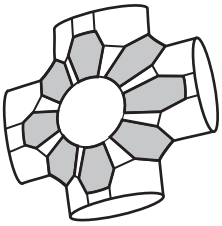


ARS MATHEMATICA CONTEMPORANEA

Volume 9, Number 2, Fall/Winter 2015, Pages 145–344

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

The University of Primorska

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

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

Coping with backlog

If we consider the quality of our journal there is one weakness that we find more severe than anything else: the *backlog*. This is the span of time between the moment a paper is submitted up to the time the printed copy of the journal reaches the reader. Each article is assigned three dates: received (R), accepted (A), and published on-line (P). Another important date is the date of printed copy.

Usually the on-line version of the paper is available very much earlier than the printed version, and because the electronic version of our journal is freely available (under so-called *diamond open access*), readers are not hurt by the difference in time between the publication of the electronic and paper versions of the paper. Part of this lag is legitimate and is depends on the technology: it is possible to put papers on-line separately, but one has to wait for the last article to be ready before the whole issue is sent to the printer.

When we set up the journal in 2007, our primary concern was to ensure about 20 high-quality papers per year. It was not clear that we would get sufficiently many enthusiasts who would be willing to submit their good papers for publication in an unknown journal with uncertain future. At first we relied on papers arising from conference series such as the 4-yearly Slovenian Graph Theory conference, GEMS, and SIGMAP. The rigidity of special issues, however, proved to be a far greater problem than we initially envisaged.

When the journal's visibility increased and its high quality became apparent, the flow of manuscripts increased, and very quickly we had more papers accepted than we needed for a single year. We adopted three strategies for reducing the backlog:

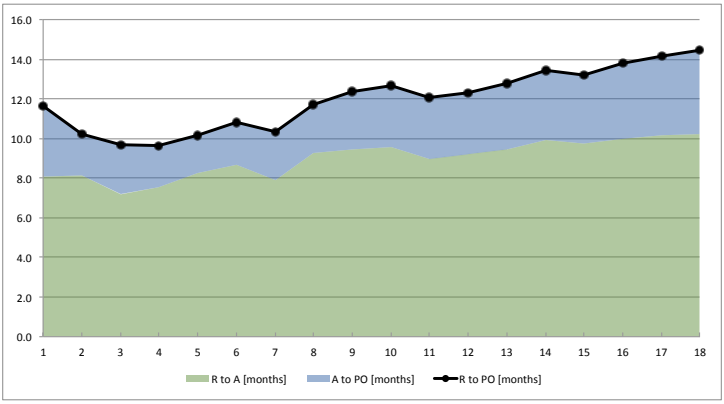
(a) We started opening future issues. This moved the problem into the future, and also gave us more flexibility in numbering the papers. But this has to be done with care: once page numbers are assigned to the electronic version of an article, they have to remain the same in the printed version, and so we cannot open the second issue of a given volume until the first one is completed.

(b) We increased the number of papers publishable each year (from the initial 20 to 60 or more), by producing two volumes each year, and publishing more articles per issue.

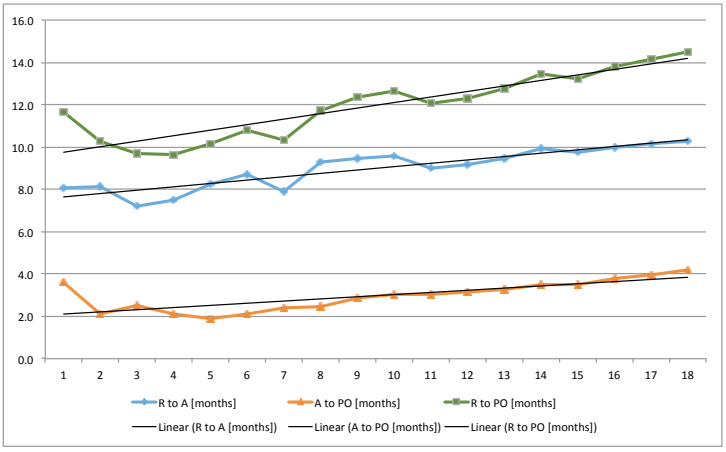
(c) We also raised the standards for acceptance, and so now the rate of acceptance of papers has been reduced to 20 per cent.

The increased volume of submissions has increased the workload for our editorial staff, and in turn this has increased the length of time taken for the review stage. Next year we will involve more Editors in managing papers. Hopefully this will reduce the backlog at this stage of the process.

Nino Bašić, Selena Praprotnik and Gordon Williams have collected and analysed data on all papers published so far in our journal. Back in 2008, it took on average less than 12 months to process a paper for the first issue of our journal. In contrast, it took over 21 months on the average to process a paper for issue Vol. 8, no. 2 in 2015. Our goal is to reduce the average time to under 12 months, as soon as possible.



The first figure (above) is a stacked plot showing the cumulative average processing time by issue: green in the bottom is the time from receipt to accepted (RA), blue at the top is the additional time from accepted to published online (AP), and the thick black line indicates the total time from receipt to published online (RP).



The second figure (above) gives plots of receipt to accepted (RA), accepted to published online (AP) and total time from receipt to published online (RP), with linear regression lines superimposed.

Unfortunately, the length of time from accepted to published on-line continues to grow as well, so our policy of slow growth and opening additional volumes is not enough. We hope that moving from 20 papers per year to the current 60 papers per year, better handling of special issues, and further management of the acceptance rate, will stabilise the backlog at an appropriate level.

Dragan Marušič and Tomaž Pisanski

Editors In Chief

Contents

Volume 9, Number 2, Fall/Winter 2015, Pages 145–344

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 145–149

Construction of planar 4-connected triangulations

Gunnar Brinkmann

Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281 S9, 9000 Ghent, Belgium

Craig Larson

Department of Mathematics and Applied Mathematics, Virginia Commonwealth University, 4106 Grace E. Harris Hall, 1015 Floyd Avenue, Richmond, VA 23284-2014

Jasper Souffriau

Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281 S9, 9000 Ghent, Belgium

Nico Van Cleemput

Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281 S9, 9000 Ghent, Belgium

Received 20 March 2013, accepted 14 August 2013, published online 21 November 2014

Abstract

In this article we describe a recursive structure for the class of 4-connected triangulations or – equivalently – cyclically 4-connected plane cubic graphs.

Keywords: Planar triangulation, cubic graph, generation, recursive structure.

Math. Subj. Class.: 05C10, 05C30, 05C75

Introduction

A recursive structure for a class C of graphs is a base set $B \subset C$ of initial graphs together with a set of *operations* on graphs that transform a graph in C to another graph in C so that each graph in C can be constructed from a graph in B by a sequence of these operations.

E-mail addresses: Gunnar.Brinkmann@UGent.be (Gunnar Brinkmann), clarson@vcu.edu (Craig Larson), Jasper.Souffriau@UGent.be (Jasper Souffriau), Nicolas.VanCleemput@UGent.be (Nico Van Cleemput)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

An *operation* is typically the replacement of a finite substructure by another – larger – substructure. In the ideal case, the set B as well as the set of operations are finite and small. All graphs discussed in this article are simple.

The two main applications for recursive structures are structure generation programs and inductive proofs, where the recursive structures describe the induction step. In this paper we discuss planar triangulations – that is plane graphs where every face is a triangle. For several classes of triangulations, recursive structures have been published: for all triangulations (that is: 3-connected triangulations) [6], for 5-connected triangulations [1][5], for triangulations with minimum degree 4 [2], for 3- and 4- connected triangulations with minimum degree 5 [3], and for Eulerian triangulations [2]. In the dual, these are constructions for 3-connected planar cubic graphs, cyclically 5-connected planar cubic graphs, 3-connected planar cubic graphs with girth 4, 3- resp. cyclically 4-connected planar cubic graphs with girth 5 and 3-connected bipartite planar cubic graphs.

In this article we will add the missing link between 3-connected triangulations and 5-connected triangulations and give a recursive structure for 4-connected triangulations. The operations necessary to construct all 4-connected triangulations are in fact the same as the ones used in [4] to construct all triangulations with minimum degree 4 – except for the operation inducing separating triangles. While it is obvious that an operation introducing separating triangles does not lead to 4-connected triangulations, it is not obvious that all 4-connected triangulations can be obtained with the remaining two operations.

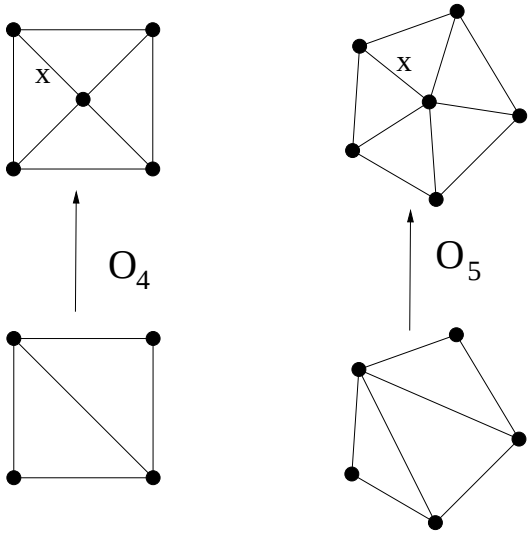


Figure 1: Two of the operations used by Eberhard [6] to generate all triangulations. Edges and vertices outside of the bounding 4-, or 5-cycle in the figure are not drawn.

Two of the operations given by Eberhard to construct all triangulations are given in Figure 1. We will show:

Theorem 0.1. *The class $\mathcal{C}_4$ of all 4-connected triangulations can be generated from the octahedron graph (depicted in Figure 2) by operations O_4 and O_5.*

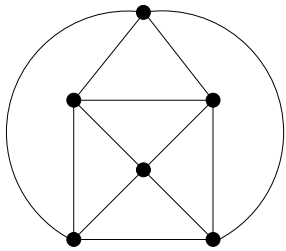


Figure 2: The octahedron graph.

Proof. We will write $\bar{\mathcal{C}}_4$ for the class $\mathcal{C}_4$ without the octahedron graph.

The operations O_4 and O_5 are in fact similar to special cases of the edge expansion operation used by Batagelj in [2]. This can best be seen when looking at the reduction – that is the inverse of the construction operation. If one compresses the edges marked with an x (that is: removes the edge and identifies the endpoints) in Figure 1, the resulting graph is the same as after replacing the vertices and their adjacent edges by one, resp. two edges.

To prove this theorem, note first that in a triangulation being 4-connected is equivalent to not having a separating – that is: non-facial – 3-cycle. We will show that for each element of the class $\bar{\mathcal{C}}_4$ an inverse operation can be applied that does not introduce separating 3-cycles and therefore leads to an element of $\mathcal{C}_4$.

This is the consequence of 3 observations;

(a) In a 4-connected triangulation no two edges in the same facial triangle belong to the same separating 4-cycle.

This follows immediately as in that case the other edges of the separating 4-cycle together with the third edge of the triangle would form a separating 3-cycle.

(b) In a 4-connected triangulation that is not the octahedron graph, no two edges in the same facial triangle with a common vertex v of degree 4 belong to different separating 4-cycles C, C'.

Suppose that this was the case. Then – due to (a) – the two separating 4-cycles must cross each other and there is an edge $\{v, y_1\}$ belonging to (w.l.o.g.) C so that the next edges $\{v, x_1\}$, $\{v, x_2\}$ in counterclockwise, resp. clockwise direction around v belong to the separating 4-cycle C' formed by the vertices x_1, v, x_2, a. This situation is depicted in Figure 3.

From the previous observation it follows that C cannot contain x_1 or x_2, so the Jordan curve theorem gives that it must contain a and that the situation is as with the dotted edges in Figure 3. This implies the presence of 8 triangles which must all be facial triangles – as no non-facial triangles exist – and implies that there are no more edges than those depicted. So the graph was the octahedron graph.

(c) In a 4-connected triangulation without vertices of degree 4, for each edge $\{v, x_1\}$ containing a vertex v of degree 5 that belongs to a separating 4-cycle C, either the previous or the next edge in the cyclic order around v or both do not belong to a separating 4-cycle.

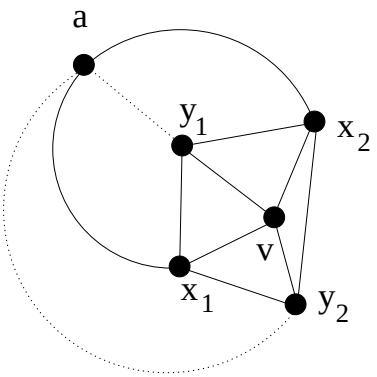


Figure 3: Two separating 4-cycles crossing in a vertex of degree 4.

By choosing the neighboring edge as the one that shares a triangle with both edges of C containing v, we can follow the same line of arguments as before to get – up to symmetry – the situation in Figure 4. In this case we don't have 8 triangles, but we do have the triangles (a, y_1, x_2), (x_1, y_1, a), (v, y_1, x_1) and (x_2, y_1, v) which must all be facial. This implies that the degree of y_1 is 4 – contradicting the assumption.

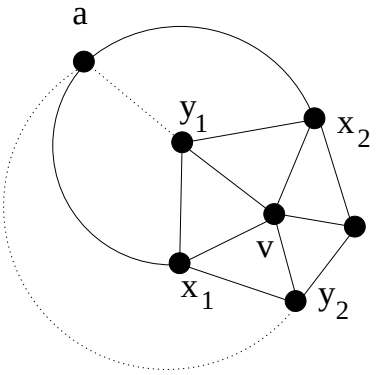


Figure 4: Two separating 4-cycles crossing in a vertex of degree 5 in a triangulation with minimum degree 5.

As in a 4-connected triangulation there are always vertices with degree 4 or degree 5, (a),(b),(c) together imply that a triangulation in $\bar{\mathcal{C}}_4$ contains an edge adjacent to a vertex of degree 4 or 5 that does not lie on a separating 4-cycle. Using this edge as the edge x in Figure 1 we can reduce such a triangulation to a smaller one without separating triangles. □

References

[1] D. Barnette, On generating planar graphs, *Discrete Math.* **7** (1974), 199–208.

[2] V. Batagelj, An improved inductive definition of two restricted classes of triangulations of the plane, *Combinatorics and Graph Theory, Banach Center Publications* **25** (1989), 11–18.

[3] G. Brinkmann and B. D. McKay, Construction of planar triangulations with minimum degree 5, *Discrete Math.* **301** (2005), 147–163.

[4] G. Brinkmann and B. D. McKay, Fast generation of planar graphs, *MATCH Commun. Math. Comput. Chem.* **58** (2007), 323–357, see `http://cs.anu.edu.au/~bdm/index.html`.

[5] J. W. Butler, A generation procedure for the simple 3-polytopes with cyclically 5-connected graphs, *Can. J. Math.* **26** (1974), 686–708.

[6] V. Eberhard, *Zur Morphologie der Polyeder*. Teubner, 1891.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 151–163

On minimal forbidden subgraphs for the class of EDM-graphs

Gašper Jaklič
FMF and IMFM, University of Ljubljana,
Jadranska 19, 1000 Ljubljana, Slovenia
and
IAM, University of Primorska, Slovenia

Jolanda Modic *
XLAB d.o.o., Pot za Brdom 100, 1000 Ljubljana, Slovenia
and
FMF, University of Ljubljana, Slovenia

Received 2 April 2013, accepted 2 April 2014, published online 21 November 2014

Abstract

In this paper, a relation between graph distance matrices and Euclidean distance matrices (EDM) is considered. Graphs, for which the distance matrix is not an EDM (NEDM-graphs), are studied. All simple connected non-isomorphic graphs on $n \leq 8$ nodes are analysed and a characterization of the smallest NEDM-graphs, i.e., the minimal forbidden subgraphs, is given. It is proven that bipartite graphs and some subdivisions of the smallest NEDM-graphs are NEDM-graphs, too.

Keywords: Graph, Euclidean distance matrix, distance, eigenvalue.

Math. Subj. Class.: 15A18, 05C50, 05C12

1 Introduction

A matrix $D \in \mathbb{R}^{n\times n}$ is *Euclidean distance matrix (EDM)*, if there exist $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \mathbb{R}^r$, such that $d_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2$, $i, j = 1, 2, \ldots, n$. The minimal possible r is called *the embedding dimension* (see [2], e.g.).

*Corresponding author. This research was funded in part by the European Union, European Social Fund, Operational Programme for Human Resources, Development for the Period 2007-2013. The author would like to thank XLAB d.o.o., Dr. Daniel Vladušič and Dr. Gregor Berginc for all the support.

E-mail addresses: gasper.jaklic@fmf.uni-lj.si (Gašper Jaklič), jolanda.modic@gmail.com (Jolanda Modic)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

Euclidean distance matrices were introduced by Menger in 1928 and have received a considerable attention. They were studied by Schoenberg [13], Young and Householder [14], Gower [4], and many other authors. In recent years many new results were obtained (see [5, 7, 8, 11] and the references therein).

They are used in various applications in linear algebra, graph theory, geodesy, bioinformatics, chemistry, e.g., where frequently a question arises, what can be said about a set of points, if only interpoint distance information is known. Some examples can be found in [2].

EDMs have many interesting properties. They are symmetric, hollow (i.e., with only zeros on the diagonal) and nonnegative. The sum of their eigenvalues is zero and they have exactly one positive eigenvalue (for a nonzero matrix). Schoenberg ([13]), Hayden, Reams and Wells ([5]) gave the following characterization of EDMs.

Theorem 1.1. *Let $D \in \mathbb{R}^{n\times n}$ be a nonzero symmetric hollow matrix and let $\boldsymbol{e} \in \mathbb{R}^n$ be the vector of ones. The following propositions are equivalent:*

(a) The matrix D is EDM.

(b) For all $\boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{x}^T\boldsymbol{e} = 0$, $\boldsymbol{x}^T D\boldsymbol{x} \leq 0$.

(c) The matrix D has exactly one positive eigenvalue and there exists $\boldsymbol{w} \in \mathbb{R}^n$ such that

$$D\boldsymbol{w} = \boldsymbol{e} \tag{1.1}$$

and $\boldsymbol{w}^T\boldsymbol{e} \geq 0$.

Throughout the paper we will use the notation $\boldsymbol{e}$ for the vector of ones of appropriate size. Vectors $\boldsymbol{e}_i$ will denote the standard basis.

Let $\mathcal{G}$ be a graph with a vertex set $\mathcal{V}(\mathcal{G})$ and an edge set $\mathcal{E}(\mathcal{G})$. Let the distance $d(u, v)$ between vertices $u, v \in \mathcal{V}(\mathcal{G})$ be defined as their graph distance, i.e., the length of the shortest path between them. Let $G := [d(u, v)]_{u,v\in\mathcal{V}(\mathcal{G})}$ be the distance matrix of $\mathcal{G}$.

If the graph distance matrix of a graph is EDM, the graph is called *an EDM-graph*. Otherwise the graph is *a NEDM-graph*.

Graph distance matrices of EDM-graphs were studied in several papers. Path and cycles were analysed in [9]. Star graphs and their generalizations were considered in [6, 10]. Some results on Cartesian products of EDM-graphs are also known (see [11]). However, the characterization of EDM-graphs in general is still an open problem.

In this paper, all simple connected non-isomorphic graphs on $n \leq 8$ nodes are analysed and a characterization of the smallest NEDM-graphs, i.e., the minimal forbidden subgraphs, is given.

In algebraic graph theory, a lot is known on the adjacency matrix and the Laplacian matrix of a graph. Many results on their eigenvalues exist, but not much is known on the graph distance matrix. Hopefully, this paper will provide a deeper insight into the relation between general graphs or networks and EDM theory.

There are some interesting possibilities of application. Molecular conformation in bioinformatics, dimensionality reduction in statistics, 3D reconstruction in computer vision, just to name a few.

The structure of the paper is as follows. In Section 2, all NEDM-graphs on $n \leq 8$ nodes are considered. Analysis of their properties enables us to find some larger NEDM-graphs, which are presented in sections 3 and 4. A proof that bipartite graphs are NEDM-graphs

is given. We present two families of subdivision graphs of the smallest NEDM-graphs that are NEDM-graphs, too.

There exist graphs, for which the system (1.1) has no solution. Such graphs are studied in Section 5.

The paper is concluded with an example, where we show that not all subdivisions of graphs result in NEDM-graphs.

2 The smallest NEDM-graphs

In this section we consider simple connected non-isomorphic graphs on $n \leq 5$ nodes and find the smallest NEDM-graphs.

There is one simple connected graph on 2 nodes, the path graph $\mathcal{P}_2$, and there exist only two simple connected graphs on 3 nodes, the path graph $\mathcal{P}_3$ and the cycle graph $\mathcal{C}_3$. In [9] it was proven that path graphs and cycle graphs are EDM-graphs.

For $n = 4$, there are 6 simple connected graphs (see Fig. 1). First four of them are the star graph $\mathcal{S}_4$, the path graph $\mathcal{P}_4$, the cycle graph $\mathcal{C}_4$ and the complete graph $\mathcal{K}_4$, respectively, which are EDM-graphs (see [9, 10]). Therefore we only need to consider the last two graphs, $\mathcal{G}_4^{(5)}$ and $\mathcal{G}_4^{(6)}$.

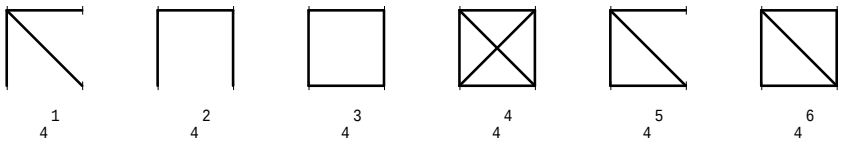


Figure 1: Simple connected graphs on 4 nodes.

Let us denote vertices of graphs $\mathcal{G}_4^{(5)}$ and $\mathcal{G}_4^{(6)}$ counterclockwise by 1, 2, 3 and 4 starting with the upper right vertex. The characteristic polynomials of the corresponding graph distance matrices

$$G_4^{(5)} = \begin{bmatrix} 0 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad G_4^{(6)} = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

are

$$p_{G_4^{(5)}}(\lambda) = (\lambda+1)(\lambda^3 - \lambda^2 - 11\lambda - 7),$$
$$p_{G_4^{(6)}}(\lambda) = (\lambda+1)(\lambda+2)(\lambda^2 - 3\lambda - 2).$$

Thus matrices $G_4^{(5)}$ and $G_4^{(6)}$ have eigenvalues

$$\sigma_{G_4^{(5)}} \doteq \{4.1, -0.7, -1, -2.4\} \quad \text{and} \quad \sigma_{G_4^{(6)}} = \left\{ \frac{3+\sqrt{17}}{2}, \frac{3-\sqrt{17}}{2}, -1, -2 \right\}.$$

Eigenvalues for $G_4^{(5)}$ were calculated numerically. Exact values can be calculated by using Cardano's formula.

One can easily verify that vectors

$$\boldsymbol{w}_{G_4^{(5)}} = [3/7,\ -1/7,\ 2/7,\ 2/7]^T \quad \text{and} \quad \boldsymbol{w}_{G_4^{(6)}} = [1/2,\ 0,\ 1/2,\ 0]^T$$

satisfy the equation $G_4^{(i)}\boldsymbol{w}_{G_4^{(i)}} = \boldsymbol{e}$, $i = 5, 6$. Since $\boldsymbol{w}^T_{G_4^{(i)}}\boldsymbol{e} > 0$, $i = 5, 6$, by Theorem 1.1 graphs $\mathcal{G}_4^{(5)}$ and $\mathcal{G}_4^{(6)}$ are EDM-graphs. Thus there are no NEDM-graphs on 4 nodes.

In the case $n = 5$, there are 21 simple connected graphs (see Fig. 2).

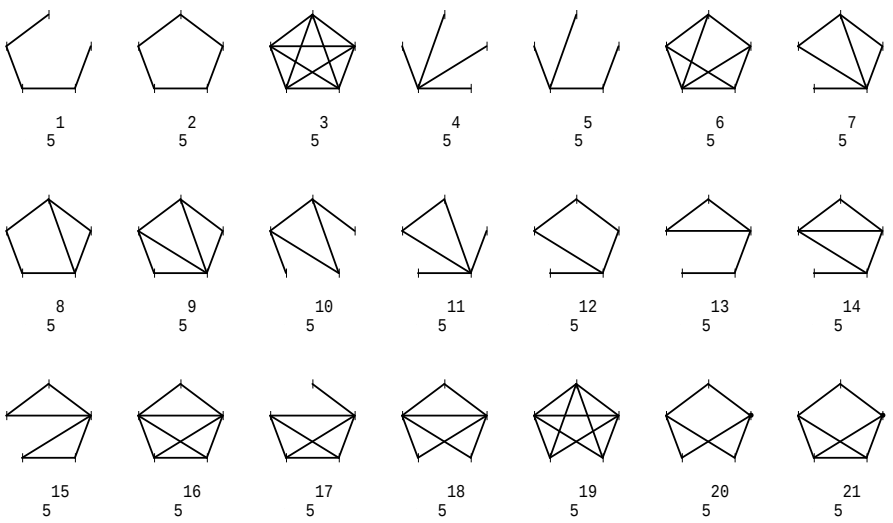


Figure 2: Simple connected graphs on 5 nodes.

Graphs $\mathcal{G}_5^{(i)}$, $i \leq 5$, are the path graph $\mathcal{P}_5$, the cycle graph $\mathcal{C}_5$, the complete graph $\mathcal{K}_5$, the star graph $\mathcal{S}_5$ and the tree $\mathcal{T}_5$, respectively. Since they are EDM-graphs (see [1]), we only need to analyse graphs $\mathcal{G}_5^{(i)}$, $i = 6, 7, \ldots, 21$.

A straightforward calculation shows that the graph distance matrix $G_5^{(i)}$ of the graph $\mathcal{G}_5^{(i)}$, $i = 6, 7, \ldots, 19$, has exactly one positive eigenvalue and that there exists $\boldsymbol{w}_{G_5^{(i)}} \in \mathbb{R}^5$, such that $G_5^{(i)}\boldsymbol{w}_{G_5^{(i)}} = \boldsymbol{e}$ and $\boldsymbol{w}^T_{G_5^{(i)}}\boldsymbol{e} \geq 0$. By Theorem 1.1, graphs $\mathcal{G}_5^{(6)}, \mathcal{G}_5^{(7)}, \ldots, \mathcal{G}_5^{(19)}$ are EDM-graphs.

We are left with graphs $\mathcal{G}_5^{(20)}$ and $\mathcal{G}_5^{(21)}$ (see Fig. 3). The characteristic polynomials of the corresponding graph distance matrices

$$G_5^{(20)} = \begin{bmatrix} 0 & 2 & 2 & 1 & 1 \\ 2 & 0 & 2 & 1 & 1 \\ 2 & 2 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 2 & 0 \end{bmatrix} \quad \text{and} \quad G_5^{(21)} = \begin{bmatrix} 0 & 2 & 2 & 1 & 1 \\ 2 & 0 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 2 & 0 \end{bmatrix}$$

are

$$p_{G_5^{(20)}}(\lambda) = -(\lambda + 2)^3(\lambda^2 - 6\lambda + 2),$$
$$p_{G_5^{(21)}}(\lambda) = -(\lambda + 1)(\lambda + 2)(\lambda^3 - 3\lambda^2 - 12\lambda + 2).$$

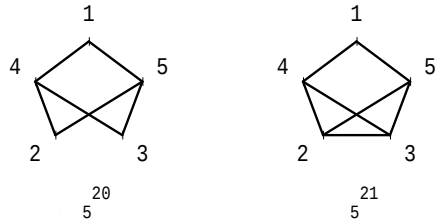


Figure 3: The graphs $\mathcal{G}_5^{(20)}$ and $\mathcal{G}_5^{(21)}$.

Thus matrices $G_5^{(20)}$ and $G_5^{(21)}$ have spectra

$$\sigma_{G_5^{(20)}} = \{3+\sqrt{7}, 3-\sqrt{7}, -2, -2, -2\} \quad \text{and} \quad \sigma_{G_5^{(21)}} \doteq \{5.2, 0.2, -1, -2, -2.4\}\,.$$

Exact eigenvalues for $G_5^{(21)}$ can be calculated by using Cardano's formula. Here they were calculated numerically. Since matrices $G_5^{(20)}$ and $G_5^{(21)}$ have two positive eigenvalues, graphs $\mathcal{G}_5^{(20)}$ and $\mathcal{G}_5^{(21)}$ are NEDM-graphs. These are the smallest NEDM-graphs.

An *induced subgraph* $\mathcal{H}$ of a graph $\mathcal{G}$ is a subset of the vertices $\mathcal{V}(\mathcal{G})$ together with all edges whose endpoints are both in this subset.

Proposition 2.1. *Let $\mathcal{G}$ be a simple connected graph and let $\mathcal{H}$ be its induced subgraph. If $\mathcal{H}$ is a NEDM-graph, the graph $\mathcal{G}$ is a NEDM-graph as well.*

Proof. Let n and $m < n$, denote the number of nodes in graphs $\mathcal{G}$ and $\mathcal{H}$, respectively. Let us order vertices of the graph $\mathcal{G}$ in such a way that the first m vertices are the vertices of the graph $\mathcal{H}$. Thus the distance matrix G of the graph $\mathcal{G}$ is of the form

$$G = \begin{bmatrix} H & * \\ * & * \end{bmatrix},$$

where H is the distance matrix of the graph $\mathcal{H}$. Every principal submatrix of an EDM has to be an EDM as well. Thus since H is not an EDM, neither is G. Therefore $\mathcal{G}$ is a NEDM-graph. □

All NEDM-graphs form a set of *forbidden subgraphs* of the class of EDM-graphs. Graphs $\mathcal{G}_5^{(20)}$ and $\mathcal{G}_5^{(21)}$ are the *minimal forbidden subgraphs*. All minimal forbidden subgraphs on 6 and 7 nodes can be seen in Fig. 4 and Fig. 5.

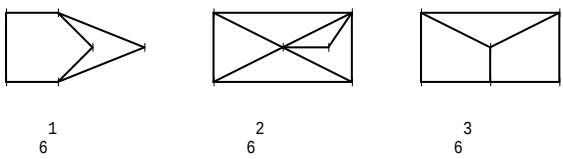


Figure 4: NEDM-graphs for $n = 6$.

Let $m(n)$ be the number of NEDM-graphs on n nodes and let $m_{new}(n)$ be the number of NEDM-graphs on n nodes for which none of the induced subgraphs is NEDM-graph.

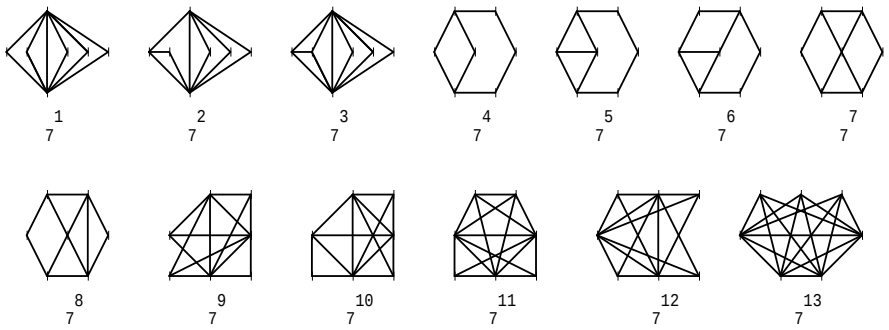


Figure 5: NEDM-graphs for $n = 7$.

We denote the number of non-isomorphic simple connected graphs on n nodes by $g(n)$. Table 1 shows how numbers $m(n)$ and $m_{new}(n)$ grow with n.

The calculations were done in the following way. By using program *geng* in Nauty ([12]) we generated all simple connected non-isomorphic graphs on $n \leq 8$ nodes. Then we applied Theorem 1.1 to determine whether a graph is an EDM-graph. Computations were done in Mathematica.

n	$g(n)$	$m(n)$	$m_{new}(n)$
5	21	2	2
6	112	27	3
7	853	341	13
8	11117	7946	48

Table 1: Number of NEDM-graphs compared to the number of all graphs on n nodes.

3 Bipartite graphs

A quick observation shows that the graph $\mathcal{G}_5^{(20)}$ is bipartite (see Fig. 3).

Let $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$ be a simple connected bipartite graph on $n \geq 5$ nodes, whose vertices are divided into two disjoint sets $\mathcal{U}_k = \{u_1, u_2, \ldots, u_k\}$, $\mathcal{Z}_{n-k} = \{u_{k+1}, u_{k+2}, \ldots, u_n\}$, $k = 2, 3, \ldots, n-2$, such that every edge connects a vertex in $\mathcal{U}_k$ to a vertex in $\mathcal{Z}_{n-k}$ (see Fig. 6). The sets $\mathcal{U}_k$ and $\mathcal{Z}_{n-k}$ are called *the partition sets*.

A graph join $\mathcal{G}_1 + \mathcal{G}_2$ of graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ with disjoint vertex sets $\mathcal{V}(\mathcal{G}_1)$, $\mathcal{V}(\mathcal{G}_2)$ and edge sets $\mathcal{E}(\mathcal{G}_1)$, $\mathcal{E}(\mathcal{G}_2)$ is the graph with the vertex set $\mathcal{V}(\mathcal{G}_1) \cup \mathcal{V}(\mathcal{G}_2)$ and the edge set $\mathcal{E}(\mathcal{G}_1) \cup \mathcal{E}(\mathcal{G}_2) \cup \{(u, v);\, u \in \mathcal{V}(\mathcal{G}_1), v \in \mathcal{V}(\mathcal{G}_2)\}$. It is the graph union $\mathcal{G}_1 \cup \mathcal{G}_2$ with all the edges that connect the vertices of the first graph with the vertices of the second graph.

The graph $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$ can also be written as the graph join of two empty graphs on k and $n-k$ vertices, i.e., $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}} = \mathcal{O}_k + \mathcal{O}_{n-k}$. The corresponding graph distance matrix is

$$G_{k,n-k} = \begin{bmatrix} 2(E_{k,k} - I_k) & E_{k,n-k} \\ E_{n-k,k} & 2(E_{n-k,n-k} - I_{n-k}) \end{bmatrix} \in \mathbb{R}^{n \times n},$$

where $E_{p,q} \in \mathbb{R}^{p\times q}$ and $I_p \in \mathbb{R}^{p\times p}$ are the matrix of ones and the identity matrix, respectively.

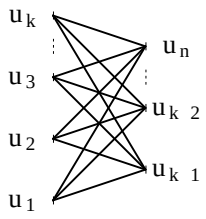


Figure 6: The graph $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$.

Theorem 3.1. *A simple connected bipartite graph $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$ on $n \geq 5$ nodes and with partition sets $\mathcal{U}_k$ and $\mathcal{Z}_{n-k}$ is a NEDM-graph.*

Proof. Since graphs $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$ and $\mathcal{G}_{\mathcal{U}_{n-k},\mathcal{Z}_k}$ are isomorphic, it is enough to see that the theorem holds true for $k = 2, 3, \ldots, \lfloor n/2 \rfloor$.

Let us analyse the eigenvalues of the graph distance matrix of $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$. A simple computation shows that $\boldsymbol{u}_{1,i} = \left[\boldsymbol{e}_1^T - \boldsymbol{e}_i^T, \boldsymbol{0}^T\right]^T$ solves the equation $G_{k,n-k}\boldsymbol{u}_{1,i} = -2\boldsymbol{u}_{1,i}$ for all $i = 2, 3, \ldots, k$, and that $\boldsymbol{u}_{2,j} = \left[\boldsymbol{0}^T, \boldsymbol{e}_1^T - \boldsymbol{e}_j^T\right]^T$, solves the equation $G_{k,n-k}\boldsymbol{u}_{2,j} = -2\boldsymbol{u}_{2,j}$ for all $j = 2, 3, \ldots, n-k$. Therefore $G_{k,n-k}$ has an eigenvalue -2 with multiplicity $n-2$.

Now let us take $\boldsymbol{u} = \left[\alpha\, \boldsymbol{e}^T, \boldsymbol{e}^T\right]^T$. The relation $G_{k,n-k}\boldsymbol{u} = \lambda\boldsymbol{u}$ yields the system of equations

$$\begin{aligned} 2(k-1)\alpha + n - k &= \lambda\,\alpha, \\ k\alpha + 2(n-k-1) &= \lambda, \end{aligned}$$

which has solutions

$$\begin{aligned} \alpha_{1,2} &= \frac{2k - n \pm \sqrt{(n-2k)^2 + k(n-k)}}{k}, \\ \lambda_{1,2} &= n - 2 \pm \sqrt{(n-2k)^2 + k(n-k)}. \end{aligned}$$

Relations $n \geq 5$ and $2 \leq k \leq \lfloor n/2 \rfloor$ imply that $\alpha_{1,2}$ and $\lambda_{1,2}$ are well-defined. Since $\lambda_1 > 0$ and

$$\lambda_1 \cdot \lambda_2 = 3(k-2)(n-2-k) + 2(n-4) > 0,$$

we conclude that $\lambda_2 > 0$. Thus, by Theorem 1.1, the graph $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$ is a NEDM-graph. □

Remark 3.2. For $k = 1$, the graph $\mathcal{G}_{\mathcal{U}_k,\mathcal{Z}_{n-k}}$ is the star graph $\mathcal{S}_n$, which is an EDM-graph.

4 Graph subdivision

Let $\mathcal{G}$ be a graph. A *subdivision* of an edge in $\mathcal{G}$ is a substitution of the edge by a path. For example, an edge of the cycle $\mathcal{C}_n$ can be subdivided into three edges, resulting in the cycle graph $\mathcal{C}_{n+2}$.

Recall the NEDM-graph $\mathcal{G}_5^{(20)}$. It contains a 4-cycle c connecting nodes 2, 3, 4 and 5 (see Fig. 7). We can construct larger NEDM-graphs by performing a subdivision of the cycle c. Let $\mathcal{G}_{5,n}^{(20)}$ be a graph on n nodes, obtained by subdividing the cycle c in the graph $\mathcal{G}_5^{(20)}$ as seen in Fig. 7. Such graphs are $\mathcal{G}_{5,6}^{(20)} = \mathcal{G}_6^{(1)}$ and $\mathcal{G}_{5,7}^{(20)} = \mathcal{G}_7^{(4)}$ (see Fig. 4 and Fig. 5).

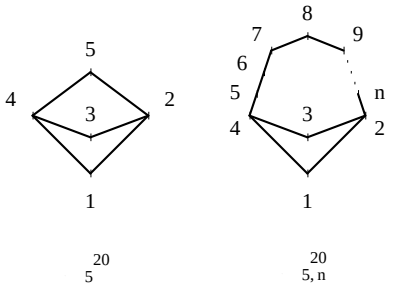


Figure 7: Construction of graphs $\mathcal{G}_{5,n}^{(20)}$.

Let $\boldsymbol{e}_i$ denote the standard basis and let C_n be the graph distance matrix of the cycle graph $\mathcal{C}_n$ (see [9]). The matrix C_n is a circulant matrix (see [3]), generated by its first row:

$$0,\ 1,\ \ldots,\ \frac{n-1}{2},\ \frac{n-1}{2},\ \frac{n-3}{2},\ \ldots,\ 1, \quad n \text{ odd},$$
$$0,\ 1,\ \ldots,\ \frac{n-2}{2},\ \frac{n}{2},\ \frac{n-2}{2},\ \ldots,\ 1, \quad n \text{ even}.$$

We will use the notation $C_n^{(i,j)}$ for the (i,j)-th element of the matrix C_n. The structure of the matrix C_n implies

$$C_n^{(1,2)} = C_n^{(2,3)} = 1, \quad C_n^{(1,3)} = 2, \quad n \geq 4, \tag{4.1}$$

and

$$C_n^{(\ell,\lfloor (n+4)/2 \rfloor)} = \begin{cases} \lfloor (n-1)/2 \rfloor, & \ell = 1, \\ \lfloor n/2 \rfloor, & \ell = 2, \\ \lfloor (n-2)/2 \rfloor, & \ell = 3, \end{cases} \qquad n \geq 3. \tag{4.2}$$

Theorem 4.1. *Graphs $\mathcal{G}_{5,n}^{(20)}$, $n \geq 5$, are NEDM-graphs.*

Proof. The graph distance matrix of the graph $\mathcal{G}_{5,n}^{(20)}$, $n \geq 5$, is

$$G_{5,n}^{(20)} = \begin{bmatrix} 0 & \boldsymbol{e}_2^T(C_{n-1} + 2I) \\ (C_{n-1} + 2I)\boldsymbol{e}_2 & C_{n-1} \end{bmatrix}.$$

By Theorem 1.1 it is enough to show that there exists $\boldsymbol{x} \in \mathbb{R}^n$, such that $\boldsymbol{x}^T\boldsymbol{e} = 0$ and $\boldsymbol{x}^T G_{5,n}^{(20)} \boldsymbol{x} > 0$.

Let us take $\boldsymbol{x} = \left[-\boldsymbol{y}^T\boldsymbol{e}, \boldsymbol{y}^T\right]^T$, with

$$\boldsymbol{y} = \begin{cases} \frac{n-1}{2}\Big(-\boldsymbol{e}_1 + \boldsymbol{e}_2 - \boldsymbol{e}_3\Big) + \boldsymbol{e}_{(n+3)/2}, & n \text{ odd}, \\ \frac{n-2}{2}\Big(-\boldsymbol{e}_1 + \frac{n-1}{n}(\boldsymbol{e}_2 - \boldsymbol{e}_3)\Big) + \frac{n-1}{n}\boldsymbol{e}_{(n+2)/2}, & n \text{ even}. \end{cases}$$

We will show that

$$\boldsymbol{x}^T G^{(20)}_{5,n}\boldsymbol{x} = \boldsymbol{y}^T C_{n-1}\boldsymbol{y} - 2(\boldsymbol{y}^T\boldsymbol{e})\Big(\boldsymbol{e}_2^T C_{n-1}\boldsymbol{y} + 2(\boldsymbol{y}^T\boldsymbol{e}_2)\Big) > 0. \tag{4.3}$$

From

$$\boldsymbol{y}^T\boldsymbol{e} = \begin{cases} \frac{3-n}{2}, & n \text{ odd}, \\ \frac{-n^2+4n-2}{2n}, & n \text{ even}, \end{cases} \quad \text{and} \quad \boldsymbol{y}^T\boldsymbol{e}_2 = \begin{cases} \frac{n-1}{2}, & n \text{ odd}, \\ \frac{(n-2)(n-1)}{2n}, & n \text{ even}, \end{cases}$$

it follows that

$$\boldsymbol{x}^T G^{(20)}_{5,n}\boldsymbol{x} = \boldsymbol{y}^T C_{n-1}\boldsymbol{y} + \begin{cases} (n-3)\Big(\boldsymbol{e}_2^T C_{n-1}\boldsymbol{y} + n - 1\Big), & n \text{ odd}, \\ \frac{n^2-4n+2}{n}\Big(\boldsymbol{e}_2^T C_{n-1}\boldsymbol{y} + \frac{(n-2)(n-1)}{n}\Big), & n \text{ even}. \end{cases} \tag{4.4}$$

Firstly, let n be odd. Terms in the relation (4.4) simplify to

$$\begin{aligned} \boldsymbol{y}^T C_{n-1}\boldsymbol{y} &= -\frac{(n-1)^2}{2}\Big(C^{(1,2)}_{n-1} - C^{(1,3)}_{n-1} + C^{(2,3)}_{n-1}\Big) - \\ &\quad - (n-1)\Big(C^{(1,(n+3)/2)}_{n-1} - C^{(2,(n+3)/2)}_{n-1} + C^{(3,(n+3)/2)}_{n-1}\Big), \\ \boldsymbol{e}_2^T C_{n-1}\boldsymbol{y} &= C^{(2,(n+3)/2)}_{n-1} - \frac{n-1}{2}\Big(C^{(1,2)}_{n-1} + C^{(2,3)}_{n-1}\Big). \end{aligned}$$

By (4.1) and (4.2),

$$\boldsymbol{y}^T C_{n-1}\boldsymbol{y} = -\frac{(n-1)(n-5)}{2}, \qquad \boldsymbol{e}_2^T C_{n-1}\boldsymbol{y} = -\frac{n-1}{2},$$

and

$$\boldsymbol{x}^T G^{(20)}_{5,n}\boldsymbol{x} = n - 1,$$

which satisfies the requirement (4.3) for all $n \geq 5$.

When n is even, the terms in the relation (4.4) simplify to

$$\begin{aligned} \boldsymbol{y}^T C_{n-1}\boldsymbol{y} &= -\frac{(n-2)(n-1)}{2n^2}\Bigg((n-2)\big(nC^{(1,2)}_{n-1} - nC^{(1,3)}_{n-1} + (n-1)C^{(2,3)}_{n-1}\big) + \\ &\quad + 2nC^{(1,(n+2)/2)}_{n-1} - 2(n-1)\big(C^{(2,(n+2)/2)}_{n-1} - C^{(3,(n+2)/2)}_{n-1}\big)\Bigg), \\ \boldsymbol{e}_2^T C_{n-1}\boldsymbol{y} &= \frac{n-1}{n}C^{(2,(n+2)/2)}_{n-1} - \frac{n-2}{2}\Big(C^{(1,2)}_{n-1} + \frac{n-1}{n}C^{(2,3)}_{n-1}\Big). \end{aligned}$$

By (4.1) and (4.2),

$$\boldsymbol{y}^T C_{n-1}\boldsymbol{y} = -\frac{(n-1)^2(n-2)(n-4)}{2n^2}, \qquad \boldsymbol{e}_2^T C_{n-1}\boldsymbol{y} = -\frac{n-2}{2}$$

and

$$\boldsymbol{x}^T G_{5,n}^{(20)} \boldsymbol{x} = \frac{n-2}{2n},$$

which satisfies the requirement (4.3) for all $n \geq 5$. □

Similarly, we can subdivide cycles of the graph $\mathcal{G}_5^{(21)}$ and produce NEDM-graphs (see Fig. 8). The graph $\mathcal{G}_5^{(21)}$ contains a 3-cycle c connecting nodes 3, 4 and 5. Let $\mathcal{G}_{5,n}^{(21)}$ be a graph on n nodes, obtained by subdividing the cycle c in the graph $\mathcal{G}_5^{(21)}$ as seen in Fig. 8. Such graphs are $\mathcal{G}_{5,5}^{(21)} = \mathcal{G}_5^{(21)}$, $\mathcal{G}_{5,6}^{(21)} = \mathcal{G}_6^{(3)}$ and $\mathcal{G}_{5,7}^{(21)} = \mathcal{G}_7^{(6)}$ (see Fig. 4 and Fig. 5).

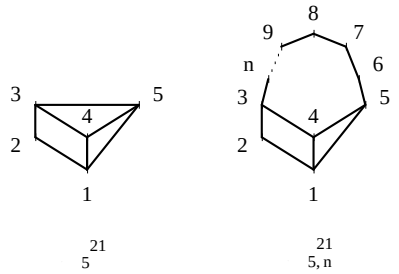


Figure 8: Graphs $\mathcal{G}_{5,n}^{(21)}$.

Theorem 4.2. *Graphs $\mathcal{G}_{5,n}^{(21)}$, $n \geq 5$, are NEDM-graphs.*

Proof. The graph distance matrix of the graph $\mathcal{G}_{5,n}^{(21)}$, $n \geq 5$, is

$$G_{5,n}^{(21)} = \begin{bmatrix} 0 & 1 & \boldsymbol{u}^T \\ 1 & 0 & \boldsymbol{v}^T \\ \boldsymbol{u} & \boldsymbol{v} & C_{n-2} \end{bmatrix},$$

where

$$\boldsymbol{u} = (C_{n-2} + I)\boldsymbol{e}_2 + \boldsymbol{e} - \boldsymbol{y} - \frac{1 + (-1)^n}{2}\boldsymbol{e}_{(n+2)/2}, \qquad \boldsymbol{v} = (C_{n-2} + I)\boldsymbol{e}_2 + \boldsymbol{y},$$

and $\boldsymbol{y} = \sum_{k=2}^{\lfloor (n+1)/2 \rfloor} \boldsymbol{e}_k$. Analogous to the proof of Theorem 4.1, we can show that for $\boldsymbol{x} = \left[\alpha, -\alpha, \boldsymbol{z}^T\right]^T$, where

$$\alpha = \begin{cases} \frac{n}{2}, & n \text{ odd}, \\ \frac{n-3}{2}, & n \text{ even}, \end{cases} \quad \text{and} \quad \boldsymbol{z} = \begin{cases} \frac{n-1}{2}\boldsymbol{e}_1 - \frac{n-3}{2}\boldsymbol{e}_2 - \boldsymbol{e}_{(n+1)/2}, & n \text{ odd}, \\ \frac{n-2}{2}\boldsymbol{e}_1 - \frac{n-4}{2}\boldsymbol{e}_2 - \boldsymbol{e}_{(n+2)/2}, & n \text{ even}, \end{cases}$$

the expression $\boldsymbol{x}^T G_{5,n}^{(21)} \boldsymbol{x} = \boldsymbol{z}^T C_{n-2} \boldsymbol{z} + 2\alpha \left(\boldsymbol{u}^T \boldsymbol{z} - \boldsymbol{v}^T \boldsymbol{z} - \alpha\right)$ is positive.

Relations

$$C_{n-2}^{(1,2)} = 1, \qquad C_{n-2}^{(1,(n+1)/2)} = C_{n-2}^{(2,(n+1)/2)} = \frac{n-3}{2},$$
$$C_{n-2}^{(1,(n+2)/2)} = \frac{n-4}{2}, \qquad C_{n-2}^{(2,(n+2)/2)} = \frac{n-2}{2},$$

imply

$$\boldsymbol{u}^T\boldsymbol{z} = \begin{cases} 2, & n \text{ odd}, \\ 1, & n \text{ even}, \end{cases} \qquad \boldsymbol{v}^T\boldsymbol{z} = \begin{cases} 3-n, & n \text{ odd}, \\ 4-n, & n \text{ even}, \end{cases}$$

and

$$\boldsymbol{z}^T C_{n-2}\boldsymbol{z} = \begin{cases} -\frac{(n-3)(n+1)}{2}, & n \text{ odd}, \\ -\frac{(n-2)(n-4)}{2}, & n \text{ even}, \end{cases}$$

which yields

$$\boldsymbol{x}^T G_{5,n}^{(21)}\boldsymbol{x} = \begin{cases} \frac{3}{2}, & n \text{ odd}, \\ \frac{1}{2}, & n \text{ even}. \end{cases}$$

Thus by Theorem 1.1, the matrix $G_{5,n}^{(21)}$ is a NEDM. □

5 Systems with no solution

When verifying whether a graph $\mathcal{G}$ with the corresponding graph distance matrix G is an EDM-graph, by Theorem 1.1 one can check if there exists a solution of the equation $G\boldsymbol{w} = \boldsymbol{e}$, such that $\boldsymbol{w}^T\boldsymbol{e} \geq 0$. For $n \geq 7$ there exist graphs, for which the equation $G\boldsymbol{w} = \boldsymbol{e}$ has no solution.

Let $\mathcal{G}_{k,n-k}$ be the graph join of a complete graph $\mathcal{K}_k$ and an empty graph $\mathcal{O}_{n-k}$, $n \geq 7$, $k = 2, 3, \ldots, n-3$, i.e.,

$$\mathcal{G}_{k,n-k} = \mathcal{K}_k + \mathcal{O}_{n-k}.$$

The graph $\mathcal{K}_k$ contains vertices $1, 2, \ldots, k$ and the graph $\mathcal{O}_{n-k}$ contains vertices $k+1, k+2, \ldots, n$. Thus the corresponding graph distance matrix is

$$G_{k,n-k} = \begin{bmatrix} E_{k,k} - I_k & E_{k,n-k} \\ E_{n-k,k} & 2(E_{n-k,n-k} - I_{n-k}) \end{bmatrix}.$$

For $n = 7$ and $k = 3$ the equation $G_{3,4}\boldsymbol{w} = \boldsymbol{e}$ has no solution since the ranks of the matrix $G_{3,4}$ and its augmented matrix $[G_{3,4}|\boldsymbol{e}]$ are different, $\operatorname{rank}(G_{3,4}) = 6$ and $\operatorname{rank}([G_{3,4}|\boldsymbol{e}]) = 7$. The same holds true if $n = 7$ and $k = 4$. Thus by Theorem 1.1 matrices $G_{3,4}$ and $G_{4,3}$ are not EDMs. On the other hand, for $n = 8$ the equation $G_{k,8-k}\boldsymbol{w} = \boldsymbol{e}$ has a solution for all $k \in \{3, 4, 5\}$. In general, the matrix $G_{k,n-k}$ is a NEDM.

Theorem 5.1. *The graph $\mathcal{K}_k + \mathcal{O}_{n-k}$, $n \geq 7$, $k = 2, 3, \ldots, n-3$, is a NEDM-graph.*

Proof. Let $G_{k,n-k}$ be the graph distance matrix of the graph $\mathcal{K}_k + \mathcal{O}_{n-k}$, $n \geq 7$, $k = 2, 3, \ldots, n-3$. For $k = 2$ we take

$$\boldsymbol{w} = \frac{1}{2}\big[4-n, 4-n, 1, 1, \ldots, 1\big]^T.$$

We can verify that $G_{2,n-2}\boldsymbol{w} = \boldsymbol{e}$ and $\boldsymbol{w}^T\boldsymbol{e} = (6-n)/2 < 0$. Thus by Theorem 1.1 the matrix $G_{2,n-2}$ is a NEDM.

Now let $k = 3, 4, \ldots, n-3$. For $n = 7$ the proof has already been done above. For $n \geq 8$ let $\boldsymbol{u} = \big[\alpha\,\boldsymbol{e}^T, \boldsymbol{e}^T\big]^T$, where vectors $\boldsymbol{e}$ are of sizes k and $n-k$, respectively. The relation $G_{k,n-k}\boldsymbol{u} = \lambda\boldsymbol{u}$ yields the system of equations

$$\begin{aligned} \alpha(k-1) + n - k &= \lambda\alpha, \\ \alpha k + 2(n-k-1) &= \lambda, \end{aligned}$$

with solutions

$$\alpha_{1,2} = \frac{1}{2k}\left(3k - 2n + 1 \pm \sqrt{4(n-k)(n-k-1) + (k+1)^2}\right),$$
$$\lambda_{1,2} = \frac{1}{2}\left(2n - k - 3 \pm \sqrt{4(n-k)(n-k-1) + (k+1)^2}\right).$$

Relations $n \geq 8$ and $3 \leq k \leq n-3$ imply that $\alpha_{1,2}$ and $\lambda_{1,2}$ are well-defined. Since $\lambda_1 > 0$ and

$$\lambda_1 \cdot \lambda_2 = (n-3-k)(k-3) + n - 7 > 0,$$

we conclude that $\lambda_2 > 0$. Thus, by Theorem 1.1, graph $\mathcal{K}_k + \mathcal{O}_{n-k}$ is a NEDM-graph. □

Remark 5.2. For $k = 1$ and $k = n-1$, the graphs $\mathcal{K}_k + \mathcal{O}_{n-k}$ are the star graph $\mathcal{S}_n$ and the complete graph $\mathcal{K}_n$, respectively, which are EDM-graphs.

Remark 5.3. For $k = n-2$, the graph $\mathcal{K}_{n-2} + \mathcal{O}_2$ is an EDM-graph. The graph distance matrix $G_{n-2,2}$ has eigenpairs

$$\left(-2, \left[\mathbf{0}^T, \boldsymbol{e}_1^T - \boldsymbol{e}_2^T\right]^T\right), \quad \left(-1, \left[\boldsymbol{e}_1^T - \boldsymbol{e}_i^T, \mathbf{0}^T\right]^T\right), \quad i = 2, 3, \ldots, n-2,$$

and

$$\left(\lambda_{1,2}, \left[\alpha_{1,2}\,\boldsymbol{e}^T, \boldsymbol{e}^T\right]^T\right)$$

with

$$\alpha_{1,2} = \frac{n - 5 \pm \sqrt{n^2 - 2n + 9}}{2(n-2)} \quad \text{and} \quad \lambda_{1,2} = \frac{n - 1 \pm \sqrt{n^2 - 2n + 9}}{2}.$$

The eigenvalue λ_1 is obviously positive. From $\lambda_1 \cdot \lambda_2 = -2$ it follows that $\lambda_2 < 0$. One can easily verify that $\boldsymbol{w} = (1/2)\left[\mathbf{0}^T, \boldsymbol{e}^T\right]^T$ solves the equation $G_{n-2,2}\boldsymbol{w} = \boldsymbol{e}$. Since $\boldsymbol{w}^T\boldsymbol{e} = 1$, Theorem 1.1 implies that $G_{n-2,2}$ is EDM.

6 Conclusion

In Section 4 we studied subdivisions of graphs. Not all graph subdivisions result in NEDM-graphs. Consider subdividing graph $\mathcal{G}_5^{(20)}$ as in Fig. 9 and denoting it by $\mathcal{H}$. The corresponding graph distance matrix

$$H = \begin{bmatrix} 0 & 1 & 2 & 2 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 & 2 & 2 \\ 2 & 2 & 1 & 0 & 1 & 2 & 1 \\ 3 & 2 & 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 2 & 2 & 1 & 0 & 3 \\ 1 & 2 & 2 & 1 & 2 & 3 & 0 \end{bmatrix}$$

has eigenvalues $\sigma_H \doteq \{10.4, 0, -0.2, -0.6, -2.2, -3.4, -4\}$, which were calculated numerically. Exact eigenvalues could be obtained using Cardano's formula. One can easily verify that vector $\boldsymbol{w}_H = [1/2,\ -1/2,\ 1/2,\ -1/2,\ 1/2,\ 0,\ 0]^T$ solves the equation $H\boldsymbol{w}_H = \boldsymbol{e}$. Since $\boldsymbol{w}_H^T\boldsymbol{e} = 1/2$, by Theorem 1.1 the graph $\mathcal{H}$ is an EDM-graph.

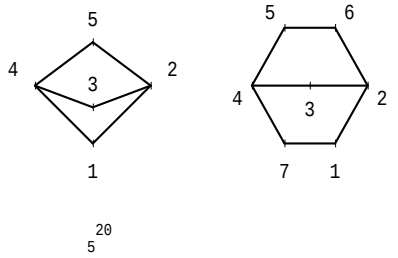


Figure 9: A subdivision of the graph $\mathcal{G}_5^{(20)}$.

References

[1] R. Balaji and R. B. Bapat, Block distance matrices, *Electron. J. Linear Algebra* **16** (2007), 435–443.

[2] J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*, Meboo, 2008.

[3] G. H. Golub and C.F. van Loan, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, third ed., Johns Hopkins University Press, 1996.

[4] J. C. Gower, Euclidean distance geometry, *Math. Sci.* **7** (1982), 1–14.

[5] T. L. Hayden, R. Reams and J. Wells, Methods for constructing distance matrices and the inverse eigenvalue problem, *Linear Algebra Appl.* **295** (1999), 97–112.

[6] G. Jaklič, J. Modic, On properties of cell matrices, *Appl. Math. Comput.* **216** (2010), 2016–2023.

[7] G. Jaklič, J. Modic, A note on "Methods for constructing distance matrices and the inverse eigenvalue problem", *Linear Algebra Appl.* **437** (2012), 2781–2792.

[8] G. Jaklič and J. Modic, Inverse eigenvalue problem for Euclidean distance matrices of size 3, *Bull. Aust. Math. Soc.* **87** (2013), 82–93.

[9] G. Jaklič and J. Modic, On Euclidean distance matrices of graphs, *Electron. J. Linear Algebra* **26** (2013), 574–589.

[10] G. Jaklič, J. Modic, Euclidean graph distance matrices of generalizations of the star graph, *Appl. Math. Comput.* **230** (2014), 650–663.

[11] G. Jaklič, J. Modic, Cartesian products of EDM-graphs, submitted.

[12] B. D. McKay, Practical Graph Isomorphism, *Congr. Numer.* **30** (1981), 45–87.

[13] I. J. Schoenberg, Metric spaces and positive definite functions, *Trans. Amer. Math. Soc.* **44** (1938), 522–536.

[14] G. Young, A. Householder, Discussion of a set of points in terms of their mutual distances, *Psychometrika* **3** (1938), 19–22.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 165–186

Sparse line deletion constructions for symmetric 4-configurations

Leah Wrenn Berman
University of Alaska Fairbanks,
513 Ambler Lane, Fairbanks, AK 99775, USA

William H. Mitchell
University of Wisconsin - Madison,
480 Lincoln Dr., Madison, WI 53706, USA

Received 15 April 2013, accepted 24 May 2014, published online 21 November 2014

Abstract

A *4-configuration* is a collection of points and lines in the Euclidean plane such that each point lies on four lines and each line passes through four points. In this paper we introduce a new family of these objects. Our construction generalizes a 2010 result of Berman and Grünbaum in which suitable 4-configurations from the well-understood *celestial* family are altered to yield new configurations with reduced geometric symmetry groups. The construction introduced in 2010 removes every other line of a symmetry class from the celestial configuration; here we we give conditions under which every p-th line can be removed, for $p \in \{2, 3, 4, \cdots\}$. The geometric symmetry groups of the new configurations we obtain are of correspondingly smaller index as subgroups of the symmetry group of the underlying celestial configuration. These *sparse* constructions can also be repeated and combined to yield a rich variety of previously unknown 4-configurations. In particular, we can begin with a configuration with very high geometric symmetry—the dihedral symmetry of an m-gon for m quite large—and produce a configuration whose only geometric symmetry is 180° rotation.

Keywords: Configurations.

Math. Subj. Class.: 05B30, 51E30, 05C60

E-mail addresses: lwberman@alaska.edu (Leah Wrenn Berman), whmitchell@math.wisc.edu (William H. Mitchell)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

1 Introduction

An n-configuration is a set of n points and n lines with the property that each point lies on n lines and each line passes through n points. Configurations can be investigated as geometric objects or more generally as combinatorial objects where the lines are abstract sets of points. In this work we take the geometric perspective and consider points and lines in the Euclidean plane. Although such geometric objects were studied in the 19th century and several theorems on 3-configurations were proved, no illustration of a geometric 4-configuration appeared in print until much more recently, in [4]. Since then many more examples have been introduced. In this paper we give a technique that produces a large new class of 4-configurations, including 4-configurations with very few symmetries. We emphasize that by a *symmetry* of a configuration we mean an isometry of the plane which maps the configuration to itself, as opposed to the more general notion of combinatorial symmetry. The collection of symmetries of a configuration, or its *symmetry group*, partitions the points and lines into orbits, called the *symmetry classes of points*, and the *symmetry classes of lines*.

One frequently studied class of 4-configurations is the *celestial* family. Its members have the property that every point lies on exactly two lines from each of two symmetry classes of lines, and every line is incident with two points from each of two symmetry classes of points. Figure 2 gives an example of a celestial 4-configuration. The first published 4-configuration, in [4], was of this class, and more examples appeared in [6]. The first discussion of celestial configurations as a family appeared in a paper called *Polycyclic Configurations* by Marko Boben and Tomaž Pisanski [2], where they were investigated as a particular class of *polycyclic* 4-configurations. Branko Grünbaum's 2009 monograph *Configurations of Points and Lines* [3] gives a detailed analysis of the construction method and theory for celestial 4-configurations. In that reference Grünbaum refers to them as k*-astral* 4-configurations. However, he also uses the term "k-astral" to describe configurations which have k symmetry classes of points and k symmetry classes of lines; while celestial 4-configurations have this property, there are many other 4-configurations with this property that are not celestial. We reserve the term "k-astral" for the more general class of configurations with k symmetry classes of points and lines, and use the term "celestial" to refer to 4-configurations with the particular symmetry restrictions described above.

In [1], one author (LWB) developed two procedures which modify suitable celestial configurations to yield new 4-configurations. In the first of these, every other line from a particular symmetry class is deleted and then an equal number of new lines that pass through the center of the configuration—diameters—are added in such a way that the resulting structure is a (noncelestial) 4-configuration. The number of points and lines remains unchanged at the end of the construction since one diameter is added for every line removed. In the second procedure, particular elements of certain symmetry classes of points and of lines are both deleted and then diameters are added in such a way that every point is incident with four lines and every line is incident with four points, with a net loss of both points and lines.

In this paper we generalize the first of those procedures. We refer to this generalized procedure as *sparse line deletion* or p*-sparse line deletion* because in general it is possible to delete a smaller number of lines than in the old construction. The new configurations obtained in this way differ qualitatively from those introduced in [1] in that they exhibit a

wider variety of symmetry groups compared to the symmetries of the underlying celestial configurations. In particular, despite beginning with a configuration with a high degree of geometric symmetry, we can obtain configurations of quite low symmetry by repeating the sparse line deletion construction, in contrast to the previous construction. Figure 1 depicts three examples of these new objects; beginning with celestial configurations with d_{18}, d_{12} and d_{16} symmetry, we develop configurations with d_6, d_4 and d_4 symmetry, respectively.

The paper is organized as follows. In Section 2 we review the theory and notation for celestial configurations. We correct a minor notational ambiguity from [1] and give new results describing the incidences of the diameters in a series of lemmas. In Section 3 we describe the p-sparse line deletion construction. In Section 4 we show how the construction may be carried out several times simultaneously to yield a rich variety of new configurations. In Section 5 we give examples of configurations obtained by a related, but poorly understood technique applicable in the case where each symmetry class contains an odd number of objects. We close by mentioning several questions that deserve further study. All figures in this paper were generated using the free software Matplotlib [5].

2 Celestial configurations

A *celestial* configuration is a 4-configuration with a high degree of geometric symmetry; specifically, such a configuration has the property that every point is incident with exactly two lines from each of two symmetry classes, and every line is incident with exactly two points from each of two symmetry classes. If a celestial configuration has k symmetry classes of points and of lines, we refer to it as a k-celestial configuration. Each k-celestial configuration consists of a composite number mk of points and mk lines for some m. The points are the vertices of k concentric regular m-gons, and the configuration exhibits m-fold dihedral symmetry (that is, d_m symmetry).

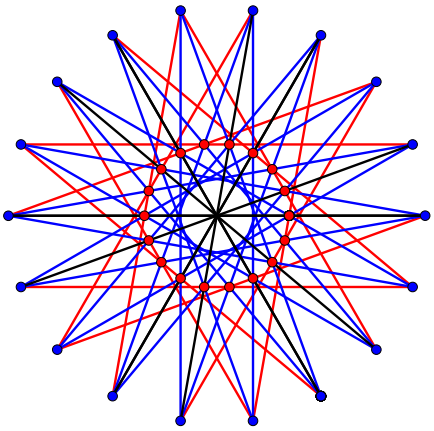
An example of a 3-celestial configuration is shown in Figure 2. In that figure, the three symmetry classes of points are distinguished by color (red, green and blue), and the three symmetry classes of lines are distinguished in the same way (also red, green, and blue). Each green line contains two red points and two green points (and similarly for the other two classes of lines), and each blue point lies on two red and two blue lines (and similarly for the other two classes of points).

Celestial configurations will serve as the building blocks of all of the new 4-configurations described in this paper. One useful feature of celestial configurations is the fact that every celestial configuration may be described by a *configuration symbol*

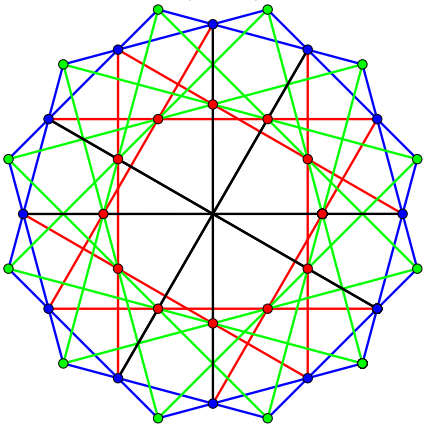
$$m\#(s_1, t_1; s_2, t_2; \cdots ; s_k, t_k)$$

which encodes a geometric construction algorithm. The integers s_i, t_i, m in the configuration symbol must satisfy several constraints for the construction to yield a 4-configuration; in this case we say the symbol is *valid*. The constraints are: $m \geq 7$, $k \geq 2$, $1 \leq s_i, t_i < \frac{m}{2}$ for all i, and

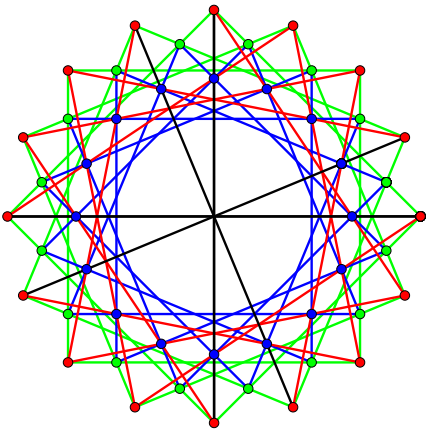
1. (order condition) adjacent entries in the sequence $(s_1, t_1, s_2, \cdots , t_k)$ (taken cyclically) are distinct;



(a) $18\#(1^{3*}, 7; 8, 6); D*$



(b) $12\#(2^{3*}, 4; 1, 2; 4, 1); D*$,



(c) $16\#(5^{4*}, 3; 4, 5; 3, 4); D*$,

Figure 1: Three new 4-configurations. (a), the 3-sparse line deletion $18\#(1^{3*}, 7; 8, 6); D*$, with d_6 symmetry. (b) the 3-sparse line deletion $12\#(2^{3*}, 4; 1, 2; 4, 1); D*$, with d_4 symmetry. (c) the 4-sparse line deletion $16\#(5^{4*}, 3; 4, 5; 3, 4); D*$, with d_4 symmetry.

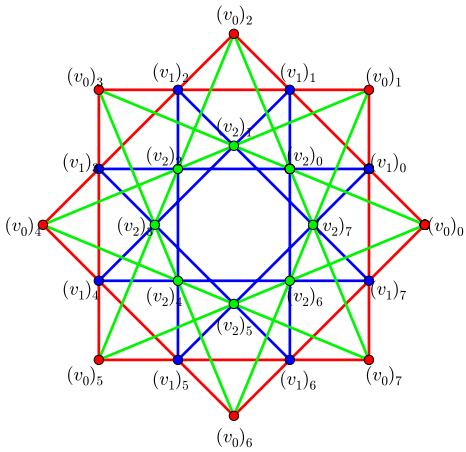


Figure 2: The 3-celestial configuration $8\#(2,1;3,2;1,3)$. The labeling updates Figure 2a in [1]. Throughout the paper we use red for v_0 and L_0, blue for v_1 and L_1, and green for v_2 and L_2.

2. (even condition) $\sum_{i=1}^{k} s_i + t_i$ is even; and
3. (cosine condition) $\prod_{i=1}^{k} \cos\left(\frac{s_i\pi}{m}\right) = \prod_{i=1}^{k} \cos\left(\frac{t_i\pi}{m}\right)$;
4. (substring condition) the symbol $m\#(L)$ is invalid whenever L is a proper contiguous substring of $(s_1, t_1; \cdots; s_k, t_k)$.

As an example illustrating contiguity, $(3,2;1,4)$ and $(4,7;5,3)$ are contiguous substrings of $(5,3;2,1;4,7)$ but $(5,2;4,7)$ is not.

The cosine condition is satisfied automatically if the sets $S = \{s_1, \ldots, s_k\}$ and $T = \{t_1, \ldots, t_k\}$ are equal, in which case the configuration is called *trivial*. All the configurations in this paper, with the exception of those in Figures 1a and 3, are formed from trivial celestial configurations. More information on these conditions can be found in [3, Chapter 3].

