-singular, so it cannot be the case by Lemma 3.3. The only other possibility for these equations to hold is $l_{i-1} = l_i$ and $\lambda_j = l_i \cot\left(\frac{\beta_i - \beta_{i-1}}{2}\right)$. Since we have such equations for all inner indices corresponding to j, we get $l_{i_1} = l_{i_2}$ for all $i_1, i_2 \in j^*$, which implies condition (2) of the theorem. Moreover, for all $i \in j^* \setminus s(j)$ we get $\cot\left(\frac{\beta_i - \beta_{i-1}}{2}\right) = \frac{n_j \lambda_j}{L_j}$, therefore $\beta_i - \beta_{i-1}$ is the same for all $i \in j^* \setminus s(j)$, which implies that there is a circle Ω_j with centre o_j such that conditions (2) and (3) of the theorem hold. It now remains to prove that P is cyclic, i.e. o_j is the same for all $j = 1, \dots, k$. If P is a smooth point of $\mathcal{M}_{sm}((1, l_1), \dots, (1, l_n)) \subset \mathcal{M}_{sm}\big(\mathbf{N}\big)$, in other words, if P does not fit in a straight line, then we are done by Theorem 2.1. Suppose that P fits in a straight line. Pick a boundary vertex $i = s(j+1)$, denote $l^j = L_j/n_j$, and apply formulae (3.1), (3.2), (4.1), (4.2) to i:

$$\begin{aligned} l^j \sin \beta_{i-1} + l^{j+1} \sin \beta_i &= \lambda_j \cos \beta_{i-1} - \lambda_{j+1} \cos \beta_i; \\ -l^j \cos \beta_{i-1} - l^{j+1} \cos \beta_i &= \lambda_j \sin \beta_{i-1} - \lambda_{j+1} \sin \beta_i. \end{aligned}$$

Since P fits in a straight line, $2(\beta_i - \beta_{i-1}) = 0$. If $\beta_{i-1} = \beta_i = \beta$, then the points $(l^j + l^{j+1})(\cos\beta, \sin\beta)$ and $(\lambda_j - \lambda_{j+1})(\cos(\beta+\pi/2), \sin(\beta+\pi/2))$ coincide which cannot be the case since $l^j, l^{j+1} > 0$. If $\beta_{i-1} = \beta_i + \pi = \beta + \pi$, then the points $(l^j - l^{j+1})(\cos\beta, \sin\beta)$ and $(\lambda_j + \lambda_{j+1})(\cos(\beta+\pi/2), \sin(\beta+\pi/2))$ coincide, which implies that $l^j = l^{j+1}$. Since this is the case for all j, P is a complete fold and thus indeed is cyclic.

Now assume that a non-singular configuration P of necklace $\mathbf{N}$ satisfies conditions (1)–(3). Let Ω be its circumscribed circle with centre o. Denote by γ_j the oriented angle $\angle p_{s(j)} o p_{s(j)+1}$ and set $\lambda_j = l_i \cot(\gamma_j/2)$ for some index i corresponding to j. Since $\gamma_j = \beta_i - \beta_{i-1}$ for inner indices i, equality $2 \cdot \operatorname{grad}_P \mathcal{A} = \lambda_1 \operatorname{grad}_P \mathcal{L}_1 + \cdots + \lambda_k \operatorname{grad}_P \mathcal{L}_k$ holds in all inner indices. For a boundary index $i = s(j+1)$ we can (performing rotation around o) assume that $\beta_{i-1} = 0$, and what we need to check then is the following two equalities:

$$l^{j+1} \sin\beta_i = l^j \cot(\gamma_{i-1}/2) - l^{j+1} \cot(\gamma_i/2) \cos\beta_i;$$
$$-l^j - l^{j+1} \cos\beta_i = -l^{j+1} \cot(\gamma_i/2) \sin\beta_i,$$

Putting the origin at o, we note that

$$p_{i+1} - p_i = l^{j+1} \cdot (\cos\beta_i, \sin\beta_i),$$
$$p_{i+1} + p_i = l^{j+1} \cot\frac{\gamma_i}{2} \cdot (-\sin\beta_i, \cos\beta_i),$$
$$p_i = \left(\frac{l_j}{2}, -\frac{l_j}{2} \cot\frac{\gamma_{i-1}}{2}\right),$$

and thus the desired equalities are just the coordinate manifestations of the obvious identity

$$\frac{p_{i+1} - p_i}{2} - \frac{p_{i+1} + p_i}{2} + p_i = (0, 0). \qquad \square$$

The following theorem provides a criterion for an admissible cyclic polygon to be a Morse point of the oriented area and gives a formula for its Morse index. It generalises Theorem 2.4 and allows one to omit the assumption in (4) of Theorem 2.7.

