

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Alfred Kobsa

*University of California, Irvine, CA, USA*

Friedemann Mattern

*ETH Zurich, Switzerland*

John C. Mitchell

*Stanford University, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

Oscar Nierstrasz

*University of Bern, Switzerland*

C. Pandu Rangan

*Indian Institute of Technology, Madras, India*

Bernhard Steffen

*TU Dortmund University, Germany*

Madhu Sudan

*Microsoft Research, Cambridge, MA, USA*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Gerhard Weikum

*Max Planck Institute for Informatics, Saarbruecken, Germany*

Alberto Márquez Pedro Ramos  
Jorge Urrutia (Eds.)

# Computational Geometry

XIV Spanish Meeting, EGC 2011  
Dedicated to Ferran Hurtado  
on the Occasion of His 60th Birthday  
Alcalá de Henares, Spain, June 27-30, 2011  
Revised Selected Papers



Volume Editors

Alberto Márquez

Universidad de Sevilla, Escuela Técnica Superior de Ingeniería Informática  
Avenida Reina Mercedes, s/n 41012 Sevilla, Spain

E-mail: almar@us.es

Pedro Ramos

Universidad de Alcalá, Escuela Politécnica Superior  
E226, Campus Universitario, 28871 Alcalá de Henares, Madrid, Spain  
E-mail: pedro.ramos@uah.es

Jorge Urrutia

Universidad Nacional Autónoma de México, Instituto de Matemáticas  
Coyoacán, 04510 México, D.F, México  
E-mail: urrutia@matem.unam.mx

Cover illustration from

*Graph of triangulations of a convex polygon and tree of triangulations.*

Permission to reproduce this image has been obtained from Elsevier, UK.

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-642-34190-8

e-ISBN 978-3-642-34191-5

DOI 10.1007/978-3-642-34191-5

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012950874

CR Subject Classification (1998): F.2, I.3.5, G.2, E.1, F.1, C.2

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

*Typesetting:* Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

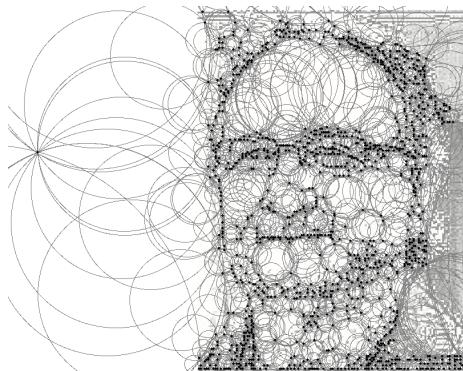
Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

# Preface

This volume contains extended versions of selected communications presented at the XIV Spanish Meeting on Computational Geometry, held in the University of Alcalá (Spain), during June 27–30, 2011.

The XIV Spanish Meeting on Computational Geometry is part of the series “Encuentros de Geometría Computacional.” The “Encuentros” have, since their start in Santander in 1990, served not only as a meeting point for computational geometers working in Spain, but also as one of the crucial contributions toward the development of a vigorous Spanish computational geometry community. The “Encuentros” made it possible for all Spanish researchers in the area to get in touch with the most relevant international figures: with the Spanish-speaking ones in the first years, and gradually with those of the entire international community.

This year, for the first time, the meeting had a fully international character, and the official language has been English. The main reason for this is that the XIV Spanish Meeting was dedicated to Prof. Ferran Hurtado, on the occasion of his 60th birthday. Professor Hurtado has played a central role in the Spanish community of computational geometry since its very beginning, and the quantity and quality of the international participants in the conference were indisputable proof of his relevance at an international level.



**Fig. 1.** This image has been provided by M. Abellanas and has been obtained with the program DEpthLAUNAY (<http://www.dma.fi.upm.es/mabellanas/delonedepth/>).

Let this volume be a token of our appreciation for Ferran’s mathematics, and the enjoyable time and friendship he has given us for many years. We are indebted to the invited speakers at our conference: Manuel Abellanas, Oswin Aichholzer, Jin Akiyama, Prosenjit Bose, Erik Demaine, Alberto Márquez, Joseph

S.B. Mitchell, János Pach, Vera Sacristán, Jorge Urrutia, all of whom were happy to join us in this celebration. We also acknowledge the enthusiastic participation of all the attendees at the XIV Spanish Meeting on Computational Geometry: their participation was essential in making this a memorable event for all of us, and in particular for Ferran. Ferran, best wishes and let us continue to enjoy friendship, mathematics and “queso, pan y vino” for many years to come!

We would like to thank all authors for their contributions to this volume. We also wish to thank the members of the Scientific Committee, for the careful revision of the papers, and the following institutions for their financial support: the spanish *Ministerio de Ciencia e Innovación*, the program *Ingenio Matemática* (i-MATH), Universidad de Alcalá, Centre de Recerca Matemàtica (CRM), Societat Catalana de Matemàtiques (SCM), Departament de Matemàtica Aplicada II (UPC), Departamento de Matemáticas (UAH), and Real Sociedad Matemática Española (RSME).

July 2012

Alberto Márquez  
Pedro Ramos  
Jorge Urrutia

# Organization

The XIV Spanish Meeting on Computational Geometry was organized by the Department of Mathematics, University of Alcalá (Spain).

## Committees

### Organizing Committee Chairs

David Orden	Universidad de Alcalá, Spain
Pedro Ramos	Universidad de Alcalá, Spain

### Scientific Committee Chairs

Alberto Márquez	Universidad de Sevilla, Spain
Pedro Ramos	Universidad de Alcalá, Spain
Jorge Urrutia	Universidad Nacional Autónoma de México

### Program Committee

Manuel Abellanas	David Rappaport
Franz Aurenhammer	Eduardo Rivera-Campo
David Avis	Günter Rote
José Miguel Díaz-Bañez	Vera Sacristán
Alfredo García	Francisco Santos
Stefan Langerman	Toni Sellàrès
Joseph O'Rourke	William Steiger
Belén Palop	David Wood

### Reviewers

Manuel Abellanas	Thomas Hackl
Greg Aloupis	Gregorio Hernández
Franz Aurenhammer	Clemens Huemer
David Avis	Elmar Langetepe
Olivier Devillers	Pat Morin
Adrian Dumitrescu	Asish Mukhopadhyay
José Miguel Díaz-Bañez	Giri Narasimhan
Ruy Fabila-Monroy	Marc Noy
Silvia Fernández-Merchant	Joseph O'Rourke
David Flores-Peñaloza	David Orden
Alfredo García	Belén Palop

## VIII Organization

Vincent Pilaud	William Steiger
David Rappaport	Javier Tejel
Eduardo Rivera-Campo	Csaba Toth
Günter Rote	Gabor F. Toth
Vera Sacristán	Masatsugu Urabe
Francisco Santos	Alexander Wolf
María Saumell	David Wood
Toni Sellarès	Mariette Yvinec

## Sponsoring Institutions

Ministerio de Ciencia e Innovación (grant MTM2010-10900-E)

Consolider Program Ingenio Matemática (i-math)

Centre de Recerca Matemàtica

Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya

Departamento de Matemáticas, Universidad de Alcalá

Societat Catalana de Matemàtiques

Universitat Politècnica de Catalunya

Universidad de Alcalá

Real Sociedad Matemática Española (RSME)

# Table of Contents

On 5-Gons and 5-Holes .....	1
<i>Oswin Aichholzer, Thomas Hackl, and Birgit Vogtenhuber</i>	
On Reversibility among Parallellohedra .....	14
<i>Jin Akiyama, Ikuro Sato, and Hyunwoo Seong</i>	
A History of Flips in Combinatorial Triangulations .....	29
<i>Prosenjit Bose and Sander Verdonschot</i>	
Tangled Thrackles .....	45
<i>János Pach, Radoš Radoičić, and Géza Tóth</i>	
Open Guard Edges and Edge Guards in Simple Polygons .....	54
<i>Csaba D. Tóth, Godfried T. Toussaint, and Andrew Winslow</i>	
String-Wrapped Rotating Disks .....	65
<i>Joseph O'Rourke</i>	
The Chromatic Number of the Convex Segment Disjointness Graph .....	79
<i>Ruy Fabila-Monroy and David R. Wood</i>	
Continuous Flattening of Convex Polyhedra .....	85
<i>Jin-ichi Itoh, Chie Nara, and Costin Vîlcu</i>	
Convexifying Monotone Polygons while Maintaining Internal Visibility .....	98
<i>Oswin Aichholzer, Mario Cetina, Ruy Fabila-Monroy, Jesús Leaños, Gelasio Salazar, and Jorge Urrutia</i>	
On the Number of Radial Orderings of Colored Planar Point Sets .....	109
<i>José Miguel Díaz-Báñez, Ruy Fabila-Monroy, and Pablo Pérez-Lantero</i>	
Notes on the Twisted Graph .....	119
<i>Elsa Omaña-Pulido and Eduardo Rivera-Campo</i>	
Locating a Service Facility and a Rapid Transit Line .....	126
<i>José Miguel Díaz-Báñez, Matias Korman, Pablo Pérez-Lantero, and Inmaculada Ventura</i>	
Simultaneously Flippable Edges in Triangulations .....	138
<i>Diane L. Souvaine, Csaba D. Tóth, and Andrew Winslow</i>	

Spiral Serpentine Polygonization of a Planar Point Set . . . . .	146
<i>Justin Iwerks and Joseph S.B. Mitchell</i>	
The 1-Center and 1-Highway Problem . . . . .	155
<i>José Miguel Díaz-Báñez, Matias Korman, Pablo Pérez-Lantero, and Inmaculada Ventura</i>	
Compact Grid Representation of Graphs . . . . .	166
<i>José Cáceres, Carmen Cortés, Clara Isabel Grima, Masahiro Hachimori, Alberto Márquez, Raiji Mukae, Atsuhiro Nakamoto, Seiya Negami, Rafael Robles, and Jesús Valenzuela</i>	
On the Heaviest Increasing or Decreasing Subsequence of a Permutation, and Paths and Matchings on Weighted Point Sets . . . . .	175
<i>Toshinori Sakai and Jorge Urrutia</i>	
A Generalization of the Source Unfolding of Convex Polyhedra . . . . .	185
<i>Erik D. Demaine and Anna Lubiw</i>	
Large Angle Crossing Drawings of Planar Graphs in Subquadratic Area . . . . .	200
<i>Patrizio Angelini, Giuseppe Di Battista, Walter Didimo, Fabrizio Frati, Seok-Hee Hong, Michael Kaufmann, Giuseppe Liotta, and Anna Lubiw</i>	
Connecting Red Cells in a Bicolour Voronoi Diagram . . . . .	210
<i>Manuel Abellanas, Antonio L. Bajuelos, Santiago Canales, Mercè Claverol, Gregorio Hernández, and Inês Matos</i>	
Covering Islands in Plane Point Sets . . . . .	220
<i>Ruy Fabila-Monroy and Clemens Huemer</i>	
Rectilinear Convex Hull with Minimum Area . . . . .	226
<i>Carlos Alegria-Galicia, Tzolkín Garduño, Areli Rosas-Navarrete, Carlos Seara, and Jorge Urrutia</i>	
Separated Matchings and Small Discrepancy Colorings . . . . .	236
<i>Viola Mészáros</i>	
A Note on the Number of Empty Triangles . . . . .	249
<i>Alfredo García</i>	
Meshes Preserving Minimum Feature Size . . . . .	258
<i>Greg Aloupis, Erik D. Demaine, Martin L. Demaine, Vida Dujmović, and John Iacono</i>	
Geometric Graphs in the Plane Lattice . . . . .	274
<i>Mikio Kano and Kazuhiro Suzuki</i>	
<b>Author Index . . . . .</b>	<b>283</b>

# On 5-Gons and 5-Holes\*

Oswin Aichholzer, Thomas Hackl, and Birgit Vogtenhuber

Institute for Software Technology, University of Technology, Graz, Austria  
`{oaiich,thackl,bvogt}@ist.tugraz.at`

**Abstract.** We consider an extension of a question of Erdős on the number of  $k$ -gons in a set of  $n$  points in the plane. Relaxing the convexity restriction we obtain results on 5-gons and 5-holes (empty 5-gons). In particular, we show a direct relation between the number of non-convex 5-gons and the rectilinear crossing number, provide an improved lower bound for the number of convex 5-holes any point set must contain, and prove that the number of general 5-holes is asymptotically maximized for point sets in convex position.

**Keywords:** discrete geometry, (empty) pentagons, Erdős-Szekeres type problem.

## Introduction

Let  $S$  be a set of  $n$  points in general position in the plane. A  $k$ -gon is a simple polygon spanned by  $k$  points of  $S$ . A  $k$ -hole is an empty  $k$ -gon, that is, a  $k$ -gon which does not contain any points of  $S$  in its interior.

Erdős [12] raised the following questions for convex  $k$ -holes and  $k$ -gons. “What is the smallest integer  $h(k)$  ( $g(k)$ ) such that any set of  $h(k)$  ( $g(k)$ ) points in the plane contains at least one convex  $k$ -hole ( $k$ -gon)?”; and more general: “What is the least number  $h_k(n)$  ( $g_k(n)$ ) of convex  $k$ -holes ( $k$ -gons) determined by any set of  $n$  points in the plane?”.

As already observed by Esther Klein, every set of 5 points determines a convex 4-hole (and thus 4-gon). Moreover, 9 points always contain a convex 5-gon and 10 points always contain a convex 5-hole, a fact proved by Harborth [16]. Only in 2007/08 Nicolás [19] and independently Gerken [15] proved that every sufficiently large point set contains a convex 6-hole, and it is well known that there exist arbitrarily large sets of points not containing any convex 7-hole [17]; see [3] for a brief survey.

In this paper we concentrate on 5-gons and 5-holes and generalize the above questions by allowing a 5-gon or 5-hole to be non-convex. Thus, when referring to a 5-gon or 5-hole, it might be convex or non-convex and we will explicitly state it when we restrict considerations to one of these two classes.

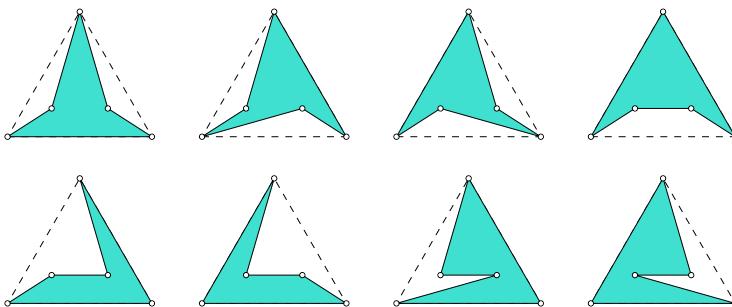
---

\* Research partially supported by the Austrian Science Fund (FWF): P23629-N18 ‘Combinatorial Problems on Geometric Graphs’, and the ESF EUROCORES programme EuroGIGA – CRP ‘ComPoSe’, Austrian Science Fund (FWF): I648-N18.

A preliminary version of this paper appeared as [6]. Similar results for 4-holes can be found in [5]. For 4-gons there is a one-to-one relation to the rectilinear crossing number of the complete graph, and thus results can be found in the respective literature (e.g. [10,13]).

## 1 Small Sets

A set of five points in convex position obviously spans precisely one convex 5-gon. In contrast, already a set of only five points (with three extremal points) can span eight different 5-gons; see Fig. 1.



**Fig. 1.** The eight different (non-convex) 5-gons spanned by a set of five points with three extremal points (fixed order type)

For small point sets, Table 1 shows the numbers of 5-gons and 5-holes, respectively. We obtained these numbers by checking all point sets (with the corresponding number of points) from the order type database [7]. Given are: the minimum number of convex 5-gons and 5-holes; the maximum number of non-convex 5-gons and 5-holes; the minimum and maximum number of (general) 5-gons and 5-holes; and, for easy comparison, the number of 5-tuples.

For counting convex 5-gons and 5-holes it is easy to see that their number is maximized by sets in convex position and gives  $\binom{n}{5}$ . Of course these sets do not contain any non-convex 5-gons or 5-holes. From Table 1 we observe that the minimum number of general 5-gons and 5-holes is  $\binom{n}{5}$  for  $5 \leq n \leq 11$ . While for 5-gons this is obviously true in general (a convex 5-tuple has exactly one polygonization, while a non-convex 5-tuple has at least four), this is not the case for 5-holes. In fact, we will show that for sufficiently large  $n$ , the convex set maximizes the number of 5-holes; see Theorem 4.

## 2 5-Gons

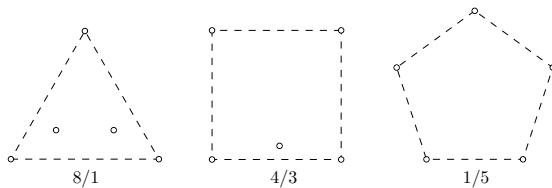
The rectilinear crossing number  $\bar{cr}(S)$  of a set  $S$  of  $n$  points in the plane is the number of proper intersections in the drawing of the complete straight line graph

**Table 1.** The number of 5-gons and 5-holes for  $n = 5 \dots 11$  points

n	numbers of 5-gons				numbers of 5-holes				$\binom{n}{5}$	
	convex		non-convex		convex		non-convex			
	min	max	min	max	min	max	min	max		
5	0		8	1	8	0		8	1	
6	0		48	6	48	0		31	6	
7	0		156	21	157	0		76	21	
8	0		408	56	410	0		157	56	
9	1		900	126	909	0		288	126	
10	2		1776	252	1790	1		492	252	
11	7		3192	462	3228	2		779	462	
								802	462	

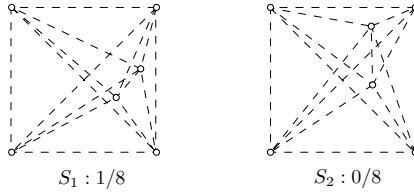
on  $S$ . It is easy to see that the number of convex 4-gons is equal to  $\bar{cr}(S)$  and is thus minimized by sets minimizing the rectilinear crossing number. This is a well known, difficult problem in discrete geometry; see [10] and [13] for details. Tight values for the minimum number of convex 4-gons are known for  $n \leq 27$  points; see e.g. [2]. Asymptotically we have at least  $c_4 \binom{n}{4} = \Theta(n^4)$  convex 4-gons, where  $c_4$  is a constant in the range  $0.379972 \leq c_4 \leq 0.380488$  [1]. As any 4 points in non-convex position span three non-convex 4-gons, we get  $3\binom{n}{4} - 3\bar{cr}(S)$  non-convex and  $3\binom{n}{4} - 2\bar{cr}(S)$  general 4-gons for a set  $S$ . Thus, sets which minimize the rectilinear crossing number also minimize the number of convex 4-gons, and maximize both the number of non-convex 4-gons and the number of general 4-gons.

Surprisingly, a similar relation can be obtained for the number of non-convex 5-gons. To see this, consider the three combinatorial different possibilities (order types) of arranging 5 points in the plane, as depicted in Fig. 2.

**Fig. 2.** The three order types for  $n = 5$ . For each set its number of different 5-gons and the number of crossings for the complete graph is shown.

**Theorem 1.** Let  $S$  be a set of  $n \geq 5$  points in the plane in general position. Then  $S$  contains  $10\binom{n}{5} - 2(n-4)\bar{cr}(S)$  non-convex 5-gons.

*Proof.* We denote with  $o_3(S)$ ,  $o_4(S)$ , and  $o_5(S)$  the number of 5-tuples of points with 3, 4, and 5, respectively, points on their convex hull. Summing over all such sets we get  $o_3(S) + o_4(S) + o_5(S) = \binom{n}{5}$ .



**Fig. 3.** Two point sets for  $n = 6$ , both with crossing number eight. One contains a convex 5-gon, the other one does not.

Note that every four points spanning a crossing pair of edges in  $S$  show up in  $(n - 4)$  5-tuples of points in  $S$ . Using the number of crossings for each order type from Fig. 2 we get  $\bar{cr}(S) = \frac{o_3(S) + 3o_4(S) + 5o_5(S)}{n-4}$ .

Considering the numbers of different 5-gons given in Fig. 2, we see that the total number of non-convex 5-gons in  $S$  is  $8o_3(S) + 4o_4(S)$ . Using these three equations, it is straight forward to obtain the following relation for the number of non-convex 5-gons in  $S$ :  $8o_3(S) + 4o_4(S) = 10\binom{n}{5} - 2(n - 4)\bar{cr}(S)$ .  $\square$

Taking the constant  $c_4$  for the rectilinear crossing number into account, we see that asymptotically we can have up to  $10\binom{n}{5} - 2(n - 4)c_4\binom{n}{4} = 10(1 - c_4)\binom{n}{5}$  non-convex 5-gons. This number is obtained for point sets minimizing the rectilinear crossing number and by a factor  $\approx 6.2$  larger than the maximum number of convex 5-gons.

For the number of convex 5-gons, no simple relation to the rectilinear crossing number is possible: There exist two different sets (order types)  $S_1$  and  $S_2$ , both of cardinality 6 with 4 extremal points, with  $\bar{cr}(S_1) = \bar{cr}(S_2) = 8$ , where  $S_1$  contains one convex 5-gon, while  $S_2$  does not contain any convex 5-gon; see Fig. 3.

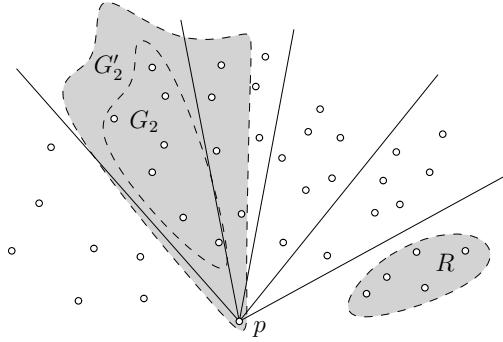
### 3 5-Holes

Recall that a 5-hole is a 5-gon which does not contain any points from the underlying set in its interior.

#### 3.1 An Improved Lower Bound for the Number of Convex 5-Holes

Let  $h_5(S)$  denote the number of convex 5-holes of a point set  $S$ , and let  $h_5(n) = \min_{|S|=n} h_5(S)$  be the number of convex 5-holes any point set of cardinality  $n$  has to have. The best upper bound  $h_5(n) \leq 1.0207n^2 + o(n^2)$  can be found in [9]. Although  $h_5(n)$  is conjectured to be quadratic in the size of  $S$ , to this date not even a super-linear lower bound exists. For quite some time, the best published lower bound was  $h_5(n) \geq \lfloor \frac{n-4}{6} \rfloor$ , obtained by Bárány and Károlyi [8]. García [14] improves this bound to  $h_5(n) \geq \frac{2}{9}n - O(1)$ . In the proceedings version [6] of the paper at hand, we presented a slightly better bound, showing  $h_5(n) \geq 3\lfloor \frac{n-4}{8} \rfloor$ .

The following theorem further improves the lower bound for  $h_5(n)$ , but still remains linear in  $n$ . It is based on an idea of Clemens Huemer [18].



**Fig. 4.** Splitting  $S_i$  into groups  $G_l$  of seven points each, plus a remainder  $R$  of at least four points

**Theorem 2.** *Any set of  $n$  points in the plane in general position contains at least  $h_5(n) \geq \lceil \frac{3}{7}(n - 11) \rceil$  convex 5-holes.*

*Proof.* Consider an arbitrary set  $S$  of  $n$  points. Assume that there is an extreme point  $p \in S$  which is incident to (at least) one convex 5-hole spanned by  $S$ . We count these convex 5-holes (solely) for  $p$ , remove  $p$  from  $S$ , and continue with  $S_1 = S \setminus \{p\}$ . Assume further that we can repeat this  $i \geq 0$  times. This way we count (at least)  $i$  different convex 5-holes, and obtain a point set  $S_i$  of cardinality  $|S_i| = n - i$ , for which all extreme points of  $S_i$  are not incident to any convex 5-hole.

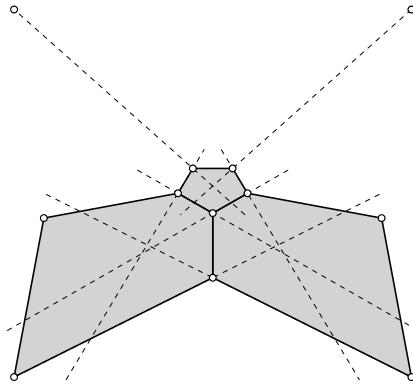
Now take any extreme point  $p \in S_i$ . Sort all other points of  $S_i$  radially around  $p$  (such that its neighbours on the convex hull  $\text{CH}(S_i)$  are the first point  $p_1$  and the last point  $p_{n-i-1}$  in the sorting, respectively). Split the sorted set  $S_i \setminus \{p\}$  into consecutive groups  $G_l$ , for  $1 \leq l \leq \lfloor \frac{n-i-5}{7} \rfloor$ , of seven points each, such that the remainder  $R$  contains at least four points; see Fig. 4. Then every group  $G_l$  together with the sorting anchor  $p$  and the first four points of  $G_{l+1}$  (or  $R$ , respectively) gives a set  $G'_l \subset S_i$  of 12 points.

We know by Dehnhardt [11] that every set of 12 points, and thus also every set  $G'_l$ , contains at least 3 convex 5-holes. As  $p$  is not incident to any convex 5-hole, all convex 5-holes in any set  $G'_l$  must be incident to at least one point of its underlying set  $G_l$  and can thus be counted (solely) for  $G_l$ . As  $R$  must have at least four points, there are exactly  $\lfloor \frac{n-i-1-4}{7} \rfloor$  groups  $G_l$ , and at least three times that many convex 5-holes in  $S_i$ . Adding the convex 5-holes we counted for points of  $S \setminus S_i$ , of which there are at least  $i$ , we obtain a lower bound of

$$\begin{aligned} i + 3 \left\lfloor \frac{n-i-5}{7} \right\rfloor &\geq i + 3 \frac{n-i-5-6}{7} \\ &= \frac{3n+4i-33}{7} \end{aligned}$$

for the total number of convex 5-holes in  $S$ . This term is minimized for  $i = 0$ , which leads to a lower bound of  $\lceil \frac{3}{7}(n - 11) \rceil$  for the minimum number  $h_5(n)$  of convex 5-holes in any set of  $n$  points.  $\square$

In the above proof we used a result by Dehnhardt [11], stating that every set of 12 points contains at least three convex 5-holes. In fact, Dehnhardt showed  $3 \leq h_5(12) \leq 4$ , and conjectured that  $h_5(12) = 4$ . Using an extension of the order type database [7], we found point sets of 12 points that contain only three convex 5-holes, disproving Dehnhardt's conjecture and thus settling  $h_5(12) = 3$ . A point set attaining this lower bound is shown in Fig. 5. Note, that this point set has 4 extreme points. This answers the question of Dehnhardt (in [11]), whether there exist sets of  $n$  points which minimize  $h_5(n)$ , whose convex hull is not a triangle.



**Fig. 5.** A set of 12 points containing only three convex 5-holes, implying  $h_5(12) = 3$

Note that on the one hand, for  $n \leq 17$  the best known lower bound is still only  $h_5(n) \geq 3$ . On the other hand, from the examples we found so far it follows that  $h_5(13) \leq 4$ ,  $h_5(14) \leq 6$ , and  $h_5(15) \leq 9$ ; see [21] for point sets attaining these bounds.

*Remark (added during revision):* Valtr [20] recently presented a bound of  $h_5(n) \geq \frac{n}{2} - O(1)$ . In a forthcoming paper (by Aichholzer, Fabila-Monroy, Hackl, Huemer, Pilz, and Vogtenhuber) this bound is further improved to  $h_5(n) \geq \frac{3}{4}n - o(n)$ .

### 3.2 A Lower Bound for the Number of (general) 5-Holes

For  $n = 10$  we obtained the following observation for general 5-holes by checking all 14 309 547 according point sets from the order type database [7].

**Observation 1.** *Let  $S$  be a set of  $n = 10$  points in the plane in general position, and  $p_1, p_2 \in S$  two arbitrary points of  $S$ . Then  $S$  contains at least 34 5-holes which have  $p_1$  and  $p_2$  among their vertices.*

Based on this simple observation we derive the following lower bound for the number of 5-holes, following the lines of a similar proof for the number of 4-holes in [5, Theorem 5].

**Theorem 3.** Let  $S$  be a set of  $n \geq 10$  points in the plane in general position. Then  $S$  contains at least  $17n^2 - O(n)$  5-holes.

*Proof.* We consider the point set  $S$  in  $x$ -sorted order,  $S = \{p_1, \dots, p_n\}$ , and sets  $S_{i,j} = \{p_i, \dots, p_j\} \subseteq S$ . The number of sets  $S_{i,j}$  having at least 10 points is

$$\sum_{i=1}^{n-9} \sum_{j=i+9}^n 1 = \frac{n^2}{2} - O(n)$$

For each  $S_{i,j}$  consider the eight points of  $S_{i,j} \setminus \{p_i, p_j\}$  which are closest to the segment  $p_ip_j$  to obtain a set of 10 points, including  $p_i$  and  $p_j$ . By Observation 1, each such set contains at least 34 5-holes which have  $p_i$  and  $p_j$  among their vertices. Moreover, as  $p_i$  and  $p_j$  are the left- and rightmost point of  $S_{i,j}$ , they are also the left- and rightmost point for each of these 5-holes. This implies that any 5-hole of  $S$  counts for at most one set  $S_{i,j}$ , which gives a lower bound of  $17n^2 - O(n)$  for the number of 5-holes in  $S$ .  $\square$

### 3.3 Maximizing the Number of (general) 5-Holes

The results for small sets shown in Table 1 suggest that the number of (general) 5-holes is minimized by sets in convex position. In this section we will not only show that this is in fact not the case, but rather prove the contrary: For sufficiently large  $n$ , sets in convex position maximize the number of 5-holes.

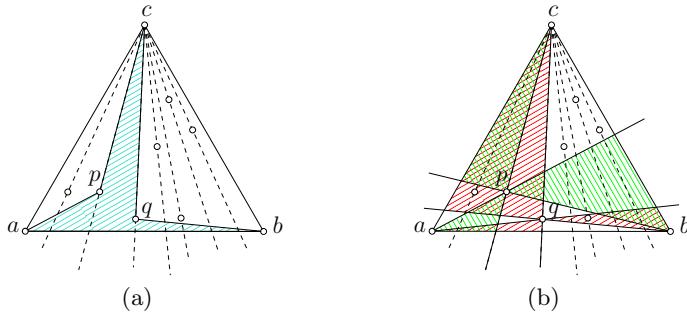
**Lemma 1.** A point set  $S$  with triangular convex hull and  $i$  interior points contains at most  $(4i+5)$  5-holes which have the three extreme points among their vertices.

*Proof.* Let  $\Delta$  be the convex hull of  $S$ ,  $a$ ,  $b$ , and  $c$  the three extreme points of  $S$  (in counterclockwise order), and  $I = S \setminus \{a, b, c\}$  the set of inner points of  $S$ ,  $|I| = i$ . As all 5-holes we consider have  $a$ ,  $b$ , and  $c$  among their vertices, they contain either one or two edges of  $\Delta$ .

First, we derive an upper bound for the number of 5-holes that contain only one edge of  $\Delta$ . If two points  $p, q \in I$  form a 5-hole that contains only the edge  $ab$  of  $\Delta$ , they have to be neighboured in a circular order of  $I$  around  $c$ ; see Fig. 6(a).

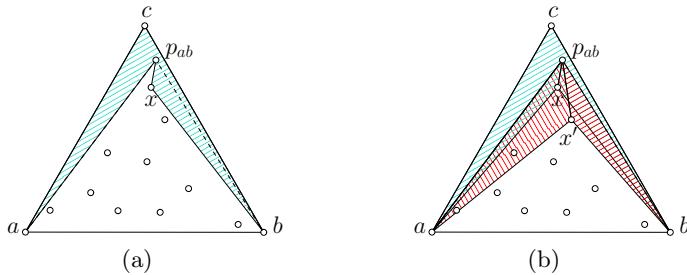
Let  $p$  be before  $q$  in the counterclockwise order around  $c$ . We say that  $p$  starts a 5-hole (with  $ab$ ). Note that  $q$  is uniquely defined by  $p$  and  $ab$ , and that the triangular area bounded by the supporting lines of  $cp$ ,  $ap$ , and  $ab$  must not contain any points of  $I$ .

Assume that  $p$  starts a 5-hole with each edge of  $\Delta$ , implying that the according areas for all three edges of  $\Delta$  have to be empty; see Fig. 6(b). Then any other point  $q \in I$  can start 5-holes with at most two edges of  $\Delta$ , as  $p$  lies in one of the three areas that would have to be empty for  $q$ ; see again Fig. 6(b). Using this fact, we conclude that at most one point of  $I$  might start three such 5-holes and all other inner points start at most two such 5-holes. This gives a total of at most  $(2i+1)$  5-holes that contain only one edge of  $\Delta$ .



**Fig. 6.** (a) A 5-hole containing only the edge  $ab$  of  $\Delta$ . (b) Shaded areas have to be empty if  $p$  (or  $q$ , respectively) starts a 5-hole with each edge of  $\Delta$ .

Second, we consider 5-holes that contain two edges of  $\Delta$  where one of the two vertices of  $I$  is reflex and the other is convex. Assume that there exists such a 5-hole without the edge  $ab$ , and with  $p_{ab}$  as reflex vertex; see Fig. 7(a).



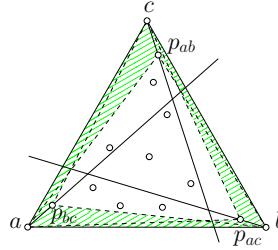
**Fig. 7.** (a) A 5-hole  $ap_{ab}cbc$  containing two edges of  $\Delta$ . (b) Only one point of  $I$  can span two 5-holes for  $ab$ .

Then the non-convex quadrilateral  $ap_{ab}bc$  must not contain any points of  $I$ , which implies that all other such 5-holes without the edge  $ab$  have  $p_{ab}$  as reflex vertex as well. Let  $x$  be the convex vertex in a 5-hole without  $ab$ . We say that  $x$  spans the 5-hole (for  $ab$ ).

Note that a point  $x$  might span two 5-holes for  $ab$ , namely  $axp_{ab}bc$  and  $ap_{ab}xbc$ . But this situation can happen for at most one point  $x$ , as all other points have to lie inside the triangle  $axb$  and thus for each of them,  $x$  lies inside exactly one of the two according possible 5-holes; see Fig. 7(b).

Now assume that for every edge  $e$  of  $\Delta$ , there exist 5-holes skipping (solely)  $e$ . Then for every edge  $e$  there is one unique point  $p_e \in I$  that is the single reflex vertex in all 5-holes for  $e$ , and each non-convex quadrilateral spanned by  $\Delta \setminus \{e\}$  and  $p_e$  is empty; see Fig. 8.

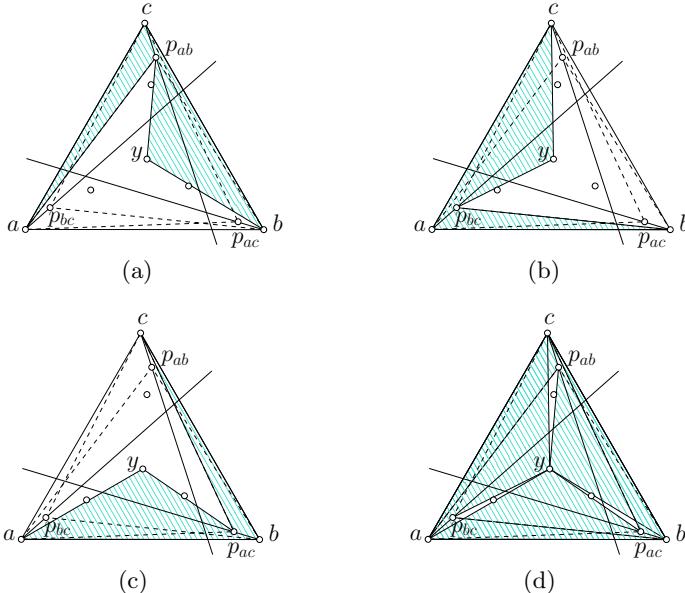
Assume further that there is a point  $y$ , that spans a 5-hole for each edge  $e$  of  $\Delta$ . Note that if a point  $x$  lies below the supporting line of  $ap_{bc}$ , then the 5-gon  $axp_{ab}bc$  contains  $p_{bc}$ . Accordingly, if  $x$  lies below the supporting line of  $bp_{ac}$ ,



**Fig. 8.** If for each combination of two sides of  $\Delta$  there is a 5-hole where one vertex of  $I$  is convex, then the shaded area must be empty

then  $p_{ac}$  lies inside  $ap_{ab}xbc$ . Thus, no point inside the triangle formed by the supporting lines of  $ap_{bc}$ ,  $bp_{ac}$ , and  $ab$  can span a 5-hole for  $ab$  because any such 5-gon contains either  $p_{bc}$  or  $p_{ac}$ . As similar statements hold for the other edges of  $\Delta$  as well,  $y$  has to lie outside all these triangles, and thus inside the triangle formed by the supporting lines of  $ap_{bc}$ ,  $bp_{ac}$ , and  $cp_{ab}$ .

Note that  $y$  can span only one 5-hole for each side, as for each reflex point there is a line  $l$  supporting one of the segments  $cp_{ab}$ ,  $ap_{bc}$ , or  $bp_{ac}$  such that  $y$  and the reflex point lie on opposite sides of  $l$ . (Recall that the shaded area in Fig. 8 must be empty of points from  $S$  and that  $y$  lies inside the triangle formed by the supporting lines of  $ap_{bc}$ ,  $bp_{ac}$ , and  $cp_{ab}$ .)