We now turn to the construction algorithm encoded by the symbol.

2.1 Geometric construction algorithm (celestial configurations)

We write $P \vee Q$ for the line passing through points P and Q and $L \wedge M$ for the intersection of lines L and M. In the symbols $(v_i)_j$ and $(L_i)_j$, the second index j is to be interpreted modulo m. The construction algorithm to produce a celestial configuration given a valid configuration symbol is as follows.

1. Begin with the vertices of a regular m-gon; e.g. take $(v_0)_i = \left(\cos\left(\frac{2\pi i}{m}\right), \sin\left(\frac{2\pi i}{m}\right)\right)$, for $0 \le i < m$. Let v_0, written without a second subscript, denote the collection of these points.

2. Given points v_j, define $(L_j)_i = (v_j)_i \vee (v_j)_{i+s_{j+1}}$, for $0 \leq i < m$. We denote by L_j the collection of these lines.

3. Given lines L_j, define $(v_{j+1})_i = (L_j)_i \wedge (L_j)_{i-t_j}$, for $0 \leq i < m$, and let v_{j+1} denote the collection of these points.

4. Repeat the previous two steps until the line class L_{k-1} is obtained using the parameter s_k. Stop before constructing the points v_k; if the symbol is valid, the set of points v_k that would be constructed in the next step would coincide setwise with the points v_0.

For future reference we list all of the incidences explicitly in Table 1.

Table 1: Incidences between members of point and line classes in the celestial configuration $m\#(s_1, t_1; \ldots; s_k, t_k)$. The quantity δ is defined by $\delta = \frac{1}{2}\sum_{i=1}^{k} s_i - t_i$.

Object	Incidences			
$(L_j)_i, 0 \leq j < k-1$	$(v_j)_i$	$(v_j)_{i+s_{j+1}}$	$(v_{j+1})_i$,	$(v_{j+1})_{i+t_{j+1}}$
$(L_{k-1})_i$	$(v_{k-1})_i$	$(v_{k-1})_{i+s_{j+1}}$	$(v_0)_{i+\delta}$	$(v_0)_{i+\delta+t_k}$
$(v_0)_i$	$(L_0)_i$	$(L_0)_{i-s_1}$	$(L_{k-1})_{i-\delta}$	$(L_{k-1})_{i-\delta-t_k}$
$(v_j)_i, 0 < j \leq k-1$	$(L_j)_i$	$(L_j)_{i-s_{j+1}}$	$(L_{j-1})_i$	$(L_{j-1})_{i-t_j}$.

2.2 Lines through the origin

The vertices in a given point class v_j of a celestial configuration form a regular m-gon. For each integer ℓ it follows that the angle $\angle(v_j)_0\mathcal{O}(v_j)_\ell$ is an integer multiple of $2\pi/m$ (that is, an *even* multiple of π/m). A slightly weaker statement holds for points in different symmetry classes: for $i \neq j$, it is still true that the angle $\angle(v_i)_0\mathcal{O}(v_j)_\ell$ is an integer multiple of π/m. In the constructions we consider we will add lines through the center of the configuration (although the center is *not* one of the points of the configuration). We denote by D_j the line through the origin that makes an angle of $j\frac{\pi}{m}$ radians with the line $\mathcal{O} \vee (v_0)_0$ (conventionally a horizontal line) for $j = 0, 1, \cdots, m-1$. For $j \geq m$ or $j < 0$ we reduce modulo m so that $D_m = D_0 = \mathcal{O} \vee (v_0)_0$. This notation is more flexible than the concept of diametral type introduced in [1] and does not require m to be even. We refer to all of the D_j as *diameters*.

With this notation we restate some useful facts on celestial configurations.

1. Suppose that m is even and $(v_j)_i$ lies on D_a. Then $(v_j)_{i+\frac{m}{2}}$ also lies on D_a so that D_a passes through two points of v_j. However, if q is odd then D_{a+q} passes through no points of v_j. Hence if m is even, each diameter passes through either zero or two points from each symmetry class.

2. Suppose that m is odd. Then each diameter is incident with exactly one point of each symmetry class.

3. Let $0 \leq j < k-1$. If $(v_j)_0$ lies on D_a then $(v_{j+1})_0$ lies on $D_{a+s_{j+1}-t_{j+1}}$.

By combining (1) and (3) we see that if m and $(s_1 + t_1)$ are even, then the even-numbered

diameters pass through two points from each of v_0 and v_1 while the odd-numbered diameters miss all of the points in v_0 and v_1.

We now give three lemmas providing specific information on the incidences of the diameters. This information is conveniently expressed in terms of the constants $\{\beta_j\}$ defined by $\beta_0 = 0$ and

$$\beta_j = \sum_{q=1}^{j} s_q - t_q, \quad j = 1, \cdots, k-1.$$

Lemma 2.1. *For all i and j, the point $(v_j)_i$ lies on the diameter D_{β_j+2i}.*

Proof. By definition, $(v_0)_0$ lies on D_0. Applying (3) repeatedly we see that $(v_j)_0$ lies on diameter D_{β_j}. It follows that $(v_j)_i$ lies on D_{β_j+2i}. □

Lemma 2.2. *For $0 \le \ell < m$, and $0 \le j < k$, the diameter D_ℓ passes through the following points of v_j:*

$$\begin{cases} \text{none} & m \text{ even, } \beta_j - \ell \text{ odd;} \\ (v_j)_{\frac{\ell-\beta_j}{2}}, (v_j)_{\frac{m+\ell-\beta_j}{2}} & m \text{ even, } \beta_j - \ell \text{ even;} \\ (v_j)_{\frac{m+\ell-\beta_j}{2}} & m \text{ odd, } \beta_j - \ell \text{ odd;} \\ (v_j)_{\frac{\ell-\beta_j}{2}} & m \text{ odd, } \beta_j - \ell \text{ even.} \end{cases}$$

Proof. Lemma 2.1 states that for each i the point $(v_j)_i$ lies on D_{β_j+2i}, so it suffices to solve the congruence $\ell \equiv \beta_j + 2i \ (m)$ for i. Equivalently we solve $2i \equiv \ell - \beta_j \ (m)$. If m is odd, this equation has one solution because 2 is a generator of the cyclic group $\mathbb{Z}/m\mathbb{Z}$. This solution depends on the parity of $\ell - \beta_j$ as indicated. If m is even then $2i$ and $2i - m$ are always even, so there is no solution if $\ell - \beta_j$ is odd. If $\ell - \beta_j$ is even then both $\frac{\ell-\beta_j}{2}$ and $\frac{m+\ell-\beta_j}{2}$ are solutions, as indicated. □

Lemma 2.3. *If $0 \le j, \ell < k$ and $0 \le i < m$, the points of v_ℓ sharing a diameter with $(v_j)_i$ are*

$$\begin{cases} \text{none} & m \text{ even, } \beta_\ell - \beta_j \text{ odd;} \\ (v_\ell)_{i+\frac{\beta_j-\beta_\ell}{2}}, (v_\ell)_{i+\frac{m+\beta_j-\beta_\ell}{2}} & m \text{ even, } \beta_\ell - \beta_j \text{ even;} \\ (v_\ell)_{i+\frac{m+\beta_j-\beta_\ell}{2}} & m \text{ odd, } \beta_\ell - \beta_j \text{ odd;} \\ (v_\ell)_{i+\frac{\beta_j-\beta_\ell}{2}} & m \text{ odd, } \beta_\ell - \beta_j \text{ even.} \end{cases}$$

Proof. Lemma 2.1 implies that $(v_j)_i$ lies on D_{β_j+2i}. Lemma 2.2 then states which points of v_ℓ lie on this diameter. Writing $\tilde{\ell} = \beta_j + 2i$ and $\tilde{j} = \ell$ to match the notation of Lemma 2.2, we find that the following points of $v_{\tilde{j}}$ lie on $D_{\tilde{\ell}}$:

$$\begin{cases} \text{none} & m \text{ even, } \beta_{\tilde{j}} - \tilde{\ell} \text{ odd;} \\ (v_{\tilde{j}})_{\frac{\tilde{\ell}-\beta_{\tilde{j}}}{2}}, (v_{\tilde{j}})_{\frac{m+\tilde{\ell}-\beta_{\tilde{j}}}{2}} & m \text{ even, } \beta_{\tilde{j}} - \tilde{\ell} \text{ even;} \\ (v_{\tilde{j}})_{\frac{m+\tilde{\ell}-\beta_{\tilde{j}}}{2}} & m \text{ odd, } \beta_{\tilde{j}} - \tilde{\ell} \text{ odd;} \\ (v_{\tilde{j}})_{\frac{\tilde{\ell}-\beta_{\tilde{j}}}{2}} & m \text{ odd, } \beta_{\tilde{j}} - \tilde{\ell} \text{ even.} \end{cases}$$

In other words, the following points of v_ℓ lie on D_{β_j+2i}:

$$\begin{cases} \text{none} & m \text{ even, } \beta_\ell - \beta_j \text{ odd;} \\ (v_\ell)_{\frac{\beta_j+2i-\beta_\ell}{2}}, (v_\ell)_{\frac{m+\beta_j+2i-\beta_\ell}{2}} & m \text{ even, } \beta_\ell - \beta_j \text{ even;} \\ (v_\ell)_{\frac{m+\beta_j+2i-\beta_\ell}{2}} & m \text{ odd, } \beta_\ell - \beta_j \text{ odd;} \\ (v_\ell)_{\frac{\beta_j+2i-\beta_\ell}{2}} & m \text{ odd, } \beta_\ell - \beta_j \text{ even.} \end{cases}$$

The result follows. □

3 Sparse line deletion

Consider the celestial configuration $18\#(5,1;4,6)$, illustrated in Figure 3a. Suppose we delete the lines $(L_0)_k$, $k = 0, 3, 6, 9, 12, 15$; the resulting structure is not a configuration because some of the points, shown larger in Figure 3b, have lost an incidence. We say that these points have been *affected* by the line deletion. Note that the affected points of v_0 lie on the same diameters as the affected points of v_1, and each diameter that has any affected point incident with it in fact is incident with two points from each of the two symmetry classes. In addition, each affected point is missing precisely one line. Therefore, if we add the six diameters $\{D_0, D_4, D_6, D_{10}, D_{12}, D_{16}\}$, we obtain the 4-configuration depicted in Figure 3c. This is an example of the 3-*sparse line deletion* construction.

We call this construction *sparse* in comparison with the construction given in [1], because we remove only one-third of the lines L_0 instead of one-half. Figure 4 shows the result of the construction technique described in [1], which was called *odd deletion* in that work and which corresponds to 2-sparse deletion in the terminology of the present work, beginning from the same celestial configuration $18\#(5,1;4,6)$. The example of Figure 4 also serves to correct an error from [1], where it was claimed incorrectly that the construction would work only for k-celestial configurations with $k \geq 3$.

The following theorem gives necessary conditions for the procedure described above to succeed, given parameters m, s_i, t_i of the celestial configuration and a sparsity p. The proof shows that the affected points all lie on a particular set of diameters, and that all points on these diameters are affected. The case $p = 2$ was proven in [1].

Theorem 3.1 (p-Sparse Line Deletion)**.** *Let $p \geq 2$, and let $\mathcal{C}$ be a celestial 4-configuration with symbol $m\#(s_1,t_1;s_2,t_2;\cdots;s_k,t_k)$ satisfying the following conditions:*

(i) *p does not divide s_1.*

(ii) *m is even, and either $\frac{m}{2} \equiv 0 \pmod p$ or $\frac{m}{2} \equiv s_1 \pmod p$.*

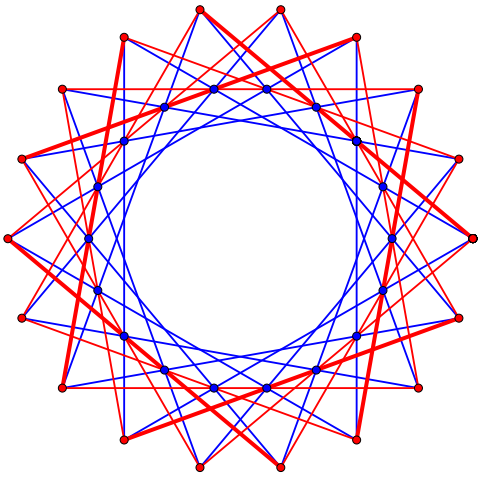
(iii) *The points lying on even-numbered diameters are precisely those of v_0 and v_1, i.e.:*

If $k = 2$, then $s_1 + t_1$ and $s_2 + t_2$ are both even.

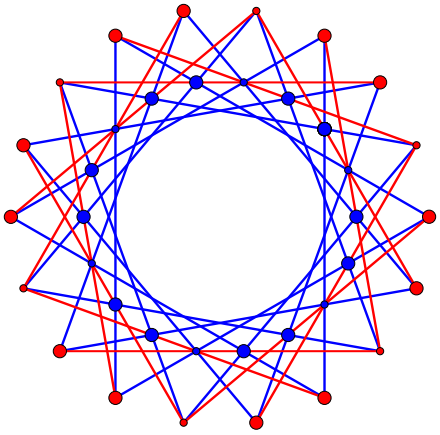
If $k \geq 3$, then $s_i + t_i$ is odd for $i = 2$, $i = k$, and even otherwise.

(iv) *The following sets coincide when reduced modulo p:*

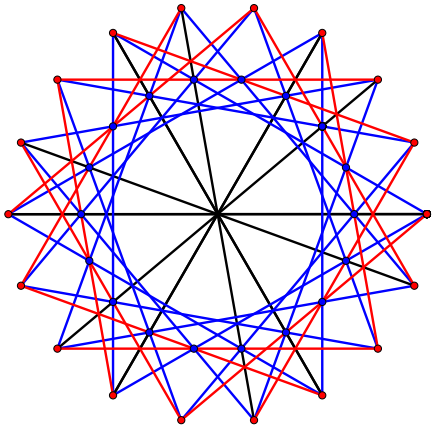
$$\{0, s_1\} = \left\{ \frac{s_1+t_1}{2}, \frac{s_1-t_1}{2} \right\}.$$



(a) $18\#(5,1;4,6)$



(b) $18\#(5^{3*},1;4,6)$



(c) $18\#(5^{3*},1;4,6);D*$

Figure 3: The 3-sparse line deletion construction. (a) The celestial configuration $18\#(5,1;4,6)$ with $(L_0)_n$ drawn thicker for $n \equiv 0 \bmod 3$. (b) Lines $(L_0)_n$ for $n \equiv 0 \bmod 3$ have been deleted and the points affected by the deletion are drawn larger. This structure is denoted $18\#(5^{3*},1;4,6)$ and is not a 4-configuration; the notation 5^{3*} is explained in Theorem 3.1. (c): The 4-configuration $18\#(5^{3*},1;4,6);D*$ obtained from (b) by adding diameters.

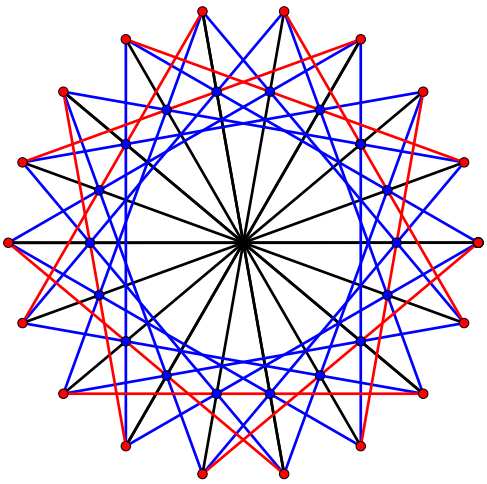


Figure 4: The 2-sparse line deletion configuration $18\#(5^{2*}, 1; 4, 6); D*$. All of the diameters have been added, so the other constructions considered in this paper are *sparse* in comparison. In the notation of [1] this would have been denoted $18\#(5^*, 1; 4, 6); D$.

Remove from $\mathcal{C}$ the lines $(L_0)_{np}$, $0 \le n < \frac{m}{p}$. Add the diameters passing through the affected points of v_0, i.e. D_{2np}, $D_{2(np+s_1)}$, for $0 \le n < \lceil \frac{m}{2p} \rceil$. Then the resulting structure $\mathcal{C}'$ is again a 4-configuration, which we denote as $m\#(s_1^{p*}, t_1; \ldots; s_k, t_k); D*$.

Proof. We verify that each object in the new structure has exactly four incidences.

Each line $(L_j)_i$ of $\mathcal{C}$ that is not deleted still has exactly four incidences in $\mathcal{C}'$ since no points are added or deleted in this construction.

The added diameters also pass through exactly four points. To see this, note that by condition (iii) the classes v_0 and v_1 and no others lie on even-numbered diameters. Condition (ii) implies that m is even, so each even-numbered diameter passes through two points from each of v_0 and v_1 and no others.

Consider now the points $(v_j)_i$ with $j > 1$. By condition (iii) these lie on odd-numbered diameters. They therefore do not gain any incidence from the added diameters, and they do not lose any incidence either since the deleted lines are chosen from L_0 and these lines are incident only with points of v_0 and v_1 (again by condition (iii)).

It remains only to show that each point of v_0 and v_1 lies on exactly four lines after diameters are added.

We begin with the points v_0. A point $(v_0)_i$ lies on two lines of L_0, namely $(L_0)_i$ and $(L_0)_{i-s_1}$. Because $s_1 \not\equiv 0 \pmod{p}$ by condition (i), at most one of these lines is deleted. Because we add a diameter if and only if it passes through an affected point of v_0, the affected points regain their lost incidence and have exactly four incidences. Hence all points $(v_0)_i$ have at least four incidences in $\mathcal{C}'$.

We must still check that none of them have five, i.e., that no unaffected point of v_0 lies

directly across the origin from an affected point of v_0 on the same diameter. We therefore suppose that $(v_0)_i$ is affected by the deletion, i.e. $i \equiv 0 \pmod p$ or $i \equiv s_1 \pmod p$, and we show that its reflection $(v_0)_{i+\frac{m}{2}}$ across the origin is also affected. To do so, we must show that either $i + \frac{m}{2} \equiv 0 \pmod p$ or $i + \frac{m}{2} \equiv s_1 \pmod p$. We consider the two cases of condition (ii). In the first case, where $\frac{m}{2} \equiv 0 \pmod p$ the desired congruence is immediate. In the second case we have $\frac{m}{2} \equiv s_1 \pmod p$, so $2s_1 \equiv m \equiv 0 \pmod p$. However, $i + \frac{m}{2}$ is congruent to $i + s_1$, since we are in the case where $\frac{m}{2} \equiv s_1 \pmod p$, and this is now congruent to either s_1 or $2s_1 \equiv 0 \pmod p$, according to whether $i \equiv 0$ or $i \equiv s_1$. Hence all of the points v_0 have exactly four incidences in the new structure $\mathcal{C}'$.

Now consider the points of v_1. We begin by showing that a point $(v_1)_i$ can lose at most one incidence when the lines $(L_0)_{np}$, $0 \le n \le \lceil \frac{m}{2p} \rceil$, are deleted. Indeed, $(v_1)_i$ lies on only two lines from the first line class, namely $(L_0)_i$ and $(L_0)_{i-t_1}$. Thus, we need to show that $t_1 \not\equiv 0 \pmod p$. This follows from condition (iv). If t_1 were congruent to $0 \pmod p$, we would have $\frac{s_1-t_1}{2} \equiv \frac{s_1-t_1}{2} + t_1 \equiv \frac{s_1+t_1}{2} \pmod p$. These numbers cannot be congruent, however, since one is congruent to 0 and the other to s_1. This shows that each point $(v_1)_i$ will lose either zero or one incidence when the lines $(L_0)_{np}$ are removed. It follows that each line deletion affects two points of v_1 as well as two points of v_0, so the same number of points are affected in each of these point classes.

Finally, we argue that the affected points of class v_1 are precisely those that lie on the added diameters. Because v_1 contains the same number of affected points as v_0, it suffices to show that each affected point of class v_1 lies on one of the diameters added previously. A counting argument then guarantees that no unaffected point lies on an added diameter. Since $\beta_1 = s_1 - t_1$ is even by condition (iii) and m is even by condition (ii), Lemma 2.3 implies that each point $(v_1)_i$ shares a diameter with $(v_0)_{i+\frac{1}{2}(s_1-t_1)}$. The affected points of v_1 are those lying on $(L_0)_q$ where $q \equiv 0 \pmod p$, namely $(v_1)_q$ and $(v_1)_{q+t_1}$. It therefore suffices to show that

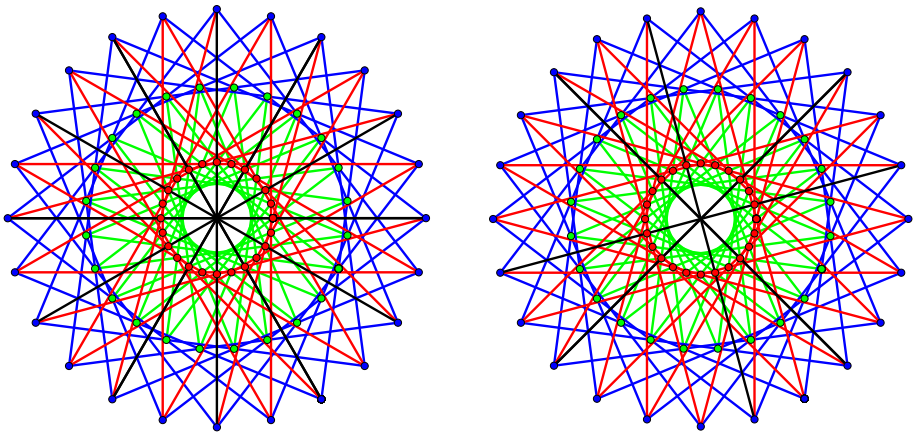
$$\text{if } i \equiv 0 \text{ or } t_1 \pmod p, \text{ then } i + \frac{1}{2}(s_1 - t_1) \equiv 0 \text{ or } s_1 \pmod p,$$

since 0 and s_1 are the remainders modulo p of the indices of affected points in v_0. But this is equivalent to condition (iv). Hence the affected points of v_1 lie on added diameters in $\mathcal{C}'$. This completes the proof. □

3.1 Notation

The notation of [1] may be extended to these generalized p-sparse constructions. If each p-th line of the class L_0 has been deleted from the celestial configuration $m\#(s_1, t_1; \ldots; s_k, t_k)$, we denote the resulting incidence structure by $m\#(s_1^{p*}, t_1; \ldots; s_k, t_k)$; it is not a configuration. The notation $m\#(s_1*, t_1; \ldots; s_k, t_k)$ that was used in [1] should now be written as $m\#(s_1^{2*}, t_1; \ldots; s_k, t_k)$ since all of those constructions were 2-sparse.

We append the symbol $D*$ to the end of the sequence to indicate that *for $0 \le i < m$ we add the diameter D_i if any of the points on D_i have been affected by the line deletion.* For brevity we do not explicitly state the indices of the added diameters. These can be recovered if necessary: under the conditions of Theorem 3.1, the added diameters are D_i with $\frac{i}{2} \equiv 0$ or $\frac{i}{2} \equiv s_1 \pmod p$. Hence if $m\#(s_1, t_1; \ldots; s_k, t_k)$ is a celestial configuration,



(a) $24\#(2^{4*}, 10; 7, 2; 10, 7); D*$ (b) $24\#(2^{6*1}, 10; 7, 2; 10, 7); D*$

Figure 5: The celestial symbol $24\#(2, 10; 7, 2; 10, 7)$ satisfies the hypotheses of Theorem 3.1 for both $p = 4$ and $p = 6$, yielding two new configurations.

then $m\#(s_1^{p*}, t_1; \ldots; s_k, t_k)$ is an incidence structure formed by removing each p-th line in L_0, and Theorem 3.1 asserts that $m\#(s_1^{p*}, t_1; \ldots; s_k, t_k); D*$ is again a configuration under certain conditions on m, s_i, t_i, and p.

We will need more powerful notation in the next section. In our examples so far, we deleted the lines $(L_0)_q$ for all $q \equiv 0 \pmod p$. The construction works equally well if we delete instead the lines $(L_0)_q$ for all $q \equiv b \pmod p$, where $0 \leq b < p$ (to see this, rotate the configuration through an angle of $-2\pi b/m$ radians, perform the same operation, then rotate back). If $b \neq 0$ we write b following the asterisk in the superscript of s_1; for clarity we may also do this even if $b = 0$.

The construction outlined in Theorem 3.1 also works if instead of deleting every p-th line in class L_0, we instead delete every p-th line in class L_{j-1}, provided the symbol satisfies the (suitably shifted) conditions of Theorem 3.1.

We therefore use the notation

$$m\#(\cdots; s_j^{p*b}, t_j; \cdots)$$

to indicate deletion of each line $(L_{j-1})_q$, with $q \equiv b \pmod p$. See Figure 5b for an example with $j = 0$ and $b = 1$.

The next section details a generalization of this deletion technique, in which several deletions on the same set of lines are performed simultaneously; to denote this, we write

$$m\#(\cdots; s_j^{p*b_1,b_2,b_3}, t_j; \cdots)$$

to indicate deletion of each line $(L_{j-1})_q$, with $q \equiv b_1$ or $q \equiv b_2$ or $q \equiv b_3 \pmod p$.

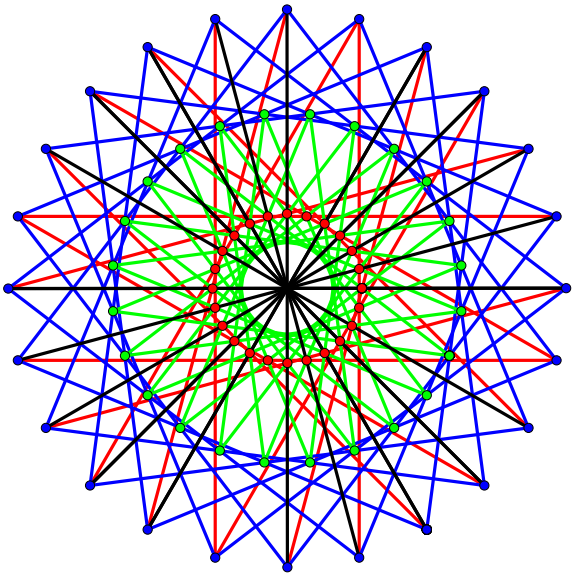


Figure 6: The constructions illustrated in Figure 5 have been carried out simultaneously to obtain $24\#(2^{4*0,6*1}, 10; 7, 2; 10, 7); D*$. This procedure degrades the symmetry group from d_{24} to d_2.

4 Repetition of sparse line deletion

4.1 Multiple deletions within the first line class

Consider the celestial configuration $24\#(2, 10; 7, 2; 10, 7)$. This symbol satisfies the conditions for Theorem 3.1 for both $p = 4$ and $p = 6$. We can delete the lines $(L_0)_q$ for $q = 0, 4, 8, \cdots$ and add diameters to obtain $24\#(2^{4*0}, 10; 7, 2; 10, 7); D*$, depicted in Figure 5a. The affected points of v_0 are those $(v_0)_i$ with $i \equiv 0$ or $i \equiv s_1 = 2 \pmod 4$. This leaves all of the $(v_0)_i$ with odd i untouched. On the other hand, if we delete all lines $(L_0)_q$ with $q \equiv 1 \pmod 6$, only points $(v_0)_i$ of odd index will be affected: see Figure 5b. We may therefore perform both constructions together to obtain the configuration $24\#(2^{4*0,6*1}, 10; 7, 2; 10, 7); D*$, depicted in Figure 6. We have now added all but two of the even-numbered diameters, and the deletion is "sparse" only in comparison with the construction given in [1]. The resulting configuration has only the four symmetries of a rectangle, compared to the 48 symmetries of the underlying celestial configuration. That is, the new symmetry group has index 12 in the original group. For 2-sparse line deletion the index is at most 4. This indicates that the more general procedure can give qualitatively novel configurations.

Many celestial configurations admit p-sparse line deletions for several values of p. A naïve exhaustion search by machine using the conditions of the theorem uncovered several extreme examples. The celestial configuration $48\#(13, 11; 20, 13; 11, 20)$ admits p-sparse line deletion with $p = 2, 3, 4, 6$ or 12. With $80\#(12, 28; 23, 12; 28, 23)$ we can take $p = 5, 8, 10$, or 20. By repeating and combining the p-sparse line deletions for some-

what larger values of p, we can obtain a large number of new 4-configurations. Even in the relatively small case of $24\#(2,10;7,2;10,7)$ we can obtain the configurations illustrated in Figure 7 in addition to those from Figures 5 and 6.

4.2 Combining deletions in the first and third classes in 4-celestial configurations

Another possibility for repetition of the p-sparse line deletion construction arises in the special case of 4-celestial configurations. Suppose that the hypotheses of Theorem 3.1 hold for the celestial symbol

$$m\#(s_1,t_1;s_2,t_2;s_3,t_3;s_4,t_4)$$

with $p=p_1$. Suppose further that they hold for the symbol

$$m\#(s_3,t_3;s_4,t_4;s_1,t_1;s_2,t_2)$$

with $p=p_2$. Beginning from the first symbol $m\#(s_1,t_1;s_2,t_2;s_3,t_3;s_4,t_4)$, we may then perform p_1-sparse deletion on the lines L_0 and add even-numbered diameters to recover a configuration (as in Theorem 3.1). We may additionally perform p_2-sparse deletion on the lines L_2 and add odd-numbered diameters to recover yet another configuration of a family not available previously.

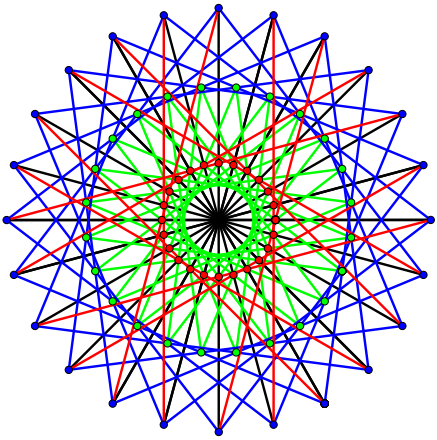
For example, the 4-celestial configuration $20\#(6,4;3,6;7,3;4,7)$ admits a 5-sparse deletion on both L_0 and L_2. We can delete the lines $(L_0)_q$ with $q\equiv 0 \pmod 5$ and add even-numbered diameters, or we can delete the lines $(L_2)_q$ with $q\equiv 0 \pmod 5$ and add odd-numbered diameters to obtain a 4-configuration. We can also do both; in this case we arrive at the configuration $20\#(6^{5*},4;3,6;7^{5*},3;4,7);D*$, depicted in Figure 8b. By rotating the first construction we obtain $20\#(6^{5*1},4;3,6;7^{5*0},3;4,7);D*$, depicted in Figure 8c. These three objects are different, at least in the geometric sense that they differ by more than an isometry, illustrating the very large number of new configurations available through this method.

Finally we note the possibility of repeating deletions within L_0 and also repeating deletions within L_2. An example is $20\#(6^{5*0,2},4;3,6;7^{5*0,4},3;4,7);D*$; see Figure 8a and note again the very small symmetry group.

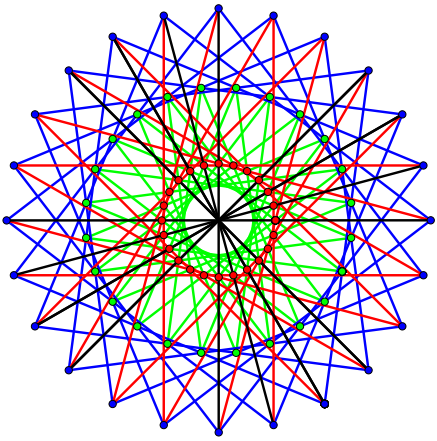
5 Constructions with an odd number of points per symmetry class

Let $\mathcal{C}$ be a k-celestial configuration with symbol $m\#(s_1,t_1;\ldots;s_k,t_k)$. Suppose that m is odd. The hypotheses of Theorem 3.1 cannot hold; in this section we ask if there is another way to remove some lines of $\mathcal{C}$ and then add an equal number of diameters to recover a 4-configuration. We will give some examples where this succeeds and suggest a classification of the resulting configurations. We leave open the task of giving explicit construction algorithms with sufficient conditions on m,s,t and p.

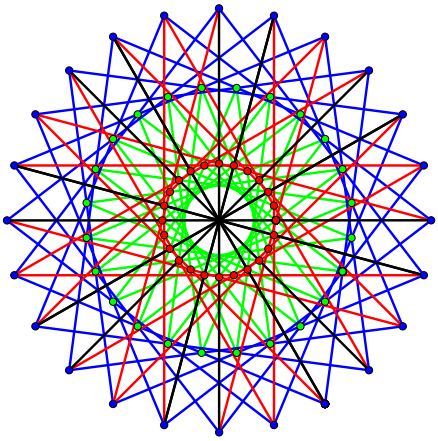
We claim that such a construction is possible only if $k=4$. Indeed, since m is odd every diameter passes through exactly one point in each symmetry class; if the added diameters are lines in a 4-configuration then there must be exactly four classes.



(a) $24\#(2^{4*0,1}, 10; 7, 2; 10, 7); D*$

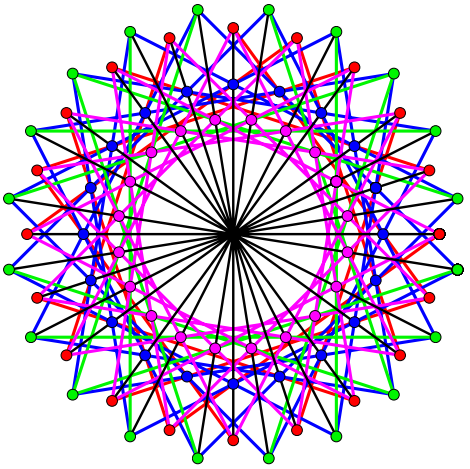


(b) $24\#(2^{6*0,1}, 10; 7, 2; 10, 7); D*$

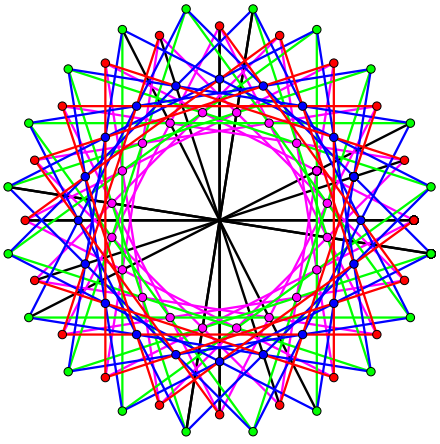


(c) $24\#(2^{6*0,3}, 10; 7, 2; 10, 7); D*$

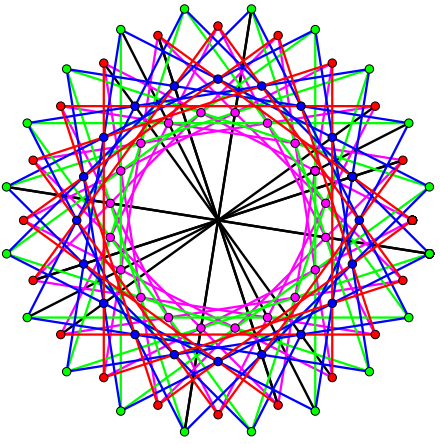
Figure 7: Three more configurations arising from multiple modifications to the same celestial configuration as in Figures 5 and 6. Note that in the configuration shown in (a) we have deleted every red line $(L_0)_q$ where q is congruent to 0 or 1 (mod 4). As a result all even-numbered diameters have been added, although this configuration cannot be constructed via 2-sparse line deletion.



(a) $20\#(6^{5*0,2}, 4; 3, 6; 7^{5*0,4}, 3; 4, 7); D*$



(b) $20\#(6^{5*}, 4; 3, 6; 7^{5*}, 3; 4, 7); D*$.



(c) $20\#(6^{5*1}, 4; 3, 6; 7^{5*0}, 3; 4, 7); D*$.

Figure 8: Three configurations obtained from $20\#(6, 4; 3, 6; 7, 3; 4, 7)$ by performing 5-sparse deletion on both L_0 and L_2. Both odd- and even-numbered diameters have been added.

The examples in previous sections proceeded in steps where some of the lines in one symmetry class were removed and diameters were added to yield a new configuration; in the more complicated examples several intermediate configurations were formed and destroyed along the way. With m odd such a scheme cannot work. Because each added diameter passes through points in four symmetry classes, while the lines of any line class L_j pass only through the two point classes v_j, v_{j+1}, we must simultaneously delete lines from more than one symmetry class. The necessity of coordinating these different classes of removed lines is the main challenge in this section.

We propose the following classification for line deletion constructions. The *ray* from the origin through $(v_0)_0$ passes through either zero points or one point from each of the classes v_1, v_2, v_3. If this ray passes through a point of class v_j we say that v_j is a *cis* class; otherwise we say that v_j is a *trans* class (that is, trans classes are on the opposite side of the origin from points v_0, while cis classes are on the same side of the origin as points v_0). Hence v_0 is always a cis class, and our Theorem 3.1 addresses the case where m is even and the set of cis classes is $\{v_0, v_1\}$.

There are $2^3 = 8$ possible sets of cis classes in a 4-celestial configuration. In Figures 9, 10, and 11 we give examples where m is odd and the cis classes are $\{v_0, v_1\}$, $\{v_0, v_1, v_2, v_3\}$, and $\{v_0, v_2, v_3\}$ respectively. It may be that for each of the eight possibilities one can find sufficient conditions for some line deletion procedure in the spirit of Theorem 3.1. This problem is beyond our scope here.

6 Questions for further study

In *Configurations of Points and Lines*, Grünbaum wrote that "constructing new 4-configurations is still more of an art than a science" [3]. We now offer several possible directions for future work towards the ultimate goal of finding and classifying all 4-configurations.

The technique we have explored here, the replacement of some lines of a celestial configuration with an equal number of diameters, can be extended further. The examples given in Section 5 should be systematized with explicit construction algorithms and sufficient conditions. There are also possibilites with m even that are not covered by Theorem 3.1. Figure 12 gives an example with $m = 12$ where v_0 and v_2 are of cis type, in contrast to the situation of Theorem 3.1, where v_0 and v_1 are of cis type. This could be the first example of a new infinite family obtained by a more general construction.

We also have yet to consider the "even deletion" procedure introduced in [1], in which points as well as lines are removed. This no doubt has a p-sparse generalization and could be worth exploring since the "even deletion" construction in [1] yielded previously unknown (25_4) configurations.

We close by mentioning a related question. We say that two configurations are *(combinatorially) isomorphic* if there exists an incidence-preserving bijection between the two configurations. It is not clear how many of the configurations introduced here belong to new isomorphism classes in this combinatorial sense. For example, it is not known whether or not the configurations depicted in Figures 8b and 8c are combinatorially isomorphic. Even for the celestial configurations this question has not been solved.

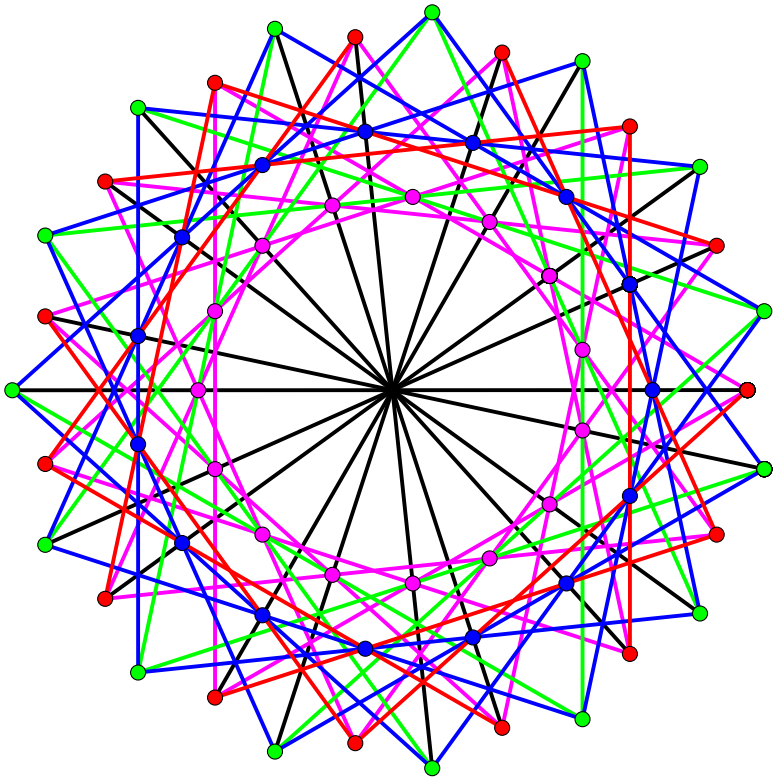


Figure 9: The configuration $15\#(4^{3*}, 2; 1, 4; 5^{3*}, 1; 2, 5); D*$. The *ray* from the origin through $(v_0)_0$ (red) passes through $(v_1)_{-1}$ (blue) but no points of v_2 (green) or v_3 (magenta), so the cis classes are v_0 and v_1.

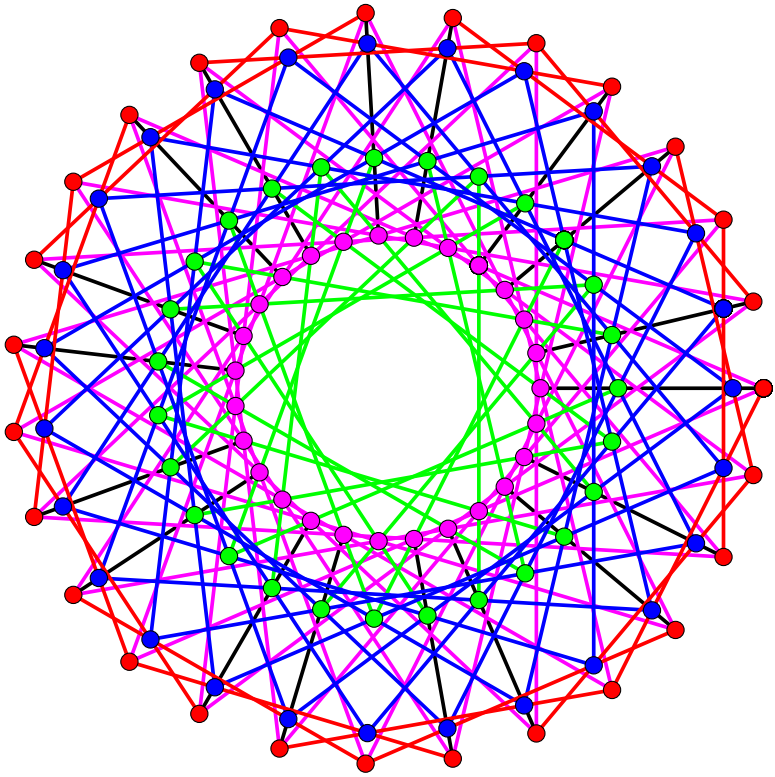


Figure 10: The configuration $27\#(4^{3*}, 2; 8, 4; 10^{3*}, 8; 2, 10); D*$. All four points on each added diameter lie on the same side of the origin, so all point classes are of cis type. The diameters could be extended through the origin without hitting other points because m is odd.

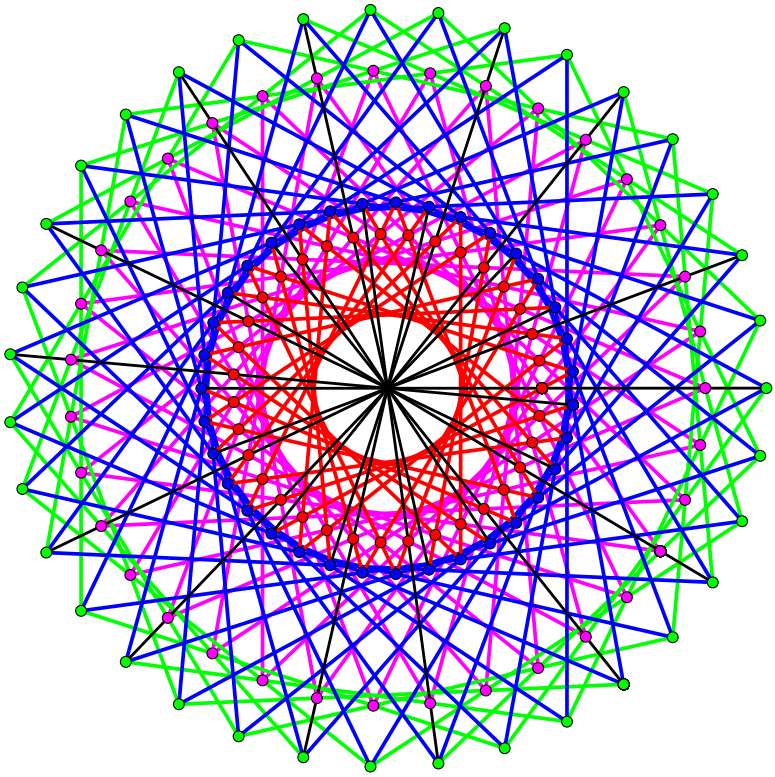


Figure 11: The configuration $35\#(12^{5*}, 13; 3, 12; 7^{5*}, 3; 13, 7); D*$. The cis classes are v_0, v_2, and v_3.

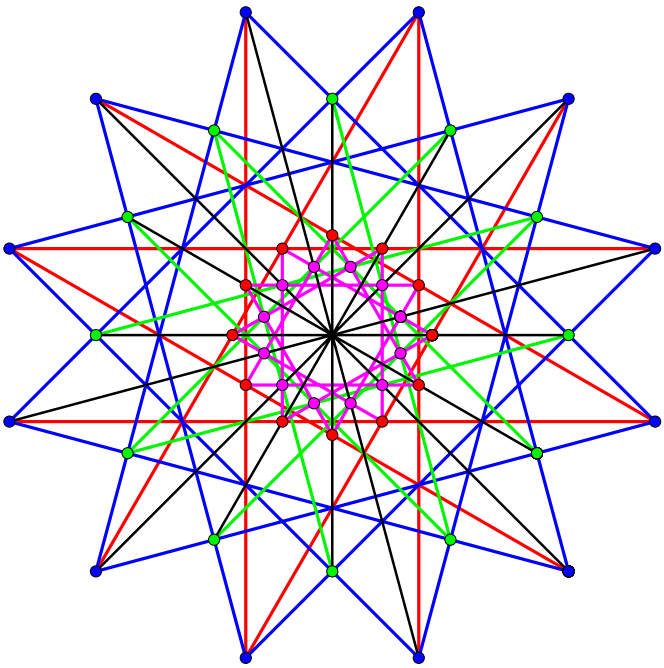


Figure 12: The configuration $12\#(2^{3*0}, 5; 4, 3; 5^{3*1}, 2; 3, 4); D*$. Here the cis classes are v_0 and v_2.

References

[1] L. W. Berman and B. Grünbaum, Deletion constructions of symmetric 4-configurations. Part I, *Contrib. Discrete Math.* **5** (2010), 18–33.

[2] M. Boben and T. Pisanski, Polycyclic configurations, *European J. Combin.* **24** (2003), 431–457, doi:10.1016/S0195-6698(03)00031-3.

[3] B. Grünbaum, *Configurations of Points and Lines*, volume 103 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 2009.

[4] B. Grünbaum and J. F. Rigby, The real configuration (21_4), *J. London Math. Soc. (2)* **41** (1990), 336–346, doi:10.1112/jlms/s2-41.2.336.

[5] J. D. Hunter, Matplotlib: A 2d graphics environment, *Computing In Science & Engineering* **9** (2007), 90–95.

[6] D. Marušič and T. Pisanski, Weakly flag-transitive configurations and half-arc-transitive graphs, *European J. Combin.* **20** (1999), 559–570, doi:10.1006/eujc.1999.0302.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 187–195

Bounds on the domination number of Kneser graphs

Patric R. J. Östergård *

Department of Communications and Networking,
Aalto University School of Electrical Engineering,
P.O. Box 13000, 00076 Aalto, Finland

Zehui Shao †

Key Laboratory of Pattern Recognition and Intelligent Information Processing,
Institutions of Higher Education of Sichuan Province,
School of Information Science and Technology,
Chengdu University, Chengdu, 610106, China

Xiaodong Xu

Guangxi Academy of Science, Nanning,
Guangxi 530007, China

Received 14 May 2013, accepted 17 August 2014, published online 28 November 2014

Abstract

The Kneser graph $\mathrm{KG}_{n,k}$ has one vertex for each k-subset of an n-set and edges between vertices whenever the corresponding subsets are disjoint. A dominating set in a graph $G = (V, E)$ is a subset $S \subseteq V$ such that each vertex in $V \setminus S$ is adjacent to at least one vertex in S. The domination number of $\mathrm{KG}_{n,k}$, denoted by $\gamma(n, k)$, is the minimum size of a dominating set in that graph. Combinatorial and computer-aided techniques for obtaining bounds on $\gamma(n, k)$ are here considered, and several new bounds are obtained. An updated table of bounds on $\gamma(n, k)$ is presented for $n \leq 21$ and $k \leq 5$.

Keywords: Dominating set, domination number, Kneser graph.

Math. Subj. Class.: 05C69, 05C35

*Corresponding author. Supported in part by the Academy of Finland, Grant No. 132122.
†Supported by the National Natural Science Foundation of China, Grant No. 61309015.
E-mail addresses: patric.ostergard@aalto.fi (Patric R. J. Östergård), zshao@cdu.edu.cn (Zehui Shao), xxdmaths@sina.com (Xiaodong Xu)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

1 Introduction

Let $G = (V, E)$ be a simple graph, that is, a graph having neither loops nor multiple edges. A *dominating set* in G is a subset $S \subseteq V$ such that each vertex in $V \setminus S$ is adjacent to at least one vertex in S. The *domination number* $\gamma(G)$ of G is the minimum size of a dominating set in G. The domination number has been extensively studied in the general case [11]. Due to a variety of applications, the case of n-cubes, $G \cong Q_n$, is of particular interest [3]; most of the early work considered such graphs. In the current work, another specific type of graphs is considered, namely Kneser graphs.

The *Kneser graph* $\mathrm{KG}_{n,k}$ has one vertex for each k-subset of an n-set and edges between vertices whenever the corresponding subsets are disjoint. If $n < 2k$, then $\mathrm{KG}_{n,k}$ has no edges, so we assume that $n \geq 2k$. We further denote the domination number of $\mathrm{KG}_{n,k}$ by $\gamma(n, k)$. See [6, Chap. 7] for an in-depth discussion of Kneser graphs.

General and specific bounds on $\gamma(n, k)$ have been considered in a sequence of studies, including [2, 8, 10, 12, 18]. However, several of the best known bounds for small parameters were rather weak prior to this study. Indeed, the aim of the current work is to apply combinatorial and computer-aided techniques to the problem of improving upper and lower bounds on the domination number of Kneser graphs—and occasionally even find the exact value when the bounds meet.

A *total dominating set* in a graph $G = (V, E)$ is a subset $S \subseteq V$ such that each vertex in V is adjacent to at least one vertex in S. The minimum size of a total dominating set is called the *total domination number*, and the total domination number of $\mathrm{KG}_{n,k}$ is denoted by $\gamma_t(n, k)$. It is obvious that

$$\gamma(n, k) \leq \gamma_t(n, k). \tag{1.1}$$

Let $C(v, k, t)$ denote the smallest number of k-subsets of a v-set, such that every t-subset of the v-set occurs in at least one of the k-subsets. Then $\gamma_t(n, k) = C(n, n - k, k)$, so by (1.1),

$$\gamma(n, k) \leq C(n, n - k, k). \tag{1.2}$$

Exact values of and upper bounds on $C(v, k, t)$ for $v \leq 32$, $k \leq 16$, and $t \leq 5$ can be found in [7].

The paper is organized as follows. Methods for obtaining upper and lower bounds on $\gamma(n, k)$ are considered in Sections 2 and 3, respectively. The results are summarized in Section 4, where an updated table of bounds on $\gamma(n, k)$ is presented for $n \leq 21$ and $k \leq 5$.

2 Upper Bounds and Exact Values

Upper bounds for the domination number are commonly constructive, that is, explicit dominating sets prove the bounds. We here present various general results for upper bounds on $\gamma(n, k)$; in some of the theorems, the exact value is in fact obtained. If $n < 2k$, then the Kneser graph consists of isolated vertices only. In the first nontrivial case, $n = 2k$, a dominating set must contain one vertex from each pair of disjoint k-sets.

Theorem 2.1. *For any k,*

$$\gamma(2k, k) = \frac{1}{2}\binom{2k}{k}.$$

It is easy to show that if n is large enough, then the smallest dominating set is obtained by taking $k + 1$ disjoint k-sets.

Theorem 2.2. *If $n \geq k^2 + k$, then $\gamma(n,k) = k+1$.*

The exact value of $\gamma(n,k)$ has also been determined for a range of values of n smaller than those covered by Theorem 2.2.

Theorem 2.3 ([10])**.** *If $k \geq 3$ and $\frac{3}{4}k^2 + k \leq n < k^2 + k$, then*

$$\gamma(n,k) = k + 1 + \left\lceil \frac{k^2 + k - n}{\lfloor k/2 \rfloor} \right\rceil .$$

With increasing n, $\gamma(n,k)$ turns out to be nonincreasing.

Theorem 2.4 ([8, Proposition 4.2.4])**.** *If $n \geq 2k+1$, then $\gamma(n+1,k) \leq \gamma(n,k)$.*

Theorem 2.5 ([8, Theorem 4.5.1])**.** *If $k \geq 4$ and $\gamma(n,k) \leq \min\{2k, n-k\}$, then $\gamma(n,k) = \gamma_t(n,k)$.*

We shall next see how dominating sets in certain Kneser graphs are related to a coloring problem for hypergraphs that has been extensively studied. We consider the case $n = 2k+1$—such Kneser graphs are known as *odd graphs*—and view a dominating set S of the graph $\mathrm{KG}_{2k+1,k}$ as the set of hyperedges in a hypergraph $G' = (V', E')$ with $|V'| = n$, $|E'| = |S|$, and edges of size k (so the hypergraph is k-*uniform*).

Now consider an arbitrary *balanced* coloring of the vertices in V' with two colors, that is, k of the vertices are colored with one color and $k+1$ with the other [21]. Since the vertex in the original Kneser graph that is labelled by the subset of the k vertices with the first color is dominated by some vertex $s \in S$, the hyperedge in E' corresponding to s is unicolor. Hence, G' does not have a balanced coloring with two colors so that no hyperedge is unicolor, that is, G' is not 2-colorable in a balanced way.

Hypergraphs that are 2-colorable (without requiring that the colorings be balanced) are said to have property B [13, Sect. 15.1]. Consequently, a hypergraph with appropriate parameters that does not have property B gives a dominating set in $\mathrm{KG}_{2k+1,k}$. Actually, this implication goes in the other direction as well.

The *upward shadow* of a $(k-1)$-subset of an n-set is the collection of all k-subsets of the n-set that contain the $(k-1)$-subset.

Lemma 2.6 ([8, Lemma 4.2.3])**.** *If $n \geq 2k+1$, then there exists a dominating set attaining $\gamma(n,k)$ that does not contain the upward shadow of any $(k-1)$-subset of the n-set.*

Theorem 2.7. *There exists a dominating set S attaining $\gamma(2k+1,k)$ that can be transformed into a k-uniform hypergraph $G' = (V,E)$ with $|V| = 2k+1$ and $|E| = |S|$ without property B.*

Proof. Consider the hypergraph $G' = (V', E')$ obtained from a dominating set attaining $\gamma(2k+1,k)$ that does not contain the upward shadow of any $(k-1)$-set. Such a dominating set exists by Lemma 2.6.

If the vertices in V' are colored in a balanced way, then since the hypergraph was constructed from a dominating set in $\mathrm{KG}_{2k+1,k}$, there exists a unicolor hyperedge.

Now consider a coloring of the vertices V' with $k+2$ and $k-1$ vertices of the two colors and denote the vertices in the former color class by U. If no hyperedge in E' is a subset of U, then the same holds for each $(k+1)$-subset of U, so by the existence of a balanced coloring we get that $(V' \setminus U) \cup v \in E'$ for all $v \in U$. But this is then an

upward shadow and we have a contradiction (so a hyperedge that is a subset of U is indeed unicolor).

The colorings with classes of size a and $n-a$ where $a \geq k+3$ are handled by considering an arbitrary subset of the larger class of size $k+2$ and using the previous argument. □

By Theorem 2.7 and results by Abbott and Liu [1] we now get that $24 \leq \gamma(9,4) \leq 26$. Note, however, that in certain studies on uniform hypergraphs without property B a further assumption is made that the hyperedges must contain all pairs of vertices. Results for that variant of the problem, which is motivated by a more general question regarding property B, are not directly applicable here. This includes the results in [16, 21].

For certain parameters, we can say a lot about $\gamma(2k+1,k)$. Let $2 \leq t < k < v$. An $S(t,k,v)$ *Steiner system* is a collection of k-sets out of a v-set with the property that every t-subset of the v-set occurs in exactly one of the k-sets. The following result has been discussed both in the context of hypergraphs without property B [4] and dominating sets in $\mathrm{KG}_{2k+1,k}$ [9]; see also [5, Lemma 11.8.3].

Theorem 2.8. *There is a Steiner system $S(k-1,k,2k+1)$ if and only if*

$$\gamma(2k+1,k) = \frac{1}{k}\binom{2k+1}{k-1}.$$

One may further use partial or exhaustive computational methods to determine bounds on $\gamma(n,k)$. Since upper bounds can be proven by finding a structure attaining the bound, nonexhaustive methods can be applied to such cases. For lower bounds, on the other hand, exhaustive methods are required; we shall consider such methods in the next section.

In [19], the tabu search metaheuristic is applied to the problem of finding dominating sets in n-cubes. The algorithm takes the parameters of the instance and the desired size of the dominating set (which is, for example, one less than the best known upper bound), and searches for such a dominating set. The algorithm is also applicable to Kneser graphs—in fact, it is applicable to arbitrary graphs. The reader should consult [19] for details. Structures obtained in the current work and leading to new upper bounds on $\gamma(n,k)$ are listed in the Abstract. Some of the best known structures turn out to have nontrivial symmetries; these could further be used to narrow down the search space.

We consider symmetries of dominating sets in terms of the labels (subsets) of the vertices. Two dominating sets are said to be *equivalent* if there is a permutation of the n-set that maps the vertices of one dominating set onto the vertices of the other. Such a mapping from a dominating set onto itself is an *automorphism*, and the set of all automorphisms of a dominating set forms the *automorphism group* of the dominating set. Such an automorphism group is isomorphic to a subgroup of the stabilizer subgroup of the dominating set in $\mathrm{Aut}(\mathrm{KG}_{n,k})$. Note that it may happen that the automorphism group is a *proper* such subgroup; consider, for example, the mapping of the k-subsets of a $2k$-set to their complements. The *nauty* software [17] is a useful computational tool in this context.

3 Lower Bounds

Several of the bounds in the previous section can be used also to get lower bounds. We shall here state two more general results that can be used to get best known lower bounds for small parameters. The first of these is the well-known *volume bound* obtained by dividing the total number of vertices with the number of vertices dominated by a single vertex.

Theorem 3.1. *For any n and k,*

$$\gamma(n,k) \geq \frac{\binom{n}{k}}{1+\binom{n-k}{k}}.$$

Theorem 3.2 ([8, Lemma 4.5.3])**.** *Assume that $n = \alpha k$, where $\alpha, k \geq 2$. Then*

$$\gamma(n,k) \geq \frac{\alpha}{\alpha-1}(\gamma(n+k,k)-1).$$

To simplify the discussion of the techniques to obtain lower bounds on $\gamma(n,k)$, it is useful to think of a dominating set as a *constant weight code* [20]. The *codewords* of this code are of length n and have 1s in the coordinates given by the corresponding k-subset and 0s in the other coordinates. Theory and terminology from coding theory can then be directly applied.

There are three general approaches to exhaustively search for codes with prescribed parameters [15, Chap. 7]: via subcodes, codeword by codeword, and coordinate by coordinate. There is no obvious way of constructing the current type of codes via subcodes, that is, codes obtained by considering the codewords with a 1 (alternatively, 0) in a given coordinate and deleting that coordinate. Some results can indeed be obtained in a backtrack search constructing the code word by word as in [24], which can be consulted for general details.

The origins of the method of constructing codes coordinate by coordinate can be traced back to the 1960s [14], after which it has been developed further and become an efficient tool in the study of dominating sets, in particular in n-cubes and related graphs [22, 23]. See also [15, Sect. 7.2.2]. A version that has been applied to the hypergraph coloring problem discussed in the previous section can be found in [21].

The idea in the coordinate by coordinate approach can be described as a generalization of the following theorem.

Theorem 3.3. *Let D be a dominating set in $\mathrm{KG}_{n,k}$, and let $D = D_0 \cup D_1$ such that D_i consists of the vertices whose label has an i in the first coordinate. Then*

1. $|D_1| + \binom{n-k-1}{k-1}|D_0| \geq \binom{n-1}{k-1}$,
2. $\binom{n-k}{k}|D_1| + \left(1+\binom{n-1-k}{k}\right)|D_0| \geq \binom{n-1}{k}$.

Proof. Let $G = (V,E)$ be the $\mathrm{KG}_{n,k}$ Kneser graph, and let $V = V_0 \cup V_1$ so that the labels of the vertices in V_i have an i in the first coordinate. Then $|V_0| = \binom{n-1}{k}$ and $|V_1| = \binom{n-1}{k-1}$. The result now follows as each vertex in D_0 dominates $1+\binom{n-1-k}{k}$ vertices in V_0 and $\binom{n-k}{k}$ vertices in V_1, and each vertex in D_1 dominates $\binom{n-k-1}{k-1}$ vertices in V_0 and 1 vertex in V_1. □

Theorem 3.3 can be generalized to an arbitrary number of specified coordinates. For example, with two specified coordinates we let $D = D_{00} \cup D_{01} \cup D_{10} \cup D_{11}$, $V = V_{00} \cup V_{01} \cup V_{10} \cup V_{11}$ and get four inequalities. For a small number of coordinates, the inequalities can be derived by hand, but when the number gets larger, it is convenient to form these computationally.

When a code is constructed coordinate by coordinate, one first fixes the number of codewords and then start from the distributions of 0s and 1s in the first coordinate given by

Theorem 3.3. In a backtrack exhaustive search, one may for the next couple of coordinates solve larger and larger systems of equations, but at some point one may start checking all possible candidates for the next coordinate and see whether the inequalities are fulfilled. At each level of the search tree, isomorph rejection should be carried out. For the sake of efficiency, one may also require that the number of 1s in the coordinates is either increasing or decreasing. Except for minor differences in details, the approach in [21] can be used.

4 Results

Table 1 summarizes the best known bounds on and exact values for $\gamma(n,k)$, $n \leq 21$, $k \leq 5$. Indices are added to the entries, to give explanations of lower and upper bounds. If a bound can be obtained in several ways, we pick the explanation that in some sense is the nicest. We omit the index when the bound follows from Theorem 2.2.

Table 1: Bounds on $\gamma(n,k)$ for $n \leq 21$, $k \leq 5$

$n\backslash k$	2	3	4	5
4	${}^{a}3^{a}$			
5	${}^{c}3^{c}$			
6	3	${}^{a}10^{a}$		
7	3	${}^{e}7^{e}$		
8	3	${}^{c}7^{c}$	${}^{a}35^{a}$	
9	3	${}^{m}7^{c}$	${}^{j}26^{l}$	
10	3	${}^{b}6^{b}$	${}^{c}15^{k}$	${}^{a}126^{a}$
11	3	${}^{b}5^{b}$	${}^{i}15^{c}$	${}^{e}66^{e}$
12	3	4	${}^{i}12^{h}$	${}^{i}37\text{–}56^{k}$
13	3	4	${}^{m}10^{h}$	${}^{j}23\text{–}39^{k}$
14	3	4	${}^{d}9^{h}$	${}^{f}16\text{–}31^{k}$
15	3	4	${}^{d}8^{h}$	${}^{g}15\text{–}27^{h}$
16	3	4	${}^{b}7^{b}$	${}^{i}12\text{–}22^{h}$
17	3	4	${}^{b}7^{b}$	${}^{c}11\text{–}17^{h}$
18	3	4	${}^{b}6^{b}$	${}^{c}11\text{–}15^{h}$
19	3	4	${}^{b}6^{b}$	${}^{c}11\text{–}14^{h}$
20	3	4	5	${}^{c}11\text{–}12^{h}$
21	3	4	5	${}^{d}11\text{–}12^{h}$

Key to Table 1.
Unmarked bounds are from Theorem 2.2.
Bounds:

- ^{a} Theorem 2.1
- ^{b} Theorem 2.3
- ^{c} Theorem 2.4
- ^{d} Theorem 2.5 and [7]
- ^{e} Theorem 2.8
- ^{f} Theorem 3.1
- ^{g} Theorem 3.2
- ^{h} Eq. (1.2) and [7]
- ^{i} Exhaustive search, coordinate by coordinate
- ^{j} Exhaustive search, word by word
- ^{k} New constructive result, see Appendix
- ^{l} Abbott and Liu [1] (Theorem 2.7)
- ^{m} Gorodezky [8]

Appendix

We here list the structures that lead to new upper bounds on $\gamma(n, k)$. We first present structures that can be described as a set of orbits under the action of a permutation group, and finally list some explicit structures (we do not exclude the possibility that these, or better, bounds could also be obtained by a structure with some symmetry).

$\gamma(10, 4) \leq 15$:
Generator of group: $(1\ 2\ 3\ 4\ 5)(6\ 7\ 8\ 9\ 10)$
Orbit representatives: 1110010000, 1010000110, 1000001101

$\gamma(13, 5) \leq 39$:
Generator of group: $(1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13)$
Orbit representatives: 1101011000000, 1110001001000, 1101000100010

$\gamma(12, 5) \leq 56$:
111000110000, 100000111100, 010001011001, 010000111010, 001010000111, 011100010001, 110101000010, 100011010100, 000011111000, 110100100001, 110011001000, 011101001000, 101011001000, 000001100111, 011000100110, 110000010110, 001001100011, 100110001100, 111110000000, 000110010110, 100001101001, 010011100001, 010100001101, 000010011011, 000111100100, 001110110000, 010101010100, 101010010001, 001110001010, 001000111010, 011011000010, 001100001101, 100101001010, 001001011001, 010110001010, 100001010011, 011010100001, 101000101100, 000100101011, 110000101100, 010010000111, 101101000010, 011001000101, 100100110001, 100010000111, 011000011100, 110010010001, 010011100100, 101000010110, 000100011101, 000101110010, 101100100001, 001011100100, 111000001010, 001101010100, 010110110000

$\gamma(14,5) \leq 31$:
10000011001100, 00011110100000, 11100100100000, 00110101000010,
00110110000001, 00001011000110, 00010010100110, 01000011000011,
00000101101100, 01011000010100, 10110000010001, 00100000100111,
10001010011000, 01101010000001, 01100010011000, 01010100011000,
11000000001101, 10010000101010, 01100001010100, 00001001110001,
10001111000000, 01001010110000, 00000001011011, 10000100010110,
11001100000010, 00101000101010, 00001100001101, 10111001000000,
01010001100001, 10000110110000, 00011010001100

References

[1] H. L. Abbott and A. C. Liu, On property B of families of sets, *Canad. Math. Bull.* **23** (1980), 429–435.

[2] E. Clark, *Domination in Kneser graphs*, unpublished manuscript, 1988.

[3] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein, *Covering Codes*, North-Holland, Amsterdam, 1997.

[4] H. L. de Vries, On property B and on Steiner systems, *Math. Z.* **153** (1977), 155–159.

[5] C. D. Godsil, *Algebraic Combinatorics*, Chapman & Hall, New York, 1993.

[6] C. Godsil and G. Royle, *Algebraic Graph Theory*, Springer-Verlag, New York, 2001.

[7] D. M. Gordon and D. R. Stinson, Coverings, in: C. J. Colbourn and J. H. Dinitz (eds.), *Handbook of Combinatorial Designs*, 2nd ed., Chapman & Hall/CRC, Boca Raton, 2007, 365–373.

[8] I. Gorodezky, *Domination in Kneser graphs*, Master's thesis, University of Waterloo, Canada, 2007.

[9] P. Hammond and D. H. Smith, Perfect codes in the graphs O_k, *J. Combin. Theory Ser. B* **19** (1975), 239–255.

[10] C. Hartman and D. B. West, *Covering designs and domination in Kneser graphs*, unpublished manuscript, 2003.

[11] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater, *Fundamentals of Domination in Graphs*, Marcel Dekker, New York, 1998.

[12] J. Ivančo and B. Zelinka, Domination in Kneser graphs, *Math. Bohem.* **118** (1993), 147–152.

[13] T. R. Jensen and B. Toft, *Graph Coloring Problems*, Wiley, New York, 1995.

[14] H. J. L. Kamps and J. H. van Lint, The football pool problem for 5 matches, *J. Combin. Theory* **3** (1967), 315–325.

[15] P. Kaski and P. R. J. Östergård, *Classification Algorithms for Codes and Designs*, Springer, Berlin, 2006.

[16] G. Manning, *The $M(4)$ problem of Erdős and Hajnal*, Ph.D. dissertation, Northern Illinois University, 1997.

[17] B. D. McKay, *nauty user's guide (version 1.5)*, Technical Report TR-CS-90-02, Computer Science Department, Australian National University, Canberra, 1990.

[18] J. C. Meyer, Quelques problèmes concernant les cliques des hypergraphes h-complets et q-parti h-complets, in: C. Berge and D. Ray-Chaudhuri (eds.), *Hypergraph Seminar*, Lecture Notes in Mathematics, Vol. 411. Springer-Verlag, Berlin, 1974, 127–139.

[19] P. R. J. Östergård, Constructing covering codes by tabu search, *J. Combin. Des.* **5** (1997), 71–80.

[20] P. R. J. Östergård, Classification of binary constant weight codes, *IEEE Trans. Inform. Theory* **56** (2010), 3779–3785.

[21] P. R. J. Östergård, On the minimum size of 4-uniform hypergraphs without property B, *Discrete Appl. Math.* **163** (2014), 199–204.

[22] P. R. J. Östergård and U. Blass, On the size of optimal binary codes of length 9 and covering radius 1, *IEEE Trans. Inform. Theory* **47** (2001), 2556–2557.

[23] P. R. J. Östergård and A. Wassermann, A new lower bound for the football pool problem for 6 matches, *J. Combin. Theory Ser. A* **99** (2002), 175–179.