Theorem 4.2 (Morse indices in the general case). *Let* $\mathbf{N} = \big((n_1, L_1), \ldots, (n_k, L_k)\big)$ *be a realisable necklace (see Definition 3.6), and* $P \in \mathcal{M}_{sm}(\mathbf{N})$ *be an admissible (see Definition 2.3) critical point of the oriented area* $\mathcal{A}$*. Then* P *is a Morse point of* $\mathcal{A}$ *if and only if it is not a bifurcating polygon (see Definition 2.5). In this case its Morse index can be computed as follows:*

$$\mu_P(\mathcal{A}) = \frac{1}{2} \sum_{j=1}^{k} (2n_j - 1) \cdot (E_j + 1) - 1 - 2w_P - \begin{cases} 0, & \text{if } \sum_{j=1}^{k} n_j E_j \tan A_j > 0; \\ 1, & \text{otherwise,} \end{cases}$$

where $E_j = \varepsilon_i$ *and* $A_j = \alpha_i$ *for some* $i \in j^*$ *(due to Theorem 4.1 this does not depend on the choice of* i*).*

Proof. Let P be as in the theorem. First, let us split the tangent space of $\mathcal{M}_{sm}(\mathbf{N})$ at the critical point P into subspaces that are orthogonal with respect to the Hessian form $\operatorname{Hess}_P \mathcal{A}$. For this, given a polygon P, we introduce the following submanifolds in $\mathcal{M}_{sm}(\mathbf{N})$:

(1) $\mathcal{E}^P = \mathcal{M}_{sm}((1, l_1), \dots, (n, l_n)) \subset \mathcal{M}_{sm}(\mathbf{N})$ is the space of all polygons having the same edge length as P;

(2) $\mathcal{C}^P = \mathcal{M}_{sm}(\mathbf{N}) \cap \mathcal{C}$ is the subspace of cyclic polygons;

(3) $\mathcal{C}_j^P = \left\{ Q \in \mathcal{M}_{sm}(\mathbf{N}) : \begin{matrix} (q_{s(j)}, \dots, q_{s(j+1)}) \text{ is cyclic} \\ q_i = p_i \text{ for } i \notin j^* \setminus \{s(j)\} \end{matrix} \right\}$ for $j = 1, \dots, k$.

We will deduce the theorem from Lemmata 4.3, 4.4, and 4.5 (see Sections 5 and 6 for their proofs).

Lemma 4.3. *Let P be as in Theorem 4.2. Then*

(1) $\mathcal{E}^P \subset \mathcal{M}_{sm}(\mathbf{N})$ is a smooth $(n-3)$-dimensional submanifold in a neighbourhood of P;

(2) $\mathcal{C}^P \subset \mathcal{M}_{sm}(\mathbf{N})$ is a smooth $(n-k)$-dimensional submanifold in a neighbourhood of P;

(3) $\mathcal{E}^P$ and $\mathcal{C}^P$ intersect transversally at P, i.e. $T_P\mathcal{M}_{sm}(\mathbf{N}) = T_P\mathcal{E}^P \oplus T_P\mathcal{C}^P$;

(4) $T_P\mathcal{E}^P$ and $T_P\mathcal{C}^P$ are orthogonal with respect to the bilinear form $\operatorname{Hess}_P \mathcal{A}$.

One can note that none of the $\mathcal{C}_j^P$ are contained in $\mathcal{C}^P$. Nonetheless, from the following lemma one sees that in the first approximation they very much are.

Lemma 4.4. *Let P be as in Theorem 4.2. Then*

(1) $\mathcal{C}_j^P \subset \mathcal{M}_{sm}(\mathbf{N})$ is a smooth $(n_j - 1)$-dimensional submanifold in a neighbourhood of P;

(2)

$$T_P\mathcal{C}^P = \bigoplus_{j=1}^{k} T_P\mathcal{C}_j^P$$

(3) $T_P\mathcal{C}_j^P$ are pairwise orthogonal with respect to the bilinear form $\operatorname{Hess}_P \mathcal{A}$.

It remains to compute the Morse index of P with respect to $\mathcal{A}$ on each $\mathcal{C}_j^P$.

Lemma 4.5. *Suppose that $P \in \mathcal{C}_{n+1}$ is such that $l_1 = \dots = l_n = L/n$ and $\varepsilon_i = 1$ $(\varepsilon_i = -1)$ for $i = 1, \dots, n$. Then P is a non-degenerate local maximum (minimum) of the oriented area on $\mathcal{M}_{sm}((n, L), (1, l_n)) \cap \mathcal{C}_{n+1}$.*

Now we are ready to prove the theorem. From Lemmata 4.3 and 4.4, P is a Morse point of $\mathcal{A}$ on $\mathcal{M}_{sm}$ if and only if it is a Morse point of $\mathcal{A}$ on $\mathcal{E}^P$ and all of $\mathcal{C}_j^P$. Since P is always a Morse point on each $\mathcal{C}_j^P$ (because by Lemma 4.5 it is a non-degenerate local extremum), it is a Morse point of $\mathcal{A}$ on $\mathcal{M}_{sm}$ if and only if it is a Morse point of $\mathcal{A}$ on $\mathcal{E}^P$, which is equivalent to P not being bifurcating by Theorem 2.4.