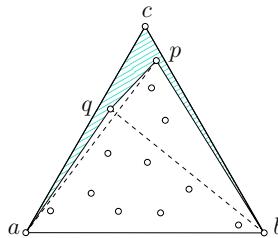


**Fig. 9.** Three 5-holes spanned by  $y$ , each one leaving out a different side of  $\Delta$

Fig. 9 shows the three possible 5-holes spanned by  $y$  both separately and altogether. As by assumption, the whole shaded area in Fig. 9(d) does not contain any points of  $I$ , all other points must be located in the non-shaded wedges.

Now, if a point lies in the wedge from  $y$  towards  $p_{ab}$ , then it cannot span a 5-hole for  $ac$ , as  $y$  lies inside one candidate and  $p_{ab}$  lies inside the other. Accordingly, a point in the wedge from  $y$  to  $p_{bc}$  cannot span a 5-hole for  $ab$ , and a point in the wedge from  $y$  to  $p_{ac}$  cannot span a 5-hole for  $bc$ . Thus, at most one point might span a 5-hole for  $e$ , for each edge  $e$  of  $\Delta$ . We obtain an upper bound of  $(2i+4)$  for the number of 5-holes that contain two edges of  $\Delta$  where one of the two vertices of  $I$  is reflex and the other is convex: at most two such 5-holes per point of  $I$ , plus one for the special point spanning a 5-hole for each edge of  $\Delta$ , plus one additional 5-hole per edge of  $\Delta$ .

Finally, consider 5-holes that contain two edges of  $\Delta$ , where the two additional vertices are both reflex, like the one shown in Fig. 10.



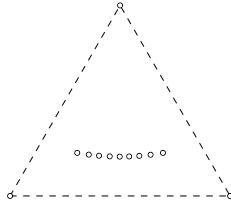
**Fig. 10.** Remaining points of  $I$  have to be located in the white areas

There is at most one such 5-hole per non-used side of  $\Delta$ . Moreover, the existence of such a 5-hole for an edge  $e$  of  $\Delta$  implies that there is no 5-hole for  $e$  where one of the vertices of  $I$  is convex. Thus, the upper bound for all 5-holes using two edges of  $\Delta$  (with and without a point of  $I$  being convex) is still  $(2i+4)$ . Hence, together with the  $(2i+1)$  5-holes using one edge of  $\Delta$  we obtain an upper bound for the total number of 5-holes of  $(4i+5)$ .  $\square$

Note that the upper bound from Lemma 1 is most likely not tight. The best example we found so far is depicted in Fig. 11. It spans  $3i+2$  (non-convex) 5-holes (of all eight types indicated in Fig. 1), where  $i$  is the number of points inside the triangle.

**Lemma 2.** *Let  $\Gamma$  be a non-empty convex quadrilateral in  $S$ . There are at most four (non-convex) 5-holes spanned by the four vertices of  $\Gamma$  plus a point of  $S$  in the interior of  $\Gamma$ .*

*Proof.* Let  $p_1, \dots, p_4$  be the vertices of  $\Gamma$ . Observe that any non-convex 5-hole has to use three edges of  $\Gamma$ . Thus there are four choices for the unused edge of  $\Gamma$ , and for each choice there is at most one way to complete the three used edges of  $\Gamma$  to a 5-hole. Assume to the contrary that two different 5-holes avoid the edge  $p_1p_2$  and use points  $q_1$  and  $q_2$ , respectively, in the interior. Then  $q_2$  lies in the triangle formed by  $p_1p_2q_1$ . But then  $q_1$  must lie inside the polygon  $p_1q_2p_2p_3p_4$ , a contradiction.  $\square$



**Fig. 11.** A point set where the extreme triangle spans  $3i+2$  (non-convex) 5-holes

Taking into account the size of the convex hull of each 5-tuple, these two lemmas lead to the following theorem.

**Theorem 4.** *For  $n \geq 86$  the number of 5-holes is maximized by a set of  $n$  points in convex position.*

*Proof.* In the following we assign every non-convex 5-tuple to the (three or four) vertices of its convex hull, and call this convex hull the *representing triangle* (or quadrilateral) of the potential non-convex 5-holes.

From Lemma 1 we know that a non-empty triangle  $\Delta$  with  $i > 0$  interior points represents at most  $4i+5$  non-convex 5-holes. In addition, each of the  $o = n - 3 - i$  points outside  $\Delta$  might form a convex quadrilateral  $\Gamma$  with  $\Delta$ . According to Lemma 2, each such  $\Gamma$  represents at most 4 non-convex 5-holes. Thus, altogether we obtain (1) as an upper bound for the number of non-convex 5-holes which have the vertices of  $\Delta$  on their convex hull.

$$4o + 4i + 5 = 4n - 7 \quad (1)$$

Note that if a (convex) quadrilateral is non-empty, then its vertices form at least one triangle which is non-empty as well. Thus, summing (1) for all non-empty triangles, we obtain an upper bound on the number of non-convex 5-holes.

Considering convex 5-holes, observe that every 5-tuple gives at most one convex 5-hole. Denote with  $N$  the number of 5-tuples that do *not* form a convex 5-hole, and with  $T$  the number of non-empty triangles. Then we get (2) as a first upper bound on the number of (general) 5-holes of a point set.

$$\binom{n}{5} - N + (4n - 7) \cdot T \quad (2)$$

To obtain an improved upper bound from (2), we need to derive a good lower bound for  $N$ . For this, consider again a non-empty triangle  $\Delta$ . As  $\Delta$  is not empty, each of the  $\binom{n-3}{2}$  5-tuples that contain all three vertices of  $\Delta$  is either not convex or not empty. On the other hand, for such a 5-tuple, all of its  $\binom{5}{3}$  contained triangles might be non-empty. Thus, we obtain  $T \binom{n-3}{2} / \binom{5}{3}$  as a lower bound for  $N$ , and (3) as an upper bound for the number of 5-holes.

$$\binom{n}{5} + \left( 4n - 7 - \frac{\binom{n-3}{2}}{\binom{5}{3}} \right) \cdot T \quad (3)$$

For  $n \geq 86$  this is at most  $\binom{n}{5}$ , the number of 5-holes for a set of  $n$  points in convex position, which proves the theorem.  $\square$

Examples show that at least for  $n \leq 16$  the number of general 5-holes is not maximized by convex sets. Hence, the truth for the lower bound in Theorem 4 of the cardinality  $n$  of the point sets lies somewhere in the range from 17 to 86.

## 4 Conclusion

In this paper we presented several results for a variant of a classical Erdős-Szekeres type problem for the case of 5-gons and 5-holes.

During the preparation of the full version of this paper we have been able to extend some of the presented results to  $k$ -gons and  $k$ -holes for  $k > 5$ . A preliminary version of these results has been presented at [4]. The thesis [21] summarizes all obtained results for  $k \geq 4$ .

Several questions remain unsettled, among which we specifically want to mention the following. Is there a super-linear lower bound for the number of convex 5-holes (cf. Theorem 2)? And does there exist a super-quadratic lower bound for the number of general 5-holes (cf. Theorem 3)?

**Acknowledgments.** We thank Clemens Huemer for helpful discussions, especially concerning the proof of Theorem 2, and Alexander Pilz for the drawing of Fig. 5.

## References

1. Ábrego, B.M., Fernández-Merchant, S., Leaños, J., Salazar, G.: A central approach to bound the number of crossings in a generalized configuration. *Electronic Notes in Discrete Mathematics* 30, 273–278 (2008)
2. Aichholzer, O.: On the rectilinear crossing number, <http://www.ist.tugraz.at/aichholzer/research/rp/triangulations/crossing/>
3. Aichholzer, O.: [Empty] [colored]  $k$ -gons - Recent results on some Erdős-Szekeres type problems. In: Proc. XIII Encuentros de Geometría Computacional, ECG 2009, Zaragoza, Spain, pp. 43–52 (2009)
4. Aichholzer, O., Fabila-Monroy, R., González-Aguilar, H., Hackl, T., Heredia, M.A., Huemer, C., Urrutia, J., Valtr, P., Vogtenhuber, B.: On  $k$ -gons and  $k$ -holes in point sets. In: Proc. 23th Canadian Conference on Computational Geometry, CCCG 2011, Toronto, Canada, pp. 21–26 (2011)
5. Aichholzer, O., Fabila-Monroy, R., González-Aguilar, H., Hackl, T., Heredia, M.A., Huemer, C., Urrutia, J., Vogtenhuber, B.: 4-holes in point sets. In: Proc. 27th European Workshop on Computational Geometry, EuroCG 2011, Morschach, Switzerland, pp. 115–118 (2011)
6. Aichholzer, O., Hackl, T., Vogtenhuber, B.: On 5-holes and 5-gons. In: Proc. XIV Encuentros de Geometría Computacional ECG2011, Alcalá de Henares, Spain, pp. 7–10 (2011)
7. Aichholzer, O., Krasser, H.: The point set order type data base: A collection of applications and results. In: Proc. 13th Canadian Conference on Computational Geometry, CCCG 2001, vol. 13, pp. 17–20 (2001)

8. Bárány, I., Károlyi, G.: Problems and Results around the Erdős-Szekeres Convex Polygon Theorem. In: Akiyama, J., Kano, M., Urabe, M. (eds.) JCDCG 2000. LNCS, vol. 2098, pp. 91–105. Springer, Heidelberg (2001)
9. Bárány, I., Valtr, P.: Planar point sets with a small number of empty convex polygons. *Studia Scientiarum Mathematicarum Hungarica* 41(2), 243–266 (2004)
10. Brass, P., Moser, W., Pach, J.: Research Problems in Discrete Geometry. Springer (2005)
11. Dehnhardt, K.: Leere konvexe Vielecke in ebenen Punktmengen. Ph.D. thesis, TU Braunschweig, Germany (1987)
12. Erdős, P.: Some more problems on elementary geometry. *Australian Mathematical Society Gazette* 5, 52–54 (1978)
13. Erdős, P., Guy, R.: Crossing number problems. *The American Mathematical Monthly* 88, 52–58 (1973)
14. García, A.: A Note on the Number of Empty Triangles. In: Marquez, A., Ramos, P., Urrutia, J. (eds.) EGC 2011 (Hurtado Festschrift). LNCS, vol. 7579, pp. 249–257. Springer, Heidelberg (2012)
15. Gerken, T.: Empty convex hexagons in planar point sets. *Discrete and Computational Geometry* 39(1-3), 239–272 (2008)
16. Harborth, H.: Konvexe Fünfecke in ebenen Punktmengen. *Elemente der Mathematik* 33, 116–118 (1978)
17. Horton, J.: Sets with no empty convex 7-gons. *Canadian Mathematical Bulletin* 26(4), 482–484 (1983)
18. Huemer, C.: Personal communication (2011)
19. Nicolás, C.: The empty hexagon theorem. *Discrete and Computational Geometry* 38(2), 389–397 (2007)
20. Valtr, P.: On empty pentagons and hexagons in planar point sets. In: Proc. 18th Computing: Australasian Theory Symposium, CATS 2012, Melbourne, Australia, pp. 47–48 (2012)
21. Vogtenhuber, B.: Combinatorial Aspects of [Colored] Point Sets in the Plane. Ph.D. thesis, Institute for Software Technology, Graz University of Technology, Graz, Austria (2011)

# On Reversibility among Parallellohedra

Jin Akiyama<sup>1</sup>, Ikuro Sato<sup>2</sup>, and Hyunwoo Seong<sup>3</sup>

<sup>1</sup> Research Institute for Mathematics Education, Tokyo University of Science,  
1-3 Kagurazaka, Shinjuku, Tokyo 162-8601, Japan  
[ja@jin-akiyama.com](mailto:ja@jin-akiyama.com)

<sup>2</sup> Department of Pathology, Research Institute, Miyagi Cancer Center,  
47-1 Medeshima-shiote (Azanodayama), Natori-city, Miyagi 981-1293, Japan  
[sato-ik510@miyagi-pho.jp](mailto:sato-ik510@miyagi-pho.jp)

<sup>3</sup> The University of Tokyo  
[hwseong@hotmail.com](mailto:hwseong@hotmail.com)

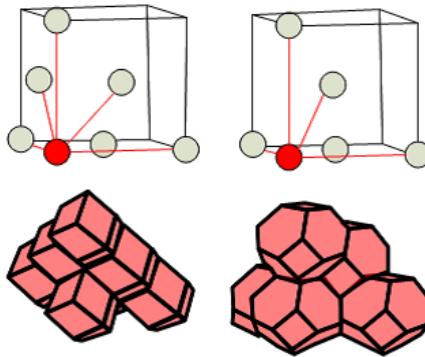
**Abstract.** Given two convex polyhedra  $\alpha$  and  $\beta$ , we say that  $\alpha$  and  $\beta$  are a **reversible pair** if  $\alpha$  has a dissection into a finite number of pieces which can be rearranged to form  $\beta$  in such a way that no face of the dissection of  $\alpha$  includes any part of an edge of  $\alpha$ , no face of the dissection of  $\beta$  includes any part of an edge of  $\beta$ , the pieces are hinged on some of their edges so that the pieces of the dissection are connected as in a tree-structure, all of the exterior surface of  $\alpha$  is in the interior of  $\beta$ , and all of the exterior surface of  $\beta$  comes from the interior of  $\alpha$ . Let  $\mathfrak{P}_i$  denote one of the five families of parallellohedra (see Section 2 for the corresponding definitions). In this paper, it is shown that given an arbitrary canonical parallellohedron  $P$ , there exists a canonical parallellohedron  $Q \in \mathfrak{P}_i$  such that the pair  $P$  and  $Q$  is reversible for each  $\mathfrak{P}_i$ .

## 1 Introduction

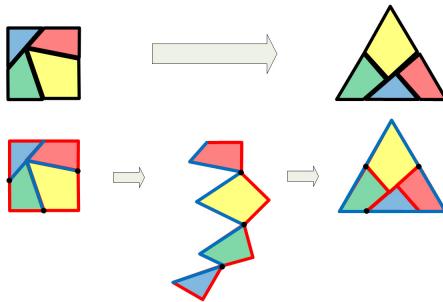
We often come across the “phase transition” phenomenon when observing natural objects in various fields including physics, crystallography, and chemistry. A good example of phase transition is the change by a metal crystal from a fcc (face-centered cubic) structure to a bcc (body-centered cubic) one. Fcc and bcc correspond to the closest packing and loosest covering of spheres, respectively. Corresponding Voronoi cells are re-arranged from a rhombic dodecahedron to truncated octahedron arrangement (Fig. 1; see [1] for details).

The molecular behavior of a form transition of this type is so micro-sized that we cannot observe the precise mechanism; however, there should exist some principle explaining the phenomenon. The problem of form transition is challenging, not only for physicists but also mathematicians. It is the question of “how” that is the motivation for this paper. In the last section of the paper, some properties of form transition between the five families of parallellohedra will be explained.

Reversibility is a special case of the dissection problem. See [2–7] for some of the background research, including Dehn invariants to solve Hilbert’s third problem, the folding and unfolding problem and Dudeney’s puzzle (Fig. 2). Reversibility between several polygons (reversibility can be extended to the 2-dimensional plane) and polyhedra are discussed in [8, 9].



**Fig. 1.** Voronoi cells of a face-centered cubic structure and a body-centered cubic structure



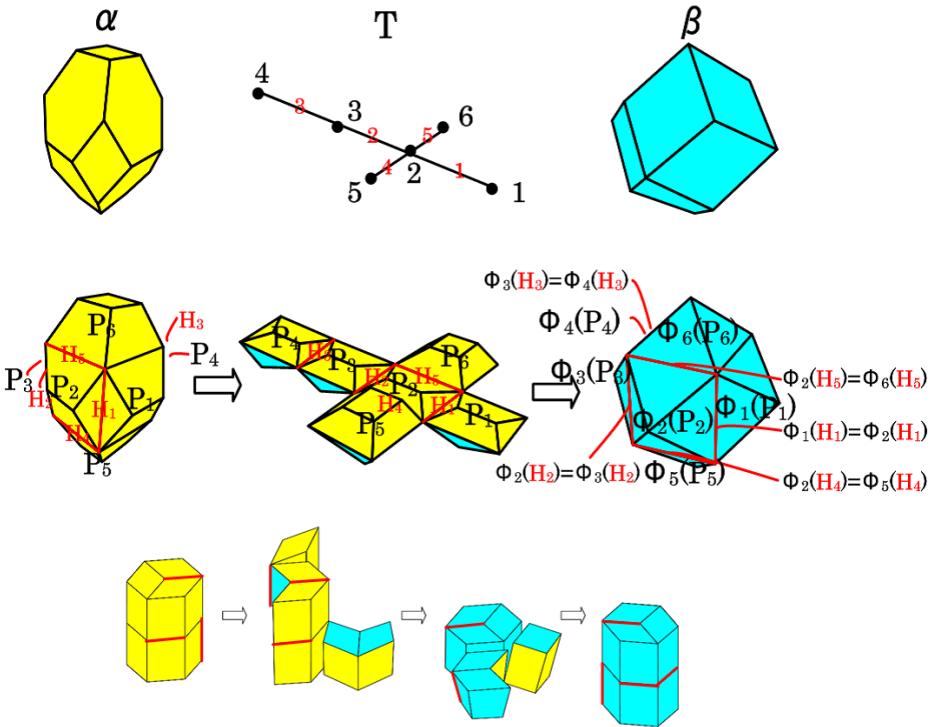
**Fig. 2.** The answer to Dudeney's puzzle and corresponding reversibility

Given two convex polyhedra  $\alpha$  and  $\beta$ , we say that  $\alpha$  and  $\beta$  are a **reversible pair** if the following exist: a finite tree  $T = (V; E)$ , a convex polyhedra  $P_v$  (where  $v \in V$ ), line segments  $H_e$  (where  $e \in E$ ), and orthogonal transformations  $\Phi_v$  (where  $v \in V$ ) which satisfy the following conditions:

- (1)  $\alpha$  can be dissected into the pieces  $P_v$  (where  $v \in V$ ).
- (2) No face of the dissection of  $\alpha$  includes any part of an edge of  $\alpha$ .
- (3) For any edge  $e \in E$  between two vertices  $v, w \in V$ , the line segment  $H_e$  is a common edge of the two pieces  $P_v$  and  $P_w$  ( $P_v$  and  $P_w$  are hinged on the line segment  $H_e$ ).
- (4) For any edge  $e \in E$  between two vertices  $v, w \in V$ ,  $\Phi_v(H_e) = \Phi_w(H_e)$ .
- (5)  $\beta$  can be dissected into the pieces  $\Phi_v(P_v)$  (where  $v \in V$ ).
- (6) No face of the dissection of  $\beta$  includes any part of an edge of  $\beta$ .
- (7) For any face  $F$  of a piece  $P_v$  (where  $v \in V$ ) lying on the exterior surface of  $\alpha$ , the corresponding face  $\Phi_v(F)$  of the rearranged piece  $\Phi_v(P_v)$  lies in the interior of  $\beta$ .

- (8) For any face  $F'$  of a rearranged piece  $\Phi_v(P_v)$  (where  $v \in V$ ) lying on the exterior surface of  $\beta$ , the corresponding face  $\Phi_v^{-1}(F')$  of the original piece  $P_v$  lies in the interior of  $\alpha$ .

Note that reversibility is symmetric; if two convex polyhedra  $\alpha$  and  $\beta$  are a reversible pair,  $\beta$  and  $\alpha$  are also a reversible pair, because the two convex polyhedra  $\beta$  and  $\alpha$ , the finite tree  $T = (V; E)$ , convex polyhedra  $\Phi_v(P_v)$  (where  $v \in V$ ), line segments  $\Phi_v(H_e)$  (where  $e \in E$ ), and orthogonal transformations  $\Phi_v^{-1}$  (where  $v \in V$ ) all satisfy all eight conditions in the definition of reversibility.



**Fig. 3.** The upper half shows an example of a reversible pair. The lower half shows an example which does not satisfy condition (2) in the definition of reversibility because the faces of dissection of  $\alpha$  include some edges of  $\alpha$ .

Note that the two convex polyhedra of a reversible pair have the same volume and the same Dehn invariant. When a convex polyhedron  $\alpha$  and its congruent copy are reversible, we say that  $\alpha$  is **self-reversible**.

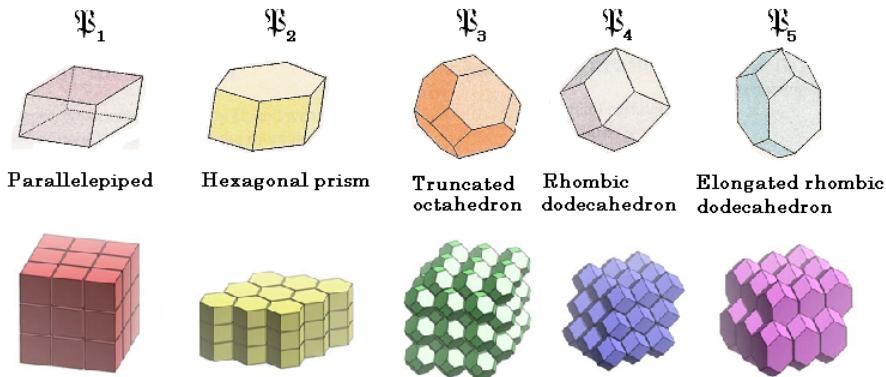
When two convex polyhedra  $\alpha$  and  $\beta$ , a finite tree  $T = (V; E)$ , convex polyhedra  $P_v$  (where  $v \in V$ ), line segments  $H_e$  (where  $e \in E$ ), and orthogonal transformations  $\Phi_v$  (where  $v \in V$ ) satisfy conditions (1), (2), and (5) to (8) in the definition

of reversibility (which means the pieces are not hinged on the line segments  $H_e$  (where  $e \in E$ )), we say that  $\alpha$  and  $\beta$  are a **weakly reversible pair**.

## 2 Parallellohedra and Their Lattice Structures

A **parallellohedron** is a convex polyhedron which fills the 3-dimensional space (i.e., a **space-filler**) by translations only. In 1885, a Russian crystallographer, Evgraf Fedorov [10], established that there are exactly five families of parallellohedra ([11]); namely, parallelepipeds, rhombic dodecahedrons, hexagonal prisms, elongated rhombic dodecahedrons, and truncated octahedrons (Fig. 4). They fill the space as shown in Figure 4. Each copy of a parallellohedron used in a tessellation is called a **cell** of the tessellation.

Label each of the five families of parallellohedra as follows: parallelepipeds by  $\mathfrak{P}_1$ , hexagonal prisms by  $\mathfrak{P}_2$ , truncated octahedrons by  $\mathfrak{P}_3$ , rhombic dodecahedrons by  $\mathfrak{P}_4$ , and elongated rhombic dodecahedrons by  $\mathfrak{P}_5$ .

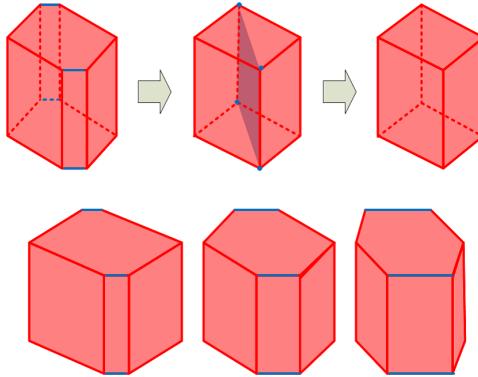


**Fig. 4.** The five families of parallellohedra

A **canonical** parallellohedron  $P$  is defined to be a axis-symmetric parallellohedron with respect to an orthogonal coordinate system denoted by  $C_P$ .

Each of the five families contains infinitely many non-similar canonical parallellohedra with the same combinatorial structure. A **scaling transformation** is defined to be a linear transformation that enlarges or shrinks a canonical parallellohedron  $P$  in the direction of each of the axes of the corresponding orthogonal coordinate system  $C_P$ . A series  $Z$  of parallel sides of a parallellohedron  $P$  is said to be **contractible** if contracting all these sides in  $Z$  results in another parallellohedron  $P'$ . A **stretching transformation** is defined to be a transformation changing the length of a contractible series of parallel sides of a parallellohedron preserving the volume and its lattice (Fig. 5; see [12] and [13] for details). Apply a **scaling stretching transformation**  $f$ , which is a composition of a scaling transformation and a stretching transformation, to a canonical parallellohedron

$P \in \mathfrak{P}_i$ . Then  $f(P)$  is a canonical parallelohedron in the same family  $\mathfrak{P}_i$ . Also, for any canonical parallelohedron  $B \in \mathfrak{P}_i$  there exists a scaling stretching transformation  $f$  such that  $B = f(P)$  because the scaling stretching transformation does not affect parallelism, the space-filling property, and canonicity.



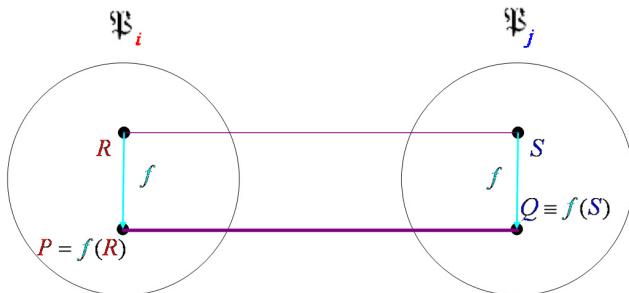
**Fig. 5.** The upper half shows an example of contractible series of parallel sides of a parallelohedron. The lower half shows an example of parallelohedra transformed by a scaling stretching transformation.

**Lemma 1.** Suppose that a pair of two canonical parallelohedra  $R \in \mathfrak{P}_i$  and  $S \in \mathfrak{P}_j$  ( $1 \leq i, j \leq 5$ ) is reversible.

Then for an arbitrary canonical parallelohedron  $P \in \mathfrak{P}_i$ , there exists a canonical parallelohedron  $Q \in \mathfrak{P}_j$  such that the pair  $P$  and  $Q$  is reversible.

*Proof.* Consider an arbitrary canonical parallelohedron  $P \in \mathfrak{P}_i$ . Because  $P$  and  $R$  belong to the same family  $\mathfrak{P}_i$ ,  $P$  can be obtained by applying a scaling stretching transformation  $f$  to  $R$ . Thus  $P = f(R)$ .

From the assumption that the pair  $R$  and  $S$  is reversible, the pair  $f(R) = P$  and  $f(S)$  is reversible because we can use the faces of dissection and rearrangement derived from  $f$  (see [13] for details). Here,  $f(S)$  and  $S$  belong to the same family  $\mathfrak{P}_j$ . The proof is completed by letting  $Q \equiv f(S)$ . Q.E.D.



Thus we know that there exist infinitely many reversible pairs between two families if we find even one reversible pair between the two families.

Let  $P \in \mathfrak{P}_i$ ,  $Q \in \mathfrak{P}_j$ , and  $T(P)$ ,  $T(Q)$  be two space tessellations, one by  $P$  and the other by  $Q$ . Construct the lattices  $L(T(P))$  and  $L(T(Q))$  by taking the center of each cell of  $T(P)$  and  $T(Q)$ , respectively. We say that two lattices  $L(T(P))$  and  $L(T(Q))$  are **identical** when the point set of  $L(T(P))$  and the point set of  $L(T(Q))$  become equal after applying a translation. When the faces of  $T(P)$  cut the cells of  $T(Q)$  and vice versa so that each cell of  $T(P)$  and  $T(Q)$  contains the same set of pieces, we say that  $T(P)$  and  $T(Q)$  **overlap uniformly**.

We define standard canonical parallelohedra, their standard tessellations, and associated lattices as shown in Fig. 6. A **standard parallelepiped**  $P_1$  is a cube with side length 1. A **standard hexagonal prism**  $P_2$  is a hexagonal prism consisting of two regular hexagons and six squares, with side length 1. A **standard truncated octahedron**  $P_3$  is a truncated octahedron consisting of six squares and eight regular hexagons, with side length  $\frac{1}{2}$ . A **standard rhombic dodecahedron**  $P_4$  is a rhombic dodecahedron consisting of twelve rhombi whose two diagonal lengths are 1 and  $\sqrt{2}$ , with side length  $\frac{\sqrt{3}}{2}$ . A **standard elongated rhombic dodecahedron**  $P_5$  is an elongated rhombic dodecahedron with four regular hexagons and eight rhombi whose two diagonal lengths are 1 and  $\sqrt{2}$ , with side length  $\frac{\sqrt{3}}{2}$ . The **standard tessellation**  $T(P_i)$  by each standard parallelohedron  $P_i$  is defined to be the tessellations by  $P_i$  where all the cells are connected by faces. Label the lattice derived from  $T(P_i)$  by  $L_i = L(T(P_i))$ .

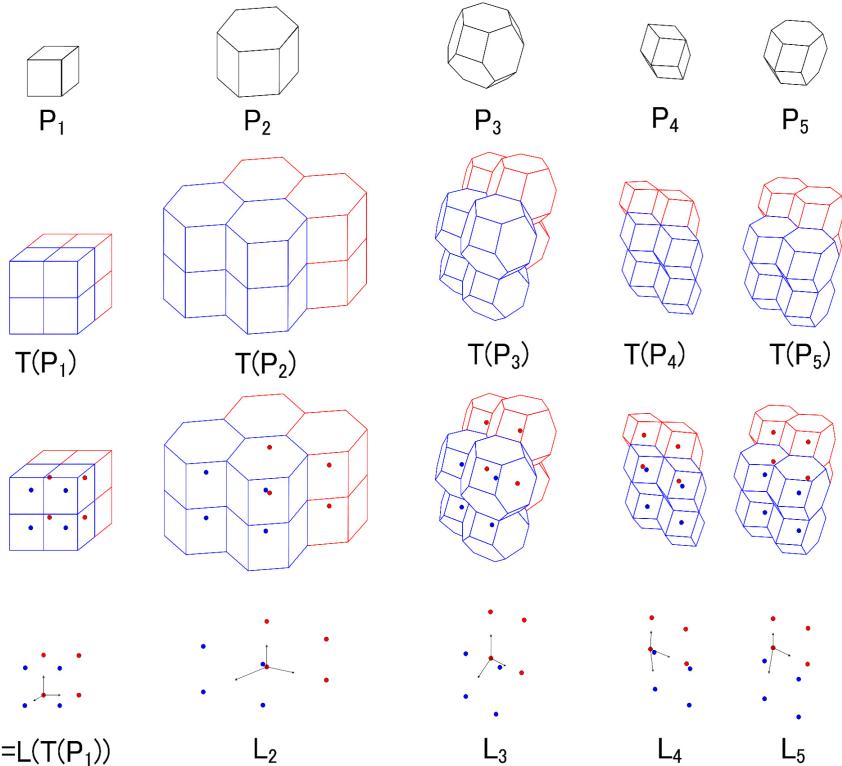
**Lemma 2.** *If two lattices  $L(T(P))$  and  $L(T(Q))$  are identical, and associated tessellations  $T(P)$  and  $T(Q)$  overlap uniformly in each cell, then the pair of two canonical parallelohedra  $P$  and  $Q$  is reversible.*

*Proof.* Since the tessellations  $T(P)$  and  $T(Q)$  overlap uniformly in each cell, each cell of  $T(P)$  contains the same set of pieces of  $T(Q)$ . Also, each cell of  $T(P)$  contains the pieces from exactly one cell of  $T(Q)$  because the lattices  $L(T(P))$  and  $L(T(Q))$  are identical and one cell of  $T(P)$  corresponds to one cell of  $T(Q)$ .

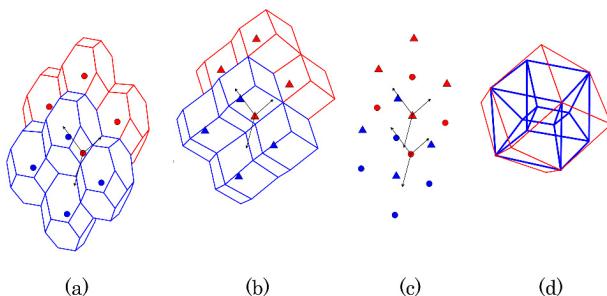
Therefore we can rearrange the pieces in a cell of  $T(P)$  to be the same arrangement with the pieces as a cell of  $T(Q)$ , so that the pieces are hinged (any graph in the definition of the reversibility where line segments  $H_e$  ( $e \in E$ ) are disjoint can be chosen because  $P$  and  $Q$  are canonical; see [12] for details), and the exterior surface of  $P$  goes into the interior of  $Q$ . Thus the pair  $P$  and  $Q$  is reversible. Q.E.D.

*Remark 1.* For example, a reversible pair of a truncated octahedron and a rhombic dodecahedron is found as follows. Apply a scaling transformation so that the height of  $P_3$  is enlarged by  $\sqrt{2}$ . The resultant canonical parallelohedron  $P'_3$ , the standard tessellation  $T(P'_3)$  of  $P'_3$  and associated lattice  $L'_3 = L(T(P'_3))$  are as shown in Fig. 7. Then the lattices  $L'_3$  and  $L_4$  are identical.

When  $T(P'_3)$  and  $T(P_4)$  overlap uniformly in each cell,  $T(P_4)$  determines the way to dissect  $P_3$  and  $T(P'_3)$  determines the way to dissect  $P_4$ , as shown in Fig. 7(d).



**Fig. 6.** Standard canonical parallellohedra, their standard tessellations, and associated lattices



**Fig. 7.** (a) The tessellation  $T(P'_3)$  and associated lattice  $L'_3$  (b) The tessellation  $T(P_4)$  and associated lattice  $L_4$  (c) The lattices  $L'_3$  and  $L_4$  when  $T(P'_3)$  and  $T(P_4)$  overlap uniformly in each cell (d) The way to dissect the two canonical parallellohedra

In order to find a reversible pair  $P$  and  $Q$ , we apply a scaling transformation to the tessellations  $T(P)$  and  $T(Q)$  and rearrange the cells of each tessellation so that the lattices  $L(T(P))$  and  $L(T(Q))$  become identical. Then we place the tessellations  $T(P)$  and  $T(Q)$  and apply a stretching transformation so that they overlap uniformly. This idea about tessellations and associated lattices plays an important role in studying reversibility in other cases.

### 3 Self-reversibilities of Parallellohedra

We discuss five kinds of self-reversibility of all canonical parallellohedra in this section.

**Theorem 1.** *Every canonical parallellohedron is self-reversible.*

*Proof.* An arbitrary canonical parallellohedron  $P$  belongs to a certain family  $\mathfrak{P}_i$  ( $1 \leq i \leq 5$ ).

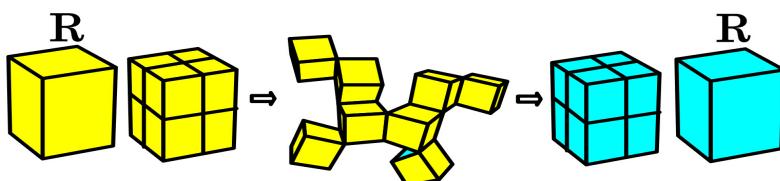
We find a self-reversible canonical parallellohedron  $R$  for each family  $\mathfrak{P}_i$ .

For  $i = 1$ , the parallelepiped is dissected once in each direction into eight pieces. For  $i = 2$ , the hexagonal prism is first dissected into four parts by the faces of dissection based on the lines of dissection used to self-reverse a convex hexagon in which one pair of opposite sides is parallel and of the same length ([9]), and then dissected in the other direction into eight pieces. After dissecting the parallellohedron, we apply an appropriate rearrangement to conceal the exterior surface.

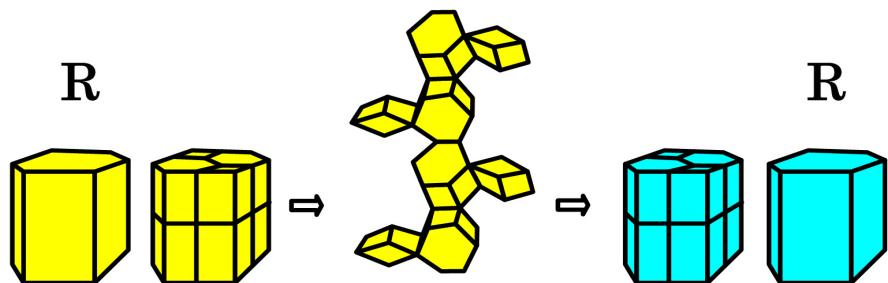
In the cases  $i = 3, 4, 5$ , after dissecting parallellohedra into a parallelepiped and six roof-shaped polyhedra (two hexahedra and four pentahedra for  $i = 3$ , six quadrilateral pyramids for  $i = 4$ , two quadrilateral pyramids and four pentahedra for  $i = 5$ ), we make one piece for each roof together with its mirror image against its bottom. After dissecting the parallellohedron, we expand the pieces so that the planes of symmetry form a net of a parallelepiped and fold the pieces in the other direction. The pieces can be folded to form a parallelepiped as shown in Fig. 13.

Thus the pair  $R \in \mathfrak{P}_i$  and  $R \in \mathfrak{P}_i$  is reversible.

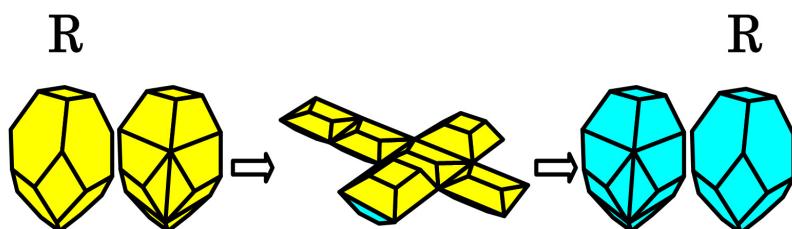
Because  $P$  and  $R$  belong to the same family  $\mathfrak{P}_i$ , the pair  $P = f(R)$  and  $P = f(R)$  is reversible from Lemma 1. Thus  $P$  is self-reversible. Q.E.D.