[24] P. R. J. Östergård and W. D. Weakley, Classification of binary covering codes, *J. Combin. Des.* **8** (2000), 391–401.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 197–207

The expected values of Kirchhoff indices in the random polyphenyl and spiro chains*

Guihua Huang
College of Mathematics and Computer Science, Hunan Normal University, Changsha, Hunan 410081, P. R. China

Meijun Kuang
College of Mathematics and Computer Science, Hunan Normal University, Changsha, Hunan 410081, P. R. China

Hanyuan Deng †
College of Mathematics and Computer Science, Hunan Normal University, Changsha, Hunan 410081, P. R. China

Received 14 March 2013, accepted 22 April 2014, published online 28 November 2014

Abstract

The Kirchhoff index $Kf(G)$ of a graph G is the sum of resistance distances between all pairs of vertices in G. In this paper, we obtain exact formulas for the expected values of the Kirchhoff indices of the random polyphenyl and spiro chains, which are graphs of a class of unbranched multispiro molecules and polycyclic aromatic hydrocarbons. Moreover, we obtain a relation between the expected values of the Kirchhoff indices of a random polyphenyl and its random hexagonal squeeze, and the average values for the Kirchhoff indices of all polyphenyl chains and all spiro chains with n hexagons, respectively.

Keywords: Expected value, average value, Kirchhoff index, resistance distance, polyphenyl chain, spiro chain.

Math. Subj. Class.: 05C12, 05C80, 05C90, 05D40

*Project supported by Hunan Provincial Natural Science Foundation of China(13JJ3053).
†Corresponding Author.
E-mail addresses: 380026412@qq.com (Guihua Huang), 1075998525@qq.com (Meijun Kuang), hydeng@hunnu.edu.cn (Hanyuan Deng)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

1 Introduction

Based on the electrical network theory, Klein and Randić [13] introduced the concept of resistance distance. A connected graph G with vertex set $\{v_1, v_2, \cdots, v_n\}$ is viewed as an electrical network N by replacing each edge of G with a unit resistor, the resistance distance between v_i and v_j, denoted by $r_G(v_i, v_j)$ or $r(v_i, v_j)$, is the elective resistance between them as computed by the methods of the theory of resistive electrical networks based on Ohm's and Kirchhoff's laws in N.

The Kirchhoff index of G, denoted by $Kf(G)$, is the sum of resistance distances between all pairs of vertices in G, namely,

$$Kf(G) = \sum_{i<j} r_G(v_i, v_j)$$

Like many topological indices, Kirchhoff index is a structure descriptor. The resistance distance is also intrinsic to the graph with some nice purely mathematical and physical interpretations [14] [15]. Also, the Kirchhoff index has been found very useful in chemistry, such as in assessing cyclicity of polycyclic structures including fullerenes, linear hexagonal chains and some special molecular graphs such as circulant graphs, distance-regular graphs and Möbius ladders [1] [18] [22] [24]. Bonchev et al. [4] used it in polymer science and found that the Kirchhoff index in their approach is especially useful for defining the topological radius $R_{top} = \frac{Kf}{n^2}$ of macromolecules containing cyclic fragments. Some closed-form formulae for Kirchhoff index have been given for circulant graphs, linear hexagonal chains and so on [1] [16] [19] [22]. The resistance distance is also well studied in mathematical literatures. Much work has been done to compute Kirchhoff index of some classes of graphs, or give some bounds for Kirchhoff index of graphs and characterize extremal graphs. For instance, unicyclic and bicyclic graphs with extremal Kirchhoff index are characterized and sharp bounds for Kirchhoff index of such graphs are obtained [6] [12] [21] [25] [26].

Polyphenyls and their derivatives, which can be used in organic synthesis, drug synthesis, heat exchangers, etc., attracted the attention of chemists for many years [11] [17] [20]. Spiro compounds are an important class of cycloalkanes in organic chemistry. A spiro union in spiro compounds is a linkage between two rings that consists of a single atom common to both rings and a free spiro union is a linkage that consists of the only direct union between the rings. Some results on energy, Merrifield-Simmons index, Hosoya index and Wiener index of the spiro and polyphenyl chains were reported in [2] [5] [9] [27]. Recently, Deng [7] [8] [10] gave the recurrences or explicit formulae for computing the Wiener index and Kirchhoff index of spiro and polyphenyl chains. Yang and Zhang [23] obtained a simple exact formula for the expected value of the Wiener index of a random polyphenyl chain. In this paper, we will consider the expected values of the Kirchhoff index of random polyphenyl and spiro chains.

A polyphenyl chain PPC_n with n hexagons can be regarded as a polyphenyl chain PPC_{n-1} with $n-1$ hexagons to which a new terminal hexagon has been adjoined by a cut edge, see Figure 1.

Let $PPC_n = H_1H_2\cdots H_n$ be a polyphenyl chain with $n(n \geq 2)$ hexagons, where H_k is the k-th hexagon of PPC_n attached to H_{k-1} by a cut edge $u_{k-1}c_k$, $k = 2, 3, \cdots, n$. A vertex v of H_k is said to be ortho-, meta- and para-vertex of H_k if the distance between v and c_k is 1, 2 and 3, denoted by o_k, m_k and p_k, respectively. Examples of ortho-, meta-, and

para-vertices are shown in Figure 1. Except the first hexagon, any hexagon in a polyphenyl chain has two ortho-vertices, two meta-vertices and one para-vertex.

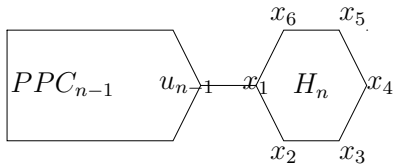


Figure 1: A polyphenyl chain PPC_n with n hexagons, $c_n = x_1$ and ortho-vertices $o_n = x_2, x_6$, meta-vertices $m_n = x_3, x_5$, and para-vertex $p_n = x_4$ in H_n.

A polyphenyl chain PPC_n is a polyphenyl ortho-chain if $u_k = o_k$ for $2 \leq k \leq n-1$. A polyphenyl chain PPC_n is a polyphenyl meta-chain if $u_k = m_k$ for $2 \leq k \leq n-1$. A polyphenyl chain PPC_n is a polyphenyl para-chain if $u_k = p_k$ for $2 \leq k \leq n-1$. The polyphenyl ortho-, meta- and para-chain with n hexagons are denoted by $\overline{O_n}$, $\overline{M_n}$ and $\overline{P_n}$, respectively.

For $n \geq 3$, the terminal hexagon can be attached to meta-, ortho-, or para-vertex in three ways, which results in the local arrangements we describe as PPC^1_{n+1}, PPC^2_{n+1}, PPC^3_{n+1}, see Figure 2.

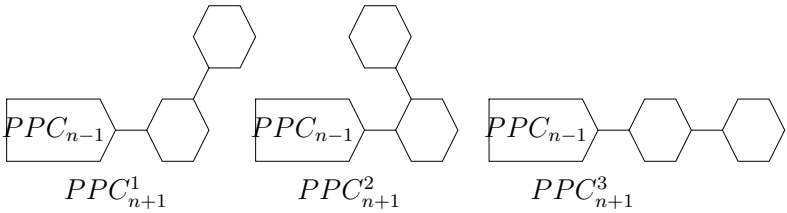


Figure 2: The three types of local arrangements in polyphenyl chains.

A random polyphenyl chain $PPC(n, p_1, p_2)$ with n hexagons is a polyphenyl chain obtained by stepwise addition of terminal hexagons. At each step $k(= 3, 4, \cdots, n)$, a random selection is made from one of the three possible constructions:

(i)$PPC_{k-1} \to PPC^1_k$ with probability p_1,

(ii)$PPC_{k-1} \to PPC^2_k$ with probability p_2,

(iii)$PPC_{k-1} \to PPC^3_k$ with probability $1 - p_1 - p_2$

where the probabilities p_1 and p_2 are constants, irrespective to the step parameter k.

Specially, the random polyphenyl chain $PPC(n, 1, 0)$ is the polyphenyl meta-chain $\overline{M_n}$, $PPC(n, 0, 1)$ is the polyphenyl orth-chain $\overline{O_n}$, and $PPC(n, 0, 0)$ is the polyphenyl para-chain $\overline{P_n}$, respectively.

Also, a spiro chain SPC_n with n hexagons can be regarded as a spiro chain SPC_{n-1} with $n-1$ hexagons to which a new terminal hexagon has been adjoined, see Figure 3.

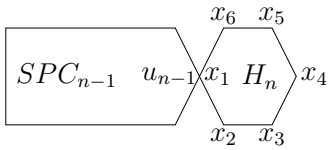


Figure 3: A spiro chain SPC_n with n hexagons.

For $n \geq 3$, the terminal hexagon can also be attached in three ways, which results in the local arrangements we describe as SPC^1_{n+1}, SPC^2_{n+1}, SPC^3_{n+1}, see Figure 4.

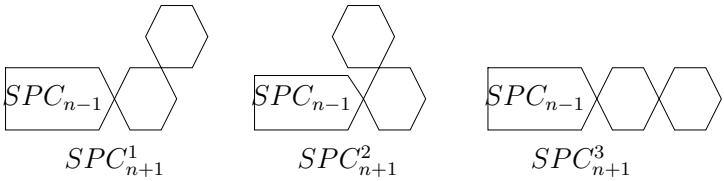


Figure 4: The three types of local arrangements in spiro chains.

A random spiro chain $SPC(n, p_1, p_2)$ with n hexagons is a spiro chain obtained by stepwise addition of terminal hexagons. At each step $k(= 3, 4, \cdots, n)$, a random selection is made from one of the three possible constructions:

(i)$SPC_{k-1} \to SPC^1_k$ with probability p_1,

(ii)$SPC_{k-1} \to SPC^2_k$ with probability p_2,

(iii)$SPC_{k-1} \to SPC^3_k$ with probability $1 - p_1 - p_2$

where the probabilities p_1 and p_2 are constants, irrelative to the step parameter k.

Similarly, the random spiro chain $SPC(n, 1, 0)$, $PPC(n, 0, 1)$ and $PPC(n, 0, 0)$ are the spiro meta-chain M_n, the spiro orth-chain O_n and the spiro para-chain P_n, respectively.

For a random polyphenyl chain $PPC(n, p_1, p_2)$ and a random spiro chain $SPC(n, p_1, p_2)$, their Kirchhoff indices are random variables. In this paper, we will obtain exact formulas for the expected values $E(Kf(PPC(n, p_1, p_2)))$ and $E(Kf(SPC(n, p_1, p_2)))$ of the Kirchhoff indices of random polyphenyl and spiro chains, respectively.

2 Main results

2.1 The Kirchhoff index of the random polyphenyl chain

In this section, we will consider the Kirchhoff index of the random polyphenyl chain.

Theorem 2.1. *For $n \geq 1$, the expected value of the Kirchhoff index of the random polyphenyl chain $PPC(n, p_1, p_2)$ is*

$$E(Kf(PPC(n, p_1, p_2))) = (15 - p_1 - 4p_2)n^3 + (3p_1 + 12p_2 + 8)n^2 - (2p_1 + 8p_2 + \frac{11}{2})n$$

Proof. Note that the polyphenyl chain PPC_n is obtained by attaching PPC_{n-1} a new terminal hexagon by an edge, we suppose that the terminal hexagon is spanned by vertices x_1, x_2, x_3, x_4, x_5, x_6, and the new edge is $u_{n-1}x_1$ (see Fig.1). Then

(i) For any $v \in PPC_{n-1}$,

$r(x_1,v) = r(u_{n-1},v) + 1, r(x_2,v) = r(u_{n-1},v) + 1 + \frac{5}{6}$,
$r(x_3,v) = r(u_{n-1},v) + 1 + \frac{4}{3}, r(x_4,v) = r(u_{n-1},v) + 1 + \frac{3}{2}$,
$r(x_5,v) = r(u_{n-1},v) + 1 + \frac{4}{3}, r(x_6,v) = r(u_{n-1},v) + 1 + \frac{5}{6}$;

(ii) PPC_{n-1} has $6(n-1)$ vertices;

(iii) For $k \in \{1,2,3,4,5,6\}$, $\sum_{i=1}^{6} r(x_k,x_i) = \frac{35}{6}$.

So, we have

$r(x_1|PPC_n) = r(u_{n-1}|PPC_{n-1}) + 1 \times 6(n-1) + \frac{35}{6}$
$r(x_2|PPC_n) = r(u_{n-1}|PPC_{n-1}) + (1+\frac{5}{6}) \times 6(n-1) + \frac{35}{6}$
$r(x_3|PPC_n) = r(u_{n-1}|PPC_{n-1}) + (1+\frac{4}{3}) \times 6(n-1) + \frac{35}{6}$
$r(x_4|PPC_n) = r(u_{n-1}|PPC_{n-1}) + (1+\frac{3}{2}) \times 6(n-1) + \frac{35}{6}$
$r(x_5|PPC_n) = r(x_3|PPC_{n-1})$
$r(x_6|PPC_n) = r(x_2|PPC_{n-1})$

where $r(x|G) = \sum_{y \in V(G)} r(x,y)$, and

$$\begin{aligned} Kf(PPC_n) &= Kf(PPC_{n-1}) + 6r(u_{n-1}|PPC_{n-1}) + 71n - 36 - \frac{1}{2}\sum_{i=1}^{6}\sum_{j=1}^{6} r(v_i,v_j) \\ &= Kf(PPC_{n-1}) + 6r(u_{n-1}|PPC_{n-1}) + 71n - 36 - \frac{35}{2} \end{aligned}$$

Then

$$Kf(PPC_{n+1}) = Kf(PPC_n) + 6r(u_n|PPC_n) + 71n + \frac{35}{2} \tag{2.1}$$

For a random polyphenyl chain $PPC(n,p_1,p_2)$, the resistance number $r(u_n|PPC(n, p_1,p_2))$ is a random variable, and its expected value is denoted by

$$U_n = E(r(u_n|PPC(n,p_1,p_2))).$$

By the expectation operator and (1), we can obtain a recursive relation for the expected value of the Kirchhoff number of a random polyphenyl chain $PPC(n,p_1,p_2)$

$$E(Kf(PPC(n+1,p_1,p_2)) = E(Kf(PPC(n,p_1,p_2))) + 6U_n + 71n + \frac{35}{2} \tag{2.2}$$

Now, we consider computing U_n.

(i) If $PPC_n \to PPC_{n+1}^1$ with probability p_1, then u_n coincides with the vertex x_3 or x_5. Consequently, $r(u_n|PPC_n)$ is given by $r(x_3|PPC_n)$ with probability p_1.

(ii) If $PPC_n \to PPC_{n+1}^2$ with probability p_2, then u_n coincides with the vertex x_2 or x_6. Consequently, $r(u_n|PPC_n)$ is given $r(x_2|PPC_n)$ with probability p_2.

(iii) If $PPC_n \to PPC_{n+1}^3$ with probability $1-p_1-p_2$, then u_n coincides with the vertex x_4. Consequently, $r(u_n|PPC_n)$ is given by $r(x_4|PPC_n)$ with probability $1-p_1-p_2$.

From (i)-(iii) above, we immediately obtain

$$\begin{aligned}U_n =& r(x_3|PPC_n)p_1 + r(x_2|PPC_n)p_2 + r(x_4|PPC_n)(1-p_1-p_2)\\ =& p_1[r(u_{n-1}|PPC(n-1,p_1,p_2)) + 14(n-1) + \frac{35}{6}]\\ &+ p_2[r(u_{n-1}|PPC(n-1,p_1,p_2)) + 11(n-1) + \frac{35}{6}]\\ &+ (1-p_1-p_2)[r(u_{n-1}|PPC(n-1,p_1,p_2)) + 15(n-1) + \frac{35}{6}]\end{aligned}$$

By applying the expectation operator to the above equation, we obtain

$$U_n = U_{n-1} + (15 - p_1 - 4p_2)n + p_1 + 4p_2 - \frac{55}{6}$$

And $U_1 = E(r(u_1|PPC(1,p_1,p_2))) = \frac{35}{6}$, using the above recurrence relation, we have

$$U_n = \frac{(15-p_1-4p_2)}{2}n^2 + (\frac{p_1}{2} + 2p_2 - \frac{5}{3})n$$

From (2),

$$\begin{aligned}&E(Kf(PPC(n+1,p_1,p_2))\\ &= E(Kf(PPC(n,p_1,p_2))) + 6[\tfrac{(15-p_1-4p_2)}{2}n^2 + (\tfrac{p_1}{2} + 2p_2 - \tfrac{5}{3})n] + 71n + \tfrac{35}{2}\\ &= E(Kf(PPC(n,p_1,p_2))) + (45 - 3p_1 - 12p_2)n^2 + (3p_1 + 12p_2 + 61)n + \tfrac{35}{2}\end{aligned}$$

and $E(Kf(PPC(1,p_1,p_2))) = \frac{35}{2}$.

Using the above recurrence relation, we have

$$E(Kf(PPC(n,p_1,p_2))) = (15-p_1-4p_2)n^3 + (3p_1+12p_2+8)n^2 - (2p_1+8p_2+\frac{11}{2})n.$$

□

Specially, by taking $(p_1,p_2) = (1,0),(0,1)$ or $(0,0)$, respectively, and Theorem 2.1, we have

Corollary 2.2. *([8]) The Kirchhoff indices of the polyphenyl meta-chain $\overline{M_n}$, the polyphenyl ortho-chain $\overline{O_n}$ and the polyphenyl para-chain $\overline{P_n}$ are*

$$Kf(\overline{M_n}) = 14n^3 + 11n^2 - \frac{15}{2}n$$

$$Kf(\overline{O_n}) = 11n^3 + 20n^2 - \frac{27}{2}n$$

$$Kf(\overline{P_n}) = 15n^3 + 8n^2 - \frac{11}{2}n$$

2.2 The Kirchhoff index of the random spiro chain

In this section, we will consider the Kirchhoff index of the random spiro chain.

Theorem 2.3. *For $n \geq 1$, the expected value of the Kirchhoff index of the random spiro chain $SPC(n, p_1, p_2)$ is*

$$E(Kf(SPC(n,p_1,p_2))) = (\frac{25}{4} - \frac{25}{36}p_1 - \frac{25}{9}p_2)n^3 + (\frac{25}{12}p_1 + \frac{25}{3}p_2 + \frac{125}{12})n^2 - (\frac{25}{18}p_1 + \frac{50}{9}p_2 - \frac{5}{6})n.$$

Proof. Note that the spiro chain SPC_n is obtained by attaching SPC_{n-1} a new terminal hexagon, we suppose that the terminal hexagon is spanned by vertices x_1, x_2, x_3, x_4, x_5, x_6, and the vertex x_1 is u_{n-1} (see Fig.3). Then

(i) For any $v \in SPC_{n-1}$,
$r(x_1, v) = r(u_{n-1}, v), r(x_2, v) = r(u_{n-1}, v) + \frac{5}{6}$,
$r(x_3, v) = r(u_{n-1}, v) + \frac{4}{3}, r(x_4, v) = r(u_{n-1}, v) + \frac{3}{2}$,
$r(x_5, v) = r(u_{n-1}, v) + \frac{4}{3}, r(x_6, v) = r(u_{n-1}, v) + \frac{5}{6}$;
(ii) SPC_{n-1} has $5(n-1)+1$ vertices;
(iii) For $k \in \{1,2,3,4,5,6\}$, $\sum_{i=1}^{6} r(x_k, x_i) = \frac{35}{6}$.

So, we have
$r(x_1|SPC_n) = r(u_{n-1}|SPC_{n-1}) + \frac{35}{6}$
$r(x_2|SPC_n) = r(u_{n-1}|SPC_{n-1}) + \frac{5}{6} \times (5n-4) + \frac{5}{6} + \frac{4}{3} + \frac{3}{2} + \frac{4}{3} = r(u_{n-1}|SPC_{n-1}) + \frac{25}{6} \times (n-1) + \frac{35}{6}$
$r(x_3|SPC_n) = r(u_{n-1}|SPC_{n-1}) + \frac{20}{3} \times (n-1) + \frac{35}{6}$
$r(x_4|SPC_n) = r(u_{n-1}|SPC_{n-1}) + \frac{15}{2} \times (n-1) + \frac{35}{6}$
$r(x_5|SPC_n) = r(x_3|SPC_{n-1})$
$r(x_6|SPC_n) = r(x_2|SPC_{n-1})$
where $r(x|G) = \sum_{y \in V(G)} r(x,y)$, and

$$\begin{aligned} Kf(SPC_n) &= Kf(SPC_{n-1}) + 5r(u_{n-1}|SPC_{n-1}) + \\ &\quad \frac{175(n-1)}{6} + 35 - \frac{1}{2}\sum_{i=1}^{6}\sum_{j=1}^{6} r(v_i, v_j) \\ &= Kf(SPC_{n-1}) + 5r(u_{n-1}|SPC_{n-1}) + \frac{175n}{6} - \frac{35}{3} \end{aligned} \tag{2.3}$$

Then

$$Kf(SPC_{n+1}) = Kf(SPC_n) + 5r(u_n|SPC_n) + \frac{175n}{6} + \frac{35}{2} \tag{2.4}$$

For a random spiro chain $SPC(n, p_1, p_2)$, the resistance number $r(u_n|SPC(n, p_1, p_2))$ is a random variable, and its expected value is denoted by

$$U_n = E(r(u_n|SPC(n, p_1, p_2))).$$

By the expectation operator and (3), we can obtain a recursive relation for the expected value of the Kirchhoff number of a random spiro chain $SPC(n, p_1, p_2)$

$$E(Kf(SPC(n+1, p_1, p_2)) = E(Kf(SPC(n, p_1, p_2))) + 5U_n + \frac{175n}{6} + \frac{35}{2} \tag{2.5}$$

Now, we consider computing U_n.

(i) If $SPC_n \to SPC^1_{n+1}$ with probability p_1, then u_n is the vertex x_3 or x_5. Consequently, $r(u_n|SPC_n)$ is given by $r(x_3|SPC_n)$ with probability p_1.

(ii) If $SPC_n \to SPC^2_{n+1}$ with probability p_2, then u_n is the vertex x_2 or x_6. Consequently, $r(u_n|SPC_n)$ is given $r(x_2|SPC_n)$ with probability p_2.

(iii) If $SPC_n \to SPC^3_{n+1}$ with probability $1-p_1-p_2$, then u_n is the vertex x_4. Consequently, $r(u_n|SPC_n)$ is given by $r(x_4|SPC_n)$ with probability $1-p_1-p_2$.

From (i)-(iii) above, we immediately obtain

$$\begin{aligned}U_n =&r(x_3|SPC_n)p_1 + r(x_2|SPC_n)p_2 + r(x_4|SPC_n)(1-p_1-p_2)\\ =&p_1[r(u_{n-1}|SPC(n-1,p_1,p_2)) + \frac{20}{3}(n-1) + \frac{35}{6}]\\ &+ p_2[r(u_{n-1}|SPC(n-1,p_1,p_2)) + \frac{25}{6}(n-1) + \frac{35}{6}]\\ &+ (1-p_1-p_2)[r(u_{n-1}|SPC(n-1,p_1,p_2)) + \frac{15}{2}(n-1) + \frac{35}{6}]\end{aligned}$$

By applying the expectation operator to the above equation, we obtain

$$U_n = U_{n-1} + (\frac{15}{2} - \frac{5}{6}p_1 - \frac{10}{3}p_2)n + \frac{5}{6}p_1 + \frac{10}{3}p_2 - \frac{5}{3}$$

And $U_1 = E(r(u_1|SPC(1,p_1,p_2))) = \frac{35}{6}$, using the above recurrence relation, we have

$$U_n = (\frac{15}{4} - \frac{5}{12}p_1 - \frac{5}{3}p_2)n^2 + (\frac{25}{12} + \frac{5}{12}p_1 + \frac{5}{3}p_2)n$$

From (4),

$$\begin{aligned}&E(Kf(SPC(n+1,p_1,p_2)) =\\ &= E(Kf(SPC(n,p_1,p_2))) + 5[(\tfrac{15}{4} - \tfrac{5}{12}p_1 - \tfrac{5}{3}p_2)n^2 +\\ &\quad (\tfrac{25}{12} + \tfrac{5}{12}p_1 + \tfrac{5}{3}p_2)n] + \tfrac{175}{6}n + \tfrac{35}{2}\end{aligned}$$

and $E(Kf(SPC(1,p_1,p_2))) = \frac{35}{2}$.

Using the above recurrence relation, we have

$$\begin{aligned}E(Kf(SPC(n,p_1,p_2))) = (\frac{25}{4} - \frac{25}{36}p_1 - \frac{25}{9}p_2)n^3 + (\frac{25}{12}p_1 + \frac{25}{3}p_2 + \frac{125}{12})n^2\\ -(\frac{25}{18}p_1 + \frac{50}{9}p_2 - \frac{5}{6})n.\end{aligned}$$

□

Specially, by taking $(p_1,p_2) = (1,0),(0,1)$ or $(0,0)$, respectively, and Theorem 2.3, we have

Corollary 2.4. *([8]) The Kirchhoff indices of the spiro meta-chain M_n, the spiro ortho-chain O_n and the spiro para-chain P_n are*

$$\begin{aligned}Kf(M_n) &= \frac{50}{9}n^3 + \frac{25}{2}n^2 - \frac{5}{9}n\\ Kf(O_n) &= \frac{125}{36}n^3 + \frac{75}{4}n^2 - \frac{85}{18}n\\ Kf(P_n) &= \frac{25}{4}n^3 + \frac{125}{12}n^2 + \frac{5}{6}n.\end{aligned}$$

2.3 A relation between $E(Kf(PPC))$ and $E(Kf(SPC))$

Since a spiro chain can be obtained from a polyphenyl chain by squeezing off its cut edges, it is straightforward by Rayleigh short-cut principle in the classical theory of electricity that the Kirchhoff index of the spiro chain is less than the polyphenyl chain. In fact, a relation between the Kirchhoff indices of a polyphenyl chain and its corresponding spiro chain obtained by squeezing off its cut edges was given in [8]. Here, we can also obtain a relation between the expected values of their Kirchhoff indices of the random polyphenyl chain $PPC(n, p_1, p_2)$ and the random spiro chain $SPC(n, p_1, p_2)$ with the same probabilities p_1 and p_2 from Theorems 2.1 and 2.3.

Theorem 2.5. *For a random polyphenyl chain $PPC(n, p_1, p_2)$ and a random spiro chain $SPC(n, p_1, p_2)$ with n hexagons, the expected values of their Kirchhoff indices are related as*

$$50E(Kf(PPC(n, p_1, p_2))) = 72E(Kf(SPC(n, p_1, p_2))) + 300n^3 - 350n^2 - 335n.$$

Theorem 2.5 also shows that the expected value of Kirchhoff index of the random spiro chain is less than the random polyphenyl chain. In fact, for $n \geq 2$, $E(Kf(SPC(n, p_1, p_2))) < \frac{25}{36}E(Kf(PPC(n, p_1, p_2)))$. The reason is quite obvious. Dividing both sides of the equation in Theorem 2.5 yields

$$E(Kf(PPC(n, p_1, p_2))) = \frac{36}{25}E(Kf(SPC(n, p_1, p_2))) + 6n^3 - 7n^2 - \frac{67}{10}n$$

and it is easily seen that for $n \geq 2$, $6n^3 - 7n^2 - \frac{67}{10}n > 0$.

2.4 The average value of the Kirchhoff index

Let $\overline{\mathcal{G}}_n$ is the set of all polyphenyl chains with n hexagons. The average value of the Kirchhoff indices with respect to $\overline{\mathcal{G}}_n$ is

$$Kf_{avr}(\overline{\mathcal{G}}_n) = \frac{1}{|\overline{\mathcal{G}}_n|} \sum_{G \in \overline{\mathcal{G}}_n} Kf(G).$$

In order to obtain the average value of the Kirchhoff indices with respect to $\overline{\mathcal{G}}_n$, we only need to take $p_1 = p_2 = \frac{1}{3}$ in the random polyphenyl chain $PPC(n, p_1, p_2)$, i.e., the average value of the Kirchhoff indices with respect to $\overline{\mathcal{G}}_n$ is just the expected value of the Kirchhoff index of the random polyphenyl chain $PPC(n, p_1, p_2)$ for $p_1 = p_2 = \frac{1}{3}$. From Theorem 2.1, we have

Theorem 2.6. *The average value of the Kirchhoff indices with respect to $\overline{\mathcal{G}}_n$ is*

$$Kf_{avr}(\overline{\mathcal{G}}_n) = \frac{40}{3}n^3 + 13n^2 - \frac{53}{6}n.$$

Similarly, let $\mathcal{G}_n$ is the set of all spiro chains with n hexagons. The average value of the Kirchhoff indices with respect to $\mathcal{G}_n$ is

$$Kf_{avr}(\mathcal{G}_n) = \frac{1}{|\mathcal{G}_n|} \sum_{G \in \mathcal{G}_n} Kf(G).$$

And the average value of the Kirchhoff indices with respect to $\mathcal{G}_n$ is just the the expected value of the Kirchhoff index of the random spiro chain $SPC(n, p_1, p_2)$ for $p_1 = p_2 = \frac{1}{3}$. From Theorem 2.3, we have

Theorem 2.7. *The average value of the Kirchhoff indices with respect to $\mathcal{G}_n$ is*

$$Kf_{avr}(\mathcal{G}_n) = \frac{275}{54}n^3 + \frac{125}{9}n^2 - \frac{40}{27}n.$$

Acknowledgments. The authors would like to thank the anonymous referees for their valuable comments and suggestions on the original manuscript.

References

[1] D. Babić, D. J. Klein, I. Lukovits, S. Nikolić, N. Trinajstić, Resistance- distance matrix: A computational algorithm and its application, *Int. J. Quantum Chem.* **90** (2002), 166–176.

[2] Y. Bai, B. Zhao, P. Zhao, Extremal Merrifield-Simmons index and Hosoya index of polyphenyl chains, *MATCH Commun. Math. Comput. Chem.* **62** (2009), 649–656.

[3] R. B. Bapat, I. Gutman, W. J. Xiao, A simple method for computing resistance distance, *Z. Naturforsch* **58a** (2003), 494–498.

[4] D. Bonchev, E. J. Markel, A. H. Dekmezian, Long chain branch polymer chain dimensions: application of topology to the Zimm-Stockmayer model, *Polymer* **43** (2002), 203–222.

[5] X. Chen, B. Zhao, P. Zhao, Six-membered ring spiro chains with extremal Merrifild-Simmons index and Hosoya index, *MATCH Commun. Math. Comput. Chem.* **62** (2009), 657–665.

[6] H. Deng, On the minimal Kirchhoff indices of graphs with a given number of cut edges, *MATCH Commun. Math. Comput. Chem.* **63** (2010), 171–180.

[7] H. Deng, Wiener indices of spiro and polyphenyl hexagonal chains, *Mathematical and Computer Modelling* **55** (2012), 634–644.

[8] H. Deng, Z. Tang, Kirchhoff indices of spiro and polyphenyl hexagonal chains, accepted by *Util. Math.*

[9] T. Došlić, F. Måløy, Chain hexagonal cacti: Matchings and independent sets, *Discrete Math.* **310** (2010), 1676–1690.

[10] T. Došlić, M. S. Litz, Matchings and independent sets in polyphenylene chains, *MATCH Commun. Math. Comput. Chem.* **67** (2012), 313–330.

[11] D. R. Flower, On the properties of bit string-based measures of chemical similarity, *J. Chem. Inf. Comput. Sci.* **38** (1998), 379–386.

[12] Q. Guo, H. Deng, The extremal Kirchhoff index of a class of unicyclic graphs, *MATCH Commun. Math. Comput. Chem.* **61** (2009), 713–722.

[13] D. J. Klein, M. Randić, Resistance distance, *J. Math. Chem.* **12** (1993), 81–95.

[14] D. J. Klein, Graph geometry, graph metrics and Wiener, *MATCH Commun. Math. Comput. Chem.* **35** (1997), 7–27.

[15] D. J. Klein, H. Y. Zhu, Distances and volumina for graphs, *J. Math. Chem.* **23** (1998), 179–195.

[16] D. J. Klein, Resistance-distance sum rules, *Croat. Chem. Acta* **75** (2002), 633–649.

[17] Q. R. Li, Q. Yang, H. Yin, S. Yang, Analysis of by-products from improved Ullmann reaction using TOFMS and GCTOFMS, *J. Univ. Sci. Technol. China* **34** (2004), 335–341.

[18] J. L. Palacios, Closed-form formulas for Kirchhoff index, *Int. J. Quantum Chem.* **81** (2001), 135–140.

[19] J. L. Palacios, Resistance distance in graphs and random walks, *Int. J. Quantum Chem.* **81** (2001), 29–33.

[20] S. Tepavčević, A. T. Wroble, M. Bissen, D. J. Wallace, Y. Choi, L. Hanley, Photoemission studies of polythiophene and polyphenyl films produced via surface polymerization by ion-assisted deposition, *J. Phys. Chem. B* **109** (2005), 7134–7140.

[21] Y. J. Yang, X. Y. Jiang, Unicyclic graphs with extremal Kirchhoff index, *MATCH Commun. Math Comput. Chem.* **60** (2008), 107–120.

[22] Y. Yang, H. Zhang, Kirchhoff index of linear hexagonal chains, *Int. J. Quantum Chem.* **108** (2008), 503–512.

[23] W. Yang, F. Zhang, Wiener index in random polyphenyl chains, *MATCH Commun. Math. Comput. Chem.* **68** (2012), 371–376.

[24] H. Zhang, Y. Yang, Resistance distance and Kirchhoff index in circulant graphs, *Int. J. Quantum Chem.* **107** (2007), 330–339.

[25] W. Zhang, H. Deng, The second maximal and minimal Kirchhoff indices of unicyclic graphs, *MATCH Commun. Math. Comput. Chem.* **61** (2009), 683–695.

[26] H. Zhang, X. Jiang, Y. Yang, Bicyclic graphs with extremal Kirchhoff index, *MATCH Commun. Math. Comput. Chem.* **61** (2009), 697–712.

[27] P. Zhao, B. Zhao, X. Chen, Y. Bai, Two classes of chains with maximal and minimal total π-electron energy, *MATCH Commun. Math. Comput. Chem.* **62** (2009), 525–536.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 209–222

Extending patches to fullerenes

Christina Graves
The University of Texas at Tyler, Department of Mathematics
Tyler, TX 75799, USA

Jennifer McLoud-Mann
University of Washington Bothell, School of STEM
Bothell, WA 98011, USA

Kristen Stagg Rovira
The University of Texas at Tyler, Department of Mathematics
Tyler, TX 75799, USA

Received 10 May 2013, accepted 7 June 2014, published online 8 December 2014

Abstract

In this paper we consider fullerene patches that can be extended to pseudconvex patches. We show that all fullerene disks with three or fewer pentagons can be extended to pseudoconvex patches, and that all pseudoconvex patches can be extended to fullerenes.

Keywords: Fullerenes, fullerene patches, boundary codes, pseudoconvex patches.

Math. Subj. Class.: 05C10, 05C75, 92E10

1 Preliminaries

A *fullerene* is a trivalent planar graph whose faces consist solely of hexagons and pentagons. A *fullerene patch*, or *patch*, is similar; it is a planar graph where all faces are hexagons and pentagons except one outer face, with vertices not on the outer face having degree 3 and vertices on the outer face having degree 2 or 3. An easy way to create a fullerene patch from a fullerene is to trace a closed circuit on a fullerene and delete all faces on one side of the circuit. However, it is not so easy to create a fullerene from a fullerene patch. In fact, there are many fullerene patches that cannot be extended to a fullerene as we will see later. We do find a family of patches that can be extended to fullerenes.

E-mail addresses: cgraves@uttyler.edu (Christina Graves), jmcloud@uw.edu (Jennifer McLoud-Mann), kstagg@uttyler.edu (Kristen Stagg Rovira)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

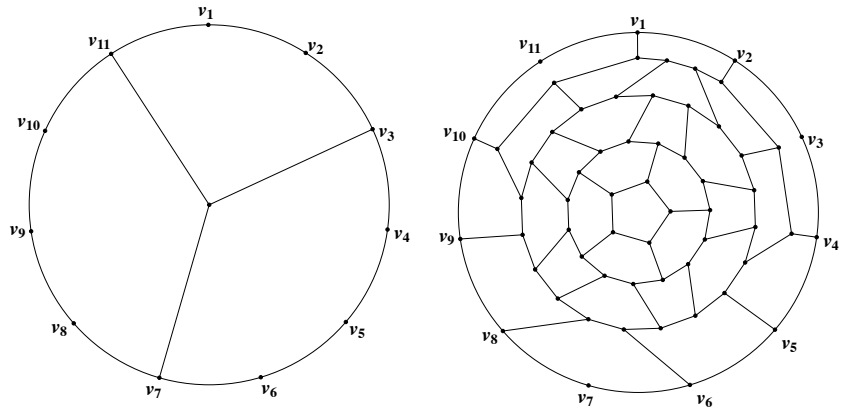


Figure 1: A patch and its complement. The patch on the left has boundary code $223(2223)^2$ and side parameters $[1, 0, 1, 0, 1]$ and the patch on the right has boundary code $332(3332)^2$.

An obvious characteristic of interest of a fullerene patch is the boundary. The *boundary code* of a patch is a sequence of 2's and 3's corresponding to the valences of the vertices on the outer face listed in cyclic order. It does not matter which vertex we start with nor which direction we travel around the patch; hence we make no distinction between a boundary code and its cyclic permutations or inverse. Given a boundary code $\mathcal{S}$, the *complement* of $\mathcal{S}$, denoted $\mathcal{S}^c$ is a sequence of 2's and 3's with a 2 every place $\mathcal{S}$ has a 3, and a 3 every place $\mathcal{S}$ has a 2.

To determine if a fullerene patch extends to a fullerene, we need to consider the following question: Given a fullerene patch Π with boundary code $\mathcal{S}$, does there exist a fullerene patch Π^c with boundary code $\mathcal{S}^c$? If Π^c exists, we can identify the vertices and edges on the boundary of Π with the corresponding vertices and edges on the boundary of Π^c to get a fullerene (see Figure 1).

The general question of interest is known in the literature as the *PentHex Puzzle*: Given a sequence of 2's and 3's, does there exist a fullerene patch with that sequence as its boundary code? Some variations on this question were explored in [3], [1], [6], and [4] among other places.

For large patches, the boundary code can be unwieldy to work with. Generalizing the definitions from [8], [7], [5] and [2], we define the following.

Definition 1.1.

1. A *break edge* is an edge on the boundary whose endpoints are both of degree two.
2. A *bend edge* is an edge on the boundary whose endpoints are both of degree three.
3. A *side* of a patch is a path on the boundary between a consecutive pair of break edges, including the break edges. The *length* of a side is the number of degree three vertices on the side.

4. A *straight side* is a side with no bend edges.

5. A *bent side* is a side containing at least one bend edge. A *straight segment* of a bent side is a subpath of a bent side between either a break edge and the closest bend edge, or between two consecutive bend edges.

We can now think of the boundary of a patch as sections of straight sides and straight segments connected by break edges or bend edges. A patch with no bend edges, called a *pseudoconvex patch*, has only straight sides and thus has boundary code $2(23)^{\ell_1}2(23)^{\ell_2}\cdots 2(23)^{\ell_s}$. Rather than writing a sequence of 2's and 3's to describe our boundary, we describe the boundary by the lengths of the straight sides $[\ell_1, \ell_2, \ldots, \ell_s]$, called *side parameters* of the patch.

If the patch has bend edges, we would still like to consolidate the information found in the boundary code. A bent side with consecutive straight segments of lengths $a_1, a_2, \ldots a_t$ can be described by $(a_1, \ldots, a_t)$. Thus, a patch with five straight sides of length 5 followed by a bent side with five straight segments of length 1 (see Figure 3) can be described by the side parameters $[5, 5, 5, 5, 5, (1, 1, 1, 1, 1)]$.

A patch with no break edges has zero sides and thus cannot be described by side parameters. If such a patch also has no bend edges, then we describe it by its boundary code $(23)^{\ell_0}$ where ℓ_0 is the number of faces on the boundary. If a patch has no break edges but does have bend edges, then we describe it by its boundary code $3(32)^{a_1-1}3(32)^{a_2-1}\cdots 3(32)^{a_t-1}$ where $a_1, a_2, \ldots, a_t$ represent the lengths of the straight segments in cyclic order.

The advantage of describing a patch by its side parameters rather than its boundary code is that this notation makes it effortless to find the number of sides (and hence break edges) of a patch. In fact, we can also tell exactly how many pentagons a patch must have by using Euler's formula.

Lemma 1.2. *In a fullerene patch, the number of break edges s, the number of bend edges e, and the number of pentagons p are related by*

$$p = 6 - s + e.$$

Proof. It is well-known (see for instance [1]) that the number of pentagons in a patch is equal to $6 - d_2 + d_3$ where d_i is the number of degree i vertices on the boundary. By definition, the number of degree 2 vertices and the number of degree three vertices on the boundary are the same except on a break edge or a bend edge. Each break edge increases the number of degree 2 vertices by 1, and each bend edge increases the number of degree 3 vertices by 1. □

In this paper, we investigate two specific types of fullerene patches: pseudoconvex patches and *fullerene disks* or simply *disks*. A disk of radius r is a fullerene patch with a central face, and every face on the boundary is distance (measured as graph distance in the dual) exactly r from the central face.

Our ultimate goal is to find a family of patches that can be extended to fullerenes. We begin by showing that a pseudoconvex patch can be extended to a fullerene. We then show that all disks with three or fewer pentagons can be extended to fullerenes.

2 Pseudoconvex Patches

Because a pseudoconvex patch has no bend edges, its boundary can be described by its side parameters $[\ell_1, \cdots, \ell_s]$ if it has $s \geq 1$ sides or by the boundary code $(23)^{\ell_0}$ if it has zero sides. The relationships between the lengths of these sides were explored in detail and summarized in Theorem 3.6 of [8]. We include a weaker lemma here.

Lemma 2.1. *The side parameters $[\ell_1, \cdots, \ell_s]$ of a pseudoconvex patch, excluding the patch with side parameters $[0, 0, 0, 0, 0]$, satisfy the inequality $\ell_1 + \cdots + \ell_s \geq 6 - s$ for $1 \leq s \leq 6$.*

Proof. If $s = 1$ or $s = 2$, the result follows directly from [8]. If $s = 3$, the side parameters are not of the form $[0, 1, k]$ or $[0, 0, k]$ for $k \geq 0$ so the result holds. In the case with $s = 4$ sides, at least two of the side lengths are non-zero and the result holds. Finally, if $s = 5$ or $s = 6$, all parameters are nonnegative so the sum is nonnegative. Thus the inequality is satisfied except in the patch consisting of a single pentagon with side parameters $[0, 0, 0, 0, 0]$. □

To extend a pseudoconvex patch to a fullerene, we need to find a complement for each pseudoconvex patch. The following lemma gives a constructive method for finding the complement patch.

Lemma 2.2. *There exists a patch with boundary code*

$$3(32)^{\ell_1}3(32)^{\ell_2}\cdots 3(32)^{\ell_s}$$

if $\ell_1 + \ell_2 + \cdots + \ell_s \geq 6 - s$ and $1 \leq s \leq 6$.

Proof. Start with the patch having boundary code $2(23)^5$ (i.e. one side of length 5) as shown in Figure 2. Add $\ell_1 + \ell_2 + \cdots + \ell_s + s - 6$ rings of hexagons to this patch to create a patch with side parameters $[\ell_1 + \cdots + \ell_s + s - 1]$. Next, add a pentagon to the break edge and add hexagons on the boundary everywhere else to yield a patch with boundary code $(23)^{\ell_1+\cdots+\ell_s+s}$.

We now add hexagons and pentagons to the boundary in the following way. First put a pentagon somewhere on the boundary. Moving clockwise around the boundary, place ℓ_1 hexagons followed by a pentagon; then place ℓ_2 hexagons followed by a pentagon, and continue this process until the patch has a completely new outer ring of hexagons and pentagons (see Figure 2). This constructed patch has the desired boundary. □

Theorem 2.3. *All pseudoconvex patches can be extended to fullerenes.*

Proof. For a pseudoconvex patch consisting of one pentagon with side parameters $[0, 0, 0, 0, 0]$, we use a stereographic projection of the dodecahedron onto the plane as the complement patch. For a pseudoconvex patch having side parameters $[\ell_1, \ldots, \ell_s]$ satisfying the condition

$$\ell_1 + \cdots + \ell_s \geq 6 - s,$$

we create the complement patch with boundary code $3(32)^{\ell_1}3(32)^{\ell_2}\cdots 3(32)^{\ell_s}$ as described in Lemma 2.2 and identify the boundaries to create a fullerene. Given a pseudoconvex patch with no sides and boundary code $(23)^{\ell_0}$, we create a second patch $(32)^{\ell_0}$ identical to the first and then identify corresponding edges and vertices appropriately to create a fullerene.

□

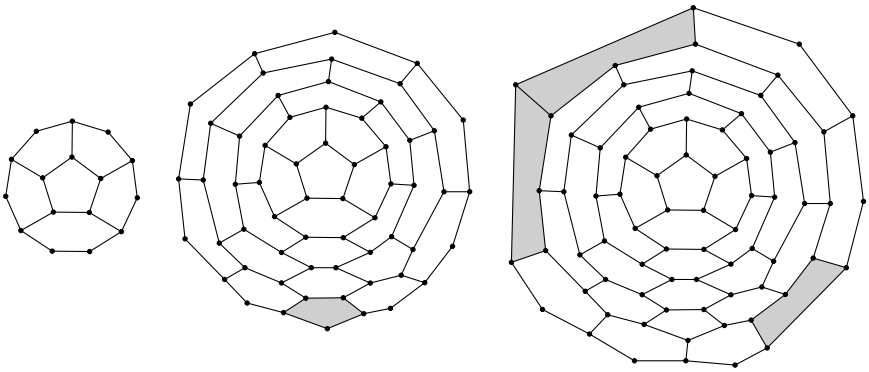


Figure 2: Creating a patch with boundary code $3(32)^2 3(32)^3 3(32)^0$. The $2(23)^5$ patch is shown on the left, the $(23)^8$ patch in the middle, and the desired patch on the right.

3 Disks

Because pseudoconvex patches extend to fullerenes, a patch that extends to a pseudoconvex patch also extends to a fullerene. One type of patch to explore is a fullerene disk. We will show that all disks with three or fewer pentagons extend to pseudoconvex patches; however, a disk with four or more pentagons may not extend to a fullerene. Consider a disk with four pentagons having side parameters $[5, 5, 5, 5, 5, (1, 1, 1, 1, 1)]$ as shown in Figure 3. If this patch could be extended to a fullerene, we would need to place a face adjacent to the four shaded pentagons. Such a face would have to have at least seven edges which is not allowed in a fullerene. Thus we restrict our attention to disks with three or less pentagons.

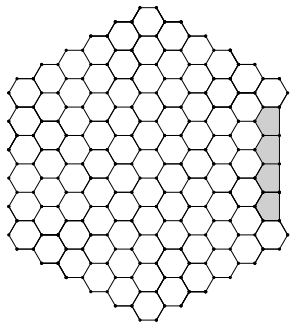


Figure 3: A disk with side parameters $[5, 5, 5, 5, 5, (1, 1, 1, 1, 1)]$. This patch cannot be extended to a fullerene.

Disks are a nice family of patches to study because all disks can be constructed by repeatedly adding layers of faces to a smaller disk. A *layer* is a collection of faces placed on the boundary of a disk of a radius r so that each new face is distance $r + 1$ from the

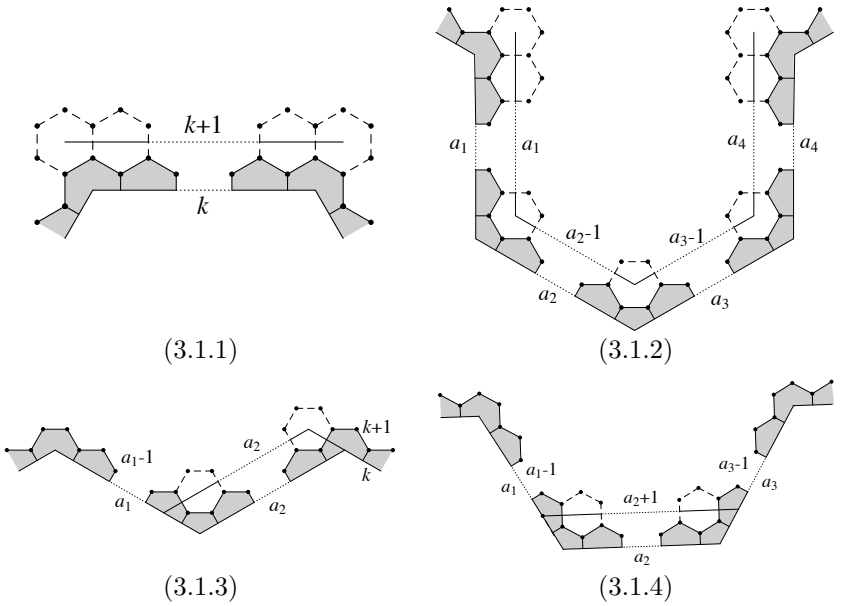


Figure 4: Depiction of Lemma 3.1 and how adding hexagons changes side parameters.

central face and and no faces from the original disk are on the new boundary. The word layer can be slightly misleading, though, because there can be faces added which are not on the boundary of the new patch. Figures 5, 6, and 7 show partial patches with one layer added.

Because many of the remaining arguments involve adding layers of hexagons to existing patches, it is convenient to discuss how such an action affects the side parameters. Adding a row of hexagons to a side of a patch means that each edge of the original boundary is incident with a newly added hexagon. Figure 4 and Lemma 3.1 demonstrate this action.

Lemma 3.1.

3.1.1. Adding a row of hexagons to the boundary of a straight side increases its length by 1.

3.1.2. Adding a partial layer of hexagons to a bent side with middle parameters larger than 1 keeps the outer parameters the same and decreases the middle parameters by 1.

3.1.3. Adding a row of hexagons to a straight segment a_2 of the bent side (a_1, a_2) with $a_1 > 1$, not including the break edge, decreases a_1 by 1, keeps a_2 the same, and increases the side following a_2 by 1.

3.1.4. Adding a row of hexagons to the straight segment a_2 of a bent side (a_1, a_2, a_3) with $a_1, a_3 > 1$, decreases a_1 and a_3 by 1 and increases a_2 by 1.

The lemma tells us how the side parameters of a patch change when we add a layer of hexagons for almost all situations. We will consider special cases for adding a layer of hexagons for patches having a bent side containing a middle parameter equal to one in later proofs.

The next three lemmas give the side parameters for all disks with three or fewer pentagons.

Lemma 3.2. *A disk with at most one pentagon has side parameters satisfying one of the following:*

A1. $[\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6]$

B1. $[\ell_1, \ell_2, \ell_3, \ell_4, \ell_5]$

B2. $[\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, (a_1, a_2)]$

Proof. A disk consisting only of hexagons has the form $[r, r, r, r, r, r]$. For a disk containing one pentagon on the boundary, there are two cases. Either the disk consists of only one face, a pentagon, and has side parameters $[0, 0, 0, 0, 0]$, or the disk can be constructed by adding faces to a disk containing no pentagons. Starting with a disk satisfying condition $A1$, adding a layer of faces with a pentagon on a break edge and hexagons everywhere else yields a $B1$ patch, and adding a layer with a pentagon on a straight side yields a $B2$ patch.

If a disk with one pentagon does not have its pentagon on the boundary, then the patch can be viewed as a disk containing one pentagon on the boundary with layers of hexagons added to it. Using Lemma 3.1, adding layers of hexagons to a $B1$ or $B2$ patch yields a $B1$ or $B2$ patch respectively. □

Lemma 3.3. *A disk with two pentagons has side parameters satisfying one of the following:*

C1. $[\ell_1, \ell_2, \ell_3, \ell_4]$

C2. $[\ell_1, \ell_2, \ell_3, \ell_4, (a_1, a_2)]$

C3. $[\ell_1, \ell_2, \ell_3, \ell_4, (a_1, a_2), (b_1, b_2)]$

C4. $[\ell_1, \ell_2, \ell_3, (a_1, a_2), \ell_4, (b_1, b_2)]$

C5. $[\ell_1, \ell_2, (a_1, a_2), \ell_3, \ell_4, (b_1, b_2)]$

C6. $[\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, (a_1, a_2, a_3)]$

Proof. We start by considering disks containing two pentagons with different distances from the central face, and the pentagon farthest from the central face on the boundary. Starting with a disk containing one pentagon satisfying condition $B1$, adding a layer of faces with a pentagon on a break edge yields a $C1$ patch, and adding a layer with a pentagon on a straight side would yield a $C2$ patch. Starting with a disk containing one pentagon satisfying condition $B2$, adding a layer of faces with a pentagon on a break edge results in a $C2$ patch, adding a layer with a pentagon on a straight side yields a $C3, C4,$ or $C5$ patch, and adding a layer with a pentagon on a straight segment of the bent side would yield a $C6$ patch.

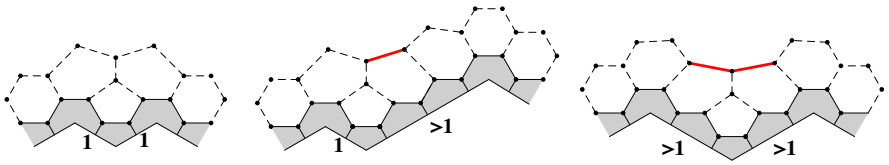


Figure 5: Adding a pentagon to a bend edge. The red edges are bend edges in the new patch.

When adding a layer of faces with a pentagon on the bend edge of a $B2$ patch, there are three cases to consider. If $a_1 = a_2 = 1$, the resulting patch is a $C1$ patch. If $1 = a_1 < a_2$, the new patch is a $C2$ patch. Lastly, if $a_1, a_2 > 1$, the resulting patch is a $C6$ patch, with the middle component of the bent edge having length 1 (see Figure 5).

Now consider disks where both pentagons are the same distance from the central face and are both on the boundary. There are no disks of radius 0 containing two pentagons, so we can construct these patches by starting with a disk containing no pentagons and adding the pentagons to the same layer. Starting with a patch with no pentagons satisfying condition $A1$, adding a layer with two pentagons on two different break edges yields a $C1$ patch, adding a layer with one pentagon on a break edge and one on a straight side would result in a $C2$ patch, adding a layer with two pentagons on two different straight sides would yield a $C3, C4,$ or $C5$ patch, and adding a layer with two pentagons on the same straight side would result in a $C6$ patch.

Notice every disk we have constructed has a pentagon on the boundary. To construct disks with two pentagons neither of which is on the boundary, we simply construct a disk with two pentagons at least one of which is on the boundary, and add layers of hexagons. By Lemma 3.1, adding layers of hexagons to a Ci patch yields a Ci patch for $1 \leq i \leq 5$. When considering a $C6$ patch, we assume without loss of generality that $a_1 \leq a_3$. Adding less than $a_2 + a_1 - 1$ layers of hexagons to a $C6$ patch results in a $C6$ patch, and adding at least $a_2 + a_1 - 1$ layers of hexagons to a $C6$ patch would yield a $C1$ or $C2$ patch. The $C1$ patch results if $a_1 = a_3$ and the $C2$ patch results otherwise. □

Lemma 3.4. *A disk with three pentagons has side parameters satisfying one of the following:*

D1. $[\ell_1, \ell_2, \ell_3]$

D2. $[\ell_1, \ell_2, \ell_3, (a_1, a_2)]$

D3. $[\ell_1, \ell_2, \ell_3, (a_1, a_2), (b_1, b_2)]$

D4. $[\ell_1, \ell_2, (a_1, a_2), \ell_3, (b_1, b_2)]$

D5. $[\ell_1, \ell_2, \ell_3, (a_1, a_2), (b_1, b_2), (c_1, c_2)]$

D6. $[\ell_1, \ell_2, (a_1, a_2), \ell_3, (b_1, b_2), (c_1, c_2)]$

D7. $[\ell_1, (a_1, a_2), \ell_2, (b_1, b_2), \ell_3, (c_1, c_2)]$

D8. $[\ell_1, \ell_2, \ell_3, \ell_4, (a_1, a_2, a_3)]$

D9. $[\ell_1, \ell_2, \ell_3, \ell_4, (a_1, a_2, a_3), (b_1, b_2)]$

D10. $[\ell_1, \ell_2, \ell_3, (a_1, a_2, a_3), \ell_4, (b_1, b_2)]$

D11. $[\ell_1, \ell_2, (a_1, a_2, a_3), \ell_3, \ell_4, (b_1, b_2)]$

D12. $[\ell_1, \ell_2, \ell_3, (a_1, a_2, a_3, a_4)]$

Proof. We begin by showing that that all disks with three pentagons having at least one pentagon on the boundary satisfy the criteria given. To show that all three-pentagon disks have the listed side parameters, we add layers of hexagons to three-pentagon disks with a pentagon on the boundary.

First, consider a three-pentagon disk with exactly one pentagon on the boundary. Such disks are created by starting with a two-pentagon disk and adding a layer of faces containing exactly one pentagon. Starting with a $C1$ patch, adding a layer of faces with a pentagon on a break edge yields a $D1$ patch, and adding a layer with a pentagon on a straight side yields a $D2$ patch. Starting with a $C2$ patch, adding a layer of faces with a pentagon on a break edge yields a $D2$ patch, and adding a layer of faces with a pentagon on a straight side yields a $D3$ or $D4$ patch. If a layer of faces is added to a $C2$ patch with a pentagon on a bend edge, then (following the proof of Lemma 3.3) a $D1$, $D2$, or $D8$ patch is obtained. If a layer is added with a pentagon on a straight straight segment of a bent side of a $C2$ patch, the resulting disk is a $D8$ patch.

Starting with a $C3 - C5$ patch, adding a layer of faces with a pentagon on a straight side yields a $D5 - D7$ patch, and adding a layer with a pentagon on a straight segment of a bent side yields a $D9$, $D10$, or $D11$ patch. Adding a layer of faces with a pentagon on a break edge of a $C3$ patch yields a $D3$, $D4$, or $D8$ patch, with a $D8$ patch occurring if the break edge is between the two bent sides. Adding a layer with a pentagon on a break edge of a $C4$ or $C5$ patch yields a $D3$ or $D4$ patch. Finally, adding a layer of faces with a pentagon on a bend edge on a $C3 - C5$ patch yields a $D2$, $D3$, $D4$, $D9$, $D10$, or $D11$ patch.

The last case involves adding a layer of faces with exactly one pentagon to a $C6$ patch. Adding a layer of faces with a pentagon on a break edge results in a $D8$ patch, adding a layer with a pentagon on a straight side yields a $D8 - D11$ patch, and adding a layer with a pentagon on a straight segment of a bent side results in a $D12$ patch. When adding a layer of faces with a pentagon on a bend edge of a $C6$ patch there are a few cases to consider. Without loss of generality, assume that the pentagon is added to the bend edge between the straight segment of length a_1 and the straight segment of length a_2. If $a_1 > 1$ and $a_2 > 2$ the resulting patch is a $D12$ patch with the new bent side having lengths $(a_1 - 1, 1, a_2 - 2, a_3)$ (see Figure 6). In the cases were either $a_1 = 1$ or $a_2 \in \{1, 2\}$, some care needs to be taken. In these cases, the length of a_3 can affect the type of new patch; however, in every case, the resulting patch is a $D2$ or $D8$ patch. Figure 7 shows these cases in full detail.

Now consider the three-pentagon disks with exactly two pentagons on the boundary. To create such disks, we start with either a $B1$ or $B2$ patch. Adding a layer of faces with two pentagons to a $B1$ patch is very similar to adding two pentagons to an $A1$ patch, which was explored in the proof of Lemma 3.3. Thus adding a layer with two pentagons on a $B1$ patch results in a $D1 - D4$ or $D8$ patch. Using similar arguments as before, adding a layer

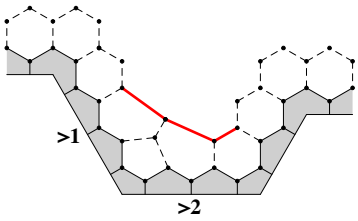


Figure 6: Adding a pentagon to a bend edge in a $C6$ patch. The red edges are bend edges in the new patch.

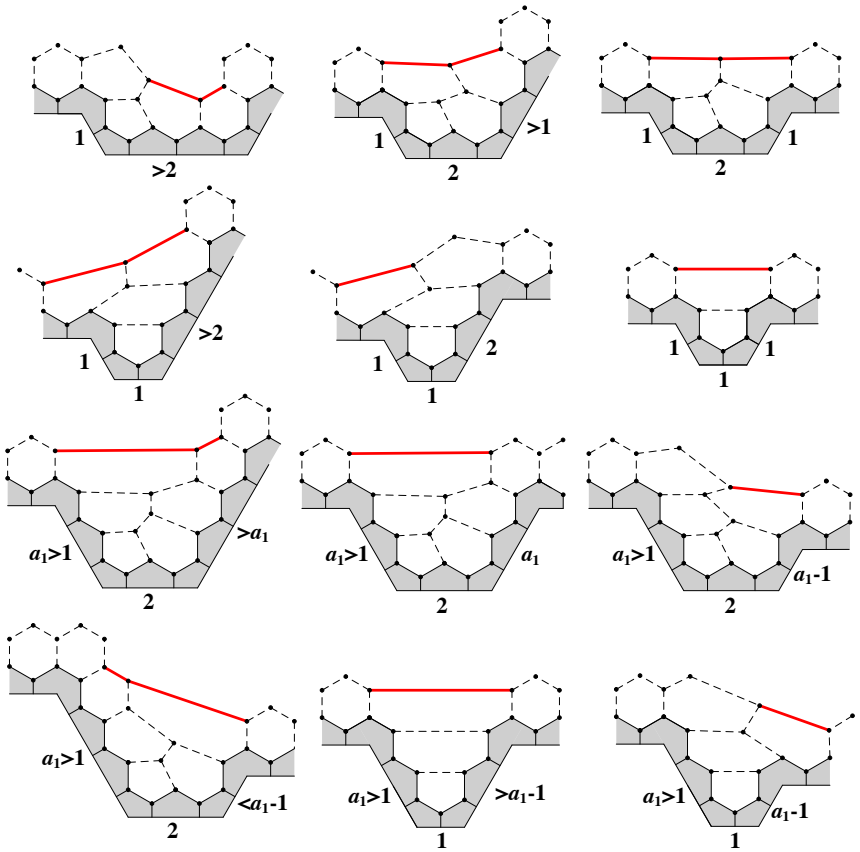


Figure 7: Adding a pentagon to a bend edge in a $C6$ patch. The red edges are bend edges in the new patch.

with two pentagons to a $B2$ patch results in a $D2 - D12$ patch if neither pentagon is on a bend edge. Adding a layer with one of the pentagons on a bend edge requires the cases seen in Figure 5. Adding a layer with one pentagon on the bend edge and one pentagon on a break edge yields a $D1$, $D2$, or $D8$ patch. Adding a layer with one pentagon on the bend edge and one pentagon on a straight side results in a $D2 - D4$ or $D9 - D11$ patch. Finally, adding a layer with one pentagon on the bend edge and one pentagon on a straight segment of the bent side (which can only be done if the straight segment is greater than 1) results in a $D8$ patch.

Three-pentagon disks with all three pentagons on the boundary are formed by adding a layer to a disk of hexagons. Starting with an $A1$ patch, adding a layer with three pentagons results in a $D1 - D11$ patch using similar arguments to those above.

Notice every patch we have constructed has a pentagon on the boundary. For those that do not, we simply construct a previous patch and add layers of hexagons. By Lemma 3.1, adding layer of hexagons to a Di patch yields a Di patch for $1 \leq i \leq 7$. For disks that have a bent side with at least three straight segments, the resulting patch's side parameters differ based on the number of layers added. For instance, adding one layer of hexagons to a patch with the side parameters $[\ell_1, \ell_2, \ell_3, \ell_4, (a_1, a_2, a_3)]$ and $a_2 > 1$ results in a patch with side parameters $[\ell_1 + 1, \ell_2 + 1, \ell_3 + 1, \ell_4 + 1, (a_1, a_2 - 1, a_3)]$. Furthermore, adding a layer of hexagons to a patch with side parameters $[\ell_1, \ell_2, \ell_3, \ell_4, (a_1, 1, a_3)]$ yields a patch with side parameters $[\ell_1 + 1, \ell_2 + 1, \ell_3 + 1, \ell_4 + 1, (a_1 - 1, 1, a_3 - 1)]$. Thus, adding less than $a_2 + \min\{a_1, a_3\} - 1$ rings to a Dj patch will yield a Dj patch for $8 \leq j \leq 11$.

Adding at least $a_2 + \min\{a_1, a_3\} - 1$ rings to a $D8$ patch yields a $D1$ patch if $a_1 = a_3$ or $D2$ patch otherwise. Adding at least $a_2 + \min\{a_1, a_3\} - 1$ rings to a $D9$ patch yields a $D2$ patch if $a_1 = a_3$, a $D3$ patch if $a_1 < a_3$, or a $D8$ patch if $a_1 > a_3$. Adding at least $a_2 + \min\{a_1, a_3\} - 1$ rings to a $D10$ or $D11$ patch yields a $D2$ patch if $a_1 = a_3$ or $D4$ patch otherwise.

Now let us consider adding rings of hexagons to a $D12$ patch; without loss of generality assume $a_2 \leq a_3$. Adding less than $a_2 + \min\{a_1 - 1, \lfloor \frac{a_3 - a_2}{2} \rfloor\}$ rings yields another $D12$ patch. Adding exactly $a_2 + \min\{a_1 - 1, \lfloor \frac{a_3 - a_2}{2} \rfloor\}$ rings yields a $D1$, $D2$, or $D8$ patch. From arguments above, adding further rings of hexagons gives a $D1$, $D2$, or $D8$ patch. □

Now that we have shown the side parameters of disks with three or fewer pentagons fall into different classes, we show that any patch with these side parameters can be extended to a pseudoconvex patch.

Lemma 3.5. *Any patch with one of the following descriptions of its side parameters can be extended to a pseudoconvex patch:*

1. $[\ell_1, \ell_2, \ldots, \ell_s]$
2. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2)]$
3. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2)]$
4. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), \ell_{s+1}, \ldots, \ell_t, (b_1, b_2)]$
5. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2), (c_1, c_2)]$
6. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2), \ell_{s+1}, \ldots, \ell_t, (c_1, c_2)]$
7. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), \ell_{s+1}, \ldots, \ell_t, (b_1, b_2), \ell_{t+1}, \ldots, \ell_u, (c_1, c_2)]$

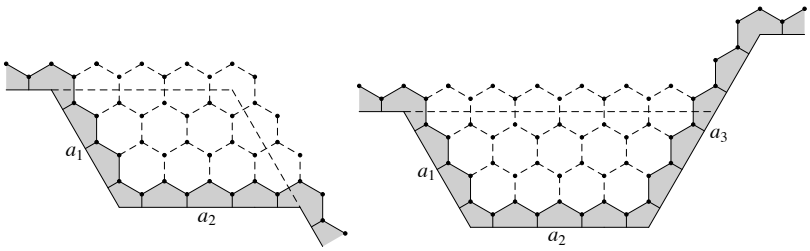


Figure 8: Filling in the bent side (a_1, a_2) and partially filling in the bent side (a_1, a_2, a_3).

8. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2, a_3)]$

9. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2, b_3)]$

10. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), \ell_{s+1}, \ldots, \ell_t, (b_1, b_2, b_3)]$

11. $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2, a_3, a_4)]$.

Proof. Let Π be a patch with one of the following descriptions above. If $\Pi = [\ell_1, \ell_2, \ldots, \ell_s]$, then it has no bend edges and is pseudoconvex. If $\Pi = [\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2)]$, then adding a_1 rows of a_2 hexagons to the straight segment a_2, not including the break edge, yields the pseudoconvex patch $[\ell_1 + a_1, \ell_2, \ldots, \ell_{s-1}, \ell_s + a_2]$ by Lemma 3.1 (see Figure 8).

If the side parameters of Π are of the form $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2)]$ or $[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), \ell_{s+1}, \ldots, \ell_t, (b_1, b_2)]$, then adding a_1 rows of a_2 hexagons to the straight segment a_2, not including the break edge, yields either

$$[\ell_1, \ell_2, \ldots, \ell_s + a_2, (b_1 + a_1, b_2)]$$

or

$$[\ell_1, \ell_2, \ldots, \ell_s + a_2, \ell_{s+1} + a_1, \ldots, \ell_t, (b_1, b_2)]$$

which are both extendable by a previous case. If Π has one of the forms

$$[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2), (c_1, c_2)],$$

$$[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2), \ell_{s+1}, \ldots, \ell_t, (c_1, c_2)],$$

or

$$[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), \ell_{s+1}, \ldots, \ell_t, (b_1, b_2), \ell_{t+1}, \ldots, \ell_u, (c_1, c_2)],$$

then adding a_1 rows of a_2 hexagons to the straight segment a_2, not including the break edge, yields a patch with one of the descriptions

$$[\ell_1, \ell_2, \ldots, \ell_s + a_2, (b_1 + a_1, b_2), (c_1, c_2)],$$

$$[\ell_1, \ell_2, \ldots, \ell_s + a_2, (b_1 + a_1, b_2), \ell_{s+1}, \ldots, \ell_t, (c_1, c_2)],$$

or

$$[\ell_1, \ell_2, \ldots, \ell_s + a_2, \ell_{s+1} + a_1, \ldots, \ell_t, (b_1, b_2), \ell_{t+1}, \ldots, \ell_u, (c_1, c_2)]$$

which are all extendable by previous cases.

Suppose $\Pi = [\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2, a_3)]$. Without loss of generality, assume $a_1 \leq a_3$. Adding $a_1 - 1$ rows of hexagons to the a_2 straight segment results in a patch with parameters

$$[\ell_1, \ell_2, \ldots, \ell_s, (1, a_1 + a_2 - 1, a_3 - a_1 + 1)]$$

by Lemma 3.1. Adding one more row yields a patch whose parameters are

$$[\ell_1, \ell_2, \ldots, \ell_{s-1}, (\ell_s + a_1 + a_2, a_3 - a_1)]$$

if $a_1 < a_3$ or

$$[\ell_s + a_1 + a_2 + \ell_1, \ell_2, \ell_3, \ldots, \ell_{s-1}]$$

if $a_1 = a_3$. Either of these is extendable by a previous case.

If Π has the form

$$[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), (b_1, b_2, b_3)]$$

or

$$[\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2), \ell_{s+1}, \ldots, \ell_t, (b_1, b_2, b_3)],$$

then adding a_1 rows of a_2 hexagons to the straight segment a_2, not including the break edge, yields a patch satisfying the form

$$[\ell_1, \ell_2, \ldots, \ell_s + a_2, (b_1 + a_1, b_2, b_3)]$$

or

$$[\ell_1, \ell_2, \ldots, \ell_s + a_2, \ell_{s+1} + a_1, \ldots, \ell_t, (b_1, b_2, b_3)]$$

which are both extendable by previous cases.

Suppose $\Pi = [\ell_1, \ell_2, \ldots, \ell_s, (a_1, a_2, a_3, a_4)]$. Without loss of generality, we may assume that $a_1 \leq a_4$. Note that $a_1 \leq a_3 + a_4$. We have the following four cases to consider: (i) $0 < a_1 < a_3$, (ii) $a_1 = a_3$, (iii) $a_3 < a_1 < a_3 + a_4$, and (iv) $a_1 = a_3 + a_4$. In each of the four cases we will add a_1 rows of hexagons to the straight segment a_2 in order to create a patch which is extendable by a previous case.