Moreover, Lemmata 4.3 and 4.4 imply that if P is a Morse point of $\mathcal{A}$ on $\mathcal{M}_{sm}$, then its Morse index is

$$\mu_P^{n_1,\dots,n_k}(\mathcal{A}) = \mu_P^{\mathcal{E}^P}(\mathcal{A}) + \mu_P^{\mathcal{C}^P}(\mathcal{A}) = \mu_P^{1,\dots,1}(\mathcal{A}) + \sum_{j=1}^{k} \mu_P^{\mathcal{C}_j^P}(\mathcal{A}).$$

From Theorem 2.4 we know that

$$\mu_P^{1,\dots,1}(\mathcal{A}) = \frac{1}{2}\sum_{j=1}^{k} n_j(E_j+1) - 1 - 2\omega - \begin{cases} 0, & \text{if } \sum_{j=1}^{k} n_j E_j \tan A_j > 0; \\ 1, & \text{otherwise.} \end{cases}$$

From Lemma 4.5 and (1) of Lemma 4.4 we get

$$\mu_P^{\mathcal{C}_j^P}(\mathcal{A}) = \frac{1}{2}(n_j - 1)\cdot(E_j+1).$$

Summing all up, we obtain the desired formula. □

5 Orthogonality with respect to the Hessian form of the oriented area

Let us remind that $\mathcal{C}_n$ is the configuration space of cyclic polygons with at least three different vertices (see (2.2)). First, we parametrise $\mathcal{C}_n$ smoothly. For this we introduce

$$\mathcal{H}_n = \Big\{(\theta_1,\dots,\theta_n) \in \left(S^1\right)^n : \#\{\theta_1,\dots,\theta_n\} \geq 3\Big\}\Big/ S^1,$$

where S^1 acts on $\left(S^1\right)^n$ diagonally by rotations. Consider the following map

$$\begin{aligned} \widetilde{\varphi}\colon \Big(\left(S^1\right)^n \setminus \mathit{Diag}\Big) \times \mathbb{R}_{>0} &\to \left(\mathbb{R}^2\right)^n \setminus \mathit{Diag}, \\ (\theta_1,\dots,\theta_n,R) &\mapsto R\cdot\left((\cos\theta_1,\sin\theta_1),\dots,(\cos\theta_n,\sin\theta_n)\right). \end{aligned}$$

Lemma 5.1. *$\widetilde{\varphi}$ induces a diffeomorphism $\varphi\colon \mathcal{H}_n \times \mathbb{R}_{>0} \to \mathcal{C}_n$.*

Proof. φ is obviously a bijection, so the only thing we need to check is that the Jacobian of $\widetilde{\varphi}$ has rank $(n+1)$ at every point. In fact, it is just a statement of the form '$S^1 \times \mathbb{R}_{>0}$ is diffeomorphic to $\mathbb{R}^2 \setminus \{0\}$ via polar coordinates', but we compute the Jacobian for the sake of completeness:

$$\operatorname{Jac}\varphi = \begin{pmatrix} \operatorname{Jac}_1\varphi \\ \vdots \\ \operatorname{Jac}_n\varphi \\ \operatorname{Jac}_{n+1}\varphi \end{pmatrix} = \begin{pmatrix} -R\sin\theta_1 & R\cos\theta_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -R\sin\theta_n & R\cos\theta_n \\ \cos\theta_1 & \sin\theta_1 & \cos\theta_2 & \dots & \cos\theta_n & \sin\theta_n \end{pmatrix}$$

The first n rows are obviously linearly independent. Suppose one has $\operatorname{Jac}_{n+1}\varphi = \lambda_1 \operatorname{Jac}_1\varphi + \dots + \lambda_n \operatorname{Jac}_n\varphi$. Then for any $i = 1,\dots,n$ one gets

$$(\cos\theta_i,\sin\theta_i) = \lambda_1(-R\sin\theta_i, R\cos\theta_i) = \lambda_1 R\left(\cos\left(\theta_i+\pi/2\right),\cos\left(\theta_i+\pi/2\right)\right),$$

which implies $\lambda_i = 0$, a contradiction. □

We now provide local coordinates for $\mathcal{C}_n$.

Lemma 5.2. *Let $P \in \mathcal{C}_n$ be an admissible non-bifurcating cyclic polygon with edge lengths $l_1,\dots,l_n > 0$. For $Q \in \mathcal{C}_n$ let $t_i(Q) = l_i(Q) - l_i$. Then $(t_1,\dots,t_n)$ are smooth local coordinates for $\mathcal{C}_n$ around P.*

Proof. In view of Lemma 5.1 we just need to show that for

$$\psi\colon \mathcal{H}_n \times \mathbb{R}_{>0} \to \mathbb{R}^n,$$
$$(\theta_1, \ldots, \theta_n, R) \mapsto R \cdot \left(\sqrt{2 - 2\cos(\theta_2 - \theta_1)}, \ldots, \sqrt{2 - 2\cos(\theta_1 - \theta_n)}\right)$$