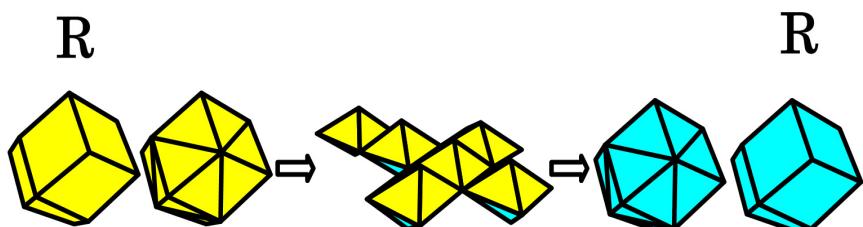
**Fig. 8.**  $i = 1$  (parallelepiped). The parallellohedron  $R$  is self-reversible as shown.



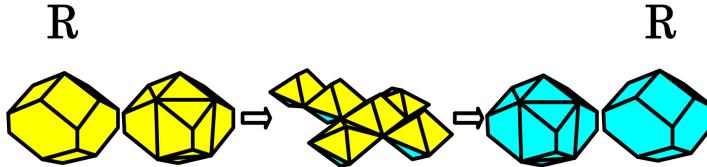
**Fig. 9.**  $i = 2$  (hexagonal prism). The parallelohedron  $R$  is self-reversible as shown.



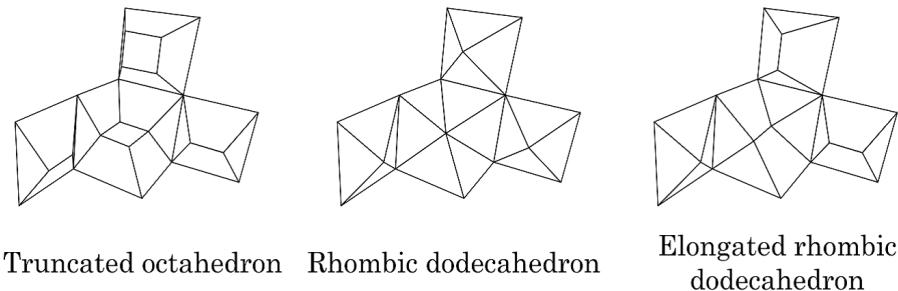
**Fig. 10.**  $i = 3$  (truncated octahedron). The parallelohedron  $R$  is self-reversible as shown.



**Fig. 11.**  $i = 4$  (rhombic dodecahedron). The parallelohedron  $R$  is self-reversible as shown.



**Fig. 12.**  $i = 5$  (elongated rhombic dodecahedron). The parallelohedron  $R$  is self-reversible as shown.



**Fig. 13.** The roof-shaped polyhedra form a parallelepiped

## 4 Reversibility among Parallelohedra

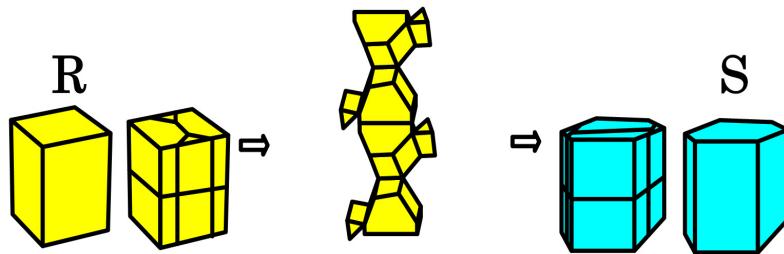
We discuss ten kinds of reversibility between two non-congruent canonical parallelohedra in this section.

**Theorem 2.** *For an arbitrary canonical parallelohedron  $P \in \mathfrak{P}_i$  and an arbitrary family  $\mathfrak{P}_j$  of parallelohedra, there exists a canonical parallelohedron  $Q \in \mathfrak{P}_j$  such that the pair  $P$  and  $Q$  is reversible, where  $i$  and  $j$  satisfy  $1 \leq i, j \leq 5$ .*

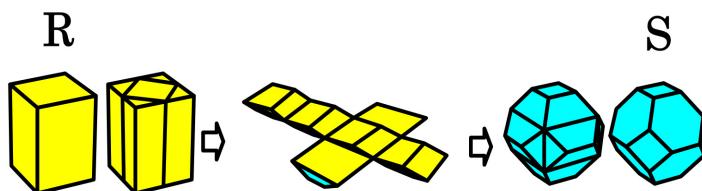
*Proof.* When  $i = j$ , the pair  $P$  and  $Q$  become reversible by letting  $Q \equiv P$  from Theorem 1. By symmetry, we can suppose that  $i < j$ . For each case  $(i, j)$ , we find a reversible pair of two canonical parallelohedra  $R \in \mathfrak{P}_i$  and  $S \in \mathfrak{P}_j$ .

In this case, the parallelepiped is dissected into eight pieces by the faces of dissection in the case  $i = 2$  in the proof of Theorem 1. After dissecting the parallelepiped, we apply an appropriate rearrangement to conceal the exterior surface, similarly to the cases  $i = 2$  in the proof of the theorem.

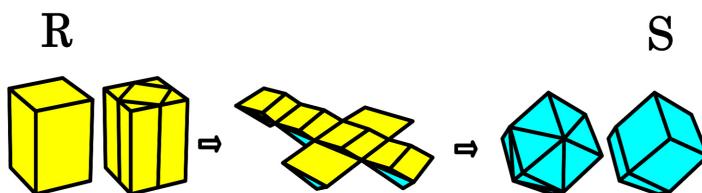
In the cases  $(i, j) = (1, 3), (1, 4), (1, 5)$ , the parallelepiped is first dissected by separating two roof-shaped polyhedra with a common apex. The remaining polyhedron is dissected into four roof-shaped polyhedra by the faces of dissection based on the lines of dissection used to self-reverse a parallelogram. After dissecting the parallelepiped, we expand the pieces like a net of a parallelepiped



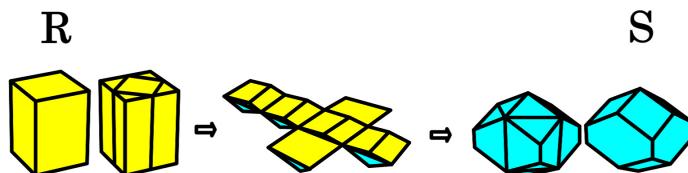
**Fig. 14.**  $(i, j) = (1, 2)$  (between parallelepiped and hexagonal prism). The pair  $R$  and  $S$  is reversible as shown.



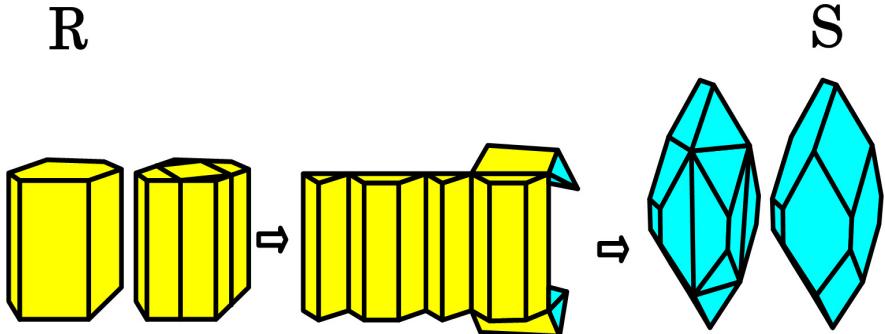
**Fig. 15.**  $(i, j) = (1, 3)$  (between parallelepiped and truncated octahedron). The pair  $R$  and  $S$  is reversible as shown.



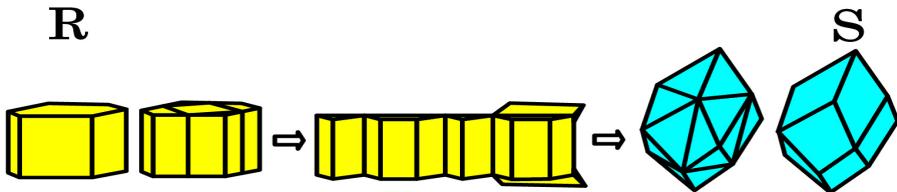
**Fig. 16.**  $(i, j) = (1, 4)$  (between parallelepiped and rhombic dodecahedron). The pair  $R$  and  $S$  is reversible as shown.



**Fig. 17.**  $(i, j) = (1, 5)$  (between parallelepiped and elongated rhombic dodecahedron). The pair  $R$  and  $S$  is reversible as shown.



**Fig. 18.**  $(i, j) = (2, 3)$  (between hexagonal prism and truncated octahedron). The pair  $R$  and  $S$  is reversible as shown.



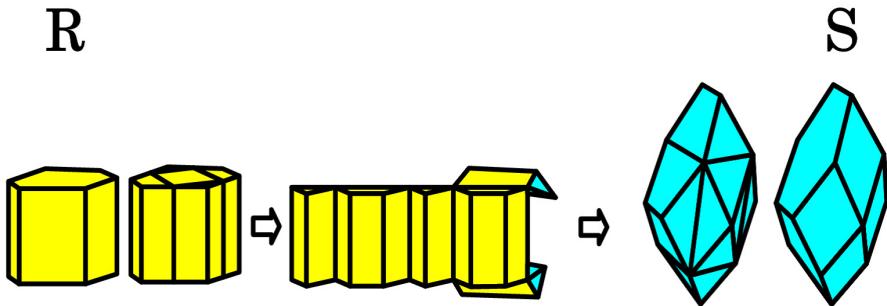
**Fig. 19.**  $(i, j) = (2, 4)$  (between hexagonal prism and rhombic dodecahedron). The pair  $R$  and  $S$  is reversible as shown.

and fold the pieces in the other direction. The pieces can be folded to form a truncated octahedron for  $j = 3$ , a rhombic dodecahedron for  $j = 4$ , and an elongated rhombic dodecahedron for  $j = 5$ .

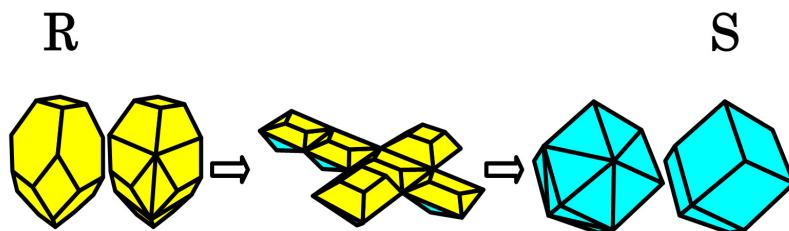
In the cases  $(i, j) = (2, 3), (2, 4), (2, 5)$ , the hexagonal prism is first dissected by separating two roof-shaped polyhedra with a common apex. The remaining polyhedron is dissected into four roof-shaped polyhedra by the faces of dissection based on the line of dissection used to self-reverse a convex hexagon in which one pair of opposite sides is parallel and of the same length. After dissecting the hexagonal prism, we expand the pieces so that the planes of symmetry form a net of a parallelepiped and fold the pieces in the other direction. The pieces can be folded to form a truncated octahedron for  $j = 3$ , a rhombic dodecahedron for  $j = 4$ , and an elongated rhombic dodecahedron for  $j = 5$ .

The cases  $(i, j) = (3, 4), (3, 5), (4, 5)$  are similar to the cases  $i = 3, 4, 5$  in the proof of Theorem 1. Here we take one roof from each parallellohedron to make a piece.

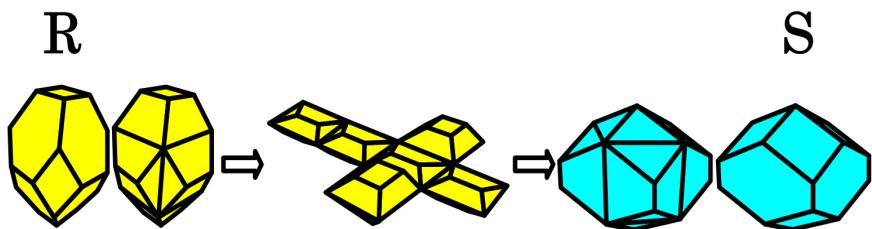
Thus the pair of two canonical parallellohedra  $R \in \mathfrak{P}_i$  and  $S \in \mathfrak{P}_j$  is reversible.



**Fig. 20.**  $(i, j) = (2, 5)$  (between hexagonal prism and elongated rhombic dodecahedron). The pair  $R$  and  $S$  is reversible as shown.



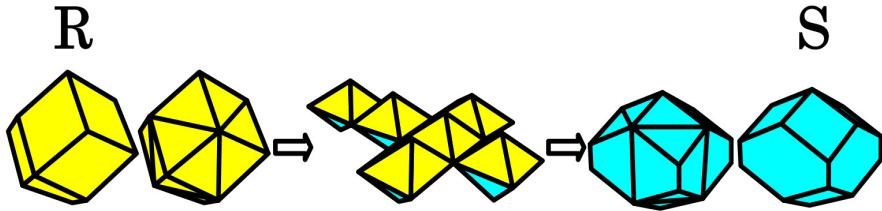
**Fig. 21.**  $(i, j) = (3, 4)$  (between truncated octahedron and rhombic dodecahedron). The pair  $R$  and  $S$  is reversible as shown.



**Fig. 22.**  $(i, j) = (3, 5)$  (between truncated octahedron and elongated rhombic dodecahedron). The pair  $R$  and  $S$  in the figure above is reversible as shown.

Because  $P$  and  $R$  belong to the same family  $\mathfrak{P}_i$ , there exists a canonical parallelohedron  $Q \in \mathfrak{P}_j$  such that the pair  $P$  and  $Q$  is reversible from Lemma 1. Q.E.D.

From Theorem 2, we obtain the following theorem about weak reversibility among parallelohedra.



**Fig. 23.**  $(i, j) = (4, 5)$  (between rhombic dodecahedron and elongated rhombic dodecahedron). The pair  $R$  and  $S$  is reversible as shown.

**Theorem 3.** *For an arbitrary parallelohedron  $\alpha \in \mathfrak{P}_i$  and an arbitrary family  $\mathfrak{P}_j$  of parallelohedra, there exists a parallelohedron  $\beta \in \mathfrak{P}_j$  such that the pair  $P$  and  $Q$  is weakly reversible, where  $i$  and  $j$  satisfy  $1 \leq i, j \leq 5$ .*

*Proof.* A parallelohedron  $\alpha \in \mathfrak{P}_i$  can be obtained by applying an affine transformation  $f$  to a canonical parallelohedron  $P \in \mathfrak{P}_i$  (see [12] for details). From Theorem 2, there exists a canonical parallelohedron  $Q \in \mathfrak{P}_j$  such that the pair  $P$  and  $Q$  is reversible. Thus  $\alpha = f^{-1}(P)$  and  $\beta \equiv f^{-1}(Q)$  are weakly reversible because Lemma 1 holds for affine transformation and weak reversibilities.

It seems that no line segments  $H_e$  ( $e \in E$ ) satisfy condition (7) and (8) in the definition of reversibility.

We end with the following conjecture.

*Conjecture 1.* Two non-canonical parallelohedra  $\alpha$  and  $\beta$  can never be reversible.

**Acknowledgements.** The authors would like to thank referees and Mari-Jo Ruiz for giving us various invaluable comments.

## References

1. Conway, J.H., Sloane, N.J.A.: Sphere packings, lattices and groups, 3rd edn. A Series of Comprehensive Studies in Mathematics, vol. 290. Springer (1999)
2. Boltyanskii, V.G.: Equivalent and Equidecomposable Figures. D. C. Heath and Co. (1963), Translated and adapted from the first Russian edition (1956) by Henn, A.K., Watts, C.E.
3. Boltyanskii, V.G.: Hilbert's third problem. V. H. Winton & Sons (1978), Translated by Silverman, R.A.
4. Frederickson, G.N.: Dissections: Plane and Fancy. Cambridge University Press, New York (1997)
5. Frederickson, G.N.: Hinged Dissections: Swinging and Twisting. Cambridge University Press, New York (2002)
6. Frederickson, G.N.: Piano-Hinged Dissections: Time to Fold. Cambridge University Press, New York (2006)

7. Demaine, E.D., O'Rourke, J.: Geometric Folding Algorithms: Linkages, Origami, Polyhedra. Cambridge University Press, New York (2007)
8. Akiyama, J., Nakamura, G.: Dudeney Dissection of Polygons. In: Akiyama, J., Kano, M., Urabe, M. (eds.) JCDCG 1998. LNCS, vol. 1763, pp. 14–29. Springer, Heidelberg (2000)
9. Akiyama, J., Nakamura, G.: Congruent Dudeney dissections of triangles and convex quadrilaterals – all hinge points interior to the sides of the polygons. In: Aronov, B., Basu, S., Pach, J., Sharir, M. (eds.) Discrete and Computational Geometry. Algorithms and Combinatorics, vol. 25, pp. 43–63 (2003)
10. Fedorov, E.S.: An introduction to the theory of figures. In: Notices of the Imperial Mineralogical Society (St. Petersburg) Ser. 2, vol. 21, pp. 1–279 (1885); Republished with comments by Akad. Nauk. SSSR, Moscow (1953) (in Russian)
11. Alexandrov, A.D.: Convex Polyhedra. Springer Monographs in Mathematics (2005)
12. Dolbilin, N., Itoh, J., Nara, C.: Geometric realization on affine equivalent 3-parallelohedra (to be published)
13. Akiyama, J., Seong, H.: On the reversibilities among quasi-parallellogons (to be published)

# A History of Flips in Combinatorial Triangulations<sup>\*,\*\*</sup>

Prosenjit Bose and Sander Verdonschot

School of Computer Science, Carleton University  
`jit@scs.carleton.ca, sander@cg.scs.carleton.ca`

**Abstract.** Given two combinatorial triangulations, how many edge flips are necessary and sufficient to convert one into the other? This question has occupied researchers for over 75 years. We provide a comprehensive survey, including full proofs, of the various attempts to answer it.

## 1 Introduction

A triangulation is a simple planar graph where every face including the outer face is a cycle of length 3. In any triangulation, an edge  $e = xy$  is adjacent to two faces:  $xya$  and  $xyb$ . An edge flip consists of deleting the edge  $e$  from the triangulation and adding the unique edge  $e' = ab$  to the graph such that it remains a triangulation. In other words, the edge  $e$  is flippable provided that  $ab$  is not currently an edge of the triangulation. If the vertices have fixed coordinates in the plane, the restriction that the new edge may not introduce any crossings is usually added. This is commonly referred to as the *geometric* setting. However, we focus on the problem in the *combinatorial* setting, where we are only given a combinatorial embedding of the graph (the clockwise order of edges around each vertex). Even in this setting, not all edges in a triangulation are flippable. Gao *et al.* [4] showed that in every  $n$ -vertex triangulation at least  $n - 2$  edges are always flippable and that there exist some triangulations where at most  $n - 2$  edges are flippable. Moreover, if the triangulation has minimum degree at least 4, then they showed that there are at least  $2n + 3$  flippable edges and the bound is tight in certain cases. Note that by flipping an edge  $e$ , we transform one triangulation into another. But this triangulation could be isomorphic to the triangulation prior to the flip, or we could end up in a cycle of a few triangulations. This gives rise to the following question: Can any  $n$ -vertex triangulation be transformed into any other  $n$ -vertex triangulation through a finite sequence of flips? To our knowledge, Wagner [8] was the first to address this question directly and he answered it in the affirmative. Although it is well known that the number of  $n$ -vertex triangulations is exponential in  $n$ , Wagner's inductive proof gives rise to a construction algorithm that can achieve this transformation using at most

---

\* As flips are a topic close to Ferran Hurtado's heart, we would like to dedicate this article to him on the occasion of his 60th birthday.

\*\* Research supported in part by NSERC.

$2n^2$  edge flips. The surprising element of Wagner’s proof is that he circumvents the issue of graph isomorphism by showing how to convert any given triangulation into a *canonical* triangulation that can be easily recognized. The canonical triangulation is the unique triangulation that consists of two dominant vertices (vertices that are adjacent to all other vertices). Unfortunately, the curse of this approach is that one may use many more flips than necessary to convert one triangulation into another. In fact, it is possible that two triangulations are one edge flip away from each other but Wagner’s approach uses a quadratic number of flips to convert one into the other.

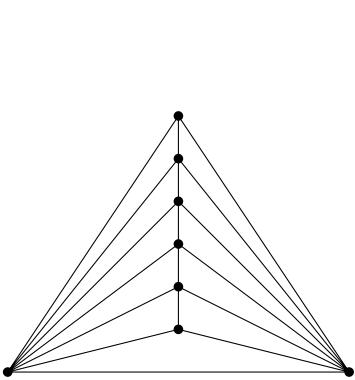
The notion of two triangulations being “close” to each other in terms of number of flips can be expressed through a *flip graph*. The *flip graph* has as vertex set all possible  $n$ -vertex triangulations and two vertices in the flip graph are joined by an edge provided that the respective triangulations differ by exactly one flip. Questions about the flip operation can be viewed as questions on the flip graph. Asking whether any  $n$ -vertex triangulation can be converted into any other via flips is asking whether the flip graph is connected. Asking for the smallest number of flips required to convert one triangulation into another is asking for the shortest path in the flip graph between the two vertices representing the given triangulations. The maximum, minimum and average degree in the flip graph almost correspond to the maximum, minimum and average number of flippable edges. The caveat is that one needs to account for isomorphic triangulations when computing the degree of a vertex in the flip graph. One can also ask what the chromatic number of the flip graph is, whether it is hamiltonian etc. Many of these questions have been addressed in the literature [2]. However, we focus mainly on attempts to determine the diameter of the flip graph. In other words, how many edge flips are sufficient and sometimes necessary to transform a given triangulation into any other? We go beyond merely stating the results by providing a substantial amount of detail on the proofs.

## 2 Wagner’s Bound

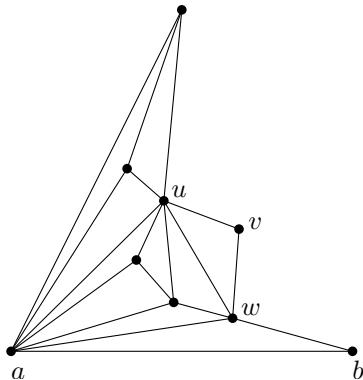
Wagner in 1936 [8] first addressed the problem of determining whether one can convert a given triangulation into another via edge flips. Although his paper is entitled “Remarks on the four-color problem”, it contains a proof that every planar graph has a straight-line embedding, defines the edge flip operation (or diagonal transformation, as Wagner calls it) and shows that any two triangulations can be transformed into each other by a finite series of edge flips before finishing with a result on the number of valid colorings of a graph.

To prove that any pair of triangulations can be transformed into each other via flips, Wagner first introduces the *canonical triangulation*, which is the unique triangulation with two dominant vertices (see Figure 1). We will denote the canonical triangulation on  $n$  vertices by  $\Delta_n$ .

**Lemma 1. (Wagner [8], Theorem 4)** *Any triangulation on  $n$  vertices can be transformed into  $\Delta_n$  by a sequence of at most  $n^2 - 7n + 12$  flips.*



**Fig. 1.** The canonical triangulation on 8 vertices



**Fig. 2.** A face  $uvw$  such that  $u$  and  $w$  are neighbours of  $a$ , while  $v$  is not. Flipping the edge  $uw$  brings us closer to the canonical triangulation.

*Proof.* To transform a given triangulation into the canonical one, we fix an outer face and pick two of its vertices, say  $a$  and  $b$ , to become the dominant vertices in the canonical triangulation. If  $a$  is not adjacent to all other vertices, there exists a face  $uvw$  such that  $u$  and  $w$  are neighbours of  $a$ , while  $v$  is not. This situation is illustrated in Figure 2. We flip the edge  $uw$ .

In his original proof, Wagner argues that this gives a finite sequence of flips that increases the degree of  $a$  by one. He simply states that this sequence is finite and does not argue why  $uw$  is flippable in the first place. We provide these additional arguments below.

We consider two cases:

- $a uw$  is a face. In this case, the flip will result in the edge  $av$ , increasing the degree of  $a$  by one. This flip is valid, as  $v$  was not adjacent to  $a$  before the flip.
- $a uw$  is not a face. In this case the flip is also valid, since  $a uw$  forms a triangle that separates  $v$  from the vertices inside. The flip does not increase the degree of  $a$ , but it does increase the degree of  $v$  and since the number of vertices is finite, the degree of  $v$  cannot increase indefinitely. Therefore, we must eventually arrive to the first case, where we increase the degree of  $a$  by one.

Since the same strategy can be used to increase the degree of  $b$  as long as it is not dominant, this gives us a sequence of flips that transforms any triangulation into the canonical one. Every vertex of a triangulation has degree at least 3, so the degree of  $a$  and  $b$  needs to increase by at most  $n - 4$ . Since we might need to increase the degree of  $v$  from 2 until it is adjacent to all but one of the

neighbours of  $a$  or  $b$ , the total flip sequence has length at most

$$2 \sum_{i=3}^{n-2} (i-2) = n^2 - 7n + 12$$

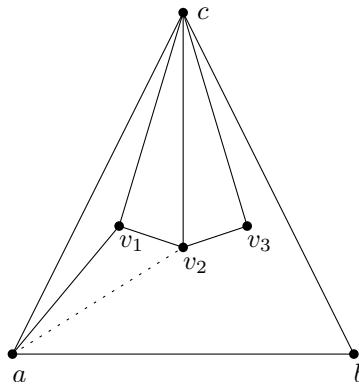
□

By using the canonical triangulation as an intermediate form, the main result follows.

**Theorem 2. (Wagner [8], Theorem 4)** *Any pair of triangulations  $T_1$  and  $T_2$  on  $n$  vertices can be transformed into each other by a sequence of at most  $2n^2 - 14n + 24$  flips.*

*Proof.* By Lemma 1, we have two sequences of flips,  $S_1$  and  $S_2$ , that transform  $T_1$  and  $T_2$  into the canonical triangulation, respectively. Since a flip can be reversed, we can use  $S_1$ , followed by the reverse of  $S_2$  to transform  $T_1$  into  $T_2$ . Since both  $S_1$  and  $S_2$  have length at most  $n^2 - 7n + 12$ , the total sequence uses at most  $2n^2 - 14n + 24$  flips. □

A simpler and more precise proof that also gives a quadratic upper bound was given by Negami and Nakamoto [7].



**Fig. 3.** The exterior triangle  $abc$  with the first three neighbours of  $c$  in counter-clockwise order. Depending on the presence of edge  $av_2$ , either  $cv_1$  or  $cv_2$  is flipped.

**Lemma 3. (Negami and Nakamoto [7], Theorem 1)** *Any triangulation on  $n$  vertices can be transformed into  $\Delta_n$  by a sequence of  $O(n^2)$  flips.*

*Proof.* Let  $abc$  be the outer face. Suppose we wish to make both  $a$  and  $b$  dominant. Instead of showing that a sequence of flips can always increase the degree of  $a$  or  $b$ , we will show that it is always possible to find one flip that decreases the

degree of  $c$ . Once  $c$  has degree 3, the same argument can be used to find a flip that decreases the degree of  $c$ 's neighbour inside the triangle until it has degree 4, and so on.

To determine which edge to flip, let  $a, v_1, v_2, \dots, b$  be the neighbours of  $c$  in counter-clockwise order. This situation is illustrated in Figure 3. If  $a$  and  $v_2$  are not adjacent, we can flip  $cv_1$  into  $av_2$ , reducing the degree of  $c$ . If  $a$  and  $v_2$  are adjacent,  $av_2c$  forms a cycle that separates  $v_1$  and  $v_3$ , so we can flip  $cv_2$  to reduce  $c$ 's degree. We continue this until  $c$  has degree 3, at which point we apply the same argument to reduce the degree of  $c$ 's remaining neighbour inside the triangle until it has degree 4. Then we continue with the neighbour of  $v_1$  inside the triangle  $av_1b$ , and so on, until all vertices except for  $a$  and  $b$  have degree 3 or 4, at which point we have obtained the canonical triangulation.  $\square$

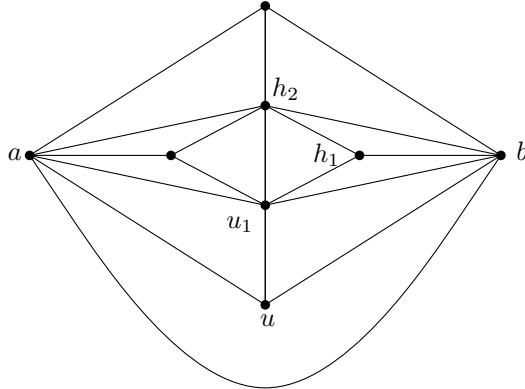
### 3 Komuro's Bound

Since Wagner's result, it remained an open problem whether the diameter of the flip graph was indeed quadratic in the number of vertices. Komuro [5] showed that in fact the diameter was linear by proving a linear upper and lower bound. We present the argument for the upper bound in this section and discuss the lower bound in Section 6.

Komuro used Wagner's approach of converting a given triangulation into the canonical triangulation. Given an arbitrary triangulation, the key is to bound the number of flips needed to make two vertices, say  $a, b$ , dominant. If there always exists one edge flip that increases the degree of  $a$  or  $b$  by 1, then at most  $2n - 8$  flips are sufficient since dominant vertices have degree  $n - 1$  and all vertices in a triangulation have degree at least 3. However, this is not always the case. Figure 4 shows a triangulation where no single flip increases the degree of  $a$  or  $b$ . Komuro used the following function to bound the number of flips:  $d_G(a, b) = 3\deg(a) + \deg(b)$ . He showed that there always exists either one edge flip where  $d_G(a, b)$  goes up by at least 1 or two edge flips where  $d_G(a, b)$  goes up by at least 2. The cleverness of the function is that in some cases, two edge flips increase the degree of  $a$  by 1 but decrease the degree of  $b$  by 1. However, since the function increases by 2, it still increases by at least 1 per flip. Since  $d_G(a, b) \leq 4n - 4$ , we have that  $4n - 4 - d_G(a, b)$  is an upper bound on the number of flips required to make  $a$  and  $b$  dominant.

**Lemma 4. (Komuro [5], Lemma 2)** *Let  $G$  be a triangulation on  $n$  vertices and let  $a, b$  be any pair of adjacent vertices of  $G$ . Then  $G$  can be transformed into the canonical triangulation  $\Delta_n$  with  $a$  and  $b$  as dominant vertices with at most  $4n - 4 - (3\deg(a) + \deg(b))$  edge flips.*

*Proof.* In a triangulation, every vertex must have degree at least 3. Let  $uab$  be a face adjacent to  $ab$ . We consider two cases:  $\deg(u) = 3$  and  $\deg(u) > 3$ . We begin with the latter. Since  $\deg(u) \geq 4$ , let  $a, b, w_1, w_2$  be four consecutive neighbours of  $u$  in counter-clockwise order. If  $b$  is not adjacent to  $w_2$ , then flipping edge  $uw_1$



**Fig. 4.** No single edge can be flipped to increase the degree of  $a$  or  $b$

increases  $\deg(b)$  by 1 and thus  $d_G(a, b)$  by 1. If  $b$  is adjacent to  $w_2$ , then  $ubw_2$  is a separating triangle (a cycle of length 3 whose removal disconnects the graph) that separates  $a$  from  $w_1$ . Therefore, flipping edge  $ub$  decreases  $\deg(b)$  by 1 and increases  $\deg(a)$  by 1. Thus, with one flip  $d_G(a, b)$  increases by 2.

Now consider the case when  $\deg(u) = 3$ . Let  $u_1$  be the unique vertex adjacent to  $u, a, b$ . We now have 3 cases to consider:  $\deg(u_1) = 3, \deg(u_1) \geq 5, \deg(u_1) = 4$ . If  $\deg(u_1) = 3$ , then the graph is isomorphic to  $K_4$  which is  $\Delta_4$ . If  $\deg(u_1) \geq 5$ , let  $a, u, b, h_1, h_2$  be five consecutive neighbours of  $u_1$  in counter-clockwise order. If  $b$  is not adjacent to  $h_2$ , then flipping edge  $u_1h_1$  increases  $\deg(b)$  by 1 and thus  $d_G(a, b)$  by 1. If  $b$  is adjacent to  $h_2$ , then  $u_1bh_2$  is a separating triangle that separates  $u, a$  from  $h_1$  (see Figure 4). Therefore, flipping edge  $u_1b$  and  $u_1u$  decreases  $\deg(b)$  by 1 and increases  $\deg(a)$  by 1. Thus, with two flips  $d_G(a, b)$  increases by 2.

Finally, if  $\deg(u_1) = 4$ , then there is unique vertex  $u_2$  adjacent to  $a, u_1, b$ . If  $\deg(u_2) = 3$ , the graph is isomorphic to  $\Delta_5$ . If  $\deg(u_2) \geq 5$  we apply the same argument as when  $\deg(u_1) \geq 5$ . If  $\deg(u_2) = 4$ , we obtain another unique vertex  $u_3$ . This process ends with  $u_{n-3}$ , at which point  $a$  and  $b$  are dominant.

Since  $d_G(a, b)$  increases by at least 1 for one flip and at least 2 for two flips, we note that the total number of flips does not exceed  $d_{\Delta_n}(a, b) - d_G(a, b) = 4n - 4 - (3\deg(a) + \deg(b))$  as required.  $\square$

Using this lemma, Komuro proved the following theorem.

**Theorem 5. (Komuro [5], Theorem 1)** *Any two triangulations with  $n$  vertices can be transformed into each other by at most  $8n - 54$  edge flips if  $n \geq 13$  and at most  $8n - 48$  edge flips if  $n \geq 7$ .*

*Proof.* Given a triangulation  $G$  on  $7 \leq n \leq 12$  vertices, one can prove by contradiction that either  $G$  is one flip from  $\Delta_n$  or there exists an edge  $ab$  where both vertices have degree at least 5 implying that  $d_G(a, b) \geq 20$ . This gives an upper

bound of  $4n - 24$  to convert  $G$  to  $\Delta_n$ , which gives an upper bound of  $8n - 48$  to convert any triangulation to any other via the canonical triangulation. Moreover, for  $n \geq 13$ , either  $G$  is one flip from canonical or there exists an edge  $ab$  where  $a$  has degree at least 6 and  $b$  has degree at least 5. This means that  $d_G(a, b) \geq 23$ . The result follows.  $\square$

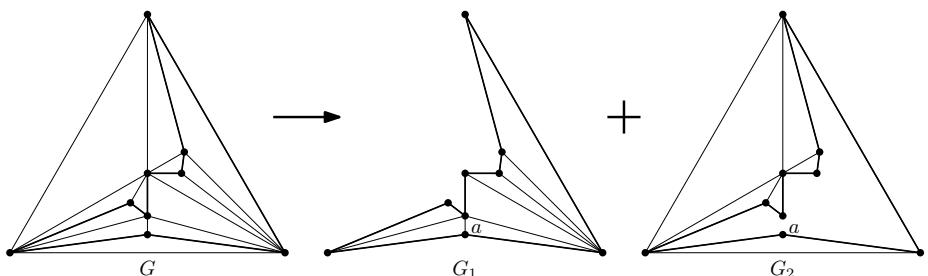
## 4 Mori *et al.*'s Bound

In 2001, Mori, Nakamoto and Ota [6] improved the bound by Komuro to  $6n - 30$ . They used a two-step approach by finding a short path to a strongly connected kernel, which consists of all Hamiltonian triangulations. An  $n$ -vertex triangulation is Hamiltonian if it contains a Hamiltonian cycle, i.e. a cycle of length  $n$ . The general idea of the proof is to find a fast way to make any triangulation Hamiltonian and then use the Hamiltonian cycle to decompose the graph into two outerplanar graphs. These have the following nice property.

**Lemma 6. (Mori *et al.* [6], Lemma 8 and Proposition 9)** *Any vertex  $v$  in a maximal outerplanar graph on  $n$  vertices can be made dominant by  $n - 1 - \deg(v)$  flips.*

*Proof.* If  $v$  is not dominant, there is a triangle  $vxy$  where  $xy$  is not an edge of the outer face. Then we can flip  $xy$  into  $vz$ , where  $z$  is the other vertex of the quadrilateral formed by the two triangles that share  $xy$ . This flip must be legal, since if  $vz$  was already an edge, the graph would have  $K_4$  as a subgraph, which is impossible for outerplanar graphs. Since each such flip increases the degree of  $v$  by one,  $n - 1 - \deg(v)$  flips are both necessary and sufficient.  $\square$

With this property, Mori *et al.* showed that it is possible to quickly transform any Hamiltonian triangulation into the canonical form by decomposing it along the Hamiltonian cycle into two outerplanar graphs.



**Fig. 5.** The decomposition of a Hamiltonian graph  $G$  into two outerplanar graphs  $G_1$  and  $G_2$ . The vertex  $a$  has degree 2 in  $G_2$ .