For (i) and (ii), we begin by adding $a_1 - 1$ rows of hexagons to create a patch with parameters $[\ell_1, \ell_2, \ldots, \ell_s, (1, a_1 + a_2 - 1, a_3 - a_1 + 1, a_4)]$ by Lemma 3.1. Adding one more row will yield

$$[\ell_1, \ell_2, \ldots, \ell_{s-1}, (\ell_s + a_1 + a_2, a_3 - a_1, a_4)]$$

in case (i) and

$$[\ell_1, \ell_2, \ldots, \ell_{s-1}, (\ell_s + a_1 + a_2 - 1, 1, a_4 - 1)]$$

in case (ii).

For (iii) and (iv), we begin by adding $a_3 - 1$ rows of hexagons to straight side a_2 to yield $[\ell_1, \ell_2, \ldots, \ell_s, (a_1 - a_3 + 1, a_2 + a_3 - 1, 1, a_4)]$ by Lemma 3.1. Adding another row yields a $[\ell_1, \ell_2, \ldots, \ell_s, (a_1 - a_3, a_2 + a_3 - 1, 1, a_4 - 1)]$ patch. For (iii), adding $a_1 - a_3 - 1$ rows gives

$$[\ell_1, \ell_2, \ldots, \ell_s, (1, a_2 + a_3 - 1, 1, a_4 + a_3 - a_1)]$$

and finally adding one more row gives a

$$[\ell_1, \ell_2, \ldots, \ell_{s-1}, (\ell_s + a_2 + a_3 - 1, 1, a_4 + a_3 - a_1 - 1)]$$

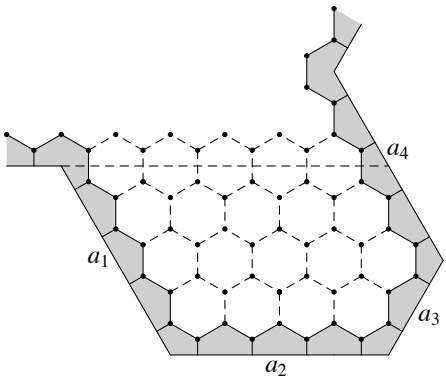


Figure 9: Partially filling in the bent side (a_1, a_2, a_3, a_4).

patch (see Figure 9). For (iv), adding $a_1 - a_3 - 2$ rows of hexagons gives $[\ell_1, \ell_2, \ldots, \ell_s, (2, a_2 + a_3 - 1, 1, 1)]$, an additional row yields

$$[\ell_2, \ell_3, \ldots, \ell_s, (1, a_2 + a_3, \ell_1)],$$

and finally adding another row gives $[\ell_2, \ell_3, \ldots, \ell_{s-1}, (\ell_s + a_2 + a_3, \ell_1)]$.

□

Theorem 3.6. *All disks containing at most three pentagons extend to a fullerene.*

Proof. All disks containing at most three pentagons will have side parameters satisfying the conditions of Lemma 3.2, Lemma 3.3, or Lemma 3.4. By Lemma 3.5, these disks can be extended to pseudoconvex patches. By Theorem 2.3, those pseudoconvex patches can be extended to fullerenes. □

References

[1] P. Bonsma and F. Breuer, *Finding fullerene patches in polynomial time.* In ISAAC 2009, volume 5878 of *LNCS* (2009), 750–759.

[2] J. Bornhoft, G. Brinkmann and J. Greinus, Pentagon-hexagon-patches with short boundaries, *European Journal of Combinatorics* **24** (2003), 517–529.

[3] G. Brinkmann and A. W. M. Dress, PentHex puzzles: A reliable and efficient top-down approach to fullerene-structure enumeration, *Advance in Applied Math* **21** (1998), 473–480.

[4] G. Brinkmann, U.V. Nathusius and A.H.R. Palser, A constructive enumeration of nanotube caps, *Discrete Applied Mathematcis* **116** (2002), 55–71.

[5] G. Brinkmann and N. Van Cleemput, Classification and Generation of Nanocones, *Discrete Applied Mathematics* **159** (2011), 1528–1539.

[6] M. Deza, P. W. Fowler, and V. Grishukhin. Allowed boundary sequences for fused polycyclic patches and related algorithmic problems. *J. Chem. Inf. Comput. Sci.* **41** (2001), 300–308.

[7] J.E. Graver and C. Graves, Fullerene Patches I, *Ars Mathematica Contemporanea* **3** (2010), 104–120.

[8] C. Graves and J. McLoud-Mann, Side lengths of pseudoconvex fullerene patches, *Ars Mathematica Contemporanea* **5** (2012), 291–302.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 223–242

Fast recognition of partial star products and quasi cartesian products*

Marc Hellmuth †
Center for Bioinformatics, Saarland University,
D - 66041 Saarbrücken, Germany

Wilfried Imrich
Chair of Applied Mathematics, Montanuniversität,
A-8700 Leoben, Austria

Tomas Kupka
Department of Applied Mathematics, VSB-Technical University of Ostrava,
Ostrava, 70833, Czech Republic

Received 12 August 2013, accepted 16 December 2013, published online 8 December 2014

Abstract

This paper is concerned with the fast computation of a relation $\mathfrak{d}$ on the edge set of connected graphs that plays a decisive role in the recognition of approximate Cartesian products, the weak reconstruction of Cartesian products, and the recognition of Cartesian graph bundles with a triangle free basis.

A special case of $\mathfrak{d}$ is the relation δ^*, whose convex closure yields the product relation σ that induces the prime factor decomposition of connected graphs with respect to the Cartesian product. For the construction of $\mathfrak{d}$ so-called Partial Star Products are of particular interest. Several special data structures are used that allow to compute Partial Star Products in constant time. These computations are tuned to the recognition of approximate graph products, but also lead to a linear time algorithm for the computation of δ^* for graphs with maximum bounded degree.

Furthermore, we define *quasi Cartesian products* as graphs with non-trivial δ^*. We provide several examples, and show that quasi Cartesian products can be recognized in linear time for graphs with bounded maximum degree. Finally, we note that quasi products can be recognized in sublinear time with a parallelized algorithm.

*We thank Lydia Ostermeier for her insightful comments on graph bundles, as well as for the suggestion of the term "quasi product". This work was supported in part by ARRS Slovenia and the Deutsche Forschungsgemeinschaft (DFG) Project STA850/11-1 within the EUROCORES Program EuroGIGA (project GReGAS) of the European Science Foundation. This paper is based on part of the dissertation of the third author.

†Corresponding Author

This work is licensed under http://creativecommons.org/licenses/by/3.0/

Keywords: Cartesian product, quasi product, graph bundle, approximate product, partial star product, product relation.

Math. Subj. Class.: 05C15, 05C10

1 Introduction

Cartesian products of graphs derive their popularity from their simplicity, and their importance from the fact that many classes of graphs, such as hypercubes, Hamming graphs, median graphs, benzenoid graphs, or Cartesian graph bundles, are either Cartesian products or closely related to them [5]. As even slight disturbances of a product, such as the addition or deletion of an edge, can destroy the product structure completely [2], the question arises whether it is possible to restore the original product structure after such a disturbance. In other words, given a graph, the question is, how close it is to a Cartesian product, and whether one can find this product algorithmically. Unfortunately, in general this problem can only be solved by heuristic algorithms, as discussed in detail in [8]. That paper also presents several heuristic algorithms for the solution of this problem.

One of the main steps towards such algorithms is the computation of an equivalence relation $\mathfrak{d}_{|S_v}(W)^*$ on the edge-set of a graph. The complexity of the computation of $\mathfrak{d}_{|S_v}(W)^*$ in [8] is $O(n\Delta^4)$, where n is the number of vertices, and Δ the maximum degree of G. Here we improve the recognition complexity of $\mathfrak{d}_{|S_v}(W)^*$ to $O(m\Delta)$, where m is the number of edges of G, and thereby improve the complexity of the just mentioned heuristic algorithms.

A special case is the computation of the relation $\delta^* = \mathfrak{d}_{|S_v}(V(G))^*$. This relation defines the so-called quasi Cartesian product, see Section 3. Hence, quasi products can be recognized in $O(m\Delta)$ time. As the algorithm can easily be parallelized, it leads to sublinear recognition of quasi Cartesian products.

When the given graph G is a Cartesian product from which just one vertex was deleted, things are easier. In that case, the product is uniquely defined and can be reconstructed in polynomial time from G, see [1] and [3]. In other words, if G is given, and if one knows that there is a Cartesian product graph H such that $G = H \setminus x$, then H is uniquely defined. Hagauer and Žerovnik showed that the complexity of finding H is $O(mn(\Delta^2 + m))$. The methods of the present paper will lead to a new algorithm of complexity $O(m\Delta^2 + \Delta^4)$ for the solution of this problem. This is part of the dissertation [13] of the third author, and will be the topic of a subsequent publication.

Another class of graphs that is closely related to Cartesian products are Cartesian graph bundles, see Section 3. In [11] it was proved that Cartesian graph bundles over a triangle-free base can be effectively recognized, and in [14] it was shown that this can be done in $O(mn^2)$ time. With the methods of this paper, we suppose that one can improve it to $O(m\Delta)$ time. This too will be published separately.

2 Preliminaries

We consider finite, connected undirected graphs $G = (V, E)$ without loops and multiple edges. The *Cartesian product* $G_1 \Box G_2$ of graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$

E-mail addresses: marc.hellmuth@bioinf.uni-sb.de (Marc Hellmuth), imrich@unileoben.ac.at (Wilfried Imrich), tomas.kupka@teradata.com (Tomas Kupka)

is a graph with vertex set $V_1 \times V_2$, where the vertices (u_1, v_1) and (u_2, v_2) are adjacent if $u_1u_2 \in E_1$ and $v_1 = v_2$, or if $v_1v_2 \in E_2$ and $u_1 = u_2$. The Cartesian product is associative, commutative, and has the one vertex graph K_1 as a unit [5]. By associativity we can write $G_1 \Box G_2 \Box \cdots \Box G_k$ for a product G of graphs $G_1, G_2, \ldots, G_k$ and can label the vertices of G by the set of all k-tuples $(v_1, v_2, \ldots, v_k)$, where $v_i \in G_i$ for $1 \leq i \leq k$. If v is labeled $(v_1, v_2, \ldots, v_k)$, then we call v_i its ith coordinate. One says two edges have the same *Cartesian color* if their endpoints differ in the same coordinate.

A graph G is *prime* if it is non-trivial, and if the identity $G = G_1 \Box G_2$ implies that G_1 or G_2 is the one-vertex graph K_1. A representation of a graph G as a product $G_1 \Box G_2 \Box \cdots \Box G_k$ of prime graphs is called a *prime factorization* of G. It is well known that every connected graph G has a prime factor decomposition with respect to the Cartesian product, and that this factorization is unique up to isomorphisms and the order of the factors, see Sabidussi [15]. Furthermore, the prime factor decomposition can be computed in linear time, see [10].

Following the notation in [8], an induced cycle on four vertices is called *chordless square*. Let the edges $e = vu$ and $f = vw$ span a chordless square $vuxw$. Then f is the *opposite* edge of the edge xu. The vertex x is called *top vertex* (w.r.t. the square spanned by e and f). A top vertex x is *unique* if $|N(x) \cap N(v)| = 2$, where $N(u)$ denotes the (open) 1-neighborhood of vertex u. In other words, a top vertex x is not unique if there are further squares with top vertex x spanned by the edges e or f together with a third distinct edge g. Note that the existence of a unique top vertex x does not imply that e and f span a unique square, as there might be another square $vuyw$ with a possible unique top vertex y. Thus, e and f span a unique square $vuxw$ only if $|N(u) \cap N(w)| = 2$. The *degree* $\deg(u) := |N(u)|$ of a vertex u is the number of edges that contain u. The maximum degree of a graph is denoted by Δ and a path on n vertices by P_n.

We now recall the Breadth-First Search (BFS) ordering of the vertices $v_0, v_1, \ldots, v_{n-1}$ of a graph: select an arbitrary, but fixed vertex $v_0 \in V(G)$, called the *root*, and create a sorted list of vertices. Begin with v_0; append all neighbors $v_1, \ldots, v_{\deg(v_0)}$ of v_0 to the list; then append all neighbors of v_1 that are not already in the list; and continue recursively with $v_2, v_3, \ldots$ until all vertices of G are processed.

2.1 The Relations δ, σ and the Square Property.

There are two basic relations δ and σ, among other relations that are defined on the edge set of a given graph, that play an important role in the field of Cartesian product recognition. In the sequel we shall also use the notation R^* for the *transitive closure* of a relation R, that is, R^* is the smallest transitive relation containing R.

Definition 2.1. Two edges $e, f \in E(G)$ are in the *relation* δ_G, if one of the following conditions in G is satisfied:

(i) e and f are adjacent and it is not the case that there is a unique square spanned by e and f, and that this square is chordless.

(ii) e and f are opposite edges of a chordless square.

(iii) $e = f$.

Clearly, this relation is reflexive and symmetric but not necessarily transitive. The transitive closure δ_G^* is an equivalence relation.

If adjacent edges e and f are not in relation δ, that is, if Condition (i) of Definition 2.1 is not fulfilled, then they span a unique square, and this unique square spanned by e and f is chordless. We call such a square the *unique chordless square (spanned by e and f)*.

Two edges e and f are in the *product relation* σ_G if they have the same Cartesian colors with respect to the prime factorization of G. The product relation σ_G is a uniquely defined equivalence relation on $E(G)$ that contains all information about the prime factorization[1]. Furthermore, δ_G and δ_G^* are contained in σ_G. If there is no risk of confusion we write δ or σ for δ_G or σ_G, respectively.

We say an equivalence relation ρ defined on the edge set of a graph G has the *square property* if the following three conditions hold:

(a) For any two edges $e = uv$ and $f = uw$ that belong to different equivalence classes of ρ there exists a unique vertex $x \neq u$ of G that is adjacent to v and w.

(b) The square $uvxw$ is chordless.

(c) The opposite edges of any chordless square belong to the same equivalence class of ρ.

From the definition of δ it easily follows that δ is a refinement of any such ρ. It also implies that δ^*, and thus also σ, have the square property. This property is of fundamental importance, both for the Cartesian and the quasi Cartesian product. We note in passing that σ is the convex hull of δ^*, see [12].

2.2 The Partial Star Product

This section is concerned with the *partial star product*, which plays a decisive role in the local approach. As it was introduced in [8], we will only define it here, list some of its most basic properties, and refer to [8] for details.

Let $G = (V, E)$ be a given graph and E_v the set of all edges incident to some vertex $v \in V$. We define the local relation $\mathfrak{d}_v$ as follows:

$$\mathfrak{d}_v = ((E_v \times E) \cup (E \times E_v)) \cap \delta_G \subseteq \delta_{\langle N_2^G[v]\rangle},$$

where $\langle N_2^G[v]\rangle$ denotes the induced closed 2-neigborhood of v in G. In other words, $\mathfrak{d}_v$ is the subset of δ_G that contains all pairs $(e, f) \in \delta_G$, where at least one of the edges e and f is incident to v. Clearly $\mathfrak{d}_v^*$, which is not necessarily a subset of δ, is contained in δ^*, see [8].

Let S_v be a subgraph of G that contains all edges incident to v and all squares spanned by edges $e, e' \in E_v$ where e and e' are not in relation $\mathfrak{d}_v^*$. Then S_v is called *partial star product* (PSP for short). To be more precise:

Definition 2.2 (Partial Star Product (PSP))**.** Let $F_v \subseteq E \setminus E_v$ be the set of edges which are opposite edges of (chordless) squares spanned by $e, e' \in E_v$ that are in different $\mathfrak{d}_v^*$ classes, that is, $(e, e') \notin \mathfrak{d}_v^*$.

Then the *partial star product* is the subgraph $S_v \subseteq G$ with edge set $E' = E_v \cup F_v$ and vertex set $\cup_{e\in E'} e$, which consists of the end vertices of the edges in E'. We call v the *center* of S_v, E_v the set of *primal edges*, F_v the set of *non-primal edges*, and the vertices adjacent to v *primal vertices* of S_v.

[1]For the properties of σ that we will cite or use, we refer the reader to [5] or [9].

As shown in [8], a partial star product S_v is always an isometric subgraph or even isomorphic to a Cartesian product graph H, where the factors of H are so-called stars $K_{1,n}$. These stars can directly be determined by the respective $\mathfrak{d}_v^*$ classes, see [8].

Now we define a *local coloring* of S_v as the restriction of the relation $\mathfrak{d}_v^*$ to S_v:

$$\mathfrak{d}_{|S_v} := \mathfrak{d}^*_{v|S_v} = \{(e,f) \in \mathfrak{d}_v^* \mid e,f \in E(S_v)\}.$$

In other words, $\mathfrak{d}_{|S_v}$ is the subset of $\mathfrak{d}_v^*$ that contains all pairs of edges $(e,f) \in \mathfrak{d}_v^*$ where both e and f are in S_v and edges obtain the same local color whenever they are in the same equivalence class of $\mathfrak{d}_{|S_v}$. As an example consider the PSP S_v in Figure 1(d). The relation $\mathfrak{d}_{|S_v}$ has three equivalence classes (highlighted by thick, dashed and double-lined edges). Note, δ^* just contains one equivalence class. Hence, $\mathfrak{d}_{|S_v} \neq \delta^*_{S_v}$.

For a given subset $W \subseteq V$ we set

$$\mathfrak{d}_{|S_v}(W) = \cup_{v \in W} \mathfrak{d}_{|S_v} .$$

The transitive closure of $\mathfrak{d}_{|S_v}(W)$ is then called the *global coloring* with respect to W. As shown in [8], we have the following theorem.

Theorem 2.3. *Let $G = (V,E)$ be a given graph and $\mathfrak{d}_{|S_v}(V) = \cup_{v \in V} \mathfrak{d}_{|S_v}$. Then*

$$\mathfrak{d}_{|S_v}(V)^* = \delta^*_G.$$

For later reference and for the design of the recognition algorithm we list the following three lemmas about relevant properties of the PSP.

Lemma 2.4 ([8])**.** *Let G=(V,E) be a given graph and S_v be a PSP of an arbitrary vertex $v \in V$. If $e,f \in E_v$ are primal edges that are not in relation $\mathfrak{d}_v^*$, then e and f span a unique chordless square with a unique top vertex in G.*

Conversely, suppose that x is a non-primal vertex of S_v. Then there is a unique chordless square in S_v that contains x, and that is spanned by edges $e,f \in E_v$ with $(e,f) \notin \mathfrak{d}_v^*$.

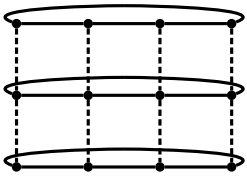
Lemma 2.5 ([8])**.** *Let G=(V,E) be a given graph and $f \in F_v$ be a non-primal edge of a PSP S_v of an arbitrary vertex $v \in V$. Then f is opposite to exactly one primal edge $e \in E_v$ in S_v, and $(e,f) \in \mathfrak{d}_{|S_v}$.*

Lemma 2.6 ([8])**.** *Let G=(V,E) be a given graph and $W \subseteq V$ such that $\langle W \rangle$ is connected. Then each vertex $x \in W$ meets every equivalence class of $\mathfrak{d}_{|S_v}(W)^*$ in $\cup_{v \in W} S_v$.*

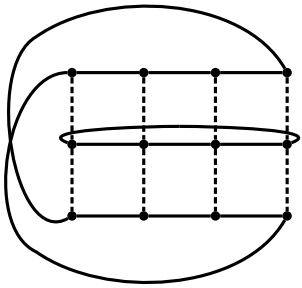
3 Quasi Cartesian Products

Given a Cartesian product $G = A \Box B$ of two connected, prime graphs A and B, one can recover the factors A and B as follows: the product relation σ has two equivalence classes, say E_1 and E_2, and the connected components of the graph $(V(G), E_1)$ are all isomorphic copies of the factor A, or of the factor B, see Figure 1(a). This property naturally extends to products of more than two prime factors.

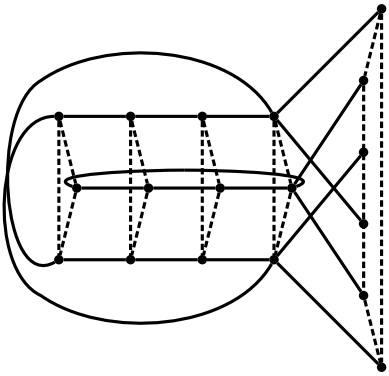
We already observed that δ is finer than any equivalence relation ρ that satisfies the square property. Hence the equivalence classes of ρ are unions of δ^*-classes. This also



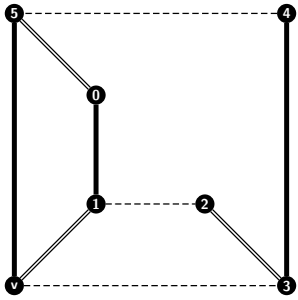
(a) The Cartesian product $G = P_3 \square C_4$.



(b) A quasi Cartesian product, which is also a graph bundle.



(c) A quasi Cartesian product, which is not a graph bundle.



(d) The approximate product and PSP S_v, which is neither a quasi product nor a graph bundle.

Figure 1: Shown are several quasi Cartesian products, graph bundles and approximate products.

holds for σ. It is important to keep in mind that σ can be trivial, that is, it consists of a single equivalence class even when δ^* has more than one equivalence class.

We call all graphs G with a non-trivial equivalence relation ρ that is defined on $E(G)$ and satisfies the square property *quasi (Cartesian) products*. Since $\delta^* \subseteq \rho$ for every such relation ρ, it follows that δ^* must have at least two equivalence classes for any quasi product. By Theorem 2.3 we have $\mathfrak{d}_{|S_v}(V(G))^* = \delta^*$. In other words, quasi products can be defined as graphs where the PSP's of all vertices are non-trivial, that is, none of the PSP's is a star $K_{1,n}$, and in addition, where the union over all $\mathfrak{d}_{|S_v}$ yields a non-trivial δ^*.

Consider the equivalence classes of the relation δ^* of the graph G of Figure 1(b). It has two equivalence classes, and locally looks like a Cartesian product, but is actually reminiscent of a Möbius band. Notice that the graph G in Figure 1(b) is prime with respect to Cartesian multiplication, although δ^* has two equivalence classes: all components of the first class are paths of length 2, and there are two components of the other δ^*-class, which do not have the same size. Locally this graph looks either like $P_3 \square P_3$ or $P_2 \square P_3$.

In fact, the graph in Figure 1(b) is a so-called Cartesian graph bundle [11], where *Cartesian graph bundles* are defined as follows: Let B and F be graphs. A graph G is a (Carte-

sian) graph bundle with fiber F over the base B if there exists a weak homomorphism[2] $p : G \to B$ such that

(i) for any $u \in V(B)$, the subgraph (induced by) $p^{-1}(u)$ is isomorphic to F, and

(ii) for any $e \in E(B)$, the subgraph $p^{-1}(e)$ is isomorphic to $K_2 \square F$.

The graph of Figure 1(c) shows that not all quasi Cartesian products are graph bundles. On the other hand, not every graph bundle has to be a quasi product. The standard example is the complete bipartite graph $K_{3,3}$. It is a graph bundle with base K_3 and fiber K_2, but has only one δ^*-class.

Note, in [8] we considered "approximate products" which were first introduced in [7, 6]. As approximate products are the graphs that have a (small) edit distance to a non-trivial product graph, it is clear that every bundle and quasi product can be considered as an approximate product, while the converse is not true. For example, consider the graph in Figure 1(d). Here, δ^* has only one equivalence class. However, the relation $\mathfrak{d}_{|S_v}$ has, in this case, three equivalence classes (highlighted by thick, dashed and double-lined edges).

Because of the local product-like structure of quasi Cartesian products we are led to the following conjecture:

Conjecture 3.1. *Quasi Cartesian products can be reconstructed in essentially the same time from vertex-deleted subgraphs as Cartesian products.*

4 Recognition Algorithms

4.1 Computing the Local and Global Coloring

For a given graph G, let $W \subseteq V(G)$ be an arbitrary subset of the vertex set of G such that the induced subgraph $\langle W \rangle$ is connected. Our approach for the computation is based on the recognition of all PSP's S_v with $v \in W$, and subsequent merging of their local colorings. The subroutine computing local colorings calls the vertices in BFS-order with respect to an arbitrarily chosen root $v_0 \in W$.

Let us now briefly introduce several additional notions used in the PSP recognition algorithm. At the start of every iteration we assign pairwise different *temporary local colors* to the primal edges of every PSP. These colors are then merged in subroutine processes to compute *local colors* associated with every PSP. Analogously, we use *temporary global colors* that are initially assigned to every edge incident with the root v_0.

For any vertex v of distance two from a PSP center c we store attributes called *first and second primal neighbor*, that is, references to adjacent primal vertices from which v was "visited" (in pseudo-code attributes are accessed by $v.FirstPrimalNeighbor$ and $v.SecondPrimalNeighbor$). When v is found to have at least two primal neighbors we add v to $\mathbb{T}_c$, which is a stack of candidates for non-primal vertices of S_c. Finally, we use *incidence* and *absence lists* to store recognized squares spanned by primal edges. Whenever we recognize that two primal edges span a square we put them into the incidence list. If we find out that a pair of primal edges cannot span a unique chordless square with unique top vertex, then we move it into the absence list. Note that the above structures are local and are always associated with a certain PSP recognition subroutine (Algorithm 4.1). Finally, we will "map" local colors to temporary global colors via temporary vectors which helps us to merge local with global colors.

[2] A weak homomorphism maps edges into edges or single vertices.

Algorithm 4.1 computes a local coloring for a given PSP and merges it with the global coloring $\mathfrak{d}_{|S_v}(W)^*$ where $W \subseteq V(G)$ is the set of treated centers. Algorithm 4.2 summarizes the main control structure of the local approach.

Algorithm 4.1 (*PSP recognition*)

Input: *Connected graph $G = (V, E)$, PSP center $c \in V$, global coloring $\mathfrak{d}_{|S}(W)^*$, where $W \subseteq V$ is the set of treated centers and where the subgraph induced by $W \cup c$ is connected.*

Output: *New temporary global coloring $\mathfrak{d}_{|S}(W \cup c)^*$.*

1. *Initialization.*
2. `FOR` *every neighbor u of c* `DO:`
   (a) `FOR` *every neighbor w of u (except c)* `DO:`
      i. `IF` *w is primal w.r.t. c* `THEN` *add pair of primal edges (cu, cw) to absence list.*
      ii. `ELSEIF` *w was not visited* `THEN` *set $w.FirstPrimalNeighbor = u$.*
      iii. `ELSE` *(w is not primal and was already visited)* `DO:`
         A. `IF` *only one primal neighbor v ($v \neq u$) of w was recognized so far, then* `DO:`
            - *Set $w.SecondPrimalNeighbor = u$.*
            - `IF` *(cu, cv) is not in incidence list, then add w to the stack $\mathbb{T}_c$ and add the pair (cu, cv) to incidence list.*
            - `ELSE` *(cu and cv span more squares) add pair (cu, cv) to absence list.*
         B. `ELSE:`
            - *Add all pairs formed by primal edges cv_1, cv_2, cu to absence list, where v_1, v_2 are first and second primal neighbors of w.*
3. *Assign pairwise different temporary local colors to primal edges.*
4. `FOR` *any pair (cu, cv) of primal edges cu and cv* `DO:`
   (a) `IF` *(cu, cv) is contained in absence list* `THEN` *merge temporary local colors of cu and cv.*
   (b) `IF` *(cu, cv) is not contained in incidence list* `THEN` *merge temporary local colors of cu and cv.*

   (Resulting merged temporary local colors determine local colors of primal edges in S_c. We will reference them in the following steps.).
5. `FOR` *any primal edge cu* `DO:`
   (a) `IF` *cu was already assigned some temporary global color d_1* `THEN`
      i. `IF` *local color b of cu was already mapped to some temporary global color d_2, where $d_2 \neq d_1$,* `THEN` *merge d_1 and d_2.*
      ii. `ELSE` *map local color b to d_1.*
6. `FOR` *any vertex v from the stack $\mathbb{T}_c$* `DO:`
   (a) *Check local colors of primal edges cw_1 and cw_2 (where w_1, w_2 are first and second primal neighbor of v, respectively).*
   (b) `IF` *they differ in local colors* `THEN`
      i. `IF` *there was defined temporary global color d_1 for vw_1* `THEN`

A. `IF` local color b of cw_2 was already mapped to some temporary global color d_2, where $d_2 \neq d_1$ `THEN` merge d_1 and d_2.
B. `ELSE` map local color b to d_1.

ii. `IF` there was already defined temporary global color d_1 for vw_2 `THEN`:

A. `IF` local color b of cw_1 was already mapped to some temporary global color d_2, where $d_2 \neq d_1$ `THEN` merge d_1 and d_2.
B. `ELSE` map local color b to d_1.

7. Take every edge e of the PSP S_c that was not colored by any temporary global color up to now and assign it d, where d is the temporary global color to which the local color of e or the local color of its opposite primal edge e' was mapped.
(If there is a local color b that was not mapped to any temporary global color, then we create a new temporary global color and assign it to all edges of color b.)

Algorithm 4.2 (*Computation of* $\mathfrak{d}_{|S_v}(W)^*$)

Input: A connected graph G, $W \subseteq V(G)$ s.t. the induced subgraph $\langle W \rangle$ is connected, and an arbitrary vertex $v_0 \in W$.
Output: Relation $\mathfrak{d}_{|S_v}(W)^*$.

1. Initialization.
2. Set sequence Q of vertices $v_0, v_1, \ldots, v_n$ that form W in BFS-order with respect to v_0.
3. Set $W' := \emptyset$.
4. Assign pairwise different temporary global colors to edges incident to v_0.
5. `FOR` any vertex v_i from sequence Q `DO`:

(a) Use Algorithm 4.1 to compute $\mathfrak{d}_{|S_v}(W' \cup v_i)^*$.
(b) Add v_i to W'.

In order to show that Algorithm 4.1 correctly recognizes the local coloring, we define the (temporary) relations α_c and β_c for a chosen vertex c: Two *primal* edges of S_c are

- in *relation* α_c if they are contained in the incidence list and
- in *relation* β_c if they are contained in the absence list

after Algorithm 4.1 is executed for c. Note, we denote by $\overline{\alpha}_c$ the complement of α_c, which contains all pairs of primal edges of PSP S_c that are not listed in the incidence list.

Lemma 4.1. *Let e and f be two primal edges of the PSP S_c. If e and f span a square with some non-primal vertex w as unique top-vertex, then $(e, f) \in \alpha_c$.*

Proof. Let $e = cu_1$ and $f = cu_2$ be primal edges in S_c that span a square cu_1wu_2 with unique top-vertex w, where w is non-primal. Note, since w is the unique top vertex, the vertices u_1 and u_2 are its only primal neighbors. W.l.o.g. assume that for vertex w no first primal neighbor was assigned and let first u_1 and then u_2 be visited. In Step 2a vertex w is recognized and the first primal neighbor u_1 is determined in Step 2(a)ii. Take the next vertex u_2. Since w is not primal and was already visited, we are in Step 2(a)iii. Since only one primal neighbor of w was recognized so far, we go to Step 2(a)iiiA. If (cu_1, cu_2) is not already contained in the incidence list, it will be added now and thus, $(cu_1, cu_2) \in \alpha_c$. □

Corollary 4.2. *Let e and f be two adjacent distinct primal edges of the PSP S_c. If $(e, f) \in \overline{\alpha}_c$, then e and f do not span a square or span a square with non-unique or primal top vertex. In particular, $\overline{\alpha}_c$ contains all pairs (e, f) that do not span any square.*

Proof. The first statement is just the contrapositive of the statement in Lemma 4.1. For the second statement observe that if $e = cx$ and $f = cy$ are two distinct primal edges of S_c that do not span a square, then the vertices x and y do not have a common non-primal neighbor w. It is now easy to verify that in none of the substeps of Step 2 the pair (e, f) is added to the incidence list, and thus, $(e, f) \in \overline{\alpha}_c$. □

Lemma 4.3. *Let e and f be two primal edges of the PSP S_c that are in relation β_c. Then e and f do not span a unique chordless square with unique top vertex.*

Proof. Let $e = cu_1$ and $f = cu_2$ be primal edges of S_c. Then pair (e, f) is added to the absence list in:

a) Step 2(a)i, when u_1 and u_2 are adjacent. Then no square spanned by e and f can be chordless.

b) Step 2(a)iiiA (`ELSE`-condition), when (e, f) is already listed in the incidence list and another square spanned by e and f is recognized. Thus, e and f do not span a unique square.

c) Step 2(a)iiiB, when e and f span a square with top vertex w that has more than two primal neighbors and at least one of the primal vertices u_1 and u_2 are recognized as first or second primal neighbor of w. Thus e and f span a square with non-unique top vertex.

□

Lemma 4.4. *Relation β_c^* contains all pairs of primal edges (e, f) of S_c that satisfy at least one of the following conditions:*

a) e and f span a square with a chord.

b) e and f span a square with non-unique top vertex.

c) e and f span more than one square.

Proof. Let $e = cu_1$ and $f = cu_2$ be primal edges of the PSP S_c.

a) If e and f span a square with a chord, then u_1 and u_2 are adjacent or the top vertex w of the spanned square is primal and thus, there is a primal edge $g = cw$. In the first case, we can conclude analogously as in the proof of Lemma 4.3 that $(e, f) \in \beta_c$. In the second case, we analogously obtain $(e, g), (f, g) \in \beta_c$ and therefore, $(e, f) \in \beta_c^*$.

b) Let e and f span a square with non-unique top vertex w. If at least one of the primal vertices u_1, u_2 is a first or second neighbor of w then e and f are listed in the absence list, as shown in the proof of Lemma 4.3. If u_1 and u_2 are neither first nor second primal neighbors of w, then both edges e and f will be added to the absence list in Step 2(a)iiiB, together with the primal edge $g = cu_3$, where u_3 is the first recognized primal neighbor of w. In other words, $(e, g), (f, g) \in \beta_c$ and hence, $(e, f) \in \beta_c^*$.

c) Let e and f span two squares with top vertices w and w', respectively and assume w.l.o.g. that first vertex w is visited and then w'. If both vertices u_1 and u_2 are recognized as first and second primal neighbors of w and w', then (cu_1, cu_2) is added to the incidence list when visiting w in Step 2(a)iiiA. However, when we visit w', then we insert (cu_1, cu_2) to the absence list in Step 2(a)iiiA, because this pair is already included in the incidence list. Thus, $(e, f) \in \beta_c$. If at least one of the vertices w, w' does not have u_1 and u_2 as first or second primal neighbor, then e and f must span a square with non-unique top vertex. Item b) implies that $(e, f) \in \beta_c^*$.

□

Lemma 4.5. *Let f be a non-primal edge and e_1, e_2 be two distinct primal edges of S_c. Let $(e_1, f), (e_2, f) \in \mathfrak{d}_c$. Then $(e_1, e_2) \in \beta_c^*$.*

Proof. Since the edge f is non-primal, f is not incident with the center c. Recall, by the definition of $\mathfrak{d}_c$, two distinct edges can be in relation $\mathfrak{d}_c$ only if they have a common vertex or are opposite edges in a square. To prove our lemma we need to investigate the three following cases, which are also illustrated in Figure 2:

a) Suppose both edges e_1 and e_2 are incident with f. Then e_1 and e_2 span a triangle and consequently (e_1, e_2) will be added to the absence list in Step 2(a)i.

b) Let e_1 and e_2 be opposite to f in some squares. There are two possible cases (see Figure 2 b)). In the first case e_1 and e_2 span a square with non-unique top vertex. By Lemma 4.4, $(e_1, e_2) \in \beta_c^*$. In the second case e_1 and e_2 span triangles with other primal edges e_3 and e_4. As in Case a) of this proof, we have $(e_1, e_3) \in \beta_c$, $(e_3, e_4) \in \beta_c$, $(e_4, e_2) \in \beta_c$ and consequently, $(e_1, e_2) \in \beta_c^*$

c) Suppose only e_1 has a common vertex with f and e_2 is opposite to f in a square. Again we need to consider two cases (see Figure 2 c)). Since e_1 and f are adjacent and $(e_1, f) \in \mathfrak{d}_c$, we can conclude that either no square is spanned by e_1 and f, or that the square spanned by e_1 and f is not chordless or not unique. It is easy to see that in the first case the edges e_1 and e_2 are contained in a common triangle and thus will be added to the absence list in Step 2(a)i. In the second case e_1, e_2 span a square which has a chord or has a non-unique top vertex. In both cases Lemma 4.4 implies that e_1 and e_2 are in relation β_c^*.

□

Lemma 4.6. *Let e and f be distinct primal edges of the PSP S_c. Then $(e, f) \in (\overline{\alpha}_c \cup \beta_c)^*$ if and only if $(e, f) \in \mathfrak{d}_c^*$.*

Proof. Assume first that $(e, f) \in \overline{\alpha}_c \cup \beta_c$. By Corollary 4.2, if $(e, f) \in \overline{\alpha}_c$, then e and f do not span a common square, or span a square with non-unique or primal top vertex. In the first case, e and f are in relation δ_G and consequently also in relation $\mathfrak{d}_c$. On the other hand, if e and f span square with non-unique top vertex then, by Lemma 2.4, e and f are in relation $\mathfrak{d}_c^*$ as well. Finally, if e and f span a square with primal top vertex w, then this square has a chord cw and $(e, f) \in \mathfrak{d}_c^*$. If $(e, f) \in \beta_c$, then Lemma 4.3 implies that e and f do not span a unique chordless square with unique top vertex. Again, by Lemma 2.4, we infer that $(e, f) \in \mathfrak{d}_c^*$. Hence, $\overline{\alpha}_c \cup \beta_c \subseteq \mathfrak{d}_c^*$, and consequently, $(\overline{\alpha}_c \cup \beta_c)^* \subseteq \mathfrak{d}_c^*$.

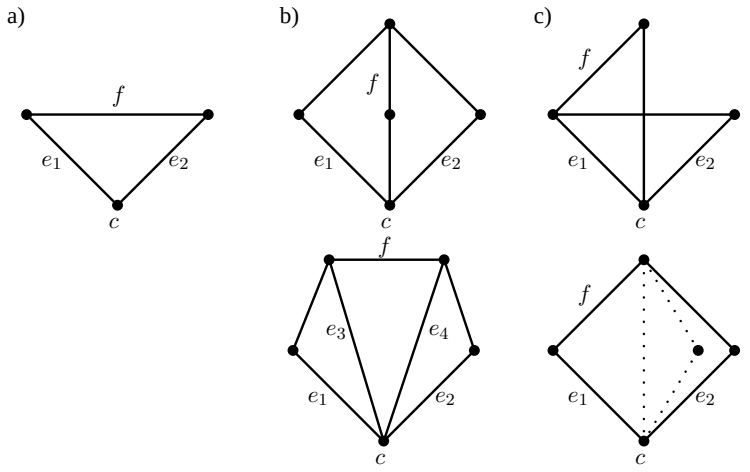


Figure 2: The three possible cases a), b), and c) that are investigated in the proof of Lemma 4.5.

Now, let $(e, f) \in \mathfrak{d}_c^*$. Then there is a sequence $U = (e = e_1, e_2, \ldots, e_k = f)$, $k \geq 2$, with $(e_i, e_{i+1}) \in \mathfrak{d}_c$ for $i \in \{1, 2, \ldots, k-1\}$. By definition of $\mathfrak{d}_c$, two primal edges are in relation $\mathfrak{d}_c$ if and only if they do not span a unique and chordless square. Corollary 4.2 and Lemma 4.4 imply that all these pairs are contained in $(\overline{\alpha}_c \cup \beta_c)^*$. Hence, any two consecutive primal edges e_i and e_{i+1} contained in the sequence U are in relation $(\overline{\alpha}_c \cup \beta_c)^*$. Assume that there is an edge $e_i \in U$ that is not incident to the center c and thus, non-primal. By the definition of $\mathfrak{d}_c$, and since $(e_{i-1}, e_i), (e_i, e_{i+1}) \in \mathfrak{d}_c$, we can conclude that the edges e_{i-1} and e_{i+1} must be primal in S_c. Lemma 4.5 implies that e_{i-1} and e_{i+1} must be in relation β_c^*. Thus, if we remove the edge e_i from U, we still can claim that all consecutive primal edges in $U \setminus \{e_i\}$ are in relation $(\overline{\alpha}_c \cup \beta_c)^*$. By removing all non-primal edges from U we therefore obtain a sequence $U' = e = e_1, e_2', \ldots, e_j' = f$ of primal edges. By analogous arguments as before, all pairs (e_i', e_{i+1}') of U' must be contained in $(\overline{\alpha}_c \cup \beta_c)^*$. By transitivity, e and f are also in $(\overline{\alpha}_c \cup \beta_c)^*$. □

Corollary 4.7. *Let e and f be primal edges of the PSP S_c. Then $(e, f) \in (\overline{\alpha}_c \cup \beta_c)^*$ if and only if e and f have the same local color in S_c.*

Proof. This is an immediate consequence of Lemma 4.6, the local color assignment, and the merging procedure (Step 3 and 4) in Algorithm 4.1. □

Lemma 4.8. *Let $\mathfrak{d}_{|S_v}(W)^*$ be a global coloring associated with a set of treated centers W and assume that the induced subgraph $\langle W \rangle$ is connected. Let c be a vertex that is not contained in W but adjacent to a vertex in W. Then Algorithm 4.1 computes the global coloring $\mathfrak{d}_{|S_v}(W \cup c)^*$ by taking W and c as input.*

Proof. Let $W \subseteq V(G)$ be a set of PSP centers and let $c \in V(G)$ be a given center of PSP S_c where $c \notin W$ and $\langle W \cup c \rangle$ is connected. In Step 2 of Algorithm 4.1 we compute the absence and incidence lists. In Step 3, we assign pairwise different temporary local colors to any primal edge adjacent to c. Two temporary local colors b_1 and b_2 are then merged in Step 4 if and only if there exists some pair of primal edges $(e_1, e_2) \in (\overline{\alpha}_c \cup \beta_c)$ where

e_1 is colored with b_1 and e_2 with b_2. Therefore, merged temporary local colors reflect equivalence classes of $(\overline{\alpha}_c \cup \beta_c)^*$ containing the primal edges incident to c. By Corollary 4.7, $(\overline{\alpha}_c \cup \beta_c)^*$ classes indeed determine the local colors of primal edges in S_c.

Note, if one knows the colors of primal edges incident to c, then it is very easy to determine the set of non-primal edges of S_c, as any two primal edges of different equivalence classes span a unique and chordless square. In Step 6, we investigate each vertex v from stack $\mathbb{T}_c$ and check the local colors of primal edges cw_1 and cw_2, where w_1 and w_2 are the first and second recognized primal neighbors of v, respectively. If cw_1 and cw_2 differ in their local colors, then vw_1 and vw_2 are non-primal edges of S_c, as follows from the PSP construction. Recall that the stack contains all vertices that are at distance two from center c and which are adjacent to at least two primal vertices. In other words, the stack contains all non-primal top vertices of all squares spanned by primal edges. Consequently, we claim that all non-primal edges of the PSP S_c are treated in Step 6. Note that non-primal edges have the same local color as their opposite primal edge, which is unique by Lemma 2.5.

As we already argued, after Step 4 is performed we know, or can at least easily determine all edges of S_c and their local colors. Recall that local colors define the local coloring $\mathfrak{d}_{|S_c}$. Suppose, temporary global colors that correspond to the global coloring $\mathfrak{d}_{|S_v}(W)^*$ are assigned. Our goal is to modify and identify temporary global colors such that they will correspond to the global coloring $\mathfrak{d}_{|S_v}(W \cup c)^*$. Let $B_1, B_2, \ldots, B_k$ be the classes of $\mathfrak{d}_{|S_c}$ (*local classes*) and $D_1, D_2, \ldots, D_l$ be the classes of $\mathfrak{d}_{|S_v}(W)^*$ (*global classes*). When a local class B_i and a global class D_j have a nonempty intersection, then we can infer that all their edges must be contained in a common class of $\mathfrak{d}_{|S_v}(W \cup c)^*$. Note, by means of Lemma 2.6, we can conclude that for each local class B_i there is a global class D_j such that $B_i \cap D_j \neq \emptyset$, see also [8]. In that case we need to guarantee that edges of B_i and D_j will be colored by the same temporary global color. Note, in the beginning of the iteration two edges have the same temporary global color if and only if they lie in a common global class.

In Step 5 and Step 6, we investigate all primal and non-primal edges of S_c. When we treat first edge e that is colored by some local color b_i, that is $e \in B_i$, and has already been assigned some temporary global color d_j, and therefore $e \in D_j$, then we map b_i to d_j. Thus, we keep the information that $e \in B_i \cap D_j$. In Step 7, we then assign temporary global color d_j to any edge of S_c that is colored by the local color b_i. If the local color b_i is already mapped to some temporary global color d_j, and if we find another edge of S_c that is colored by b_i and simultaneously has been assigned some different temporary global color $d_{j\prime}$, then we merge d_j and $d_{j\prime}$ in Step 5(a)i. Obviously this is correct, since $B_i \cap D_j \neq \emptyset$ and $B_i \cap D_{j\prime} \neq \emptyset$, and hence $D_j, D_{j\prime}$ and B_i must be contained in a common equivalence class of $\mathfrak{d}_{|S_v}(W \cup c)^*$. Recall, for each local class B_i there is a global class D_j such that $B_i \cap D_j \neq \emptyset$. This means that every local color is mapped to some global color, and consequently there is no need to create a new temporary global color in Step 7.

Therefore, whenever local and global classes share an edge, then all their edges will have the same temporary global color at the end of Step 7. On the other hand, when edges of two different global classes are colored by the same temporary global color, then both global classes must be contained in a common class of $\mathfrak{d}_{|S_v}(W \cup c)^*$.

Hence, after the performance of Step 7, the merged temporary global colors determine the equivalence classes of $\mathfrak{d}_{|S_v}(W \cup c)^*$. □

Lemma 4.9. *Let G be a connected graph, $W \subseteq V(G)$ s.t. $\langle W \rangle$ is connected, and v_0*

an arbitrary vertex of G. Then Algorithm 4.2 computes the global coloring $\mathfrak{d}_{|S_v}(W)^*$ by taking G, W, and v_0 as input.

Proof. In Step 2 we define the BFS-order in which the vertices will be processed and store this sequence in Q. In Step 4 we assign pairwise different temporary global colors to all edges that are incident with v_0. In Step 5 we iterate over all vertices of the given induced connected subgraph $\langle W\rangle$ of G. For every vertex we execute Algorithm 4.1. Lemma 4.8 implies that in the first iteration we correctly compute the local colors for S_{v_0}, and consequently also $\mathfrak{d}_{|S_v}(\{v_0\})^*$. Obviously, whenever we merge two temporary local colors of two primal edges in the first iteration, then we also merge their temporary global colors. Consequently, the resulting temporary global colors correspond to the global coloring $\mathfrak{d}_{|S_v}(\{v_0\})^*$ after the first iteration. Lemma 4.8 implies that after all iterations are performed, that is, all vertices in Q are processed, the resulting temporary global colors correspond to $\mathfrak{d}_{|S_v}(W)^*$ for the given input set $W \subseteq V(G)$. □

For the global coloring, Theorem 2.3 implies that $\mathfrak{d}_{|S_v}(V(G))^* = \delta^*_G$. This leads immediately to the following theorem.

Theorem 4.10. *Let G be a connected graph and v_0 an arbitrary vertex of G. Then Algorithm 4.2 computes the global coloring δ^*_G by taking G, $V(G)$, and v_0 as input.*

4.2 Time Complexity

We begin with the complexity of merging colors. We have global and local colors, and will define *local* and *global color graphs*. Both graphs are acyclic temporary structures. Their vertex sets are the sets of temporary colors in the initial state. In this state the color graphs have no edges. Every component is a single vertex and corresponds to an initial temporary color. Recall that we color edges of graphs, for example the edges of G or S_v. The color of an edge is indicated by a pointer to a vertex of the color graph. These pointers are not changed, but the colors will correspond to the components of the color graph. When two colors are merged, then this will be reflected by adding an edge between their respective components.

The color graph is represented by an adjacency list as described in [5, Chapter 17.2] or [9, pp. 34 -37]. Thus, working with the color graph needs $O(k)$ space when k colors are used. Furthermore, for every vertex of the color graph we keep an index of the connected component in which the vertex is contained. We also store the actual size of every component, that is, the number of vertices of this component.

Suppose we wish to merge temporary colors of edges e and f that are identified with vertices a, respectively b, in the color graph. We first check whether a and b are contained in the same connected component by comparing component indices. If the component indices are the same, then e and f already have the same color, and no action is necessary. Otherwise we insert an edge between a and b in the color graph. As this merges the components of a and b we have to update component indices and the size. The size is updated in constant time. For the component index we use the index of the larger component. Thus, no index change is necessary for the larger component, but we have to assign the new index to all vertices of the smaller component.

Notice that the color graph remains acyclic, as we only add edges between different components.

Lemma 4.11. *Let $G_0 = (V, E)$ be a graph with $V = \{v_1, \ldots, v_k\}$ and $E = \emptyset$. The components of G_0 consist of single vertices. We assign component index j to every component $\{v_j\}$. For $i \in \{1, \ldots, k-1\}$ let G_{i+1} denote the graph that results from G_i by adding an edge between two distinct connected components, say C and C'. If $|C| \leq |C'|$, we use the component index of C' for the new component and assign it to every vertex of C.*

Then every G_i is acyclic, and the total cost of merging colors is $O(k \log_2 k)$.

Proof. Acyclicity is true by construction.

A vertex is assigned a new component index when its component is merged with a larger one. Thus, the size of the component at least doubles at every such step. Because the maximum size of a component is bounded by k, there can be at most $\log_2 k$ reassignments of the component index for every vertex. As there are k vertices, this means that the total cost of merging colors is $O(k \log_2 k)$. □

The color graph is used to identify temporary local, resp., global colors. Based on this, we now define the *local* and *global color graph*.

Assigned labels of the vertices of the global color graph are stored in the edge list, where any edge is identified with at most one such label. Notice that the original graph is represented by an extended adjacency list, where for any vertex and its neighbor a reference to the edge (in the edge list) that connects them is stored. This reference allows to access a global temporary color from adjacency list in constant time.

In every iteration of Algorithm 4.2, we recognize the PSP for one vertex by calling Algorithm 4.1. In the following paragraph we introduce several temporary attributes and matrices that are used in the algorithm.

Suppose we execute an iteration that recognizes some PSP S_c. To indicate whether a vertex was treated in this iteration we introduce the attribute $visited$, that is, when vertex v is visited in this iteration we set $v.visited = c$. Any value different from c means that vertex v was not yet treated in this iteration. Analogously, we introduce the attribute $primal$ to indicate that a vertex is adjacent to the current center c. The attribute $tempLabel$ maps primal vertices to the indices of rows and columns of the matrices $incidenceList$ and $absenceList$. For any vertex v that is at distance two from the center c we store its first and second primal neighbor w_1 and w_2 in the attributes $FirstPrimalNeighbor$ and $SecondPrimalNeighbor$. Furthermore, we need to keep the position of vw_1 and vw_2 in the edge list to get their temporary global colors. For this purpose, we use attributes $firstEdge$ and $secondEdge$. Attribute $mapLocalColor$ helps us to map temporary local colors to the vertices of the global color graph. Any vertex that is at distance two from the center and has a least two primal neighbors is a candidate for a non-primal vertex. We insert them to the $stack$. The temporary structures help to access the required information in constant time:

- $v.visited = c$
  vertex v has been already visited in the current iteration.
- $v.primal = c$
  vertex v is adjacent to center c.
- $incidenceList[v.tempLabel, u.tempLabel] = 0$
  pair of primal edges (cv, cu) is missing in the incidence list.
- $absenceList[v.tempLabel, u.tempLabel] = 1$
  pair of primal edges (cv, cu) was inserted to the absence list.

- $v.firstPrimalNeighbor = u$
  u is the first recognized primal neighbor of the non-primal vertex v.
- $v.firstEdge = e$
  edge e joins the non-primal vertex v with its first recognized primal neighbor (it is used to get the temporary global color from the edge list).
- $b.mapLocalColor = d$
  local color b is mapped to temporary global color d (i.e. there exists an edge that is colored by both colors).

Note that the temporary matrices $incidenceList$ and $absenceList$ have dimension $\deg(c) \times \deg(c)$ and that all their entries are set to zero in the beginning of every iteration.

Theorem 4.12. *For a given connected graph $G = (V, E)$ with maximum degree Δ and $W \subseteq V$, Algorithm 4.2 runs in $O(|E|\Delta)$ time and $O(|E| + \Delta^2)$ space.*

Proof. Let G be a given graph with m edges and n vertices. In Step 1 of Algorithm 4.2 we initialize all temporary attributes and matrices. This consumes $O(m+n) = O(m)$ time and space, since G is connected, and hence, $m \geq n - 1$. Moreover, we set all temporary colors of edges in the edge list to zero, which does not increase the time and space complexity of the initial step. Recall that we use an extended adjacency list, where every vertex and its neighbors keep the reference to the edge in the edge list that connects them. To create an extended adjacency list we iterate over all edges in the edge list, and for every edge $uv = e \in E(G)$ we set a new entry for the neighbor v for u and, simultaneously, we add a reference $v.edge = e$. The same is done for vertex v. It can be done in $O(m)$ time and space.

In Step 2 of Algorithm 4.2, we build a sequence of vertices in BFS-order starting with v_0, which is done in $O(m + n)$ time in general. Since G is connected, the BFS-ordering can be computed in $O(m)$ time. Step 3 takes constant time. In Step 4 we initialize the global color graph that has $\deg(v_0)$ vertices (bounded by Δ in general). As we already showed, all operations on the global color graph take $O(\Delta \log_2 \Delta)$ time and $O(\Delta)$ space. We proceed to traverse all neighbors $u_1, u_2, \ldots, u_{\deg(v_0)}$ of the root $v_0 \in V(G)$ (via the adjacency list) and assign them unique labels $1, 2, \ldots, \deg(v_0)$ in edge list, that is, every edge v_0u_i gets the label i. In this way, we initialize pairwise different temporary global colors of edges incident with v_0 , that is, to vertices of the global color graph. Using the extended adjacency list, we set the label to an edge in the edge list in constant time. In Step 5 we run Algorithm 4.1 for any vertex from the defined BFS-sequence.

In the remainder of this proof, we will focus on the complexity of Algorithm 4.1. Suppose we perform Algorithm 4.1 for vertex c to recognize the PSP S_c. The recognition process is based on temporary structures. We do not need to reset any of these structures, for any execution of Algorithm 4.1 for a new center c, except $absenceList$ and $incidenceList$. This is done in Step 1. Further, we set here the attribute $tempLabel$ for every primal vertex v, such that every vertex has assigned a unique number from $\{1, 2, \ldots, \deg(c)\}$. Finally, we traverse all neighbors of the center c and for each of them we set $primal$ to c. Hence, the initial step of Algorithm 4.1 is done in $O(\deg(c)^2)$ time.

Step 2a is performed for every neighbor of every primal vertex. The number of all such neighbors is at most $\deg(c)\Delta$. For every treated vertex, we set attribute $visited$ to c.

This allows us to verify in constant time that a vertex was already visited in the recognition subroutine Algorithm 4.1.

If the condition in Step 2(a)i is satisfied, then we add primal edges cu and cw to the absence list. By the previous arguments, this can be done in constant time by usage of $tempLabel$ and $absenceList$.

If the condition in Step 2(a)ii is satisfied, we set vertex u as first primal neighbor of vertex w. For this purpose, we use the attribute $firstPrimalNeighbor$. We also set $w.firstEdge = e$, where e is a reference to the edge in the edge list that connects u and w. This reference is obtained from the extended adjacency list in constant time. Recall, the edge list is used to store the labels of vertices of the global color graph for the edges of a given graph, that is, the assignment of temporary global colors to the edges. Using $w.firstEdge$, we are able to directly access the temporary global color of edge uw in constant time.

Step 2(a)iii is performed when we try to visit a vertex w from some vertex u where w has been already visited before from some vertex v. If v is the only recognized primal neighbor of w, then we perform analogous operations as in the previous step. Moreover. if (cu, cv) is not contained in the incidence list, then we set u as second primal neighbor of w, add (cu, cv) to the incidence list and add w to the stack. Otherwise we add (cu, cv) to the absence list. The number of operations in this step is constant.

If w has more recognized primal neighbors we process case B. Here we just add all pairs formed by cv_1, cv_2, cu to absence list. Again, the number of operations is constant by usage of $tempLabel$ and matrices $incidenceList$ and $absenceList$.

In Step 3 we assign pairwise different temporary local colors to the primal edges. Assume the neighbors of the center c are labeled by $1, 2, \ldots, \deg(c)$, then we set value $u.tempLabel$ to cu. In Step 4a we iterate over all entries of the $absenceList$. For all pairs of edges that are in the absence list we check whether they still have different temporary local colors and if so, we merge their temporary local colors by adding a respective edge in the local color graph. Analogously we treat all pairs of edges contained in the $incidenceList$ in Step 4b. Here we merge temporary local colors of primal edges cu and cv when the pair (cu, cv) is missing. To treat all entries of the $absenceList$ and $incidenceList$ we need to perform $\deg(c)^2$ iterations. Recall, the temporary local color of the primal edge cu is equal to the index of the connected component in the local color graph, in which vertex $u.tempLabel$ is contained. Thus, the temporary local color of this primal edge can be accessed in constant time. As we already showed, the number of all operations on the local color graph is bounded by $O(\deg(c) \log_2 \deg(c))$. Hence, the overall time complexity of both Steps 3 and 4 is $O(\deg(c)^2)$.

In Step 5 we map temporary local colors of primal edges to temporary global colors. For this purpose, we use the attribute $mapLocalColor$. The temporary global color of every edge can be accessed by the extended adjacency list, the edge list and the global color graph in constant time. Since we need to iterate over all primal vertices, we can conclude that Step 5 takes $O(\deg(c))$ time.

In Step 6 we perform analogous operations for any vertex from Stack $\mathbb{T}_c$ as in Step 5. In the worst case, we add all vertices that are at distance two from the center to the stack. Hence, the size of the stack is bounded by $O(\deg(c)\Delta)$. Recall that the first and second primal neighbor w_1 and w_2 of every vertex v from the stack can be directly accessed by the attributes $firstPrimalNeighbor$ and $secondPrimalNeighbor$. On the other hand, the temporary global colors of non-primal edges vw_1 and vw_2 can be accessed directly by the

attributes $firstEdge$ and $secondEdge$. Thus, all needful information can be accessed in constant time. Consequently, the time complexity of this step is bounded by $O(\deg(c)\Delta)$.

In the last step, Step 7, we iterate over all edges of the recognized PSP. Note, the list of all primal edges can be obtained from the extended adjacency list. To get all non-primal edges we iterate over all vertices from the stack and use the attributes $firstEdge$ and $secondEdge$, which takes $O(\deg(c)\Delta)$ time. The remaining operations can be done in constant time.

To summarize, Algorithm 4.1 runs in $O(\deg(c)\Delta)$ time. Consequently, Step 5 of Algorithm 4.2 runs in $O(\sum_{c\in W}\deg(c)\Delta) = O(m\Delta)$ time, which defines also the total time complexity of Algorithm 4.2. The most space consuming structures are the edge list and the extended adjacency list ($O(m)$ space) and the temporary matrices $absenceList$ and $incidenceList$ ($O(\Delta^2)$ space). Hence, the overall space complexity is $O(m + \Delta^2)$. □

Since quasi Cartesian products are defined as graphs with non-trivial δ^*, Theorem 4.10 and 4.12 imply the following corollary.

Corollary 4.13. *For a given connected graph $G = (V, E)$ with* bounded *maximum degree Algorithm 4.2 (with slight modifications) determines whether G is a quasi Cartesian product in $O(|E|)$ time and $O(|E|)$ space.*

4.3 Parallel Processing

The local approach allows the parallel computation of $\delta^*(G)$ on multiple processors. Consider a graph G with vertex set $V(G)$. Suppose we are given a decomposition of $V(G) = W_1 \cup W_2 \cup \cdots \cup W_k$ into k parts such, that $|W_1| \approx |W_2| \approx \cdots \approx |W_k|$, where the subgraphs induced by $W_1, W_2, \ldots, W_k$ are connected, and the number of edges whose endpoints lie in different partitions is small (we call such a decomposition *good*).

Algorithm 4.3 (*Parallel recognition of δ^**)

Input: *A graph G, and a good decomposition $V(G) = W_1 \cup W_2 \cup \cdots \cup W_k$.*
Output: *Relation δ^*_G.*

1. *For every partition W_i concurrently compute global coloring $\mathfrak{d}_{|S_v}(W_i)$ ($i \in \{1, 2, \ldots, k\}$):*
   (a) *Take all vertices of W_i and order them in BFS to get sequence Q_i.*
   (b) *Set $W' := \emptyset$.*
   (c) *Assign pairwise different temporary global colors to edges incident to first vertex in Q_i.*
   (d) *For any vertex v from sequence Q_i do:*
      i. *Use Algorithm 4.1 to compute $\mathfrak{d}_{|S_v}(W' \cup v)^*$.*
      ii. *Move all edges that were treated in previous step and have at least one endpoint not in partition W_i to stack $\mathbb{T}_i$.*
      iii. *Add v to W'.*
2. *Run concurrently for every partition W_i to merge all global colorings ($i \in \{1, 2, \ldots, k\}$):*
   (a) *For each edge from stack $\mathbb{T}_i$, take all its assigned global colors and merge them.*

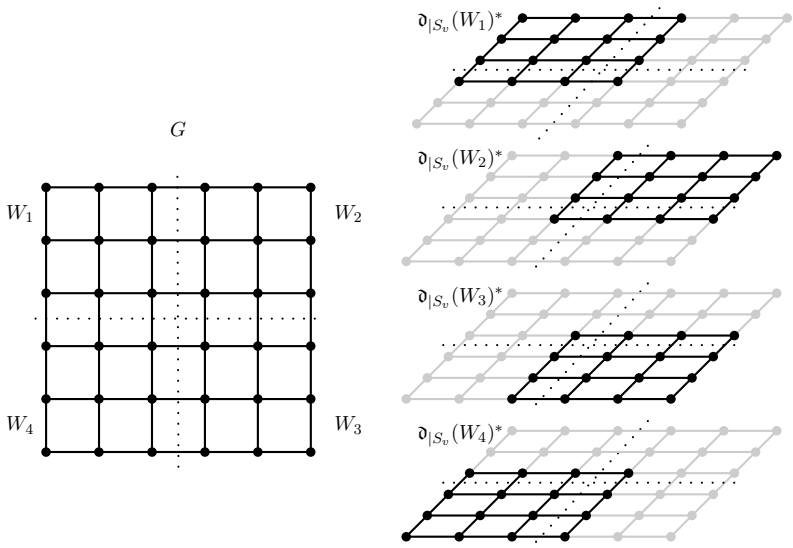


Figure 3: Example - Parallel recognition of δ^*.

Then algorithm 4.2 can be used to compute the colorings $\mathfrak{d}_{|S_v}(W_1)^*, \mathfrak{d}_{|S_v}(W_2)^*, \ldots, \mathfrak{d}_{|S_v}(W_k)^*$, where every instance of the algorithm can run in parallel. The resulting global colorings are used to compute $\mathfrak{d}_{|S_v}(V(G))^* = (\mathfrak{d}_{|S_v}(W_1)^* \cup \mathfrak{d}_{|S_v}(W_2)^* \cup \cdots \cup \mathfrak{d}_{|S_v}(W_k)^*)^*$. The sketch of the parallelization is summarized in Algorithm 4.3.

Figure 3 shows an example of decomposed vertex set of a given graph G. The computation of global colorings associated with the individual sets of the partition can be done then in parallel. The edges that are colored by global color when the partition is treated are highlighted by bold black color. Thus we can observe that many edges will be colored by more then one color.

Notice that we do not treat the task of finding a good partition. With the methods of [4] this is possible with high probability in $O(\log n)$ time, where n is the number of vertices.

References

[1] W. Dörfler, Some results on the reconstruction of graphs, in: *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. I*, North-Holland, Amsterdam, pp. 361–363. Colloq. Math. Soc. János Bolyai, Vol. 10, 1975.

[2] J. Feigenbaum, *Product graphs: some algorithmic and combinatorial results*, Technical Report STAN-CS-86-1121, Stanford University, Computer Science, 1986, phD Thesis.

[3] J. Hagauer and J. Žerovnik, An algorithm for the weak reconstruction of Cartesian-product graphs, *J. Combin. Inform. System Sci.* **24** (1999), 87–103.

[4] S. Halperin and U. Zwick, Optimal randomized EREW PRAM algorithms for finding spanning forests, *J. Algorithms* **39** (2001), 1–46, doi:10.1006/jagm.2000.1146.

[5] R. Hammack, W. Imrich and S. Klavžar, *Handbook of Product Graphs*, Discrete Mathematics and its Applications, CRC Press, 2nd edition, 2011.

[6] M. Hellmuth, A local prime factor decomposition algorithm, *Discrete Math.* **311** (2011), 944–965.

[7] M. Hellmuth, W. Imrich, W. Klöckl and P. F. Stadler, Approximate graph products, *European J. Combin.* **30** (2009), 1119 – 1133.

[8] M. Hellmuth, W. Imrich and T. Kupka, Partial star products: A local covering approach for the recognition of approximate Cartesian product graphs, *Math. Comput. Sci* **7** (2013), 255–273.

[9] W. Imrich and S. Klavžar, *Product graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[10] W. Imrich and I. Peterin, Recognizing Cartesian products in linear time, *Discrete Math.* **307** (2007), 472 – 483.

[11] W. Imrich, T. Pisanski and J. Žerovnik, Recognizing Cartesian graph bundles, *Discrete Math.* **167-168** (1997), 393–403.

[12] W. Imrich and J. Žerovnik, Factoring Cartesian-product graphs, *J. Graph Theory* **18** (1994), 557–567, doi:10.1002/jgt.3190180604.

[13] T. Kupka, *A local approach for embedding graphs into Cartesian products*, Ph.D. thesis, VSB-Technical University of Ostrava, 2013.

[14] T. Pisanski, B. Zmazek and J. Žerovnik, An algorithm for k-convex closure and an application, *Int. J. Comput. Math.* **78** (2001), 1–11, doi:10.1080/00207160108805092.

[15] G. Sabidussi, Graph multiplication, *Math. Z.* **72** (1960), 446–457.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 243–259

Rational sums of hermitian squares of free noncommutative polynomials

Kristijan Cafuta *
University of Ljubljana, Faculty of Electrical Engineering,
Laboratory of Applied Mathematics,
Tržaška cesta 25, 1000 Ljubljana, Slovenia

Igor Klep †
The University of Auckland, Department of Mathematics,
Private Bag 92019, Auckland 1142, New Zealand

Janez Povh ‡
Faculty of Information studies in Novo mesto,
Novi trg 5, 8000 Novo mesto, Slovenia

Received 25 July 2013, accepted 13 March 2014, published online 11 January 2015

Abstract

In this paper we consider polynomials in noncommuting variables that admit sum of hermitian squares and commutators decompositions. We recall algorithms for finding decompositions of this type that are based on semidefinite programming. The main part of the article investigates how to find such decomposition with *rational coefficients* if the original polynomial has rational coefficients. We show that the numerical evidence, obtained by the Gram matrix method and semidefinite programming, which is usually an almost feasible point, can be frequently tweaked to obtain an *exact certificate* using rational numbers. In the presence of Slater points, the Peyrl-Parrilo rounding and projecting method applies. On the other hand, in the absence of strict feasibility, a variant of the facial reduction is proposed to reduce the size of the semidefinite program and to enforce the existence of Slater points. All these methods are implemented in our open source computer algebra package `NCSOStools`. Throughout the paper many worked out examples are presented to illustrate our results.

*Partially supported by the Slovenian Research Agency grants P1-0222 and L1-6722.

†Supported by the Marsden Fund Council of the Royal Society of New Zealand. Partially supported by the Slovenian Research Agency grants P1-0222, L1-4292 and L1-6722. Part of this research was done while the author was on leave from the University of Maribor.

‡The author wishes to thank to Slovenian research agency for support via program P1-0383 and project L74119 and to Creative Core FISNM-3330-13-500033 Simulations project funded by the European Union.

This work is licensed under http://creativecommons.org/licenses/by/3.0/

Keywords: Sum of squares, semidefinite programming, noncommutative polynomial, Matlab toolbox, commutator, cyclic equivalence, free positivity, real algebraic geometry, Motzkin polynomial, Bessis-Moussa-Villani (BMV) conjecture, NCSOStools.

Math. Subj. Class.: Primary 13J30, 90C22; Secondary 08B20, 11E25, 90C90

1 Introduction

In this paper we consider free noncommutative (nc) polynomials that are sums of hermitian squares (and commutators). We focus on the following important question: how to obtain a rational certificate (i.e., a symbolic proof) for such a decomposition when the given nc polynomial has rational coefficients and we have numerical (approximate) evidence of a sum of hermitian squares (and commutators) decomposition obtained by mathematical optimization methods (e.g. by using open-source software package `NCSOStools`)?

1.1 Notation

Nc polynomials with real coefficients, denoted by $\mathbb{R}\langle \underline{X}\rangle$, are (real) linear combinations of words in letters $X_1, \ldots, X_n$, including the empty word 1. We shortly denote by $\underline{X}$ the n-tuple of letters $(X_1, \ldots, X_n)$. These nc polynomials form a free algebra, which we equip with the *involution* $*$ that fixes $\mathbb{R}$ and letters point-wise and thus reverses words, e.g. $(X_1X_2X_3 - X_3^2X_1)^* = X_3X_2X_1 - 2X_1X_3^2$. Hence $\mathbb{R}\langle \underline{X}\rangle$ is the $*$-algebra freely generated by n symmetric letters. The subset of $\mathbb{R}\langle \underline{X}\rangle$ consisting of all symmetric nc polynomials is denoted by

$$\operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle := \{f \in \mathbb{R}\langle \underline{X}\rangle \mid f = f^*\}.$$

If $V = (v_i)$ is a (column) vector of nc polynomials $v_i \in \mathbb{R}\langle \underline{X}\rangle$, then V^* is the row vector with components v_i^* and V^t denotes the row vector with components v_i.

The length of the longest word in an nc polynomial $f \in \mathbb{R}\langle \underline{X}\rangle$ is the *degree* of f and is denoted by $\deg f$. The degree of f in X_i, $\deg_i f$, is the largest number of occurrences of the letter X_i in a monomial appearing in f. Similarly, the length of the shortest word appearing in $f \in \mathbb{R}\langle \underline{X}\rangle$ is called the *min-degree* of f and denoted by $\operatorname{mindeg} f$. Likewise, $\operatorname{mindeg}_i f$ is introduced. If the variable X_i does not occur in any monomial of f, then $\operatorname{mindeg}_i f = 0$. The set of all nc polynomials of degree $\leq d$ will be denoted by $\mathbb{R}\langle \underline{X}\rangle_{\leq d}$. Whenever an nc polynomial f involves only two variables, we write $f \in \mathbb{R}\langle X, Y\rangle$.