$\operatorname{Jac}\psi$ is of rank n at points where $\theta_1 \neq \theta_2 \neq \cdots \neq \theta_n \neq \theta_1$. Indeed, $\operatorname{Jac}\psi$ is

$$\begin{pmatrix}
\frac{R\sin(\theta_1-\theta_2)}{\sqrt{2-2\cos(\theta_1-\theta_2)}} & 0 & \cdots & 0 & \frac{R\sin(\theta_1-\theta_n)}{\sqrt{2-2\cos(\theta_1-\theta_n)}} \\
\frac{R\sin(\theta_2-\theta_1)}{\sqrt{2-2\cos(\theta_2-\theta_1)}} & \frac{R\sin(\theta_2-\theta_3)}{\sqrt{2-2\cos(\theta_2-\theta_3)}} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{R\sin(\theta_{n-1}-\theta_n)}{\sqrt{2-2\cos(\theta_{n-1}-\theta_n)}} & 0 \\
0 & 0 & \cdots & \frac{R\sin(\theta_n-\theta_{n-1})}{\sqrt{2-2\cos(\theta_n-\theta_{n-1})}} & \frac{R\sin(\theta_n-\theta_1)}{\sqrt{2-2\cos(\theta_n-\theta_1)}} \\
\sqrt{2-2\cos(\theta_2-\theta_1)} & \sqrt{2-2\cos(\theta_3-\theta_2)} & \cdots & \sqrt{2-2\cos(\theta_n-\theta_{n-1})} & \sqrt{2-2\cos(\theta_1-\theta_n)}
\end{pmatrix}$$

Since $2(\theta_{i+1} - \theta_i) \neq 0$, all the entries are well-defined and non-zero. Consider a vanishing non-trivial linear combination of columns. The form of the first n rows forces the coefficient at the i-th column to be equal (up to the common multiplier) to $\frac{\sqrt{2-2\cos(\theta_i-\theta_{i+1})}}{\sin(\theta_i-\theta_{i+1})}$, but then for the last row we have

$$0 = \sum_{i=1}^{n} \frac{2 - 2\cos(\theta_i - \theta_{i+1})}{\sin(\theta_i - \theta_{i+1})} = 2\sum_{i=1}^{n} \tan\left(\frac{\theta_i - \theta_{i+1}}{2}\right),$$

which means exactly that P is bifurcating and contradicts the assumptions of the lemma. Thus, $\operatorname{Jac}\psi$ has rank n as desired. □

Proof of Lemma 4.3. To prove the first two claims let us note that smooth structures on $\mathcal{E}^P$, $\mathcal{C}^P$, and $\mathcal{M}_{sm}(N)$ come from the smooth structure on $\operatorname{Poly}_n = \left(\mathbb{R}^2\right)^n$. Thus, the first claim immediately follows from Lemma 3.3, as the only cyclic polygon fitting into a straight line is a complete fold, which is not admissible. The dimension of $\mathcal{E}^P$ is computed according to (2) in Proposition 3.7. From Lemma 5.1 it follows that $\mathcal{C}_n$ around P is a smooth submanifold in $\operatorname{Poly}_n/\operatorname{Iso}_+$, and from Lemma 5.2 we deduce that $\mathcal{C}^P$ around P is a smooth $(n-k)$-dimensional submanifold of $\mathcal{C}_n$ as it is a preimage of the linear subspace of codimension k in $\mathbb{R}^n$ under the map $Q \mapsto (t_1(Q), \ldots, t_n(Q))$. Thus the second claim is also proved.

The third claim is equivalent (by dimension count) to representability of every vector in $T_P\mathcal{M}_{sm}(\mathbf{N})$ as a sum of two vectors from $T_P\mathcal{E}^P$ and $T_P\mathcal{C}^P$ respectively, but this is indeed the case since every polygon Q near P in $\mathcal{M}_{sm}(\mathbf{N})$ can be obtained first by a move in $\mathcal{C}^P$ making the edges of desired length (by Lemma 5.2) and then by a move inside $\mathcal{E}^Q$.

Finally, we establish the fourth claim. Consider $v \in T_P C$ and $w \in T_P E$. To compute $\operatorname{Hess}_P \mathcal{A}(v,w)$, we choose a curve $\gamma\colon (-\varepsilon,\varepsilon) \to T_P C$ such that $\gamma(0) = P$ and $\gamma'(0) = v$, then we extend w to a vector field $W(t) \in T_{\gamma(t)} E^{\gamma(t)}$ along γ. Then

$$\operatorname{Hess}_P \mathcal{A}(v,w) = \left.\frac{d}{dt}\right|_{t=0} d_{\gamma(t)}\mathcal{A}(W(t)).$$

But $d_{\gamma(t)}\mathcal{A}$ vanishes on $T_{\gamma(t)} E^{\gamma(t)}$ by Theorem 4.1. □

The following lemma allows one to relate $\mathcal{C}_j^P$ with $\mathcal{C}^P$.

Lemma 5.3. *Let $P \in \mathcal{C}_n$ be an admissible non-bifurcating cyclic polygon such that $l_1 = l_2$ and $\angle p_1 o p_2 = \angle p_2 o p_3$, where o is the centre of the circumscribed circle Ω. Let V be a local vector field around P equal to $\left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2}\right)$ in the coordinates from Lemma 5.2. Then $(VR)(P) = 0$ and $(V d_{ab})(P) = 0$ for $a, b \in \{1,\dots,n\} \setminus \{2\}$, where Vf is the derivative of a function f along V, $R(Q)$ is the radius of the circumscribed circle of Q and $d_{ab}(Q) = |q_b - q_a|$.*