**Theorem 7. (Mori et al. [6], Proposition 9)** *Any Hamiltonian triangulation on  $n$  vertices can be transformed into  $\Delta_n$  by at most  $2n - 10$  flips, preserving the existence of Hamiltonian cycles.*

*Proof.* Given a Hamiltonian triangulation  $G$  with Hamiltonian cycle  $C$ , we can decompose it into two outerplanar graphs  $G_1$  and  $G_2$ , such that each contains  $C$  and all edges on one side of  $C$ . This is illustrated in Figure 5. Let  $a$  be a vertex of degree 2 in  $G_2$ . We are going to make  $a$  dominant in  $G_1$ . Since  $G$  is 3-connected and  $a$  has no additional neighbours in  $G_2$ , the degree of  $a$  in  $G_1$  is at least 3. Thus by Lemma 6, we can make  $a$  dominant by at most  $n - 4$  flips. Each of these flips is valid, as  $a$  is not connected to anyone in  $G_2$ , except for its neighbours on  $C$ .

Now consider the subgraph  $G'_2 = G_2 \setminus \{a\}$ . Since  $a$  has degree 2 in  $G_2$ ,  $G'_2$  is still outerplanar, so by applying Lemma 6 again we can make a vertex of  $G'_2$  dominant as well, which gives us the canonical triangulation. Since  $G'_2$  has  $n - 1$  vertices and it always has a vertex of degree at least 4 (provided that  $n \geq 6$ ), we need at most  $n - 6$  flips for this. Since we did not flip any of the edges on  $C$ , the theorem follows.  $\square$

This shows that the Hamiltonian triangulations are closely connected, so all we need to figure out is how we can quickly make a triangulation Hamiltonian. Here, we turn to an old result by Whitney [9] that shows that all 4-connected triangulations are Hamiltonian. Since a triangulation is 4-connected if and only if it does not have any separating triangles (cycles of length 3 whose removal disconnects the graph), by removing all separating triangles from a triangulation, we make it 4-connected and therefore Hamiltonian. Fortunately, separating triangles are easy to remove using flips, as the following lemmas show.

**Lemma 8. (Mori et al. [6], Lemma 11)** *In a triangulation with  $n \geq 6$  vertices, flipping any edge of a separating triangle  $D = abc$  will remove that separating triangle. This never introduces a new separating triangle, provided that the selected edge belongs to multiple separating triangles or none of the edges of  $D$  belong to multiple separating triangles.*

*Proof.* Since  $D$  is separating and the newly created edge connects a vertex on the inside to a vertex on the outside, the flip is always legal. Since the flip removes an edge of  $D$ , it is no longer a separating triangle. Now suppose that we flipped  $ab$  to a new edge  $xy$  and introduced a new separating triangle  $D'$ . Then  $D'$  must be  $xyc$ . But since  $n \geq 6$  and our construction so far uses only 5 vertices, one of the faces  $ayc$ ,  $byc$ ,  $axc$ , or  $bxc$  must be a separating triangle as well. This means that either  $ac$  or  $bc$  is an edge that belongs to multiple separating triangles, while  $ab$  only belongs to  $D$ , which contradicts the choice of  $ab$ .  $\square$

**Lemma 9. (Mori et al. [6], Lemma 11)** *Any triangulation on  $n$  vertices can be made 4-connected by at most  $n - 4$  flips.*

*Proof.* We will show that a triangulation can have at most  $n - 4$  separating triangles, the result follows by Lemma 8. The proof is by induction on  $n$ . For the base case, let  $n = 4$ . Then our graph must be  $K_4$ , which has no separating triangles as required. For the induction we can assume that our graph  $G$  has a separating triangle  $T$  which partitions  $G$  into two components  $G_1$  and  $G_2$ . By induction,  $G_1$  and  $G_2$  have at most  $n_1 - 4$  and  $n_2 - 4$  separating triangles, where  $n_1$  and  $n_2$  are the number of vertices in  $G_1$  and  $G_2$ , respectively, including the vertices of  $T$ . Therefore  $G$  can have at most  $n_1 - 4 + n_2 - 4 + 1 = (n_1 + n_2 - 3) - 4 = n - 4$  separating triangles.  $\square$

Now we can prove the main result.

**Theorem 10. (Mori et al. [6], Theorem 4)** *Any two triangulations on  $n$  vertices can be transformed into each other by at most  $6n - 30$  flips.*

*Proof.* The connection between Lemma 9 and Theorem 7 is an old proof by Whitney [9] that any 4-connected triangulation is Hamiltonian. Therefore we can transform any triangulation into the canonical form by at most  $n - 4 + 2n - 10 = 3n - 14$  flips. By looking carefully at the proof of Theorem 7, we see that if the graph is 4-connected, the first vertex (vertex  $a$ ) is guaranteed to have degree at least 4, which brings the bound down to  $3n - 15$  flips to the canonical triangulation and  $6n - 30$  flips between any pair of triangulations.  $\square$

## 5 Bose et al.'s Bound

Mori's bound on the number of flips required to remove all separating triangles in a triangulation is not tight. Recently, Bose et al. [3] showed that  $(3n - 6)/5$  flips suffice and are sometimes necessary to make a triangulation 4-connected. We give a summary of the proof below, the full details can be found in the original paper. The matching lower bound is described in Section 6. First, we divide the edges of the graph into two types; edges that belong to separating triangles and edges that do not. The latter are called *free edges* and they have the following nice property.

**Lemma 11. (Bose et al. [3], Lemma 2)** *In a triangulation  $T$ , every vertex  $v$  of a separating triangle  $D$  is incident to at least one free edge inside  $D$ .*

*Proof.* Consider one of the edges of  $D$  incident to  $v$ . Since  $D$  is separating, its interior cannot be empty and since  $D$  is part of  $T$ , there is a triangular face inside  $D$  that uses this edge. Now consider the other edge  $e$  of this face that is incident to  $v$ .

The remainder of the proof is by induction on the number of separating triangles contained in  $D$ . For the base case, assume that  $D$  does not contain any other separating triangles. Then  $e$  must be a free edge and we are done.

For the induction step, there are two further cases. If  $e$  does not belong to a separating triangle, we are again done, so assume that  $e$  belongs to a separating

triangle  $D'$ . Since  $D'$  is itself a separating triangle contained in  $D$  and containment is transitive, the number of separating triangles contained by  $D'$  must be strictly smaller than that of  $D$ . Since  $v$  is also a vertex of  $D'$ , our induction hypothesis tells us that there is a free edge incident to  $v$  inside  $D'$ . Since  $D'$  is contained in  $D$ , this edge is also inside  $D$ .  $\square$

Separating triangles can be contained in other separating triangles. A *deepest* separating triangle is one that is contained in the maximum number of separating triangles. We will remove all separating triangles by repeatedly flipping an edge of a deepest separating triangle.

**Theorem 12. (Bose et al. [3], Theorem 3)** *A triangulation on  $n \geq 6$  vertices can be made 4-connected using at most  $\lfloor (3n - 6)/5 \rfloor$  flips.*

*Proof.* We prove this using a simple charging scheme; we place one coin on every edge at the start and charge 5 coins per flip. This guarantees that we perform at most  $\lfloor (3n - 6)/5 \rfloor$  flips. To prove that we can actually charge 5 coins per flip, we need two invariants:

- Every edge of a separating triangle has a coin.
- Every vertex of a separating triangle has an incident free edge that is inside the triangle and that has a coin.

We restrict ourselves to flipping edges of a deepest separating triangle  $D$ . This gives us four types of edges we can charge:

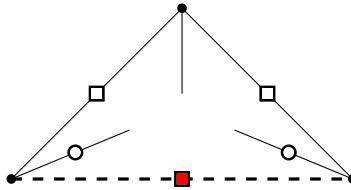
**Type 1 (■).** The flipped edge  $e$ . The flip removes all separating triangles that  $e$  belongs to and does not introduce any new ones, so both invariants are still satisfied if we remove  $e$ 's coin.

**Type 2 (□).** An edge  $e$  of  $D$  that is not shared with any other separating triangle. Again, the flip removes the separating triangle that  $e$  belongs to and does not introduce any new ones, so we can safely charge  $e$ 's coin.

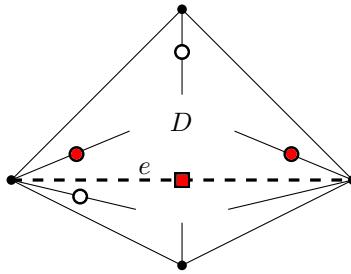
**Type 3 (○).** A free edge  $e$  of a vertex of  $D$  that is not shared with any containing separating triangle. Since the flip removed  $D$  and  $e$  is not incident to a vertex of another separating triangle that contains it, it is no longer required to have a coin to satisfy the second invariant.

**Type 4 (●).** A free edge  $e$  incident to a vertex  $v$  of  $D$ , where  $v$  is an endpoint of an edge  $e'$  of  $D$  that is shared with a non-containing separating triangle  $B$ , provided that we flip  $e'$ . Any separating triangle that contains  $D$  but not  $B$  must share  $e'$  and is therefore removed by the flip. So every separating triangle after the flip that shares  $v$  and contains  $D$  also contains  $B$ . Since the second invariant requires only one free edge with a coin for each vertex, we can safely charge the free edge inside  $D$ , as long as we do not charge the one in  $B$ .

To decide which edge we flip and how we pay for each flip, we distinguish five cases for  $D$ , based on the number of edges shared with other separating triangles and whether any of these triangles contain  $D$ . These cases are illustrated in Figures 6, 7, and 8.



**Fig. 6.** The edges that are charged if the deepest separating triangle does not share any edges with other separating triangles. The flipped edge is dashed and the charged edges are marked with shaded boxes (Type 1), white boxes (Type 2), white disks (Type 3) or shaded disks (Type 4).

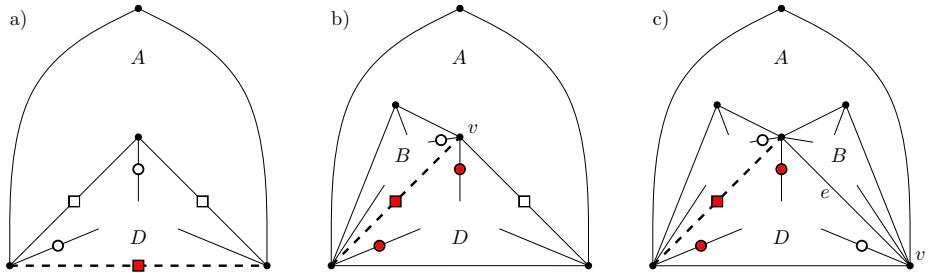


**Fig. 7.** The edges that are charged if the deepest separating triangle only shares edges with non-containing separating triangles

Case 1.  $D$  does not share any edges with other separating triangles (Figure 6). In this case, we flip any one of  $D$ 's edges and charge all of them. Since  $D$  can share at most one vertex with a containing triangle, we charge the remaining two coins from two free edges, each incident to one of the other two vertices.

Case 2.  $D$  does not share any edge with a containing separating triangle, but shares one or more edges with non-containing separating triangles (Figure 7). In this case, we flip one of the shared edges  $e$ . We charge  $e$  and two free edges inside  $D$  that are incident to the vertices of  $e$ . Now consider the quadrilateral formed by  $D$  and the non-containing separating triangle that shares  $e$  with  $D$ . Since  $D$  does not share an edge with a containing separating triangle, at most two vertices of this quadrilateral can be shared with containing separating triangles. Therefore we charge two free edges, each incident to one of the other two vertices, for the last two coins.

Case 3.  $D$  shares an edge with a containing triangle  $A$  and does not share the other edges with any separating triangle (Figure 8a). In this case, we flip the shared edge and charge all of  $D$ 's edges. The vertex of  $D$  that is not shared with  $A$  cannot be shared with any containing triangle and at most one of the vertices of the shared edge can be shared with containing separating triangles that are not removed by the flip, so we charge two free edges incident to the unshared vertices for the remaining coins.



**Fig. 8.** The edges that are charged if the deepest separating triangle shares an edge with a containing triangle

Case 4.  $D$  shares an edge with a containing triangle  $A$  and one other edge with a non-containing separating triangle  $B$  (Figure 8b). In this case, we flip the edge that is shared with  $B$ . Let  $v$  be the vertex of  $D$  that is not shared with  $A$ . We charge the flipped edge, the unshared edge of  $D$  and two free edges inside  $D$  that are incident to the vertices of the flipped edge. We charge the last coin from a free edge in  $B$  that is incident to  $v$ . Every separating triangle that contains  $D$  must contain  $B$  as well and since  $D$  is deepest,  $B$  must be deepest too. Therefore  $v$  cannot be shared with any other separating triangle that contains this free edge and we can safely charge it.

Case 5.  $D$  shares one edge with a containing triangle  $A$  and the other two with non-containing separating triangles (Figure 8c). In this case we flip one of the edges shared with non-containing triangles. The charged edges are identical to the previous case, except that there is no unshared edge any more. Instead, we charge the last free edge in  $D$ . This is allowed, as there is still an uncharged free edge incident to this vertex inside the separating triangle  $B$  whose edge we do not flip.

This shows that we can charge 5 coins for every flip, while maintaining the invariants. As long as our triangulation has a separating triangle, we can always find a deepest separating triangle  $D$ , which must fit into one of the cases above. This gives us an edge of  $D$  to flip and five edges to charge, each of which is guaranteed by the invariants to have a coin. Therefore the process stops only after all separating triangles have been removed.  $\square$

By combining this result with the bound for 4-connected triangulations by Mori *et al.*, we obtain a better bound on the diameter of the flip graph.

**Corollary 1.** *Any pair of triangulations on  $n$  vertices can be transformed into each other by at most  $5.2n - 24.4$  flips.*

*Proof.* Mori *et al.* [6] showed that any two 4-connected triangulations can be transformed into each other by at most  $4n - 22$  flips. By Theorem 12, we can make a triangulation 4-connected using at most  $\lfloor(3n-6)/5\rfloor$  flips. Hence, we can transform any triangulation into any other using at most  $2 \cdot \lceil(3n-6)/5\rceil + 4n - 22 = 5.2n - 24.4$  flips.  $\square$

## 6 Lower Bounds

The best known lower bound on the diameter of the flip graph is by Komuro [5] and is based on the maximum degree of the vertices in the graph.

**Theorem 13. (Komuro [5], Theorem 5)** *Let  $G$  be a triangulation on  $n$  vertices. Then at least  $2n - 2\Delta(G) - 3$  flips are needed to transform  $G$  into the canonical triangulation, where  $\Delta(G)$  denotes the maximum degree of  $G$ .*

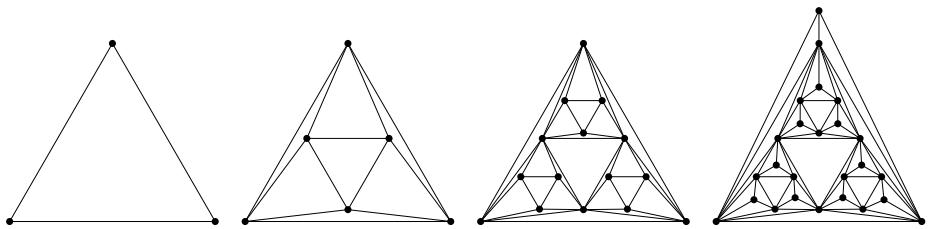
*Proof.* Let  $a$  and  $b$  be the two vertices of degree  $n - 1$  in the canonical triangulation. Each flip increases the degree in  $G$  of either  $a$  or  $b$  by at most one. The only possible exception is the flip that creates the edge  $ab$ , which increases the degree of both vertices by one. Since the initial degree of  $a$  and  $b$  is at most  $\Delta(G)$ , we need at least  $2(n - 1 - \Delta(G)) - 1 = 2n - 2\Delta(G) - 3$  flips.  $\square$

Since there are triangulations that have maximum degree 6, this gives a lower bound of  $2n - 15$  flips. It is interesting that one of the triangulations in the lower bound is the canonical form. This implies that either the lower bound is very far off, or the canonical triangulation is a bad choice of intermediate triangulation. It also means that as long as we use this canonical form, the best we can hope for is an upper bound of  $4n - 30$  flips. Komuro also gave a lower bound on the number of flips required to transform between any pair of triangulations, again based on the degrees of the vertices.

**Theorem 14. (Komuro [5], Theorem 4)** *Let  $G$  and  $G'$  be triangulations on  $n$  vertices. Let  $v_1, \dots, v_n$  and  $v'_1, \dots, v'_n$  be the vertices of  $G$  and  $G'$ , respectively, ordered by increasing degree. Then at least  $\frac{1}{4}D(G, G')$  flips are needed to transform  $G$  into  $G'$ , where  $D(G, G') = \sum_{i=1}^n |\deg(v_i) - \deg(v'_i)|$ .*

*Proof.* Let  $\sigma$  be a mapping between the vertices of  $G$  and  $G'$  and suppose we transform  $G$  into  $G'$  using flips, such that  $v_i \in G$  becomes  $v'_{\sigma(i)} \in G'$ . Since every flip changes the degree of a vertex by one, we need at least  $|\deg(v_i) - \deg(v'_{\sigma(i)})|$  flips to obtain the correct degree for  $v'_{\sigma(i)}$ . However, each flip affects the degrees of 4 vertices, giving a bound of  $\frac{1}{4} \sum_{i=1}^n |\deg(v_i) - \deg(v'_{\sigma(i)})|$  flips. Our actual lower bound is the minimum of this bound over all mappings  $\sigma$ . Mapping every vertex to a vertex with the same rank when ordered by degree (i.e.  $\sigma(i) = i$ ) achieves this minimum.  $\square$

There are stricter lower bounds if we restrict ourselves to the approach used by Mori *et al.*. First, since there are Hamiltonian triangulations with maximum degree 6, the  $2n - 15$  lower bound holds for the transformation of a Hamiltonian triangulation to the canonical one as well. The technique used by Mori *et al.* uses at most  $2n - 10$  flips, so it is only 5 flips removed from the lower bound. There is also a tight lower bound by Bose *et al.* for the number of flips required to make a triangulation 4-connected.



**Fig. 9.** Recursive construction of a triangulation with a large number of edge-disjoint separating triangles

**Theorem 15. (Bose et al. [3], Theorem 5)** *There are triangulations that require  $\lceil (3n - 10)/5 \rceil$  flips to make them 4-connected.*

*Proof.* This bound is based on the recursive construction illustrated in Figure 9. It starts with a single triangle, adds an inverted triangle and connects each vertex to both vertices of the opposing edge. Then it recurses on three of the new triangles incident to the original triangle vertices. In the final step of the recursion, instead of an inverted triangle, a single vertex is added and connected to all three triangle vertices. One vertex is also added to the exterior face, so the original triangle becomes separating as well. This generates a triangulation with  $(3n - 10)/5$  edge-disjoint separating triangles. Since removing all separating triangles is the only way to make a graph 4-connected, this requires at least  $\lceil (3n - 10)/5 \rceil$  flips. This differs less than a single flip from the upper bound of  $\lfloor (3n - 6)/5 \rfloor$  and since the number of flips is necessarily integer, these bounds are tight.  $\square$

This gives a strong indication that we have reached the limit of Mori *et al.*'s approach of converting a triangulation to the canonical form by first making it 4-connected. However, 4-connectedness is only a sufficient condition for Hamiltonicity, not a necessary one. There are a lot of examples (especially small ones) of triangulations that are Hamiltonian, while not being 4-connected. Therefore it might be possible to make a triangulation Hamiltonian with fewer flips. The best known lower bound on this number is by Aichholzer, Huemer and Krasser [1].

**Theorem 16. (Aichholzer et al. [1], Corollary 6)** *There are triangulations on  $n$  vertices that require at least  $(n - 8)/3$  flips to make them Hamiltonian.*

*Proof.* Take any triangulation on  $k \geq 5$  vertices, then add a new vertex to each face and connect it to all three vertices of the face. The new triangulation has  $n = 3k - 4$  vertices. Call the original vertices white and the new vertices black. Since every black vertex is adjacent only to white vertices, a Hamilton cycle would have to pass through at least one white vertex between every pair of black vertex. However, there are  $2k - 4$  black vertices and only  $k$  white vertices, so no such cycle can exist. Furthermore, a single flip can reduce the number of black components by at most one. Therefore at least  $k - 4 = (n - 8)/3$  flips are required to make this triangulation Hamiltonian.  $\square$

This shows that there is still a small improvement possible, but we most likely need new techniques to reduce the gap between the upper and lower bounds further.

## 7 Conclusions

We presented a comprehensive overview including full proofs of the following problem: Given two  $n$ -vertex triangulations, how many edge flips are necessary and sufficient to transform one triangulation into the other? Currently, the best known upper bound is  $5.2n - 24.4$  edge flips, while for the lower bound, there exist pairs of triangulations that require at least  $2n - 15$  edge flips. There remain a number of open problems in this area; we outline a few of them here. The obvious one is to reduce the gap between the upper and lower bound. There are strong indications, as outlined in Section 6, that the upper bound cannot be improved much further with the current techniques. Moreover, it seems counter-intuitive that the canonical triangulation used for all the upper bounds is actually one of the triangulations used to prove the best known lower bound. We believe that a new approach will be needed to improve the current upper bound.

The main graph theoretic property used to prove the lower bounds is vertex degree. The best known lower bound is achieved by showing that converting a triangulation whose maximum degree is 6 into one with two dominant vertices requires at least  $2n - 15$  flips. We feel that other graph theoretic structures will have to be exploited to improve the lower bound, although the question of graph isomorphism is always lurking beneath the surface.

The main open problem in this area is to try to find a way to determine what is the smallest number of flips needed to convert a given triangulation into another. Currently, all of the known approaches can be made to use at least a linear number of flips even when two triangulations differ by one flip. The problem is that all of the approaches first convert a given triangulation into a canonical one. Any algorithm that uses a number of flips that is sensitive to the smallest number of required flips would be a major step forward and could lead to a better understanding of this problem.

## References

- [1] Aichholzer, O., Huemer, C., Krasser, H.: Triangulations without pointed spanning trees. *Comput. Geom.* 40(1), 79–83 (2008)
- [2] Bose, P., Hurtado, F.: Flips in planar graphs. *Comput. Geom.* 42(1), 60–80 (2009)
- [3] Bose, P., Jansens, D., van Renssen, A., Saumell, M., Verdonschot, S.: Making triangulations 4-connected using flips. In: Proceedings of the 23rd Canadian Conference on Computational Geometry, pp. 241–247 (2011); A full version of this paper can be found at arXiv:1110.6473
- [4] Gao, Z., Urrutia, J., Wang, J.: Diagonal flips in labelled planar triangulations. *Graphs Combin.* 17(4), 647–657 (2001)
- [5] Komuro, H.: The diagonal flips of triangulations on the sphere. *Yokohama Math. J.* 44(2), 115–122 (1997)

- [6] Mori, R., Nakamoto, A., Ota, K.: Diagonal flips in Hamiltonian triangulations on the sphere. *Graphs Combin.* 19(3), 413–418 (2003)
- [7] Negami, S., Nakamoto, A.: Diagonal transformations of graphs on closed surfaces. *Sci. Rep. Yokohama Nat. Univ. Sect. I Math. Phys. Chem.* 40, 71–97 (1993)
- [8] Wagner, K.: Bemerkungen zum vierfarbenproblem. *Jahresber. Dtsch. Math.-Ver.* 46, 26–32 (1936)
- [9] Whitney, H.: A theorem on graphs. *Ann. of Math.* (2) 32(2), 378–390 (1931)

# Tangled Thrackles

János Pach<sup>1,\*</sup>, Radoš Radoičić<sup>2,\*\*</sup>, and Géza Tóth<sup>3,\*\*\*</sup>

<sup>1</sup> Ecole Polytechnique Fédérale de Lausanne  
[pach@renyi.hu](mailto:pach@renyi.hu)

<sup>2</sup> Baruch College, City University of New York  
[rados.radoicic@baruch.cuny.edu](mailto:rados.radoicic@baruch.cuny.edu)

<sup>3</sup> Rényi Institute of Mathematics, Budapest  
[geza@renyi.hu](mailto:geza@renyi.hu)

**Abstract.** A *tangle* is a graph drawn in the plane so that any pair of edges have precisely one point in common, and this point is either an endpoint or a point of tangency. If we allow a third option: the common point may be a proper crossing between the two edges, then the graph is called a *tangled thrackle*. We establish the following analogues of Conway’s thrackle conjecture: The number of edges of a tangle cannot exceed its number of vertices,  $n$ . We also prove that the number of edges of an  $x$ -monotone tangled thrackle with  $n$  vertices is at most  $n+1$ . Both results are tight for  $n > 3$ . For not necessarily  $x$ -monotone tangled thrackles, we have a somewhat weaker, but nearly linear, upper bound.

## 1 Introduction

A *drawing* of a simple undirected graph  $G$  is a mapping  $f$  that assigns to each vertex a distinct point in the plane and to each edge  $uv$  a simple continuous curve (i.e., a homeomorphic image of a closed interval) connecting  $f(u)$  and  $f(v)$ , not passing through the image of any other vertex. For simplicity, the point  $f(u)$  assigned to vertex  $u$  is also called a vertex of the drawing, and if it leads to no confusion, it is also denoted by  $u$ . In the same vein, the curve assigned to  $uv$  is called an *edge* of the drawing and it is also denoted by  $uv$ .  $V(G)$  and  $E(G)$  will stand for the vertex set and edge set of the underlying graph  $G$ , as well as of its drawing. Throughout the paper, we assume that no three edges have an interior point in common. Paths and cycles on  $n$  vertices will be denoted by  $P_n$  and  $C_n$ , respectively.

A drawing of  $G$  is a *thrackle* if every pair of edges have precisely one point in common, either a common vertex or a *proper crossing*.<sup>1</sup> In other words, in

---

\* Supported by NSF grant CCF-08-30272, Hungarian Science Foundation EuroGIGA Grant OTKA NN 102029, and Swiss National Science Foundation Grant 200021-125287/1.

\*\* Supported by Hungarian Science Foundation Grant OTKA T 046246.

\*\*\* Supported by Hungarian Science Foundation Grants OTKA K 83767 and NN 102029.

<sup>1</sup> At a proper crossing of two edges, one edge passes from one side of the other edge to its other side.

a thrackle, every two nonadjacent edges cross exactly once, and adjacent edges do not cross. If it creates no confusion, the underlying abstract graph  $G$  is also called a thrackle. In the late sixties, Conway [2], [19], [21] conjectured that every thrackle has at most as many edges as vertices. In spite of considerable efforts, this conjecture is still open. If true, the conjecture would be tight, as any cycle other than  $C_4$  is a thrackle [22]. Lovász, Pach, and Szegedy established the first linear upper bound of  $2n - 3$  on the number of edges in a thrackle on  $n$  vertices, by proving that (the underlying graph of) every bipartite thrackle is actually planar. This bound has been improved since [3], and the current record of  $\frac{167}{117}n < 1.43n$  is due to Fulek and Pach [8]. For related results, see [1], [4], [5], [10], [13], [14], and for applications of thrackles, consult [1], [9].

Assuming the aforementioned conjecture is true, Woodall characterized all thrackles: a graph is a thrackle if and only if it has at most one odd cycle, it contains no  $C_4$ , and each of its connected components contains at most one cycle. This reduces Conway's conjecture to verifying that each graph consisting of two even cycles that share a single vertex is not a thrackle [14], [22]. Erdős resolved the conjecture for thrackles drawn by straight-line edges (see [15] for an elegant proof of Perles, and its relation to some classical work on diameters of point sets [11]). Cairns and Nikolayevsky [6] proved that every outerplanar thrackle has at most as many edges as vertices.<sup>2</sup> In [15], Pach and Sterling verified the conjecture for the case of *x-monotone* thrackles, that is, thrackles whose edges are curves that meet every vertical line in at most one point.

Inspired by recent work on the number of tangencies in families of curves in various settings (cf. [7], [16]), we propose two new variants of thrackles. A drawing of a graph is called a *tangle* if every pair of edges have precisely one point in common: either a common vertex or a *touching point* (*a point of tangency*). In other words, in a tangle, any two nonadjacent edges touch at exactly one interior point, at which the two edges do not cross. We prove the analogue of Conway's conjecture for this variant.

**Theorem 1.** *Let  $n \geq 3$ . The maximum number of edges that a tangle of  $n$  vertices can have is  $n$ .*

A drawing of a graph is called a *tangled thrackle* if every pair of edges have precisely one point in common: either a common vertex, or a point of tangency, or a *proper crossing* (at which an edge passes from one side of the other edge to the other side). In other words, any two nonadjacent edges of a tangled thrackle either touch exactly once, or cross exactly once.

We conjecture the following.

*Conjecture 1.* Every tangled thrackle on  $n$  vertices has  $O(n)$  edges.

We confirm our conjecture in the case of *x-monotone* drawings. Moreover, in this case we have a sharp bound.

---

<sup>2</sup> A thrackle is called *outerplanar* if its vertices lie on a circle whose interior contain all other edges.

**Theorem 2.** Let  $n \geq 4$ . The maximum number of edges that an  $x$ -monotone tangled thrackle of  $n$  vertices can have is  $n + 1$ .

In the general case, the best upper bound we have is slightly superlinear.

**Theorem 3.** Let  $tt(n)$  denote the maximum number of edges that a tangled thrackle of  $n$  vertices can have. Then we have

$$\left\lfloor \frac{7n}{6} \right\rfloor \leq tt(n) \leq cn \log^{12} n,$$

for some constant  $c$ .

## 2 Proof of Theorem 1

Our proof of Theorem 1 is based on the fact that cycle  $C_k$  is a tangle if and only if  $k \in \{3, 4\}$  (see Corollary 2), which stands in sharp contrast to the fact that every cycle, except  $C_4$ , is a thrackle.

First, we prove the following lemma.

**Lemma 1.** If  $G$  is a tangle that contains  $P_5$  or  $C_4$  as a subgraph, then  $G$  has no other edges.

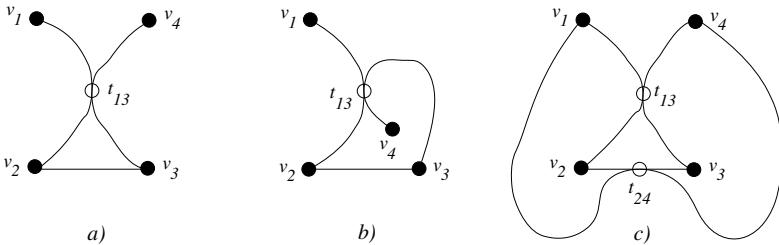
*Proof.* Let  $G$  be a tangle, and let  $H$  be its subgraph isomorphic to either  $P_5$  or  $C_4$ . Let  $v_i, i = 1, \dots, 5$  denote the vertices of  $H$ , and  $e_i = v_i v_{i+1}$  denote the edges of  $H$ . For  $(i, j) \in \{(1, 3), (1, 4), (2, 4)\}$  let  $t_{ij}$  denote the point of tangency of  $e_i$  and  $e_j$ . Note that if  $H \cong C_4$ , then  $v_1$  and  $v_5$  are identical, and  $t_{14}$  is not defined. Let  $\tilde{H}$  be the (drawing of the) planar graph, obtained from  $G$  by introducing new vertices of degree four at the points of tangency  $t_{ij}$ , and defining the edges of  $\tilde{H}$  maximal pieces of the edges of  $G$  that connect two vertices in  $V(\tilde{H})$  and contain no other point from  $V(\tilde{H})$ . If  $H \cong P_5$ , then  $|V(\tilde{H})| = 8$  and  $|E(\tilde{H})| = 10$ . Similarly, if  $H \cong C_4$ , then  $|V(\tilde{H})| = 6$  and  $|E(\tilde{H})| = 8$ . Hence, in both cases,  $\tilde{H}$  has four faces.

Given a face  $f$  of  $\tilde{H}$ , let the *border* of  $f$  be defined as the set  $B(f)$  of all edges  $e_i \in E(\tilde{H})$  that contribute infinitely many points to the boundary of  $f$ . We claim that the borders of the four faces of  $\tilde{H}$  are precisely

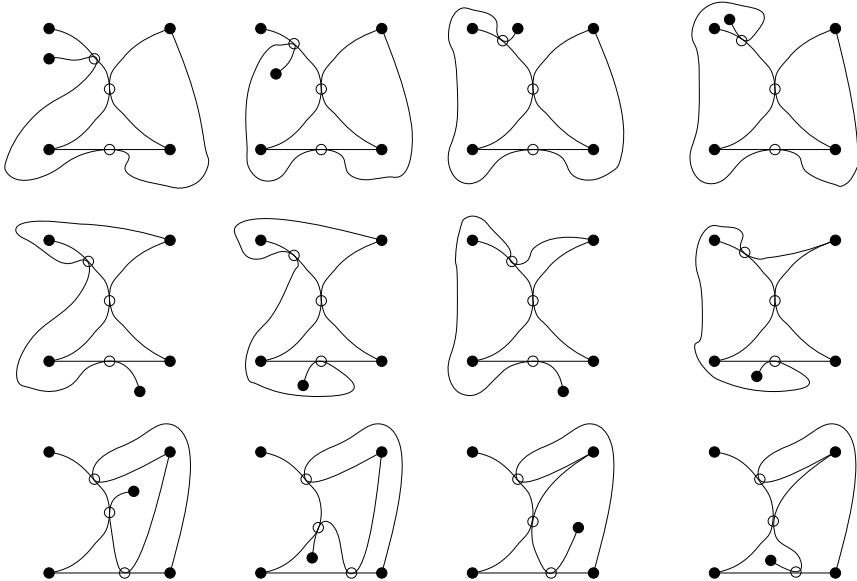
$$\{e_1, e_2, e_3\}, \{e_1, e_2, e_4\}, \{e_1, e_3, e_4\}, \{e_2, e_3, e_4\}. \quad (*)$$

Indeed, this is trivial in the case  $H \cong C_4$ , since the tangle drawing of  $H$  on the sphere has to be topologically equivalent (on the sphere) to Figure 1(c). If  $H \cong P_5$ , then let  $H' \cong P_4$  be the subgraph of  $H$  induced by vertices  $v_i, i = 1, \dots, 4$ . The tangle drawing of  $H'$  has to be topologically equivalent to either (a) or (b) in Figure 1.

According to the order of  $v_4, v_5, t_{24}$ , and  $t_{14}$  along the edge  $e_4$ , and according to which face of  $\tilde{H}$  the vertex  $v_5$  belongs to, we have several cases, depicted in Figure 2. It is easy to check that in each case  $\tilde{H}$  has four faces, and their borders are the triples listed in (\*).



**Fig. 1.** (a), (b) tangled drawings of  $P_4$ ; (c) tangled drawing of  $C_4$



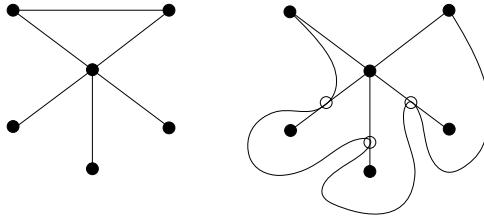
**Fig. 2.** Tangled drawings of  $P_5$

Suppose by contradiction that  $G$  contains another edge  $e$ . Since  $G$  is a tangle,  $e$  has precisely one point in common with each edge  $e_i$ ,  $i = 1, \dots, 4$ . On the other hand,  $e$  must be contained in a face of  $\tilde{H}$ . However, according to (\*), no border  $B(f)$  of a face of  $\tilde{H}$  contains all edges  $e_i$ ,  $i = 1, \dots, 4$ . Using our assumption that no three edges have an interior point in common, this is a contradiction.  $\square$

The following is an immediate corollary to Lemma 1.

**Lemma 2.** *Let  $k \geq 3$ . A cycle  $C_k$  is a tangle if and only if  $k = 3$  or  $4$ .*

Now we are in a position to complete the proof of Theorem 1. Assume  $G$  is a tangle with  $n$  vertices and  $e \geq n+1$  edges. We can assume  $G$  is connected; otherwise, we can find a component of  $G$  with more edges than vertices and continue working with it. Note that  $G$  cannot contain a  $C_4$  as a subgraph; otherwise,  $G$  would contain an additional edge, contradicting Lemma 1. Since  $G$  contains at



**Fig. 3.** A star with an additional edge (on the left) and its tangle drawing (on the right)

least two more edges than its spanning tree,  $G$  has two cycles,  $C$  and  $C'$ . In view of Lemma 2,  $C$  and  $C'$  must be triangles. They cannot share an edge; otherwise,  $G$  would have a  $C_4$ . Since  $G$  is connected, there exists a shortest path  $\ell$  (possibly of length 0) between a vertex  $v$  of  $C$  and a vertex  $v'$  of  $C'$ . Taking a path of length 2 in  $C$  and in  $C'$ , which starts at  $v$  and  $v'$ , respectively, and connecting them by  $\ell$ , we obtain a copy of  $P_5$  in  $G$ . Moreover, the vertices of this path  $P$  span at least one additional edge (e.g., the third edge of  $C$  that does not belong not to  $P$ ). This contradicts Lemma 1.

It is easy to see that Theorem 1 is tight for every  $n \geq 3$ . Indeed, all stars with an additional edge are tangles (see Figure 3: the additional edge can be drawn so that it touches every edge not adjacent to it precisely once).<sup>3</sup>

### 3 Proof of Theorem 2

Let  $G(V, E)$  be an  $x$ -monotone tangled thrackle on  $n$  vertices. For any vertex  $v$ , let  $x(v)$  denote the  $x$ -coordinate of  $v$ . We can also assume that  $G$  has no isolated vertex.

Call vertex  $v$  of  $G$  a *right vertex* (resp. *left vertex*) if for every edge  $uv$  incident to it we have  $x(u) < x(v)$  (resp.  $x(u) > x(v)$ ). Any vertex that is neither a right vertex nor a left one is said to be *two-sided*. Obviously,  $G$  has at most one two-sided vertex. Indeed, if  $v$  and  $v'$  were two such vertices with  $x(v) \leq x(v')$ , then any edge whose right endpoint is  $v$  would be disjoint from all edges whose left endpoint is  $v'$ , contradicting the definition of a tangled thrackle.