Example 1.1. Let $f = 3Y^2X + 2XYXY - 5Y^3 \in \mathbb{R}\langle X, Y\rangle$. Then

$$\deg f = 4,\ \deg_X f = 2,\ \deg_Y f = 3,\ \operatorname{mindeg} f = 3,\ \operatorname{mindeg}_X f = 0,\ \operatorname{mindeg}_Y f = 2,$$

$$f^* = 3XY^2 + 2YXYX - 5Y^3.$$

Positivity of nc polynomials is a core part of *free real algebraic geometry*. In this paper we consider two types of positivity: (i) *positivity via eigenvalues*, i.e., $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ is positive if $f(\underline{A})$ is a positive semidefinite matrix for every n-tuple of real symmetric matrices $\underline{A}$ of the same order; (ii) *trace positivity*, i.e., $f \in \mathbb{R}\langle \underline{X}\rangle$ is trace positive if

E-mail addresses: kristijan.cafuta@fe.uni-lj.si (Kristijan Cafuta), igor.klep@auckland.ac.nz (Igor Klep), janez.povh@fis.unm.si (Janez Povh)

$\operatorname{tr} f(\underline{A}) \geq 0$ for every n-tuple of real symmetric matrices $\underline{A}$ of the same order. Note that positivity implies trace positivity while the converse is not true.

Helton [16] and McCullough [27] proved that a symmetric nc polynomial f is positive if and only if it can be decomposed as a *sum of hermitian squares* (SOHS), that is, there exist nc polynomials $g_1, \ldots, g_m$ such that $f = \sum_{i=1}^{m} g_i^* g_i$. We denote all nc polynomials that admit SOHS decompositions as

$$\Sigma^2 := \Big\{ f \in \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle \mid f = \sum_{i=1}^{m} g_i^* g_i,\ g_i \in \mathbb{R}\langle \underline{X} \rangle,\ m \geq 1 \Big\}.$$

For trace positivity there is no necessary and sufficient condition of this type but there exists an important sufficient condition, obtained using cyclic equivalence to SOHS [18]; for a more example specific approach to certificates for trace positivity we refer to [36]. Nc polynomials $f, g \in \mathbb{R}\langle \underline{X} \rangle$ are *cyclically equivalent* ($f \overset{\text{cyc}}{\sim} g$) if and only if there exist nc polynomials $p_i, q_i \in \mathbb{R}\langle \underline{X} \rangle$ such that

$$f - g = \sum_{i=1}^{k} (p_i q_i - q_i p_i).$$

We call an element of the form $[p, q] := pq - qp$, where $p, q \in \mathbb{R}\langle \underline{X} \rangle$, a *commutator*. Cyclically equivalent nc polynomials have equal trace if they are evaluated at the same n-tuple of real symmetric matrices, since the trace of every commutator of matrices is zero. Therefore if f is cyclically equivalent to SOHS, it is trace positive. We denote the set of nc polynomials of this type by

$$\Theta^2 := \big\{ f \in \mathbb{R}\langle \underline{X} \rangle \mid \exists g \in \Sigma^2 :\ f \overset{\text{cyc}}{\sim} g \big\}.$$

By definition, the elements in Θ^2 are exactly the nc polynomials which can be written as sums of hermitian squares with commutators.

Although any bivariate nc polynomial of degree at most 4 is trace positive if and only if it is a sum of (four) squares with commutators [5, 8], there are trace positive nc polynomials which are *not* members of Θ^2. Probably the easiest example is the noncommutative Motzkin polynomial, $XY^4X + YX^4Y - 3XY^2X + 1$ [18, Example 4.4]; see also Subsection 3.3.2. We also refer the reader to [19, Example 3.5] for more sophisticated examples obtained by considering the BMV conjecture.

Cyclic equivalence is obviously an equivalence relation. It can be easily detected by the following remark.

Remark 1.2 ([18])**.**

(a) For words $v, w \in \langle \underline{X} \rangle$, we have $v \overset{\text{cyc}}{\sim} w$ if and only if there are words $v_1, v_2 \in \langle \underline{X} \rangle$ such that $v = v_1 v_2$ and $w = v_2 v_1$. That is, $v \overset{\text{cyc}}{\sim} w$ if and only if w is a cyclic permutation of v.

(b) Nc polynomials $f = \sum_{w \in \langle \underline{X} \rangle} a_w w$ and $g = \sum_{w \in \langle \underline{X} \rangle} b_w w$ ($a_w, b_w \in \mathbb{R}$) are cyclically equivalent if and only if for each word $v \in \langle \underline{X} \rangle$,

$$\sum_{\substack{w \in \langle \underline{X} \rangle \\ w \overset{\text{cyc}}{\sim} v}} a_w = \sum_{\substack{w \in \langle \underline{X} \rangle \\ w \overset{\text{cyc}}{\sim} v}} b_w. \tag{1.1}$$

Example 1.3. Let $f = 1 + X^2 + 2X^2Y - 2XY + 2XY^2X \in \mathbb{R}\langle X, Y\rangle$. Since

$$f = (X + XY)^*(X + XY) + (1 - YX)^*(1 - YX) + [X^2 - X, Y] + [XY, YX]$$

it follows that

$$f \overset{\text{cyc}}{\sim} (X + XY)^*(X + XY) + (1 - YX)^*(1 - YX),$$

and therefore $f \in \Theta^2$.

1.2 Motivation and related work

There is s surge of interest in free real algebraic geometry in the last decade, partially due to many facets of applications. A nice survey on connections to control theory, systems engineering and optimization is given by de Oliveira, Helton, McCullough, Putinar [13]. Applications to quantum physics are explained e.g. by Pironio, Navascués, Acín [32] who also consider computational aspects related to sums of hermitian squares. On the theoretical level, trace positive nc polynomials arise e.g. in the Lieb-Seiringer reformulation of the famous Bessis-Moussa-Villani (BMV) conjecture [2] from statistical quantum mechanics, which was recently proved by Stahl [39]. This connection will be explained in detail later to demonstrate the usage of our proposed algorithm. In addition, trace positive nc polynomials occur naturally in von Neumann algebras and functional analysis. For instance, Connes' embedding problem [12] on finite II_1-factors is a question about the existence of a certain type of sum of hermitian squares certificates for trace positive nc polynomials [18]. Motivated by this intensive research in free real algebraic geometry we have developed `NCSOStools` [10] – an open source Matlab toolbox for solving such problems using *semidefinite programming*. As a side product our toolbox implements symbolic computation with free noncommuting variables in Matlab.

1.3 Contribution

The main contribution of this paper is the following. Once we know that a given *rational* nc polynomial f can be decomposed as a sum of hermitian squares (with commutators), i.e., we have numerical evidence for the existence of such a decomposition, we aim to obtain an *exact* (rational) certificate. Following ideas from [31] (see also [17]) we propose an algorithm which under a strict feasibility assumption theoretically and practically always yields a rational certificate. On the other hand, in the absence of strict feasibility, a variant of the facial reduction [3] (in our case projecting onto the orthogonal complement of the nullspace of the analytic center) is used to reduce the size of the semidefinite program and enforce the existence of Slater points. We employ the noncommutative version of Motzkin's polynomial to demonstrate how the proposed algorithm as implemented in `NCSOStools` is used and provide new rational certificates for some instances of nc polynomials related to the Bessis-Moussa-Villani conjecture.

2 Nc polynomials and semidefinite programming

2.1 Semidefinite programming

Semidefinite programming (SDP) is a generalization of linear programming (LP) where one looks for the optimum of a linear function over the intersection of an affine subspace

with the cone of positive semidefinite matrices. Although this is a far-reaching extension of LP, there exists several methods that can solve semidefinite programs efficiently in theory and practice. Given $s \times s$ self-adjoint matrices $C,\ A_1, \ldots,\ A_m$ of the same size over $\mathbb{R}$ and a vector $b \in \mathbb{R}^m$, we formulate a *semidefinite program in standard primal form* as follows:

$$\begin{array}{lrcl} \inf & \langle C, G\rangle & & \\ \text{s.t.} & \langle A_i, G\rangle & = & b_i, \quad i = 1, \ldots, m \\ & G & \succeq & 0. \end{array} \tag{PSDP}$$

Here $\langle \cdot, \cdot \rangle$ stands for the standard inner product of matrices: $\langle A, B\rangle = \operatorname{tr}(B^*A)$, and $G \succeq 0$ means that G is positive semidefinite. If $C = 0$ or if C is not important, we call such a problem a semidefinite programming feasibility problem:

$$\begin{array}{lrcl} & G & \succeq & 0, \\ \text{s.t.} & \langle A_i, G\rangle & = & b_i, \quad i = 1, \ldots, m. \end{array} \tag{FSDP}$$

The complexity of solving semidefinite programs is mainly determined by the order s of matrix variable G and the number of linear constraints m. Given $\varepsilon > 0$, the interior point methods can find an ε-optimal solution with polynomially many iterations, where each iteration takes polynomially many real number operations, provided that (PSDP) and its dual both have non-empty interiors of feasible sets and we have good initial points. The variables appearing in these polynomial bounds are s, m and $\log \varepsilon$ (cf. [40, Chapter 10.4.4]).

Many problems in control theory, system identification and signal processing can be formulated using SDPs [4, 30, 1]. Combinatorial optimization problems can often be modeled or approximated by SDPs [14, 23, 34, 35, 33]. SDP has important role in real algebraic geometry, where it is used e.g. for finding sums of squares decompositions of polynomials or approximating the moment problem [22, 21, 26, 24], and in free real algebraic geometry [18, 20, 6], as is recalled in the following subsection.

2.2 Sums of hermitian squares (with commutators) and semidefinite programming

Testing whether a given nc polynomial $f \in \mathbb{R}\langle \underline{X}\rangle$ is an element of Σ^2 can be done efficiently by using semidefinite programming [20, 10]. This is the Gram matrix method, which is based on the following proposition [16, 28], the noncommutative version of the classical result for commuting variables.

Proposition 2.1. *Suppose the nc polynomial $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ is of degree $\leq 2d$ and let W_d be the vector of all words $w \in \langle \underline{X}\rangle$ of degree $\leq d$. Then $f \in \Sigma^2$ if and only if there exists a positive semidefinite matrix G_f (called* a Gram matrix *for f) satisfying $f = W_d^* G_f W_d$.*

Example 2.2. Take $f = 1 + X^2 + XY + YX + 4YX^2Y + Y^2$ and let $V = [1 \quad X \quad Y \quad XY]^t$ be a subvector of W_2. Then the Gram matrix for nc polynomial f corresponding to the vector V is

$$G(u) := \begin{bmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 1-u & 0 \\ 0 & 1-u & 1 & 0 \\ u & 0 & 0 & 4 \end{bmatrix}.$$

The question is: does there exist (at least one) u such that $G(u)$ is a positive semidefinite matrix? Since $G(2) = C^tC$ for

$$C = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 \end{bmatrix},$$

it follows that $f = (1+2XY)^*(1+2XY) + (X-Y)^*(X-Y) \in \Sigma^2$.

As we saw in the last example we can sometimes replace W_d with a smaller subvector in the Gram matrix method. An algorithm (the Newton chip method) for reducing the size of needed word vector is presented in [20] and is implemented in `NCSOStools`. See also [29] for a strengthening.

Similarly we can use semidefinite programming to test whether a given nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is an element of Θ^2 as first observed in [19], see also [10, 7, 6]. The method behind it is a variant of the Gram matrix method:

Proposition 2.3. *Suppose that an nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is of degree $\leq 2d$ and let W_d be as above. Then $f \in \Theta^2$ if and only if there exists a positive semidefinite matrix G_f (called* a tracial Gram matrix *for f) such that $f \overset{\text{cyc}}{\sim} W_d^* G_f W_d$.*

Again we can sometimes replace the full word vector W_d with a smaller subvector. An algorithm (the Newton cyclic chip method) for reducing the size of needed word vector is presented in [6] and is implemented in `NCSOStools`.

Following Proposition 2.1, we can decide whether an nc polynomial f is a sum of hermitian squares by solving a semidefinite programming feasibility problem in the matrix variable G, where the constraints $\langle A_i, G \rangle = b_i$ are implied by the fact that for each product of monomials $w \in \{p^*q \mid p, q \in W\}$ the following must be true:

$$\sum_{\substack{p,q \in W \\ p^*q = w}} G_{p,q} = a_w, \tag{2.1}$$

where a_w is the coefficient of w in f ($a_w = 0$ if the monomial w does not appear in f). Since any input nc polynomial f is symmetric (so $a_w = a_{w^*}$ for all w), the corresponding SDP feasibility problem is as follows:

$$\begin{array}{rcl} G & \succeq & 0 \\ \text{s.t.} \quad \langle A_w, G \rangle & = & a_w + a_{w^*} \quad \forall w \in \{p^*q \mid p, q \in W\}, \end{array} \tag{SOHS$_{\text{SDP}}$}$$

where $A_w = A_{w^*}$ is the symmetric matrix defined by

$$(A_w)_{u,v} = \begin{cases} 2; & \text{if } u^*v \in \{w, w^*\},\ w^* = w, \\ 1; & \text{if } u^*v \in \{w, w^*\},\ w^* \neq w, \\ 0; & \text{otherwise.} \end{cases}$$

Similarly, following Proposition 2.3, an nc polynomial f is cyclically equivalent to a sum of hermitian squares if and only if there exists a positive semidefinite matrix G such that $f \overset{\text{cyc}}{\sim} W^*GW$. Again, this is an SDP feasibility problem (FSDP) in the matrix variable G, where the constraints $\langle A_i, G \rangle = b_i$ are essentially equations (1.1), i.e., for each product of monomials $v \in \{p^*q \mid p, q \in W\}$ the following must be true:

$$\sum_{\substack{p,q \in W \\ p^*q \overset{\text{cyc}}{\sim} v}} G_{p,q} = \sum_{w \overset{\text{cyc}}{\sim} v} a_w. \tag{2.2}$$

The SDP feasibility problem is as follows [6, Corollary 4.5]:

$$\begin{aligned} & & G &\succeq 0, \\ \text{s.t.} & \sum_{\substack{p,q,\ p^*q \overset{\text{cyc}}{\sim} v \\ \vee\ p^*q \overset{\text{cyc}}{\sim} v^*}} \langle A_v,\, G\rangle &= \sum_{w \overset{\text{cyc}}{\sim} v} (a_w + a_{w^*}),\ \ \forall v \in W \end{aligned} \qquad (\text{CSOHS}_{\text{SDP}})$$

where $A_v = A_{v^*}$ is the symmetric matrix defined by

$$(A_v)_{p,q} = \begin{cases} 2; & \text{if } p^*q \overset{\text{cyc}}{\sim} v \ \&\ p^*q \overset{\text{cyc}}{\sim} v^*, \\ 1; & \text{if } p^*q \overset{\text{cyc}}{\sim} v \ \&\ p^*q \overset{\text{cyc}}{\not\sim} v^*, \\ 0; & \text{otherwise.} \end{cases}$$

Remark 2.4. Finding a Gram matrix for the sum of hermitian squares (and commutators) decomposition problem by solving (SOHS_{SDP}) and ($\text{CSOHS}_{\text{SDP}}$) gives a solution of *highest rank* since under a strict feasibility assumption the interior point methods yield solutions in the relative interior of the optimal face, which is in our case the whole feasibility set. If strict complementarity is additionally provided, the interior point methods lead to the analytic center of the feasibility set [15].

Alternately, we can consider these SDP problems as usual SDP problems by using a non-zero choice of C. The choice $C = I$ is a commonly used heuristic for matrix rank minimization [37], and it tends to give sum of hermitian squares (and commutators) with a small number of hermitian squares.

Even though the above assumptions do not always hold for the instances of SDPs we construct, in our experiments the choice $C = 0$ in the objective function almost always gave a solution of higher rank than the choice $C = I$. High ranks are desired and exploited when trying to compute a rational (exact) Gram matrix from numerical solution of (SOHS_{SDP}) and ($\text{CSOHS}_{\text{SDP}}$).

3 Rational sums of hermitian squares and facial reduction

In this section particular emphasis is given to the extraction of *rational* certificates if the input data is rational. We present several examples illustrating our results, e.g. concerning the recently proven BMV conjecture [39] from statistical physics (Subsection 3.3.1) and the noncommutative Motzkin polynomial (Subsection 3.3.2).

3.1 Rational sums of hermitian squares

Consider a feasibility SDP in primal form (FSDP) and assume the input data A_i, b_i is rational for $i = 1, \ldots, m$. If the problem is feasible, does there exist a *rational* solution? If so, can one use a combination of numerical and symbolic computation to produce one?

Example 3.1. Some caution is necessary, as a feasible SDP of the form (FSDP) need not admit a rational solution. For a simple concrete example, note that

$$\begin{bmatrix} 2 & x \\ x & 1 \end{bmatrix} \oplus \begin{bmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x \end{bmatrix} \succeq 0 \quad \Leftrightarrow \quad x = \sqrt{2}.$$

In fact there are commutative polynomials with rational coefficients that are sums of squares of polynomials over the reals, but not over the rationals (see [38]). Adapting an example of

Scheiderer, we obtain an nc polynomial with rational coefficients that is cyclically equivalent to a sum of hermitian squares of nc polynomials over the reals, but not over the rationals:

$$f = 1 + X^3 + X^4 - \frac{3}{2}XY - \frac{3}{2}YX - 4XYX + 2Y^2 + Y^3 + \frac{1}{2}XY^3 + \frac{1}{2}Y^3X + Y^4.$$

This is a dehomogenized and symmetrized noncommutative version of the (commutative) polynomial from [38, Theorem 2.1] (setting $x_0 = 1$, $x_1 = X$ and $x_2 = Y$). So f is not cyclically equivalent to a sum of hermitian squares with rational coefficients. By [38, Theorem 2.1], $f|_{\mathbb{R}^2} \geq 0$. Together with the fact that f is cyclically sorted, [18, Proposition 4.2] implies that f is trace positive. Since f is of degree 4 in two variables it is a sum of hermitian squares with commutators [5, 8] (with real coefficients).

On the other hand, if (FSDP) admits a feasible *positive definite* solution, then it admits a (positive definite) *rational* solution. More exactly, we have the following:

Theorem 3.2 (Peyrl & Parrilo [31])**.** *If an approximate feasible point G_0 for* (FSDP) *satisfies*

$$\delta := \min(\mathrm{eig}(G_0)) > \|(\langle A_i, G_0\rangle - b_i)_i\| =: \varepsilon, \tag{3.1}$$

then a (*positive definite*) *rational feasible point G exists. It can be obtained from G_0 in the following two steps* (*cf. Figure* 1)*:*

(1) *compute a rational approximation $\tilde{G}$ of G_0 with $\tau := \|\tilde{G} - G_0\|$ satisfying $\tau^2 + \varepsilon^2 < \delta^2$;*

(2) *project $\tilde{G}$ onto the affine subspace $\mathcal{L}$ given by the equations $\langle A_i, G\rangle = b_i$ to obtain G.*

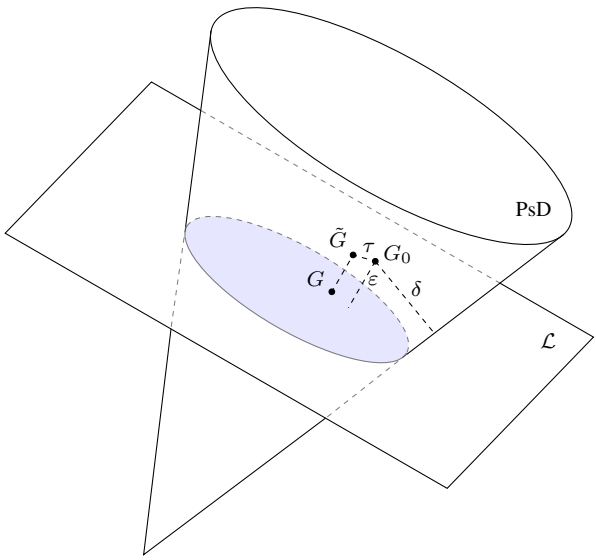


Figure 1: Rounding and projecting to obtain a rational solution

Note that the results in [31] are stated for SDPs arising from sum of squares problems, but their results carry over verbatim to the setting of (the seemingly more) general SDPs.

The rationalization scheme based on this Peyrl-Parrilo technique has been implemented in `NCSOStools`; see Example 3.5 for a demonstration.

3.2 Facial reduction

Not all is lost, however, if the SDP solver gives a *singular* feasible point G_0 for (FSDP). Suppose that z is a *rational* nullvector for G_0. Let P be a change of basis matrix containing z as a first column and a (rational) orthogonal basis for the orthogonal complement $\{z\}^\perp$ as its remaining columns. Then

$$P^t G_0 P = \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix},$$

i.e.,

$$G_0 = P^{-t} \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix} P^{-1}$$

for some symmetric $\hat{G}_0$. Hence

$$b_i = \langle A_i, G_0 \rangle = \operatorname{tr}(A_i G_0) = \operatorname{tr}\left(A_i P^{-t} \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix} P^{-1} \right) = \operatorname{tr}\left(P^{-1} A_i P^{-t} \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix} \right).$$

So if

$$P^{-1} A_i P^{-t} = \begin{bmatrix} a_i & c_i^t \\ c_i & \hat{A}_i \end{bmatrix}$$

then $\hat{A}_i$ is a symmetric matrix with rational entries and

$$b_i = \operatorname{tr}\left(\begin{bmatrix} a_i & c_i^t \\ c_i & \hat{A}_i \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix} \right) = \operatorname{tr}(\hat{A}_i \hat{G}_0) = \langle \hat{A}_i, \hat{G}_0 \rangle.$$

We have established a variant of the facial reduction [3] which applies whenever the original SDP is given by rational data and has a singular feasible point with a rational nullvector:

Theorem 3.3. *Let* (FSDP)*, $\hat{A}_i$ and $\hat{G}_0$ be as above. Consider the feasibility SDP*

$$\begin{array}{rcl} \hat{G} & \succeq & 0 \\ \text{s.t.} \quad \langle \hat{A}_i, \hat{G} \rangle & = & b_i, \quad i = 1, \ldots, m \end{array} \tag{FSDP'}$$

(1) (FSDP') *is feasible if and only if* (FSDP) *is feasible.*

(2) (FSDP') *admits a rational solution if and only if* (FSDP) *does.*

3.3 Examples

3.3.1 BMV conjecture

In their 2004 paper [25], Lieb and Seiringer gave the following purely algebraic reformulation of the Bessis-Moussa-Villani (BMV) conjecture [2] from quantum statistical physics, which was recently proved in the original formulation by Stahl [39]:

Conjecture 3.4. *For all positive semidefinite matrices A and B and all $m \in \mathbb{N}$, the polynomial $p(t) := \operatorname{tr}((A+tB)^m) \in \mathbb{R}[t]$ has only nonnegative coefficients.*

The coefficient of t^k in $p(t)$ for a given m is the trace of $S_{m,k}(A,B)$, where $S_{m,k}(A,B)$ is the sum of all words of length m in the letters A and B in which B appears exactly k times. For example, $S_{4,2}(A,B) = A^2B^2 + ABAB + AB^2A + BABA + B^2A^2 + BA^2B$. Thus $S_{m,k}(X,Y)$ is an nc polynomial; it is the sum of all words in two variables X, Y of degree m in which Y appears exactly k times.

Even though the motivating conjecture was proved, the related questions concerning nc polynomials remain interesting. In the last few years there has been much activity around the following question: which pairs (m,k) does $S_{m,k}(X^2,Y^2) \in \Theta^2$ or $S_{m,k}(X,Y) \in \Theta^2$ hold for? An affirmative answer (for all m,k) to the former would imply the BMV conjecture. This question has been resolved completely (see e.g. [19, 11, 9]), however only finitely many nontrivial $S_{m,k}(X^2,Y^2)$ admit a Θ^2-certificate. Adding to the current state of knowledge (nicely summarized in [11]), we shall use our computer algebra system `NCSOStools` to establish $S_{10,2}(X,Y) \in \Theta^2$ and $S_{14,6}(X,Y) \notin \Theta^2$. We also show that $S_{2m,2}(X,Y) \in \Theta^2$ holds for all $m \in \mathbb{N}$.

Example 3.5. Consider the nc polynomial $f = S_{10,2}(X,Y)$, i.e., the sum of all words of degree 10 in the nc variables X and Y in which Y appears exactly twice. To prove that $f \in \Theta^2$ with the aid of `NCSOStools`, proceed as follows:

(1) Define two noncommuting variables:

```
>> NCvars x y
```

(2) Our nc polynomial f is constructed using `BMV(10,2)`. For a numerical test whether $f \in \Theta^2$, run

```
>> p.obj = 0;
>> [IsCycEq,G0,W,sohs,g,SDP_data] = NCcycSos(BMV(10,2),p);
```

Using the SDP solver SDPT3, this yields a *floating point* Gram matrix G_0

$$G_0 = \begin{bmatrix} 5.0000 & 2.5000 & -1.8851 & 0.8230 & -0.0899 \\ 2.5000 & 8.7702 & 1.6770 & -2.7313 & 0.8230 \\ -1.8851 & 1.6770 & 10.6424 & 1.6770 & -1.8851 \\ 0.8230 & -2.7313 & 1.6770 & 8.7702 & 2.5000 \\ -0.0899 & 0.8230 & -1.8851 & 2.5000 & 5.0000 \end{bmatrix}$$

for the word vector

$$W = \begin{bmatrix} X^4Y & X^3YX & X^2YX^2 & XYX^3 & YX^4 \end{bmatrix}^t.$$

The rest of the output: `IsCycEq` $= 1$ since f is (numerically) an element of Θ^2; `sohs` is a vector of nc polynomials g_i with $f \overset{\text{cyc}}{\sim} \sum_i g_i^* g_i =$ `g`; `SDP_data` is the SDP data for (2.2) constructed from f.

(3) To round and project the obtained floating point solution `G0` following Theorem 3.2, feed `G0` and `SDP_data` into `RprojRldlt`:

```
>> [G,L,D,P,err]=RprojRldlt(G0,SDP_data,true)
```

This produces a rational Gram matrix `G` for f with respect to W and its LDU decomposition $PLDL^tP^t$, where P is a permutation matrix, L lower unitriangular, and D a diagonal matrix with positive entries. We caution the reader that `L`, `D`, and `G` are cells, each containing numerators and denominators separately as a matrix. Finally, the obtained rational sum of hermitian squares certificate for $f = S_{10,2}(X,Y)$ is

$$f \overset{\text{cyc}}{\sim} \sum_{i=1}^{5} \lambda_i g_i^* g_i$$

for

$$\begin{aligned}
g_1 &= X^2YX^2 + \frac{7}{44}X^3YX + \frac{7}{44}XYX^3 - \frac{2}{11}X^4Y - \frac{2}{11}YX^4\\
g_2 &= X^3YX - \frac{577}{1535}XYX^3 + \frac{408}{1535}X^4Y + \frac{188}{1535}YX^4\\
g_3 &= XYX^3 + \frac{11909}{45984}X^4Y + \frac{7613}{15328}YX^4\\
g_4 &= X^4Y - \frac{296301}{647065}YX^4\\
g_5 &= YX^4
\end{aligned}$$

and

$$\lambda_1 = 11, \quad \lambda_2 = \frac{1535}{176}, \quad \lambda_3 = \frac{11496}{1535}, \quad \lambda_4 = \frac{647065}{183936}, \quad \lambda_5 = \frac{1242629}{647065}.$$

This example is not surprising, as it is a particular instance of a larger pattern:

Proposition 3.6. *For all $m \in \mathbb{N}$ we have: $S_{2m,2}(X,Y) \in \Theta^2$.*

Proof. We first point out that for all $m \in \mathbb{N}$ we have

$$\begin{aligned}
S_{2m,2}(X,Y) &= \sum_{\alpha+\beta\le 2m-2} X^\alpha YX^\beta YX^{2m-2-\alpha-\beta}\\
&\overset{\text{cyc}}{\sim} \sum_{t=0}^{2m-2} (2m-2-t+1)YX^tYX^{2m-2-t}\\
&\overset{\text{cyc}}{\sim} \frac{1}{2}\sum_{t=0}^{2m-2} (2m-2-t+1)(YX^tYX^{2m-2-t} + YX^{2m-2-t}YX^t)\\
&= \frac{1}{2}\sum_{t=0}^{2m-2} \Big((2m-2-t+1)YX^tYX^{2m-2-t} + (t+1)YX^tYX^{2m-2-t}\Big)\\
&\overset{\text{cyc}}{\sim} m\sum_{t=0}^{2m-2} YX^tYX^{2m-2-t}.
\end{aligned}$$

Note that for $t = 2s$ we have $YX^tYX^{2m-2-t} \overset{\text{cyc}}{\sim} X^{m-s-1}YX^{2s}YX^{m-s-1} \in \Sigma^2$, hence we next turn our attention to words YX^tYX^{2m-2-t} for odd t. In such cases we write $t = 2s+1$ and observe that

$$\begin{aligned} &YX^{2s+1}YX^{2m-3-2s} \\ &\overset{\text{cyc}}{\sim} \frac{1}{2}\Big((X^{s+1}YX^{m-s-2} + X^sYX^{m-s-1})^*(X^{s+1}YX^{m-s-2} + X^sYX^{m-s-1})\Big) \\ &\qquad - \frac{1}{2}X^{m-s-2}YX^{2s+2}YX^{m-s-2} - \frac{1}{2}X^{m-s-1}YX^{2s}YX^{m-s-1}. \end{aligned}$$

Therefore each word with odd t is cyclically equivalent to a hermitian square minus two hermitian squares. These two negative hermitian squares cancel out with the "even" words for $t = 2s$ and $t = 2s+2$. In fact, each word with odd t cancels one half of these two even terms, hence all even terms finally cancel out and only one half of the first and the last even term remains (these two terms are cyclically equivalent). Finally we get

$$\begin{aligned} &S_{2m,2}(X,Y) \\ &\overset{\text{cyc}}{\sim} \frac{m}{2}\sum_{t=0}^{2m-2}(X^{s+1}YX^{m-s-2} + X^sYX^{m-s-1})(X^{s+1}YX^{m-s-2} + X^sYX^{m-s-1})^* \\ &\qquad + X^{m-1}Y^2X^{m-1}. \quad \square \end{aligned}$$

Example 3.7. We conclude this subsubsection by showing $S_{14,6}(X,Y) \notin \Theta^2$. We define two noncommuting variables and run `NCcycSos` as in the previous examples:

```
>> NCvars x y
>> [IsCycEq,G0,V,sohs,g,SDP_data] = NCcycSos(BMV(14,6));
```

However, this seems to be an infeasible problem. In fact, we shall use the generated data `SDP_data` to prove it is strongly infeasible by computing a rational hyperplane separating Θ^2 and $S_{14,6}(X,Y)$. Let $\mathcal{P}$ be the set of all nc polynomials p with $\deg_X p = \operatorname{mindeg}_X p = 8$ and $\deg_Y p = \operatorname{mindeg}_Y p = 6$. Obviously, $S_{14,6}(X,Y) \in \mathcal{P}$. Each $p \in \mathcal{P}$ can be represented by a 35×35 Gram matrix using the basis V from given as output of `NCcycSos`. An important observation is that $p \in \Theta^2$ if and only if there is a positive semidefinite matrix G satisfying $p \overset{\text{cyc}}{\sim} V^*GV$, cf. Proposition 2.3.

Let $L : \mathcal{P} \to \mathbb{R}$ be a linear $*$-map nonnegative on $\Theta^2 \cap \mathcal{P}$. It can be represented as $p \mapsto \langle M, G_p\rangle$ for a symmetric 35×35 matrix M, where G_p is a Gram matrix for p. Since $L(\Sigma^2) \subseteq [0,\infty)$, the matrix M is positive semidefinite. The fact that $L(f) = 0$ for all $f \overset{\text{cyc}}{\sim} 0$, can be modeled with constraints $\langle M, H\rangle = 0$ for all $H \in A^\perp$, cf. [9, Section 2.2]. Here, $A^\perp$ is the orthogonal complement of the span of the A_v from Section 2.2 in the set of symmetric matrices. Clearly, it suffices to consider H from a linearly independent generating subset $\mathcal{C}$ of $A^\perp$.

To express $L(S_{14,6}(X,Y)) < 0$, we first compute a Gram matrix for $S_{14,6}(X,Y)$. The matrix $A =$ `SDP_data.A` and vector $b =$ `SDP_data.b` model the linear constraints $\langle A_v, G\rangle = b_v$ for $v \in \langle X,Y\rangle$ with $\deg_X v = 8, \deg_Y v = 6$. Hence a symmetrized solution of the linear system

```
>> SDP_data.A\SDP_data.b
```

will be a Gram matrix G for $S_{14,6}(X,Y)$. Now consider the feasibility SDP

$$\begin{array}{rcl} M & \succeq & 0 \\ \text{s.t. } \langle M, G\rangle & = & -35, \quad \forall H \in \mathcal{C} : \langle M, H\rangle \ = \ 0. \end{array}$$

(Here, -35 is just a convenient scaling factor.) Every feasible point induces a hyperplane separating Θ^2 and $S_{14,6}(X,Y)$. Solving this SDP with SeDuMi (using the trivial objective function $C = 0$) yields a floating point solution `M0` in the relative interior of the optimal face, see Remark 2.4, with minimal eigenvalue $\delta = 0.3426$ and residual norm $\varepsilon = 6.8 \cdot 10^{-9}$. Thus we can find a rational feasible solution `M` as explained in Theorem 3.2, using `RprojRldlt`. This proves $S_{14,6}(X,Y) \notin \Theta^2$.

3.3.2 Noncommutative Motzkin polynomial

The nc polynomial

$$f_{\mathrm{Mot}}(X,Y) = XY^4X + YX^4Y - 3XY^2X + 1 \in \mathbb{R}\langle X, Y\rangle$$

is a noncommutative version of the (commutative) Motzkin polynomial. The Motzkin polynomial is a well-known example of a (commutative) polynomial which is nonnegative on $\mathbb{R}^2$ but is not a sum of squares of polynomials. Similarly, f_{Mot} is an example of trace positive nc polynomial which is not a member of Θ^2 [18, Example 4.4]. Indeed, since the (commutative) Motzkin polynomial is not a sum of squares of polynomials, f_{Mot} is not a member of Θ^2. An alternative proof for trace positivity of $f_{\mathrm{Mot}}(X,Y)$ follows from the fact that $f_{\mathrm{Mot}}(X^3,Y^3) \in \Theta^2$, as we can show with the aid of the facial reduction procedure from Subsection 3.2.

Example 3.8. Consider $f = f_{\mathrm{Mot}}(X^3,Y^3) = X^3Y^{12}X^3 + Y^3X^{12}Y^3 - 3X^3Y^6X^3 + 1$. To prove that $f \in \Theta^2$ with the aid of `NCSOStools`, proceed as follows:

(1) Define two noncommuting variables and the nc polynomial f:

```
>> NCvars x y
>> f = x^3*y^12*x^3 + y^3*x^12*y^3 - 3*x^3*y^6*x^3 + 1;
```

(2) Define a custom vector of monomials W

```
>> W = {''; 'x*y*y'; 'x*x*y'; 'x*x*y*y*y*y';
 'x*x*x*x*y*y'; 'x*x*x*y*y*y*y*y*y'; 'x*x*x*x*y*y*y*y*y';
 'x*x*x*x*x*y*y*y*y'; 'x*x*x*x*x*x*y*y*y'}';
```

(3) For a numerical test whether $f \in \Theta^2$, run

```
>> param.V = W;
[IsCycEq,G0,W,sohs,g,SDP_data] = NCcycSos(f,param);
```

This yields a floating point Gram matrix G_0 that is *singular*.

(4) Try to round and project the obtained floating point solution `G0`, feed `G0` and `SDP_data` into `RprojRldlt`:

```
>> [G,L,D,P,err] = RprojRldlt(G0,SDP_data)
```

This exits with an error, since unlike in Example 3.5, the rounding and projecting alone does not yield a rational feasible point.

(5) Instead, let us reexamine G_0. A detailed look at the matrix reveals three nullvectors. We thus run our interactive procedure which aids the computer in reducing the size of the SDP as in Theorem 3.3.

```
>> [G,SDP_data] = fac_reduct(f,param)
```

This leads the computer to return a floating point feasible point $G_0 \in \mathbb{R}^{9\times 9}$ and the data for this SDP, `SDP_data`. It also stays in interactive mode and the user can inspect the matrix and enter the nullvector z to be used in the dimension reduction. We feed in three nullvectors as a matrix of three columns:

```
K>> z = [0 -1 0; -1 0 0; 0 0 1; 0 -1 0; 0 -1 0; -1 0 0;
 0 0 1; -1 0 0; 0 0 1];
    return
```

Inside the interactive routine this enables the computer to produce a positive definite feasible $\hat{G}_0 \in \mathbb{R}^{6\times 6}$. Hence we exit the interactive routine.

```
K>> stop = 1; return
```

Now, `NCSOStools` uses $\hat{G}_0$ to produce a rational positive semidefinite Gram matrix G for f, which proves $f \in \Theta^2$. Like in the Example 3.5, the solution G is a cell containing two matrices with numerators and denominators of the rational entries of G. The reader can verify that $f \overset{\text{cyc}}{\sim} W^*GW$ exactly by doing rational arithmetic or approximately by computing floating point approximation for G and using floating point arithmetic.

(6) To compute the LDU decomposition $PLDL^tP^t$ for the rational Gram matrix `G` of f with respect to W (where `G,L,D` are cells, each containing numerators and denominators separately as a matrix) run

```
>> [L,D,P] = Rldlt(G)
```

The obtained rational sum of hermitian squares certificate for $f_{\text{Mot}}(X^3,Y^3)$ is then

$$f_{\text{Mot}}(X^3,Y^3) \overset{\text{cyc}}{\sim} \sum_{i=1}^{6} \lambda_i g_i^* g_i$$

for

$$\begin{aligned} g_1 &= 1 - \frac{1}{2}X^2Y^4 - \frac{1}{2}X^4Y^2 \\ g_2 &= XY^2 - \frac{1}{2}X^3Y^6 - \frac{1}{2}X^5Y^4 \\ g_3 &= X^2Y - \frac{1}{2}X^4Y^5 - \frac{1}{2}X^6Y^3 \\ g_4 &= X^2Y^4 - X^4Y^2 \\ g_5 &= X^3Y^6 - X^5Y^4 \\ g_6 &= X^4Y^5 - X^6Y^3 \end{aligned}$$

and

$$\lambda_1 = \lambda_2 = \lambda_3 = 1, \quad \lambda_4 = \lambda_5 = \lambda_6 = \frac{3}{4}.$$

Remark 3.9. We point out that this yields a rational sum of squares certificate for $\check{f}(x^3, y^3)$ where $\check{f}(x, y) = 1 + x^4y^2 + x^2y^4 - 3x^2y^2$ is the commutative Motzkin polynomial.

4 Conclusions

In this paper we considered nc polynomials p in freely noncommuting variables which can be decomposed as a sum of hermitian squares (and commutators) with a special focus on nc polynomials with rational coefficients that admit rational decompositions.

We explained how to obtain rational decompositions in theory and practice: if the related semidefinite programming problems have strictly feasible solutions then the algorithm we proposed - a variant of Peyrl-Parrilo rounding and projecting method - always yields a rational (i.e., exact symbolic) decomposition. In the absence of strict feasibility we proposed a variant of the facial reduction to reduce the size of the semidefinite program and enforce the existence of Slater points.

We implemented both methods in our open source software package `NCSOStools` [10] and demonstrated them on several illustrative examples.

References

[1] M. Anjos and J. Lasserre, *Handbook of Semidefinite, Conic and Polynomial Optimization*, volume 166 of *International Series in Operational Research and Management Science*, Springer, 2012.

[2] D. Bessis, P. Moussa and M. Villani, Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics, *J. Mathematical Phys.* **16** (1975), 2318–2325.

[3] J. Borwein and H. Wolkowicz, Facial reduction for a cone-convex programming problem, *J. Austral. Math. Soc. Ser. A* **30** (1980/81), 369–380.

[4] S. Boyd, L. E. Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Studies in Applied Mathematics, SIAM, 1994.

[5] S. Burgdorf and I. Klep, Trace-positive polynomials and the quartic tracial moment problem, *C. R. Math. Acad. Sci. Paris* **348** (2010), 721–726.

[6] S. Burgdorf, K. Cafuta, I. Klep and J. Povh, Algorithmic aspects of sums of Hermitian squares of noncommutative polynomials, *Comput. Optim. Appl.* **55** (2013), 137–153.

[7] S. Burgdorf, K. Cafuta, I. Klep and J. Povh, The tracial moment problem and trace-optimization of polynomials, *Math. Program.* **137** (2013), 557–578.

[8] K. Cafuta, On matrix algebras associated to sum-of-squares semidefinite programs, *Linear Multilinear Algebra* **61** (2013), 1496–1509.

[9] K. Cafuta, I. Klep and J. Povh, A note on the nonexistence of sum of squares certificates for the Bessis-Moussa-Villani conjecture., *J. Math. Phys.* **51** (2010), 083521, 10.

[10] K. Cafuta, I. Klep and J. Povh, NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials, *Optim. Methods and Softw.* **26** (2011), 363–380, `http://ncsostools.fis.unm.si/`

[11] B. Collins, K. Dykema and F. Torres-Ayala, Sum-of-squares results for polynomials related to the Bessis-Moussa-Villani conjecture, *J. Stat. Phys.* **139** (2010), 779–799.

[12] A. Connes, Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$., *Ann. of Math. (2)* **104** (1976), 73–115.

[13] M. de Oliveira, J. Helton, S. McCullough and M. Putinar, Engineering systems and free semi-algebraic geometry, in: M. Putinar and S. Sullivant (eds.), *Emerging applications of algebraic geometry*, Springer, New York, volume 149 of *IMA Vol. Math. Appl.*, pp. 17–61, 2008.

[14] M. Goemans and D. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of the Association for Computing Machinery* **42** (1995), 1115–1145.

[15] M. Halická, E. de Klerk and C. Roos, On the convergence of the central path in semidefinite optimization, *SIAM J. Optim.* **12** (2002), 1090–1099.

[16] J. Helton, "Positive" noncommutative polynomials are sums of squares, *Ann. of Math. (2)* **156** (2002), 675–694.

[17] E. Kaltofen, B. Li, Z. Yang and L. Zhi, Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients, *J. Symbolic Comput.* **47** (2012), 1–15.

[18] I. Klep and M. Schweighofer, Connes' embedding conjecture and sums of Hermitian squares, *Adv. Math.* **217** (2008), 1816–1837.

[19] I. Klep and M. Schweighofer, Sums of Hermitian squares and the BMV conjecture, *J. Stat. Phys* **133** (2008), 739–760.

[20] I. Klep and J. Povh, Semidefinite programming and sums of hermitian squares of noncommutative polynomials, *J. Pure Appl. Algebra* **214** (2010), 740–749.

[21] J. B. Lassere, *Moments, positive polynomials and their applications*, volume 1 of *Imperial College Press Optimization Series*, Imperial College Press, London, 2009.

[22] J. B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Optim.* **11** (2000/01), 796–817.

[23] M. Laurent and F. Rendl, Semidefinite programming and integer programming, in: G. N. K. Aardal and R. Weismantel (eds.), *Discrete Optimization*, Elsevier, volume 12 of *Handbooks in Operations Research and Management Science*, pp. 393 – 514, 2005.

[24] M. Laurent, Sums of squares, moment matrices and optimization over polynomials, in: *Emerging applications of algebraic geometry*, Springer, New York, volume 149 of *IMA Vol. Math. Appl.*, pp. 157–270, 2009.

[25] E. Lieb and R. Seiringer, Equivalent forms of the Bessis-Moussa-Villani conjecture, *J. Stat. Phys.* **115** (2004), 185–190.

[26] M. Marshall, *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*, American Mathematical Society, Providence, RI, 2008.

[27] S. McCullough, Factorization of operator-valued polynomials in several non-commuting variables, *Linear Algebra Appl.* **326** (2001), 193–203.

[28] S. McCullough and M. Putinar, Noncommutative sums of squares, *Pacific J. Math.* **218** (2005), 167–171.

[29] C. Nelson, *A real nullstellensatz for matrices of non-commutative polynomials*, http://arxiv.org/abs/1305.0799

[30] P. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, Ph.D. thesis, California Institute of Technology, 2000.

[31] H. Peyrl and P. Parrilo, Computing sum of squares decompositions with rational coefficients, *Theoret. Comput. Sci.* **409** (2008), 269–281.

[32] S. Pironio, M. Navascués and A. Acín, Convergent relaxations of polynomial optimization problems with noncommuting variables, *SIAM J. Optim.* **20** (2010), 2157–2180.

[33] J. Povh, Contribution of copositive formulations to the graph partitioning problem, *Optimization: A Journal of Mathematical Programming and Operations Research* (2011), 1–13.

[34] J. Povh and F. Rendl, A copositive programming approach to graph partitioning, *SIAM Journal on Optimization* **18** (2007), 223–241.

[35] J. Povh and F. Rendl, Copositive and semidefinite relaxations of the quadratic assignment problem, *Discrete Optimization* **6** (2009), 231–241.

[36] R. Quarez, *Some examples of trace-positive quaternary quartics*, to appear in Proc. Amer. Math. Soc., http://hal.archives-ouvertes.fr/hal-00685397

[37] B. Recht, M. Fazel and P. A. Parrilo, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization, *SIAM Rev.* **52** (2010), 471–501.

[38] C. Scheiderer, *Sums of squares of polynomials with rational coefficients*, to appear in J. Eur. Math. Soc., http://arxiv.org/abs/1209.2976

[39] H. R. Stahl, Proof of the BMV conjecture, *Acta Math.* **211** (2013), 255–290.

[40] H. Wolkowicz, R. Saigal and L. Vandenberghe, *Handbook of semidefinite programming: Theory, Algorithms, and Applications*, volume 27 of *International Series in Operations Research & Management Science*, Kluwer Academic Publishers, Boston, MA, 2000.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 261–266

On mixed discriminants of positively definite matrix*

Chang-Jian Zhao †
Department of Mathematics, China Jiliang University, Hangzhou 310018, P. R. China

Xiao-Yan Li
Department of Mathematics, Hunan Normal University, Changsha 410000, P. R. China

Received 2 November 2012, accepted 21 February 2014, published online 11 January 2015

Abstract

In the paper, some new inequalities for the mixed discriminants of positively definite matrix are established, which are the matrix analogues of inequalities of the well-known mixed volumes function.

Keywords: Mixed discriminants, symmetric matrix, mixed volume, Aleksandrov's inequality.

Math. Subj. Class.: 15A09, 52A40

1 Introduction

Let $x_1, \ldots, x_n$ be a set of nonnegative quantities and let $E_i(x)$ be the i-th elementary symmetric function of an n-tuple $x = x(x_1, \ldots, x_n)$ of non-negative reals is defined by $E_0(x) = 1$ and

$$E_i(x) = \sum_{1<j_1<\cdots<j_i\leq n} x_{j_1}x_{j_2}\cdots x_{j_i},\ 1 \leq i \leq n.$$

An interesting inequality for the symmetric function was established ([1], also see [2], p. 33) as follows.

$$\frac{E_i(x+y)}{E_{i-1}(x+y)} \geq \frac{E_i(x)}{E_{i-1}(x)} + \frac{E_i(y)}{E_{i-1}(y)}. \tag{1.1}$$

*Research is supported by National Natural Science Foundation of China (11371334).

†Author to whom correspondence should be addressed.
E-mail addresses: chjzhao@163.com chjzhao@aliyun.com chjzhao315@sohu.com (Chang-Jian Zhao), lixy-77@163.com (Xiao-Yan Li)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

A matrix analogue of (1.1) is the following result of Bergstrom [3].

Let K and L be positive definite matrix, and let K_i and L_i denote the sub-matrices obtained by deleting the i-th row and column. Then

$$\frac{\det(K+L)}{\det(K_i+L_i)} \geq \frac{\det(K)}{\det(K_i)} + \frac{\det(L)}{\det(L_i)}. \tag{1.2}$$

An interesting proof is due to Bellman [4] (also see [2], p. 67). A generalization of (1.2) was established by Ky Fan [5] (also see [6-7]). Moreover, we assume all positive definite matrix are supposed to be symmetric in the article.

There is a remarkable similarity between inequalities for symmetric functions (or determinants of symmetric matrices) and inequalities for the mixed volumes of convex bodies. In 1991, V. Milman asked if there is version of (1.1) or (1.2) in the theory of mixed volumes and it was stated as the following open question (see [8]):

Question 1.1. For which values of i is it true that for very pair of convex bodies K and L in $\mathbb{R}^n$,

$$\frac{W_i(K+L)}{W_{i+1}(K+L)} \geq \frac{W_i(K)}{W_{i+1}(K)} + \frac{W_i(L)}{W_{i+1}(L)} \ ? \tag{1.3}$$

The convex body is the compact and convex subsets with non-empty interiors in $\mathbb{R}^n$. $W_i(K)$ denotes the quermassintegral of convex body K and $W_{i+1}(K)$ denotes the mixed volumes $V(\underbrace{K,\dots,K}_{n-i-1},\underbrace{B,\dots,B}_{i+1})$. The sum $+$ is the usual Minkowski vector sum and B denotes the unit ball.

A theorem by Minkowski provides a fundamental relation between volume and operations of addition and multiplication of convex bodies by nonnegative reals: If $K_1,\dots,K_m$ are convex bodies, $m \in \mathbb{N}$, then the volume of $t_1K_1+\cdots+t_mK_m$ is a homogeneous polynomial of degree n in $t_i > 0$ (see [14]). That is

$$V(t_1K_1+\cdots+t_mK_m) = \sum_{1\leq i_1,\dots,i_n\leq m} V(K_{i_1},\dots,K_{i_n})t_{i_1}\cdots t_{i_n},$$

where the coefficients $V(K_{i_1},\dots,K_{i_n})$ are chosen to be invariant under permutations of their arguments. The coefficient $V(K_{i_1},\dots,K_{i_n})$ is called the *mixed volume* of the n-tupe $(K_{i_1},\dots,K_{i_n})$. Steiner's formula is a special case of Minkowski's theorem; the volume of $K+tB$, $t\geq 0$, can be expanded as a polynomial in t:

$$V(K+tB) = \sum_{i=0}^{n} \binom{n}{i} W_i(K)t^i,$$

where $W_i(K) := V(\underbrace{K,\dots,K}_{n-i},\underbrace{B,\dots,B}_{i})$ is the *quermassintegral* of convex body K.

A partial answer (L must be a ball) of (1.3) was established by Gianopoulos, Hartzoulaki and Paouris [9]).

If K is a convex body and D is a ball in $\mathbb{R}^n$, then for $i=0,\dots,n-1$

$$\frac{W_i(K+D)}{W_{i+1}(K+D)} \geq \frac{W_i(K)}{W_{i+1}(K)} + \frac{W_i(D)}{W_{i+1}(D)}. \tag{1.4}$$

The answer to the above question is negative; it can be proved that (1.3) is true in full generality only when $i = n-1$ or $i = n-2$ (the details see [10]). Moreover, a dual inequality of (1.4) for the dual quermassintegral of star bodies was proved by Li and Leng [11].

In the paper, we establish some inequalities for mixed discriminants of positively definite matrix which are matrix analogues of some mixed volumes inequalities.

2 Mixed discriminants and Aleksandrov's inequality

Recall that for positive definite $n \times n$ matrices $K_1, \ldots, K_N$ and $\lambda_1, \ldots, \lambda_N \geq 0$, the determinant of the linear combination $\lambda_1 K_1 + \cdots + \lambda_N K_N$ is a homogeneous polynomial of degree n in the λ_i (see e.g. [12]),

$$\det(\lambda_1 K_1 + \cdots + \lambda_N K_N) = \sum_{1 \leq i_1, \ldots, i_n \leq N} D(K_{i_1}, \ldots, K_{i_n}) \lambda_{i_1} \cdots \lambda_{i_n}, \tag{2.1}$$

where the coefficient $D(K_{i_1}, \ldots, K_{i_n})$ are chosen to be invariant under permutations of their arguments. The coefficient $D(K_{i_1}, \ldots, K_{i_n})$ is called the *mixed discriminant* of $K_{i_1}, \ldots, K_{i_n}$.

The mixed discriminant $D(K, \ldots, K, I, \ldots, I)$, with $n-k$ copies of K and k copies of the identity matrix, I, will be abbreviated by $D_k(K)$. From (2.1), we have

$$D_i(K + \lambda I) = \sum_{j=0}^{n-i} \binom{n-i}{j} \lambda^j D_{i+j}(K). \tag{2.2}$$

Note that the elementary mixed discriminants $D_0(K), \ldots, D_n(K)$ are thus defined as the coefficients of the polynomial

$$\det(K + \lambda I) = \sum_{i=0}^{n} \binom{n}{i} \lambda^i D_i(K). \tag{2.3}$$

Obviously, $D_0(K) = \det(K)$ while $nD_{n-1}(K)$ is the trace of K.

The well-known Aleksandrov's inequality for mixed discriminants can state as follows (see [13], also see [14], p.383 or [15], p.35):

Lemma 2.1. *If $K_1, K_2, \ldots, K_n$ are real symmetric positively definite $n \times n$ matrices, then*

$$D(K_1, K_2, K_3, \ldots, K_n)^2 \geq D(K_1, K_1, K_3, \ldots, K_n) D(K_2, K_2, K_3, \ldots, K_n), \tag{2.4}$$

with equality if and only if K_1=λK_2 with positive number λ.

3 Inequalities for mixed discriminants of positively definite matrix

Theorem 3.1. *Let K be symmetric positively definite matrix and I stands for the identity matrix and $t \geq 0$. If $0 \leq i \leq n-1$ and $i \in \mathbb{N}$, then the function*

$$g(t) = \frac{D_i(K + tI)}{D_{i+1}(K + tI)} \tag{3.1}$$

is an increasing and concave function on $[0, +\infty)$.

Proof. If $f_i(s) = D_i(K + sI)$, then by the linearity of the mixed discriminant we see that

$$f_i(s+\varepsilon) = \sum_{j=0}^{n-i} \binom{n-i}{j} \varepsilon^j D_{i+j}(K+sI)$$

$$= f_i(s) + \varepsilon(n-i) f_{i+1}(s) + o(\varepsilon^2).$$

Hence

$$\frac{df_i(s)}{ds} = \lim_{\varepsilon \to 0} \frac{f(s+\varepsilon) - f(s)}{\varepsilon}$$

$$= (n-i) f_{i+1}(s). \tag{3.2}$$

Similarly, we obtain

$$\frac{df_{i+1}(s)}{ds} = (n-i-1) f_{i+2}(s). \tag{3.3}$$

From (2.4), we obtain for all $0 \le i < n$

$$f_{i+1}^2(s) - f_i(s) f_{i+2}(s) \ge 0, \tag{3.4}$$

with equality if and only if $K = \mu I$.

From (3.2), (3.3) and (3.4), we have

$$\frac{df_i(s)}{ds} f_{i+1}(s) - f(s) \frac{df_{i+1}(s)}{ds}$$

$$= f_{i+1}^2(s) + (n-i-1)(f_{i+1}^2(s) - f_i(s) f_{i+2}(s))$$

$$\ge f_{i+1}^2(s).$$

Therefore

$$\frac{dg(s)}{ds} = \left(\frac{f_i(s)}{f_{i+1}(s)} \right)'$$

$$= \frac{f_i'(s) f_{i+1}(s) - f(s) f_{i+1}'(s)}{f_{i+1}^2(s)}$$

$$= (n-i) - (n-i-1) \frac{f_i(s) f_{i+2}(s)}{f_{i+1}^2(s)}. \tag{3.5}$$

Hence

$$f(t) = \frac{D_i(K+tI)}{D_{i+1}(K+tI)}$$

is an increasing and concave function on $[0, +\infty)$. □

Theorem 3.2. *Let K be symmetric positively definite matrix and I stands for the identity matrix. If $0 \le i < n$, then*

$$(n-i) D_{i+2}(K)(D_{i+1}(K)^2 - D_i(K) D_{i+2}(K))$$

$$\ge (n-i-2) D_i(K)(D_{i+2}(K)^2 - D_{i+1}(K) D_{i+3}(K)). \tag{3.6}$$

Proof. Let $f_i(t) = D_i(K + tI)$ for $t \geq 0$ and $g(t) = \dfrac{f_i(t)}{f_{i+1}(t)}$, then

$$\frac{dg(s)}{ds} = (n-i) - (n-i-1)\frac{f_i(t)f_{i+2}(t)}{f_{i+1}^2(t)}. \tag{3.7}$$

By differentiating the both sides of (3.6) again, we have

$$\frac{d^2g(s)}{ds^2} = -(n-i-1)$$

$$\times \frac{(n-i)f_{i+2}(t)f_{i+1}^2(t) + (n-i-2)f_i(t)f_{i+1}(t)f_{i+3}(t) - 2(n-i-1)f_i(t)f_{i+2}^2(t)}{f_{i+1}^3(t)}. \tag{3.8}$$

From (3.8) and in view of $g(s)$ being a concave function, we obtain

$$(n-i)f_{i+2}(t)f_{i+1}^2(t) + (n-i-2)f_i(t)f_{i+1}(t)f_{i+3}(t) - 2(n-i-1)f_i(t)f_{i+2}^2(t) \geq 0,$$

for $t \in (0, +\infty)$.

This can be equivalently written in the form

$$(n-i)f_{i+2}(t)\Big(f_{i+1}^2(t) - f_i(t)f_{i+2}(t)\Big) \geq (n-i-2)f_i(t)\Big(f_{i+2}^2(t) - f_{i+1}(t)f_{i+3}(t)\Big). \tag{3.9}$$

Hence

$$(n-i)D_{i+2}(K+tI)\Big(D_{i+1}(K+tI)^2 - D_i(K+tI)D_{i+2}(K+tI)\Big)$$

$$\geq (n-i-2)D_i(K+tI)\Big(D_{i+2}(K)^2 - D_{i+1}(K+tI)D_{i+3}(K+tI)\Big). \tag{3.10}$$

Notice that $f_i(t)$ is continuous function, letting $t \to 0^+$ in (3.10), (3.10) reduces to the inequality in Theorem 3.2. □

References

[1] A. D. Aleksandrov, Zur Theorie der gemischten Volumina von konvexen Körpen, Die gemischten Diskriminanten und die gemischten Volumina (in Russian), *Mat. Sbornik* **3** (1938), 227–252.

[2] E. F. Bechenbach, R. Bellman, *Inequalities*, first ed., Springer-Verlag, Berlin-Göttingen, Heidelberg, 1961.

[3] R. Bellman, Notes on matrix theory-IV: an inequality due to Bergstrom, *Amer. Math. Monthly* **62** (1955), 172–173.

[4] H. Bergstrom, *A triangle inequality for matrices. Den Elfte Skandinaviski Matematikerkongress, Trondheim*, first ed., 1949, Oslo: John Grundt Tanums Forlag, 1952.

[5] A. Dembo, T. M. Cover and J. A. Thomas, Information theoretic inequalities, *IEEE Trans. Infor. Theory* **37** (1991), 1501–1518.

[6] K. Fan, Some inequalities conerning positive-definite hermitian matrices, *Proc. Cambridge Phil. Soc.* **51** (1955), 414–421.

[7] K. Fan, Problem 4786, *Amer. Math. Monthly* **65** (1958), 289.

[8] M. Fradelizi, A. Giannopoulos and M. Meyer, Some inequalities about mixed volumes, *Israel J. Math.* **135** (2003), 157–179.

[9] A. Giannopoulos, M. Hartzoulaki and G. Paouris, On a local version of the Aleksandrov-Fenchel inequality for the quermassintegrtals of a convex body, *Proc. Amer. Math. Soc.* **130** (2002), 2403–2412.

[10] X. Y. Li and G. S. Leng, Some inequalities about dual mixed volumes of star bodies, *Acta Math. Sci.* **25** (2005), 505–510.

[11] E. Lutwak, D. Yang and G. Zhang, A new affine invariant for polytopes and Schneider's projection problem, *Trans. Amer. Math. Soc.* **353** (2001), 1767–1779.

[12] M. Marcus and M. Marcus, Inequalities for symmetric functions and Hermitian matrices, *Canad J. Math.* **8** (1956), 524–531.

[13] L. Mirsky, Maximum principles in matrix theory, *Proc. Glasgow Math. Assoc.* **4** (1958), 34–37.

[14] P. Pranayanuntana, *Elliptic Brunn-Minkowski Theory*, first ed., UMT Dissertations Publishing, 2003.

[15] R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory*, first ed., Cambridge Univ. Press, 1993.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 267–277

Odd edge coloring of graphs

Borut Lužar
Faculty of Information Studies, 8000 Novo mesto, Slovenia
and
Institute of Mathematics, Physics and Mechanics, 1000 Ljubljana, Slovenia

Mirko Petruševski
Department of Mathematics and Informatics, Faculty of Mechanical Engineering, Skopje, Republic of Macedonia

Riste Škrekovski
Institute of Mathematics, Physics and Mechanics, 1000 Ljubljana, Slovenia
and
Faculty of Information Studies, 8000 Novo mesto, Slovenia
and
University of Primorska, FAMNIT, 6000 Koper, Slovenia

Received 24 November 2013, accepted 1 October 2014, published online 11 January 2015

Abstract

An edge coloring of a graph G is said to be an odd edge coloring if for each vertex v of G and each color c, the vertex v uses the color c an odd number of times or does not use it at all. In [5], Pyber proved that 4 colors suffice for an odd edge coloring of any simple graph. Recently, some results on this type of colorings of (multi)graphs were successfully applied in solving a problem of facial parity edge coloring [3, 2]. In this paper we present additional results, namely we prove that 6 colors suffice for an odd edge coloring of any loopless connected (multi)graph, provide examples showing that this upper bound is sharp and characterize the family of loopless connected (multi)graphs for which the bound 6 is achieved. We also pose several open problems.

Keywords: Edge coloring, odd subgraph, Shannon triangle.

Math. Subj. Class.: 05C15

E-mail addresses: borut.luzar@gmail.com (Borut Lužar), mirko.petrushevski@gmail.com (Mirko Petruševski), skrekovski@gmail.com (Riste Škrekovski)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

1 Introduction

Throughout the article we mainly follow the terminology and notation used in [1]. A graph is denoted by $G = (V(G), E(G))$, where $V(G)$ is the vertex set and $E(G)$ is the edge set. A graph G is always regarded as being finite (i.e. having finite number of vertices $n(G)$, and finite number of edges $m(G)$) with loops and multiple edges allowed. The parameters $n(G)$ and $m(G)$ are usually called *order* and *size* of G, respectively. Whenever $n(G) = 1$ we say that G is *trivial* and whenever $m(G) = 0$ we say that G is *empty*. For $X \subseteq V(G) \cup E(G)$, the subgraph obtained by removing the vertices and edges from the set X is denoted by $G - X$. We refer to the vertices having even (resp. odd) degree as *even* (resp. *odd*) vertices. A graph is called *even* (resp. *odd*) whenever all of its vertices are even (resp. odd).

An *odd edge coloring* of G is a (not necessarily proper) edge coloring such that each color class induces an odd subgraph of G. An odd edge coloring of G using at most k colors is referred to as an *odd k-edge-coloring*, and we say that G is *odd k-edge-colorable*. If G admits an odd edge coloring, the *odd chromatic index* $\chi'_o(G)$ is defined to be the minimum integer k for which G is odd k-edge-colorable.

By definition, each loop at a vertex v colored with c contributes 2 to the count of appearances of c at v. Thus, it is obvious that a necessary and sufficient condition for the existence of an odd edge coloring of G is the absence of vertices incident only to loops. Apart from this, the presence of loops does not influence the existence nor changes the value of the index $\chi'_o(G)$. Therefore, we choose to restrict our attention to loopless connected graphs throughout the article.

As a notion, odd edge coloring was first introduced by Pyber in his survey on graph coverings [5] as an edge decomposition of a graph into (edge disjoint) odd subgraphs. Such decompositions represent a counterpart to decompositions into even subgraphs, which were mainly used while proving various flow problems (see e.g. [6]).

In the mentioned work, Pyber considered simple graphs and proved the following result.

Theorem 1.1 (Pyber, 1991)**.** *For every simple graph G, it holds that*

$$\chi'_o(G) \leq 4\,.$$

He also remarked that the upper bound is realized by a wheel on four spokes W_4 and asked whether there is an infinite family of connected graphs for which this bound is achieved. In 2006, Mátrai [4] presented such a construction, taking an even number of copies of W_4 and an additional vertex v. Choosing an arbitrary edge from the wheel, he removed the same edge from every copy and connected its two end-vertices with v (see Fig. 1). From $\chi'_o(W_4) = 4$ readily follows that the obtained graph has χ'_o equal to 4.

A generalization of Theorem 1.1 was successfully applied in solving a problem of facial parity edge colorings in [2], and its improvement in [3]. In this paper, an analogous result to Theorem 1.1 is proved for loopless graphs. Namely, in Theorem 11 we prove that 6 colors suffice for an odd edge coloring of any loopless connected graph, and characterize the family of loopless connected graphs needing 6 colors.

2 Preliminary results

It is an easy matter to characterize the graphs having $\chi'_o \leq 1$: namely, $\chi'_o(G) = 0$ if and only if G is empty, while $\chi'_o(G) = 1$ if and only if G is nonempty and the subgraph of G

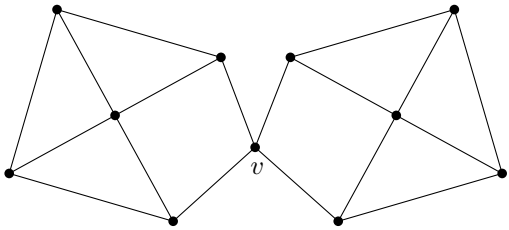


Figure 1: A graph with odd chromatic index equal to 4.

induced by its non-isolated vertices is odd. The following result was initially stated in [5], but for the sake of completeness we present it here with a proof.

Proposition 2.1. *For every forest F, it holds that $\chi'_o(F) \leq 2$.*

Proof. It is enough to prove this for an arbitrary tree T. If T is trivial or odd, then $\chi'_o(T) \leq 1$, so suppose this is not the case. We construct an odd 2-edge-coloring of T. Take an even vertex r as the root of T. To begin with, color the edges incident to r by using the color 1 for all but one such edge, and color this remaining edge by the color 2.

Continue by coloring the incident edges to each vertex v which has one incident edge already colored as follows:

- if v has even degree in T, then we complete the coloring of its incident edges by coloring them in the other color (if color 1 was used for the already colored edge, then we use color 2 for the remaining edges, and vice versa);
- if v has odd degree in T, then we complete the coloring of its incident edges by coloring them with the same color as the already colored edge.

Since there are no cycles in T, every vertex $u \neq r$ eventually is in a position to have exactly one of its incident edges colored. Namely, consider the vertices along the unique ru-path in T and suppose the opposite, i.e. suppose there exist at least one vertex on this path that never gets in the stated position. Choose the first such vertex after r (denote it by w) on the tracing of this path. Thus, the predecessor of w gets in the stated position. But this implies that w also gets in position, a contradiction.

Therefore, the above procedure produces an odd 2-edge-coloring of T. □

Let G be a graph and T be an even-sized subset of $V(G)$. Following [1], a spanning subgraph H of G is said to be a *T-join* if $d_H(v)$ is odd for all $v \in T$ and even for all $v \in V(G)\backslash T$. For example, if P is an xy-path in G, the spanning subgraph of G with edge set $E(P)$ is an $\{x, y\}$-join. Note that by removing (resp. adding) a cycle (as an edge set) from (resp. to) a T-join we again obtain a T-join. Thus, whenever a T-join exists, there also exists such a forest (resp. coforest).

A classical result about T-joins (see [1]) is that whenever G is connected, there exists a T-join for every even-sized subset T of $V(G)$. Consider a connected graph G of even order, and let T be the set of its even vertices. The handshake lemma assures that T has even size, hence there exists a T-join H. By putting $K := G - E(H)$ we obtain an odd factor of G, i.e. a spanning odd subgraph. Note that if the T-join H was chosen to be a forest, then the obtained odd factor K satisfies the following statement.

Proposition 2.2. *Given a connected graph G of even order, there exists an odd factor K of G such that $G - E(K)$ is a forest.*

3 Tight upper bound for χ'_o

In this section, through a number of propositions we derive the main result of the paper, a general tight upper bound $\chi'_o \leq 6$ with a characterization of the loopless connected graphs for which the bound is achieved.

Proposition 3.1. *Given a loopless connected graph G of even order, it holds that*

$$\chi'_o(G) \leq 3\,.$$

If furthermore G is even, then $\chi'_o(G) \leq 2$.

Proof. By Proposition 2.2, we can take an odd factor K of G such that $G - E(K)$ is a forest denoted by F. From Proposition 2.1 we infer that

$$\chi'_o(G) \leq \chi'_o(F) + \chi'_o(K) \leq 2 + 1 = 3\,.$$

If in addition G is even, then F is a spanning odd forest of G, giving

$$\chi'_o(G) \leq \chi'_o(F) + \chi'_o(K) \leq 1 + 1 = 2\,,$$

which completes the proof. □

Let G be a loopless graph. By a *bouquet of parallel edges in G* we refer to a subset of $E(G)$ consisting of all the edges linking a pair of adjacent vertices.

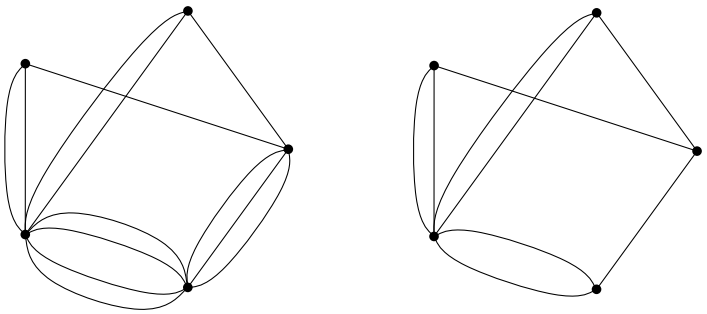


Figure 2: A loopless graph (left) and its reduction (right).

The *reduction* $\mathrm{red}(G)$ of a loopless graph G is defined to be a spanning subgraph of G obtained by the following change at every bouquet of parallel edges: remove maximum possible even number of edges without altering the adjacency relation in $V(G)$ (see Fig. 2 for an example). Obviously, up to isomorphism, each loopless graph has a unique reduction. We say that a loopless graph G is *reduced* whenever its multiplicity is at most 2, i.e. when $G \cong \mathrm{red}(G)$.

It was already remarked in [2] that an odd k-edge-coloring of its reduction readily provides an odd k-edge-coloring of any loopless graph G, since the removed edges between

any two adjacent vertices may all adopt one arbitrary color used on the remaining edges between them in $\mathrm{red}(G)$. Hence,

$$\chi'_o(G) \leq \chi'_o(\mathrm{red}(G))\,. \tag{3.1}$$

Remark 1. Regarding the inequality (3.1), suppose that a graph G satisfies $\chi'_o(G) < \chi'_o(\mathrm{red}(G))$. In respect of the reduced graph $\mathrm{red}(G)$, assume G is minimal such graph in terms of size. Consider an arbitrary optimal odd edge coloring of G. Then, by the minimality of G, on each bouquet of parallel edges no color appears more than once, unless possibly on a bouquet of size 2. In other words, in every optimal odd edge coloring, on any bouquet of parallel edges in G that reduces in size for $\mathrm{red}(G)$, no color is repeated. Hence, whenever the inequality (3.1) is strict, it holds that $\chi'_o(\mathrm{red}(G)) \geq 4$.