Proof. Consider a curve $P(s)\colon (-\varepsilon,\varepsilon) \to \mathcal{C}_n$, $(t_1,\dots,t_n)(P(s)) = (s,-s,0,\dots,0)$. We choose representatives $\widetilde{P}(s) \in \mathrm{Poly}_n$ in such a way that $o_{\widetilde{P}(s)} = (0,0)$ and $(p_3 - p_1)$ is codirectional with x-axes. Notice that $\widetilde{P}(-s)$ is obtained from $\widetilde{P}(s)$ by the following procedure: $p_i(-s) = p_i(s)$ for $i \neq 2$ and $p_2(-s)$ is symmetric to $p_2(s)$ relative to y-axes. From this it follows that $\widetilde{P}(s) - \widetilde{P}(-s)) = (0,0,2\eta,0,\dots,0)$ for some $\eta > 0$. Hence all p_i for $i \neq 2$ are not moving in the first approximation, which implies the statement of the lemma. □

Proof of Lemma 4.4. The space $\left\{Q \in \mathcal{M}_{sm}(\mathbf{N}) : q_i = p_i \text{ for } i \notin j^* \setminus \{s(j)\}\right\}$ is a smooth submanifold in $\mathcal{M}_{sm}(\mathbf{N})$ diffeomorphic to $\mathcal{M}_{sm}\big((n_j, L_j), (1, |p_{s(j+1)} - p_{s(j)}|)\big)$. Under this identification, $\mathcal{C}_j^P$ is just $\mathcal{C}^P$. Applying (2) of Lemma 4.3 to $\mathcal{M}_{sm}\big((n_j, L_j), (1, |p_{s(j+1)} - p_{s(j)}|)\big)$, we get the first claim.

To establish the second claim it suffices to find bases in every $T_P\mathcal{C}_j^P$ such that their disjoint union forms a basis of $T_P\mathcal{C}^P$. Consider the coordinates from Lemma 5.2. On the one hand, when we consider cyclic polygons coordinatised by $(t_1,\dots,t_n)$, the vectors $\left(\frac{\partial}{\partial t_{i-1}} - \frac{\partial}{\partial t_i}\right)$ for inner i form a basis of $T_P\mathcal{C}^P$. On the other hand, when we consider $\mathcal{C}_j^P$ coordinatised by $(s_i)_{i \in j^* \setminus \{s(j)\}}$, where $s_i = l_i(Q) - l_i(P)$, the vectors $\left(\frac{\partial}{\partial s_{i-1}} - \frac{\partial}{\partial s_i}\right)$ for $i \in j^* \setminus \{s(j)\}$ form a basis of $T_P\mathcal{C}_j^P$. It now follows from Lemma 5.3 that $\left(\frac{\partial}{\partial t_{i-1}} - \frac{\partial}{\partial t_i}\right) = \left(\frac{\partial}{\partial s_{i-1}} - \frac{\partial}{\partial s_i}\right)$ for $i \in j^* \setminus \{s(j)\}$ thus the second claim is proven.

Now we pass to the third claim. Consider $v \in T_P\mathcal{C}_j^P$ and $w \in T_P\mathcal{C}_h^P$, take a curve $\gamma\colon (-\varepsilon,\varepsilon) \to \mathcal{C}_j^P$ such that $\gamma(0) = P$ and $\gamma'(0) = v$, and a curve $\sigma\colon (-\varepsilon,\varepsilon) \to \mathcal{C}_h^P$, such that $\sigma(0) = P$ and $\sigma'(0) = w$. Then extend w to a vector field $W(t) \in T_{\gamma(t)}\mathcal{M}_{sm}(\mathbf{N})$ along γ by setting $W(t) = \sigma_t'(0)$, where $\sigma_t\colon (-\varepsilon,\varepsilon) \to \mathcal{C}_h^{\gamma(t)}$ is such that $\sigma_t(0) = \gamma(t)$ and for all $i \in j^* \setminus s(j)$ the i-th vertex of $\sigma_t(s)$ is the same as the $i - th$ vertex of $\sigma(s)$. Then

$$\operatorname{Hess}_P \mathcal{A}(v,w) = \left.\frac{d}{dt}\right|_{t=0} W(t)\mathcal{A},$$

and it vanishes since $W(t)\mathcal{A}$ does not depend on t. □

6 Configuration spaces of polygons with perimeter and one edge length fixed

These are the spaces $\mathcal{M}\big((n, L), (1, l)\big)$ for $L \geq l$. Vividly speaking, it is *the space of broken lines of given length with fixed endpoints*. One can think that the first and the last vertices have coordinates $(0, 0)$ and $(l, 0)$ respectively. Our interest in these spaces was first motivated by the fact that they are simple enough to be studied completely, but then it turned out that they are important for understanding the case of general necklaces.

Proposition 6.1 (Configuration space in the 'two consecutive beads are fixed' case)**.** *Let $L > l$ and $n \geq 2$. Then $\mathcal{M}\big((n, L), (1, l)\big)$ is homeomorphic to the sphere S^{2n-3}.*

Proof. By setting $p_1 = (0, 0)$ and $p_{n+1} = (l, 0)$ we identify $\mathcal{M}\big((n, L), (1, l)\big)$ with the level set

$$F^{-1}(L) = \Big\{(p_2, \dots, p_n) \in \left(\mathbb{R}^2\right)^{n-1} : F(p_2, \dots, p_n) = L\Big\},$$
$$\text{where}\quad F(p_2, \dots, p_n) = |p_2| + |p_3 - p_2| + \dots + |p_n - p_{n-1}| + |(l, 0) - p_n|.$$