We distinguish two cases.

**Case 1.**  $G$  has no two-sided vertex.

Among all edges  $e$  that share the same left (or right) endpoint  $v$ , there is a *highest* edge, that is, one that runs above all other edges  $e$  in a small nonempty open interval  $(x(v), x(v) + \varepsilon)$ . (The *lowest* edge can be defined analogously.)

For each left vertex, delete the highest edge incident to it, and for each right vertex delete the lowest edge. In this way, we removed at most  $n$  edges. Suppose that there is a remaining edge  $uv$  with  $x(u) < x(v)$ . Then  $G$  must have an

<sup>3</sup> In all figures in this paper, vertices marked by empty circles are proper points of tangency, while the original vertices of the graph are represented by black dots.

edge  $uu'$  running above  $uv$ , and an edge  $v'v$  running below it. Clearly, the edges  $uu'$  and  $v'v$  cannot have any point in common, contradicting the definition of a tangled thrackle. Therefore,  $G$  has at most  $n$  edges.

**Case 2.**  $G$  has a two-sided vertex  $v$ .

Replace  $v$  by two vertices,  $v_1$  and  $v_2$ , very close to the original position of  $v$ , such that  $v_1$  is to the left of  $v_2$ . Slightly modify the drawing of  $G$  by reconnecting every edge  $uv \in E(G)$  to the vertex  $v_2$  if  $x(u) < x(v)$  and to  $v_1$  if  $x(u) > x(v)$ , in such a way that every edge  $u_2v_2$  crosses all edges  $v_1u_1$ , and the resulting drawing  $G'$  remains an  $x$ -monotone tangled thrackle.  $G'$  has  $n+1$  vertices, and none of them is two-sided. Therefore, by the previous case, we have  $|E(G')| = |E(G)| \leq n+1$ , as required.

It remains to prove that Theorem 2 is tight, that is, for every  $n \geq 4$  there exist  $x$ -monotone tangled thrackles with  $n$  vertices and  $n+1$  edges.

**Lemma 3.** *Let  $G$  be an  $x$ -monotone tangled thrackle, and let  $uv$  be an edge of  $G$  with  $x(u) < x(v)$  which does not touch any other edge. Suppose that  $uv$  is the lowest among all edges whose left endpoint is  $u$ , and the lowest among all edges whose right endpoint is  $v$ . Let  $G'$  denote the graph obtained from  $G$  by adding two new vertices,  $u'$  and  $v'$ , and replacing the edge  $uv$  by the path  $uv'u'v$  consisting of the edges  $uv'$ ,  $u'v'$ , and  $u'v$ .*

*Then  $G'$  can also be drawn as an  $x$ -monotone tangled thrackle.*

*Proof.* Place  $u'$  above  $u$ , very close to it, and place  $v'$  above  $v$ , very close to it. Draw the new edges  $uv'$ ,  $u'v'$ , and  $u'v$  so that

- (a) they all run very close to the original edge  $uv$ ;
- (b) they all cross every edge that used to cross  $uv$  in  $G$ ;
- (c) every edge whose left endpoint is  $u$  crosses both  $u'v$  and  $u'v'$ ;
- (d) every edge whose right endpoint is  $v$  crosses both  $uv'$  and  $u'v'$ . □

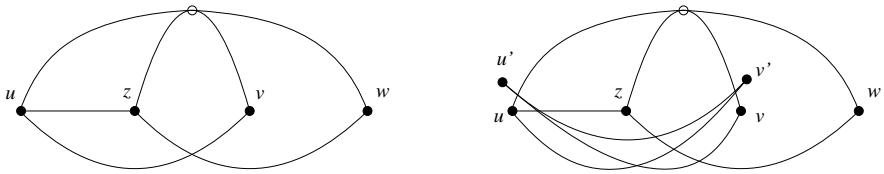
A cycle of length 4 with a diagonal can be drawn as an  $x$ -monotone tangled thrackle. It has  $n=4$  vertices and  $n+1=5$  edges. Repeatedly applying Lemma 3 (first with the edge  $uv$ , then for  $uv'$ , say, etc.), for every even  $n \geq 6$  we obtain an  $x$ -monotone tangled thrackle with  $n$  vertices and  $n+1$  edges. See Figure 4.

Another construction, suggested by Nikolai Hähnle, is depicted on Figure 5. It consists of a cycle of length 4 with a diagonal  $uz$ , plus a number of additional vertices of degree one connected to  $u$ .

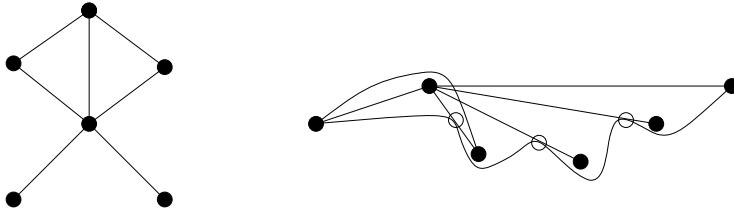
## 4 Proof of Theorem 3

**Lemma 4.** *There are no five curves in the plane with disjoint endpoints such that any two of them have precisely one point in common, a point of tangency, and all of these points are distinct.*

*Proof.* Suppose there exist five such curves. Fix a different point on each of them, and connect each pair of points using two pieces of the corresponding curves that meet at their point of tangency. This way we obtain a planar drawing of  $K_5$ ,



**Fig. 4.**  $C_4$  with diagonal  $uz$ , drawn as an  $x$ -monotone tangled thrackle (on the left); edge  $uv$  has been replaced by path  $uv'u'v$  (on the right).



**Fig. 5.** A graph with  $n$  vertices and  $n + 1$  edges (on the left) and its drawing as an  $x$ -monotone tangled thrackle (on the right)

which may be degenerate in the sense that two adjacent edges may overlap. By slightly perturbing this drawing, if necessary, we can eliminate the common arcs and produce a crossing-free proper drawing of  $K_5$ , contradicting Kuratowski's theorem.  $\square$

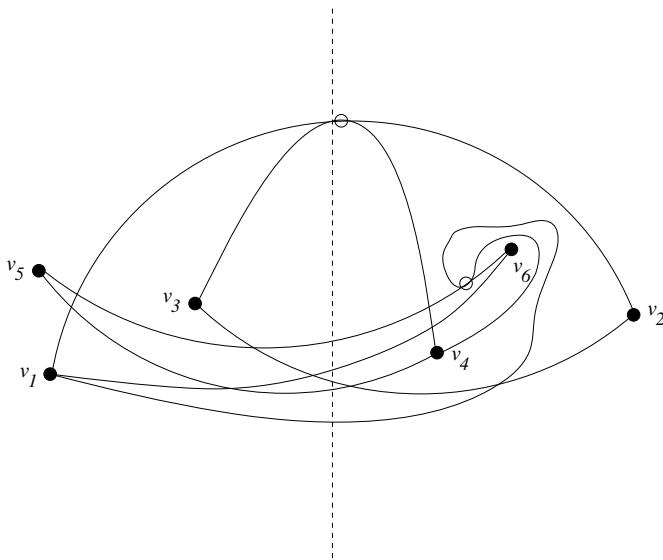
A graph  $G$  drawn in the plane so that any two edges have at most one point in common, which is either a common endpoint or a proper crossing (but not a touching point) is called a *simple topological graph*. Two edges of  $G$  are said to be *disjoint* if they do not share an endpoint or an interior point. We need the following result from [18].

**Lemma 5.** [18] *For any  $k > 0$ , there is a constant  $c_k$  such that every simple topological graph with  $n$  vertices and no  $k$  pairwise disjoint edges has at most  $c_k n \log^{4k-8} n$  edges.*

**Proof of Theorem 3.** Let  $G$  be a tangled thrackle with  $n$  vertices and more than  $c_5 n \log^{12} n$  edges, where  $c_5 > 0$  is the constant that appears in Lemma 5.

Slightly modifying the edges of  $G$  near their points of tangencies, we can attain that no two edges touch each other, and in the process we do not lose any proper crossings. The resulting drawing is a simple topological graph that has no five pairwise disjoint edges. Indeed, the corresponding five edges of  $G$  would be pairwise touching, which contradicts Lemma 4. Thus, the upper bound follows from Lemma 5.

For the lower bound, start with the tangled thrackle drawing of  $C_6$  together with one of its main diagonals, shown in Figure 6. It has the property that there



**Fig. 6.** A tangled thrackle drawing of  $C_6$  with its main diagonal  $v_1v_4$

is a vertical line  $\ell$  that intersects every edge exactly once. Pick a point  $p$  on  $\ell$ . Using an affine transformation, “squash” this drawing parallel to the direction of the  $y$ -axis, to obtain a very “flat” copy of this drawing that lies in a small neighborhood of a horizontal segment. By rotating this drawing about  $p$  through  $k - 1$  different small angles, we can obtain a tangled thrackle. Each copy alone satisfies the conditions, and any pair of edges from different copies cross exactly once. The resulting drawing has  $6k$  vertices and  $7k$  edges, which proves the lower bound.  $\square$

**Remark.** We can modify the notion of tangles and tangled thrackles by allowing several edges to touch one another at the same point.

A drawing of a graph is called a *degenerate tangle* if every pair of edges have precisely one point in common, either a common vertex or a touching point (point of tangency), where several edges may touch one another at the same point. In a *degenerate tangled thrackle*, there is a third option: two edges are also allowed to properly cross each other. It is easy to see that the underlying graph of a degenerate tangle is a planar graph. Therefore, the number of edges of a degenerate tangle of  $n$  vertices is at most  $3n - 6$ . Our proof of Theorem 1 breaks down in this case. Not every degenerate tangle can be redrawn as a tangle (consider, for example, a cycle of length four together with one of its main diagonals).

On the other hand, the proof of Theorem 2 goes through without any change for  $x$ -monotone degenerate tangled thrackles. It yields that any such graph with  $n$  vertices has at most  $n + 1$  edges. We believe that a linear upper bound may hold even if we drop the assumption of  $x$ -monotonicity.

## References

1. Ackerman, E., Fox, J., Pach, J., Suk, A.: On grids in topological graphs. In: Proceedings of the 25th Annual Symposium on Computational Geometry, pp. 403–412. ACM Press (2009)
2. Braß, P., Moser, W., Pach, J.: Research Problems in Discrete Geometry. Springer, New York (2005)
3. Cairns, G., Nikolayevsky, Y.: Bounds for generalized thrackles. *Discrete and Computational Geometry* 23, 191–206 (2000)
4. Cairns, G., McIntyre, M., Nikolayevsky, Y.: The thrackle conjecture for  $K_5$  and  $K_{3,3}$ . In: Towards a Theory of Geometric Graphs, *Contemp. Math.*, vol. 342, pp. 35–54. Amer. Math. Soc., Providence (2004)
5. Cairns, G., Nikolayevsky, Y.: Generalized thrackle drawings of non-bipartite graphs. *Discrete and Computational Geometry* 41, 119–134 (2009)
6. Cairns, G., Nikolayevsky, Y.: Outerplanar thrackles. *Graphs and Combinatorics* 28, 85–96 (2012)
7. Fox, J., Frati, F., Pach, J., Pinchasi, R.: Crossings between Curves with Many Tangencies. In: Rahman, M. S., Fujita, S. (eds.) WALCOM 2010. LNCS, vol. 5942, pp. 1–8. Springer, Heidelberg (2010)
8. Fulek, R., Pach, J.: A computational approach to Conway’s thrackle conjecture. *Computational Geometry: Theory and Applications* 44, 345–355 (2011)
9. Graham, R.L.: The largest small hexagon. *Journal of Combinatorial Theory, Series A* 18, 165–170 (1975)
10. Green, J.E., Ringeisen, R.D.: Combinatorial drawings and thrackle surfaces. In: Graph Theory, Combinatorics, and Algorithms (Kalamazoo, MI, 1992), vol. 2, pp. 999–1009. Wiley-Intersci. Publ., Wiley, New York (1995)
11. Hopf, H., Pannwitz, E.: Aufgabe Nr. 167. *Jahresbericht Deutsch. Math.-Verein.* 43, 114 (1934)
12. Lovász, L., Pach, J., Szegedy, M.: On Conway’s thrackle conjecture. *Discrete and Computational Geometry* 18, 369–376 (1998)
13. Perlstein, A., Pinchasi, R.: Generalized thrackles and geometric graphs in  $\mathbb{R}^3$  with no pair of strongly avoiding edges. *Graphs and Combinatorics* 24, 373–389 (2008)
14. Piazza, B.L., Ringeisen, R.D., Stueckle, S.K.: Subthrackable graphs and four cycles. In: Graph Theory and Applications (Hakone, 1990). *Discrete Mathematics*, vol. 127, pp. 265–276 (1994)
15. Pach, J., Sterling, E.: Conway’s conjecture for monotone thrackles. *American Mathematical Monthly* 118, 544–548 (2011)
16. Pach, J., Suk, A., Treml, M.: Tangencies between families of disjoint regions in the plane. *Computational Geometry: Theory and Applications* 45, 131–138 (2012)
17. Pach, J., Töröcsik, J.: Some geometric applications of Dilworth’s theorem. *Discrete and Computational Geometry* 12, 1–7 (1994)
18. Pach, J., Tóth, G.: Disjoint edges in topological graphs. *Journal of Combinatorics* 1, 335–344 (2010)
19. Ringeisen, R.D.: Two old extremal graph drawing conjectures: progress and perspectives. *Congressus Numerantium* 115, 91–103 (1996)
20. Tóth, G.: Note on geometric graphs. *Journal of Combinatorial Theory, Series A* 89, 126–132 (2000)
21. Erdős, P.: Unsolved problem. In: Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972), pp. 351–363. Inst. Math. Appl., Southend-on-Sea (1972)
22. Woodall, D.R.: Thrackles and deadlock. In: Welsh, D.J.A. (ed.) *Combinatorial Mathematics and Its Applications*, pp. 335–347. Academic Press (1969)

# Open Guard Edges and Edge Guards in Simple Polygons<sup>\*</sup>

Csaba D. Tóth<sup>1,2,\*\*\*</sup>, Godfried T. Toussaint<sup>3,4</sup>, and Andrew Winslow<sup>2,\*\*\*</sup>

<sup>1</sup> University of Calgary, Calgary, AB, Canada

<sup>2</sup> Tufts University, Medford, MA, USA

<sup>3</sup> New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

<sup>4</sup> McGill University, Montreal, QC, Canada

cdtoth@ucalgary.ca, gt42@nyu.edu, awinslow@cs.tufts.edu

**Abstract.** An *open edge* of a simple polygon is the set of points in the relative interior of an edge. We revisit several art gallery problems, previously considered for closed edge guards, using open edge guards. A *guard edge* of a polygon is an edge that sees every point inside the polygon. We show that every simple non-starshaped polygon admits at most one open guard edge, and give a simple new proof that it admits at most three closed guard edges. We also characterize open guard edges using a special type of kernel. Finally, we present lower bound constructions for simple polygons with  $n$  vertices that require  $\lfloor n/3 \rfloor$  open edge guards, and conjecture that this bound is tight.

**Keywords:** art gallery, illumination, visibility, mobile guards.

## 1 Introduction

Let  $P$  be a simply connected closed polygonal domain with  $n$  vertices. Two points  $p, q \in P$  are mutually visible to each other if the closed line segment  $pq$  lies in  $P$ . In a starshaped polygon  $P$ , all points in  $P$  are visible from a single point  $x \in P$ , which is called a *guard point* for  $P$ . The set of all guard points is the *kernel* of  $P$ .

For a set  $S \subseteq P$  of multiple guards, or the trajectories of mobile guards, we adopt the notion of weak visibility [2]. A point  $p \in P$  is (weakly) visible to a set  $S \subseteq P$  if it is visible from some point in  $S$ . If every point  $p \in P$  is (weakly) visible from  $S$ , then  $S$  is a *guard set*.

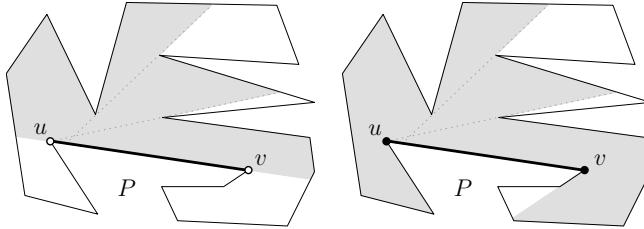
Park *et al.* [9] considered guard sets restricted to (closed) edges of a polygon. They proved that a non-starshaped simple polygon has at most three closed guard edges, and this bound is tight. They also designed an  $O(n)$  time algorithm for finding all closed guard edges in a simple  $n$ -gon. Later, it was shown that

---

<sup>\*</sup> Dedicated to Ferran Hurtado on the occasion of his 60th birthday.

<sup>\*\*</sup> Research of Csaba D. Tóth supported in part by NSERC grant RGPIN 35586.

<sup>\*\*\*</sup> Research of Andrew Winslow supported in part by NSF grants CCF-0830734 and CBET-0941538.



**Fig. 1.** The region visible by an open edge  $uv$  (left) and a closed edge  $uv$  (right) in a simple polygon

a shortest guard segment along the boundary of  $P$ , or anywhere in  $P$  can also be found in optimal  $O(n)$  time [3,4]. A *watchman tour* is a closed curve  $\gamma \subset P$  which is a guard set for  $P$ . Tan [13] gave an  $O(n^5)$  time algorithm for finding a *shortest* watchman tour.

If several guards are available, we are interested in the minimum number of guards that can jointly cover any simple polygon with  $n$  vertices. By a classical result of Chvátal [5], a set of  $\lfloor n/3 \rfloor$  vertex guards are always sufficient and sometimes necessary to cover a simple  $n$ -gon. It is known that  $\lfloor n/4 \rfloor$  closed edge guards are sometimes necessary, and  $\lfloor 3n/10 \rfloor + 1$  are always sufficient [10,11]. It is a longstanding conjecture that  $\lfloor n/4 \rfloor + O(1)$  closed edge guards are always sufficient. However,  $\lfloor n/4 \rfloor$  (open or closed) segment guards are always sufficient and sometimes necessary [8].

Viglietta [14] recently suggested the use of open edge guards for various visibility problems. A *closed edge* includes the endpoints, and an *open edge* does not. See Fig. 1. Intuitively, a closed edge can “see around the corner” if its endpoint is a reflex vertex, while an open edge cannot. In this note, we examine two art gallery problems involving edges of polygons. First, *guard edges* of a polygon; single edges that guard the entire polygon. Then we consider *edge guards*; sets of edges that together guard the entire polygon.

See Table 1 for a summary of our results. We show that every non-starshaped simple polygon admits at most one open guard edge and that this bound is tight. We then use a similar technique to give a new simple proof of the result

**Table 1.** A summary of new and related results

Maximum number of guard edges in an non-starshaped simple polygon		
Guard edge type	Lower bound	Upper bound
Closed	3 [9](Section 4)	3 [9](Section 4)
Open	1 (Section 3)	1 (Section 3)

Minimum number of edge guards needed to guard any simple polygon		
Edge guard type	Lower bound	Upper bound
Closed	$\lfloor n/4 \rfloor$ [8]	$\lfloor 3n/10 \rfloor + 1$ [10,11]
Open	$\lfloor n/3 \rfloor$ (Section 6)	$\lfloor n/2 \rfloor$ (Section 6)

of Park et al. [9] that every non-starshaped simple polygon has at most three closed guard edges. Finally, we show that guarding some simple  $n$ -gons requires at least  $\lfloor n/3 \rfloor$  open edge guards, and that no more than  $\lfloor n/2 \rfloor$  such guards are ever necessary.

## 2 Preliminaries

It is easy to express visibility in terms of shortest paths in a simple polygon (*c.f.*, [1]). Given two points,  $p$  and  $q$ , in a simple polygon  $P$  including its boundary, the geodesic path( $p, q$ ) is the shortest directed path from  $p$  to  $q$  that lies entirely in  $P$ . Points  $p$  and  $q$  see each other iff path( $p, q$ ) is a straight line segment. Every interior vertex of path( $p, q$ ) is a reflex vertex of  $P$ . We characterize weak visibility between a point and an edge in terms of geodesics.

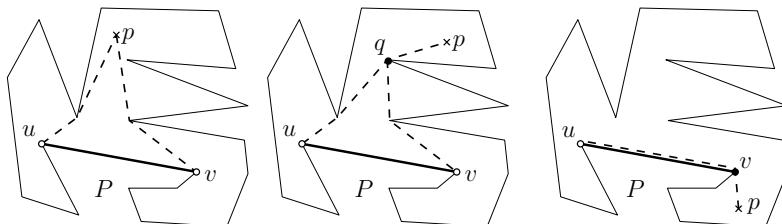
**Lemma 1.** *Let  $p$  be a point inside a simple polygon  $P$ .*

- (a) *Point  $p$  is visible from an open edge  $uv$  iff  $p$  is the only common vertex of path( $p, u$ ) and path( $p, v$ );*
- (b)  *$p$  is visible from a closed edge  $uv$  iff all common vertices of path( $p, u$ ) and path( $p, v$ ) are in  $\{p, u, v\}$ .*

*Proof.* We define a *pseudo-triangle* to be a simple polygon whose boundary consists of three reflex chains.

(a) If  $p$  is the only common vertex of the two geodesics, then  $uv$ , path( $p, u$ ), and path( $p, v$ ) form a pseudo-triangle lying in  $P$  with corners  $p$ ,  $u$  and  $v$ . Each corner of a pseudo-triangle is weakly visible from the opposite side, hence  $p$  is visible from a point in  $uv$  (Fig. 2, left). If  $q \neq p$  is the last vertex in common on the two geodesics path( $p, u$ ) and path( $p, v$ ), then  $q$  is an interior point of every geodesic from  $p$  to any  $w \in uv$ . Hence  $p$  is not visible from any point of the open edge  $uv$  (Fig. 2, middle).

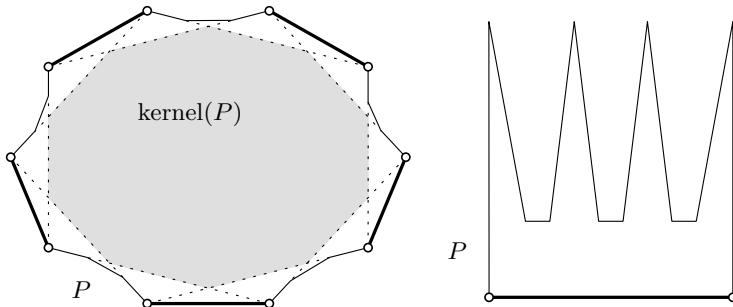
(b) If  $p$  is the only common vertex of the two geodesics, then  $p$  is visible from an interior point of  $uv$  as in part (a). If  $u$  or  $v$  is the only common vertex (apart from  $p$ ) of the two geodesics, then point  $p$  is directly visible from  $u$  or  $v$  (Fig. 2, right). Finally, if  $q \notin \{p, u, v\}$  is a common vertex of the two geodesics, then  $q$  is an interior point of every geodesic from  $p$  to any  $w \in uv$ , and hence  $p$  is not visible from any point of the closed edge  $uv$ .  $\square$



**Fig. 2.** The geodesics path( $p, u$ ) and path( $p, v$ ). Left:  $p$  is the only common vertex of path( $p, u$ ) and path( $p, v$ ). Middle: the common vertices are  $p$  and  $q$ . Right: The common vertices are  $p$  and  $v$ .

### 3 Open Guard Edges

In this section we consider open guard edges. Observe that every edge of a convex polygon is a guard edge, since it lies in the kernel of the polygon; but there may be  $n/4$  or more open guard edges even if all edges are disjoint from the kernel (Fig. 3, left). In this section, we show that every non-starshaped simple polygon has at most one open guard edge. This bound is tight, as shown by the example in Fig. 3, right.



**Fig. 3.** Left: a starshaped  $n$ -gon  $P$  with  $n/4$  open guard edges where the kernel lies in the interior of  $P$ . Right: a non-starshaped polygon with one open guard edge.

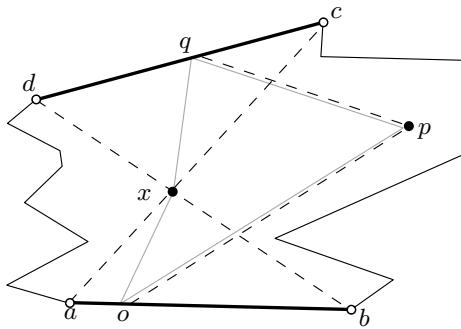
We prove the upper bound by contradiction: we prove that a simple polygon with at least two open guard edges is starshaped. Let  $P$  be a simple polygon, and suppose that edges  $ab$  and  $cd$  are open guard edges. We may assume without loss of generality that  $a, b, c, d$  are in counterclockwise order along the boundary of  $P$  (possibly,  $b = c$  or  $d = a$ ).

**Lemma 2.**  $\text{path}(b, c)$  and  $\text{path}(a, d)$  are disjoint.  $\text{path}(a, c) = ac$  and  $\text{path}(b, d) = bd$  are line segments.

*Proof.* Note that  $ab$ ,  $\text{path}(b, c)$ ,  $cd$ , and  $\text{path}(a, d)$  form a geodesic quadrilateral  $Q$ , i.e. a quadrilateral where each side is a geodesic path. Every geodesic between a point in  $ab$  and a point in  $cd$  lies in  $Q$ . If  $\text{path}(b, c)$  and  $\text{path}(a, d)$  have a common interior vertex  $q$ , then  $a$  or  $b$  is not visible from the open edge  $cd$  by Lemma 1, and so  $cd$  cannot be a guard edge. We conclude that  $\text{path}(b, c)$  and  $\text{path}(a, d)$  are disjoint, and  $Q$  is a simple polygon.

The geodesics  $\text{path}(a, c)$  and  $\text{path}(b, d)$  lie in  $Q$ , as otherwise  $\text{path}(b, c)$  or  $\text{path}(a, d)$  is not a geodesic. So any interior vertex of  $\text{path}(a, c)$  and  $\text{path}(b, d)$  is a vertex of  $Q$ . If an interior vertex of  $\text{path}(a, c)$  is in  $\text{path}(b, c)$ , then  $c$  is not visible from  $ab$ . Similarly, if an interior vertex of  $\text{path}(a, c)$  is in  $\text{path}(a, d)$ , then  $a$  is not visible from  $cd$ . Hence,  $\text{path}(a, c)$  has no interior vertices. Analogous argument shows that  $\text{path}(b, d)$  has no interior vertices, either.  $\square$

**Lemma 3.** The intersection point  $x = ac \cap bd$  is in the kernel of  $P$ .



**Fig. 4.** A schematic of the proof that a simple polygon with two open guard edges must be starshaped. The guard edges are  $ab$  and  $cd$ . The point  $x = ac \cap bd$  is in the kernel of the polygon, since every point  $p \in P$  is visible from  $x$ .

*Proof.* Refer to Fig. 4. It is enough to show that an arbitrary point  $p$  in polygon  $P$  is visible from  $x$ . By Lemma 2,  $ac$  and  $bd$  are diagonals of  $P$ . The triangles  $\Delta(abx)$  and  $\Delta(cdx)$  lie inside  $P$ . If  $p \in \Delta(abx)$  or  $p \in \Delta(cdx)$ , then segment  $px$  lies in the same triangle.

Assume now that  $p$  is outside of both triangles. Since  $ab$  and  $cd$  are open guard edges,  $p$  sees some points in their relative interiors, say  $o \in ab$  and  $q \in cd$ . So all the edges of the quadrilateral  $Q = (o, p, q, x)$  lie in  $P$ , i.e. the boundary of  $Q$  lies in  $P$ . Thus any holes in  $Q$  are also holes in  $P$ , and since  $P$  is simple,  $Q$  must also be simple and lie in  $P$ . Since the vertices  $o$  and  $q$  of  $Q$  are interior points of the edges  $ab$  and  $cd$ , they are convex vertices in  $Q$ . So the diagonal  $px$  of  $Q$  lies in the interior of  $Q$  and  $P$ .  $\square$

**Theorem 1.** *Every non-starshaped simple polygon has at most one open guard edge.*

*Proof.* If a simple polygon has two open guard edges, then it has a nonempty kernel by Lemma 3, and thus is starshaped. So every non-starshaped simple polygon has at most one open guard edge.  $\square$

**Remark.** The upper bound of Theorem 1 does not apply to polygons with holes. Note that an open edge on the boundary of a hole cannot see the entire boundary of the hole. So all open edge guards are on the boundary of the outer polygon. By the result in [9] there are at most 3 *closed* guard edges on the outer boundary of a polygon with holes. Since every open guard edge is a closed guard edge, as well, a polygon with holes has at most 3 open guard edges. This upper bound is tight, as shown by the following simple construction. Let the outer polygon and a hole be two centrally dilated triangles. Then all three open edges of the outer polygon are guard edges.

## 4 Closed Guard Edges

In this section, we extend the argument of the previous section to give a short proof for the following result of Park *et al.* [9].

**Theorem 2 ([9]).** *Every non-starshaped simple polygon has at most three closed guard edges.*

We proceed by contradiction, and show that the presence of four closed guard edges implies that the polygon is starshaped. Let  $P$  be a simple polygon where  $g_1, g_2, g_3$ , and  $g_4$ , in counterclockwise order, are guard edges. Let  $g_1 = ab$  and  $g_3 = cd$  such that  $a, b, c$ , and  $d$  are in counterclockwise order along  $P$ . Note that the vertices  $a, b, c$ , and  $d$  are distinct.

**Lemma 4.** *The geodesics path( $b, c$ ) and path( $a, d$ ) are disjoint; and all vertices of the geodesics path( $a, c$ ) and path( $b, d$ ) are in  $\{a, b, c, d\}$ .*

*Proof.* Consider the geodesic quadrilateral  $Q$  formed by  $ab$ , path( $b, c$ ),  $cd$ , and path( $a, d$ ). Every geodesic between a point in  $ab$  and a point in  $cd$  lies in  $Q$ . Suppose that an interior vertex  $q$  of path( $b, c$ ) is a vertex of path( $a, d$ ). If  $q = a$  or an interior vertex of path( $a, d$ ), then  $b$  is not visible from the closed edge  $cd$  by Lemma 1. Similarly, if  $q = d$ , then  $c$  is not visible from the closed edge  $ab$ . We conclude that path( $b, c$ ) and path( $d, a$ ) are disjoint, and  $Q$  is a simple polygon.

The geodesics path( $a, c$ ) and path( $b, d$ ) lie in  $Q$ , so any interior vertex of path( $a, c$ ) and path( $b, d$ ) is a vertex of  $Q$ . If path( $a, c$ ) and path( $b, c$ ) have a common interior vertex, then  $c$  is not visible from  $ab$ . Similarly, no two geodesics from  $\{a, b\}$  to  $\{c, d\}$  can have any common interior vertex. Hence all interior vertices of path( $a, c$ ) and path( $b, d$ ) are in  $\{a, b, c, d\}$ .  $\square$

**Corollary 1.**

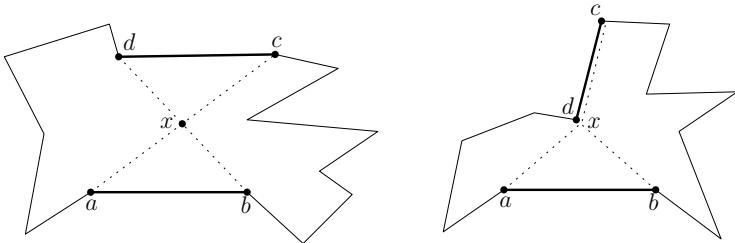
- If  $\{a, b, c, d\}$  is in convex position, then path( $a, c$ ) =  $ac$  and path( $b, d$ ) =  $bd$ .  
*Fig. 5, left.*
- Otherwise suppose w.l.o.g. that  $\text{conv}(\{a, b, c, d\}) = \Delta(abc)$ . Then path( $a, c$ ) =  $(a, d, c)$  and path( $b, d$ ) =  $bd$ . *Fig. 5, right.*

**Lemma 5.** *The intersection point  $x = \text{path}(a, c) \cap \text{path}(b, d)$  is in the kernel of  $P$ .*

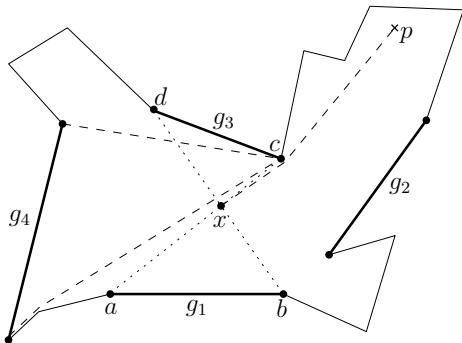
*Proof.* It is enough to show that an arbitrary point  $p$  in polygon  $P$  is visible from  $x$ . By Corollary 1, the triangles  $\Delta(abx)$  and  $\Delta(cdx)$  lie inside  $P$  (one of the triangles may be degenerate). If  $p$  is in  $\Delta(abx)$  or  $\Delta(cdx)$ , then segment  $px$  lies in the same triangle. Refer to Fig. 6.

Assume now that  $p$  is outside of both triangles and, w.l.o.g. it is on the right side of the directed geodesics path( $a, c$ ) and path( $b, d$ ). That is,  $p$  and the guard edge  $g_4$  are on opposite sides of these geodesics.

If  $\text{path}(p, x) = px$ , then  $p$  is visible from  $x$ . Suppose, to the contrary, that  $\text{path}(p, x)$  is not a straight line segment. Assume w.l.o.g. that  $\text{path}(p, x)$  makes



**Fig. 5.** The convex hull of two closed guard edges,  $ab$  and  $cd$ , is either a quadrilateral or a triangle



**Fig. 6.** A schematic of the proof that a simple polygon with four closed guard edges must be starshaped. Suppose that  $g_1 = ab$ ,  $g_2$ ,  $g_3 = cd$ , and  $g_4$  are guard edges. If a point  $p \in P$  is not visible from  $x = \text{path}(a, c) \cap \text{path}(b, d)$ , then we show that  $p$  is also not visible from  $g_2$  or  $g_4$ .

a *right* turn at its last interior vertex  $q$ . Then  $\text{path}(p, d)$  also makes a right turn at  $q$ . Since  $p$  is visible from the guard edge  $cd$ , we must have  $q = c$  by Lemma 1(b). Recall that any geodesic from  $p$  to a point in  $g_4$  crosses both  $\text{path}(a, c)$  and  $\text{path}(b, d)$ . Since we assumed that  $\text{path}(p, x)$  makes a right turn at  $c$ , every geodesic from  $p$  to a point in  $g_4$  also makes a right turn at  $c$ . However,  $c$  is disjoint from  $g_4$ , and by Lemma 1(b),  $p$  is not visible from  $g_4$ , contradicting our initial assumption. We conclude that  $\text{path}(p, x)$  is a straight line segment, and so  $p$  is visible from  $x$ .  $\square$

**Proof of Theorem 2.** If a simple polygon has four closed guard edges, then it has a nonempty kernel by Lemma 5, and thus is starshaped. So every non-starshaped simple polygon has at most three closed guard edges.  $\square$

## 5 Characterizing Open Guard Edges

In this section, we characterize the open guard edges of a simple polygon  $P$  in terms of the left and right kernels of  $P$  (defined below).

## 5.1 Left and Right Kernels

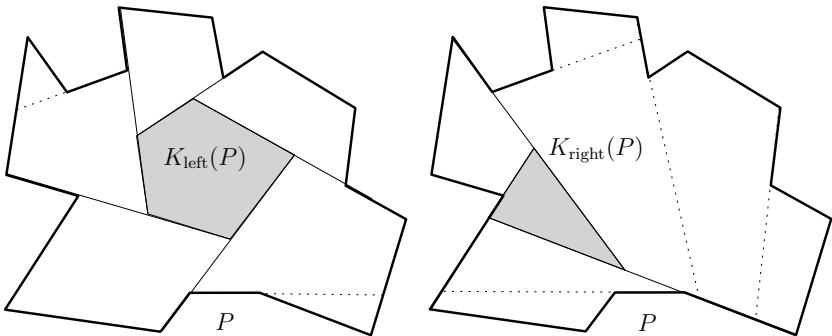
Recall that the set of points from which the entire polygon  $P$  is visible is the *kernel*, denoted  $K(P)$ , which is the intersection of all closed halfplanes bounded by a supporting line of an edge of  $P$  and facing towards the interior of  $P$ . Lee and Preparata [6] designed an optimal  $O(n)$  time algorithm for computing the kernel of simple polygon with  $n$  vertices. We now define a weaker version of the kernel: the *left* and *right kernels* of  $P$ , denoted  $K_{\text{left}}(P)$  and  $K_{\text{right}}(P)$ .

For every reflex vertex  $r$ , we define two polygons  $C_{\text{left}}(r) \subset P$  and  $C_{\text{right}}(r) \subset P$ . Shoot a ray from  $r$  in a direction collinear with the edge incident to  $r$  preceding (resp., following)  $r$  in counterclockwise order; and let  $C_{\text{left}}(r)$  (resp.,  $C_{\text{right}}(r)$ ) be the part of  $P$  on the left (resp., right) of the ray. These polygons have previously been defined in [3]. It is clear that if  $P$  is weakly visible from a set  $S \subset P$ , then  $S$  must intersect both  $C_{\text{left}}(r)$  and  $C_{\text{right}}(r)$  for every reflex vertex  $r$ .