A loopless graph G on three pairwise adjacent vertices is referred to as a *Shannon triangle*. Denote the sizes of its bouquets of parallel edges by p, q, r in non-increasing order. We say that the considered Shannon triangle is of *type* (p, q, r). In particular, whenever p, q, r are even, we speak of a Shannon triangle of *even type*. Observe that there are exactly four different types of reduced Shannon triangles and only one of them is of even type (depicted in Fig. 3).

Next, we prove that (3.1) is always an equality for the case of Shannon triangles. Furthermore, we prove that the reduced Shannon triangles of different types attain odd edge chromatic indices for all integer values between 3 and 6.

Proposition 3.2. *Given a Shannon triangle G, let (p, q, r) be the type of $\mathrm{red}(G)$. Then*

$$\chi'_o(G) = \chi'_o(\mathrm{red}(G)) = p + q + r. \tag{3.2}$$

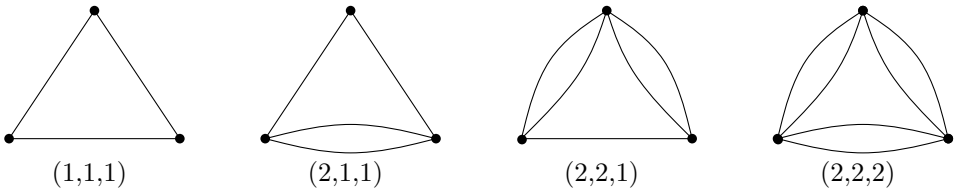


Figure 3: The four types of Shannon triangles with odd chromatic indices 3, 4, 5, and 6, respectively. The last one is of even type.

Proof. Observe that in an arbitrary odd edge coloring of a Shannon triangle, no color appears altogether an even number of times on any bouquet. Namely, denote the vertices of the triangle by u, v, and w and suppose that on the edges between u and v some color, say 1, appears in total an even number of times. This implies that 1 appears overall an odd number of times on each of the remaining two bouquets. Thus, 1 appears an even number of times at the vertex w, a contradiction.

This clearly implies that no color is repeated in an arbitrary odd edge coloring of a reduced Shannon triangle. Hence the second equality in (3.2) is established.

Next, we prove the first equality. Suppose there exists a Shannon triangle G for which $\chi'_o(G) < \chi'_o(\mathrm{red}(G))$. In respect of the reduced graph $\mathrm{red}(G)$, assume G is minimal such graph in terms of size. Then, Remark 1 and the above observation imply that no color is repeated in an optimal odd edge coloring of G. This readily gives $\chi'_o(G) > \chi'_o(\mathrm{red}(G))$, a contradiction.

□

In the following, we give several other propositions leading to the main theorem of this article, but first we introduce some additional notation. Let v be a vertex of a reduced graph G and S_v be the subgraph of G induced by the set of edges incident to v, i.e. $S_v := G[E(v)]$. Each pair of parallel edges from this subgraph is said to be a *petal of* v. Each edge of a petal is referred to as a *petal edge of* v. The other edges incident to v are called *leaf edges of* v. Denote by $p(v)$ and $l(v)$ the number of petals and number of leaf edges of v, respectively.

Proposition 3.3. *If a connected reduced graph G has a non-cutvertex v for which either $p(v)$ is odd or $l(v) \neq 0$, then $\chi'_o(G) \leq 5$.*

Proof. By Proposition 3.1, we may assume $n(G)$ is odd. Suppose v is a non-cutvertex of G such that either $p(v)$ is odd or $l(v) \neq 0$. We consider the four possible cases according to the parities of $p(v)$ and $l(v)$:

- (i) *$p(v)$ is odd and $l(v)$ is even.* By Proposition 3.1, there exists an odd 3-edge-coloring φ of $G - v$ with color set $\{1, 2, 3\}$. We extend φ to G by using two additional colors 4 and 5: color by 4 each leaf edge of v and precisely one petal edge from each petal of v; color by 5 the remaining petal edges of v. We obtain an odd 5-edge-coloring of G.
- (ii) *$p(v)$ is odd and $l(v)$ is odd.* Let e be a leaf edge of v. Let F be a forest in $G - v$ (as in the proof of Proposition 3.1), and add v together with the leaf edge e to F. We obtain a forest F' in the subgraph $G' := G - (E(v)\backslash e)$. Since two colors suffice for an odd edge coloring of F' and the edge set $E(G')\backslash E(F')$ induces an odd subgraph of G', by using a third color for the edges from $E(G')\backslash E(F')$ we obtain an odd 3-edge-coloring φ of G'. We extend φ to an odd 5-edge-coloring of G as in (i).
- (iii) *$p(v)$ is even and $l(v)$ is odd.* In this case, we use a similar approach as in (ii), but take a petal edge of v as the edge e, instead. Again, we finish as in (i).
- (iv) *$p(v)$ is even and $l(v)$ is even and positive.* Let φ be an odd 3-edge-coloring of $G - v$ with color set $\{1, 2, 3\}$. Extend φ to G by coloring with 4 one leaf edge of v and precisely one petal edge of each petal of v, and moreover, by coloring with 5 all the remaining uncolored edges incident to v. We obtain an odd 5-edge-coloring of G.

□

Whenever $l(v) = 0$, we refer to S_v as the *orchid at* v. According to the parity of $p(v)$, we distinguish between even and odd orchids.

A graph obtained from a path of length $k \geq 1$ in which every edge is replaced by two parallel edges is called an *open k-necklace*. Observe that every open k-necklace is an even graph of order $k + 1$. Similarly, a graph obtained from a cycle of length $k \geq 2$ in which every edge is replaced by two parallel edges is called a *closed k-necklace*. Every closed k-necklace is an even graph of order k.

Proposition 3.4. *The odd chromatic index of an open k-necklace G satisfies*

$$\chi'_o(G) = \begin{cases} 2 & \text{if } k \text{ is odd}, \\ 4 & \text{if } k \text{ is even}. \end{cases}$$

Proof. Assume k is odd. Since G needs at least two colors for an odd edge coloring, Proposition 3.1 implies $\chi'_o(G) = 2$. We construct a particular odd 2-edge-coloring of G that we will use in Proposition 8. Fix one of the two natural orders for the bouquets of parallel edges of G, and color as follows:

1) for the edges of the first, third, fifth,..., k-th such bouquet use each of the colors 1 and 2 once;
2) for the edges of the second, fourth, sixth,..., $(k-1)$-st such bouquet use the color 1 twice.

Assume now k is even. First we establish the inequality $\chi'_o(G) \leq 4$ by constructing an odd 4-edge-coloring of G. Again, by fixing one of the two natural orders for the bouquets of G: on the edges of the first $(k-1)$ bouquets apply the odd 2-edge-coloring of 1) and 2); for the edges of the k-th bouquet use each of the colors 3 and 4 once.

Second we prove $\chi'_o(G) = 4$. Suppose $\chi'_o(G) < 4$ and consider an optimal odd edge coloring of G. Fix one of the two natural orders for the bouquets of G. Obviously, the first bouquet is dichromatic, i.e. two colors are used for its edges. Hence, $\chi'_o(G) < 4$ implies that the second bouquet is monochromatic. But then the third bouquet is dichromatic, etc. We deduce that the k-th bouquet is monochromatic. This is a contradiction, hence $\chi'_o(G) = 4$.

□

Proposition 3.5. *The odd chromatic index of a closed k-necklace G satisfies*

$$\chi'_o(G) = \begin{cases} 2 & \text{if } k \text{ is even}, \\ 4 & \text{if } k \text{ is odd}, k \geq 5, \\ 6 & \text{if } k = 3. \end{cases}$$

Proof. Assume k is even. Since G needs at least two colors for an odd edge coloring, Proposition 3.1 implies $\chi'_o(G) = 2$.

Assume next k is odd and $k \geq 5$. First we prove $\chi'_o(G) \leq 4$ by constructing an odd 4-edge-coloring of G. Remove a vertex v from G to obtain the open $(k-2)$-necklace $G - v$. Use the odd 2-edge-coloring of $G - v$ constructed in Proposition 3.4 with a single change of color: for one edge from the last bouquet instead of using the color 2 use the color 3. Now color the edges of the orchid at v:

- for the uncolored petal of v "neighboring" a bouquet from $G - v$ colored by 1 and 2, use the color 2 for both edges;
- for the remaining uncolored petal of v use the colors 2 and 4 once each.

Now, suppose $\chi'_o(G) < 4$ and consider an optimal odd edge coloring of G. Obviously, at each vertex the orchid is dichromatic with three of its edges having the same color. Thus, any two consecutive bouquets in G are such that one is monochromatic and the other is

dichromatic. Hence, precisely half of the k bouquets in G are monochromatic. This is a contradiction with the parity of k.

Finally, note that for $k = 3$, the closed k-necklace G is the Shannon triangle of type $(2, 2, 2)$ and thus $\chi'_o(G) = 6$ (see Proposition 3.2).

□

We are now in a position to prove our main result for loopless 2-connected graphs.

Proposition 3.6. *Let G be a loopless 2-connected graph which is not isomorphic to a Shannon triangle of even type. Then $\chi'_o(G) \leq 5$.*

Proof. By Proposition 3.1, inequality (3.1) and Proposition 3.3, we may assume that G has odd order n and is a 2-connected reduced graph with an even orchid at every vertex. Denote by v a vertex of maximum degree in G. If the orchid S_v has precisely two petals, then G is a closed n-necklace with $n \geq 5$ (namely, if $n = 3$ then G would be the Shannon triangle of type $(2, 2, 2)$). Hence, in this case Proposition 3.5 implies $\chi'_o(G) = 4$.

So, we may assume S_v has at least four petals. Consider the graph $G - v$ and denote by $u_1, u_2, \ldots, u_{2s}$ the neighbors of v in G. By Proposition 3.1, we have $\chi'_o(G - v) = 2$. Consider an initial odd 2-edge-coloring of $G - v$ with color set $\{1, 2\}$ such that the edges of the spanning odd forest F constructed in the proof of Proposition 3.1 are colored by 1.

Denote by $\mathcal{M}$ the collection of maximal paths in F that have non-empty intersection with the set $\{u_1, u_2, \ldots, u_{2s}\}$. Note that every member of $\mathcal{M}$ is non-trivial and has both end-vertices among the leaves of F. We distinguish between the following three cases:

(i) *At least one $P \in \mathcal{M}$ does not have both end-vertices in the set $\{u_1, u_2, \ldots, u_{2s}\}$.* Start a tracing of one such path P from an end-vertex not belonging to $\{u_1, u_2, \ldots, u_{2s}\}$. Let u be the first vertex from $V(P) \cap \{u_1, u_2, \ldots, u_{2s}\}$ met on this tracing, and denote by P_0 the traced subpath of P. Beginning at the edge incident to u, recolor the edges of P_0 by the colors 3 and 4, alternatingly. Color the petal edges of v in the following way: use the color 4 for one vu-edge, while using the color 1 for all the remaining petal edges incident to v. Thus, we obtain an odd 4-edge-coloring of G.

(ii) *Each member of $\mathcal{M}$ has both end-vertices in the set $\{u_1, u_2, \ldots, u_{2s}\}$ and there is such a path $P \in \mathcal{M}$ of length at least 2.* Denote by u_i and u_j the end-vertices of P. We recolor the edges of P by using the colors $3, 4$ and 5 as follows: the edge incident to u_i is recolored by 3 and the edge incident to u_j is recolored by 4. The other edges along P are recolored in such a way to obtain a proper edge coloring of P, which is clearly achievable. Color the petal edges of v in the following way:

- use color 4 for both vu_j-edges;
- use color 4 for one vu_i-edge;
- use color 1 for the remaining petal edges of v.

Thus, we obtain an odd 5-edge-coloring of G.

(iii) *Each member of $\mathcal{M}$ has both end-vertices in the set $\{u_1, u_2, \ldots, u_{2s}\}$ and there is no such path of length at least 2.* In this case, the edges of F incident to at least one of the vertices $u_1, u_2, \ldots, u_{2s}$ form a matching on the set $\{u_1, u_2, \ldots, u_{2s}\}$. Without loss of generality, suppose this matches u_{2i-1} with u_{2i}, for every $i \in \{1, 2, \ldots, s\}$. From the initial odd 2-edge-coloring of $G - v$, we obtain an odd 5-edge-coloring of G as follows:

- recolor the edge u_1u_2 in F with the color 5;
- use both the colors 1 and 4 once for the petal vu_1-edges;
- use both the colors 3 and 4 once for each of the petals $vu_2, vu_3, \ldots, vu_{2s-1}$;
- use both the colors 3 and 5 once for the petal vu_{2s}-edges (here we make use of the fact that $s \geq 2$).

This establishes the inequality $\chi'_o(G) \leq 5$, which completes the proof of the statement.

□

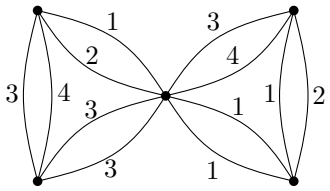


Figure 4: An odd 4-edge-coloring.

Next we prove that 5 colors suffice for an odd edge coloring of any loopless connected graph which is not a block.

Proposition 3.7. *If G is a loopless graph of connectivity $\kappa = 1$, then $\chi'_o(G) \leq 5$.*

Proof. We may restrict to reduced graphs of connectivity $\kappa = 1$. Suppose the statement is not valid, and let G be a minimal counterexample in terms of the number of blocks. By Propositions 3.1 and 3.3, G has odd order and is a connected reduced graph with an even orchid at every non-cutvertex. Observe that $n(G) \geq 4$ (for otherwise $\kappa = 1$ implies $m(G) \leq 4$).

Choose an end-block B of G, with s being the relevant cut-vertex, such that the graph $G' := G - V(B - s)$ satisfies the inequality $\chi'_o(G') \leq 5$. Namely, if the graph G has more than two blocks then we merely choose B to be an arbitrary end-block of G: since G' is of connectivity 1 and has one block less than G, the choice of G assures that $\chi'_o(G') \leq 5$. And, if G has only two blocks, Proposition 3.6 and Fig. 4 assure that at least one of these two blocks will do: G cannot be the odd 4-edge-colorable graph depicted in Fig. 4.

Observe that B is an even graph having odd order. Namely, every $v \in V(B)\backslash\{s\}$ has an even orchid S_v. Hence, the same is true for the vertex s in B, proving that B is an even graph. Regarding the order, an even $n(B)$ would imply that $n(G')$ is also even and hence the inequalities $\chi'_o(B) \leq 2$ and $\chi'_o(G') \leq 3$ would yield the bound $\chi'_o(G) \leq 5$, which is a contradiction.

So, $B - s$ is an even graph of even order. Take an odd factor K of $B - s$ and color the edges from $B - s$ by using the color 1 for $E(K)$ and the color 2 for $E(B - s - E(K))$. Extend this to an edge coloring of B by using the color 1 for each edge incident to s in B. Use an odd 5-edge-coloring of G' with color set $\{1, 2, \ldots, 5\}$ such that the color 1 is used for at least one edge incident to s in G'. These two edge colorings of B and G' together constitute an odd 5-edge-coloring of G. This is in contradiction with the choice of G, which completes the proof of the statement.

□

By now, we have assembled all the parts for a proof of the main theorem in this article. It characterizes the Shannon triangles of even type as the only loopless connected graphs needing the maximum 6 colors for an odd edge coloring.

Theorem 3.8. *For every loopless connected graph G, it holds that*

$$\chi'_o(G) \leq 6\,.$$

Equality is achieved only for the Shannon triangles of even type.

Proof. Straightforward from Propositions 3.6 and 3.7. □

4 Concluding remarks and further work

Based on several additional observations, we propose three conjectures.

Conjecture 4.1. *For every loopless graph G, it holds that*

$$\chi'_o(G) = \chi'_o(\mathrm{red}(G))\,. \tag{4.1}$$

From Remark 1, we infer that Conjecture 4.1 is true whenever $\chi'_o(\mathrm{red}(G)) \leq 3$ and Theorem 3.8 assures its validity whenever $\chi'_o(\mathrm{red}(G)) = 6$. Furthermore, Propositions 3.4 and 3.5 imply that whenever $\mathrm{red}(G)$ is a necklace (open or closed), (4.1) stands. Namely, it is readily deduced from Remark 1 that if $\chi'_o(\mathrm{red}(G)) \leq 4$ and every bouquet of $\mathrm{red}(G)$ has size 2, then the equality (4.1) is fulfilled.

In regard to Proposition 3.6, we propose the following

Conjecture 4.2. *For every loopless 2-connected graph G whose reduction is neither the Shannon triangle of type $(2,2,1)$ nor of type $(2,2,2)$, it holds that*

$$\chi'_o(G) \leq 4\,.$$

Furthermore, we believe that an even stronger statement is true. Namely, we propose that the bound in Theorem 3.8 could be further reduced, by excluding the graphs whose reductions are isomorphic to the Shannon triangle of type $(2,2,1)$.

Conjecture 4.3. *For every loopless connected graph G whose reduction is neither the Shannon triangle of type $(2,2,1)$ nor of type $(2,2,2)$, it holds that*

$$\chi'_o(G) \leq 4\,.$$

Assuming the validity of Conjecture 3, note that any possible counterexample G to Conjecture 4.1 must satisfy $\chi'_o(G) = 3$ and $\chi'_o(\mathrm{red}(G)) = 4$.

Another direction for further work is to provide an answer to the following open problem.

Problem 4.4. Characterize the loopless graphs which are odd 2-edge-colorable.

Acknowledgements. We thank the anonymous referees for helpful comments. This work is partially supported by ARRS Program P1-0383.

References

[1] J. A. Bondy and U. S. R. Murty, *Graph Theory*, Graduate Texts in mathematics, Springer, New York **244**, 2008.

[2] J. Czap, S. Jendrol', F. Kardoš and R. Soták, Facial parity edge colouring of plane pseudographs, *Disc. Math.* **312** (2012), 2735–2740.

[3] B. Lužar and R. Škrekovski, Improved bound on facial parity edge coloring, *Discrete Math.* **313** (2013), 2218–2222.

[4] T. Mátrai, Covering the edges of a graph by three odd subgraphs, *J. Graph Theory* **53** (2006), 75–82.

[5] L. Pyber, Covering the edges of a graph by..., *Graphs and Numbers, Colloq. Math. Soc. János Bolyai* **60** (1991), 583–610.

[6] P. D. Seymour, Nowhere-zero 6-flows, *J. Comb. Theory Ser. B* **30** (1981), 130–135.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 279–286

Edge looseness of plane graphs

Július Czap

Department of Applied Mathematics and Business Informatics, Faculty of Economics
Technical University of Košice, Němcovej 32, 040 01 Košice, Slovakia

Received 7 February 2014, accepted 15 April 2015, published online 8 July 2015

Abstract

A face of an edge colored plane graph is called *e-loose* if the number of colors used on its edges is at least three. The *e-looseness* of a plane graph G is the minimum positive integer k such that any edge coloring of G with k colors involves an e-loose face. In this paper we determine tight lower and upper bounds for the e-looseness of connected plane graphs. These bounds are expressed by linear polynomials of the number of faces.

Keywords: Plane graph, edge coloring.

Math. Subj. Class.: 05C10, 05C15

1 Introduction

We use the standard terminology according to Bondy and Murty [1]. All considered graphs are undirected, finite, loopless, multiple edges are allowed.

Let $G = (V, E, F)$ denote a connected plane graph with the vertex set V, the edge set E and the face set F. We say that two edges of G are *face-independent* if they are not incident with the same face. Two edges of G are *face-adjacent* if they are consecutive edges of a facial trail of some face. The *medial graph* $M(G)$ of G is the simple graph obtained as follows. For each edge e of G insert a vertex $m(e)$ in $M(G)$. Join two vertices of $M(G)$ if the corresponding edges are face-adjacent in G. The embedding of G induces the embedding of $M(G)$.

Edge colorings of graphs embedded on surfaces with face-constrains have recently drawn a substantial amount of attention, see [5, 6, 7, 12] and references therein. There are two questions derived from hypergraph colorings that one may ask in this setting:

Question 1.1. What is the minimum number of colors needed to color the edges of a connected plane graph in such a way that each of its faces is incident with edges of at least two different colors?

E-mail address: julius.czap@tuke.sk (Július Czap)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

Question 1.2. What is the maximum number of colors that can be used in an edge coloring of a connected plane graph in such a way that each of its faces is incident with edges of at most two different colors?

The following result gives the answer to Question 1.1.

Theorem 1.3. *Every connected plane graph (on at least two edges) has an edge coloring with at most 3 colors such that each of its faces is incident with edges of at least two different colors. Moreover, if G is simple, then 2 colors are sufficient.*

Proof. First we show that every connected plane graph without faces of size two has an edge coloring with 2 colors such that these two colors appear on every face. Let G be such a plane graph and let $M(G)$ be its medial graph. Add edges to $M(G)$ to obtain a plane triangulation T. By the Four Color Theorem, T has a proper vertex coloring which uses 3 or 4 colors. Combine the first color class with the second, and combine the third with the fourth. This yields a 2-coloring of the graph T. It is easy to check that each face of T is incident with vertices of two different colors. This vertex coloring induces a required edge coloring of G.

Now assume that the claim does not hold for a connected plane graph with faces of size two. Let G be a counterexample on minimum number of edges. Let e be an edge which is incident with faces f and g, where f has size two. Let $G - e$ be the graph obtained from G by removing the edge e. The graph $G - e$ is not a counterexample because it has fewer edges than G. Consequently, it has a required coloring. Let h be the face in $G - e$ corresponding to the faces f and g in G. Extend the coloring of $G - e$ to a coloring of G in the following way. If two colors appear on h, then color e with the third color. Otherwise (three colors appear on h) we color e with a color which does not appear on the second edge of f. This means that the minimum counterexample does not exist.

There are graphs which require three colors for such a coloring, for example the connected graph on two vertices and three edges. □

In this paper, we focus on Question 1.2. A face of an edge colored plane graph is called *monochromatic* or *bichromatic* if the number of colors used on its edges is one or two, respectively. A face which is neither monochromatic nor bichromatic is called *edge loose* (or shortly e-loose). The *edge looseness* (or shortly e-looseness) of a plane graph G is the minimum positive integer k such that any surjective k-edge-coloring involves an e-loose face. This parameter of G will be denoted by $\mathrm{els}(G)$. The e-looseness is well defined for all plane graphs having at least one face incident with at least three different edges. (Note that every connected plane graph on at least four vertices has such a face.) Throughout the paper, we will consider only such graphs.

2 Upper bounds

2.1 1-connected plane graphs

Theorem 2.1. *Let $G = (V, E, F)$ be a connected plane graph. Then $\mathrm{els}(G) \leq |F| + 2$. Moreover, this bound is sharp.*

Proof. Let φ be an edge coloring of G such that each of its faces is either monochromatic or bichromatic. First we show that φ uses at most $|F| + 1$ colors. Clearly, we can assume that φ uses at least two colors. This means that at least one face is bichromatic, say f_1. Let

f_2 be a face of G adjacent to f_1. There is at most one color which appears on f_2 and does not appear on f_1, since the faces f_1 and f_2 share an edge and at most two colors occur on f_2. The graph G is connected, therefore we can number the faces $f_1, f_2, \ldots, f_{|F|}$ of G such that for every $i \geq 2$ the face f_i is adjacent to a face f_j for some $j < i$. Using the argument described above, there is at most one color which appears on f_j and does not appear on $f_1, \ldots, f_{j-1}$. Consequently, φ uses at most $2 + (|F| - 1)$ colors. This means that every edge coloring with at least $|F| + 2$ colors involves an e-loose face.

Now we prove that this upper bound is tight. Let H be a connected plane graph on at least three vertices. We insert a new vertex into each face of H and join each new vertex by an edge inside the corresponding face to exactly one of its vertices. In this way we obtain a 1-connected plane graph $\widetilde{H}$. Clearly, we added $|F(H)|$ edges to H and these new edges are face-independent in $\widetilde{H}$. If we color these face-independent edges with distinct colors and all other edges with the same color, then we obtain an edge coloring of $\widetilde{H}$ such that every face is bichromatic. Hence, $\mathrm{els}(\widetilde{H}) \geq |F(H)| + 2 = |F(\widetilde{H})| + 2$. □

2.2 2-connected plane graphs

When G is 2-connected, the bound from Theorem 2.1 can be improved by one.

Let us recall that the (geometric) *dual* $G^* = (V^*, E^*, F^*)$ of a plane graph $G = (V, E, F)$ can be defined as follows: Corresponding to each face f of G there is a vertex f^* of G^*, and corresponding to each edge e of G there is an edge e^* of G^*; two vertices f^* and g^* are joined by the edge e^* in G^* if and only if their corresponding faces f and g are separated by the edge e in G (an edge separates the faces incident with it).

Theorem 2.2. *Let $G = (V, E, F)$ be a 2-connected plane graph. Then* $\mathrm{els}(G) \leq |F| + 1$. *Moreover, this bound is sharp.*

Proof. Let φ be an edge coloring of G such that every face is either monochromatic or bichromatic. This coloring induces a coloring of the dual G^* in a natural way. Observe that at most 2 colors appear at any vertex of G^*. Let us choose one edge from each color class, and let the chosen edges induce the subgraph H of G^*. Each component of H is either a path or a cycle, since the graph G^* does not contain any loop. Therefore, $2|E(H)| = \sum_{v \in V(H)} \deg_H(v) \leq 2|V(H)| \leq 2|V(G^*)| = 2|F(G)|$. Since the number of colors used by φ equals $|E(H)|$, we deduce that φ uses at most $|F(G)|$ colors.

To see that the bound is tight consider the plane embedding of the complete bipartite graph $K_{2,n}$, see Figure 1. This plane graph has n faces and an n-edge-coloring such that every face is bichromatic. □

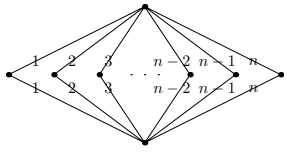


Figure 1: A plane drawing of $K_{2,n}$.

2.3 3-connected plane graphs

Lemma 2.3. *Let $G = (V, E, F)$ be a 3-connected plane triangulation. Let t denote the maximum number of disjoint cycles in its dual. Then* $\mathrm{els}(G) = \frac{1}{2}|F| + 1 + t$.

Proof. Any M_2-edge coloring (being an edge coloring of a graph such that at most 2 colors appear at any vertex, see [2, 3]) of the dual graph G^* corresponds to an edge coloring of G with property that every face of G is incident with edges of at most two different colors and vice versa. The graph G^* is cubic, since G is a plane triangulation. In [3], it was proved that the maximum number of colors that can be used in an M_2-edge coloring of a cubic plane graph H is equal to $\frac{|V(H)|}{2} + t$, where t is the maximum number of disjoint cycles in H. Consequently, $\mathrm{els}(G) = \frac{|V(G^*)|}{2} + t + 1 = \frac{|F(G)|}{2} + t + 1$. □

Theorem 2.4. *Let $G = (V, E, F)$ be a 3-connected plane triangulation. Let g denote the girth of its dual. Then*

- $\mathrm{els}(G) \leq \frac{5}{6}|F| + 1$ *if* $g \geq 3$,
- $\mathrm{els}(G) \leq \frac{3}{4}|F| + 1$ *if* $g \geq 4$,
- $\mathrm{els}(G) \leq \frac{7}{10}|F| + 1$ *if* $g \geq 5$.

Moreover, these bounds are sharp.

Proof. It follows from Lemma 2.3, since the number of disjoint cycles in the dual G^* is not greater than $\frac{|V(G^*)|}{g} = \frac{|F(G)|}{g}$.

By Lemma 2.3 we can easily prove that the bounds are attained on the duals of graphs shown in Figure 2. □

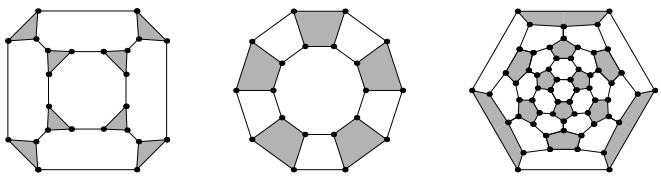


Figure 2: The duals of these graphs show that the bounds are tight.

Conjecture 2.5. *Let $G = (V, E, F)$ be a 3-connected plane graph. Then* $\mathrm{els}(G) \leq \frac{5}{6}|F| + 1$.

3 Lower bounds

A *matching* of G is a set of pairwise disjoint edges, and a maximum matching is one of maximum cardinality.

Let $c(H)$ denote the number of components of a graph H.

Lemma 3.1. *Let G be a connected plane graph and let G^* be its dual. Let M^* be a matching in G^*. Then* $\mathrm{els}(G) \geq |M^*| + c(G^* - M^*) + 1$. *Moreover, this bound is sharp.*

Proof. We color the edges of the matching M^* with distinct colors and use an additional new color for the edges of each component of $G^* - M^*$. Thus, we obtain an edge coloring of the dual graph G^* such that at most 2 colors appear at any vertex of G^*. This coloring induces a coloring of G in which every face is either monochromatic or bichromatic.

To see that the bound is sharp, let G be a plane graph whose dual is a $2r$-sided prism. The $2r$-sided prism H_{2r}, $r \geq 2$, consists of the vertex set $V = \{u_1, u_2, \ldots, u_{2r}, v_1, v_2, \ldots, v_{2r}\}$ and the edge set $E = \{u_i u_{i+1}, v_i v_{i+1}, u_i v_i | i = 1, \ldots, 2r\}$, where $2r + 1 := 1$. The set of faces consists of two $2r$-gonal faces $f_1 = [u_1, \ldots, u_{2r}]$ and $f_2 = [v_1, \ldots, v_{2r}]$ and $2r$ quadrangles $[u_i u_{i+1} v_{i+1} v_i]$ for $i = 1, 2, \ldots, 2r$, see Figure 3 for illustration.

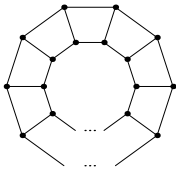


Figure 3: A $2r$-sided prism.

Let $M = \{u_i u_{i+1}, v_i v_{i+1} | i = 1, 3, 5, \ldots, 2r - 1\}$ be a matching of H_{2r}. Observe that $|M| = 2r$ and $c(H_{2r} - M) = r$. Therefore $\mathrm{els}(G) \geq 3r + 1$. On the other hand, the graph H_{2r} is a simple 3-connected cubic plane graph, hence its dual (the graph G) is a simple triangulation. By Theorem 2.4 we have $\mathrm{els}(G) \leq \frac{3}{4} \cdot 4r + 1$. □

Corollary 3.2. *Let G be a connected plane graph and let G^* be its dual. Let M^* be a maximum matching in G^*. Then* $\mathrm{els}(G) \geq |M^*| + 2$.

Proof. It immediately follows from Lemma 3.1, since $c(G^* - M^*) \geq 1$. □

Since there are 2-connected (and 1-connected) graphs G with arbitrarily many faces which have $\mathrm{els}(G) \leq 4$ (Take a plane drawing of the cycle $C = v_1 v_2 \ldots v_{3n}$. Add n vertices $u_1, u_2, \ldots, u_n$ to the inner part of C and join u_i with v_{3i-2} and v_{3i} as it is depicted in Figure 4. It is easy to see that the e-looseness of the obtained graph is four.), there is no nontrivial lower bound on $\mathrm{els}(G)$ expressed by a linear polynomial of $|F|$ if G is not 3-connected. Hence, in the remaining part of the paper we will investigate e-looseness of 3-connected plane graphs G. Since the dual of G is a simple plane graph, we may apply structural properties of planar graphs on the dual graph; in particular, we will use the following one.

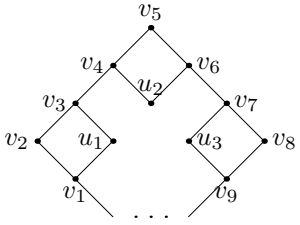


Figure 4: A graph G with "many" faces and $\mathrm{els}(G) = 4$.

Theorem 3.3. [11] *Let G be an n-vertex 3-connected planar graph with $n \geq 78$. Let δ denote the minimum degree of G and let M be a maximum matching in G. Then,*

- $|M| \geq \frac{n+4}{3}$ *if* $\delta \geq 3$,
- $|M| \geq \frac{3n+8}{7}$ *if* $\delta \geq 4$,
- $|M| \geq \frac{9n+20}{19}$ *if* $\delta \geq 5$.

Note that, if a planar graph G is 4-connected, then it has a matching of size $\left\lfloor \frac{|V(G)|}{2} \right\rfloor$. This immediately follows from Tutte's result [14]: every 4-connected planar graph contains a Hamiltonian cycle.

Lemma 3.4. *Let $G = (V, E, F)$ be a 3-connected plane graph. Then it has at least $\frac{|V|}{2} + 2$ faces. Moreover, this bound is tight.*

Proof. The minimum degree of G is at least 3, since it is 3-connected. Using the handshaking lemma we have $2|E| = \sum_{v \in V} \deg(v) \geq 3|V|$. Consequently, $|E| \geq \frac{3}{2}|V|$. Using this fact and the Euler's polyhedral formula $|V| - |E| + |F| = 2$ we obtain $|F| \geq 2 + \frac{|V|}{2}$.

This bound is attained for every 3-connected cubic plane graph. □

Using Corollary 3.2, Theorem 3.3 and Lemma 3.4 we can prove the following.

Theorem 3.5. *Let $G = (V, E, F)$ be a 3-connected plane graph on at least 152 vertices. Let g be the girth of G. Then,*

- $\mathrm{els}(G) \geq \frac{|F|+4}{3} + 2$ *if* $g \geq 3$,
- $\mathrm{els}(G) \geq \frac{3|F|+8}{7} + 2$ *if* $g \geq 4$,
- $\mathrm{els}(G) \geq \frac{9|F|+20}{19} + 2$ *if* $g \geq 5$.

Moreover, these bounds are sharp.

Proof. The 3-connectedness of G implies that its dual G^* is also 3-connected, hence we can use Theorem 3.3 for G^*. By Lemma 3.4, $|V(G^*)| = |F(G)| \geq \frac{|V(G)|}{2} + 2 \geq 78$. It is easy to check that $\frac{n+4}{3} \leq \frac{3n+8}{7} \leq \frac{9n+20}{19}$ for $n \geq 2$. As $g(G) \leq \delta(G^*)$, the result follows from Corollary 3.2 and Theorem 3.3. The sharpness of these bounds follows from Theorem 3.6. □

Theorem 3.6. *For any integer $n \geq 4$ there exists a 3-connected cubic plane graph $G = (V, E, F)$ with girth g such that*

- $g = 3$, $|F| = 3n - 4$ *and* $\mathrm{els}(G) = n + 2$,
- $g = 4$, $|F| = 7n - 12$ *and* $\mathrm{els}(G) = 3n - 2$,
- $g = 5$, $|F| = 19n - 36$ *and* $\mathrm{els}(G) = 9n - 14$.

Proof. Let T be a simple plane triangulation on $n \geq 4$ vertices. Let G_i be the graph obtained from T by inserting the configuration H_i, shown in Figure 5, into each of its faces, for $i = 3, 4, 5$.

Any plane triangulation on n vertices has $2n - 4$ faces, therefore
$|F(G_3^*)| = |V(G_3)| = n + 2n - 4 = 3n - 4$,

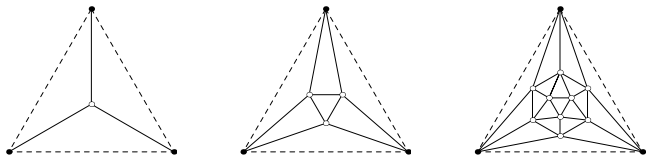


Figure 5: A construction of triangulations with edge-connectivity 3, 4 and 5.

$|F(G_4^*)| = |V(G_4)| = n + 3(2n-4) = 7n - 12$ and
$|F(G_5^*)| = |V(G_5)| = n + 9(2n-4) = 19n - 36$.

In [2] it was proved that $K_2(G_3) = n+1$, $K_2(G_4) = 3n-3$ and $K_2(G_5) = 9n-15$, where $K_2(G_i)$ denotes the maximum number of colors used in an M_2-edge coloring of G_i, $i = 3, 4, 5$. Consequently, for the duals of these graphs it holds $\mathrm{els}(G_3^*) = n+2$, $\mathrm{els}(G_4^*) = 3n-2$ and $\mathrm{els}(G_5^*) = 9n-14$.

Observe, that each minimum edge-cut of size g in G_i corresponds to a cycle in G_i^* and vice versa, therefore, the edge connectivity of the graph G_i is equal to the girth of the dual graph G_i^*. The edge connectivity of G_i is equal to i, for $i = 3, 4, 5$, since the tetrahedron, octahedron and icosahedron have edge connectivities $3, 4$ and 5, respectively. □

Finally, note that the vertex version of Question 1.2 was investigated in [4], where it was proved that the vertex looseness of a connected plane graph G equals the maximum number of vertex disjoint cycles in the dual graph G^* increased by 2. Vertex looseness of triangulations on closed surfaces was studied in [8, 9, 10, 13].

References

[1] J.A. Bondy and U.S.R. Murty, *Graph theory*, Springer, New York, 2008.

[2] K. Budajová and J. Czap, M_2-edge coloring and maximum matching of graphs, *Int. J. Pure Appl. Math.* **88** (2013), 161–167.

[3] J. Czap, M_i-edge colorings of graphs, *Appl. Math. Sciences* **5** (2011), 2437–2442.

[4] J. Czap, S. Jendroľ, F. Kardoš and J. Miškuf, Looseness of plane graphs, *Graphs Combin.* **27** (2011), 73–85.

[5] S. Jendroľ, J. Miškuf, R. Soták and E. Škrabuľáková, Rainbow faces in edge-colored plane graphs, *J. Graph Theory* **62** (2009), 84–99.

[6] B. Lužar, M. Mockovčiaková, R. Soták, R. Škrekovski and P. Šugerek, ℓ-facial edge colorings of graphs, *Discrete Appl. Math.* **181** (2015), 193–200.

[7] B. Lužar and R. Škrekovski, Improved bound on facial parity edge coloring, *Discrete Math.* **313** (2013), 2218–2222.

[8] K. Matsuoka, Looseness width of 5-connected triangulations on the torus, *Electron. Notes Discrete Math.* **31** (2008), 105–108.

[9] S. Negami, Looseness ranges of triangulations on closed surfaces, *Discrete Math.* **303** (2005), 167–174.

[10] S. Negami and T. Midorikawa Loosely-tightness of triangulations of closed surfaces, Sci. Rep. Yokohama Nat. Univ., Sect. I, *Math. Phys. Chem.* **43** (1996), 25–41.

[11] T. Nishizeki and I. Baybars, Lower bounds on the cardinality of the maximum matchings of planar graphs, *Discrete Math.* **28** (1979), 255–267.

[12] J. Przybylo, On the facial Thue choice index via entropy compression, *J. Graph Theory* **77** (2014), 180–189.

[13] T. Tanuma, One-loosely tight triangulations of closed surfaces, *Yokohama Math. J.* **47** (1999), 203–211.

[14] W.T. Tutte, A theorem on planar graphs, *Trans. Am. Math. Soc.* **82** (1956), 99–116.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 287–300

Levels in bargraphs

Aubrey Blecher , Charlotte Brennan *, Arnold Knopfmacher †

University of the Witwatersrand, The John Knopfmacher Centre for Applicable Analysis and Number Theory, Johannesburg, South Africa

Received 20 January 2014, accepted 8 February 2015, published online 8 July 2015

Abstract

Bargraphs are lattice paths in $\mathbb{N}_0^2$, which start at the origin and terminate immediately upon return to the x-axis. The allowed steps are the up step $(0,1)$, the down step $(0,-1)$ and the horizontal step $(1,0)$. The first step is an up step and the horizontal steps must all lie above the x-axis. An up step cannot follow a down step and vice versa. In this paper we consider levels, which are maximal sequences of two or more adjacent horizontal steps. We find the generating functions that count the total number of levels, the leftmost x-coordinate and the height of the first level and obtain the generating function for the mean of these parameters. Finally, we obtain the asymptotics of these means as the length of the path tends to infinity.

Keywords: Bargraphs, levels, generating functions, asymptotics.

Math. Subj. Class.: 05A15, 05A16

1 Introduction

Bargraphs are lattice paths in $\mathbb{N}_0^2$, starting at the origin and ending upon first return to the x-axis. The allowed steps are the up step, $u = (0,1)$, the down step, $d = (0,-1)$ and the horizontal step, $h = (1,0)$. The first step has to be an up step and the horizontal steps must all lie above the x-axis. An up step cannot follow a down step and vice versa. It is clear that the number of down steps must equal the number of up steps. Related lattice paths such as Dyck paths and Motzkin paths have been studied extensively (see [4, 9]) whereas until now bargraphs which are fundamental combinatorial structures, have not attracted the same amount of interest.

*Corresponding author. This material is based upon work supported by the National Research Foundation under grant number 86329

†This material is based upon work supported by the National Research Foundation under grant number 81021

E-mail addresses: Aubrey.Blecher@wits.ac.za (Aubrey Blecher), Charlotte.Brennan@wits.ac.za (Charlotte Brennan), Arnold.Knopfmacher@wits.ac.za (Arnold Knopfmacher)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

Bousquet-Mélou and Rechnitzer in [2] and Geraschenko in [8] have studied bargraphs which were named skylines in the latter, and wall polyominoes as per the study of Feretić, in [6]. Bargraphs models arise frequently in statistical physics, see for example [3, 5, 10, 12, 15, 17]. In addition, bargraphs are commonly used in probability theory to represent frequency diagrams and are also related to compositions of integers [11].

In this paper, we consider levels, which are maximal sequences of two or more adjacent horizontal steps. We find different generating functions in each of the following sections where x counts the horizontal steps, y counts the up vertical steps and w counts one of the following parameters: the total number of levels and the horizontal position or the height of the first level. To facilitate these computations, we also find the generating function for paths with no levels.

The study of levels in bargraphs is related to the modelling of tethered polymers under pulling forces, see [13, 14]. These pulling forces have vertical and horizontal components and tend to be resisted by what is known as the stiffness of the polymers. The polymers undergo phase changes, called the stretched (adsorption) phase, where the polymer is stretched vertically. The free (desorbed) phase occurs only when the vertical force is zero. In the bargraph models of polymers positive or negative energy is added to points in levels on the bargraph (called stiffness sites), they tend to keep the polymer horizontal or cause it to bend.

As an example of a bargraph we have

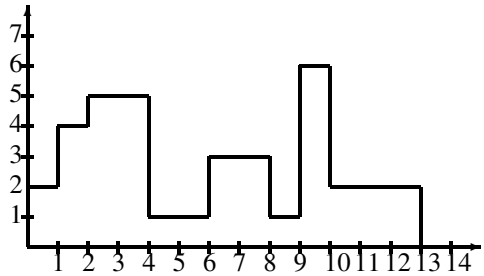


Figure 1: A bargraph with 12 up steps, 13 horizontal steps and 4 levels

Often in the lattice walk and polygon literature, "bargraphs" refer to polygon structures (which would be obtained from the objects considered here by joining the first and last vertices with horizontal steps). The objects discussed here are sometimes called "partially directed walks above a wall" depending on the context (in polymer modelling work for example).

The main tool for elucidating the statistics of interest in this study is a decomposition of bargraphs which is based on the first return to level one. This was described initially by Prellberg and Brak in [16] and more recently in [2], where it is called the *wasp-waist decomposition*. The present authors have also discussed it in [1].

It follows from the wasp-waist decomposition that the generating function $B(x, y)$ which counts all bargraphs is

$$B := B(x, y) = \frac{1 - x - y - xy - \sqrt{(1 - x - y - xy)^2 - 4x^2y}}{2x}. \tag{1.1}$$

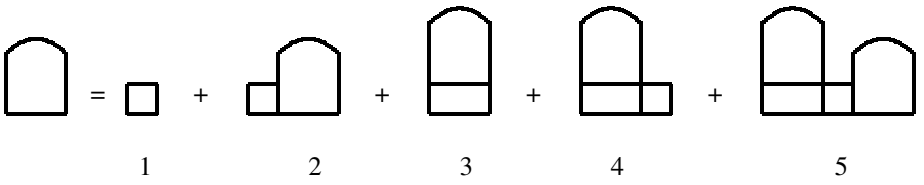


Figure 2: Wasp-waist decomposition of bargraphs

Here x counts the number of horizontal steps and y counts the number of up steps (see Theorem 1 in [1]) or [2, 7]).

The series expansion, $B(x, y)$ begins

$$x(y + y^2 + y^3 + y^4) + x^2(y + 3y^2 + 5y^3 + 7y^4) + x^3(y + 6y^2 + 16y^3 + 31y^4) + x^4(y + \mathbf{10}y^2 + 40y^3 + 105y^4 + 219y^4).$$

The bold coefficient of x^4y^2 is illustrated below with the full set of 10 bargraphs with 4 horizontal steps and 2 vertical up steps.

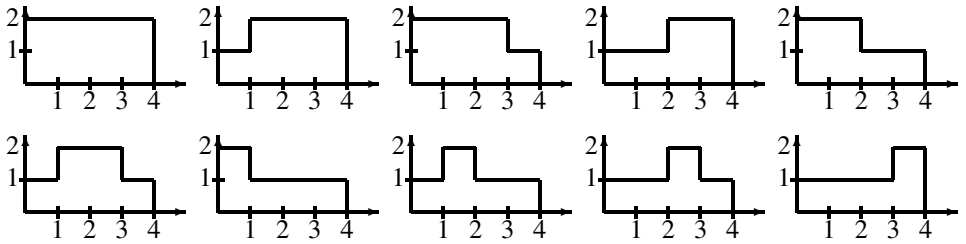


Figure 3: The 10 graphs with 4 horizontal steps and 2 vertical up steps

In [1, 2] the authors found an asymptotic expression for $B(z, z)$, where z marks the semi-perimeter of the bargraphs. This is known as the generating function for the isotropic case. The dominant singularity ρ is the positive root of

$$D := 1 - 4z + 2z^2 + z^4 = 0, \tag{1.2}$$

given by

$$\rho = \frac{1}{3}\left(-1 - \frac{4 \times 2^{2/3}}{(13 + 3\sqrt{33})^{1/3}} + \left(2(13 + 3\sqrt{33})\right)^{1/3}\right) = 0.295598\cdots. \tag{1.3}$$

We have $B(z, z) \sim -\frac{\sqrt{1-\rho-\rho^3}}{\sqrt{\rho}}\,(1 - \frac{z}{\rho})^{1/2}$ as $z \to \rho$. Hence

$$[z^n]B(z, z) \sim \frac{\sqrt{1 - \rho - \rho^3}}{2\sqrt{\pi\,\rho\,n^3}}\,\rho^{-n}. \tag{1.4}$$

The following definitions will be used:

A *level* in a bargraph is a maximal sequence of two or more adjacent horizontal steps denoted by h^r where $r \geq 2$. It is preceded and followed by either an up step or a down step. The *length of the level* is the number r of horizontal steps in the sequence. The *height* of a level is the y-coordinate of the horizontal steps in the sequence.

Thus, the graph in Figure 1 has four levels, three of length 2 and one of length 3.

In all the generating functions of the following sections, the horizontal steps are counted by x, the vertical up steps are counted by y and the parameter that is under investigation by w. In each section, we use $G(x, y, w)$ or $F(x, y, w)$ for the generating function where the definition of G or F applies only to the section under consideration.

2 Total number of levels

2.1 Generating function for the number of levels

A level is a sequence of two or more adjacent horizontal steps as defined in the previous section. Let $F(x, y, w)$ be the generating function where w marks the total number of levels. Using the wasp-waist decomposition in Figure 2, we have

$$F := F(x, y, w) = \underbrace{xy}_{1} + \underbrace{F_2}_{2} + \underbrace{yF}_{3} + \underbrace{xyF}_{4} + \underbrace{FF_2}_{5} \tag{2.1}$$

The numbers below the terms refer to the cases in the wasp-waist decomposition. This will be done throughout the paper. The generating function $F_2 := F_2(x, y, w)$ is the analogous function restricted to case 2. We use the following symbolic decomposition for F_2

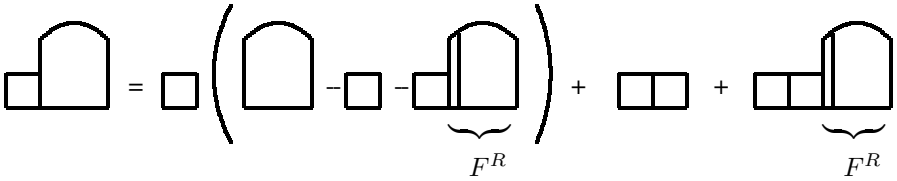


Figure 4: Decomposition for F_2

where F^R is the generating function for bargraphs in which the first column is of height 2 or more. The function F^R is easily obtained by considering all bargraphs except those starting with a column of height one. Thus

$$F^R = F - xy - F_2. \tag{2.2}$$

From Figure 4, we get

$$F_2 = x(F - xy - xF^R) + wx^2y + wx^2F^R. \tag{2.3}$$

So, combining equations (2.1), (2.2) and (2.3), we find

$$F = \frac{1}{2(x - x^2 + wx^2)} \Big(1 - x - y - xy + 2x^2y - 2wx^2y - \sqrt{4(-x + x^2 - wx^2)(xy - x^2y + wx^2y) + (1 - x - y - xy + 2x^2y - 2wx^2y)^2} \Big). \tag{2.4}$$

In order to find the generating function for the total number of levels in bargraphs, we differentiate F with respect to w and then put $w = 1$ to obtain

$$F_{Levels} := \left.\frac{\partial F}{\partial w}\right|_{w=1} = \frac{(1-x)(1-y)\left(1 - x - y - xy - \sqrt{(1 - x - y - xy)^2 - 4x^2y}\right)}{2\sqrt{(1 - x - y - xy)^2 - 4x^2y}},$$

where z marks the semiperimeter.

The series expansion begins

$$x^2(y + y^2 + y^3 + y^4) + x^3(y + 5y^2 + 9y^3 + 13y^4) + x^4(y + \mathbf{12}y^2 + 38y^3 + 79y^4).$$

There are in total 12 levels in our example in Figure 3. This is shown in bold in the series expansion.

2.2 Asymptotics in the isotropic case

We consider bargraphs with respect to the semiperimeter by substituting z for x and y in F to obtain

$$F_{Levels}(z, z) = \frac{(1-z)^2(1 - 2z - z^2 - \sqrt{1 - 4z + 2z^2 + z^4})}{2\sqrt{1 - 4z + 2z^2 + z^4}}.$$

In order to compute the asymptotics for the coefficients, we use singularity analysis as described in [7]. Let ρ be as in (1.2) and (1.3). We find that as $z \to \rho$

$$F_{Levels} \sim \frac{1 - 4\rho + 4\rho^2 - \rho^4}{4\sqrt{\rho(1 - \rho - \rho^3)}\sqrt{1 - \frac{z}{\rho}}}.$$

By singularity analysis we have

$$[z^n]F_{Levels} \sim \frac{1 - 4\rho + 4\rho^2 - \rho^4}{4\sqrt{\pi\, n}\sqrt{\rho\,(1 - \rho - \rho^3)}}\, \rho^{-n}.$$

Then after dividing by the asymptotic expression for the total number of bargraphs found in (1.4), we get the following result:

Theorem 2.1. *The average number of levels in bargraphs of semiperimeter n is asymptotic to*

$$\frac{1 - 4\rho + 4\rho^2 - \rho^4}{2(1 - \rho - \rho^3)}\, n = C\, n,$$

as $n \to \infty$ where $C = 0.117516\cdots$.

3 Bargraphs with no levels

3.1 Generating function for the number of graphs with no level

Because we require it later, we begin by enumerating a special class of bargraphs, namely one in which an adjacent sequence of horizontal steps does not occur (i.e. the only sequences of horizontals are single). This is denoted by $F_0 := F(x, y, 0)$ where F is the generating function (2.4) from the previous section.

We use the wasp-waist decomposition in Figure 2 to obtain

$$F_0 = \underbrace{xy}_{1} + \underbrace{F_{0,2}}_{2} + \underbrace{yF_0}_{3} + \underbrace{yF_0x}_{4} + \underbrace{F_0F_{0,2}}_{5}. \tag{3.1}$$

Case 2 is explained below in Figure 5.

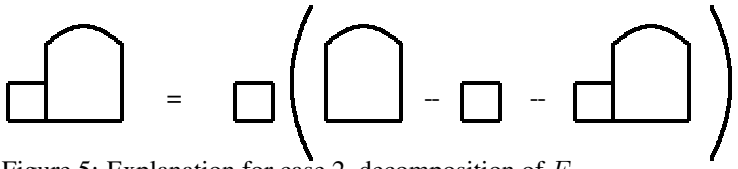


Figure 5: Explanation for case 2, decomposition of $F_{0,2}$

Thus

$$F_{0,2} = x\big(F_0 - xy - F_{0,2}\big),$$

which leads to

$$F_{0,2} = \frac{x\big(F_0 - xy\big)}{1+x}. \tag{3.2}$$

The exclusions in case 2 are because we are not allowing adjacent horizontal steps.

Hence, from (3.1) and (3.2), we have:

$$F_0 = \underbrace{xy}_{1} + \underbrace{\frac{x(F_0 - xy)}{1+x}}_{2} + \underbrace{yF_0}_{3} + \underbrace{yF_0x}_{4} + \underbrace{\frac{F_0x(F_0 - xy)}{1+x}}_{5}.$$

Solving this for F_0, we obtain

$$F_0 = \frac{1 - y - 2xy - \sqrt{1-y}\sqrt{1 - y - 4xy - 4x^2y}}{2x}. \tag{3.3}$$

The series expansion for F_0 begins

$$\begin{aligned} &x(y + y^2 + y^3 + y^4) + x^2(2y^2 + 4y^3 + 6y^4) \\ &+ x^3(y^2 + 7y^3 + 18y^4) + x^4(6y^3 + 32y^4 + 92y^5). \end{aligned}$$

Our example in Figure 3, shows that indeed there are no bargraphs having 4 horizontal and 2 up steps and no levels, which is confirmed by the lack of x^4y^2 term.

3.2 Asymptotics in the isotropic case

As before we substitute z for x and y in F_0 and obtain

$$F_0(z,z) = \frac{1-z-2z^2-\sqrt{1-z}\sqrt{1-z-4z^2-4z^3}}{2z}.$$

Let τ be the dominant root of $1-z-4z^2-4z^3=0$, its value is

$$\tau = \frac{1}{12}\left(-4+(224-24\sqrt{87})^{1/3}+2(28+3\sqrt{87})^{1/3}\right) = 0.34781\cdots.$$

Using singularity analysis we have as $z \to \tau$

$$F_0(z,z) \sim -\frac{\sqrt{1-\tau}\sqrt{\tau(1+8\tau+12\tau^2)}\sqrt{1-\frac{z}{\tau}}}{2\tau}.$$

Extracting coefficients will yield the asymptotic number of bargraphs with no levels.

$$[z^n]F_0(z,z) \sim \frac{\sqrt{1-\tau}\sqrt{\tau(1+8\tau+12\tau^2)}}{4\sqrt{\pi\, n^3}}\,\tau^{-n},$$

as $n \to \infty$.

For $n=100$, there are 3.20775×10^{42} bargraphs whereas the asymptotics give 3.24376×10^{42}.

4 Horizontal position of the first level

4.1 Generating function for the mean

Now we derive a generating function G_x for bargraphs in which the leftmost x-coordinate of the first level is counted by w. In the case where the bargraph has no level, we define the horizontal position to be 0. In Figure 6, the start of the first level is the point with coordinates $(2,5)$ and therefore the x-coordinate of the start of the first level here is 2.

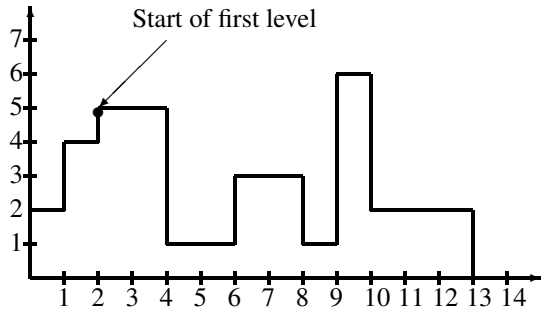


Figure 6: Horizontal position of the start of the first level

By the wasp-waist decomposition we have

$$G_x = \underbrace{xy}_{1} + \underbrace{F_2}_{2} + \underbrace{yG_x}_{3} + \underbrace{yG_xx}_{4} + \underbrace{F_5}_{5} \tag{4.1}$$

To calculate the generating function for case 2, we use Figure 7 below. The part labelled L in Figure 7 indicates a bargraph with at least one level.

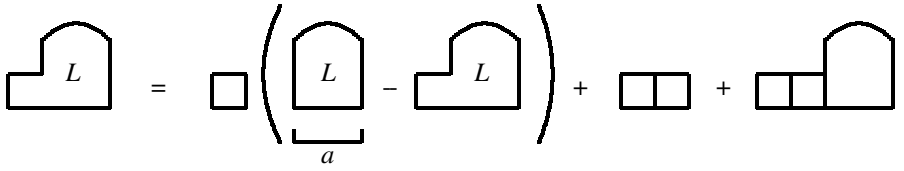


Figure 7: Decomposition for $F_{L,2}$

Note that $F_{L,2}$ is the generating function for case 2, (paths which have at least 1 level).

The generating function for the graph labelled "a" in Figure 7 is therefore $G_x - F_0$, since F_0 is the generating function for graphs with no levels from Section 3.

Thus, using Figure 7, we have:

$$F_{L,2} := F_{L,2}(x,y,w) = wx\big(G_x - F_0 - F_{L,2}\big) + x^2y + x^2B$$

where B is the generating function for all bargraphs from equation (1.1). Hence,

$$F_{L,2} = \frac{wxG_x - wxF_0 + x^2y + x^2B}{1+wx}, \tag{4.2}$$

and from (3.2)

$$F_{0.2} = \frac{x}{1+x}(F_0 - xy) = \frac{x\left(-xy + \frac{1-y-2xy+\sqrt{1-y}\sqrt{1-y-4xy-4x^2y}}{2x}\right)}{1+x}.$$

So, for case 2

$$F_2 = F_{L,2} + F_{0,2}.$$

Thus finally, the decomposition for case 5 requires Figure 8 below:

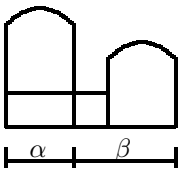


Figure 8: Case 5

For case 5, we have the concatenation of two bargraphs labelled α and β. There are three cases depending on whether the graphs α and β have levels or not.

i. Graph α has levels with generating function $y(G_x - F_0)$, in which case the generating function for β is $\frac{xB}{y}$.

ii. Neither graph has levels, thus the generating function is $F_0F_{0,2}$ where $F_{0,2}$ is as in (3.2) or

iii. Graph α has no levels but graph β has, so the generating function is $F_0(xw, y)F_2$ where $F_0(w) := F_0(xw, y)$ indicates that x has been replaced by xw in $F_0(x, y)$.

Thus

$$G_x = \underbrace{xy}_{1} + \underbrace{F_2}_{2} + \underbrace{yG_x}_{3} + \underbrace{yG_xx}_{4} + \underbrace{\big((G_x - F_0)xB + F_0F_{0,2} + F_0(xw, y)F_2\big)}_{5}$$

where in all but one case, the parameters have been omitted.

We solve for G_x, leading to

$$G_x(x, y, w) = \frac{-\frac{Bx^2F_0}{wx+1} + BxF_0 - \frac{Bx^2}{wx+1} - F_0F_{0,2} - \frac{x^2yF_0(w)}{wx+1} + \frac{wx(F_0)^2}{wx+1} + \frac{F_0wx}{wx+1} - F_{0,2} - \frac{x^2y}{wx+1} - xy}{Bx + \frac{wxF_0(w)}{wx+1} + \frac{wx}{wx+1} + xy + y - 1} \tag{4.3}$$

where

$$F_0(w) = \frac{1 - y - 2wxy - \sqrt{1-y}\sqrt{1 - y - 4wxy - 4w^2x^2y}}{2wx}.$$

Remark: We note that from (3.3) $F_0(w)|_{w=1} = F_0$.

Now, in order to find the mean horizontal position, we calculate:

$$\begin{aligned}
&\left.\frac{\partial G_x}{\partial w}\right|_{w=1}\\
&= \frac{x}{\left(Bx(x+1)+F_0x+(x+1)^2y-1\right)^2}\\
&\times\{F_0((x+1)F_0'+F_0+1)(F_{0,2}+xy+y-1)+F_{0,2}((x+1)F_0'+F_0+1)\\
&+B^2x^2(x(-F_0'+F_0+1)-F_0')+Bx(2x^2y(-F_0'+F_0+1)-(y-1)F_0'\\
&+x(F_0^2-3yF_0'+F_0'+F_0y+F_0+y))+xy(x^2y(-F_0'+F_0+1)-yF_0'+2F_0'\\
&+x(F_0^2-2yF_0'+2F_0'+F_0y+F_0+y)+F_0+1)\}
\end{aligned}$$

where

$$\begin{aligned}
F_0' &= \left.\frac{\partial F_0(w)}{\partial w}\right|_{w=1}\\
&= \frac{y\left(-2x\sqrt{y-1}+\sqrt{(2x+1)^2y-1}-\sqrt{y-1}\right)-\sqrt{(2x+1)^2y-1}+\sqrt{y-1}}{2x\sqrt{(2x+1)^2y-1}}.
\end{aligned} \tag{4.4}$$

The series expansion of $\left.\frac{\partial G_x}{\partial w}\right|_{w=1}$ begins

$$x^3\left(2y^2+4y^3+6y^4\right)+x^4\left(\mathbf{5}y^2+25y^3+60y^4\right).$$

In our example in Figure 3, the sum of the horizontal positions of the first levels is 5.

4.2 Asymptotics in the isotropic case

Using singularity analysis and computer algebra we find that

$$\left.\frac{\partial G_x}{\partial w}\right|_{w=1} \sim -2\,c_1(\rho)\sqrt{\rho\,(1-\rho-\rho^3)}\left(1-\frac{z}{\rho}\right)^{1/2}$$

where ρ is as in (1.3) and

$$\begin{aligned}
c_1(\rho) = &\frac{1-\rho}{\left((-1+\rho)\rho^2+\sqrt{-1+\rho}\sqrt{Y(\rho)}\right)^3\sqrt{Y(\rho)}}\\
&\times\left(\sqrt{-1+\rho}(1-\rho-12\rho^2-4\rho^3+13\rho^4+27\rho^5+18\rho^6+18\rho^7+4\rho^8)\right.\\
&\left.+(-1+\rho+4\rho^2+8\rho^3-5\rho^4+\rho^5-6\rho^6-2\rho^7)\sqrt{Y(\rho)}\right)
\end{aligned}$$

as $z\to\rho$ and $Y(\rho)=-1+\rho+4\rho^2+4\rho^3$.

The coefficient is

$$[z^n]\left.\frac{\partial G_x}{\partial w}\right|_{w=1} \sim \frac{c_1(\rho)\,\sqrt{\rho\,(1-\rho-\rho^3)}}{\sqrt{\pi\,n^3}}\,\rho^{-n}.$$

After dividing by the asymptotic number of bargraphs we get

Theorem 4.1. *The average horizontal position of the first level in bargraphs is asymptotic to the constant*

$$2\,\rho\,c_1(\rho) = 2.38298, \quad \textit{as } n \to \infty.$$

For $n = 200$, the exact average is $2.35787\cdots$.

5 Height of the first level

5.1 Generating function for the mean

Let $G_y(x, y, w)$ be the generating function for the y-coordinate of the first level for bargraphs where w marks this coordinate. If there are no levels then there is no w, so we have a contribution to w^0. As in the previous section, the first level in Figure 9 begins at the point $(2, 5)$, with y-coordinate 5.

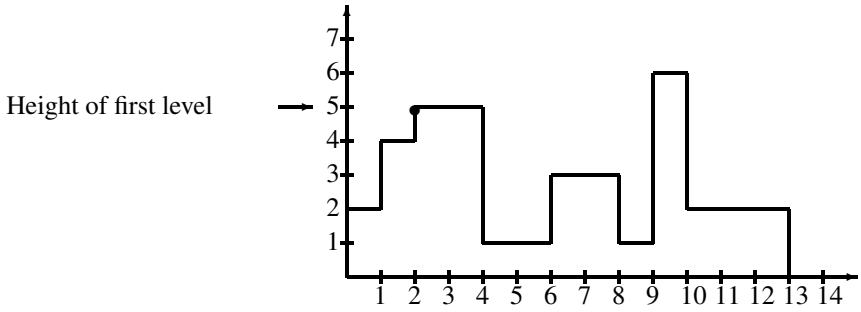


Figure 9: Height of the first level

Using the wasp-waist decomposition, this yields:

$$G_y = \underbrace{xy}_{1} + \underbrace{F_2}_{2} + \underbrace{F_3}_{3} + \underbrace{xF_3}_{4} + \underbrace{F_5}_{5}. \tag{5.1}$$

Considering case 2 separately, we have for F_2:

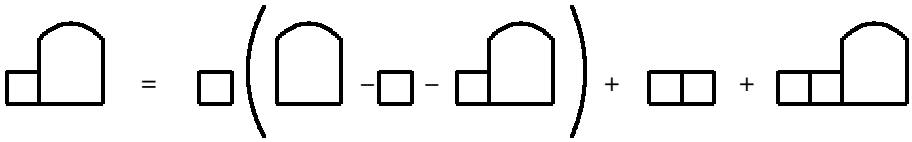


Figure 10: Decomposition for F_2

Thus

$$F_2 = x(G_y - xy - F_2) + x^2yw + x^2wB.$$

So

$$F_2 = \frac{x(G_y - xy + xyw + xwB)}{1+x} \tag{5.2}$$

and

$$F_3 = yw\big(G_y - F_0\big) + yF_0 \tag{5.3}$$

where the first and second terms distinguish between the cases where there are levels (which are therefore multiplied by w) and no levels.

Also separately, for the last case F_5 we can use Figure 8. If α has levels, then the generating functions for α and β are $w\big(G_y - F_0\big)$ and $xB(x,y)$ respectively. On the other hand, if α has no levels, the generating functions are yF_0 and F_2.

Thus

$$F_5 = w(G_y - F_0)xB(x,y) + yF_0F_2. \tag{5.4}$$

Substituting (5.2), (5.3), and (5.4) in (5.1) and solving for G_y, we obtain

$$G_y = \frac{T}{Bwx + \frac{F_0x}{x+1} + w(x+1)y + \frac{x}{x+1} - 1} \tag{5.5}$$

where

$$\begin{aligned} T = &- \frac{BF_0wx^2}{x+1} + BF_0wx - \frac{Bwx^2}{x+1} - \frac{F_0wx^2y}{x+1} + F_0w(x+1)y \\ &+ \frac{F_0x^2y}{x+1} - F_0(x+1)y - \frac{wx^2y}{x+1} + \frac{x^2y}{x+1} - xy. \end{aligned}$$

The generating function for the sum of the heights of the first levels is obtained from the derivative of G_y with respect to w and then setting $w = 1$.

Using the following substitutions

$$\begin{cases} X(x,y) = -1 + (1+2x)^2y, \\ Y(x,y) = (-1+y)(-1+x^2(-1+y) + y + 2x(1+y)), \end{cases} \tag{5.6}$$

we have

$$\begin{aligned} &\left.\frac{\partial G_y}{\partial w}\right|_{w=1} \\ &= \frac{\left(-1+x+y-xy+\sqrt{Y(x,y)}\right)}{\left(x^2(1-y) + x\sqrt{Y(x,y)} - \sqrt{X(x,y)}\sqrt{y-1} + \sqrt{Y(x,y)}\right)^2} \\ &\times \left(4x^2(y-1)y + x\left(-2\sqrt{X(x,y)}\sqrt{y-1}y + \sqrt{X(x,y)}\sqrt{y-1} + 4y^2 - 3y - 1\right)\right. \\ &\qquad \left. + y\left(-\sqrt{X(x,y)}\sqrt{y-1} + y - 1\right)\right). \end{aligned} \tag{5.7}$$

The series expansion of $\left.\frac{\partial G_y}{\partial w}\right|_{w=1}$ begins

$$\begin{aligned} x^2\left(y + 2y^2 + 3y^3 + 4y^4\right) + x^3\left(y + 8y^2 + 21y^3 + 40y^4\right) + \\ x^4\left(y + \mathbf{15}y^2 + 71y^3 + 198y^4\right). \end{aligned}$$

Figure 3 illustrates that the sum of the heights of the first levels is 15 as shown in bold above.

5.2 Asymptotics in the isotropic case

Substituting z for both x and y in the above equation (5.7) and using $X(z,z) := X(z) = -1+z+4z^2+4z^3$ and $Y(z,z) := Y(z) = 1-4z+2z^2+z^4$, we obtain

$$\begin{aligned}
&\left.\frac{\partial G_y}{\partial w}\right|_{w=1} \\
&= \frac{\left(-1+2z-z^2+\sqrt{Y(z)}\right)}{\left((1-z)z^2-\sqrt{z-1}\sqrt{X(z)}+\sqrt{Y(z)}+z\sqrt{Y(z)}\right)^2} \\
&\quad \times \left[4z^3(z-1)+z\left(-1+z-\sqrt{z-1}\sqrt{X(z)}\right)\right. \\
&\qquad \left.+z\left(-1-3z+4z^2+\sqrt{z-1}\sqrt{X(z)}-2z\sqrt{z-1}\sqrt{X(z)}\right)\right] \\
&\sim -2\,c_2(\rho)\sqrt{\rho(1-\rho-\rho^3)}\left(1-\frac{z}{\rho}\right)^{1/2},
\end{aligned}$$

by using computer algebra as $z \to \rho$, where

$$c_2(\rho) = 2\rho\frac{\left(-2+2\rho+\rho^2-\rho^3+\sqrt{-1+\rho}\sqrt{X(\rho)}\right)\left(1+\rho-2\rho^3+\rho\sqrt{-1+\rho}\sqrt{X(\rho)}\right)}{\left(\rho^2(-1+\rho)+\sqrt{-1+\rho}\sqrt{X(\rho)}\right)^3}.$$

Hence

$$[z^n]\left.\frac{\partial G_y}{\partial w}\right|_{w=1} \sim \frac{c_2(\rho)\sqrt{\rho(1-\rho-\rho^3)}}{\sqrt{\pi}\,n^3}\rho^{-n} \quad \text{as } n\to\infty$$

where

Thus after dividing by the asymptotic number of bargraphs we obtain

Theorem 5.1. *The average height of the first level in bargraphs is asymptotic to the constant*

$$2\,\rho\,c_2(\rho) \approx 6.15883\cdots, \quad \textit{as } n\to\infty.$$

For $n = 300$, the exact average is $6.00066\cdots$.

References

[1] A. Blecher, C. Brennan and A. Knopfmacher, Combinatorial parameters in bargraphs (2015), submitted.

[2] M. Bousquet-Mélou and A. Rechnitzer, The site perimeter of bargraphs, *Adv. Appl. Math.* **31** (2003), 86–112.

[3] M. Bousquet-Mélou and R. Brak, Exactly solved models of polyominoes and polygons, Chapter 3 of Polygons, Polyominoes and Polycubes, *Lecture notes in physics, volume 775*, Springer, Berlin, Heidelberg, 2009, 43–78.