F is a convex function as sum of convex functions. The sublevel set $F^{-1}((-\infty, L])$ is bounded since if any of $|p_i|$ is greater than L, then $F(p_2, \dots, p_n) \geq L$ by triangle inequality. Also, the set $F^{-1}((-\infty, L))$ is non-empty, since if all of the p_i are in the disk of radius δ around $(l/2, 0)$, then $F(p_2, \dots, p_n) < (l/2 + \delta) + (n - 3)\delta + (l/2 + \delta) = l + (n - 1)\delta$, which is less than L for small δ. So, $F^{-1}(L)$ is a boundary of the compact convex set $F^{-1}((-\infty, L]) \subset \left(\mathbb{R}^2\right)^{n-1}$ with non-empty interior and thus is homeomorphic to S^{2n-3}. □

The following two propositions are easily deduced from Theorems 4.1 and 4.2 respectively.

Proposition 6.2 (Critical points in the 'two consecutive beads are fixed' case)**.** *Let $L > l$ and $n \geq 2$. Then critical points of $\mathcal{A}$ on $\mathcal{M}_{sm}\big((n, L), (1, l)\big)$ are in bijection with the solutions of*

$$|U_{n-1}(x)| = \frac{nl}{L} \tag{6.1}$$

where U_{n-1} is the $(n-1)$-th Chebyshev polynomial of second kind, that is, $U_{n-1}(\cos\varphi) = \frac{\sin n\varphi}{\sin\varphi}$.

Proof. By Theorem 4.1 a configuration $P \in \mathcal{M}_{sm}\big((n, L), (1, l)\big)$ is a critical point of $\mathcal{A}$ if and only if it is inscribed in a circle Ω with centre o and radius R in such a way that $\angle p_1 o p_2 = \dots = \angle p_n o p_{n+1} =: \alpha(P) = \alpha$. Let us construct a bijection

$$\{\text{critical points of } \mathcal{A} \text{ on } \mathcal{M}_{sm}\big((n, L), (1, l)\big)\} \to \{\text{solutions of (6.1)}\},\ \ P \mapsto c_P.$$

Let $c_P = \cos(\alpha/2)$, where $\alpha/2 \in (0, \pi)$. Since $L/n = R\sqrt{2 - 2\cos\alpha} = 2R\sin(\alpha/2)$ and $l = R\sqrt{2 - 2\cos(n\alpha)} = 2R|\sin(n\alpha/2)|$, we get $U_{n-1}(c_P) = nl/L$. Since the map

$$\frac{\mathbb{R}}{2\pi\mathbb{Z}} \setminus \{0\} \to (-1, 1),\ \ \alpha \mapsto \cos(\alpha/2)$$

is a bijection and P is uniquely determined by $\alpha(P)$, the constructed map $P \mapsto c_P$ is indeed a bijection. □

Proposition 6.3 (Morse index in the 'two consecutive beads are fixed' case)**.** *If P is an admissible non-bifurcating critical configuration of $\mathcal{A}$ on $\mathcal{M}_{sm}((n, L), (1, l))$, then its Morse index is*

$$\mu_P^{n,1}(\mathcal{A}) = \begin{cases} 2n-2-i, & \text{if } c_P \text{ is the } i\text{-th largest positive solution of (6.1);} \\ i-1, & \text{if } c_P \text{ is the } i\text{-th smallest negative solution of (6.1).} \end{cases}$$

Proof. By symmetry reasons, to prove the claim, it suffices to prove it only for P with $c_P > 0$. Then by Theorem 4.2 one has

$$\mu_P^{n,1}(\mathcal{A}) = (2n-1) + \frac{1}{2}(\varepsilon_{n+1}+1) - 1 - w_P - \begin{cases} 0, & \text{if } n\tan(\alpha/2) > \varepsilon_{n+1}\tan(n\alpha/2); \\ 1, & \text{otherwise.} \end{cases}$$

The roots and extrema of $U_{n-1}(t)$ are interchanging. Let us start from $t = 1$ and move to the right. The extrema correspond to the bifurcating polygons (i.e. those with $n\tan(\alpha/2) = \varepsilon_{n+1}\tan(n\alpha/2)$) and the roots correspond to polygons with $l_{n+1} = 0$. So, when t passes a root, ε_{n+1} changes from 1 to -1 and whenever t passes an extrema, the last summand changes from 0 to 1. When p_1p_{n+1} passes through o, w_P increases by 1, and ε_{n+1} changes from -1 to 1, which does not change the Morse index. The right-most t corresponds to the global maximum, so the above argument completes the prove. □

Finally, we check the last yet unproven ingredient in the proof of Theorem 4.2.