Now we define  $K_{\text{left}}(P)$  as the intersection of polygons  $C_{\text{left}}(r)$  for all reflex vertices  $r$ ; and  $K_{\text{right}}(P)$  as the intersection of polygons  $C_{\text{right}}(r)$  for all  $r$ . See Fig. 7 for an example. Clearly, we have

$$K(P) = K_{\text{left}}(P) \cap K_{\text{right}}(P).$$

By construction, both  $K_{\text{left}}(P)$  and  $K_{\text{right}}(P)$  are convex polygons, whose edges are collinear with some of the edges of  $P$ .



**Fig. 7.** The left and right kernels of a polygon. The dotted lines bound some polygons  $C_{\text{left}}(r)$  and  $C_{\text{right}}(r)$ , but they are not part of the kernel decompositions.

## 5.2 Left and Right Kernel Decompositions

In the following lemma we use two decompositions of a polygon based on its left and right kernels. The *left kernel decomposition* is the decomposition of the polygonal domain  $P$  in the following way: One cell of the decomposition is the left kernel  $K_{\text{left}}(P)$ . The region inside  $P$  but in the exterior of  $K_{\text{left}}(P)$  is decomposed by extending each edge of  $K_{\text{left}}(P)$  in clockwise direction. Refer to Fig. 7. Since  $K_{\text{left}}(P)$  lies on the left side of rays emitted from reflex vertices of

$P$ , the clockwise extension of every edge of  $K_{\text{left}}(P)$  reaches a collinear edge of  $P$ . The right kernel decomposition is defined analogously: one cell is  $K_{\text{right}}(P)$ , and the rest of  $P$  is decomposed by counter-clockwise extensions of the edges of  $K_{\text{right}}(P)$ . Note that if an open edge of  $P$  is disjoint from the left kernel, then it is adjacent to a unique region of the left kernel decomposition. Additionally, each region of the decomposition, except for  $K_{\text{left}}(P)$ , has exactly one common edge with the left kernel.

**Lemma 6.** *An open edge  $e$  of a simple polygon  $P$  is a guard edge of  $P$  iff  $e$  intersects both the left and the right kernels of  $P$ .*

*Proof.* Let  $e = uv$  be an open edge of  $P$ . First assume that  $e$  is disjoint from the left kernel  $K_{\text{left}}(P)$ . Then  $e$  is adjacent to a unique region in the left kernel decomposition of  $P$ . This region is adjacent to a unique edge  $k$  of  $K_{\text{left}}(P)$ , and  $k$  lies on a ray emitted by a reflex vertex  $r$  on  $P$ . Then  $e$  and the polygon  $C_{\text{left}}(r)$  lies on opposite sides of the supporting line of  $k$ . Hence  $e$  does not intersect  $C_{\text{left}}(r)$ , and so it is not a guard edge.

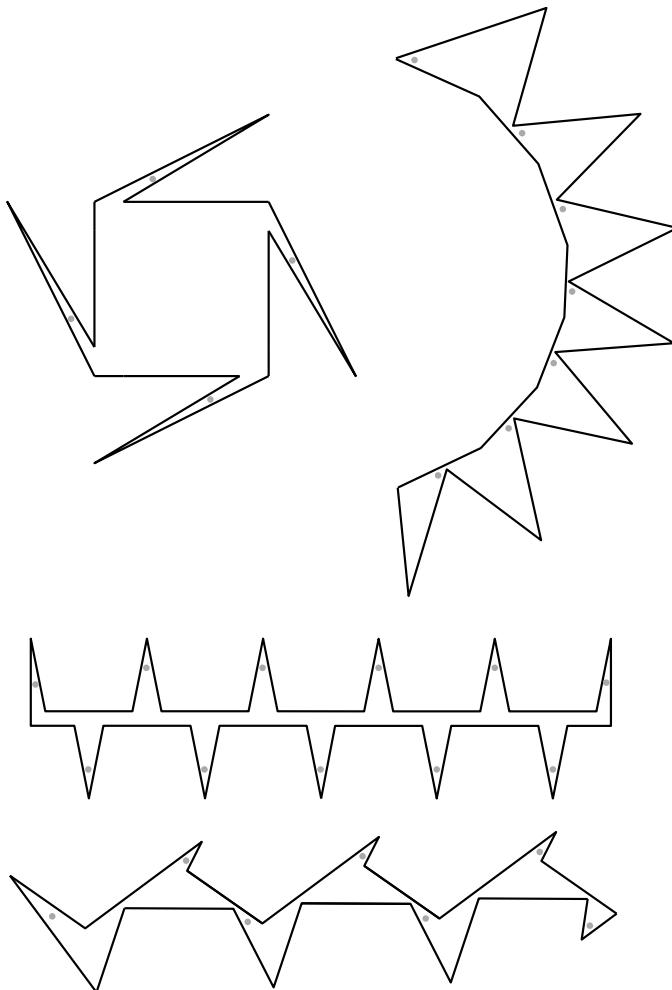
Now assume that  $e = uv$  is not a guard edge, that is, there is a point  $p \in P$  such that  $p$  is not visible from  $e$ . By Lemma 1(a), the geodesics path( $p, u$ ) and path( $p, v$ ) have common interior vertices. Let  $r$  be their last common vertex, which is necessarily a reflex vertex of  $P$ , and assume w.l.o.g. that both geodesics make a right turn at  $r$ . Then  $p \in C_{\text{left}}(r)$ , but  $e$  is disjoint from  $C_{\text{left}}(r)$ . That is,  $e$  is disjoint from the left kernel of  $P$ .  $\square$

## 6 Open Edge Guards

Recall that every simple polygon with  $n$  vertices can be covered by  $\lfloor 3n/10 \rfloor + 1$  closed edge guards, and there are  $n$ -gons that require at least  $\lfloor n/4 \rfloor$  closed edge guards. It turns out that the endpoints of the closed edge guards are crucial for these bounds. Significantly more edge guards may be necessary if the endpoints are removed.

We construct four different infinite families of polygons that require  $\lfloor n/3 \rfloor$  open edge guards for  $n$  vertices. Refer to Fig. 8. The lower bounds for all our constructions can be verified by a standard “hidden point” argument. We hide  $\lfloor n/3 \rfloor$  points (gray dots in Fig. 8) in the interior of a polygon such that each open edge guard sees exactly one such point. That is, each hidden point requires a unique open edge guard, and any set of fewer than  $\lfloor n/3 \rfloor$  open edge guards would miss at least one hidden point.

It is not difficult to see that  $\lfloor n/2 \rfloor$  open edge guards are always sufficient. Partition the set of edges of the polygon into two subsets for which the interior normals of the edges have either a positive or negative  $y$ -component. Each subset of open edges jointly covers the entire polygon. We conjecture this upper bound is weak, and that  $\lfloor n/3 \rfloor$  is the tight bound for the number of open edge guards necessary and sufficient to guard any simple polygon with  $n$  vertices.



**Fig. 8.** Examples of polygons requiring  $n/3$  open edge guards. The gray dots in each polygon indicate a set of points that require a distinct edge to guard each point.

## References

1. Avis, D., Gum, T., Toussaint, G.T.: Visibility between two edges of a simple polygon. *The Visual Computer* 2, 342–357 (1986)
2. Avis, D., Toussaint, G.: An optimal algorithm for determining the visibility of a polygon from an edge. *IEEE Tran. Comput.* C-30, 910–914 (1981)
3. Bhattacharya, B.K., Das, G., Mukhopadhyay, A., Narasimhan, G.: Optimally computing a shortest weakly visible line segment inside a simple polygon. *Comput. Geom. Theory Appl.* 23, 1–29 (2002)
4. Chen, D.Z.: Optimally computing the shortest weakly visible subedge of a simple polygon. *J. Algorithms* 20, 459–478 (1996)
5. Chvátal, V.: A combinatorial theorem in plane geometry. *J. Combin. Theory Ser. B* 28, 39–41 (1975)
6. Lee, D.T., Preparata, F.: An optimal algorithm for finding the kernel of a polygon. *J. ACM* 26, 415–421 (1979)
7. Lu, B.-K., Hsu, F.-R., Tang, C.Y.: Finding the shortest boundary guard of a simple polygon. *Theor. Comp. Sci.* 263, 113–121 (2001)
8. O'Rourke, J.: Galleries need fewer mobile guards: A variation on Chvátal's theorem. *Geometriae Dedicata* 14, 273–283 (1983)
9. Park, J., Shin, S.Y., Chwa, K., Woo, T.C.: On the number of guard edges of a polygon. *Discrete Comput. Geom.* 10, 447–462 (1993)
10. Shermer, T.C.: Recent results in art galleries. *Proc. IEEE* 80, 1384–1399 (1992)
11. Shermer, T.C.: A tight bound on the combinatorial edge guarding problem. *Snapshots of Comp. and Discrete Geom.* 3, 191–223 (1994)
12. Shin, S.Y., Woo, T.: An optimal algorithm for finding all visible edges in a simple polygon. *IEEE Tran. Robotics and Automation* 5, 202–207 (1989)
13. Tan, X.: Fast computation of shortest watchman routes in simple polygons. *Inf. Proc. Lett.* 77, 27–33 (2001)
14. Viglietta, G.: Searching polyhedra by rotating planes, manuscript, arXiv:1104.4137 (2011)

# String-Wrapped Rotating Disks

Joseph O'Rourke

Smith College, Northampton, MA 01063, USA  
[orourke@cs.smith.edu](mailto:orourke@cs.smith.edu)  
<http://cs.smith.edu/~orourke/>

**Abstract.** Let the centers of a finite number of disjoint, closed disks be pinned to the plane, but with each free to rotate about its center. Given an arrangement of such disks with each labeled + or −, we investigate the question of whether they can be all wrapped by a single loop of string so that, when the string is taut and circulates, it rotates by friction all the  $\oplus$ -disks counterclockwise and all the  $\ominus$ -disks clockwise, without any string-rubbing conflicts. We show that although this is not always possible, natural disk-separation conditions guarantee a solution. We also characterize the hexagonal “penny-packing” arrangements that are wrappable.

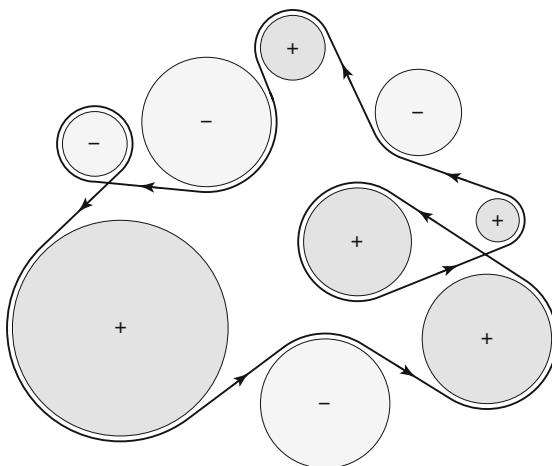
## 1 Introduction

Let  $\mathcal{A}$  be a collection of  $n$  disjoint closed disks in the plane, each labeled + or −, called  $\oplus$ - and  $\ominus$ -disks respectively. We seek to wrap them all in one continuous loop of string so that, were one of the disks rotated by a motor, all the others would spin by friction with the string/belt in a direction consistent with the labeling: counterclockwise (ccw) for  $\oplus$ -disks and clockwise (cw) for  $\ominus$ -disks. See Figure 1. We call a wrapping *proper* if it satisfies these conditions:

1. The string is *taut*: it follows arcs of disk boundaries and disk-disk bi-tangents only.
2. Each disk boundary circle has a positive-length arc in contact with the string. (It is acceptable for the string to wrap around a disk more than once.)
3. One of the two possible circulation directions (i.e., orientations) for the string loop rotates each disk in the direction consistent with its labeling.
4. If the string contacts a point of a disk boundary circle, its circulation there must be in the direction consistent with that disk’s label, i.e., there is no *rubbing conflict*.

We permit the string to cross itself. Indeed such crossings are necessary: for a pair of  $\oplus$ - and  $\ominus$  disks, the string must form a crossing figure-8 shape regardless of their radii and placement. Although the conditions for a proper wrapping are suggested by physical analogy, the pursuit here is not driven by any application. Henceforth a proper wrapping will often be called simply a *wrapping*.

Proper string wrappings are a variation on the “conveyor-belt” wrappings introduced by Abellanas in 2001 (but not published until [1]), and further studied



**Fig. 1.** A proper wrapping of disks with a loop of string: each  $\oplus$ -disk rotates counter-clockwise, each  $\ominus$ -disk clockwise.

in [2]. The belts in Abellanas' model differ from string wrappings in that the disk rotation directions are not pre-specified, and the belt cannot self-cross. These differences considerably change the character of the problem, although the questions raised are analogous.

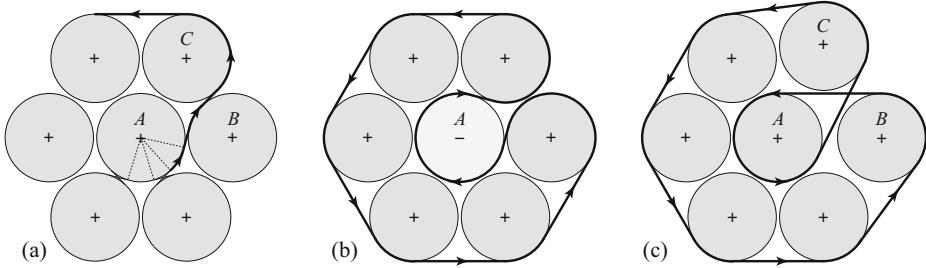
We show that not all arrangements of disks have a proper wrapping, but that various separation conditions guarantee proper wrappings. For example, every collection of unit disks has a proper wrapping when each pair is separated by a distance of 0.31 or more. We focus particularly on “penny-packing” arrangements of congruent disks, where we obtain a characterization of the wrappable arrangements. A characterization of wrappability for arbitrary arrangements remains for future work.

## 2 Unwrappable Arrangements

An example of an unwrappable arrangement is shown in Figure 2(a). It consists of one unit disk surrounded by six others, arranged in a hexagonal penny-packing pattern, except the disks are just barely disjoint. We now argue that this configuration is unwrappable.

The central disk  $A$  must have a positive-length arc of ccw string touching it. Because a taut string can only leave the boundary of a disk along a tangent, the string follows at least the arc between two adjacent tangents. In order for the string to reach another disk, say  $C$ , and contribute a ccw arc, it must first touch another disk,  $B$  in the figure, but now rubbing it in a cw arc. Thus a rubbing conflict is unavoidable.

Without moving the disks, this arrangement can be properly wrapped with a different pattern of  $\pm$  labels. For example, reversing the central disk  $A$  enables



**Fig. 2.** (a) An unwrappable arrangement of seven unit  $\oplus$ -disks. (b) Proper wrapping with  $A$  an  $\ominus$ -disk. (c) Proper wrapping with  $B$  and  $C$  displaced slightly.

a proper wrapping: Figure 2(b). Indeed all other  $\pm$  patterns of labels (except all  $-$ ) in this example are wrappable. Retaining the original  $+$  labels but moving two disks slightly also permits the configuration to be wrapped: Figure 2(c).

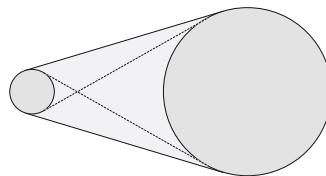
### 3 Separation Conditions

The primary impediment to a proper wrapping is the 4th no-rubbing-conflicts condition. Figure 2(c) indicates that disk-separation conditions may suffice to ensure the existence of a proper wrapping, as separation of the disks separates their tangents and avoids unwanted rubbings. In this section we offer three straightforward conditions that ensure a proper string wrapping exists.

#### 3.1 Connected Hull-Visibility Graph

Define two disks to be *hull-visible* to one another (a symmetric relation) if and only if the (closed) convex hull of the disks does not intersect any other disk; see Figure 3. If two disks can see one another in this sense, then none of their four bi-tangents are blocked (or even touched) by any other disk.

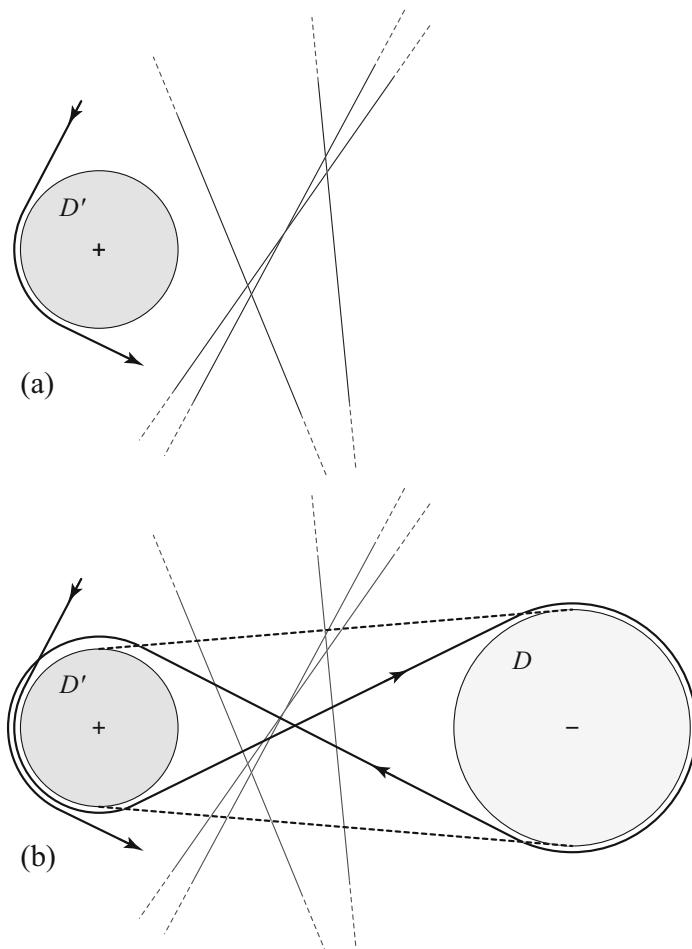
For an arrangement  $\mathcal{A}$  of disks, define their *hull-visibility graph*  $G_V(\mathcal{A}) = G_V$  to have a node for each disk, and an arc connecting two disk nodes if and only if the disks are hull-visible to one another. Call the hull of a pair of disks connected in  $G_V$  to the *edge corridor* for that edge.



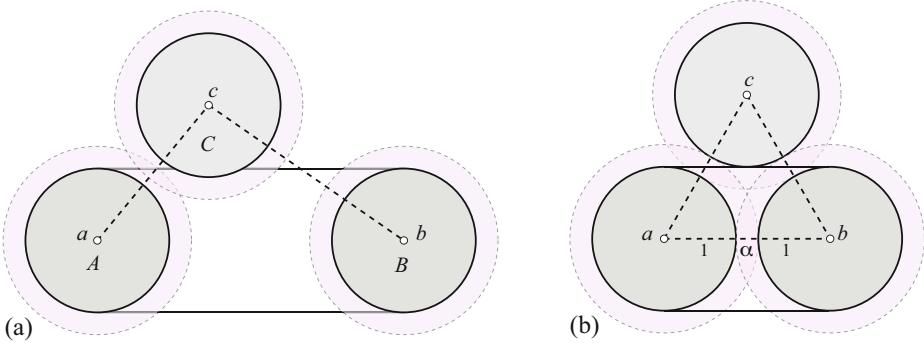
**Fig. 3.** Two disks are visible to one another if their hull does not intersect any other disk

**Lemma 1** ( $G_V$ ). *If  $G_V(\mathcal{A})$  is connected, then there is a proper wrapping of  $\mathcal{A}$ .*

*Proof.* The proof is by induction, with the hypothesis that the string lies within the union of the edge corridors for all the edges in  $G_V$ . Let  $H_k$  be a connected component of  $G_V$  that includes exactly  $k$  nodes, and assume the corresponding  $k$  disks have been properly wrapped with string loop  $S'$  by the induction hypothesis. Let  $D$  be a disk whose node is not in  $H_k$ , but which is connected by an edge  $e = (D, D')$  to  $H_k$ . Because  $e \in G_V$ ,  $D$  can see  $D'$ , and so the edge corridor of  $e$  is not blocked by any disk. It is possible that the edge corridor is crossed by segments of  $S'$  (see Figure 4), but none of those segments can intersect  $D$ , because by hypothesis  $S'$  remains within edge corridors, which  $D$  cannot intersect.



**Fig. 4.** (a)  $D'$  is a disk whose node is in  $H_k$ . (b)  $D$  can connect to  $D'$  inside their corridor by wrapping once around  $D'$ .



**Fig. 5.** (a) Whenever  $C$  blocks (intersects) the  $(A, B)$  corridor,  $c$  is closer to  $a$  and  $b$  than  $|ab|$ . (b) The limiting situation.

$S'$  must touch  $D'$  in a positive-length arc, oriented consistent with the  $\pm$  label of  $D'$ . Regardless of the position of this arc with respect to  $D$ , it is possible to wrap the string  $S$  around  $D'$  to reach the relevant bi-tangency points, and then follow those tangents within the corridor of  $e$  out to wrap around  $D$ , regardless of the sign labels of  $D'$  and  $D$ . Thus we have incorporated  $D$  into  $H_{k+1}$ , maintaining the hypothesis that the string  $S$  lies in the edge-corridor union.  $\square$

The conditions of this lemma are by no means necessary for the existence of a proper wrapping:  $G_V$  for the configuration in Figure 2(b) is completely disconnected—seven isolated nodes—and yet it can be properly wrapped.

### 3.2 Unit Disks Halo

The sufficiency condition of Lemma 1 is a global property of the arrangement  $\mathcal{A}$  of disks, not immediately evident upon inspection. Next we explore local separation conditions that allow us to conclude that  $G_V$  is connected.

Define an  $\alpha$ -halo,  $\alpha > 0$ , for a disk  $D$  of radius  $r$  to be a concentric disk  $D'$  of radius  $(1 + \alpha)$  such that no other disk of  $\mathcal{A}$  intersects  $D'$ .

**Lemma 2 (Unit Disks).** *Let  $A$ ,  $B$ , and  $C$  be three unit disks with centers at  $a$ ,  $b$ , and  $c$  respectively, each with  $\alpha$ -halos for  $\alpha = 4/\sqrt{3} - 2 \approx 0.31$ . Then, if  $C$  intersects the  $(A, B)$  corridor,  $c$  is closer to  $a$  and to  $b$  than is  $a$  to  $b$ :  $|ac| < |ab|$  and  $|bc| < |ab|$ .*

*Proof.* Figure 5(a) illustrates the claim of the lemma. The calculation of  $\alpha$  follows from the limiting configuration shown in Figure 5(b), where  $\triangle abc$  is equilateral and  $(2 + \alpha)\frac{\sqrt{3}}{2} = 2$ . In this configuration, if  $c$  is moved left or right along a horizontal, so that it still just intersects the  $(A, B)$  corridor, its  $\alpha$ -halo intersects  $A$  or  $B$  respectively.  $\square$

**Theorem 1 (Unit Disks).** *An arrangement  $\mathcal{A}$  of unit disks with  $\alpha$ -halos,  $\alpha = 4/\sqrt{3} - 2$ , has a connected visibility graph  $G_V(\mathcal{A})$ , and so can be properly wrapped.*

*Proof.* We show that  $G_V$  is connected by constructing its minimal spanning tree via Prim's algorithm. At any stage in that algorithm, a subset of the nodes have been connected into a tree  $T'$ . Then a shortest edge  $e = (a, b)$  is selected such that  $a \in T'$  and  $b \notin T'$ , and  $T'$  is updated to  $T = T' \cup (a, b)$ . (This addition of  $e$  cannot create a cycle because  $b \notin T'$ .)

In our situation,  $e = (A, B)$ , with the edge length the distance between the disk centers,  $|ab|$ . (We will use  $A$  and  $a$  interchangeably to indicate the nodes of  $G_V$ .) Now we claim that the corridor for  $e$  cannot be intersected by any other disk in  $\mathcal{A}$ . Suppose to the contrary that  $C$  obstructs the  $(A, B)$  corridor. Then Lemma 2 says that  $|ac| < |ab|$  and  $|bc| < |ab|$ . Consider two cases. First, suppose  $c \in T'$ . Then because  $|cb| < |ab|$ , we have identified a shorter edge between a node  $c$  in  $T'$  and a node  $b$  not in  $T'$ . Second, suppose  $c \notin T'$ . Then because  $|ac| < |ab|$ , we have again identified a shorter edge between  $a \in T'$  and  $c \notin T'$ . Both possibilities contradict the choice of  $(a, b)$  by the algorithm. Therefore, Prim's algorithm will indeed construct a spanning tree of  $G_V$ . Knowing that  $G_V$  is connected, we can apply Lemma 1 to conclude that  $\mathcal{A}$  can be properly wrapped.  $\square$

### 3.3 Arbitrary Radii Halo

For disks of different radii, we define the distance between them to be the distance between their bounding circles (rather than their centers). The assumption in Theorem 1 that all disks are congruent can be removed at the cost of a significant increase in the value of  $\alpha$ .

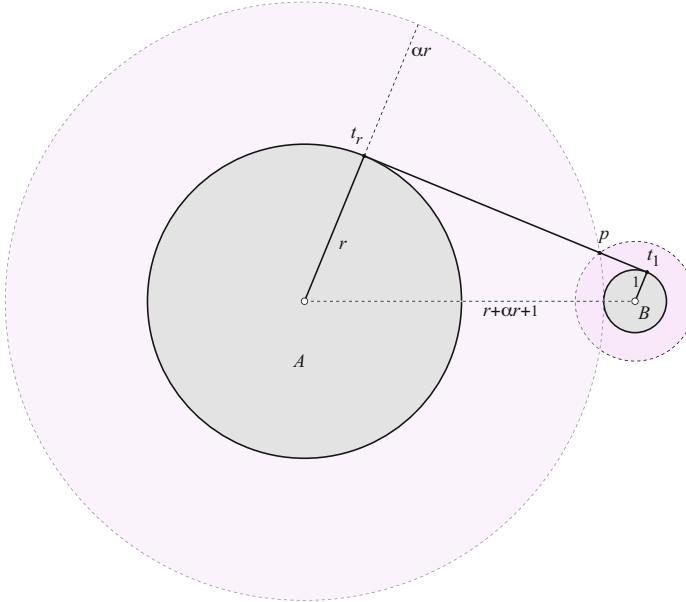
Note that, for disks of arbitrary radii, we must allow for a disk radius to be arbitrarily small, effectively a point regardless of  $\alpha$ . Thus the strategy to avoid blockage of a  $(A, B)$  corridor is to cover it entirely with the  $\alpha$ -halos of  $A$  and  $B$ , for no  $C$  can penetrate these halos by definition.

**Lemma 3 ( $\alpha=1$  Halo).** *The conclusion of Lemma 2 holds for three disks  $A$ ,  $B$ , and  $C$  of arbitrary radius if  $\alpha = 1$ .*

Without loss of generality we may assume that  $A$  has a radius  $r > 1$ ,  $B$  has unit radius, and  $C$  has an arbitrarily small radius. The only proof I have found for this lemma is a brute-force computation of the coverage of the corridor by the  $\alpha$ -halos of  $A$  and  $B$ , as illustrated in Figure 6. The calculation shows that  $\alpha$  ranges from  $2/3$  for  $r = 1$  and approaches but never exceeds 1 as  $r \rightarrow \infty$ . Because the calculation is tedious and the result of limited interest, we leave it as a claim.

The analog of Theorem 1 follows immediately by an identical proof, only invoking Lemma 3 rather than Lemma 2:

**Theorem 2 ( $\alpha=1$  Halo).** *Any arrangement  $\mathcal{A}$  disks with  $\alpha$ -halos,  $\alpha = 1$ , has a connected visibility graph  $G_V(\mathcal{A})$ , and so can be properly wrapped.*



**Fig. 6.** Covering the corridor for  $(A, B)$  by  $\alpha$ -halos

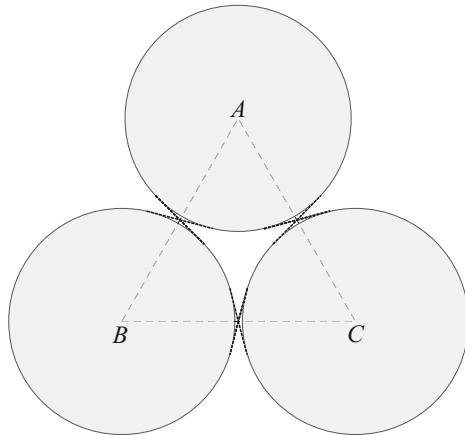
## 4 Penny Arrangements

The visibility and separation conditions in the preceding section guarantee the existence of sufficiently many unblocked bi-tangents to support a proper wrapping. Ideally we would like precise characterization of the wrappable arrangements. In some sense this is easily obtained: an arrangement is wrappable if and only if a certain graph representing the bi-tangents has a spanning cycle. But this is unsatisfying: we would prefer a characterization in terms of more local, or at least easily recognizable, characteristics of the arrangement. I have only achieved this for the penny-packing arrangements epitomized by Figure 2(a,b), as described in this section. In some sense these arrangements are the obverse of well-separated arrangements, for nearly all bi-tangents are blocked.

### 4.1 Characterization Theorem

Let  $H$  be the *hexagonal lattice*, also known as the equilateral triangle lattice (wallpaper group  $p6m$ ), with each edge of each triangle of unit length. We can think of  $H$  as an infinite (plane) graph.

Disks of radius  $\frac{1}{2} - \varepsilon$  are centered on some finite subset of the points of  $H$ , where  $\varepsilon > 0$  is small. Any  $\varepsilon < \frac{1}{2} - \sqrt{3}/4 \approx 0.07$  suffices to ensure that the only tangents between adjacent disks are the “cross” X-tangents; see Figure 7. Each disk has up to six *neighbors* following the edges of  $H$ .



**Fig. 7.** For a unit equilateral underlying lattice, the disk radii are nearly  $\frac{1}{2}$

We will define four subgraphs of  $H$ , three to describe the characterization theorem and one more used in the proof:  $G$ ,  $G^+$ ,  $G^-$ , and  $G_T$ .

Define the plane graph  $G \subset H$  to have a node for each disk, and an arc for each pair of disks that are neighbors in  $H$ . See Figures 8 and 9.

We restrict attention to a subclass of all possible arrangements of disks on  $H$ . Define a *penny arrangement* of disks to be one whose disk radii are close to  $\frac{1}{2}$  in the sense discussed previously, and which satisfies two additional assumptions. Let  $\partial G$  be the outer face of  $G$ .

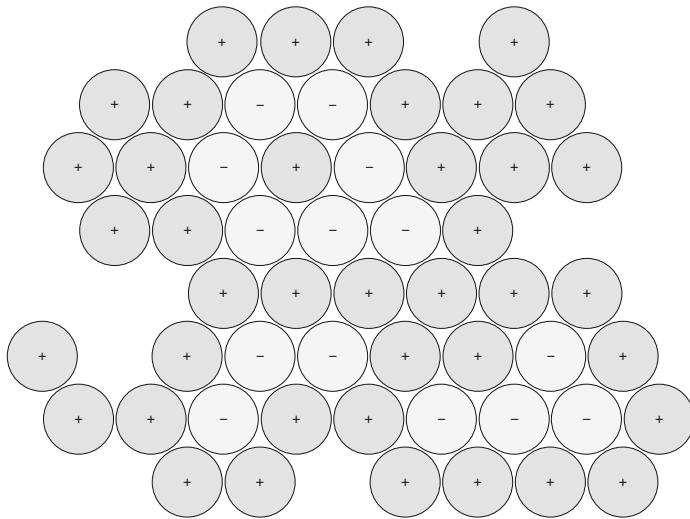
1.  $G$  is connected and has no holes, in the sense that every lattice point inside  $\partial G$  is a node of  $G$ .
2. Every node of  $\partial G$  is an  $\oplus$ -disk.

The example in Figure 8 satisfies these assumptions. I believe these assumptions are not essential, but they allow a more transparent proof. Relaxing these assumptions will be discussed at the end of Section 4.2.

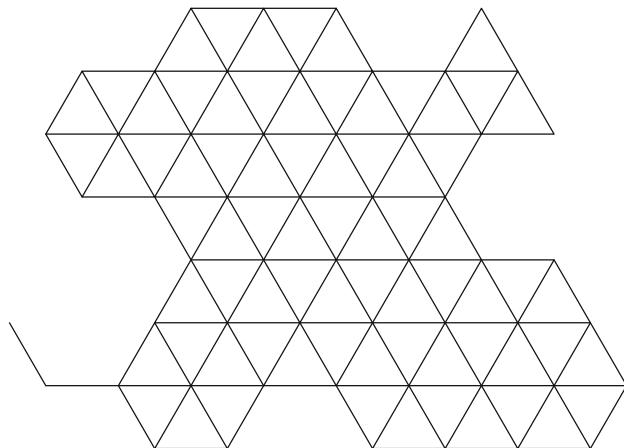
Now we define  $G^+$  and  $G^-$ .  $G^+$  is the subgraph of  $G$  restricted to nodes corresponding to  $\oplus$ -disks, and similarly  $G^-$  is the subgraph restricted to  $\ominus$ -disks. In general both  $G^+$  and  $G^-$  consist of several components. Each component of  $G^+$  is adjacent to a component of  $G^-$  via at least one edge of  $G$  (because  $G$  is connected). See Figure 10.

We now introduce notation that will permit us to characterize the wrappable penny arrangements in terms of structural constraints on  $G^+$  and  $G^-$ . We phrase these definitions in terms of  $G^+$  but the same hold for  $G^-$ .

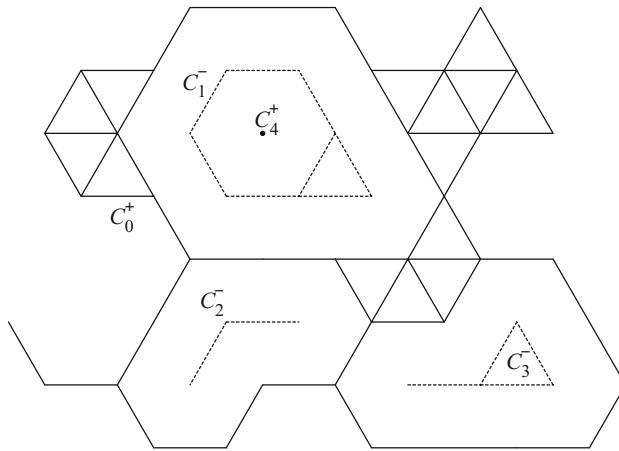
A *hole* in a component  $C^+$  of  $G^+$  is a face that contains a lattice point not occupied by a  $\oplus$ -disk. ( $C_0^+$  in Figure 10 has three holes;  $C_1^-$  has one hole.) A *gate edge* of a hole in  $C^+$  is one whose removal joins the hole with another hole of  $C^+$ , or with the exterior of  $C^+$ . A hole of  $C^+$  is *shallow* if it can be connected to the exterior of  $C^+$  by removal of a sequence of gate edges. Figure 11(b) shows an example.



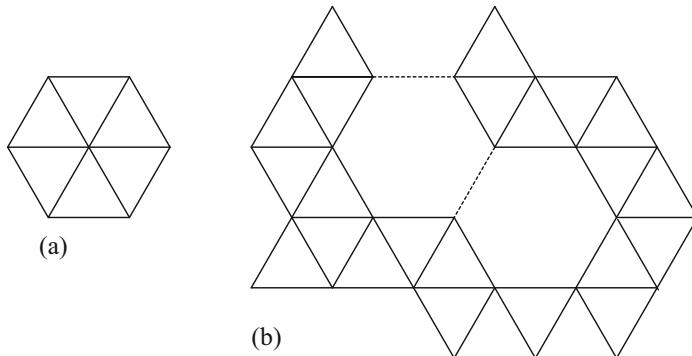
**Fig. 8.** A penny arrangement on a hexagonal lattice  $H$



**Fig. 9.** The graph  $G \subset H$  recording all disk adjacencies



**Fig. 10.**  $G^+$  and  $G^-$  for the arrangement in Figure 8.  $G^+$  consists of two components,  $C_0^+$  and  $C_4^+$ , and  $G^-$  consists of three components,  $C_1^-$ ,  $C_2^-$ , and  $C_3^-$ .



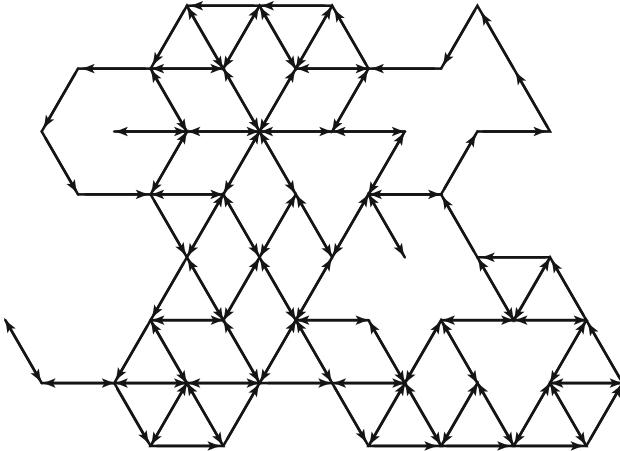
**Fig. 11.** (a) The hexagon graph. (b) A component (of  $G^+$  or  $G^-$ ) whose holes are shallow. The two gate edges are shown dashed.