[4] E. Deutsch, Dyck path enumeration, *Discrete Math.* **204** 1-3 (1999), 167–202.

[5] P. Duchon, q-grammars and wall polyominoes, *Ann. Comb.* **3** (1999), 311–321.

[6] S. Feretić, A perimeter enumeration of column-convex polyominoes, *Discrete Math. Theor. Comput. Sci.* **9** (2007), 57–84.

[7] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

[8] A. Geraschenko, An investigation of skyline polynomials. `http://people.brandeis.edu/~gessel/47a/geraschenko.pdf`

[9] K. Humphreys, A history and a survey of lattice path enumeration, *J. Statist. Plann. Inference* **140:8** (2010), 2237–2254.

[10] E. J. Janse van Rensburg and P. Rechnitzer, Exchange symmetries in Motzkin path and bargraph models of copolymer adsorption. *Electron. J. Comb.* **9** (2002), R20.

[11] D. Merlini, F. Uncini and M. C. Verri, A unified approach to the study of general and palindromic compositions, *Integers* **4** (2004), #A23.

[12] J. Osborn and T. Prellberg, Forcing adsorption of a tethered polymer by pulling, *J. Stat. Mech-Theory E.* (2010), P09018.

[13] A. Owczarek, Exact solution for semi-flexible partially directed walks at an adsorbing wall, *J. Stat. Mech.: Theor. and Exp.* (2009), P11002.

[14] A. Owczarek, Effect of stiffness on the pulling of an adsorbing polymer from a wall: an exact solution of a partially directed walk model, *J. Phys. A: Math. Theor.* **43** (2010), 225002 (16pp).

[15] A. Owczarek and T. Prellberg, Exact Solution of the Discrete (1+1)-dimensional SOS Model with Field and Surface Interactions, *J. Stat. Phys.* **70:5/6** (1993), 1175–1194.

[16] T. Prellberg and R. Brak, Critical exponents from nonlinear functional equations for partially directed cluster models. *J. Stat. Phys.* **78** (1995), 701–730.

[17] C. Richard, I. Jensen and A. J. Guttmann, Scaling Function for Self-Avoiding Polygons, *Proceedings TH2002 Supplement*, Birkhäuser Verlag, Basel, 2003, 267–277.

creative commons

ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 301–320

What can Wikipedia and Google tell us about stock prices under different market regimes?

Boris Cergol *
XLAB, Ljubljana, Slovenia

Matjaž Omladič †
Department for automation, biocybernetics and robotics, Jozef Stefan Institute, Ljubljana

Received 21 October 2013, accepted 19 December 2013, published online 10 July 2015

Abstract

In less than five years a surprisingly high level of attention has built up in the possible connection between internet search data and stock prices. It is the main aim of this paper to point out how this connection may depend heavily on different regimes of the market, i.e. the bear market vs. the bull market. We consider three types of internet search data (relative Google search frequencies of company tickers, relative Google search frequencies of company names and page visits of Wikipedia articles about individual companies) and a substantial sample of companies which are members of the S&P 500 index. We discover two inverse patterns in stock prices: in the bear market what we propose to term a "merry frown" and in bull market a "sour smile", both clearly seen especially for the Wikipedia data. We propose market neutral strategies that exploit these new patterns and yield up to 17% in average annual return during our sample period from 2008 to 2013.

Keywords: Stock returns, internet search data, market regimes, trading strategies.

Math. Subj. Class.: 91G70, 97K80

1 Introduction

A byproduct of the increasingly widespread use of the internet is the data on internet activity of individual users. While most of this data is retained by the website owners and unavailable to the public either due to privacy or business reasons there are some exceptions. One such example is the Google Trends service which enables users to view the

*Operation part financed by the European Union, European Social Fund.
†Supported by a grant from the Slovenian Research Agency - ARRS.
E-mail addresses: boris.cergol@xlab.si (Boris Cergol), matjaz@omladic.net (Matjaž Omladič)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

relative frequencies of search queries entered into Google's search engine. Since becoming publicly available in 2006 Google Trends have attracted attention of researchers in various fields. In [12] the authors show that analysis of health-related search queries can lead to accurate estimates of influenza epidemics with a reporting lag of only one day which is almost two weeks sooner than traditional surveillance systems. Choi and Varian [6] apply a similar approach to estimating a number of economic indicators such as automobile sales or unemployment claims.

The relevance of internet search data for financial data analysis was first explored by Da, Engelberg and Gao [6] who considered the relative search frequencies of company tickers and names as proxies for investor attention in the US stock market. They show that search frequencies outperform existing measures of investor attention and that an increase in a company ticker's search frequency predicts a higher stock price in the following two weeks. In [2, 14] the authors obtain similar results in terms of future returns and additionally observe that an increase in a company name's search frequency is associated with a rise in trading activity and stock liquidity. The prevailing explanation for positive correlation between future stock returns and company-related relative search frequencies is based on the theory of Barber and Odean [3]. They suggest that attention-grabbing stocks experience short-term buying pressure from individual investors. This might simply be due to the fact that a single investor faces a difficult decision when deciding which of the thousands of available stocks to buy, while the decision of which stock to sell is much easier since it is usually limited to the few stocks that are part of his existing portfolio.

Google trends data has also been used in assessing investor sentiment. In [8] the authors construct the index which is a sum of relative search frequencies of economy-related terms associated with negative sentiment. This new index is able to predict values of existing investor sentiment indicators and has a perceptible impact on short-term future stock prices. In [18] a number of stock market index strategies are tested that profit from fluctuations of relative search frequencies of individual economy-related terms. Strategies of the same type are further explored in a related work [17] where Google trends data is replaced by the numbers of page visits to economy-related Wikipedia pages.

The main contribution of our paper is the addition of market regimes into the study of the connection between stock returns and internet search. If the reason for positive correlation between future returns and search frequencies is in fact in the cognitive bias of individual investors then we would expect that the effect would be even stronger in periods when investors face greater uncertainty and are even more prone to irrational decisions. We present a two-state hidden Markov model for the returns of the S&P 500 index. The model parameters are estimated by the Baum-Welch algorithm after which the most likely sequence of hidden states is found by the Viterbi algorithm. The first of the two states is characterized by low returns and high volatility and corresponds to what is commonly refered to as the "bear market" regime by investors. Conversely, the second state is characterized by high returns and low volatility and we label it the "bull market" regime.

We choose a sample of stocks that are members of the S&P 500 index and study the relation between their future short term returns and three different types of internet search data: the page visits of company-related articles on Wikipedia, the relative frequency of Google searches for company tickers and the relative frequency of Google searches for company names. To the best of our knowledge ours is the first study of this kind that takes Wikipedia data for individual companies into consideration. We also perform our analysis on daily data which is in contrast to most of the existing literature where financial

applications of internet search is studied using weekly data.

We perform a number of cross-sectional Fama-MacBeth regressions where future stock returns are the explained variable and a single internet search variable is the regressor. This regressions are performed on a subsample of observations that belong to either the bear regime or the bull regime as well as on the entire sample. Our main result is that the market regime indeed has a strong influence on the relation between future stock returns and internet search data. In all three cases of internet search variables the future returns are higher in the bear regime compared to the bull regime given the same increase of the chosen internet search variable. This effect remains evident even after controlling for the factors of the Carhart four-factor model [5].

After controlling for the Carhart factors the Wikipedia page visits variable emerges as the one with the greatest influence on future stock returns. In fact, both of the Google search variables prove to be statistically insignificant. To our surprise, we also find very little evidence supporting the theory that an increase of investor attention to a given stock translates into a short term rise of the stock's price due to increased buying pressure. Instead we observe two different price patterns for which we propose the terms "merry frown" and "sour smile". A merry frown is a pattern of positive correlation between future stock returns and Wikipedia page visits that is observed only during the bear market. A sour smile is a pattern of negative correlation between future stock returns and Wikipedia page visits that is observed only during the bull market. Both patterns might be explained as a corrective investor counter-reaction to initial overpessimism in the bear market and to initial overoptimism in the bull market.

Economic significance of the merry frown and the sour smile is explored by constructing a market neutral strategy with long positions in stocks that are in the highest decile and short positions in stocks that are in the lowest decile with regard to Wikipedia page visits during the bear market. In the bull market the positions are reversed. We backtest the strategy for different trading frequencies (from 1 to 10 days) and observe that they generate positive returns which decrease with the length of trading frequency. The returns of the strategies are compared to random market neutral strategies generated by a Monte Carlo simulation and their statistical significance is established. We also find that returns of the strategies strongly increase if they are restricted to a subsample of stocks that are preferred by individual investors such as high volatility stocks, low market capitalization stocks or low price-to-book ratio stocks. In the best case, a trading strategy with daily trading frequency that is restricted to stocks with higher than median volatility yields an average annual return of 17% in our sample period.

The paper is organized as follows: Section 2 describes the data. The market regime model is presented in Section 3. In Section 4 we discuss the results of the Fama-MacBeth regressions. The trading strategies and backtest results are presented in Section 5. Section 6 concludes the paper.

2 Data description

The most important choice in the beginning of every statistical research is the choice of statistical population. We decided to limit our study to a sample of stocks that are included in the index S&P 500. More precisely, the stocks that were members of this index on June 7, 2013. Our choice is primarily motivated by the fact that the publicly accessible and freely available data on individual stocks is of highest quality for stocks listed on the

US stock market. Of course, the extension of our study to stocks listed on some less common stock markets remains a challenge for future investigation. We choose a sampling period from October 1, 2007, up to June 30, 2013. We restrict ourselves to this specific sampling period primarily because of the availability of website search data. Fortunately, this period includes very diverse market conditions including one of the greatest market crashes in history and the following rebounding growth. This gives us confidence that our findings would easily extend to future periods. For every stock in our sample we obtain the daily dividend and split adjusted closing prices from the Yahoo Finance website[1]. We additionally remove all stocks for which data is not available for the entire sample period. This mostly includes stocks that were members of the S&P 500 index on June 7, 2013, but were not yet publicly traded at the begining of our sample period.

We get the daily closing values of the index S&P 500 in the economic data section of the website of Federal Reserve Bank of St. Louis[2]. This data is obtained for a longer period from January 1, 2000, up to June 30, 2013, as required by the regime-switching model described in Section 3.

There are various choices for website search data that are worth testing for possible relations with fluctuations in stock prices. One of the possibilities is Google search data which is publicly available via the Google Trends Service[3]. In related studies, authors most commonly use relative search frequencies of stock tickers [7, 14] while some also consider relative search frequencies of company names [2]. The majority of studies rely on weekly data for these frequencies. This might be based on availability problems with Google data. When one requests a search frequency time series for a certain term the format of the returned series depends on the length of the period. For periods no longer than 3 months one gets the daily data, but for longer periods only weekly data is returned. An additional feature of the data so obtained is that it is normalized within the series to have a maximum value of 100. This has some advantages but also makes it difficult to compare values of series in different periods. This may have been the reason for most authors to restrict their studies to weekly data.

In order to overcome this difficulty, we acquire the three-month-period data every two months. From the data for the overlapping month we compute the quotient between the normalized factors of the two consecutive periods thus enabling us to concatenate the short period time series into one long period time series with the daily data. An additional problem arises with company names. Namely, one would need to know which name people are using for the company when searching for information about it. For instance, it is unlikely that most people would search for American Express Company by typing its full name into the search window but would instead just type American Express or simply AmEx. Accordingly, we replace company names in our sample with suitable abbreviations. By following the concatenating procedure described above we obtain a daily time series of relative search frequencies for both a company ticker and an abbreviated company name for each company in our sample. The sample period for this data is chosen to be from January 1, 2008, up to May 31, 2013.

Another source of internet search data that has recently been studied related to financial data is Wikipedia. For every article on Wikipedia, a time series of daily unique page visits

[1] http://www.finance.yahoo.com
[2] http://research.stlouisfed.org/fred2
[3] http://www.google.com.au/trends

can be obtained from the website Stats.Grok.Se[4]. This source of information has not gained as much attention as the Google Trends Data and the reason for this may lie in the fact that the available time series only span from January 2008. For every company in our sample, we find the Wikipedia article associated to the company and obtain a time series of unique daily page visits to this article in a sample period from January 1, 2008, up to May 31, 2013.

There are some comments we have to make that relate to data preprocessing of both Google search and Wikipedia page visits data. We first note that this data is not available throughout the chosen period for each company in our sampling period. Therefore, we have to exclude the companies with too much missing data. Our rule is to allow no more than 10% of missing data, while for the missing values we apply an imputation procedure that takes into account the weekly seasonality. Next we perform a detrendization of data using a sort of "longitudinal normalization", i.e. we divide the number for a given day by the average of the numbers for the last 56 days. The choice of length for this normalization period is similar to choices made by authors in related studies, for example in [7] where the length of the normalization period is 8 weeks. We also have to take care of the outliers. Their influence is reduced by taking a logarithm transformation of our data. Additionally, for each stock in our sample and each of the three variables, we perform a winsorization of the corresponding time series by limiting the range of data to its first percentile from below and to its 99th percentile from above.

There are strong seasonal effects on the weekly basis in both Google and Wikipedia data. We want to make data collected on different days of the week comparable by introducing a seasonal adjustment in the following way. We regress the data to the days-of-the-week (but one) as dummy variables to get average differences between the different days of the week which we add to the data of this particular day.

3 Market regimes

We intend to study the influence of market regimes on the relation between internet search data and stock returns. A market regime may be considered as a phase of persistent attributes observed in financial time series. This concept is most commonly used by investors when classifying the market into two phases: the bear market characterized by low returns and high volatility and the bull market characterized by high returns and low volatility. This dichotomous approach also lies at the core of our view on market dynamics. A systematic regime-switching time series model was first proposed by Hamilton [13] and the variants of this model are still being overwhelmingly used to study the regimes of economic and financial time series. However, recently, the problem of parameter estimation in financial regime-switching models has also been tackled by the Baum-Welch algorithm [19] which has previously been mostly used in engineering applications. We also decided to base our market regime estimation model on the Baum-Welch algorithm and refer the reader to [15] for a discussion of its advantages over Hamilton's algorithm.

We will determine the actual switching of the regimes only on the cumulative level and base it on the returns data of the S&P 500 index. The background of the model is a hidden Markov chain (denoted by $\mathcal{Q} = (q_t)_{t=1}^T$, where T is the length of the period) so that the market regimes are seen as states of this chain, the bear market becomes say state $i = 1$ and the bull market state $i = 2$. As a consequence the model has a 2×2

[4] http://stats.grok.se

transition matrix A defined in the usual way: $a_{ij} = p(q_{t+1} = j \,|\, q_t = i)$. Given an initial distribution of the states $\Pi = (\pi_1, \pi_2)$ we then have the Markov chain uniquely determined. Furthermore, we assume that the observations form a random sequence denoted by $\mathcal{O} = (o_t)_{t=1}^T$ where each o_t is the index return at time t and is determined randomly following a normal distribution $N(\mu_i, \sigma_i)$, for $i = 1, 2$, where the two parameters of this distribution depend on the state of the hidden Markov chain, i.e. on the market regime. We will call them the observation distributions and denote the corresponding sequence of distributions by $\mathcal{B}$. The entire model will be denoted by $\mathcal{M} := \{A, \Pi, \mathcal{B}\}$.

We first present the *forward algorithm* which helps computing the so called *forward variable*

$$\kappa_t(i) = p(o_1 o_2 \dots o_t, q_t = i \,|\, \mathcal{M}),$$

Here, κ is defined recursively:

$$\kappa_1(i) = \pi_i b_i(o_1),$$

for $i = 1, 2$, where $b_i(o_1)$ is the density of $N(\mu_i, \sigma_i)$ at the point o_1. Furthermore,

$$\kappa_{t+1}(j) = \left[\sum_{i=1}^{2} \kappa_t(i) a_{ij}\right] b_j(o_{t+1}), .$$

for $j = 1, 2$ and $t = 1, 2, \dots T-1$ and $b_j(o_{t+1})$ is defined by analogy with the above. Likelihood $p(O|\mathcal{M})$ can be computed using the forward variable in the following manner:

$$p(O|\mathcal{M}) = \sum_{i=1}^{2} \kappa_T(i).$$

In the Baum-Welch algorithm we also need the *backward variable*

$$\varrho_t(i) = p(o_{t+1} o_{t+2} \dots o_T \,|\, q_t = i, \mathcal{M})$$

which we compute recursively using the *backward algorithm*. We first initialize $\varrho_T(i) = 1$ for $i = 1, 2$, and then let

$$\varrho_t(i) = \sum_{j=1}^{2} a_{ij} b_j(o_{t+1}) \varrho_{t+1}(j)$$

for $i = 1, 2$ and $t = T-1, T-2, \dots, 1$. Likelihood $p(O|\mathcal{M})$ can be computed using the backward variable in the following manner:

$$p(O|\mathcal{M}) = \sum_{i=1}^{2} \pi_i \varrho_1(i) b_i(o_1).$$

To initialize the Baum-Welch algorithm we choose a starting estimate of the model denoted by $\widehat{\mathcal{M}}_0$, and then we compute the likelihood $p(\mathcal{O} \,|\, \widehat{\mathcal{M}}_0)$ using the forward variable.

At this stage we start the iterative procedure consisting of four steps. The first step is to compute the forward variable $\widehat{\kappa}_t(i)$ and the backward variable $\widehat{\varrho}_t(i)$ based on the estimate of the model obtained on the previous iteration step $\widehat{\mathcal{M}}_k$. The final result of this step is

the likelihood of transition from the state i to the state j given the model $\widehat{\mathcal{M}}_k$ and the observations $\mathcal{O}$ when time goes from t to $t+1$:

$$\begin{aligned}\widehat{\psi}_t(i,j) &= p(q_t = i, q_{t+1} = j|\mathcal{O}, \widehat{\mathcal{M}}_k) \\ &= \frac{p(q_t = i, q_{t+1} = j, \mathcal{O} \,|\, \widehat{\mathcal{M}}_k)}{p(\mathcal{O} \,|\, \widehat{\mathcal{M}}_k)} \\ &= \frac{\widehat{\kappa}_t(i) a_{ij} b_j(o_{t+1}) \widehat{\varrho}_{t+1}(j)}{p(\mathcal{O} \,|\, \widehat{\mathcal{M}}_k)}.\end{aligned}$$

The first equation above is just the definition, in the second one we use the conditional formula and in the third one we express the likelihood with (the estimates of) the forward and backward variable. Next, we express the denominator of the last fraction above also using (the estimates of) the forward and backward variable:

$$p(\mathcal{O} \,|\, \widehat{\mathcal{M}}_k) = \sum_{i=1}^{2} \sum_{j=1}^{2} \widehat{\kappa}_t(i) a_{ij} b_j(o_{t+1}) \widehat{\varrho}_{t+1}(j).$$

On the second step of the iteration procedure we need to estimate another likelihood, i.e.

$$\Gamma_i(i) = p(q_t = i \,|\, \mathcal{O}, \mathcal{M}) = \sum_{j=1}^{R} \psi_t(i,j).$$

Using the estimates of the first step $\widehat{\psi}_t(i,j)$ we compute the next estimate of the model $\widehat{\mathcal{M}}_{k+1}$. First, we compute the elements of the transition matrix A

$$\widehat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \widehat{\psi}_t(i,j)}{\sum_{t=1}^{T-1} \Gamma_t(i)},$$

followed by the initial distribution Π

$$\widehat{\pi}_i = \Gamma_1(i)$$

and finally the observation distributions $\mathcal{B}$

$$\widehat{b}_j(s) = \frac{\sum_{t=1}^{T} \Gamma_t(j)'}{\sum_{t=1}^{T} \Gamma_t(i)}.$$

Here we understand $\Gamma_t(j)'$ given $o_t = s$.

On the third step of the iteration procedure we compute the likelihood using the new model $\widehat{\mathcal{M}}_{k+1}$. The fourth step is decisive: we compare the estimates of these likelihoods on the last two steps. If they are close enough, we stop the algorithm. If not, we proceed with another iteration starting with step one.

So, the final result of this algorithm is an estimate of the hidden Markov model $\mathcal{M} := \{A, \Pi, \mathcal{B}\}$. Based on this estimate, we want to give a prediction of the most probable state (bull or bear regime) for each point of time in the period. This will be done using the Viterbi algorithm. We first introduce, for $i = 1, 2$, the *Viterbi variable*

$$\delta_t(i) = \max_{q_1, q_2, \ldots, q_{t-1}} p(q_1, q_2, \ldots, q_{t-1}, q_t = i, o_1, o_2, \ldots, o_t \,|\, \mathcal{M})$$

which means the conditional likelihood of the most likely path of length t ending in state i given the model, where o_t are the the actual observed values of the index under consideration. We also need the value of the last but one state in this optimal path that ends in state i which we denote by $\tau_t(i)$.

We initialize by letting $\delta_1(i) = \pi_i b_i(o_1)$ and $\tau_1(i) = 0$ (an "empty state" on which nothing really depends) for $i = 1, 2$. The inductive steps of the algorithm go for $t = 2, 3, \ldots, T$. The dynamic programming approach yields

$$\delta_t(i) = b_i(o_t) \max_{i=1,2}(\delta_{t-1}(i)a_{ij})$$

together with

$$\tau_t(j) = \arg\max_{i=1,2}(\delta_{t-1}(i)a_{ij}).$$

At the end of the algorithm we terminate with the final optimal regime

$$q_T^* = \arg\max_{1=1,2}(\delta_{T-1}(i)a_{ij})$$

and then backtrack the whole optimal path

$$q_t^* = \tau_{t+1}(q_{t+1}^*)$$

for $t = T-1, T-2, \ldots, 1$.

Both the Baum-Welch algorithm and the Viterbi algorithm are not what we usually call online algorithms that would process the input data in the sequence they would be fed to the algorithm. This is a shortcoming since we are looking for a way to determining the market regime as a stopping time in the sense of martingale theory, i.e. the decision about a certain point in time can be made only based on the data of previous points in time. We are overcoming this obstacle by implementing it in an expanding window approach. The starting window in our approach will be the period from January 2, 2000 to January 2, 2008 (because January 2 is the first trading day in a year). On each step we expand the window by one trading day until we reach May 31, 2013 where the period that we are interested in ends. In each of these windows we run the Baum-Welch and the Viterbi algorithm and retain only the final optimal regime q_T^*. This way we determine the optimal market regime for each trading day of the period we are interested in a stopping time manner.

In Figure 1 we present the results of the algorithm described above. In this figure the daily values of the S&P 500 index are superimposed over two backgrounds – the red one corresponds to days of the bear market and the blue to days of the bull market according to our estimation. It is clear that our model is able to recognize quite well the bear market of 2008/2009 as well as the more pronounced market corrections in the following years. However it does perform less admirably when recognizing the beginning of the bull market in 2009. This is not unexpected since this period was characterized by extremely high returns as well as high, albeit decreasing, volatility. Such conditions are not well aligned with our model which assumes only two market regimes - the one with high returns and low volatility and the one with low returns and high volatility. An obvious solution to this problem would be extending our model to 4 regimes. However, fitting such a model would require a much larger data sample than we have available. For example, the authors of [16] fit a 4-regime model on 123 years of data. Due to this limitation, we decided for the more parsimonous 2-regime model.

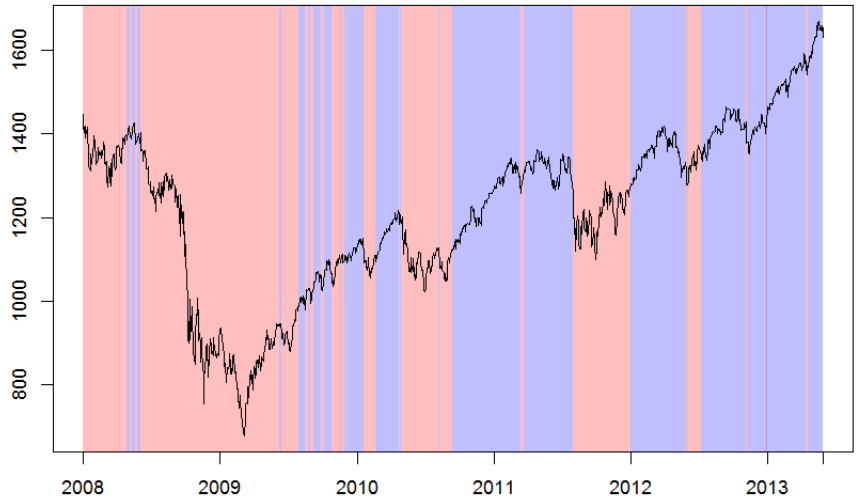


Figure 1: S&P 500 index during the bear regime (red background) and the bull regime (blue background)

4 Linear regression

In this section we study the connection between the stock returns and the data described in Section 2. Our main statistical tool will be the Fama-Macbeth cross-sectional linear regression [9]. This is a two-step procedure where a cross-sectional regression is performed for each time unit and then the time-series average of the estimated regression coefficients is calculated.

We first present the results of an analysis in which the explained variables are the cumulative future returns, while the search data is used as the regressor. For each k running from 1 to 15 we perform a Fama-Macbeth regression for the cumulative return from the time $T = t + 1$ to the time $T = t + k$ as the explained variable which we denote $r_{i,t+k}$ where i is the index spanning our entire selection of stocks. This way we allow for different periods of time that may be of interest ranging from 1 trading day to roughly 1 trading month. We perform an additional regression to test the contemporary return, i.e. the return observed on $T = t$. As described before we use three types of search data: Wikipedia page visits (denoted by $\mathrm{wiki}_{i,t}$), Google search queries for company tickers (denoted by $\mathrm{goog\_tickers}_{i,t}$) and Google search queries for company names (denoted by $\mathrm{goog\_names}_{i,t}$). All this data is taken at time $T = t$. The cross-sectional regressions performed for each time unit t and each $k = 0, 1, \ldots, 15$ are described by the following equations:

$$\begin{aligned} r_{i,t+k} &= \alpha_{i,t} + \beta_{\mathrm{wiki},t}\,\mathrm{wiki}_{i,t} + \varepsilon_{i,t+k}, \\ r_{i,t+k} &= \alpha_{i,t} + \beta_{\mathrm{goog\_tickers},t}\,\mathrm{goog\_tickers}_{i,t} + \varepsilon_{i,t+k}, \\ r_{i,t+k} &= \alpha_{i,t} + \beta_{\mathrm{goog\_names},t}\,\mathrm{goog\_names}_{i,t} + \varepsilon_{i,t+k}. \end{aligned}$$

There is an additional regressor we want to test for its influence which has a form of a dummy variable, i.e. the market regime. So, in practice we perform three actual regressions for each case of interest, one for the bear markets, one for the bull markets and one for the joint data independent of the regime. The results are presented in Table 1.

To make the presentation clearer we standardized the search data on every fixed date under consideration so that the regression coefficient has a simple interpretation. It gives an increase in the average return (positive or negative) given that the average internet search variable increases by one standard deviation. When we want to present the relation of these data to the length of the period we run into another difficulty, namely that the cumulative returns computed for different periods are not immediately comparable since their magnitude trivially depends on the period, so we decided to annualize them. There are three graphs in Figure 2. The first one presents dependence of annualized returns (based on regression coefficients) for the Wikipedia page visits, where the red line presents the data of the bear market, the blue one the data of the bull market and the purple one the joint data. A similar graph is presented for the Google search data for company tickers and the third one for the Google search data for company names.

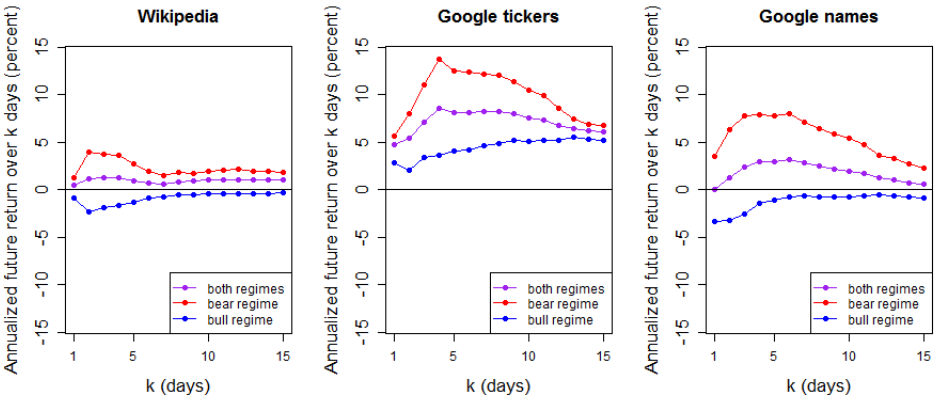


Figure 2: Changes in annualized future returns over k days after we observe a one standard deviation increase in individual search variables.

We first observe that the regression results in the case when market regimes are not taken under consideration differ substantially from the results when we do take them into account. Actually, in all the three cases of internet search variables we observe that a raise in the internet search variable is associated with higher future returns in the bear market compared to future returns in the bull market. This is confirming our starting hypothesis that market regimes have a strong influence on the connection between internet search data and stock returns. We also observe that there is a substantial difference between the importance of distinct internet search variables. Based on the analysis performed so far it seems that the possible influence of search data on stock returns is statistically the strongest for Google company tickers, followed by Google company names and finally Wikipedia page visits. It is also evident that the influence of search data on future returns is mostly short term with the largest absolute values of annualized returns (based on regression coefficients) attained for cases where $k \leqslant 10$.

	Wikipedia			Google tickers			Google names		
k	Both	Bear	Bull	Both	Bear	Bull	Both	Bear	Bull
0	**0.018***	0.010	**0.027*****	**0.017****	**0.041****	0.003	0.006	0.019	0.004
	(0.006)	(0.009)	(0.007)	(0.008)	(0.017)	(0.008)	(0.008)	(0.015)	(0.007)
1	0.002	0.005	-0.003	**0.019****	0.022	0.011	0.000	0.014	**-0.014***
	(0.004)	(0.008)	(0.004)	(0.009)	(0.020)	(0.008)	(0.008)	(0.014)	(0.007)
2	0.009	**0.031****	**-0.018*****	**0.042*****	**0.061***	0.016	0.010	**0.048****	**-0.026****
	(0.008)	(0.016)	(0.007)	(0.016)	(0.032)	(0.015)	(0.013)	(0.021)	(0.012)
3	0.015	**0.044****	**-0.022****	**0.082*****	**0.125*****	**0.040***	0.028	**0.089*****	**-0.030***
	(0.012)	(0.022)	(0.010)	(0.024)	(0.044)	(0.022)	(0.018)	(0.030)	(0.017)
4	0.020	**0.057****	**-0.026****	**0.131*****	**0.205*****	**0.056****	**0.046****	**0.121*****	-0.022
	(0.014)	(0.026)	(0.012)	(0.032)	(0.058)	(0.026)	(0.022)	(0.037)	(0.021)
5	0.018	**0.053***	**-0.027***	**0.156*****	**0.235*****	**0.079****	**0.059****	**0.148*****	-0.021
	(0.017)	(0.031)	(0.014)	(0.039)	(0.069)	(0.031)	(0.027)	(0.047)	(0.025)
6	0.018	0.046	-0.021	**0.186*****	**0.280*****	**0.098*****	**0.075****	**0.184*****	-0.019
	(0.020)	(0.035)	(0.016)	(0.047)	(0.083)	(0.037)	(0.033)	(0.056)	(0.029)
7	0.017	0.042	-0.021	**0.221*****	**0.320*****	**0.128*****	**0.078****	**0.191*****	-0.018
	(0.024)	(0.040)	(0.019)	(0.055)	(0.095)	(0.045)	(0.037)	(0.061)	(0.033)
8	0.027	0.056	-0.018	**0.251*****	**0.363*****	**0.152*****	**0.077***	**0.198*****	-0.024
	(0.026)	(0.044)	(0.022)	(0.062)	(0.104)	(0.052)	(0.040)	(0.066)	(0.036)
9	0.032	0.063	-0.018	**0.277*****	**0.387*****	**0.181*****	**0.077***	**0.205*****	-0.028
	(0.029)	(0.048)	(0.023)	(0.066)	(0.112)	(0.057)	(0.044)	(0.070)	(0.040)
10	0.040	0.076	-0.015	**0.292*****	**0.398*****	**0.198*****	0.075	**0.208*****	-0.031
	(0.031)	(0.051)	(0.025)	(0.072)	(0.122)	(0.061)	(0.046)	(0.075)	(0.043)
11	0.047	**0.091***	-0.018	**0.312*****	**0.414*****	**0.223*****	0.074	**0.206****	-0.030
	(0.033)	(0.054)	(0.027)	(0.076)	(0.128)	(0.066)	(0.049)	(0.080)	(0.046)
12	0.052	**0.100***	-0.021	**0.314*****	**0.392*****	**0.243*****	0.059	**0.169****	-0.025
	(0.035)	(0.058)	(0.029)	(0.079)	(0.135)	(0.070)	(0.052)	(0.085)	(0.048)
13	0.053	0.099	-0.020	**0.325*****	**0.371*****	**0.278*****	0.054	**0.168***	-0.035
	(0.037)	(0.061)	(0.031)	(0.084)	(0.143)	(0.076)	(0.054)	(0.089)	(0.051)
14	0.056	**0.107***	-0.022	**0.334*****	**0.372****	**0.290*****	0.042	0.148	-0.042
	(0.039)	(0.063)	(0.033)	(0.088)	(0.150)	(0.081)	(0.056)	(0.092)	(0.053)
15	0.061	0.109	-0.016	**0.353*****	**0.394****	**0.306*****	0.034	0.137	-0.049
	(0.042)	(0.068)	(0.035)	(0.094)	(0.159)	(0.087)	(0.059)	(0.097)	(0.056)

Table 1: Regression coefficients of internet search variables in Fama-MacBeth regressions where cumulative future stock returns are the explained variable. Table columns correspond to different regressor-regime combinations and table rows correspond to different horizons of future returns. Standard errors for regression coefficients are given in parentheses. Statistical significance at levels of 10%, 5% and 1% is denoted by *,**, and ***, respectively. Additionally, statistically significant results are printed in bold.

We observe another phenomenon which is best seen in the case of Wikipedia page visits. In the bear market the values of returns first go up and then go back down so that they form a shape of a frown. During the bull market, on the other hand, we observe a mirror shape, i.e. a shape of a smile. Now, the interpretation of these shapes is in some sense the opposite of the usual meaning conveyed by these shapes. While the frown noticed means good news in bear times, the smile means bad news in bull times. So, we propose the two shapes to be called the "merry frown" and the "sour smile". These shapes are not so easy to interpret. A possible explanation (taking into account also some other details of the Wiki shape) is that in bear markets investors are pessimistic and their overpessimistic reaction after increased attention perceived via the number of Wikipedia page visits on the first day, results in a counter-reaction in the days to follow and creates the merry frown. In the bull markets though investors are optimistic and their overoptimistic immediate reaction on the increased attention overturns into a sour smile.

In the next step we investigate whether the observed connection between internet search data and stock returns can be explained by including additional factors into our model. We replace our initial explained variable $r_{i,t}$ (future cumulative returns) by the so-called abnormal cumulative returns $ar_{i,t}$. These returns are obtained as residuals in a variant of the Carhart [5] four factor asset pricing model which is an extension of the well known Fama-French model [10]. The model is defined by the following equation:

$$r_{i,t+k} = r_t^{rf} + \beta_{1,i}(r_t^{mkt} - r_t^{rf}) + \beta_{2,i}\,\mathrm{HML}_t + \beta_{3,i}\,\mathrm{SMB}_t + \beta_{4,i}\,\mathrm{UMD}_t + ar_{i,t+k},$$

where r_t^{rf} is the risk-free rate of return (approximated by the daily rate of one month U.S. Treasury bills), $r_t^{mkt} - r_t^{rf}$ is the excess return of the entire stock market over the risk-free return, HML_t is the return difference between a portfolio of stocks with high and low book-to-market stocks, SMB_t is the return difference between a portfolio of small and big stocks in terms of their market capitalization and UMD_t is the return difference between a portfolio of stocks with high and low returns in the past year. The betas are estimated on a daily basis, using a rolling window of 120 days.

We repeat the cumulative return regressions described above for the case of abnormal cumulative returns and report the results in Table 2. The period dependencies of annualized abnormal returns are displayed in Figure 3. We see that results in the case of Wikipedia page visits variable are quite similar to those obtained before accounting for the Carhart factors. However the influence of Google search queries for company tickers and company names is greatly diminished. In fact, no statistically significant results at the 5% level are obtained for the company tickers variable regardless of the bear or bull market. This is in contrast to previous research which was performed on samples taken from earlier periods. Furthermore, we find little evidence that a rise in internet search variables corresponding to individual companies might directly translate into short-term buying pressure and consequently higher stock prices. The differences between the bear market and the bull market remain clearly visible, especially in the case of Wikipedia page visits. While our results show that a statistically significant dependence between future stock returns and internet search variables exists, we do note that the explanatory power of all the tested regressions as measured by the R^2 statistic is very low and only rises above 1% in a few cases. This is not unexpected if we take into account the fact that the regressions are predictive, that we are only using a single explanatory variable and that future stock returns are notoriously hard to predict. The questions whether our observations can nevertheless be used to obtain economic gains will be explored in Section 5.

	Wikipedia			Google tickers			Google names		
k	Both	Bear	Bull	Both	Bear	Bull	Both	Bear	Bull
0	**0.014*****	0.006	**0.022*****	**0.011***	0.010	**0.015****	-0.000	-0.006	0.006
	(0.004)	(0.007)	(0.006)	(0.006)	(0.012)	(0.006)	(0.005)	(0.008)	(0.006)
1	0.002	0.006	-0.003	0.002	0.004	0.003	**-0.010****	-0.012	-0.009
	(0.003)	(0.006)	(0.003)	(0.006)	(0.011)	(0.006)	(0.005)	(0.008)	(0.005)
2	0.005	**0.018***	**-0.013****	0.005	0.006	0.003	**-0.016***	-0.009	**-0.022****
	(0.006)	(0.011)	(0.006)	(0.010)	(0.019)	(0.010)	(0.009)	(0.014)	(0.009)
3	0.012	**0.034****	**-0.015****	0.003	0.001	0.005	**-0.024****	-0.018	**-0.028****
	(0.008)	(0.015)	(0.008)	(0.015)	(0.027)	(0.013)	(0.012)	(0.020)	(0.013)
4	**0.018***	**0.048****	**-0.016***	0.017	0.030	0.006	-0.022	-0.009	**-0.031***
	(0.011)	(0.019)	(0.009)	(0.019)	(0.035)	(0.016)	(0.015)	(0.026)	(0.016)
5	0.019	**0.048****	-0.017	0.019	0.034	0.005	-0.022	0.000	**-0.038****
	(0.013)	(0.021)	(0.011)	(0.023)	(0.042)	(0.019)	(0.019)	(0.031)	(0.019)
6	0.020	**0.046***	-0.013	0.024	0.042	0.009	-0.024	0.004	**-0.044****
	(0.014)	(0.024)	(0.013)	(0.026)	(0.050)	(0.023)	(0.021)	(0.036)	(0.022)
7	0.024	**0.054****	-0.015	0.038	0.064	0.018	-0.021	0.011	**-0.043***
	(0.016)	(0.027)	(0.014)	(0.030)	(0.056)	(0.027)	(0.024)	(0.040)	(0.025)
8	0.027	**0.059****	-0.015	0.046	0.072	0.027	-0.025	0.012	**-0.052***
	(0.018)	(0.029)	(0.016)	(0.033)	(0.062)	(0.032)	(0.027)	(0.045)	(0.028)
9	0.029	**0.062****	-0.014	0.059	0.084	0.044	-0.032	0.007	**-0.062****
	(0.019)	(0.031)	(0.017)	(0.038)	(0.070)	(0.036)	(0.029)	(0.048)	(0.030)
10	0.031	**0.068****	-0.016	0.057	0.076	0.050	-0.043	-0.006	**-0.070****
	(0.020)	(0.033)	(0.019)	(0.041)	(0.077)	(0.039)	(0.032)	(0.052)	(0.032)
11	0.034	**0.074****	-0.016	0.066	0.089	0.058	-0.051	-0.017	**-0.075****
	(0.022)	(0.036)	(0.021)	(0.045)	(0.084)	(0.043)	(0.034)	(0.056)	(0.034)
12	0.036	**0.077****	-0.017	0.066	0.084	0.061	-0.058	-0.029	**-0.079****
	(0.023)	(0.039)	(0.022)	(0.048)	(0.087)	(0.046)	(0.037)	(0.061)	(0.036)
13	0.039	**0.081***	-0.013	0.070	0.081	0.072	**-0.067***	-0.041	**-0.085****
	(0.025)	(0.042)	(0.023)	(0.051)	(0.094)	(0.048)	(0.039)	(0.064)	(0.039)
14	0.040	**0.086***	-0.018	0.070	0.076	0.078	**-0.077***	-0.060	**-0.087****
	(0.027)	(0.044)	(0.024)	(0.054)	(0.101)	(0.050)	(0.041)	(0.067)	(0.040)
15	0.039	**0.084***	-0.018	0.072	0.074	0.085	**-0.085***	-0.075	**-0.088****
	(0.029)	(0.047)	(0.026)	(0.057)	(0.107)	(0.053)	(0.044)	(0.070)	(0.043)

Table 2: Regression coefficients of internet search variables in Fama-MacBeth regressions where cumulative abnormal future stock returns are the explained variable. Table columns correspond to different regressor-regime combinations and table rows correspond to different horizons of future returns. Standard errors for regression coefficients are given in parentheses. Statistical significance at levels of 10%, 5% and 1% is denoted by *,**, and ***, respectively. Additionally, statistically significant results are printed in bold.

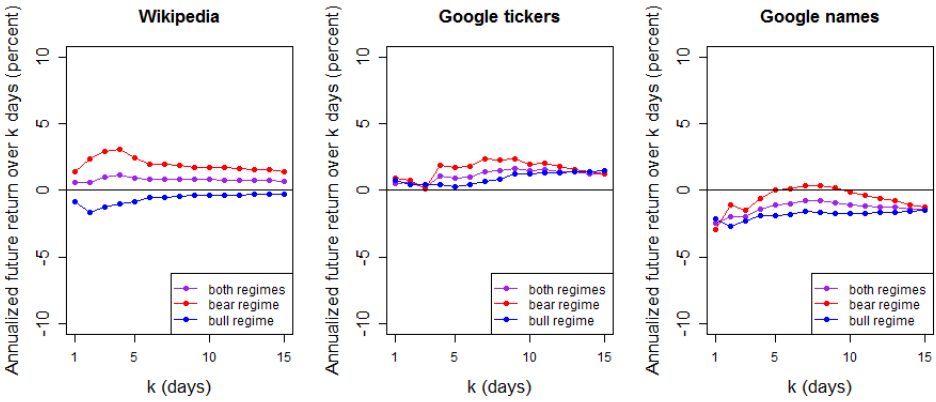


Figure 3: Changes in abnormal annualized future returns over k days after we observe a one standard deviation increase in individual search variables.

Based on these tests we conclude that the influence of the kind of attention noticed by Google search queries (either for company tickers or names) can be perceived also by other data that are more commonly applied by financial practitioners. On the other hand, it shows that Wikipedia page visits do indeed provide new information about the behavior of stock prices. Also, we perceive that the merry frown and sour smile effects persist for Wikipedia page visits even after controlling for the most commonly used asset pricing factors.

5 Trading strategies

In this section we want to verify how the results of Section 4 can be used, if at all, in forming trading strategies. In other words, we want to either statistically prove or disprove that internet data can increase our profits in financial markets. The evidence for the influence of internet search data on future stock returns is most compelling in the case of Wikipedia data, as shown in Section 4. Since we were also not able to find any examples in existing literature of this type of data being used in construction of trading strategies based on individual stocks, we decided to limit our analysis in this chapter only to Wikipedia page visits.

Our results show that in bear markets higher Wikipedia page visits are positively correlated with short term future return while in bull markets the corresponding correlation is negative. So we propose the strategy for bear markets to enter long position at the end of the trading day for stocks in the upper decile with respect to the most recent available data on Wikipedia page visits; and similarly to enter the short position at the end of the trading day for stocks in the lower decile with respect to these visits. In the bull market, the strategy is to do exactly the opposite. We propose that all the long positions and all the short positions are entered using the same weights with respect to the wealth that we are prepared to invest into this strategy. Since the data on Wikipedia page visits for any given day is only made available the following day we lag our Wikipedia variable for one day to ensure that the data would have been available at the time of our trading decision.

Of course there is a problem of determining the actual frequency of trading, this means for how long we should hold our positions. We know that we are talking about a short term

effect, but what does short really mean in this particular context? To make this dilemma as clear as possible we are making a number of tests using some alternatives. Let f be the number of trading days between two consecutive trading decisions. For $f = 1, 2, \ldots, 8$ we are testing the fth strategy and give the result for three options. The first option is that we allow only trading in the bear market, the second one is that we allow trading only in the bull market and the third one is that we allow trading during both markets. In Table 3 we present the results obtained in percentage points of the annual return. It is clear that $f = 1$ is the best of the proposed strategies in all the three cases. It is also clear that the results are getting smaller with f increasing in the case of combined strategy and the bull-only strategy. However in the case of the bear strategy $f = 2$ and $f = 3$ are slightly better than $f = 1$. For f big enough the results of the strategies seem to become more or less random. The best of the three options tested is the combined application of both bear and bull strategy. It is also clear that the results of the bear strategy are better than the results of the bull strategy.

Trading frequency (days)	1	2	3	4	5	6	7	8
Bear	5.28	5.50	6.10	4.02	2.30	0.94	-0.55	2.21
Bull	2.76	1.33	-0.15	0.91	0.59	1.02	1.08	-2.53
Joint Bear & Bull	8.19	6.91	5.94	4.97	2.91	1.97	0.52	-0.37

Table 3: Average annual returns (in percentage points) of proposed trading strategies in relation to the trading frequency.

We also want to compare our strategies to suitable benchmark strategies. However, as we believe, the most usual benchmarks such as various indices are long-only strategies and the comparative testing with our strategies which include both long and short positions would not be fair. So we decided to compare it with random strategies using a Monte Carlo approach. Our control strategy is to choose in a uniformly random way 10 % stocks to be put in a long position and 10 % to be put in a short position. We created 1000 strategies of this type and computed the average yearly return for each of them. This produces a random sample of possible average yearly returns which we compare statistically to the average return of each of the strategies under consideration. As usual in this kind of situation, we perform a one-sample one-way Student t-test where we test the null hypothesis that the mean yearly return for the population of random strategies is equal to the return of our strategies against the alternative hypothesis which states that the mean yearly return for the population of random strategies is lower to the return of our strategies. As can be seen from results given in Table 4 we can reject this hypothesis for our joint bear and bull strategies for most of the trading frequencies considered.

In Figure 4 we want to present a slightly different view on the results of our strategies compared to the random approach. Assume we invest a certain equity in the strategies above to be compared; and that we invest the same amount into each of the random strategies described in the previous paragraph. We compare the average of the randomly invested equity to the equity gained via the strategy under consideration for each day of our sample period. More interesting than the averages as such are the bands created around the averages using the daily standard deviation and its small multiples. We can see that the equity invested in our joint bear and bull strategy mostly stays in the area that is beyond the band which is three standard deviations above the average equity of random strategies.

Trading frequency (days)	1	2	3	4	5	6	7	8
μ_0	8.19	6.91	5.94	4.97	2.91	1.97	0.52	-0.37
μ	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25
t value	-64.55	-54.78	-47.34	-39.91	-24.18	-16.96	-5.89	0.92
p value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82

Table 4: The results of a one-way Student t-test for testing the null hypothesis that the mean yearly return (μ) for the population of random strategies is equal to the return (μ_0) for our joint bear and bull strategy against the alternative hypothesis $\mu < \mu_0$. The p values are given in percentage points and rounded to two decimals.

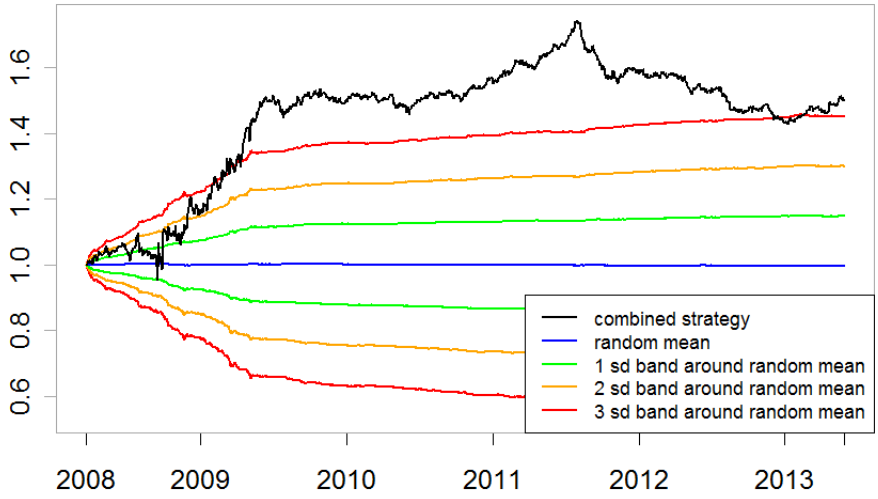


Figure 4: Equity curve of joint bear and bull strategy with trading frequency of one day compared to equities of random strategies represented by standard deviation bands around the mean equity.

In most cases the Wikipedia pages on individual S&P 500 companies contains only the most basic information. It is therefore safe to assume that this information source will mostly be utilized by individual investors since institutional investors have access to more sophisticated tools offering greater depth of information. Our hypothesis is that the influence of Wikipedia page visits on future stock returns will be higher for stocks that are likely to attract a higher proportion of individual investors. According to Barber and Odean [4] the individual investors generally have a tendency to tilt their stock investments towards high-beta, small and value stocks. In light of this result we construct three additional strategies based on our joint bear and bull strategy. In all these strategies we restrict our trading decision to a subsample of stocks that fall above or below the median of one of the following variables: volatility, market capitalization and price-to-book ratio. In the first strategy

we choose a subsample of high volatility stocks, in the second one we choose a subsample of low market capitalization stocks and in the last one we chose a subsample of low price-to-book ratio. Volatility is calculated in a 20 trading day rolling window approach. The market capitalization and price-to-book ratio variables are obtained from the ADVFN service[5]. We present the results in Table 5. The results strongly support our hypothesis since all three subsample strategies outperform the full sample joint bear and bull strategy in cases of the most relevant trading frequencies ($1 \leq f \leq 5$).

Trading frequency (days)	1	2	3	4	5	6	7	8
Full sample	8.19	6.91	5.94	4.97	2.91	1.97	0.52	-0.37
High volatility subsample	15.35	16.40	10.13	8.53	4.12	7.15	-0.78	1.32
Low market cap subsample	17.43	13.41	9.07	6.86	4.53	5.49	0.31	1.12
Low price to book subsample	13.57	9.99	7.16	8.15	5.87	0.81	-0.64	3.30

Table 5: Average yearly returns (in percentage points) of subsample strategies compared to average yearly return of full sample joint bear and bull strategy.

All of the strategies presented up to this point have included both long and short positions. Since many investors face restrictions with respect to opening short positions in stocks, the question naturally arises whether our strategies can be adapted to be long-only. Let us consider the simplest possible adaptation which is the strategy where the investment rule during the bear market is to enter long positions at the end of each trading day for stocks in the upper decile with respect to the most recent available data on Wikipedia page visits. During the bull market, the strategy enters long positions for stocks which are in the lower decile with respect to the Wikipedia page visits. The annualized return of such a strategy in our sample period is 20.36%. Since the strategy is long-only, it is reasonable to compare its performance to that of the S&P 500 index whose annualized return during our sample period is merely 4.36%. The equity curves obtained by investing the same amount of wealth in both our adapted long-only strategy and the S&P 500 index are shown in Figure 5. The backtesting results favor the conclusion that even those investors who are restricted to only opening long positions might benefit from including the information about Wikipedia page visits in their investment decisions.

6 Conclusion

The key point of our paper is that it is essential to incorporate information about the market regime when studying the influence of internet search data on stock returns. This is clearly true for all the search variables considered since all show markedly higher correlations with future stock returns in the bear regime than in the bull regime. However, the distinction between the two regimes is especially singnificant in the case of the Wikipedia variable where we observe two inverse price patterns - a merry frown in the bear regime and a sour smile in the bull regime. Our regime estimation method is based on a hidden Markov model that only accounts for information revealed to us through the price fluctuations of the S&P 500 index. We suspect that even more interesting results might be obtained if search data were somehow included into the regime switching model itself, perhaps by building upon existing research into estimation of investor sentiment by internet search data such as [8].

After controlling for the Carhart factors the Wikipedia page visits variable emerges

[5] http://www.advfn.com

Figure 5: Equity curve of adapted long-only joint bear and bull strategy with trading frequency of one day compared to equity curve of S&P 500 index.

as the one with the most significant influence on future stock returns. Until recently this data set has been largely overlooked by researchers however we believe that it holds great potential for future applications. In a surprising turn, both of the Google search variables prove to be statistically insignificant for most periods of future return for stocks in our sample. This result is at odds with previous studies performed on earlier sample periods and warrants further research that would explain this discrepancy. We suggest that this might be caused by arbitrageurs already taking advantage of the effect of company-related Google search frequencies in line with the weak-form market-efficiency hypothesis.

We would like to make an additional point about Google Trends data with regard to future research. We noticed that previous studies have almost exclusively focused on relative search frequencies which is most likely due to the fact that individual time series obtained from the Google Trends service are normalized within series so that their values always span the interval from 0 to 100. In Section 2 we describe a straightforward approach that enables us to obtain full sample daily trends data regardless of normalization. A quite similar approach might be used to obtain data where non-relative search frequencies of two different terms can be compared. It would be interesting to know whether such data provides us with an even better proxy for investors' attention.

We also believe that the results presented in our paper may be of benefit to financial practitioners in at least two ways. Firstly, we show that Wikipedia can provide investors with insights into a stock's risk profile that are overlooked by existing asset pricing models such as the Carhart four-factor model. Secondly, the trading strategies presented in Section 5 may be of interest to speculative investors who are comfortable executing trading

strategies with target investment holding periods of less than a week.

References

[1] A. Ang and A. Timmermann, Regime Changes and Financial Markets, *Annual Review of Financial Economics*, **4** (2012), 313-337.

[2] M. Bank, M. Larch and G. Peter, Google search volume and its influence on liquidity and returns of German stocks, *Financial Markets and Portfolio Management*, **25** (2011), 239-264.

[3] B. M. Barber and T. Odean, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies*, **21** (2008), 785-818.

[4] B. M. Barber and T. Odean, Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors, *Journal of Finance*, **55** (2000), 773-806.

[5] M. M. Carhart, On the persistence in mutual fund performance, *Journal of Finance*, **52** (1997), 57-82.

[6] H. Choi and H. Varian, Predicting the Present with Google Trends, *Economic Record*, **88** (2012), 2-9.

[7] Z. Da, J. Engelberg and P. Gao, In Search of Attention, *Journal of Finance*, **66** (2010), 1461-1499.

[8] Z. Da, J. Engelberg and P. Gao, The Sum of all FEARS: Investor Sentiment and Asset Prices, *Working Paper*, (2010).

[9] E. F. Fama and J. MacBeth, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy*, **81** (1973), 607-636,

[10] E. F. Fama and K. R. French, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, **33**, 3-56.

[11] M. S. Drake, D. T. Roulstone and J. R. Thornock, Investor Information Demand: Evidence from Google Searches Around Earnings Announcements, *Journal of Accounting Research*, **50** (2012), 1001-1040.

[12] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski and L. Brilliant, Detecting influenza epidemics using search engine query data, *Nature*, **457** (2009), 1012-1014.

[13] J. D. Hamilton, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, **57** (1989), 357-384.

[14] K. Joseph, M. B. Wintoki, Z. Zhang, Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search, *International Journal of Forecasting*, **27** (2011), 1116-1127.

[15] S. Mitra, P. Date, Regime switching volatility calibration by the Baum-Welch method, *Journal of Computational and Applied Mathematics*, **234** (2010), 3243-3260.

[16] J. M. Maheu, T. H. McCurdy, Y. Song, Extracting Bull and Bear Markets from Stock Returns, *Working Paper*, (2009).

[17] H. S. Moat, C. Curme, A. Avakian, D. Y. Kenett, H. E. Stanley, and T. Preis, Quantifying Wikipedia usage patterns before stock market moves, *Scientific Reports*, **3** (2013)

[18] T. Preis, H. S. Moat and H. E. Stanley, Quantifying Trading Behavior in Financial Markets Using Google Trends, *Scientific Reports*, **3** (2013)

[19] L. R. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, *Proceedings of the IEEE*, **77** (1989), 257-285.

[20] N. Vlastakis and R. N. Markellos, Information demand and stock market volatility, *Journal of Banking & Finance*, **6** (2012), 1808-1821.

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 9 (2015) 321–344

Subdivision into i-packings and S-packing chromatic number of some lattices

Nicolas Gastineau *
LE2I, UMR CNRS 6306, Université de Bourgogne, 21078 Dijon, France

Hamamache Kheddouci
Université de Lyon, CNRS, Université Lyon 1, LIRIS, UMR5205, F-69622, France

Olivier Togni
LE2I, UMR CNRS 6306, Université de Bourgogne, 21078 Dijon, France

Received 30 January 2013, accepted 3 June 2015, published online 7 August 2015

Abstract

An i-packing in a graph G is a set of vertices at pairwise distance greater than i. For a nondecreasing sequence of integers $S = (s_1, s_2, \ldots)$, the S-packing chromatic number of a graph G is the least integer k such that there exists a coloring of G into k colors where each set of vertices colored i, $i = 1, \ldots, k$, is an s_i-packing. This paper describes various subdivisions of an i-packing into j-packings ($j > i$) for the hexagonal, square and triangular lattices. These results allow us to bound the S-packing chromatic number for these graphs, with more precise bounds and exact values for sequences $S = (s_i, i \in \mathbb{N}^*)$, $s_i = d + \lfloor (i-1)/n \rfloor$.

Keywords: Packing chromatic number, i-packing, hexagonal lattice, square lattice, triangular lattice, distance coloring.

Math. Subj. Class.: 05C15, 05C63, 05C70.

1 Introduction

Let $G = (V, E)$ be a (finite or infinite) graph and let $N(u) = \{v \in V(G) |\ uv \in E(G)\}$ be the set of neighbors of vertex u. A set $X_i \subseteq V(G)$ is an i-packing if for any distinct pair u, $v \in X_i$, $d(u, v) > i$, where $d(u, v)$ denotes the usual shortest path distance between u and

*Partially supported by the Burgundy Council and the Rhône-Alpes Council.
E-mail addresses: Nicolas.Gastineau@u-bourgogne.fr (Nicolas Gastineau), hamamache.kheddouci@univ-lyon1.fr (Hamamache Kheddouci), Olivier.Togni@u-bourgogne.fr (Olivier Togni)

This work is licensed under http://creativecommons.org/licenses/by/3.0/

v. We will use X_i to refer to an i-packing in a graph G. A k-coloring c of G is a map from $V(G)$ to $\{1,\ldots,k\}$ such that for every pair of adjacent vertices (u,v), we have $c(u) \neq c(v)$. For a graph G and a k-coloring c of G, let c_i be $\{u \in V(G) |\ c(u) = i\}$. The smallest integer k such that there exists a k-coloring of G for which for every i, with $1 \leq i \leq k$, c_i is a i-packing, is called the *packing chromatic number* of G and is denoted by $\chi_\rho(G)$. This concept was introduced by Goddard *et al.* [7] under the name of broadcast chromatic number. More generally, for a nondecreasing sequence of integers $S = (s_1, s_2, \ldots)$, an *S-packing k-coloring* is a k-coloring c of $V(G)$ such that for every i, with $1 \leq i \leq k$, c_i is a s_i-packing. Such a coloring will also simply be called an $(s_1,\ldots,s_k)$-coloring. The *S-packing chromatic number* of G denoted by $\chi_\rho^S(G)$, is the smallest k such that G admits an S-packing k-coloring. For the sequences $S = (s_i, i \in \mathbb{N}^*)$, with $s_i = d + \lfloor (i-1)/n \rfloor$, we call $\chi_\rho^S(G)$ the *(d,n)-packing chromatic number* and denote it by $\chi_\rho^{d,n}(G)$. For any connected graph G such that $|V(G)| \geq d+1$, $\chi_\rho^{d,n}(G) \geq d+1$ and $\chi_\rho^{1,1}(G) = \chi_\rho(G)$. For every bipartite graph G, $\chi_\rho^{1,2}(G) = 2$ (a bipartite graph is 2-colorable). Moreover, the smallest n such that $\chi_\rho^{d,n}(G) = n$ corresponds to the d-distant chromatic number [12], *i.e.* the minimum number of d-packings that form a partition of the vertices.

Let P_∞ denote the two-way infinite path, let $\mathbb{Z}^2 = P_\infty \Box P_\infty$ denote the planar square lattice (where $\Box$ is the Cartesian product), $\mathcal{T}$ denote the planar triangular lattice and $\mathcal{H}$ denote the planar hexagonal lattice. In this article, for an $(s_1, s_2, \ldots)$-coloring of a graph, we prefer to map vertices to the color multiset $\{s_1, s_2, \ldots\}$ even if two colors can then be denoted by the same number. This notation allows the reader to directly see to which type of packing the vertex belong depending on its color. When needed, we will denote colors of vertices in different i-packings by $i_a, i_b, \ldots$.

1.1 Motivation and related work

Packing colorings in graphs are inspired from frequency planning in wireless systems. The concept of S-packing coloring emphasizes the fact that signals can have different powers similar to the packing coloring but enables the presence of several signals with the same power, providing a more realistic model for the frequency assignment problem.

The packing chromatic number of lattices has been studied by several authors: Soukal and Holub [13] proved that $\chi_\rho(\mathbb{Z}^2) \leq 17$, Ekstein *et al.* [1] that $12 \leq \chi_\rho(\mathbb{Z}^2)$; Fiala *et al.* [4] showed that $\chi_\rho(\mathcal{H}) \leq 7$, $\chi_\rho(\mathbb{Z}^2 \Box P_2) = \infty$ and $\chi_\rho(\mathcal{H} \Box P_6) = \infty$ and Finbow and Rall [5] proved that $\chi_\rho(\mathcal{T}) = \infty$.

S-packing colorings with sequences S other than $(1,2,\ldots,k)$ first appear in [7, 3]. Goddard and Xu [8] have recently studied S-packing colorings for the infinite path and for square and triangular lattices, determining conditions on the first elements of the sequence for which the graph is or is not S-packing-colorable.

Regarding the complexity, Goddard *et al.* [7] proved that the problem of $(1,2,3)$-packing coloring is polynomial while $(1,2,3,4)$-packing coloring and $(1,2,2)$-packing coloring are NP-complete. Fiala and Golovach [3] showed that the problem of $(1,2,\ldots,k)$-coloring is NP-complete for trees. The NP-completeness of $(1,1,2)$-coloring was proved by Goddard and Xu [9] and afterward by Gastineau [6].

While the packing coloring corresponds to an S-packing coloring with a strictly increasing sequence and the d-distant chromatic number corresponds to a constant one, the sequence in the (d,n)-packing coloring also tends to infinity, but the parameter n allows us to control its growth.

$d\backslash n$	1	2	3	4	5	6
1	**7**[4, 11]	**2**	**2**	**2**	**2**	**2**
2	∞	5 - 8	**5**	**4**	**4**	**4**
3	∞	15 - 35	9 - 13	8 - 10	7 - 8	**6**
4	∞	61 - ?	20 - 58	15 - 27	13 - 21	12 - 18
5	∞	∞	37 - ?	25 - ?	21 - ?	19 - ?
8	∞	∞	∞	?	?	?
11	∞	∞	∞	∞	?	?
13	∞	∞	∞	∞	∞	?
16	∞	∞	∞	∞	∞	∞

Table 1: Bounds for (d, n)-packing chromatic numbers of the hexagonal lattice.

Moreover, one can note that all the S-packing colorings of square and hexagonal lattices published so far have the property that the s_1-packing is maximum and the other s_i-packings are obtained by subdivisions of s_1-packings (and are not always maximum). Therefore, we find it interesting to study subdivision of an i-packing into j-packings, $j > i$, in lattices. These subdivisions can in turn be used to describe patterns to obtain an S-packing coloring of a lattice. However, determining the families of graphs G for which for any S such that G is S-colorable, the S-coloring satisfies the above property remains an open question. Recently Goddard and Xu [8] proved that there exist nondecreasing sequences S such that P_∞ is S-colorable and in any $(s_1, \ldots, \chi_\rho^S(P_\infty))$-packing coloring of P_∞, the s_1-packing is not maximum, showing that for P_∞, there are sequences S for which the above property is not satisfied.

1.2 Our results

The second section introduces some definitions and results related to density. The third section introduces some subdivision of the lattices into i-packings. The fourth and fifth sections give lower bounds resulting from Section 2 and upper bounds resulting from Section 3 for the S-packing chromatic number and the (d, n)-packing chromatic number of the lattices $\mathcal{H}$, $\mathbb{Z}^2$ and $\mathcal{T}$. Tables 1 , 2 and 3 summarize the values obtained in this paper for the (d, n)-packing chromatic number, giving an idea of our results. The emphasized numbers are exact values and all pairs of values are lower and upper bounds. Lower bounds have been calculated from Proposition 2.2 and Propositions 2.5, 2.7 and 2.9. Some of the results for square and triangular lattice have been found independently by Goddard and Xu [8].

2 Density of i-packings

2.1 Density of an i-packing in a lattice

Let $G = (V, E)$ be a graph, finite or infinite and let n be a positive integer. For a vertex x of G, the ball of radius n centered at x is the set $B_n(x) = \{v \in V(G) | d_G(x, v) \leq n\}$ and the sphere of radius n centered at x is the set $\partial B_n(x) = \{v \in V(G) | d_G(x, v) = n\}$. The density of a set of vertices $X \subseteq V(G)$ is $d(X) = \limsup_{l \to \infty} \max_{x \in V} \{\frac{|X \cap B_l(x)|}{|B_l(x)|}\}$.

The notion of k-area was introduced by Fiala *et al.* [4]. We propose here a slightly modified

$d\backslash n$	1	2	3	4	5	6
1	12 - 17 [1, 13]	**2**	**2**	**2**	**2**	**2**
2	∞	11 - 20	7 - 8	**6** [8]	**5** [8]	**5**
3	∞	57 - ?	16 - 33	12 - 20	10 - 17	10 - 14
4	∞	∞	44 - ?	25 - 56	20 - 34	18 - 28
5	∞	∞	199 - ?	50 - ?	35 - ?	29 - ?
6	∞	∞	∞	?	?	?
8	∞	∞	∞	∞	?	?
10	∞	∞	∞	∞	∞	?
12	∞	∞	∞	∞	∞	∞

Table 2: Bounds for (d, n)-packing chromatic numbers of the square lattice.

$d\backslash n$	1	2	3	4	5	6
1	∞[5]	5 - 6 [8]	**3**	**3**	**3**	**3**
2	∞	127 - ?	14 - ?	10 - 16	9 - 13	8 - 10
3	∞	∞	81 - ?	28 - 72	20 - 38	17 - 26
4	∞	∞	∞	104 - ?	49 - ?	36 - ?
5	∞	∞	∞	∞	?	?
7	∞	∞	∞	∞	∞	?
8	∞	∞	∞	∞	∞	∞

Table 3: Bounds for (d, n)-packing chromatic numbers of the triangular lattice.

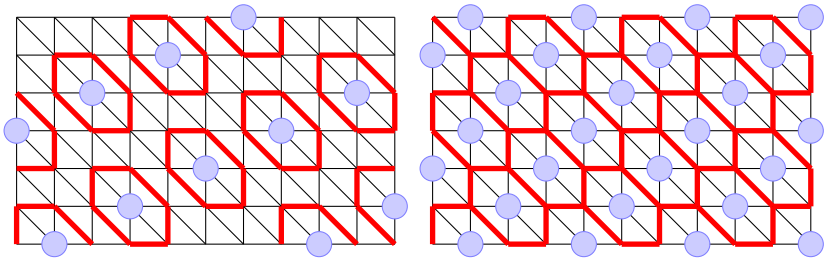


Figure 1: Examples of k-area in $\mathcal{T}$.

definition:

Definition 2.1. Let G be a graph, $x \in V(G)$, and let k be a positive integer. The k-area $A(x,k)$ assigned to G is defined by :

$$A(x,k) = \begin{cases} |B_{k/2}(x)| & \text{for } k \text{ even;} \\ |B_{\lfloor k/2 \rfloor}(x)| + \sum\limits_{u \in \partial B_{\lceil k/2 \rceil}(x)} \frac{|N(u) \cap B_{\lfloor k/2 \rfloor}(x)| + |N(u) \cap \partial B_{\lceil k/2 \rceil}(x)|/2}{deg(u)} & \text{for } k \text{ odd.} \end{cases}$$

For vertex-transitive graphs, the k-areas are the same for all vertices, hence we denote it by $A(k)$.

The motivation for our modification of the notion of k-area with the introduction of the set of neighbors inside the sphere is to have sharper density bounds than the ones obtained by the initial notion of k-area. For the square and the hexagonal lattice the notion coincide as the relation is empty. However, for the triangular lattice, the density bound is smaller: the definition of Fiala *et al.* [4] gives $A(1) = 2$ whereas $A(1) = 3$ in our case since there are adjacent vertices in the sphere (as for every $u \in \partial B_1(x)$, $|N(u) \cap \partial B_1(x)|/2 = 1$, then it adds one to the initial definition of k-area). Figure 1 illustrates this example giving a coverage of the triangular lattice by balls of radius 1. In one case (on the left) the balls are disjoint and in the second case (on the right) each sphere can be shared by several balls. Observe that in the second case, each vertex u in a sphere centered at x has two neighbors, and hence $|N(u) \cap \partial B_1(x)|/2 = 1$.

Proposition 2.2. *Let G be a vertex-transitive graph with finite degree, and i be a positive integer. If X_i is an i-packing in G, then*

$$d(X_i) \leq \frac{1}{A(i)}.$$

Proof. Observe that for arbitrary vertices x and y of an i-packing X_i, the sets $B_{\lfloor i/2 \rfloor}(x)$ and $B_{\lfloor i/2 \rfloor}(y)$ are disjoint, since the vertices x and y are at distance greater than i. Then $d(X_i) < 1/|B_{\lfloor i/2 \rfloor}(x)|$. Assume that i is an odd number and let u be a vertex at distance $\lceil i/2 \rceil$ from x, then u has $\deg(u)$ neighbors. If among these $\deg(u)$ neighbors, k neighbors are in $B_{\lfloor i/2 \rfloor}(x)$ then u can be at distance $\lceil i/2 \rceil$ from only $k/\deg(u)$ vertices in X_i. Hence $d(X_i) < 1/(|B_{\lfloor i/2 \rfloor}(x)| + \sum\limits_{u \in \partial B_{\lceil k/2 \rceil}(x)} \frac{|N(u) \cap B_{\lfloor k/2 \rfloor}(x)|}{deg(u)})$.