Proof of Lemma 4.5. Let P be as in the lemma. Without loss of generality we can assume that $\Omega_P = \Omega$ is the unit circle with centre o, and, due to symmetry, it is enough to prove the statement for P with $w_P > 0$. We should prove that the function

$$\frac{\mathcal{A}}{l_{n+1}^2}\colon \left\{\text{polygons } P \text{ inscribed in the unit circle with } \frac{\mathcal{L}_1(P)}{l_{n+1}(P)} = \frac{L}{l}\right\} \to \mathbb{R}$$

attains a non-degenerate local maximum at P. For this it suffice to prove that the function

$$\begin{gathered} G\colon \left\{\begin{matrix}\text{polygons inscribed} \\ \text{in the unit circle}\end{matrix}\right\} \to \mathbb{R}, \\ G(Q) = \frac{2\mathcal{A}(Q)}{l_{n+1}(Q)^2} - \lambda\left(\frac{\mathcal{L}_1(Q)^2}{l_{n+1}(Q)^2} - \frac{L^2}{l^2}\right) - \mu\left(\frac{\mathcal{L}_1(Q)^2}{l_{n+1}(Q)^2} - \frac{L^2}{l^2}\right)^2 \end{gathered} \tag{6.2}$$

attains a non-degenerate local maximum at P for suitable λ and μ. Set $\alpha = \angle p_1op_2 = \cdots = \angle p_nop_{n+1} \in (0, \pi)$ and introduce local coordinates by setting $t_i(Q) = \angle q_ioq_{i+1} - \alpha$ for $i = 1, \ldots, n$. First, we write the functions involved in the definition (6.2) in these coordinates:

$$l_{n+1}(t_1, \ldots, t_n) = \sqrt{2 - 2\cos\left(n\alpha + \sum_{i=1}^n t_i\right)};$$

$$\mathcal{L}_1(t_1, \ldots, t_n) = \sum_{i=1}^n \sqrt{2 - 2\cos(\alpha + t_i)};$$

$$2\mathcal{A}(t_1, \ldots, t_n) = \sum_{i=1}^n \sin(\alpha + t_i) - \sin\left(n\alpha + \sum_{i=1}^n t_i\right).$$

Secondly, we perform the computations in the 2-jets at the point P, which by the aforementioned coordinates are identified with $\mathbb{R}[t_1, \dots, t_n]/I$, where I is the ideal generated by all products $t_i t_j t_h$ with $i, j, h = 1, \dots, n$. It turns out that the 2-jets of the functions we are interested in are all contained in the subring $\mathbb{R} + \mathbb{R}T_1 + \mathbb{R}T_1^2 + \mathbb{R}T_2$, where $T_1 = \sum_{i=1}^n t_i$ and $T_2 = \sum_{i=1}^n t_i^2$. This subring is naturally identified with the ring $\mathbb{R}[T_1, T_2]/(T_1^3, T_2^2, T_1 T_2)$. With all the identifications done, the 2-jets of the functions involved in the definition (6.2) look as follows:

$$j_2 l_{n+1} = l \cdot \left(1 + \frac{1}{2}\cot\left(\frac{n\alpha}{2}\right) T_1 - \frac{1}{8} T_1^2\right);$$
$$j_2 \mathcal{L}_1 = L \cdot \left(1 + \frac{1}{2n}\cot\left(\frac{\alpha}{2}\right) T_1 - \frac{1}{8n} T_2\right);$$
$$j_2(2\mathcal{A}) = (n\sin\alpha - \sin(n\alpha)) + (\cos\alpha - \cos n\alpha)T_1 - \frac{\sin\alpha}{2} T_2 + \frac{\sin(n\alpha)}{2} T_1^2.$$

Now, setting $x = \tan\frac{\alpha}{2}$ and $y = \tan\frac{n\alpha}{2}$, we can write the 2-jets of the summands in (6.2) in more or less compact form:

$$j_2\left(\frac{\mathcal{L}_1^2}{l_{n+1}^2} - \frac{L^2}{l^2}\right) = \frac{nx(1+y^2)(y-nx)}{y^3(1+x^2)} T_1 - \frac{nx^2(1+y^2)}{4y^2(1+x^2)} T_2 + C_1(n,x,y) T_1^2;$$
$$j_2\left(\frac{\mathcal{L}_1^2}{l_{n+1}^2} - \frac{L^2}{l^2}\right)^2 = \frac{n^2x^2(1+y^2)^2(y-nx)^2}{y^6(1+x^2)} T_1^2;$$
$$j_2\left(\frac{2\mathcal{A}}{l_{n+1}^2} - \frac{2\mathcal{A}(P)}{l^2}\right) = \frac{(1+y^2)(y-nx)}{2y^3(1+x^2)} T_1 - \frac{x(1+y^2)}{4y^2(1+x^2)} T_2 + C_2(n,x,y) T_1^2.$$

To get rid of T_1 in $j_2 G$ we set $\lambda = \frac{1}{2nx}$, and then finally obtain

$$j_2\big(G - G(P)\big) = -\frac{x(1+y^2)}{8y^2(1+x^2)} T_2 + \left(C_2(n,x,y) - \frac{1}{2nx} C_1(n,x,y) - \mu \cdot \frac{n^2x^2(1+y^2)^2(y-nx)^2}{y^6(1+x^2)}\right) T_1^2.$$

Note that the first summand is a negative definite quadratic form since $x > 0$. As for the second one, $nx - y \neq 0$ as P is not bifurcating, and thus, whatever C_1 and C_2 are, when μ is big enough the second term is a non-positive definite quadratic form. Therefore, G attains a non-degenerate local maximum at P for $\lambda = \frac{1}{2nx}$ and large positive μ. □

ORCID iD

Daniil Mamaev https://orcid.org/0000-0002-7606-4276

References

[1] R. Connelly and E. D. Demaine, Geometry and topology of polygonal linkages, in: C. D. Toth, J. O'Rourke and J. E. Goodman (eds.), *Handbook of Discrete and Computational Geometry*, CRC Press, Boca Raton, Florida, pp. 233–256, 2017, https://www.taylorfrancis.com/books/e/9781315119601/chapters/10.1201/9781315119601-9.