Call the graph shown in Figure 11(a) the *hexagon* graph; it coincides with  $G^+$  for Figure 2(a). The characterization is in terms of these properties:

**Property (1):** No component of  $G^+$  or  $G^-$  contains the hexagon graph as a subgraph.

**Property (2):** All holes of every component of  $G^+$  and  $G^-$  are shallow.

**Theorem 3 (Penny Arrangements).** *A penny arrangement  $\mathcal{A}$  of disks is wrappable if and only if  $G^+$  and  $G^-$  satisfy Properties (1) and (2).*



**Fig. 12.**  $G_T^+$  for the arrangement in Figure 8.  $\partial G$  includes a counterclockwise cycle.

## 4.2 Proof of Theorem 3

Although I find it natural to see Theorem 3 by wrapping each component of  $G^+$  and  $G^-$  separately, and using the two properties to join the wrappings, this seems to lead to an awkward proof. So here we take a different approach using a tangency graph, which leads to a less natural wrapping, but a simpler a proof.

The key observation is that the string can not wrap directly between two adjacent disks of the same label, unless they are on  $\partial G$ . The reason is that the desired tangent touches a third disk, unless that third disk is not present because it would be outside of  $G$ . Examining Figure 7 again shows that only two oppositely labeled adjacent (internal) disks can directly wrapped one to the other. This suggests defining a directed *tangency graph*  $G_T$  to include for each edge of  $G$  connecting a  $\oplus$ -disk to an  $\ominus$ -disk, an edge in both directions, and, for each edge of  $\partial G$ , the ccw-directed edge between the connected  $\oplus$ -disks. (Recall our assumption that the nodes of  $\partial G$  are all  $\oplus$ -disks.) See Figure 12.  $G_T$  simply records all the bi-tangents that might be followed disk-to-disk. Note that all the edges of  $G_T$  that are not connecting two nodes of  $\partial G$  are bidirectional, connecting  $\oplus$ - and  $\ominus$ -disks, and that the only edges of  $G_T$  connecting like-labeled disks are those on  $\partial G$ .

A *directed spanning cycle* in a directed graph is a collection of oriented arcs that touch each node of the graph and form a cycle. Both nodes and arcs may be visited several times by the cycle. Here is the nearly obvious characterization mentioned previously.

**Lemma 4 ( $G_T$ ).** *A penny arrangement is wrappable if and only if  $G_T$  contains a directed spanning cycle.*

*Proof.* Suppose  $G_T$  contains a directed spanning cycle  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m)$ . We incrementally construct a string that turns each disk. Assume we have a string  $S_{k-1}$  that corresponds to  $(\vec{e}_1, \dots, \vec{e}_{k-1})$ ; we seek to extend it to include  $\vec{e}_k = (A, B)$ , where  $A$  and  $B$  are the two disks connected by the tangent corresponding to  $\vec{e}_k$ . First, wrap the string around  $A$  consistent with its label so that it reaches the tangent departure point. Then follow the tangent from  $A$  to  $B$ . By definition of  $G_T$ , this necessarily works, in that the tangent is consistent with the  $\pm$  labels of the disks and does not rub against any other disk. See Figure 13(a,b). The new string  $S_k$  is now poised for the next extension. This establishes the “if”-direction of the lemma claim.

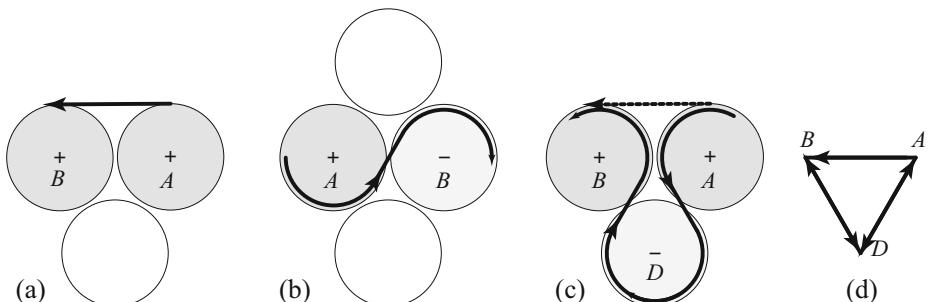
For the “only if”-direction, suppose that there is a proper wrapping of the arrangement by a string  $S$ . Then it is straightforward to map  $S$  to the tangents it follows, and so to a corresponding spanning cycle in  $G_T$ .  $\square$

Now we connect the cycle in  $G_T$  to the structural properties of  $\mathcal{A}$ .

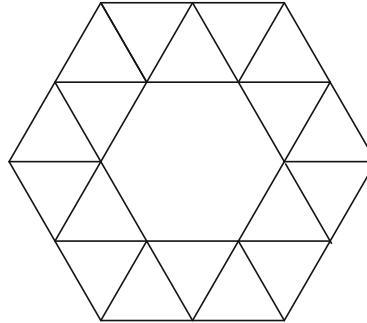
**Lemma 5 (Props.1&2).**  $G_T$  contains a directed spanning cycle for a penny arrangement  $\mathcal{A}$  if and only if  $\mathcal{A}$  satisfies Properties (1) and (2),

*Proof.* Assume first that  $\mathcal{A}$  satisfies the two properties. By the definition of a penny arrangement,  $G_T$  has a counterclockwise cycle including every node of  $\partial G$ . By Property (2) applied to the outermost component of  $G^+$ , there must be at least one edge  $\vec{e} = (A, B)$  of this cycle representing a gate edge, and so whose endpoints are connected to a  $\ominus$ -disk  $D$ . See Figure 13(c,d).

We now argue that the bidirectional arcs of  $G_T$  connect all the disks in  $G$ . Property (1) shows that every disk is connected to another: forbidding the hexagon subgraph means that no  $\oplus$ -disk can be entirely surrounded by  $\oplus$ -disks. Property (2) guarantees that every component of  $G^+$  or  $G^-$  is connected in  $G_T$  to its exterior via a series gate edges. A simple induction on the depth of nesting of one component inside another leads to the connected conclusion. Start with a deepest component (e.g.,  $C_4^+$  in Figure 10). All of its disks are connected in



**Fig. 13.** A string connecting two disks adjacent via a directed arc of  $G_T$ : (a)  $(A, B) \subseteq \partial G$ . (b) The tangent does not touch any disk other than  $A$  and  $B$ . (c) Here  $(A, B)$  is a gate edge guaranteed by Property (2), and so edges of  $G_T$  connect to an adjacent disk  $D$ , as in (d).



**Fig. 14.** Property (2) fails:  $C^+$  contains a hole without a gate edge. Here only edges of  $G^+$  are shown. None of the diagonals connecting the inner to the outer hexagon can be edges of  $G_T$ , because they connect  $\oplus$ -disks not both on  $\partial G$ .

$G_T$  to the surrounding component ( $C_1^-$  in the figure). The series of gate edges that connect the holes to the exterior of the component guarantee a connection up the nesting hierarchy.

Now take a spanning tree  $T$  of  $G_T$ , following bidirectional edges (i.e., not employing the counterclockwise edges of  $G_T$  corresponding to  $\partial G$ ). Traversing  $T$ , and connecting through  $D$  to the cycle for  $\partial G$  (cf. Figure 13(c)), results in a proper wrapping of the arrangement. This establishes the “if”-direction of the lemma.

For the “only-if” direction, assume Properties (1) and (2) fail to hold for  $\mathcal{A}$ . Certainly if a component includes a hexagon subgraph (violating Property (1)), there can be no spanning cycle, because the central disk has no incident edges in  $G_T$ . So assume Property (2) fails to hold for some component  $C^+$ . This means that removal of all gate edges for its holes leaves a merged hole  $C^-$  that still cannot connect to the exterior of  $C^+$ . This can only happen if  $C^-$  is surrounded by a “padding” of  $\oplus$ -disks thick enough so that no gate edge connects to the exterior. The simplest example is illustrated in Figure 14. In this case it is easy to see that, although the  $\ominus$ -disks in  $C^-$  connect by edges of  $G_T$  to  $C^+$ , the padding isolates these connections from the exterior of  $C^+$ , disconnecting  $G_T$ . Thus, again  $G_T$  cannot contain a spanning cycle.  $\square$

Lemmas 4 and 5 together establish Theorem 3.

I believe the assumption in Theorem 3 that  $\partial G$  is composed entirely of  $\oplus$ -disks can be removed by surrounding a given arrangement with “virtual”  $\oplus$ -disks, and later removing them. There also seems no impediment permitting  $G$  to have holes, or to be disconnected. But these extensions perhaps best await a proof technique that encompasses them without effort.

## 5 Discussion

It is natural to hope that some analog of Theorem 3 holds for arbitrary arrangements. However, I have not found a formulation that avoids devolving to

a version of Lemma 4, relying on the existence of a cycle in a tangency graph analogous to  $G_T$ . A clean characterization remains open.

A second open question is to find a shortest wrapping when proper wrappings exist. For widely spaced disks, the analogy with Euclidean TSP suggests this may be NP-hard, but the situation is less clear for congested arrangements.

**Acknowledgments.** I thank Manuel Abellanas for graciously forgiving my oversight of his priority, and I thank Erik Demaine for the observation connecting to the ETSP. This paper benefited from insightful comments from referees and attendees at the *XIV Spanish Meeting on Computational Geometry*.

## References

1. Abellanas, M.: Conectando puntos: poligonizaciones y otros problemas relacionados. *Gaceta de la Real Sociedad Matemática Española* 11(3), 543–558 (2008)
2. Demaine, E.D., Demaine, M.L., Palop, B.: Conveyer-belt alphabet. In: Aardse, H., van Baalen, A. (eds.) *Findings in Elasticity*, pp. 86–89. Pars Foundation, Lars Müller Publishers (April 2010)

# The Chromatic Number of the Convex Segment Disjointness Graph<sup>\*</sup>

Ruy Fabila-Monroy<sup>1</sup> and David R. Wood<sup>2</sup>

<sup>1</sup> Departamento de Matemáticas

Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional  
México, D.F., México

[ruyfabila@math.cinvestav.edu.mx](mailto:ruyfabila@math.cinvestav.edu.mx)

<sup>2</sup> Department of Mathematics and Statistics

The University of Melbourne

Melbourne, Australia

[woodd@unimelb.edu.au](mailto:woodd@unimelb.edu.au)

**Abstract.** Let  $P$  be a set of  $n$  points in general and convex position in the plane. Let  $D_n$  be the graph whose vertex set is the set of all line segments with endpoints in  $P$ , where disjoint segments are adjacent. The chromatic number of this graph was first studied by Araujo et al. [CGTA, 2005]. The previous best bounds are  $\frac{3n}{4} \leq \chi(D_n) < n - \sqrt{\frac{n}{2}}$  (ignoring lower order terms). In this paper we improve the lower bound to  $\chi(D_n) \geq n - \sqrt{2n}$ , achieving near-tight bounds on  $\chi(D_n)$ .

## 1 Introduction

Throughout this paper,  $P$  is a set of  $n > 3$  points in general and convex position in the plane. The *convex segment disjointness graph*, denoted by  $D_n$ , is the graph whose vertex set is the set of all line segments with endpoints in  $P$ , where two vertices are adjacent if the corresponding segments are disjoint. Obviously  $D_n$  does not depend on the choice of  $P$ . This graph and other related graphs, were introduced by Araujo, Dumitrescu, Hurtado, Noy and Urrutia [1], who proved the following bounds on the chromatic number of  $D_n$ :

$$2 \lfloor \frac{1}{3}(n+1) \rfloor - 1 \leq \chi(D_n) < n - \frac{1}{2} \lfloor \log n \rfloor .$$

Both bounds were improved by Dujmović and Wood [5] to

$$\frac{3}{4}(n-2) \leq \chi(D_n) < n - \sqrt{\frac{1}{2}n} - \frac{1}{2}(\ln n) + 4 .$$

In this paper we improve the lower bound, achieving near-tight bounds on  $\chi(D_n)$ .

### Theorem 1.

$$n - \sqrt{2n + \frac{1}{4}} + \frac{1}{2} \leq \chi(D_n) < n - \sqrt{\frac{1}{2}n} - \frac{1}{2}(\ln n) + 4 .$$

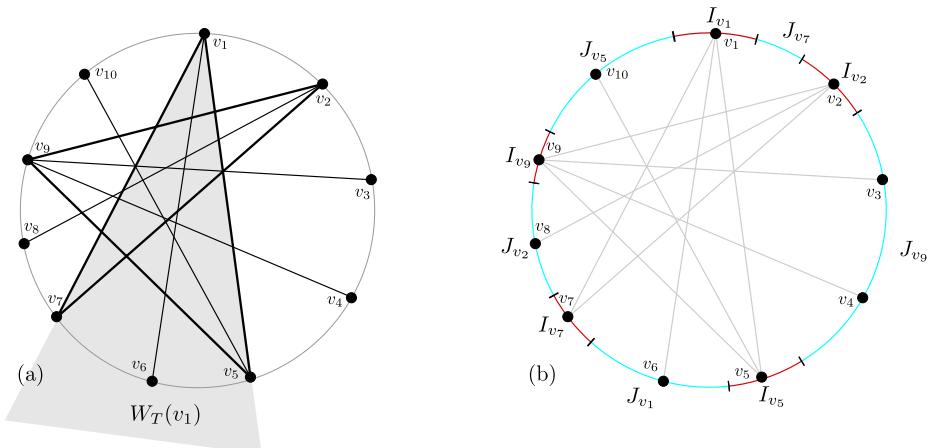
---

\* Dedicat al nostre amic i mestre Ferran Hurtado.

The proof of Theorem 1 is based on the observation that each colour class in a colouring of  $D_n$  is a convex thrackle. We then prove that two maximal convex thrackles must share an edge in common. From this we prove a tight upper bound on the number of edges in the union of  $k$  maximal convex thrackles. Theorem 1 quickly follows.

## 2 Convex Thrackles

A *convex thrackle* on  $P$  is a geometric graph with vertex set  $P$  such that every pair of edges intersect; that is, they have a common endpoint or they cross. Observe that a geometric graph  $H$  on  $P$  is a convex thrackle if and only if  $E(H)$  forms an independent set in  $D_n$ . A convex thrackle is *maximal* if it is edge-maximal. As illustrated in Figure 1(a), it is well known and easily proved that every maximal convex thrackle  $T$  consists of an odd cycle  $C(T)$  together with some degree 1 vertices adjacent to vertices of  $C(T)$ ; see [2–9]. In particular,  $T$  has  $n$  edges. For each vertex  $v$  in  $C(T)$ , let  $W_T(v)$  be the convex wedge with apex  $v$ , such that the boundary rays of  $W_T(v)$  contain the neighbours of  $v$  in  $C(T)$ . Every degree-1 vertex  $u$  of  $T$  lies in a unique wedge and the apex of this wedge is the only neighbour of  $u$  in  $T$ .



**Fig. 1.** (a) maximal convex thrackle with  $C(T) = \{v_1, v_7, v_2, v_9, v_5\}$ , (b) the intervals pairs  $(I_u, J_u)$

### 3 Convex Thrackles and Free $\mathbb{Z}_2$ -Actions of $S^1$

In this section, we prove the key lemma (Lemma 2) that a pair of convex maximal thrackles share an edge of a particular type. We offer two proofs, one topological, one combinatorial. The topological proof first needs this technical lemma.

A  $\mathbb{Z}_2$ -action on the unit circle  $S^1$  is a homeomorphism  $f : S^1 \rightarrow S^1$  such that  $f(f(x)) = x$  for all  $x \in S^1$ . We say that  $f$  is free if  $f(x) \neq x$  for all  $x \in S^1$ .

**Lemma 1.** *If  $f$  and  $g$  are free  $\mathbb{Z}_2$ -actions of  $S^1$ , then  $f(x) = g(x)$  for some point  $x \in S^1$ .*

*Proof.* For points  $x, y \in S^1$ , let  $\overrightarrow{xy}$  be the clockwise arc from  $x$  to  $y$  in  $S^1$ . Let  $x_0 \in S^1$ . If  $f(x_0) = g(x_0)$  then we are done. Now assume that  $f(x_0) \neq g(x_0)$ . Without loss of generality,  $x_0, g(x_0), f(x_0)$  appear in this clockwise order around  $S^1$ . Parameterise  $\overrightarrow{x_0g(x_0)}$  with a continuous injective function  $p : [0, 1] \rightarrow \overrightarrow{x_0g(x_0)}$ , such that  $p(0) = x_0$  and  $p(1) = g(x_0)$ . Assume that  $g(p(t)) \neq f(p(t))$  for all  $t \in [0, 1]$ , otherwise we are done. Since  $g$  is free,  $p(t) \neq g(p(t))$  for all  $t \in [0, 1]$ . Thus  $g(p([0, 1])) = g(p(0))g(p(1)) = \overrightarrow{g(x_0)x_0}$ . Also  $f(p([0, 1])) = \overrightarrow{f(x_0)f(p(1))}$ , as otherwise  $g(p(t)) = f(p(t))$  for some  $t \in [0, 1]$ . This implies that  $p(t), g(p(t)), f(p(t))$  appear in this clockwise order around  $S^1$ . In particular, with  $t = 1$ , we have  $f(p(1)) \in \overrightarrow{x_0g(x_0)}$ . Thus  $x_0 \in \overrightarrow{f(x_0)f(p(1))}$ . Hence  $x_0 = f(p(t))$  for some  $t \in [0, 1]$ . Since  $f$  is a  $\mathbb{Z}_2$ -action,  $f(x_0) = p(t)$ . This is a contradiction since  $p(t) \in \overrightarrow{x_0g(x_0)}$  but  $f(x_0) \notin \overrightarrow{x_0g(x_0)}$ .  $\square$

Assume that  $P$  lies on  $S^1$ . Let  $T$  be a maximal convex thrackle on  $P$ . As illustrated in Figure 1(b), for each vertex  $u$  in  $C(T)$ , let  $(I_u, J_u)$  be a pair of closed intervals of  $S^1$  defined as follows. Interval  $I_u$  contains  $u$  and bounded by the points of  $S^1$  that are  $1/3$  of the way towards the first points of  $P$  in the clockwise and anticlockwise direction from  $u$ . Let  $v$  and  $w$  be the neighbours of  $u$  in  $C(T)$ , so that  $v$  is before  $w$  in the clockwise direction from  $u$ . Let  $p$  be the endpoint of  $I_v$  in the clockwise direction from  $v$ . Let  $q$  be the endpoint of  $I_w$  in the anticlockwise direction from  $w$ . Then  $J_u$  is the interval bounded by  $p$  and  $q$  and not containing  $u$ . Define  $f_T : S^1 \rightarrow S^1$  as follows. For each  $v \in C(T)$ , map the anticlockwise endpoint of  $I_v$  to the anticlockwise endpoint of  $J_v$ , map the clockwise endpoint of  $I_v$  to the clockwise endpoint of  $J_v$ , and extend  $f_T$  linearly for the interior points of  $I_v$  and  $J_v$ , such that  $f_T(I_v) = J_v$  and  $f_T(J_v) = I_v$ . Since the intervals  $I_v$  and  $J_v$  are disjoint,  $f_T$  is a free  $\mathbb{Z}_2$ -action of  $S^1$ .

**Lemma 2.** *Let  $T_1$  and  $T_2$  be maximal convex thrackles on  $P$ , such that  $C(T_1) \cap C(T_2) = \emptyset$ . Then there is an edge in  $T_1 \cap T_2$ , with one endpoint in  $C(T_1)$  and one endpoint in  $C(T_2)$ .*

*Proof (Topological proof).* By Lemma 1, there exists  $x \in S^1$  such that  $f_{T_1}(x) = y = f_{T_2}(x)$ . Let  $u \in C(T_1)$  and  $v \in C(T_2)$  so that  $x \in I_u \cup J_u$  and  $x \in I_v \cup J_v$ , where  $(I_u, J_u)$  and  $(I_v, J_v)$  are defined with respect to  $T_1$  and  $T_2$  respectively. Since  $C(T_1) \cap C(T_2) = \emptyset$ , we have  $u \neq v$  and  $I_u \cap I_v = \emptyset$ . Thus  $x \notin I_u \cap I_v$ .

If  $x \in J_u \cap J_v$  then  $y \in I_u \cap I_v$ , implying  $u = v$ . Thus  $x \notin J_u \cap J_v$ . Hence  $x \in (I_u \cap J_v) \cup (J_u \cap I_v)$ . Without loss of generality,  $x \in I_u \cap J_v$ . Thus  $y \in J_u \cap I_v$ . If  $I_u \cap J_v = \{x\}$  then  $x$  is an endpoint of both  $I_u$  and  $J_v$ , implying  $u \in C(T_2)$ , which is a contradiction. Thus  $I_u \cap J_v$  contains points other than  $x$ . It follows that  $I_u \subset J_v$  and  $I_v \subset J_u$ . Therefore the edge  $uv$  is in both  $T_1$  and  $T_2$ . Moreover one endpoint of  $uv$  is in  $C(T_1)$  and one endpoint is in  $C(T_2)$ .  $\square$

*Proof (Combinatorial Proof).* Let  $H$  be the directed multigraph with vertex set  $C(T_1) \cup C(T_2)$ , where there is a *blue* arc  $uv$  in  $H$  if  $u$  is in  $W_{T_1}(v)$  and there is a *red* arc  $uv$  in  $H$  if  $u$  is in  $W_{T_2}(v)$ . Since  $C(T_1) \cap C(T_2) = \emptyset$ , every vertex of  $H$  has outdegree 1. Therefore  $|E(H)| = |V(H)|$  and there is a cycle  $\Gamma$  in the undirected multigraph underlying  $H$ . In fact, since every vertex has outdegree 1,  $\Gamma$  is a directed cycle. By construction, vertices in  $H$  are not incident to an incoming and an outgoing edge of the same color. Thus  $\Gamma$  alternates between blue and red arcs. The red edges of  $\Gamma$  form a matching as well as the blue edges, both of which are thrackles. However, there is only one matching thrackle on a set of points in convex position. Therefore  $\Gamma$  is a 2-cycle and the result follows.  $\square$

## 4 Main Results

**Theorem 2.** *For every set  $P$  of  $n$  points in convex and general position, the union of  $k$  maximal convex thrackles on  $P$  has at most  $kn - \binom{k}{2}$  edges.*

*Proof.* Let  $T_1, \dots, T_k$  be a set of  $k$  maximal convex thrackles on  $P$  with  $m$  edges in total. Recall that  $C(T_i)$  is the cycle in  $T_i$ .

First suppose that the cycles  $C(T_1), \dots, C(T_k)$  are pairwise disjoint. By Lemma 2, distinct  $T_i$  and  $T_j$  have an edge in common, with one endpoint in  $C(T_i)$  and one endpoint in  $C(T_j)$ . Hence distinct pairs of thrackles have distinct edges in common. Since every maximal convex thrackle has  $n$  edges and we overcount at least one edge for every pair,  $m \leq kn - \binom{k}{2}$ .

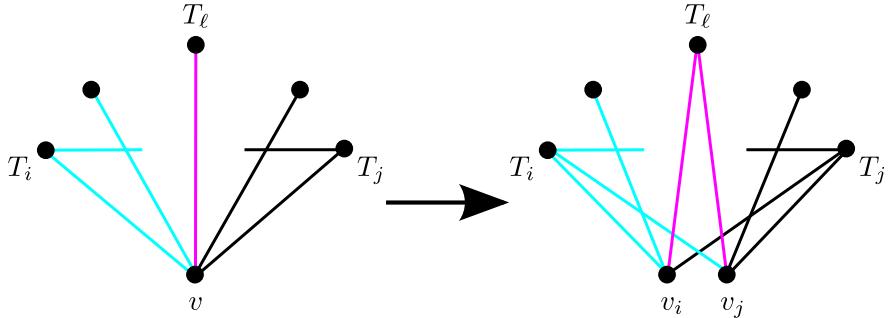
Now assume that some vertex is in distinct thrackle cycles. Say  $n_0$  vertices are in no thrackle cycle. Consider a vertex  $v$  in some thrackle cycle. Let  $I_v := \{i \in [1, k] : v \in C(T_i)\}$ . As illustrated in Figure 2, replace  $v$  by vertices  $\{v_i : i \in I_v\}$  that are placed consecutively around the circle in the spot previously occupied by  $v$ . For each  $i \in I_v$  if  $vx$  and  $vy$  are the two edges in  $C(T_i)$  incident to  $v$ , then replace  $vx$  and  $vy$  by  $v_ix$  and  $v_iy$ . This replacement of vertices and edges is done at all vertices simultaneously. Now, the thrackle cycles are pairwise disjoint. For each  $j \in [1, k] \setminus I_v$ , each vertex  $v_i$  is in the wedge of some vertex  $x$  in  $C(T_j)$ , implying the edge  $xv_i$  can be added to  $T_j$ . Hence  $k - 1$  edges incident to  $v_i$  are added in this way. Moreover, these are distinct edges because the thrackle cycles are pairwise disjoint. These modified thrackles now have  $n' = n_0 + \sum_v |I_v|$  vertices and  $m' \geq m + \sum_v (k - 1)|I_v|$  edges. By the argument in the first case,  $m' \leq kn' - \binom{k}{2}$ . That is,

$$m + \sum_v (k - 1)|I_v| \leq k(n_0 + \sum_v |I_v|) - \binom{k}{2}.$$

Since  $|I_v| \leq k$ ,

$$m \leq kn_0 + \sum_v |I_v| - \binom{k}{2} \leq kn - \binom{k}{2},$$

as desired.  $\square$



**Fig. 2.** Construction in the proof of Theorem 2

We now show that Theorem 2 is best possible for all  $n \geq 2k$ . Let  $S$  be a set of  $k$  vertices in  $P$  with no two consecutive vertices in  $S$ . If  $v \in S$  and  $x, v, y$  are consecutive in this order in  $P$ , then  $T_v := \{vw : w \in P \setminus \{v\}\} \cup \{xy\}$  is a maximal convex thrackle, and  $\{T_v : v \in S\}$  has exactly  $kn - \binom{k}{2}$  edges in total.

*Proof (of Theorem 1).* If  $\chi(D_n) = k$  then, there are  $k$  convex thrackles whose union is the complete geometric graph on  $P$ . Possibly add edges to obtain  $k$  maximal convex thrackles with  $\binom{n}{2}$  edges in total. By Theorem 2,  $\binom{n}{2} \leq kn - \binom{k}{2}$ . The quadratic formula implies the result.  $\square$

**Acknowledgements.** Thanks to Javier Cano, Alfredo García, Javier Tejel and Jorge Urrutia for pointing out an error in an earlier version of this paper. R.F.-M. was supported by an Endeavour Fellowship from the Department of Education, Employment and Workplace Relations of the Australian Government. D.W. is supported by a QEII Fellowship from the Australian Research Council.

## References

1. Araujo, G., Dumitrescu, A., Hurtado, F., Noy, M., Urrutia, J.: On the chromatic number of some geometric type Kneser graphs. *Comput. Geom. Theory Appl.* 32(1), 59–69 (2005)
2. Cairns, G., Nikolayevsky, Y.: Bounds for generalized thrackles. *Discrete Comput. Geom.* 23(2), 191–206 (2000)

3. Cairns, G., Nikolayevsky, Y.: Generalized thrackle drawings of non-bipartite graphs. *Discrete Comput. Geom.* 41(1), 119–134 (2009)
4. Cairns, G., Nikolayevsky, Y.: Outerplanar thrackles. *Graphs and Combinatorics* 28(1), 85–96 (2012)
5. Dujmović, V., Wood, D.R.: Thickness and antithickness (2010) (in preparation)
6. Fenchel, W., Sutherland, J.: Lösung der aufgabe 167. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 45, 33–35 (1935)
7. Hopf, H., Pammwitz, E.: Aufgabe no. 167. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 43 (1934)
8. Lovász, L., Pach, J., Szegedy, M.: On Conway’s thrackle conjecture. *Discrete Comput. Geom.* 18(4), 369–376 (1997)
9. Woodall, D.R.: Thrackles and deadlock. In: *Combinatorial Mathematics and its Applications* (Proc. Conf., Oxford, 1969), pp. 335–347. Academic Press, London (1971)

# Continuous Flattening of Convex Polyhedra

Jin-ichi Itoh<sup>1,\*</sup>, Chie Nara<sup>2,\*\*</sup>, and Costin Vilcu<sup>3,\*\*\*</sup>

<sup>1</sup> Faculty of Education, Kumamoto University,  
Kumamoto, 860-8555, Japan  
[j-ito@kumamoto-u.ac.jp](mailto:j-ito@kumamoto-u.ac.jp)

<sup>2</sup> Liberal Arts Education Center, Aso Campus, Tokai University,  
Aso, Kumamoto, 869-1404, Japan  
[cnara@ktmail.tokai-u.jp](mailto:cnara@ktmail.tokai-u.jp)

<sup>3</sup> Institute of Mathematics “Simion Stoilow” of the Romanian Academy,  
P.O. Box 1-764, 70700 Bucharest, Romania  
[Costin.Vilcu@imar.ro](mailto:Costin.Vilcu@imar.ro)

**Abstract.** A *flat folding* of a polyhedron is a folding by creases into a multilayered planar shape. It is an open problem of E. Demaine et al., that every flat folded state of a polyhedron can be reached by a continuous folding process. Here we prove that every convex polyhedron possesses infinitely many continuous flat folding processes. Moreover, we give a sufficient condition under which every flat folded state of a convex polyhedron can be reached by a continuous folding process.

## 1 Introduction

We use the terminology *polyhedron* for a closed polyhedral surface which is permitted to touch itself but not self-intersect (and so a doubly covered polygon is a polyhedron). A *flat folding* of a polyhedron is a folding by creases into a multilayered planar shape.

The results presented here are related to the following problem proposed by Erik Demaine et al. (see Open Problem 18.1 in [5]): *Can every flat folded state of a polyhedron be reached by a continuous folding process?* The existence of flat folded states for polyhedra homeomorphic to the 2-sphere was proved by the method of disk packing (see §18.3 in [5]), and for some special classes of convex polyhedra was also proved by the method of straight skeletons (see §18.4 in [5], and [4]). Notice that, if a polyhedron  $P$  is flattened by a continuous folding process (see Definition 1) with polyhedra  $\{P_t : 0 \leq t \leq 1\}$ , then the crease pattern in  $P$  for  $\{P_t : 0 \leq t \leq 1\}$  is an infinite set of line segments; this follows from Cauchy’s rigidity theorem and Sabitov’s result on the volume invariance under flexing [9]-[10].

Section 2 of this work is devoted to preliminaries. We also briefly present there (Theorem 1) the method to continuously flatten the Platonic polyhedra onto their original faces, proposed in [6].

---

\* Supported by Grand-in-Aid for Scientific Research (No.23540098), JSPS.

\*\* Supported by Grand-in-Aid for Scientific Research (No.23540160), JSPS.

\*\*\* Partially supported by the grant PN-II-ID-PCE-2012-4-0378 of CNCS-UEFISCDI.

In Sect. 3 we propose a method to flatten general convex polyhedra by continuous folding processes (Theorem 2). For the proofs, we employ Alexandrov's gluing theorem and the structure of cut loci (see Definition 2).

In Sect. 4 we give a sufficient condition, under which every flat folded state of a convex polyhedron can be reached by a continuous folding process (Theorem 3).

We end the paper with a few remarks and open questions (Sect. 5).

## 2 Preliminaries

We start with the definition of a continuous folding process for a polyhedron.

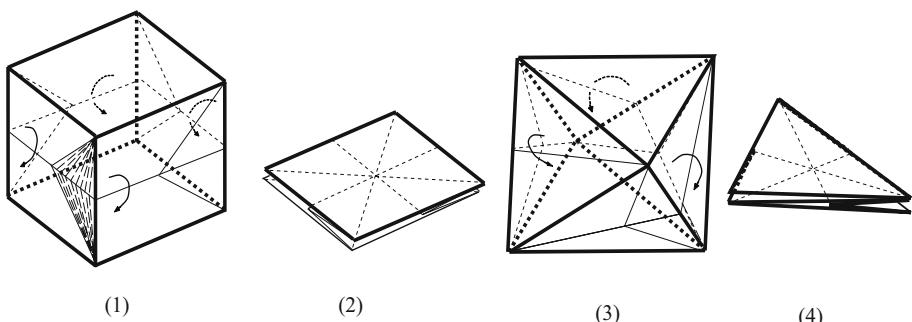
**Definition 1.** Let  $P$  be a polyhedron in the Euclidean space  $\mathbb{R}^3$ . We say that a family of polyhedra  $\{P_t : 0 \leq t \leq 1\}$  is a continuous folding process from  $P = P_0$  to  $P_1$  if it satisfies the following conditions:

- (1) for each  $0 \leq t \leq 1$ , there exists a polyhedron  $P'_t$  obtained from  $P$  by subdividing some faces of  $P$  (i.e., some faces of  $P'_t$  may be included in the same face of  $P$ , but  $P'_t$  is congruent to  $P$ ) such that  $P_t$  is combinatorially equivalent to  $P'_t$  and the corresponding faces of  $P'_t$  and  $P_t$  are congruent,
- (2) the mapping  $[0, 1] \ni \tau \mapsto P_\tau \in \{P_t : 0 \leq t \leq 1\}$  is continuous.

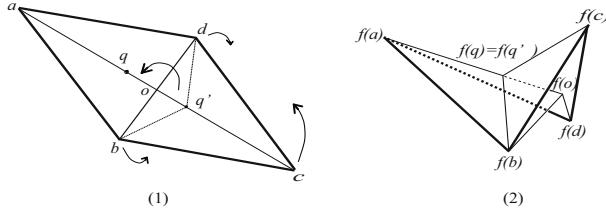
Moreover, if  $P_1$  is a flat folded polyhedron, we say that  $P$  is flattened by a continuous folding process and we call  $P_1$  a flat folded polyhedron (or state) of  $P$ .

In the case of Platonic polyhedra, two of us proved the next result [6], which will serve as a contrast in Section 4.

**Theorem 1.** For the five Platonic polyhedra there are continuous flat folding processes onto their original faces.



**Fig. 1.** (1) The cube; (2) the flattened cube on its face; (3) the regular octahedron; (4) the flattened octahedron on its face.



**Fig. 2.** An example of a folded rhombus

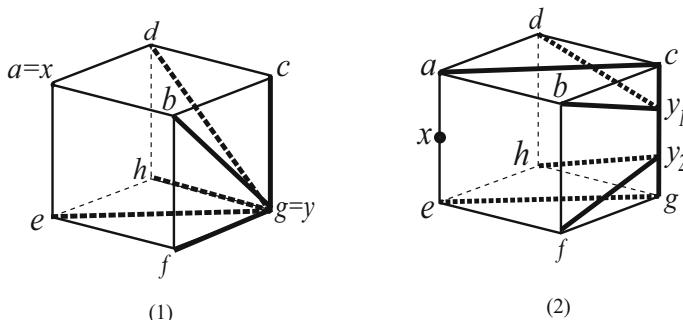
Figure 1 shows how to continuously flatten the cube and the regular octahedron on their faces (see [6] for details or for the other Platonic polyhedra). Theorem 1 was proved by using a key lemma: any rhombus can be folded into a shape as showed in Fig.2 (2), with distances  $|f(b)f(d)| = l$  and  $|f(a)f(c)| = m$  for any given  $0 \leq l \leq |bd|$  and  $0 \leq m \leq |ac|$ , where we denote by  $|xy|$  the Euclidean distance between  $x, y \in \mathbb{R}^3$ .

Our main tools here are the Alexandrov's gluing theorem (stated below) and the cut loci, to which the remaining of this section is devoted.

**Alexandrov's Gluing Theorem.** Consider a topological sphere  $S$  obtained by gluing planar polygons (i.e., naturally identifying pairs of sides of the same length) such that at most  $2\pi$  angle is glued at each point. Then  $S$ , endowed with the intrinsic metric induced by the distance in  $\mathbb{R}^2$ , is isometric to a convex polyhedron  $P \subset \mathbb{R}^3$ , possibly degenerated. Moreover,  $P$  is unique up to rigid motion and reflection in  $\mathbb{R}^3$ . (See [2], p.100.)

**Definition 2.** The cut locus  $C(x) = C(x, P)$  of the point  $x$  on the convex polyhedron  $P$  is the set of endpoints (different from  $x$ ) of all nonextendable segments (i.e., shortest paths on  $P$ ) starting at  $x$ .

Figure 3 provides two examples of cut loci for points on the cube.



**Fig. 3.** (1) The cut locus of the vertex  $x = a$  on a cube; (2) the cut locus on a cube with respect to the midpoint  $x$  of the edge  $ae$ .

We summarize in our first lemma various known properties of cut loci.

**Lemma 1.** (i)  $C(x)$  is a tree whose leaves (endpoints) are vertices of  $P$ , and all vertices of  $P$ , excepting possibly  $x$ , are included in  $C(x)$ . Notice that we allow vertices of degree two in  $C(x)$ .

(ii) The junction points in  $C(x)$  are joined to  $x$  by as many segments as their degree in the tree.

(iii) The edges of  $C(x)$  are segments on  $P$ .

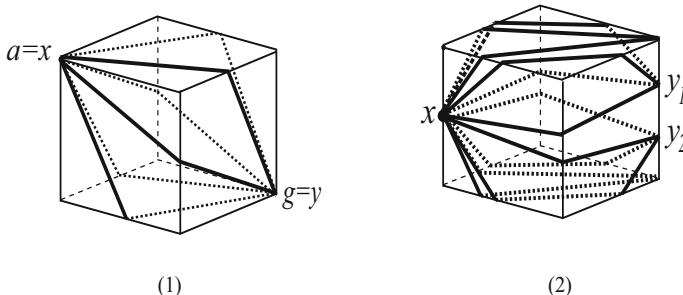
(iv) Assume the segments  $\gamma$  and  $\gamma'$  from  $x$  to  $y \in C(x)$  are bounding a domain  $D$  of  $P$ , which intersects no other segment from  $x$  to  $y$ . Then the arc of  $C(x)$  at  $y$  towards  $D$  bisects the angle of  $D$  at  $y$ .