Moreover u and a neighbor v of u in $B_{\lfloor i/2 \rfloor}(x)$ cannot be both at distance $\lceil i/2 \rceil$ from more

than 2 vertices in X_i, therefore uv can only belong to two spheres of radius $\lceil i/2 \rceil$ centered at a vertex in X_i. Hence it follows that $d(X_i) < 1/A(i)$. □

Corollary 2.3. *Let G be a vertex-transitive graph with finite degree, and i be a positive integer. If G has a finite S-packing chromatic number, then*

$$\sum_{i=1}^{\infty} \frac{1}{A(s_i)} \geq 1.$$

Corollary 2.4. *Let G be a vertex-transitive graph with finite degree, and i be a positive integer. If G has a finite (d, n)-packing chromatic number, then*

$$\sum_{i=d}^{\infty} \frac{n}{A(i)} \geq 1.$$

An i-packing X_i is called a *maximized i-packing* if for any other i-packing $X_i^{'}$, $d(X_i) \geq d(X_i^{'})$.

2.2 Density of an i-packing in the hexagonal lattice

Proposition 2.5. *Let $\mathcal{H}$ be the hexagonal lattice, x be a vertex in $V(\mathcal{H})$ and n be a positive integer. Then*

1. $|\partial B_n(x)| = 3n$;
2. $|B_n(x)| = \frac{3}{2}n^2 + \frac{3}{2}n + 1$.

Proof. 1. As the set $\partial B_n(x)$ always contains three more vertices than $\partial B_{n-1}(x)$, then $|\partial B_n(x)| = 3n$.
2. The graph $\mathcal{H}$ is 3-regular and so $|B_1(x)| = 4$. Suppose the statement is true for n, then $|B_{n+1}(x)| = |B_n(x)| + |\partial B_{n+1}(x)| = \frac{3}{2}n^2 + \frac{3}{2}n + 1 + 3(n+1) = (\frac{3}{2}n^2 + \frac{3}{2} + 3n) + (\frac{3}{2}n + \frac{3}{2}) + 1 = \frac{3}{2}(n+1)^2 + \frac{3}{2}(n+1) + 1$ and the result follows by induction. □

Proposition 2.6. *Let $\mathcal{H}$ be the hexagonal lattice and k be a positive integer. Then*

1. $A(2k) = \frac{3}{2}k^2 + \frac{3}{2}k + 1$;
2. $A(4k+1) = 6k^2 + 6k + 2$;
3. $A(4k+3) = 6k^2 + 12k + 6$.

Proof. 1. The first property results easily from Proposition 2.5.
2. If $n = 4k+1$, then Proposition 2.5 gives $|B_{2k}(x)| = \frac{3}{2}(2k)^2 + \frac{3}{2}(2k) + 1 = 6k^2 + 3k + 1$. For every vertex y in $\partial B_{2k+1}(x)$, y has no neighbor in $\partial B_{2k+1}(x)$ other than itself, so $|N(y) \cap \partial B_{2k+1}(x)| = 0$. We have to distinguish two kinds of vertices: $3k$ vertices have two neighbors in $B_{2k}(x)$ and $|\partial B_{2k+1}(x)| - 3k = 3k + 3$ vertices have one neighbor in $B_{2k}(x)$. Therefore, $|A(4k+1)| = 6k^2 + 3k + 1 + \frac{6k}{3} + \frac{3k+3}{3} = 6k^2 + 6k + 2$.
3. If $n = 4k+3$, then Proposition 2.5 gives $|B_{2k+1}(x)| = \frac{3}{2}(2k+1)^2 + \frac{3}{2}(2k+1) + 1 = 6k^2 + 9k + 4$. For every vertex y in $\partial B_{2k+2}(x)$, y has no neighbor in $\partial B_{2k+2}(x))$ other than itself, so $|N(y) \cap \partial B_{2k+2}(x)| = 0$. We have to distinguish two kinds of vertices: $3k$ vertices have two neighbors in $B_{2k+2}(x)$ and $|\partial B_{2k+2}(x)| - 3k = 3k + 6$ vertices have one neighbor in $B_{2k}(x)$. Hence, we have $|A(4k+3)| = 6k^2 + 9k + 4 + \frac{6k}{3} + \frac{3k+6}{3} = 6k^2 + 12k + 6$. □

Note that this result appeared in the article of Goddard and Xu [8].

2.3 Density of an i-packing in the square lattice

Proposition 2.7. *Let $\mathbb{Z}^2$ be the square lattice, x be a vertex in $V(\mathbb{Z}^2)$ and n be a positive integer. Then*

1. $|\partial B_n(x)| = 4n$;
2. $|B_n(x)| = 2n^2 + 2n + 1$.

Proof. 1. As the set $\partial B_n(x)$ always contains four more vertices than $\partial B_{n-1}(x)$, then $|\partial B_n(x)| = 4n$.
2. The graph $\mathbb{Z}^2$ is 4-regular and so $|B_1(x)| = 5$. Suppose the statement is true for n, then $|B_{n+1}(x)| = |B_n(x)|+|\partial B_{n+1}(x)| = 2n^2+2n+1+4n+4 = 2(n+1)^2+6n+5-4n-2 = 2(n+1)^2 + 2(n+1) + 1$ and the result follows by induction. □

Proposition 2.8. *Let $\mathbb{Z}^2$ be the square lattice and k be a positive integer. Then*

1. $A(2k) = 2k^2 + 2k + 1$;
2. $A(2k+1) = 2k^2 + 4k + 2$.

Proof. 1. The first property results easily from Proposition 2.7.
2. If $n = 2k + 1$, then Proposition 2.7 gives $|B_k(x)| = k^2 + 2k + 1$. For every vertex y in $\partial B_{k+1}(x)$, y has no neighbor in $\partial B_{k+1}(x)$ other than itself, so $|N(y) \cap \partial B_{k+1}(x)| = 0$. We have to distinguish two kinds of vertices: $4k$ vertices have two neighbors in $B_k(x)$ and 4 vertices have one neighbor in $B_k(x)$. Hence, we have $|A(2k+1)| = 2k^2 + 2k + 1 + 2\frac{4k}{4} + \frac{4}{4} = 2k^2 + 4k + 2$. □

Note that this result appeared implicitly in the article of Fiala et al. [4].

2.4 Density of an i-packing in the triangular lattice

Proposition 2.9. *Let $\mathcal{T}$ be the triangular lattice, x be a vertex in $V(\mathcal{T})$ and n be a positive integer. Then*

1. $|\partial B_n(x)| = 6n$;
2. $|B_n(x)| = 3n^2 + 3n + 1$.

Proof. 1. As the set $\partial B_n(x)$ always contains six more vertices than $\partial B_{n-1}(x)$, then $|\partial B_n(x)| = 6n$.
2. The graph $\mathcal{T}$ is 6-regular and so $|B_1(x)| = 7$. Suppose the statement is true for n, then $|B_{n+1}(x)| = |B_n(x)|+|\partial B_{n+1}(x)| = 3n^2+3n+1+6n+6 = 3(n^2+1)+3n+1+6-3 = 3(n^2 + 1) + 3(n + 1) + 1$ and the result follows by induction. □

Proposition 2.10. *Let $\mathcal{T}$ be the triangular lattice and k be a positive integer. Then*

1. $A(2k) = 3k^2 + 3k + 1$;
2. $A(2k+1) = 3k^2 + 6k + 3$.

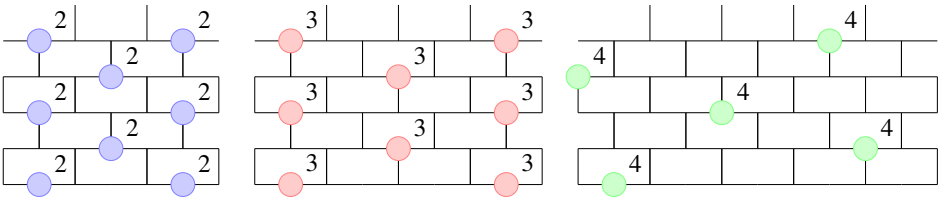


Figure 2: The sets X_2 (2-packing), X_3 (3-packing) and X_4 (4-packing) in $\mathcal{H}$.

Proof. 1. The first property result easily from Proposition 2.9.
2. If $n = 2k+1$, then Proposition 2.9 gives $|B_k(x)| = 3k^2+3k+1$. For every vertex y in $\partial B_{k+1}(x)$, y has two neighbors in $\partial B_{k+1}(x))$ other than itself, so $|N(y) \cap \partial B_{2k+1}(x)| = 2$. We have to distinguish two kinds of vertices: $6k$ vertices have two neighbors in $B_k(x)$ and six vertices have one neighbor in $B_k(x)$. Hence, we have $|A(2k+1)| = 3k^2+3k+1+\frac{6k+6}{6}+2\frac{6k}{6}+\frac{6}{6} = 3k^2+6k+3$. □

3 Subdivision of an i-packing in $\mathcal{H}$, $\mathbb{Z}^2$ and $\mathcal{T}$

3.1 Subdivision of a 2-packing in $\mathcal{H}$

Let X_2 be the (unique) maximized 2-packing in $\mathcal{H}$ represented in Figure 2. Note that $d(X_2) = 1/A(2) = 1/4$ and remark that four 2-packings form a partition of $\mathcal{H}$ if we translate X_2 three times.
The hexagonal lattice can be seen as a subgraph of the square lattice. In fact in Figure 2, $\mathcal{H}$ is represented as subgraph of the usual representation of the square lattice. In the square lattice, we can choose one vertex as the origin and all the other vertices can be nominated by a Cartesian coordinate. In every description of $\mathcal{H}$, our origin $(0,0)$ will be a vertex in the packing that we want to describe such that there is no edge between $(0,0)$ and $(0,1)$. In fact we illustrate packings with a figure in this subsection but it will not be the case after; we will use Cartesian coordinates in order to describe a packing. For example, X_2 from Figure 2 is the set of vertices: $X_2 = \{(2x+4y, x)|\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$.
In Appendix A, we recall a proposition about distance in the hexagonal lattice from Jacko and Jendrol [10]. This proposition is useful to verify that a set is an i-packing. These verifications are left to the reader in the remaining propositions.

Proposition 3.1. *Let $k > 0$ and $m > 0$ be integers. There exist:*

1. *k^2 $(3k-1)$-packings that form a partition of X_2;*
2. *$2k^2$ $(4k-1)$-packings that form a partition of X_2;*
3. *two $(3\times 2k-1)$-packings that form a partition of a $(4k-1)$-packings from Point 2;*
4. *m^2 $(3mk-1)$-packings that form a partition of a $(3k-1)$-packing from Point 1;*
5. *m^2 $(4mk-1)$-packings that form a partition of a $(4k-1)$-packing from Point 2.*

Proof. 1. Let A_k be the $(3k-1)$-packing defined by $A_k = \{(2kx+4ky, kx)|\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$. Let $\mathscr{F} = \{(2i+4j, i)|i, j \in \{0,\dots,k-1\}$ be a family of k^2 vectors. Make k^2 copies

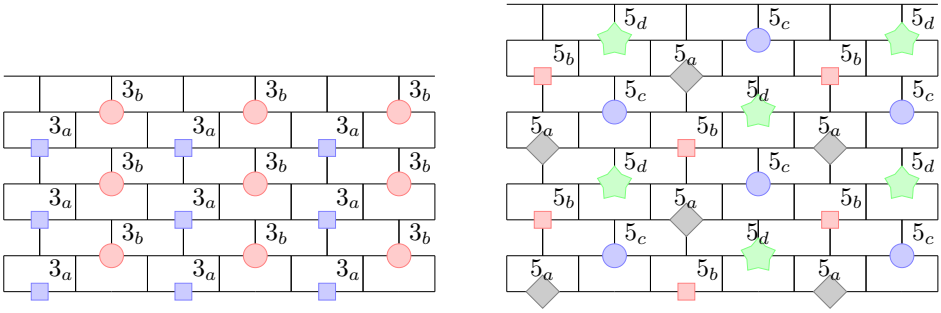


Figure 3: Two 3-packings forming a partition of X_2 (on the left) and four 5-packings forming a partition of X_2 (on the right).

of the set A_k and translate each one by a vector from $\mathscr{F}$ to obtain a partition of X_2.
2. Let B_k be the $(4k-1)$-packing defined by $B_k = \{(4kx, 2ky) |\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$. Let $\mathscr{F} = \{(4i + 2a, 2j + a) | i, j \in \{0, \ldots, k-1\},\ a \in \{0, 1\}\}$ be a family of $2k^2$ vectors. Make $2k^2$ copies of the set B_k and translate each one by a vector from $\mathscr{F}$ to obtain a partition of X_2.
3. Note that $A_{2k} \subseteq B_k$ and if A'_{2k} is A_{2k} translated by the vector $(0, 2k)$, then $A'_{2k} \cup A_{2k} = B_k$.
4. Note that $A_{mk} \subseteq A_k$. Let $\mathscr{F} = \{(2mki + 4mkj, mki) | i, j \in \{0, \ldots, m-1\}\}$ be a family of m^2 vectors. Make m^2 copies of the set A_{mk} and translate each one by a vector from $\mathscr{F}$ to obtain a partition of A_k.
5. Note that $B_{mk} \subseteq B_k$. Let $\mathscr{F} = \{(4mki, 2mkj) | i, j \in \{0, \ldots, m-1\}\}$ be a family of m^2 vectors. Make m^2 copies of the set B_{mk} and translate each one by a vector from $\mathscr{F}$ to obtain a partition of B_k. □

Figure 3 illustrates a partition of X_2 from Points 1 and 2 for $k = 1$. In the remaining of the section, the proofs of decomposition of a set X will be resumed in a table and the proofs of properties similar from those from Points 3, 4 and 5 will be left to the reader.

3.2 Subdivision of a 3-packing in $\mathcal{H}$

Let $X_3 = \{(3x + 6y, x) |\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$ be the maximized 3-packing in $\mathcal{H}$ from Figure 2. Note that $d(X_3) = 1/A(3) = 1/6$ and that six 3-packings form a partition of $\mathcal{H}$ if we translate X_3 five times.

Proposition 3.2. *Let $k > 0$ and $m > 0$ be integers. There exist:*

1. *k^2 $p_{1,k}$-packings, $p_{1,k} = (4k-1)$, that form a partition of X_3;*
2. *$3k^2$ $p_{2,k}$-packings, $p_{2,k} = (6k-1)$, that form a partition of X_3;*
3. *$8k^2$ $p_{3,k}$-packings, $p_{3,k} = (10k-1)$, that form a partition of X_3;*
4. *$24k^2$ $p_{4,k}$-packings, $p_{4,k} = (18k-1)$, that form a partition of X_3;*
5. *m^2 $p_{j,mk}$-packings that form a partition of a $p_{j,k}$-packing from Point j, for $j \in \{1, \ldots, 4\}$;*

6. *three* $(4 \times 3k - 1)$*-packings that form a partition of a* $(6k - 1)$*-packing from Point 2;*
7. *two* $(4 \times 4k - 1)$*-packings that form a partition of a* $(10k - 1)$*-packing from Point 3;*
8. *four 17-packings and six 23-packings that form a partition of every 5-packing from Point 2.*

Proof. The proof is resumed in Table B.4, this table contains: in which i-packing X will be decomposed (Column 1), the number of i-packings needed to form a partition of X (Column 2), the description of an i-packing with Cartesian coordinates (assuming x and y are integers) (Column 3) and the family of translation vectors (Column 4). We assume that if we do copies of this i-packing and we translate each one by one of these vectors. Afterward, we obtain a partition of X in i-packings. □

3.3 Subdivision of a 4-packing in $\mathcal{H}$

Let $X_4 = \{(3x + 7y, 2x + y) |\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$ be the 4-packing in $\mathcal{H}$ from Figure 2. Note that $d(X_4) = 1/11$ and that $1/A(4) = 1/10$. However, we claim that a 4-packing with density 1/10 does not exist. Note that eleven 4-packings form a partition of $\mathcal{H}$ if we translate X_4 ten times.

Proposition 3.3. *Let* $k > 0$ *and* $m > 0$ *be integers. There exist:*

1. k^2 $p_{1,k}$*-packings,* $p_{1,k} = (5k - 1)$*, that form a partition of* X_4*;*
2. $2k^2$ $p_{2,k}$*-packings,* $p_{2,k} = (6k - 1)$*, that form a partition of* X_4*;*
3. $3k^2$ $p_{3,k}$*-packings,* $p_{3,k} = (8k - 1)$*, that form a partition of* X_4*;*
4. $6k^2$ $p_{4,k}$*-packings,* $p_{4,k} = (11k - 1)$*, that form a partition of* X_4*;*
5. m^2 $p_{j,mk}$*-packings that form a partition of a* $p_{j,k}$*-packing from Point* j*, for* $j \in \{1, \ldots, 4\}$*;*
6. *two* $(5 \times 2k - 1)$*-packings that form a partition of a* $(6k - 1)$*-packing from Point 2;*
7. *two* $(6k - 1)$*-packings that form a partition of a* $(5k - 1)$*-packing from Point 1;*
8. *three* $(5 \times 3k - 1)$*-packings that form a partition of a* $(8k - 1)$*-packing from Point 2;*
9. *three* $(8k - 1)$*-packings that form a partition of a* $(5k - 1)$*-packing from Point 1.*

Proof. See Table B.5. □

3.4 Subdivision of a 2-packing in $\mathbb{Z}^2$

In the square lattice, we can choose one vertex as the origin and all the other vertices will be nominated by Cartesian coordinates. In all our representations our origin $(0, 0)$ will be in the packing that we want to describe. Let $X_2 = \{(2x + y, x + 3y) |\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$ be the maximized 2-packing in $\mathbb{Z}^2$ from Figure 4. Note that $d(X_2) = 1/A(2) = 1/5$ and that five 2-packings form a partition of $\mathbb{Z}^2$ if we translate X_2 four times.

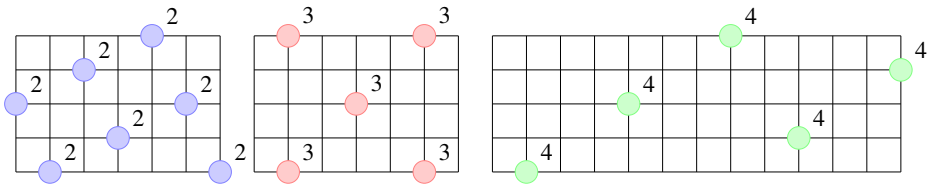


Figure 4: The sets X_2 (2-packing), X_3 (3-packing) and X_4 (4-packing) in $\mathbb{Z}^2$.

Proposition 3.4. *Let $k > 0$, $m > 0$ be integers. There exist:*

1. k^2 $(3k-1)$*-packings that form a partition of* X_2*;*
2. $2k^2$ $(4k-1)$*-packings that form a partition of* X_2*;*
3. *two* $(3 \times 2k-1)$*-packings that form a partition of a* $(4k-1)$*-packing from Point 2;*
4. *two* $(4k-1)$*-packings that form a partition of a* $(3k-1)$*-packing from Point 1;*
5. m^2 $(3mk-1)$*-packings that form a partition of a* $(3k-1)$*-packing from Point 1;*
6. m^2 $(4mk-1)$*-packings that form a partition of a* $(4k-1)$*-packing from Point 2.*

Proof. See Table B.6. □

3.5 Subdivision of a 3-packing in $\mathbb{Z}^2$

Let $X_3 = \{(2x+4y, 2y) |\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$ be the maximized 3-packing in $\mathbb{Z}^2$ from Figure 4. Note that $d(X_3) = 1/A(3) = 1/8$ and that eight 3-packings form a partition of $\mathbb{Z}^2$ if we translate X_3 seven times.

Proposition 3.5. *Let $k > 0$ and $m > 0$ be integers. There exist:*

1. k^2 $(4k-1)$*-packings that form a partition of* X_3*;*
2. m^2 $(4mk-1)$*-packings that form a partition of a* $(4k-1)$*-packing from Point 1.*

Proof. See Table B.6. □

3.6 Subdivision of a 4-packing in $\mathbb{Z}^2$

Let $X_4 = \{(3x+8y, 2x+y) |\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$ be the maximized 4-packing in $\mathbb{Z}^2$ from Figure 4. Note that $d(X_4) = 1/A(4) = 1/13$ and that thirteen 4-packings form a partition of $\mathbb{Z}^2$ if we translate X_4 twelve times.

Proposition 3.6. *Let $k > 0$, $m > 0$ be integers. There exist:*

1. k^2 $(5k-1)$*-packings that form a partition of* X_4*;*
2. $2k^2$ $(6k-1)$*-packings that form a partition of* X_4*;*
3. *two* $(5 \times 2k-1)$*-packings that form a partition of a* $(6k-1)$*-packing from Point 2;*

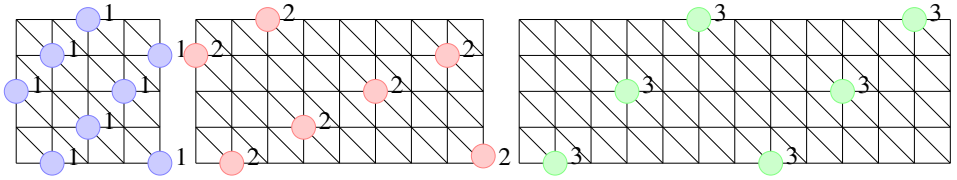


Figure 5: The sets X_1 (1-packing), X_2 (2-packing) and X_3 (3-packing) in $\mathcal{T}$.

4. *two* $(6k-1)$*-packings that form a partition of a* $(5k-1)$*-packing from Point 1;*
5. m^2 $(5mk-1)$*-packings that form a partition of a* $(5k-1)$*-packing from Point 1;*
6. m^2 $(6mk-1)$*-packings that form a partition of a* $(6k-1)$*-packing from Point 2.*

Proof. See Table B.6. □

3.7 Subdivision of an independent set in $\mathcal{T}$

The square lattice can be seen as a subgraph of the triangular lattice. In fact in Figure 5, the triangular lattice is represented as a supergraph of the square lattice. Therefore, we can choose one vertex as the origin and all the other vertices will be nominated by Cartesian coordinates. In all our representations our origin $(0,0)$ will be a vertex such that $(0,0)$ has $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$, $(-1,1)$ and $(1,-1)$ as neighbors. Let $X_1 = \{(x+3y,x)|\ x\in\mathbb{Z},\ y\in\mathbb{Z}\}$ be the (unique) maximized independent set (1-packing) in $\mathcal{T}$ from Figure 5. Note that $d(X_1) = 1/A(1) = 1/3$ and that three independent sets form a partition of $\mathcal{T}$ if we translate X_1 two times.

Proposition 3.7. *Let* $k>0$ *and* $m>0$ *be integers. There exist:*

1. k^2 $(2k-1)$*-packings that form a partition of* X_1*;*
2. $3k^2$ $(3k-1)$*-packings that form a partition of* X_1*;*
3. *three* $(3k-1)$*-packings that form a partition of a* $(2k-1)$*-packing from Point 1;*
4. m^2 $(2mk-1)$*-packings that form a partition of a* $(2k-1)$*-packing from Point 1;*
5. m^2 $(3mk-1)$*-packings that form a partition of a* $(3k-1)$*-packing from Point 2.*

Proof. See Table B.7. □

3.8 Subdivision of a 2-packing in $\mathcal{T}$

Let $X_2 = \{(2x-y,x+3y)|\ x\in\mathbb{Z},\ y\in\mathbb{Z}\}$ be the maximized 2-packing in $\mathcal{T}$ from Figure 5. Note that $d(X_2) = 1/A(2) = 1/7$ and that seven 2-packings form a partition of $\mathcal{T}$ if we translate X_2 six times.

Proposition 3.8. *Let* $k>0$ *and* $m>0$ *be integers. There exist:*

1. k^2 $(3k-1)$*-packings that form a partition of* X_2*;*
2. m^2 $(3mk-1)$*-packings that form a partition of a* $(3k-1)$*-packing from Point 1.*

Proof. See Table B.7. □

3.9 Subdivision of a 3-packing in $\mathcal{T}$

Let $X_3 = \{(2x + 6y, 2x) |\ x \in \mathbb{Z},\ y \in \mathbb{Z}\}$ be the maximized 3-packing in $\mathcal{T}$ from Figure 5. Note that $d(X_3) = 1/A(3) = 1/12$ and that twelve 3-packings form a partition of $\mathcal{T}$ if we translate X_3 eleven times.

Proposition 3.9. *Let $k > 0$ and $m > 0$ be integers. There exist:*

1. *k^2 $(4k-1)$-packings that form a partition of X_3;*
2. *$3k^2$ $(6k-1)$-packings that form a partition of X_3;*
3. *three $(4 \times 3k - 1)$-packings that form a partition of a $(6k-1)$-packing from Point 1;*
4. *m^2 $(4mk-1)$-packings that form a partition of a $(4k-1)$-packing from Point 1;*
5. *m^2 $(6mk-1)$-packings that form a partition of a $(6k-1)$-packing from Point 2.*

Proof. See Table B.7. □

4 S-packing chromatic number

4.1 General properties

In the previous section, we obtained several properties of subdivision of an i-packings in a lattice. This section illustrates general properties obtained on the S-packing chromatic number using only a small part of these properties. For a given sequence S, one can find other colorings of a lattice using properties from the previous section.

Corollary 4.1. *Let $a_0 = 1$. If $s_1 = 2$ and there exist three integers $1 < a_1 < \ldots < a_3$ and three integers $k_1, \ldots, k_3$ such that $s_{a_i} \leq 3k_i - 1$ and $a_i - a_{i-1} \geq k_i^2$ or $s_{a_i} \leq 4k_i - 1$ and $a_i - a_{i-1} \geq 2k_i^2$ for $i \in \{1, \ldots, 3\}$ then $\chi_\rho^S(\mathcal{H}) \leq a_3$.*

This corollary can be useful to find upper bounds for a given sequence. For example, if $S = (2, 2, 2, 2, \ldots)$, then taking $a_1 = 2$, $a_2 = 3$ and $a_3 = 4$, Corollary 4.1 gives us $\chi_\rho^S(\mathcal{H}) \leq 4$ (this result is in fact treated in next subsection). Similarly, for the sequence $S = (2, 3, 3, 5, 5, 5, 5, 7, 7, 7, 7, 7, 7, 7, 7, \ldots)$, then taking $a_1 = 3$, $a_2 = 7$ and $a_3 = 15$, Corollary 4.1 gives us $\chi_\rho^S(\mathcal{H}) \leq 15$. There are similar results for $s_1 = 3$ or $s_1 = 4$ using Propositions 3.3 and 3.4.
For the two remaining lattices, the two following properties are given for $\mathbb{Z}^2$ with $s_1 = 2$ and for $\mathcal{T}$ with $s_1 = 1$. There exist similar properties for $\mathbb{Z}^2$ with $s_1 = 3$ or 4 using Propositions 3.5 and 3.6 and for $\mathcal{T}$ with $s_1 = 2$ or 3 using Propositions 3.8 and 3.9.

Corollary 4.2. *Let $a_0 = 1$. If $s_1 = 2$ and there exist four integers $1 < a_1 < \ldots < a_4$ and four integers $k_1, \ldots, k_4$ such that $s_{a_i} \leq 3k_i - 1$ and $a_i - a_{i-1} \geq k_i^2$ or $s_{a_i} \leq 4k_i - 1$ and $a_i - a_{i-1} \geq 2k_i^2$ for $i \in \{1, \ldots, 4\}$ then $\chi_\rho^S(\mathbb{Z}^2) \leq a_4$.*

Corollary 4.3. *Let $a_0 = 1$. If $s_1 = 1$ and there exist two integers $1 < a_1 < a_2$ and two integers k_1 and k_2 such that $s_{a_i} \leq 2k_i - 1$ and $a_i - a_{i-1} \geq k_i^2$ or $s_{a_i} \leq 3k_i - 1$ and $a_i - a_{i-1} \geq 3k_i^2$ for $i \in \{1, \ldots, 2\}$ then $\chi_\rho^S(\mathcal{T}) \leq a_2$.*

4.2 S-packing chromatic number and distance coloring

Jacko and Jendrol [10], Fertin *et al.* [2] and Ševcıková [14] have studied distance colorings of $\mathcal{H}$, $\mathbb{Z}^2$ and $\mathcal{T}$ respectively. The following propositions comes from their work and can be translated in S-packing coloring:

Proposition 4.4 ([10])**.** *Let n and d be integers. The minimum n such that $s_1 = d$, $s_n = d$ and $\chi_\rho^S(\mathcal{H}) = n$ is given by*

$$n = \begin{cases} \lceil \frac{3}{8}(d+1)^2 \rceil & \text{for } d \text{ odd;} \\ \lceil \frac{3}{8}(d+4/3)^2 \rceil & \text{for } d \text{ even.} \end{cases}$$

Proposition 4.5 ([2])**.** *Let n and d be integers. The minimum n such that $s_1 = d$, $s_n = d$ and $\chi_\rho^S(\mathbb{Z}^2) = n$ is given by*

$$n = \begin{cases} \frac{1}{2}(d+1)^2 & \text{for } d \text{ odd;} \\ \frac{1}{2}((d+1)^2+1) & \text{for } d \text{ even.} \end{cases}$$

Proposition 4.6 ([14])**.** *Let n and d be integers. The minimum n such that $s_1 = d$, $s_n = d$ and $\chi_\rho^S(\mathcal{T}) = n$ is given by*

$$n = \left\lceil \frac{3}{4}(d+1)^2 \right\rceil.$$

5 (d, n)-packing chromatic number

5.1 Hexagonal lattice

Proposition 5.1. $\chi_\rho^{2,1}(\mathcal{H}) = \infty$, $\chi_\rho^{5,2}(\mathcal{H}) = \infty$, $\chi_\rho^{8,3}(\mathcal{H}) = \infty$, $\chi_\rho^{11,4}(\mathcal{H}) = \infty$, $\chi_\rho^{13,5}(\mathcal{H}) = \infty$ *and* $\chi_\rho^{16,6}(\mathcal{H}) = \infty$.

Proof. Let $\mathcal{H}$ be the hexagonal lattice and k be an integer, $k \geq 16$.
$\sum_{i=1}^{k} \frac{1}{A(i)} = \sum_{i=1}^{n} \frac{1}{A(2i)} + \sum_{i=0}^{n} \frac{1}{A(4i+1)} + \sum_{i=0}^{k} \frac{1}{A(4i+3)} = \sum_{i=1}^{k} \frac{1}{\frac{3}{2}i^2+\frac{3}{2}i+1} + \sum_{i=0}^{k} \frac{1}{6i^2+6i+2} + \sum_{i=0}^{k} \frac{1}{6i^2+12i+6} < \frac{2}{15}\sqrt{15}\pi \tanh(\frac{1}{6}\pi\sqrt{15}) + \frac{1}{6}\sqrt{3}\pi \tanh(\frac{1}{6}\pi\sqrt{3}) + \frac{1}{36}\pi^2 - 1 < 1.494.$
Therefore: $\sum_{i=2}^{k} \frac{1}{A(i)} < 1.494 - \frac{1}{A(1)} < 0.994 < 1$, $\sum_{i=5}^{k} \frac{2}{A(i)} < 2(1.494 - \sum_{i=1}^{4} \frac{1}{A(i)}) < 0.955 < 1$, $\sum_{i=8}^{k} \frac{3}{A(i)} < 3(1.494 - \sum_{i=1}^{7} \frac{1}{A(i)}) < 0.935 < 1$, $\sum_{i=11}^{k} \frac{4}{A(i)} < 4(1.494 - \sum_{i=1}^{10} \frac{1}{A(i)}) < 0.925 < 1$ and $\sum_{i=13}^{k} \frac{5}{A(i)} < 5(1.494 - \sum_{i=1}^{12} \frac{1}{A(i)}) < 0.986 < 1$, $\sum_{i=16}^{k} \frac{6}{A(i)} < 6(1.494 - \sum_{i=1}^{15} \frac{1}{A(i)}) < 0.968 < 1$. Corollary 2.3 allows us to conclude. □

Proposition 5.2. $\chi_\rho^{2,2}(\mathcal{H}) \leq 8$, $\chi_\rho^{2,3}(\mathcal{H}) \leq 5$ *and* $\forall n \geq 4$, $\chi_\rho^{2,n}(\mathcal{H}) = 4$.

Proof. Using Proposition 3.1, we define a $(2, n)$-packing coloring of $\mathcal{H}$ for each $n = 2, 3$ and $n \geq 4$. $\mathcal{H}$ can be partitioned into four 2-packings, the first two ones can be colored by color 2, the third one by two colors 3 and the last one by four colors under 5, it will be two 4 and two 5, to conclude $\chi_\rho^{2,2}(\mathcal{H}) \leq 8$. $\mathcal{H}$ can be partitioned into four 2-packings, the

first three ones can be colored by colors 2 and the third one by two colors 3, to conclude $\chi_\rho^{2,3}(\mathcal{H}) \leq 5$. $\mathcal{H}$ can be partitioned into four 2-packings, hence $\forall n \geq 4$, $\chi_\rho^{2,n}(\mathcal{H}) = 4$. □

The following table summarizes the colorings defined in the above proof. The symbol P in the table refers to the packings we use and how we subdivide them into i-packings (A_i is an i-packing) and the symbol C refers to the associated colors we use for each i-packing. By $k \times A_i$ we mean we use k i-packings, and by $k \times i$ we mean we use k colors i. In the rest of the paper, similar proofs will be only described by a table using the same format than this one.

(2,2)-packing	P	2×X_2	X_2	X_2
			2×A_3	4×A_5
	C	2×2	2×3	2×4, 2×5
(2,3)-packing	P	3×X_2	X_2	
			2×A_3	
	C	3×2	2×3	

Proposition 5.3. $\chi_\rho^{3,2}(\mathcal{H}) \leq 35$, $\chi_\rho^{3,3}(\mathcal{H}) \leq 13$, $\chi_\rho^{3,4}(\mathcal{H}) \leq 10$, $\chi_\rho^{3,5}(\mathcal{H}) \leq 8$ *and* $\forall n \geq 6$, $\chi_\rho^{3,n}(\mathcal{H}) = 6$.

Proof. Using Proposition 3.2, we define a $(3,n)$-packing coloring of $\mathcal{H}$ for each $n = 2, 3, 4, 5$ and $n \geq 6$. $\mathcal{H}$ can be partitioned into six 3-packings, hence $\forall n \geq 6$, $\chi_\rho^{3,n}(\mathcal{H}) = 6$. The other colorings are described in Table C.8. □

Proposition 5.4. $\chi_\rho^{4,3}(\mathcal{H}) \leq 58$, $\chi_\rho^{4,4}(\mathcal{H}) \leq 27$, $\chi_\rho^{4,5}(\mathcal{H}) \leq 21$, $\chi_\rho^{4,6}(\mathcal{H}) \leq 18$ *and from [10]* $\forall n \geq 11$, $\chi_\rho^{4,n}(\mathcal{H}) = 11$.

Proof. Using Proposition 3.3, we define a $(4,n)$-packing coloring of $\mathcal{H}$ for each $n = 3, 4, 5, 6$ and $n \geq 11$. $\mathcal{H}$ can be partitioned into eleven 4-packings. The other colorings are described in Table C.9. □

5.2 Square lattice

Proposition 5.5. $\chi_\rho^{2,1}(\mathbb{Z}^2) = \infty$, $\chi_\rho^{4,2}(\mathbb{Z}^2) = \infty$, $\chi_\rho^{6,3}(\mathbb{Z}^2) = \infty$, $\chi_\rho^{8,4}(\mathbb{Z}^2) = \infty$, $\chi_\rho^{10,5}(\mathbb{Z}^2) = \infty$ *and* $\chi_\rho^{12,6}(\mathbb{Z}^2) = \infty$.

Proof. Let $\mathbb{Z}^2$ be the square lattice and k be an integer, $k \geq 12$.
$\sum_{i=1}^{k} \frac{1}{A(i)} = \sum_{i=1}^{k} \frac{1}{A(2i)} + \sum_{i=0}^{k} \frac{1}{A(2i+1)} = \sum_{i=1}^{k} \frac{1}{2i^2+2i+1} + \sum_{i=0}^{k} \frac{1}{2i^2+4i+2} < \frac{1}{2}\pi \tanh(\frac{1}{2}\pi) + \frac{1}{12}\pi^2 - 1 < 1.264$.
Therefore: $\sum_{i=2}^{k} \frac{1}{A(i)} < 1.264 - \frac{1}{A(1)} < 0.764 < 1$, $\sum_{i=4}^{k} \frac{2}{A(i)} < 2(1.264 - \sum_{i=1}^{3} \frac{1}{A(i)}) < 0.877 < 1$, $\sum_{i=6}^{k} \frac{3}{A(i)} < 3(1.264 - \sum_{i=1}^{5} \frac{1}{A(i)}) < 0.917 < 1$, $\sum_{i=8}^{k} \frac{4}{A(i)} < 4(1.264 - \sum_{i=1}^{7} \frac{1}{A(i)}) < 0.938 < 1$, $\sum_{i=10}^{k} \frac{5}{A(i)} < 5(1.264 - \sum_{i=1}^{9} \frac{1}{A(i)}) < 0.951 < 1$ and $\sum_{i=12}^{k} \frac{6}{A(i)} < 6(1.264 - \sum_{i=1}^{11} \frac{1}{A(i)}) < 0.959 < 1$. Corollary 2.3 allows us to conclude. □

Proposition 5.6. $\chi_\rho^{2,2}(\mathbb{Z}^2) \leq 20$, $\chi_\rho^{2,3}(\mathbb{Z}^2) \leq 8$, $\chi_\rho^{2,4}(\mathbb{Z}^2) \leq 6$ *and* $\forall n \geq 5$, $\chi_\rho^{2,n}(\mathbb{Z}^2) = 5$.

1	2	1	3	1	2	1	10	1	4	1	9	1	2	1	3	1	2	1	5	1	4	1	14
7	1	5	1	6	1	3	1	2	1	3	1	8	1	5	1	4	1	3	1	2	1	3	1
1	3	1	2	1	4	1	7	1	5	1	2	1	3	1	2	1	11	1	6	1	10	1	2
4	1	9	1	3	1	2	1	3	1	6	1	4	1	7	1	3	1	2	1	3	1	5	1
1	2	1	15	1	5	1	11	1	2	1	3	1	2	1	17	1	5	1	4	1	2	1	3
6	1	3	1	2	1	3	1	4	1	14	1	5	1	3	1	2	1	3	1	7	1	8	1
1	5	1	4	1	16	1	2	1	3	1	2	1	10	1	4	1	13	1	2	1	3	1	2
3	1	2	1	3	1	6	1	5	1	7	1	3	1	2	1	3	1	9	1	5	1	4	1
1	7	1	10	1	2	1	3	1	2	1	4	1	6	1	5	1	2	1	3	1	2	1	11
2	1	3	1	5	1	4	1	8	1	3	1	2	1	3	1	7	1	4	1	6	1	3	1
1	4	1	2	1	3	1	2	1	9	1	5	1	11	1	2	1	3	1	2	1	12	1	5
3	1	6	1	13	1	7	1	3	1	2	1	3	1	4	1	8	1	5	1	3	1	2	1
1	2	1	3	1	2	1	5	1	4	1	15	1	2	1	3	1	2	1	10	1	4	1	9
8	1	5	1	4	1	3	1	2	1	3	1	7	1	5	1	6	1	3	1	2	1	3	1
1	3	1	2	1	11	1	6	1	10	1	2	1	3	1	2	1	4	1	7	1	5	1	2
4	1	7	1	3	1	2	1	3	1	5	1	4	1	9	1	3	1	2	1	3	1	6	1
1	2	1	17	1	5	1	4	1	2	1	3	1	2	1	14	1	5	1	11	1	2	1	3
5	1	3	1	2	1	3	1	7	1	8	1	6	1	3	1	2	1	3	1	4	1	15	1
1	10	1	4	1	9	1	2	1	3	1	2	1	5	1	4	1	16	1	2	1	3	1	2
3	1	2	1	3	1	12	1	5	1	4	1	3	1	2	1	3	1	6	1	5	1	7	1
1	6	1	5	1	2	1	3	1	2	1	11	1	7	1	10	1	2	1	3	1	2	1	4
2	1	3	1	7	1	4	1	6	1	3	1	2	1	3	1	5	1	4	1	8	1	3	1
1	11	1	2	1	3	1	2	1	13	1	5	1	4	1	2	1	3	1	2	1	9	1	5
3	1	4	1	8	1	5	1	3	1	2	1	3	1	6	1	12	1	7	1	3	1	2	1

Figure 6: A 24×24 pattern [13].

Proof. Using Proposition 3.4, we define a $(2, n)$-packing coloring of $\mathbb{Z}^2$ for each $n = 2, 3, 4$ and $n \geq 5$. $\mathbb{Z}^2$ can be partitioned into five 2-packings, hence $\forall n \geq 5$, $\chi_\rho^{2,n}(\mathbb{Z}^2) = 5$. The other colorings are described in Table C.10. □

Soukal and Holub [13] have proven that $\chi_\rho^{1,1}(\mathbb{Z}^2) \leq 17$, and proposed a 24×24 pattern in order to color the square lattice. Their pattern is recalled in Figure 6.

Proposition 5.7. $\chi_\rho^{3,3}(\mathbb{Z}^2) \leq 33$.

Proof. In the pattern of Figure 6, B_i denotes the set of vertices colored by i. Note that B_2 and B_3 are both 3-packings. It can be seen that $B_{16} \cup B_{17}$ form a 11-packing and that four 7-packings form a partition of B_2 or B_3. In order to color $\mathbb{Z}^2$ starting with 3, we partition B_1 into sixteen i-packings, $2 \leq i \leq 17$ (since B_1 is $\bigcup_{i=2}^{17} B_i$ translated by the vector $(1, 0)$). Let B'_i denote a copy of B_i translated by $(1, 0)$. We use two colors 3 to color B_2 and B_3, and one color i in order to color B_i for $i \in [4, 8]$. We color B'_i by one color i, for $i \in [3, 8]$ and B'_2 that is a 3-packing is colored by one color 4, one color 5, one color 6 and one color 7. We use the remaining color 8 to color B_9. We use two colors 9 in order to color B_{16}, B'_{16}, B_{17} and B'_{17}. The remaining color 9 is used to color B'_9. We use two colors i in order to color B_i and B'_i for $i \in [10, 13]$. The remaining colors 10, 11 ,12 and 13 are used to color B_{14}, B'_{14}, B_{15} and B'_{15}. □

Proposition 5.8. $\chi_\rho^{3,4}(\mathbb{Z}^2) \leq 20$, $\chi_\rho^{3,5}(\mathbb{Z}^2) \leq 17$, $\chi_\rho^{3,6}(\mathbb{Z}^2) \leq 14$ *and* $\forall n \geq 8$, $\chi_\rho^{3,n}(\mathbb{Z}^2) = 8$.

Proof. Using Proposition 3.5, we define a $(3,n)$-packing coloring of $\mathbb{Z}^2$ for each $n = 4,5,6$ and $n \geq 8$. $\mathbb{Z}^2$ can be partitioned into eight 3-packings, hence $\forall n \geq 8$, $\chi_\rho^{3,n}(\mathbb{Z}^2) = 8$. The other colorings are described in Table C.11. □

Proposition 5.9. $\chi_\rho^{4,4}(\mathbb{Z}^2) \leq 56$, $\chi_\rho^{4,5}(\mathbb{Z}^2) \leq 34$, $\chi_\rho^{4,6}(\mathbb{Z}^2) \leq 28$ *and* $\forall n \geq 13$, $\chi_\rho^{4,n}(\mathbb{Z}^2) = 13$.

Proof. Using Proposition 3.6, we define a $(4,n)$-packing coloring of $\mathbb{Z}^2$ for each $n = 4,5,6$ and $n \geq 13$. $\mathbb{Z}^2$ can be partitioned into thirteen 4-packings, hence $\forall n \geq 13$ $\chi_\rho^{4,n}(\mathbb{Z}^2) = 13$. The other colorings are described in Table C.12. □

5.3 Triangular lattice

Proposition 5.10. $\chi_\rho^{1,1}(\mathcal{T}) = \infty$, $\chi_\rho^{3,2}(\mathcal{T}) = \infty$, $\chi_\rho^{4,3}(\mathcal{T}) = \infty$, $\chi_\rho^{5,4}(\mathcal{T}) = \infty$, $\chi_\rho^{7,5}(\mathcal{T}) = \infty$ *and* $\chi_\rho^{8,6}(\mathcal{T}) = \infty$.

Proof. Let $\mathcal{T}$ be the triangular lattice and k be an integer, $k \geq 8$.
$\sum_{i=1}^{k} \frac{1}{A(i)} = \sum_{i=1}^{k} \frac{1}{A(2i)} + \sum_{i=0}^{k} \frac{1}{A(2i+1)} = \sum_{i=1}^{k} \frac{1}{3i^2+3i+1} + \sum_{i=0}^{k} \frac{1}{3i^2+6i+3} < \frac{1}{3}\sqrt{3}\pi \tanh(\frac{1}{6}\pi\sqrt{3}) + \frac{1}{18}\pi^2 - 1 < 0.854$.
Therefore: $\sum_{i=1}^{k} \frac{1}{A(i)} < 0.854 < 1$, $\sum_{i=3}^{k} \frac{2}{A(i)} < 2(0.854 - \sum_{i=1}^{2} \frac{1}{A(i)}) < 0.755 < 1$, $\sum_{i=4}^{k} \frac{3}{A(i)} < 3(0.854 - \sum_{i=1}^{3} \frac{1}{A(i)}) < 0.883 < 1$, $\sum_{i=5}^{k} \frac{4}{A(i)} < 4(0.854 - \sum_{i=1}^{4} \frac{1}{A(i)}) < 0.966 < 1$, $\sum_{i=7}^{k} \frac{5}{A(i)} < 5(0.854 - \sum_{i=1}^{6} \frac{1}{A(i)}) < 0.887 < 1$ and $\sum_{i=8}^{k} \frac{6}{A(i)} < 6(0.854 - \sum_{i=1}^{7} \frac{1}{A(i)}) < 0.940 < 1$.
Corollary 2.3 allows us to conclude. □

Proposition 5.11. $\chi_\rho^{1,2}(\mathcal{T}) \leq 6$ *and* $\forall n \geq 3$, $\chi_\rho^{1,n}(\mathcal{T}) = 3$.

Proof. Using Proposition 3.7, we define a $(1,n)$-packing coloring of $\mathcal{T}$ for each $n = 2$ and $n \geq 3$. $\mathcal{T}$ can be partitioned into three independent sets, hence $\forall n \geq 3$, $\chi_\rho^{1,n}(\mathcal{T}) = 3$. The other coloring is described in the following table.

(1,2)-packing	P	2×X_1	X_1
			4×A_3
	C	2×1	2×2, 2×3

□

Proposition 5.12. $\chi_\rho^{2,4}(\mathcal{T}) \leq 16$, $\chi_\rho^{2,5}(\mathcal{T}) \leq 13$, $\chi_\rho^{2,6}(\mathcal{T}) \leq 10$ *and* $\forall n \geq 7$, $\chi_\rho^{2,n}(\mathcal{T}) = 7$.

Proof. Using Proposition 3.8, we define a $(2,n)$-packing coloring of $\mathcal{T}$ for each $n = 4,5,6$ and $n \geq 7$. $\mathcal{T}$ can be partitioned into seven 2-packings, hence $\forall n \geq 7$, $\chi_\rho^{2,n}(\mathcal{T}) = 7$. The other colorings are described in Table C.13. □

Proposition 5.13. $\chi_\rho^{3,4}(\mathcal{T}) \leq 72$, $\chi_\rho^{3,5}(\mathcal{T}) \leq 38$, $\chi_p^{3,6}(\mathcal{T}) \leq 26$ *and* $\forall n \geq 12$, $\chi_\rho^{3,n}(\mathcal{T}) = 12$.

Proof. Using Proposition 3.9, we define a $(3, n)$-packing coloring of $\mathcal{T}$ for each $n = 4, 5, 6$ and $n \geq 12$. $\mathcal{T}$ can be partitioned into twelve 3-packings, hence $\forall n \geq 12$, $\chi_\rho^{3,n}(\mathcal{T}) = 12$. The other colorings are described in Table C.14. □

6 Conclusion

We have determined or bounded the (d, n)-packing chromatic number of three lattices $\mathcal{H}$, $\mathbb{Z}^2$ and $\mathcal{T}$ for small values of d and n. Further studies can be done with other values of d and n or improving existing values. The (d, n)-packing chromatic number can also be investigated for other lattices. As an example, we can prove, using color patterns defined in [15] for distance graphs, that for the octagonal lattice $\mathcal{O}$, *i.e* the strong product of two infinite path (which is a supergraph of $\mathcal{T}$), $\chi_\rho^{1,2}(\mathcal{O}) \leq 58$. For other finite or infinite graphs, like k-regular infinite trees, the method has to be adapted or changed since a maximized packing cannot be described as easily as those considered in this paper. Also, for each of three lattices studied, finding a sequence S such that $\chi_\rho^S = k$ and there is no S-packing k coloring where the s_1-packing is maximized could be an interesting result.

References

[1] J. Ekstein, J. Fiala, P. Holub and B. Lidický, The packing chromatic number of the square lattice is at least 12, *arXiv*:1003.2291v1 (2010).

[2] G. Fertin, E. Godard and A. Raspaud, Acyclic and k-distance coloring of the grid, *Information Processing Letters* **87** (2003), 51–58.

[3] J. Fiala and P. A. Golovach, Complexity of the packing coloring problem for trees, *Discrete Applied Mathematics* **158** (2010), 771–778.

[4] J. Fiala, S. Klavžar and B. Lidický, The packing chromatic number of infinite product graphs, *Europan Journal of Combinatorics* **30** (2009), 1101–1113.

[5] A. S. Finbow and D. F. Rall, On the packing chromatic number of some lattices, *Discrete Applied Mathematics* **158** (2010), 1224–1228.

[6] N. Gastineau, On dichotomies among the instance of the S-coloring problem, *Discrete Mathematics* **338** (2015), 1029–1041.

[7] W. Goddard, S. M. Hedetniemi, S. T. Hedetniemi, J. M. Harris and D. F. Rall, Broadcast Chromatic Numbers of Graphs, *Ars Combinatoria* **86** (2008), 33–49.

[8] W. Goddard and H. Xu, A note on S-packing colorings of lattices, *Discrete Applied Mathematics* **166** (2014), 255–262.

[9] W. Goddard, H. Xu, The S-packing chromatic number of a graph, *Discussiones Mathematicae Graph Theory* **32** (2012), 795–806.

[10] P. Jacko and S. Jendrol, Distance Coloring of the Hexagonal Lattice, *Discussiones Mathematicae Graph Theory* **25** (2005), 151–166.

[11] D. Korže and A. Vesel, On the packing chromatic number of square and hexagonal lattice, *Ars Mathematica Contemporanea* **7** (2014), 13–22.

[12] F. Kramer and H. Kramer, A survey on the distance-colouring of graphs, *Discrete Mathematics* **308** (2008), 422–426.

[13] R. Soukal and P. Holub, A Note on Packing Chromatic Number of the Square Lattice, *The Electronic Journal of Combinatorics* **17** (2010), 447–468.

[14] A. Ševcıková, Distant chromatic number of the planar graphs, *Manuscript* (2001).

[15] O. Togni, On packing Colorings of Distance Graphs, *Discrete Applied Mathematics* **167** (2014), 280–289.

A Distances in the three lattices

Definition A.1 ([10])**.** Let $v = (a, b)$ be a vertex in the hexagonal lattice. Then the type of v is

$$\tau(v) = a + b + 1 \pmod 2.$$

As $\mathcal{H} = V_0 \cup V_1$ is a bipartite graph, the type of a vertex v corresponds to the index of the set V_i to which v belongs.

Proposition A.2 ([10])**.** *Let $v_1 = (a_1, b_1)$, $v_2 = (a_2, b_2)$ be two vertices of the hexagonal lattice and assume that $b_1 \geq b_2$. Then the distance between v_1 and v_2 is*

$$d(v_1, v_2) = \begin{cases} |a_1 - a_2| + |b_1 - b_2| & \text{if } |a_1 - a_2| \geq |b_1 - b_2|; \\ 2|b_1 - b_2| - \tau(v_1) + \tau(v_2) & \text{if } |a_1 - a_2| < |b_1 - b_2|. \end{cases}$$

Example A.3. The set X_2 from Figure 2 is a 2-packing in $\mathcal{H}$.

Proof. Let x and y be integers, then
$d((2(x+1)+4y, x+1), (2x+4y, x)) = |2x+4y+2-2x-4y| + |x+1-x| = 3 > 2$ and $d((2x+4(y+1), x), (2x+4y, x)) = 4 > 2$;
let i and j be integers, then $d((2(x+i)+4(y+j), x+i), (2x+4y, x)) \geq min(d((2(x+1)+4y, x+1), (2x+4y, x)), d((2x+4(y+1), x), (2x+4y, x))) = 3$, hence X_2 is a 2-packing. □

Claim A.4. *Let $v_1 = (a_1, b_1)$ and $v_2 = (a_2, b_2)$ be two vertices of the square lattice. Then the distance between v_1 and v_2 is*

$$d(v_1, v_2) = |a_1 - a_2| + |b_1 - b_2|.$$

Example A.5. The set X_2 from Figure 4 is a 2-packing in $\mathbb{Z}^2$.

Proof. Let x and y be integers, then
$d((2(x+1)+y, x+1+3y), (2x+y, x+3y)) = |2x+y+2-2x-y| + |x+1+3y-x-3y| = 3 > 2$ and $d((2x+y+1, x-1+3(y+1)), (2x+y, x+3y)) = 4 > 2$, to conclude X_2 is a 2-packing. □

Claim A.6. *Let $v_1 = (a_1, b_1)$ and $v_2 = (a_2, b_2)$ be two vertices of the triangular lattice. Then the distance between v_1 and v_2 is*

$$d(v_1, v_2) = \begin{cases} max(|a_1 - a_2|, |b_1 - b_2|) & \text{if } ((a_1 \geq a_2) \wedge (b_1 \leq b_2)) \vee ((a_1 \leq a_2) \wedge (b_1 \geq b_2)); \\ |a_1 - a_2| + |b_1 - b_2| & \text{otherwise.} \end{cases}$$

Example A.7. The set X_1 from Figure 5 is an independent set in $\mathcal{T}$.

Proof. Let x and y be integers, then,
$d((x+1+3y, x+1), (x+3y, x)) = |x+1+3y-x-3y| + |x+1-x| = 2 > 1$ and $d((x+3(y+1), x), (x+3y, x)) = 3 > 1$, to conclude X_1 is an independent set. □

B Decomposition of an i-packing in the three lattices

i	Number of i-packings	Description of a i-packing	Family of translation vectors
$4k-1$	k^2	$\{3kx+6ky, kx)\}$	$(3i+6j, i)$ $i,j \in \{0,\dots,k-1\}$
$6k-1$	$3k^2$	$\{3kx+6ky, 3kx)\}$	$(3i+6j, 3i+2a)$ $i,j \in \{0,\dots,k-1\}$, $a \in \{0,1,2\}$
$10k-1$	$8k^2$	$\{6kx+12ky, 4kx)\}$	$(6i+12j+3b, 4i+2a+b)$ $i,j \in \{0,\dots,k-1\}$, $a \in \{0,1,2,3\}$, $b \in \{0,1\}$
$18k-1$	$24k^2$	$\{12kx+24ky, 6kx)\}$	$(12i+24j+3b, 6i+2a+b)$ $i,j \in \{0,\dots,k-1\}$, $a \in \{0,\dots,5\}$, $b \in \{0,1,2,3\}$

Table B.4: Decomposition of X_3 in $\mathcal{H}$ into i-packings.

i	Number of i-packings	Description of a i-packing	Family of translation vectors
$5k-1$	k^2	$\{3kx-ky, 2kx+3ky)\}$	$(3i-j, 2i+3j)$ $i,j \in \{0,\dots,k-1\}$
$6k-1$	$2k^2$	$\{7kx-ky, kx+3ky)\}$	$(7i+3a-j, i+2a+3j)$ $i,j \in \{0,\dots,k-1\}$, $a \in \{0,1\}$
$8k-1$	$3k^2$	$\{7kx+2ky, kx+5ky)\}$	$(7i+2j+3a, i+5j+2a)$ $i,j \in \{0,\dots,k-1\}$, $a \in \{0,1,2\}$
$11k-1$	$6k^2$	$\{-2kx+11ky, 6kx)\}$	$(-2i+11j+7a, 6i+a)$ $i,j \in \{0,\dots,k-1\}$, $a \in \{0,\dots,5\}$

Table B.5: Decomposition of X_4 in $\mathcal{H}$ into i-packings.

	i	Number of i-packings	Description of a i-packing	Family of translation vectors
X_2	$3k-1$	k^2	$\{2kx-ky, kx+2ky)\}$	$(2i-j, i+2j)$ $i,j \in \{0,\dots,k-1\}$
	$4k-1$	$2k^2$	$\{4kx+ky, 2kx+3ky)\}$	$(4i+2a+j, 2i+2a+3j)$ $i,j \in \{0,\dots,k-1\}, a \in \{0,1\}$
X_3	$4k-1$	k^2	$\{2kx+4ky, 2kx)\}$	$(2i+4j, 2i)$ $i,j \in \{0,\dots,k-1\}$
X_4	$5k-1$	k^2	$\{3kx-2ky, 2kx+3ky)\}$	$(3i-2j, 2i+3j)$ $i,j \in \{0,\dots,k-1\}$
	$6k-1$	$2k^2$	$\{6kx+ky, 4kx+5ky)\}$	$(6i+j+3a, 4i+5j+2a)$ $i,j \in \{0,\dots,k-1\}, a \in \{0,1\}$

Table B.6: Decomposition of X_2, X_3 and X_4 in $\mathbb{Z}^2$ into i-packings.

	i	Number of i-packings	Description of a i-packing	Family of translation vectors
X_1	$2k-1$	k^2	$\{kx+3ky, kx)\}$	$(i+3j, i)$ $i,j \in \{0,\dots,k-1\}$
	$3k-1$	$3k^2$	$\{3kx+3ky, 3kx)\}$	$(3i+3j+a, 3i+a)$ $i,j \in \{0,\dots,k-1\}, a \in \{0,1,2\}$
X_2	$3k-1$	k^2	$\{2kx+7ky, kx)\}$	$(2i+7j, i)$ $i,j \in \{0,\dots,k-1\}$
X_3	$4k-1$	k^2	$\{2kx+6ky, 2kx)\}$	$(2i+6j, 2i)$ $i,j \in \{0,\dots,k-1\}$
	$6k-1$	$3k^2$	$\{6kx+6ky, 6kx)\}$	$(6i+6j+2a, 6i+2a)$ $i,j \in \{0,\dots,k-1\}, a \in \{0,1,2\}$

Table B.7: Decomposition of X_1, X_2 and X_3 in $\mathcal{T}$ into i-packings.

C Decomposition and associated colors

(3,2)-packing	P	$2\times X_3$	X_3	X_3	X_3	X_3
			$3\times X_5$	$4\times X_7$	$4\times X_9$, $8\times X_{15}$	X_5, $3\times X_{11}$, $4\times X_{17}$, $6\times X_{23}$
	C	2×3	2×4, 5	2×6, 2×7	2×8, 2×9, 2×12, 2×13, 2×14, 2×15	5, 2×10, 2×11, 2×16, 2×17, 2×18, 2×19, 20
(3,3)-packing	P	$3\times X_3$	X_3	X_3	X_3	
			$3\times X_5$	$3\times X_5$	$4\times X_7$	
	C	3×3	3×4	3×5	3×6, 7	
(3,4)-packing	P	$4\times X_3$	X_3	X_3		
			$3\times X_5$	$3\times X_5$		
	C	4×3	3×4	4, 2×5		
(3,5)-packing	P	$5\times X_3$	X_3			
			$3\times X_5$			
	C	5×3	3×4			

Table C.8: Decomposition of $\mathcal{H}$ into 3-packings and associated colors.

(4,3)-packing	P	$3\times X_4$	X_4	$2\times X_4$	$2\times X_4$	X_4
			$2\times A_5$	$6\times A_7$	A_5, $6\times A_9$	$6\times A_{11}$, $4\times A_{19}$
	C	3×4	2×5	3×6, 3×7	5, 3×8, 3×9	3×10, 3×11, 18, 3×19
	P				X_4	X_4
					$9\times A_{14}$	$11\times A_{19}$, $10\times A_{23}$
	C				3×12, 3×13, 3×14	3×15, 3×16, 3×17, 2×18, 3×20, 3×21 3×22, 23
(4,4)-packing	P	$4\times X_4$	$2\times X_4$	$2\times X_4$	$2\times X_4$	X_4
			$4\times A_5$	$6\times A_7$	$8\times A_9$	$2\times A_7$, $3\times A_{14}$
	C	4×4	4×5	4×6, 2×7	4×8, 4×9	2×7, 3×10
(4,5)-packing	P	$5\times X_4$	$2\times X_4$	$3\times X_4$	X_4	
			$4\times A_5$	$9\times A_7$	A_5, $2\times A_9$	
	C	5×4	4×5	5×6, 4×7	5, 7, 8	
(4,6)-packing	P	$6\times X_4$	$3\times X_4$	$2\times X_4$		
			$6\times A_5$	$6\times A_7$		
	C	6×4	6×5	6×6		

Table C.9: Decomposition of $\mathcal{H}$ into 4-packings and associated colors.

(2,2)-packing	P	$2\times X_2$	X_2	X_2	X_2
			$2\times A_3$	$4\times A_5$	$6\times A_8$, $6\times A_{11}$
	C	2×2	2×3	2×4, 2×5	2×6, 2×7, 2×8 2×9, 2×10, 2×11
(2,3)-packing	P	$3\times X_2$	X_2	X_2	
			$2\times A_3$	A_3, $2\times A_5$	
	C	3×2	2×3	3, 2×4	
(2,4)-packing	P	$4\times X_2$	X_2		
			$2\times A_3$		
	C	4×2	2×3		

Table C.10: Decomposition of $\mathbb{Z}^2$ into 2-packings and associated colors.

(3,4)-packing	P	$4\times X_3$	$4\times X_3$
			$16\times A_7$
	C	4×3	4×4, 4×5, 4×6, 4×7
(3,5)-packing	P	$5\times X_3$	$3\times X_3$
			$12\times A_7$
	C	5×3	5×4, 5×5, 2×6
(3,6)-packing	P	$6\times X_3$	$2\times X_3$
			$8\times A_7$
	C	6×3	6×4, 2×5

Table C.11: Decomposition of $\mathbb{Z}^2$ into 3-packings and associated colors.

(4,4)-packing	P	$4\times X_4$	$2\times X_4$	$4\times X_4$	X_4	X_4	X_4
			$4\times A_5$	$16\times A_9$	$8\times A_{11}$	$9\times A_{14}$	$3\times A_{14}$, $12\times A_{17}$
	C	4×4	4×5	4×6, 4×7, 4×8, 4×9	4×10, 4×11	4×12 4×13, 14	3×14, 4×15, 4×16, 4×17
(4,5)-packing	P	$5\times X_4$	$2\times X_4$	$5\times X_4$	X_4		
			$4\times A_5$	A_5, $18\times A_9$	$2\times A_9$, $4\times A_{11}$		
	C	5×4	4×5	5, 5×6, 5×7 5×8, 3×9	2×9, 4×10		
(4,6)-packing	P	$6\times X_4$	$3\times X_4$	$4\times X_4$			
			$6\times A_5$	$16\times A_9$			
	C	6×4	6×5	6×6, 6×7, 4×8			

Table C.12: Decomposition of $\mathbb{Z}^2$ into 4-packings and associated colors.

(2,4)-packing	P	$4\times X_2$	$3\times X_2$
			$12\times A_5$
	C	4×2	4×3, 4×4, 4×5
(2,5)-packing	P	$5\times X_2$	$2\times X_2$
			$8\times A_5$
	C	5×2	5×3, 3×4
(2,6)-packing	P	$6\times X_2$	X_2
			$4\times A_5$
	C	6×2	4×3

Table C.13: Decomposition of $\mathcal{T}$ into 2-packings and associated colors.

(3,4)-packing	P	$4\times X_3$	$2\times X_3$	$2\times X_3$	$2\times X_3$
			$6\times A_5$	$8\times A_7$	$2\times A_5$, $12\times A_{11}$
	C	4×3	4×4, 2×5	4×6, 4×7	2×5, 4×9, 4×10, 4×11
	P			X_3	X_3
				$16\times A_{15}$	$4\times A_{11}$, $20\times A_{23}$
	C			4×12, 4×13 4×14, 4×15	4×8, 4×16, 4×17, 4×18, 4×19, 4×20
(3,5)-packing	P	$5\times X_3$	$3\times X_3$	$2\times X_3$	$2\times X_3$
			$9\times A_5$	$8\times A_7$	$18\times A_{11}$
	C	5×3	5×4, 4×5	5×6, 3×7	5, 2×7, 5×8, 5×9, 3×10
(3,6)-packing	P	$6\times X_3$	$4\times X_3$	$2\times X_3$	
			$12\times A_5$	$8\times A_7$	
	C	6×3	6×4, 6×5	6×6, 2×7	

Table C.14: Decomposition of $\mathcal{T}$ into 3-packings and associated colors.

Author Guidelines

Papers should be prepared in LaTeX and submitted as a PDF file.

Articles which are accepted for publication have to be prepared in LaTeX using class file amcjou.cls and bst file amcjou.bst (if you use BibTeX). These files and an example of how to use the class file can be found at

http://amc-journal.eu/index.php/amc/about/submissions#authorGuidelines

If this is not possible, please use the default LaTeX article style, but note that this may delay the publication process.

Title page. 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 available, 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. 2010 codes – see http://www.ams.org/msc.

References. References should be listed in alphabetical order by the first author's last name and formatted as it is shown below:

[1] First A. Author, Second B. Author and Third C. Author, Article title, *Journal Title* **121** (1982), 1–100.

[2] First A. Author, Book title, third ed., Publisher, New York, 1982.

[3] First A. Author and Second B. Author, Chapter in an edited book, in: First Editor, Second Editor (eds.), *Book Title*, Publisher, Amsterdam, 1999, 232–345.

Illustrations. Any illustrations included in the paper must be provided in PDF or EPS format. Make sure that you use uniform lettering and sizing of the text.

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 2015.

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

HERE'S TO JACK E. GRAVER ON HIS 80th BIRTHDAY

BRIGITTE SERVATIUS

Jack the educator. The Syracuse Mathematics Department is housed in Carnegy Library. Imagine heavy snowfall and a slim, tall figure approaching the Carnegy building, propping the heavy door open with one knee while his hands are busy shaking the snow off the Daily Orange (SU student newspaper) and the figure seemingly freezing in this position for as long as it takes to read the front page article. For SU students this is a familiar image of Jack E. Graver, always interested in students and always finding time to read their news.

Jack the administrator. A surprisingly deep dark voice resonating from the chairman's office inside Carnegie singing "Nobody knows the trouble I have seen, nobody knows my sorrows" is another image of Jack E. Graver.

Jack the researcher. The mathematician Graver is described on MathSciNet by publication in the areas of biology and other natural sciences, combinatorics, convex and discrete geometry, game theory, economics, social and behavioral sciences, geometry manifolds and cell complexes, operations research, mathematical programming, probability theory and stochastic processes.

Graver received his Ph.D. from Indiana University in 1964 under Andrew Hugh Wallace. His dissertation's title was "An Analytic Triangulation of an Arbitrary Real Analytic Variety", the field algebraic topology. It is remarkable that the 1964 paper [3] of the same title is cited more than once in this millennium. In 1966 he cashed an Erdős check for results in [8], a paper that may be called his first major work (cited more than 10 times in this millennium!). Because of this paper, written with Jim Yackel at Dartmouth (John Wesley Young Research Instructor), Jack Graver is known as a Ramsey Theorist. From 1966 to present, he teaches at Syracuse University. Why Syracuse? Because I felt at home here, he says. In 1975 he published On the foundations of linear and integer linear programming I [4], a paper that turned his name into an adjective. On Wickipedia you can read up on *Graver bases*, but if you want a more reliable source, try [9], where the relationship of Gröbner bases to Hilbert bases and Graver bases is presented. If you look for *On the foundations of linear and integer linear programming II*, you need to read Amir Fourudi's thesis, one of Jack Graver's 9 (so far) Ph.D. students - Jack himself was handicapped by his 1977-1994 chairmanship. However, despite administrative duties, his collaboration with Mark Watkins turned him to Graph Theory, with [6, 10, 7] as major contributions to the field. A fortunate diversion into Architecture produced not only a novel and unusual text [1], but awoke Jack's interest in rigidity of frameworks and he promoted matroids as a major tool in a colloquium talk that changed my life. I became his Ph.D. student, one of four simultaneously supervised by an acting chair. We did not fight over mathematical issues as Jack had all four of us work on totally disjoint topics, we merely fought for being next in line to enter the chairman's office. Combinatorial rigidity [2] (100 citations) is to this very day the focus of my research.

Meanwhile Jack has moved on to new Ph.D. students and new topics, namely, Fullerenes [5]. When asked about retirement he justifies the non-existence of plans by stating with a smile: "Normally people retire around 65, but next year will be only my 50$^{\text{th}}$ year at SU, so I have some time to think about that."

There's more to life than math—there is family, scouting, Shakespeare, gardening, fine dining, stories. Whenever you tell him a story, Jack will tell you a funnier and better one than you have ever heard. We have told but a small part of an interesting life to which we may look up to for future inspiration.

REFERENCES

[1] J. A. Baglivo and J. E. Graver, *Incidence and symmetry in design and architecture*, volume 7 of *Cambridge Urban & Architectural Studies*, Cambridge University Press, Cambridge, 1983.

[2] J. Graver, B. Servatius and H. Servatius, *Combinatorial rigidity*, volume 2 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 1993, doi:10.1090/gsm/002.

[3] J. E. Graver, An analytic triangulation of an arbitrary real analytic variety, *J. Math. Mech.* **13** (1964), 1021–1036.

[4] J. E. Graver, On the foundations of linear and integer linear programming. I, *Math. Programming* **9** (1975), 207–226.

[5] J. E. Graver and C. M. Graves, Fullerene patches. I, *Ars Math. Contemp.* **3** (2010), 109–120.

[6] J. E. Graver and M. E. Watkins, *Combinatorics with emphasis on the theory of graphs*, Springer-Verlag, New York-Berlin, 1977, graduate Texts in Mathematics, Vol. 54.

[7] J. E. Graver and M. E. Watkins, Locally finite, planar, edge-transitive graphs, *Mem. Amer. Math. Soc.* **126** (1997), vi+75, doi:10.1090/memo/0601.

[8] J. E. Graver and J. Yackel, An upper bound for Ramsey numbers, *Bull. Amer. Math. Soc.* **72** (1966), 1076–1079.

[9] B. Sturmfels, Algebraic recipes for integer programming, in: *Trends in optimization*, Amer. Math. Soc., Providence, RI, volume 61 of *Proc. Sympos. Appl. Math.*, pp. 99–113, 2004, doi:10.1090/psapm/061/2104733.

[10] M. E. Watkins and J. E. Graver, A characterization of infinite planar primitive graphs, *J. Combin. Theory Ser. B* **91** (2004), 87–104, doi:10.1016/j.jctb.2003.10.005.

MATHEMATICAL SCIENCES, WORCESTER POLYTECHNIC INSTITUTE, WORCESTER MA 01609-2280

E-mail address: bservat@wpi.edu

Printed in Slovenia by Birografika Bori d.o.o. Ljubljana