[2] M. Farber, *Invitation to Topological Robotics*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2008, doi:10.4171/054.

[3] J. Gordon, G. Panina and Y. Teplitskaya, Polygons with prescribed edge slopes: configuration space and extremal points of perimeter, *Beitr. Algebra Geom.* **60** (2019), 1–15, doi:10.1007/s13366-018-0409-3.

[4] M. Kapovich and J. Millson, On the moduli space of polygons in the euclidean plane, *J. Differential Geom.* **42** (1995), 430–464, doi:10.4310/jdg/1214457034.

[5] G. Khimshiashvili and G. Panina, Cyclic polygons are critical points of area, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **360** (2008), 238–245, doi:10.1007/s10958-009-9417-z.

[6] G. Khimshiashvili and G. Panina, On the area of a polygonal linkage, *Dokl. Akad. Nauk. Math.* **85** (2012), 120–121, doi:10.1134/s1064562412010401.

[7] G. Khimshiashvili, G. Panina and D. Siersma, Extremal areas of polygons with fixed perimeter, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **481** (2019), 136–145, doi:10.1007/s10958-020-04835-9.

[8] J. C. Leger, Aire, périmètre et polygones cocycliques, 2018, `arXiv:1805.05423 [math.CO]`.

[9] G. Panina and A. Zhukova, Morse index of a cyclic polygon, *Cent. Eur. J. Math.* **9** (2011), 364–377, doi:10.2478/s11533-011-0011-5.

[10] D. Siersma, Extremal areas of polygons, sliding along a circle, 2020, `arXiv:2001.10882 [math.CO]`.

[11] A. Zhukova, Morse index of a cyclic polygon II, *St. Petersburg Math. J.* **24** (2013), 461–474, doi:10.1090/s1061-0022-2013-01247-7.

Author Guidelines

Before submission

Papers should be written in English, prepared in LaTeX, and must be submitted as a PDF file. The title page of the submissions must contain:

- *Title*. The title must be concise and informative.
- *Author names and affiliations*. For each author add his/her affiliation which should include the full postal address and the country name. If avilable, specify the e-mail address of each author. Clearly indicate who is the corresponding author of the paper.
- *Abstract*. A concise abstract is required. The abstract should state the problem studied and the principal results proven.
- *Keywords*. Please specify 2 to 6 keywords separated by commas.
- *Mathematics Subject Classification*. Include one or more Math. Subj. Class. (2020) codes – see https://mathscinet.ams.org/mathscinet/msc/msc2020.html.

After acceptance

Articles which are accepted for publication must be prepared in LaTeX using class file amcjoucc.cls and the bst file amcjoucc.bst (if you use BibTeX). If you don't use BibTeX, please make sure that all your references are carefully formatted following the examples provided in the sample file. All files can be found on-line at:

https://amc-journal.eu/index.php/amc/about/submissions/#authorGuidelines

Abstracts: Be concise. As much as possible, please use plain text in your abstract and avoid complicated formulas. Do not include citations in your abstract. All abstracts will be posted on the website in fairly basic HTML, and HTML can't handle complicated formulas. It can barely handle subscripts and greek letters.

Cross-referencing: All numbering of theorems, sections, figures etc. that are referenced later in the paper should be generated using standard LaTeX \label{...} and \ref{...} commands. See the sample file for examples.

Theorems and proofs: The class file has pre-defined environments for theorem-like statements; please use them rather than coding your own. Please use the standard \begin{proof} ... \end{proof} environment for your proofs.

Spacing and page formatting: Please do not modify the page formatting and do not use \medbreak, \bigbreak, \pagebreak etc. commands to force spacing. In general, please let LaTeX do all of the space formatting via the class file. The layout editors will modify the formatting and spacing as needed for publication.

Figures: Any illustrations included in the paper must be provided in PDF format, or via LaTeX packages which produce embedded graphics, such as TikZ, that compile with PdfLaTeX. (Note, however, that PSTricks is problematic.) Make sure that you use uniform lettering and sizing of the text. If you use other methods to generate your graphics, please provide .pdf versions of the images (or negotiate with the layout editor assigned to your article).

Subscription

Yearly subscription: 150 EUR

Any author or editor that subscribes to the printed edition will receive a complimentary copy of *Ars Mathematica Contemporanea*.

Subscription Order Form

Name: ....
E-mail: ....
Postal Address: ....
....
....
....

I would like to subscribe to receive ...... copies of each issue of *Ars Mathematica Contemporanea* in the year 2020.

I want to renew the order for each subsequent year if not cancelled by e-mail:

☐ Yes ☐ No

Signature: ....

Please send the order by mail, by fax or by e-mail.

By mail: Ars Mathematica Contemporanea
UP FAMNIT
Glagoljaška 8
SI-6000 Koper
Slovenia

By fax: +386 5 611 75 71

By e-mail: info@famnit.upr.si

Printed in Slovenia by *IME TISKARNE*