(v) If  $P$  has  $n$  vertices then  $C(x)$  is a tree with  $O(n)$  vertices, and it can be constructed in time  $O(n^2)$ .

*Proof.* The properties (i)-(ii) and (iv) are well known, while (iii) is Lemma 2.4 in [1].

The first part of (v) is clear; for the second part, we employ the algorithm of J. Chen and Y. Han [3] (see [7] for a public implementation).  $\square$

We will use the cube to illustrate our method. This method depends upon segments from a point  $x$  on the cube to particular points on  $C(x)$ ,  $y$  in Fig. 4(1), and  $y_1$  and  $y_2$  in Fig. 4(2).



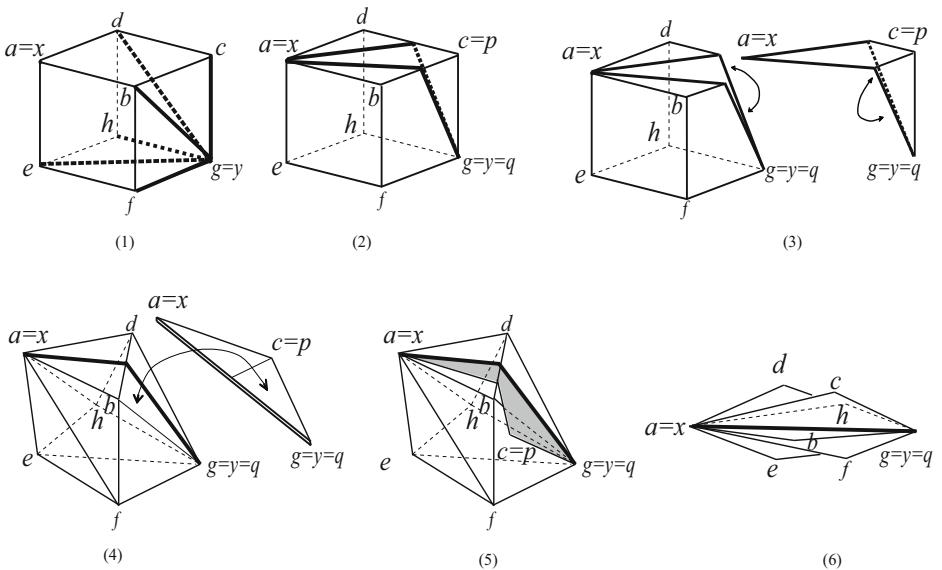
**Fig. 4.** (1) Segments joining  $x = a$  to  $y = g$ ; (2) segments joining  $x$  to  $y_1$ ,  $y_2$ , and to the cube vertices that are interior points of  $C(x)$

We call an edge of  $C(x)$  a *leaf edge* if it is incident to a leaf of  $C(x)$ .

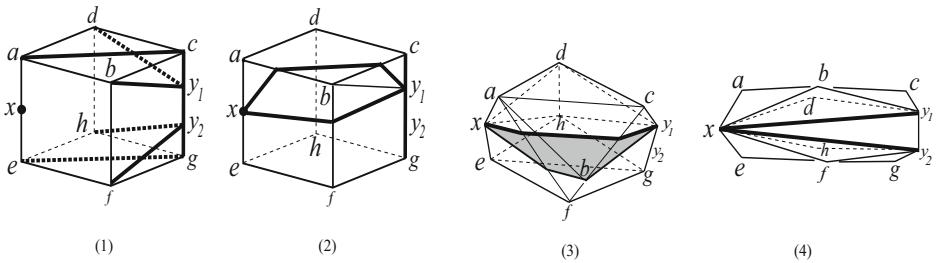
### 3 Continuous Flattening Processes for Convex Polyhedra

In this section we provide a method to continuously flatten any convex polyhedron  $P$ , based on cut loci and Alexandrov's gluing theorem. Toward this goal, we further describe the structure of cut loci.

We first give a high-level view of the method, presenting two different ways to flatten the cube, illustrated in Figs. 5 and 6. We start with an arbitrary point  $x$



**Fig. 5.** (1) The cut locus of  $x = a$ ; (2) two segments  $\gamma_1$  and  $\gamma_2$  joining  $x$  to  $g$ , enclosing precisely one leaf edge  $E = uv = gc$ ; (3) two parts of the cube, separated by  $\gamma_1 \cup \gamma_2$ ; (4) two resulting surfaces obtained by gluing (the images of)  $\gamma_1$  and  $\gamma_2$ ; (5) the resulting polyhedron from the cube after gluing  $\gamma_1$  to  $\gamma_2$ ; (6) the flat folded state of the cube finally obtained.



**Fig. 6.** (1) The cut locus of the midpoint  $x$  of  $ae$ ; (2) two segments  $\gamma_1$  and  $\gamma_2$  joining  $x$  to  $y_1$ , and enclosing precisely one leaf edge  $E = uv = y_1b$ ; (3) the resulting polyhedron after gluing  $\gamma_1$  to  $\gamma_2$ ; (4) the flat folded state of the cube finally obtained.

on  $P$ , and determine its cut locus and all segments from  $x$  to the junction points of  $C(x)$ , see Figs. 5 (1) and 6 (1). Every leaf edge  $E$  of  $C(x)$  is included in some region  $T$  of  $P$  bounded by two consecutive segments from  $x$  to a junction point of  $C(x)$  ( $E = cg$  in Fig. 5 (2/3) and  $E = by_1$  in Fig. 6 (2)).  $T$  can be flattened to a doubly covered triangle  $T_E$ , and  $P \setminus T$  can be “zipped” to some convex polyhedron  $Q_E$  (by Alexandrov’s gluing theorem). Therefore,  $P$  is isometric to  $P_E = T_E \cup Q_E$ , consisting of a bent doubly covered triangle  $T_E$  (shaded in the figures) attached to some convex polyhedron  $Q_E$  (Lemma 2), see Figs. 5 (4/5) and 6 (3). The cut locus  $C(x, Q_E)$  of  $x$  on  $Q_E$  is precisely the truncation of the cut locus  $C(x, P)$  (Lemma 3), hence we can iterate the process until  $C(x)$  is merely a path, in which case the resulting polyhedron is already flattened (Lemma 4), see Figs. 5 (6) and 6 (4). Lemma 5 shows the continuity of this procedure.

Let  $\gamma_1$  and  $\gamma_2$  be segments on  $P$  from  $x \in P$  to  $y \in C(x)$ ; cut along  $\gamma_1 \cup \gamma_2$  and keep one half-surface  $P'$ . By *gluing*  $\gamma_1$  to  $\gamma_2$  we mean to identify the points on  $\gamma_1$  and respectively  $\gamma_2$  at equal distance to  $x$ .

**Lemma 2.** *Let  $x$  be a point on a convex polyhedron  $P$ . Each leaf edge  $E = uv$  of the cut locus  $C(x)$ , starting at the leaf  $v$  of  $C(x)$ , is bounded by two segments  $\gamma_1$  and  $\gamma_2$  from  $x$  to  $u$ , whose union encloses precisely one leaf,  $v$ , of  $C(x)$ . The region  $T$  of  $P$ , enclosed by  $\gamma_1 \cup \gamma_2$  and containing  $v$ , can be flattened to a doubly covered triangle  $T_E$ , and the remaining part of  $P$  corresponds to a convex polyhedron  $Q_E$  by gluing  $\gamma_1$  to  $\gamma_2$ . The original polyhedron  $P$  is isometric to the polyhedron  $P_E = Q_E \cup T_E$ , where we attach  $T_E$  to  $Q_E$  such that  $\gamma_1$  and  $\gamma_2$  are touching each other but are included in distinct layers.*

*Proof.* Let  $E = uv$  be any leaf edge with a leaf  $v$  of the cut locus  $C(x)$ . Then there are  $d$  segments joining  $x$  to  $u$  on  $P$ , where  $d$  is the degree of the point  $u$  in the tree  $C(x)$ , and exactly two of them, say  $\gamma_1$  and  $\gamma_2$ , enclose precisely one leaf  $v$ . Figs. 5(2) and 6(2) show the regions of the cube corresponding to the leaf edges  $c g$  and  $b y_1$  in the respective cut loci. The region  $T$  of  $P$ , bounded by  $\gamma_1 \cup \gamma_2$  and containing  $v$ , has no other vertex of  $P$  inside, hence it consists of two flat congruent triangles with sides  $xv$ ,  $uv$  and  $\gamma_i$  ( $i = 1, 2$ ). Therefore,  $T$  can be flattened to some doubly covered triangle  $T_E$ , by gluing  $\gamma_1$  to  $\gamma_2$ .

The remaining part of  $P$  corresponds to a convex polyhedron  $Q_E$  by Alexandrov’s gluing theorem. Hence  $P$  is isometric to  $P_E = Q_E \cup T_E$ , where we attach  $T_E$  to  $Q_E$  such that  $\gamma_1$  and  $\gamma_2$  are touching each other but are included in distinct layers in  $P_E$ . Notice that, although flattened,  $T_E$  may not lie in a plane.  $\square$

For a leaf edge  $E$  of  $C(x, P)$ , we will use the notation  $Q_E$  and  $P_E$  for the polyhedra introduced in Lemma 2.

**Lemma 3.** *Let  $x$  be a point on a convex polyhedron  $P$ , and let  $E = uv$  be a leaf edge of the cut locus  $C(x)$ , incident to the the leaf  $v$  of  $C(x)$ . The cut locus  $C(x, Q_E)$  of  $x$  on  $Q_E$  is (isometric to) the truncation of the cut locus  $C(x, P)$  with respect to the cuts along  $\gamma_1$  and  $\gamma_2$  defined in Lemma 2, and the gluing along them.*

*Proof.* This follows from the definition of cut locus and (iv) in Lemma 1.  $\square$

**Lemma 4.** *If the cut locus  $C(x)$  of a point  $x$  on a convex polyhedron  $P$  is a path, then  $P$  is a doubly covered polygon.*

*Proof.* Since all vertices of  $P$  except possibly  $x$  are included in  $C(x)$ ,  $P$  is a doubly covered polygon, by Lemmas 2 and 3.  $\square$

**Lemma 5.** *Let  $x$  be a point in a convex polyhedron  $P$ , whose cut locus  $C(x)$  is not a path, and let  $E = uv$  be a leaf edge of  $C(x)$ , incident to the leaf  $v$  of  $C(x)$ . There is a continuous folding process from  $P$  to  $P_E$ .*

*Proof.* Let  $p_t$  be a point moving continuously from  $v$  to  $u$  along the edge  $E = uv$ , as  $t$  increases from 0 to 1, and denote  $E_t = p_tv$ . There are two segments  $\gamma_{t,1}$  and  $\gamma_{t,2}$  joining  $x$  to  $p_t$ , enclosing precisely one leaf  $v$  of  $C(x)$ . By cutting along  $\gamma_{t,1} \cup \gamma_{t,2}$  and gluing  $\gamma_{t,1}$  to  $\gamma_{t,2}$ , we obtain a doubly covered triangle  $T_t = T_{E_t}$  and a convex polyhedron  $Q_t = Q_{E_t}$ , by Alexandrov's gluing theorem. Let  $P_t = P_{E_t} = Q_t \cup T_t$  be defined similarly to  $P_E$  in Lemma 2. Here  $T_t$  is flipped clockwise about the point  $x$  and it touches  $Q_t$ , in order to avoid any conflict later.

We establish below a property of  $Q_t$ .

The structure of  $Q_t$  is given to us via Alexandrov's gluing theorem. We know the vertices of  $Q_t$  (they are  $x$ ,  $p_t$ , and vertices of  $P$ ), but do not know its edges. Nevertheless,

- (i) the edges of  $Q_t$  between vertices (corresponding to vertices) of  $P$  are a subset of the collection of all segments on  $P$  between pairs of such vertices;
- (ii) the edges of  $Q_t$  from its vertex  $x$  to vertices (corresponding to vertices) of  $P$  are a subset of the collection of all segments on  $P$  from  $x$  to such vertices;
- (iii) the edges of  $Q_t$  from its vertex  $p_t$  to vertices (corresponding to vertices) of  $P$  are a subset of the collection of all segments on  $P$  from  $p_t$  to such vertices; moreover, if  $p_t$  is joined to such a vertex by two segments then none of them corresponds to an edge of  $Q_t$ , because two vertices on the same edge of a convex polyhedron are joined by precisely one segment, the edge between them;
- (iv) the edge of  $Q_t$  between  $x$  and  $p_t$  corresponds to  $\gamma_{t,1}$  and  $\gamma_{t,2}$ .

Denote by  $G$  the subset of  $P$  consisting of all segments joining pairs of vertices of  $P$ , or joining  $x$  to vertices of  $P$ .

Consider now a neighborhood  $N$  of  $u$ , and points  $y_0, z_0 \in N$  such that the triangle  $\Delta_0 = \triangle u y_0 z_0$  intersects  $G \cup C(x)$  only at  $u$ . This is possible, because both  $G$  and  $C(x)$  are composed of finitely many segments on  $P$ , hence  $N \cap (G \cup C(x))$  consists of finitely many segments on  $N$ .

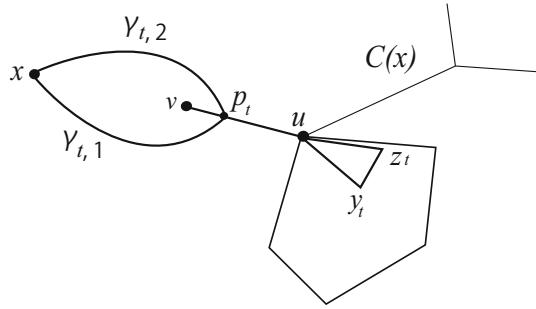
For a vertex  $w$  of  $P$  and a real number  $0 \leq t \leq 1$ , let  $G'_t(w)$  be the subset of  $P$  consisting of all segments from  $w$  to  $p_t$ , and  $G_t(w)$  be the union of  $G'_s(w)$  for all  $0 \leq s \leq t$ ,

$$G_t(w) = \bigcup_{0 \leq s \leq t} G'_s(w).$$

Notice that  $G_t(w)$  is a closed set, by the upper semi-continuity of segments<sup>1</sup>.

---

<sup>1</sup> Assume that  $\gamma_n$  is a segment joining the points  $x_n, y_n$  on  $P$ , for each  $n \in \mathbb{N}$ ; if  $x_n \rightarrow x$ ,  $y_n \rightarrow y$ , and  $\gamma_n \rightarrow \gamma$ , then  $\gamma$  is a segment between  $x$  and  $y$ .



**Fig. 7.** A triangle  $\triangle u y_t z_t$  such that  $\text{int}(\triangle u y_t z_t) \cap G_t = \emptyset$

Let  $G_t$  be the union of  $G_t(w)$  for all vertices  $w$  of  $P$ ,

$$G_t = \bigcup_{w \text{ a vertex of } P} G_t(w),$$

hence

$$G_s \subset G_t \quad \text{for } 0 \leq s < t < 1.$$

We show the following:

*Property (\*):* for  $0 \leq t < 1$  there exist points  $y_t, z_t \in N \setminus G_t$  such that the triangle  $\Delta_t = \triangle u y_t z_t$  (considered “full”, i.e., with interior points in  $P$ ) intersects  $G_t \cup C(x)$  only at  $\{u\}$  (see Fig. 7), and  $\Delta_t \subset \Delta_s$  for  $0 \leq s \leq t < 1$ .

The existence of  $\Delta_t$  as in the property (\*) guarantees we can choose a “distinguished” face  $F_t$  of  $Q_t$  for all  $t$ , such that  $F_t \subset F_s$  for  $0 \leq s \leq t < 1$ .

Case 1. Suppose that  $u$  is a vertex of  $P$ . Let  $0 \leq t < 1$ . Then, for any vertex  $w \neq u$  of  $P$  and any real number  $0 \leq s \leq t$ , no segment from  $w$  to  $p_s$  passes through  $u$ , and hence  $G_t(w)$  does not include  $u$ . Since  $G_t(w)$  is a closed set, there exists a neighborhood  $N_{t,w}$  of  $u$  which is disjoint from  $G_t(w)$ ,

$$N_{t,w} \cap G_t(w) = \emptyset.$$

Since the number of vertices of  $P$  is finite,  $N_t = \bigcap_w N_{t,w}$  is still a neighborhood of  $u$ . Now,  $G_s \subset G_t$  for  $0 \leq s \leq t < 1$  implies  $N_t \subset N_s$ , hence

$$N_t \cap (G_t \setminus G_t(u)) = \emptyset,$$

because  $G_t(u) \subset uv$ . Therefore, there exist points  $y_t, z_t \in N_t \cap \Delta_0$  such that the triangle  $\Delta_t = \triangle u y_t z_t$  intersects  $G_t \cup C(x)$  only at  $u$ . Moreover, because  $G_s \subset G_t$ , we can take  $y_t, z_t \in \bigcap_{t' < t} \Delta_{t'}$ , hence  $\Delta_t \subset \Delta_s$  for  $0 \leq s \leq t < 1$ . Since  $\Delta_0 \cap (G \cup C(x)) = \{u\}$ ,

$$\Delta_t \cap (G_t \cup G \cup C(x)) = \{u\}.$$

Case 2. Suppose that  $u$  is not a vertex of  $P$ . If  $u \notin G_t$  for all  $0 \leq t < 1$  then, just as above,  $G_t \cap (G \cup C(x)) = \{u\}$  for all  $0 \leq t < 1$ , and the property (\*) holds.

We assume in the following that  $u \in G_\tau$  for some  $0 \leq \tau < 1$ , hence there exists a vertex  $w$  of  $P$  joined to  $p_\tau$  by a segment  $\delta_\tau$  through  $u$ . Notice that such a vertex  $w$  is unique, because segments do not branch and do not pass beyond vertices. Moreover, all segments from  $w$  to  $p_s$  are subsegments of  $\delta_\tau$ , for  $\tau \leq s < 1$ , because segments do not branch.

Let  $\tau_0$  be the minimum  $\tau$  such that a segment from  $w$  to  $p_\tau$  contains  $u$ , and let  $\delta_0$  denote that segment. Since  $\delta_0$  is no longer a shortest path beyond  $p_{\tau_0}$ ,  $p_{\tau_0} \in C(w)$  and we have two subcases to consider.

(i) The point  $p_{\tau_0}$  is not a leaf of  $C(w)$ , hence it is joined to  $w$  by at least two segments. Of course,  $u \notin G_t(w)$  for  $0 \leq t < \tau_0$ . Therefore, since  $G_{\tau_0}(w)$  is closed,  $u$  is in the boundary of  $G_{\tau_0}(w)$ , because  $G_s \subset G_t$  for  $0 \leq s < t < 1$ . It follows that there exists a small half-disk  $D_{\tau_0,w}$  centered at  $u$  such that  $D_{\tau_0,w} \cap G_{\tau_0}(w) = \delta_0$ . By a similar argument to the one in Case 1, the property (\*) holds.

(ii) Assume now that  $p_{\tau_0}$  is a leaf of  $C(w)$ . Then  $p_{\tau_0}$  is a vertex of  $P$  by Lemma 1 (i), and therefore  $p_{\tau_0} = v$ . Then  $G_t(w) \subset wv$  for all  $0 \leq t < 1$ , and the property (\*) follows as above.

We show next how to realize  $P_t$  in  $\mathbb{R}^3$ , for  $0 \leq t \leq 1$ . Of course, it suffices to show how to realize  $Q_t$ .

For  $t = 0$ , we realize  $Q_0$  (and hence  $P_0 = P$ ) in  $\mathbb{R}^3$ , satisfying the conditions: (i)  $u = (0, 0, 0)$ ,  $y_0 = (y_{0,1}, 0, 0)$ , and  $z_0 = (z_{0,1}, z_{0,2}, 0)$ , for some real numbers  $y_{0,1} \geq 0$ ,  $z_{0,1}$ ,  $z_{0,2}$ ; and (ii)  $Q_0$  is included in the half-space  $z \geq 0$ .

For  $0 < t < 1$ , we realize  $Q_t$  in the half-space  $z \geq 0$  such that  $\Delta_t$  is realized as the corresponding subset of  $N$ .

Assume now that  $t = 1$ . Let  $\{t_n\}_{n \geq 1}$  be a sequence converging to 1, with  $0 < t_n < 1$ . Since the family  $\{Q_{t_n}\}_{n \geq 1}$  is bounded with respect to the Hausdorff metric on the space of all compact sets in  $\mathbb{R}^3$ , there exists a subsequence which converges to a compact set  $R_1$ , by Blashke's convergence theorem. The unicity in Alexandrov's gluing theorem shows now that  $R_1$  does not depend on the choice of the converging sequence  $t_n \rightarrow 1$ , hence we may realize  $Q_1$  by  $R_1$ .

Concluding, if  $Q_t$  is close to  $Q_s$  then their corresponding faces are close to each other, in particular those including  $\Delta_t$  and  $\Delta_s$ , and thus their realizations in  $\mathbb{R}^3$  are close to each other.

Finally, we notice that the mapping from  $0 \leq t \leq 1$  to the 1-parameter family of compact sets  $\{Q_t : 0 \leq t \leq 1\}$  is continuous with respect to the Hausdorff metric. Let  $s$  be a real number with  $0 \leq s \leq 1$ , and  $\{s_n\}_{n \geq 1}$  a sequence converging to  $s$ , with  $0 < s_n < 1$ . The family  $\{Q_{s_n}\}_{n \geq 1}$  is bounded with respect to the Hausdorff metric on the space of all compact sets in  $\mathbb{R}^3$ , hence there exists a subsequence which converges to a compact set  $R_s$ , by Blashke's convergence theorem, and the unicity in Alexandrov's gluing theorem, which completes the proof.  $\square$

**Theorem 2.** *For every convex polyhedron there exist infinitely many continuous flat folding processes.*

*Proof.* Let  $P$  be a convex polyhedron and let  $x$  be a point in  $P$ .

Step 1. Determine the cut locus  $C(x)$ , which is a tree (see Lemma 1).

Step 2. Flatten the region  $T$  of  $P$  corresponding to a leaf edge  $E$  of  $C(x)$  (see Lemma 2). The remaining part of  $P$ , after flattening  $T$  as above, is realized as a convex polyhedron  $Q_E$ , by Alexandrov's gluing theorem. Therefore, the result  $P_E$  after this flattening is isometric to  $P$ , and consists of the polyhedron  $Q_E$  attached to the doubly covered triangle  $T_E$ .  $T_E$  should be laid clockwise about the point  $x$  in order to avoid conflict.

Step 3. Iterate Step 2 for  $Q_E$  instead of  $P$ , until  $C(x, Q_E)$  is reduced to a path; i.e., until  $Q_E$  is a doubly covered polygon. Lemma 3 guarantees the iterations are possible, while Lemma 4 establishes the final form of  $Q$ .

Figs. 5(6) and 6(4) show the flat folded states of the cube after flattening all such regions corresponding to leaf edges of  $C(x, P)$ .

Since there are  $O(n)$  vertices in  $C(x)$ , where  $n$  is the number of vertices of  $P$  (see Lemma 1), we have to flatten  $O(n)$  regions of  $P$  corresponding to leaf edges of  $C(x, Q)$  one by one, and therefore the flattening process ends after  $O(n)$  iterations.

All folding processes corresponding to leaf edges are continuous by Lemma 5, so  $P$  is continuously folded to a flat folded state.  $\square$

Igor Pak [8] proved the existence of volume-increasing isometric deformations for convex polyhedra  $P$  in  $\mathbb{R}^3$ . With the above procedure, the volume of  $P$  continuously decreases to zero through isometric deformations.

## 4 Continuous Flattening Processes for Simple Flat Folded States of Convex Polyhedra

In this section we give a sufficient condition for a flat folded state of a convex polyhedron, to be reached by a continuous folding process.

**Definition 3.** A 2-covered convex polygon consists of two copies of a convex polygon glued along some of their corresponding edges (the other edges are “cut”).

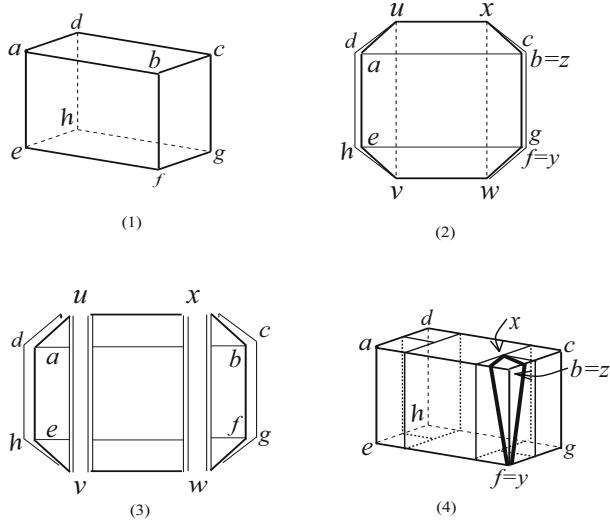
Figure 8(3) provides examples of 2-covered convex polygons. We will always regard such surfaces with boundary as having two congruent layers touching at their corresponding points, but glued along only some edges.

**Definition 4.** A flat folded state  $P_f$  of a convex polyhedron is called simple if it enjoys the following properties:

- (i)  $P_f$  can be decomposed into a finite number of 2-covered convex polygons  $\{R_i : 1 \leq i \leq k\}$ , and
- (ii) for any  $1 \leq i \leq k$ , if  $R_1, R_2, \dots, R_{i-1}$  were cut off from  $P_f$ , then  $R_i$  can also be cut off from the remaining part of  $P_f$  by precisely one cut along an edge of  $R_i$ .

Notice that all flat folded states obtained in Theorem 2 are simple.

On the other hand, Theorem 1 provides examples of non-simple flat folded states. For example, consider the flat folded state  $P_f$  of the cube  $C$  in Fig. 1(2).



**Fig. 8.** (1) A rectangular box  $B$ ; (2) a simple flat folded state of  $B$  obtained by pushing in two side faces of  $B$ ; (3) a decomposition of the flat folded state into four congruent 2-covered trapezoids and one 2-covered rectangle; (4) the region enclosed by two segments from  $x$  to  $y$ , corresponding to the doubly covered triangle  $xfb$ .

Divide each face of  $C$  into four congruent triangles by the diagonals, hence each edge of  $C$  is a common edge of two such triangles. Notice that  $P_f$  consists of twelve 2-covered such triangles  $T_i$ , each of which includes precisely one edge of the cube. Since all centers of the six faces of  $C$  are folded to the same point  $p$  in  $P_f$  (the center of an original face), all twelve 2-covered triangles meet at  $p$  in  $P_f$ . Hence neither  $T_i$ , nor any smaller 2-covered polygon  $D_i \subset T_i$ , can be cut off from  $P_f$  by precisely one straight cut ( $i = 1, \dots, 12$ ). Similarly, the flat folded state in Fig. 1(4) is not simple.

Figure 8(2) shows a flat folded state of a rectangular box  $B$ , where two side faces of  $B$  are pushed in. It is simple, because it can be decomposed into five 2-covered convex polygons as shown in Fig. 8(3), namely four congruent 2-covered trapezoids and one 2-covered rectangle. However, it cannot be obtained by the cut locus method described in the previous section.

**Theorem 3.** *Every simple flat folded state of a convex polyhedron can be reached by a continuous folding process.*

*Proof.* Let  $P_f$  be a simple flat folded state of a convex polyhedron  $P$ , decomposed into a finite collection of 2-covered convex polygons  $\{R_i : 1 \leq i \leq k\}$  by cutting  $P_f$  along some edge of  $R_i$ , one by one. By subdividing  $R_i$  ( $1 \leq i \leq k$ ) if necessary, we can assume, without loss of generality, that all  $R_i$  are 2-covered triangles.

We prove the result by induction over  $k$ .

If  $k = 1$  then  $P = P_f$  and the conclusion holds.

Suppose now that the statement is true for  $n$ , and assume  $k = n + 1$  for  $P_f$ . Let  $R_1$  be a 2-covered triangle  $\Delta xyz$  with the edge  $xy$  cut and edges  $xz$  and  $yz$  glued. Consider the cut locus  $C(x) = C(x, P)$ . Then  $E = yz$  is a leaf edge of  $C(x)$ . By Lemma 5, there exists a continuous folding process from  $P$  to the polyhedron  $P_E = Q_E \cup R_1$ . Notice that  $Q_E$  has the flat folded state  $Q_{E,f}$  consisting of 2-covered convex triangles  $\{R_i : 2 \leq i \leq n + 1\}$ , and  $Q_{E,f}$  satisfies the condition (ii) in Definition 4, with the edge  $xy$  of  $R_2$  glued. By the induction's assumption,  $Q_{E,f}$  can be reached by a continuous folding process from  $Q_E$ . Therefore,  $P_f$  can be reached by a continuous folding process from  $P$ .  $\square$

## 5 Remarks and Open Questions

Our first result in this paper proposes an algorithmic method to continuously flatten convex polyhedra. There, the assumption of convexity is essential at two points. First, for the existence of the flat folded state we used the fact that each edge of the cut locus is a segment (Lemma 1 (iii)), property which is not true on non-convex surfaces. Second, for the continuity of the process, we used the uniqueness in Alexandrov's gluing theorem, which fails for non-convex surfaces. Since essentially the same argument is employed to prove our second result, a sufficient condition for the existence of continuous flattening processes, our proof there also fails for non-convex surfaces.

Our approach raises several questions concerning the structure of cut loci.

Our flattening procedure starts with the cut locus  $C(x)$  of the point  $x$  in  $P$ , and at each step it treats (more precisely, it eliminates) all leaf edges of  $C(x)$ .

*Question 1.* For which polyhedra can one find points  $x$  whose cut locus has precisely one ramification point (i.e.,  $C(x)$  is homeomorphic to a “star graph”)? Generally, what is the minimal number of steps (i.e., of ramification points in  $C(x)$ ) to end the procedure on a convex polyhedron  $P$  with  $n$  vertices, if the point  $x$  varies on  $P$ ?

Consider now the continuous flattening process. It is based on a point continuously moving along  $C(x)$ . Suppose the movement is at constant speed. Then the flattening time is proportional to the length  $\lambda C(x)$  of  $C(x)$ , so it seems of particular interest to find lower and upper bounds on  $\lambda C(x)$ .

*Question 2.* Can one locate on each  $P$  a point  $x$  with minimal length cut locus?

The starting point of our investigation is the question of Erik Demaine et al. (see Open Problem 18.1 in [5]), on the existence of continuous folding processes for all flat folded states of (not necessarily convex) polyhedra. This problem remains open, and can be rephrased –and widened– in a different framework, as follows. Consider an *abstract convex polyhedron*  $P$  (i.e., one obtained according to Alexandrov's gluing theorem). It has a unique *isometric embedding* in  $\mathbb{R}^3$  as a convex surface, but many other non-convex *realizations* in  $\mathbb{R}^3$  (see [8] for the precise definitions).

*Question 3.* Let  $\mathcal{R}$  denote the space of all realizations of  $P$  in  $\mathbb{R}^3$ , with the topology induced by the Hausdorff metric on the space of all compact sets in  $\mathbb{R}^3$ . Is  $\mathcal{R}$  arcwise connected?

**Acknowledgement.** The authors are indebted to Joseph O'Rourke for his careful reading of a preliminary version of this paper and his valuable suggestions, particularly improving the proof of Lemma 5.

## References

1. Agarwal, P.K., Aronov, B., O'Rourke, J., Schevon, C.A.: Star unfolding of a polytope with applications. *SIAM J. Comput.* 26, 1689–1713 (1997)
2. Alexandrov, A.D.: Convex Polyhedra. Monographs in Mathematics. Springer, Berlin (2005); Translation of the 1950 Russian edition by Dairbekov, N.S., Kusteladze, S.S., Sossinsky, A.B.
3. Chen, J., Han, Y.: Shortest paths on a polyhedron. In: Proc. 6th Ann. ACM Sympos. Comput. Geom., pp. 360–369 (1990)
4. Demaine, E.D., Demaine, M.D., Lubiw, A.: Flattening polyhedra (2001) (unpublished manuscript)
5. Demaine, E.D., O'Rourke, J.: Geometric folding algorithms, Lincages, Origami, Polyhedra. Cambridge University Press (2007)
6. Itoh, J.-I., Nara, C.: Continuous Flattening of Platonic Polyhedra. In: Akiyama, J., Bo, J., Kano, M., Tan, X. (eds.) CGGA 2010. LNCS, vol. 7033, pp. 108–121. Springer, Heidelberg (2011)
7. Kaneva, B., O'Rourke, J.: An implementation of Chen and Han's segment algorithm. In: Proc. 12th Canadian Conf. Comput. Geom., pp. 139–146 (2000)
8. Pak, I.: Inflating polyhedral surfaces (2006),  
<http://www.math.ucla.edu/~pak/papers/pillow4.pdf>
9. Sabitov, I.: The volume of polyhedron as a function of its metric. *Fundam. Prikl. Mat.* 2(4), 1235–1246 (1996)
10. Sabitov, I.: The volume as a metric invariant of polyhedra. *Discrete Comput. Geom.* 20, 405–425 (1998)

# Convexifying Monotone Polygons while Maintaining Internal Visibility

Oswin Aichholzer<sup>1,\*</sup>, Mario Cetina<sup>2</sup>, Ruy Fabila-Monroy<sup>3,\*\*</sup>, Jesús Leaños<sup>4</sup>,  
Gelasio Salazar<sup>5,\*\*\*</sup>, and Jorge Urrutia<sup>6,†</sup>

<sup>1</sup> Institute for Software Technology, University of Technology, Graz, Austria  
[oaich@ist.tugraz.at](mailto:oaich@ist.tugraz.at)

<sup>2</sup> Instituto de Física, Universidad Autónoma de San Luis Potosí, México  
[mcetina@gmail.com](mailto:mcetina@gmail.com)

<sup>3</sup> Departamento de Matemáticas, Cinvestav, México  
[ruyfabila@math.cinvestav.edu.mx](mailto:ruyfabila@math.cinvestav.edu.mx)

<sup>4</sup> Unidad Académica de Matemáticas, Universidad Autónoma de Zacatecas, México  
[jleanos@mate.reduaz.mx](mailto:jleanos@mate.reduaz.mx)

<sup>5</sup> Instituto de Física, Universidad Autónoma de San Luis Potosí, México  
[gsalazar@ifisica.uaslp.mx](mailto:gsalazar@ifisica.uaslp.mx)

<sup>6</sup> Instituto de Matemáticas, Universidad Nacional Autónoma de México  
[urrutia@matem.unam.mx](mailto:urrutia@matem.unam.mx)

**Abstract.** Let  $P$  be a simple polygon on the plane. Two vertices of  $P$  are visible if the open line segment joining them is contained in the interior of  $P$ . In this paper we study the following questions posed in [8,9]: (1) Is it true that every non-convex simple polygon has a vertex that can be continuously moved such that during the process no vertex-vertex visibility is lost and some vertex-vertex visibility is gained? (2) Can every simple polygon be convexified by continuously moving only one vertex at a time without losing any internal vertex-vertex visibility during the process?

We provide a counterexample to (1). We note that our counterexample uses a monotone polygon. We also show that question (2) has a positive answer for monotone polygons.

**Keywords:** convexification, monotone polygons, visibility graph.

## 1 Introduction

Let  $P$  be a simple polygon with vertices  $\{p_1, \dots, p_n\}$ . We say that two vertices of  $P$  are *P-visible* if the relative interior of the line segment joining them is contained in the interior of  $P$ . The *visibility graph*  $VG(P)$  of  $P$  is the graph

---

\* Partially supported by FWF [Austrian Fonds zur Förderung der Wissenschaftlichen Forschung] under grant S9205-N12, NFN Industrial Geometry.

\*\* Partially supported by CONACYT grant 153984.

\*\*\* Supported by CONACYT Grant 106432.

† Partially supported by project SEP-CONACYT of México, Proyecto 80268.

with vertex set  $\{p_1, \dots, p_n\}$  in which two vertices of  $P$  are adjacent if they are  $P$ -visible. A classical problem in computational geometry is that of convexifying simple polygons; that is, using a given fixed set of transformations that can be applied to the vertices and edges of  $P$ , try to transform  $P$  into a convex polygon in such a way that some properties of  $P$  are preserved. The first formulation of a problem of this kind was proposed by Erdős [10], who proposed a strategy to convexify a non-convex polygon by using *flips*; see also [4,3,5,6,11,12].

Our starting point is the following question posed by Satyan L. Devadoss in the Open Problem Session at CCCG 2008 [8,9]:

*Question 1.* Given a simple polygon  $P$  and its visibility graph  $VG(P)$ , can the vertices of  $P$  be moved continuously (one at a time or simultaneously) along paths so that:

- the simplicity of the polygon  $P$  is maintained all the time, and
- the visibility graph of  $P$  never loses edges, only gains them.

In discussions following the workshop, the following two specific questions were raised [7]:

1. Is it true that every non-convex simple polygon has a vertex  $p$  that can be continuously moved so that while  $p$  moves,  $VG(P)$  gains at least one extra edge, and never loses any?
2. Can every simple polygon be convexified by continuously moving only one vertex at a time such that  $VG(P)$  never loses any edge?

Our aim is to prove that Question (2) has a positive answer for monotone polygons. On the other hand, we give an example that shows that the answer to Question (1) is negative, even for monotone polygons. For recent results on this topic, see also [9].

## 1.1 Polygons and Visibility

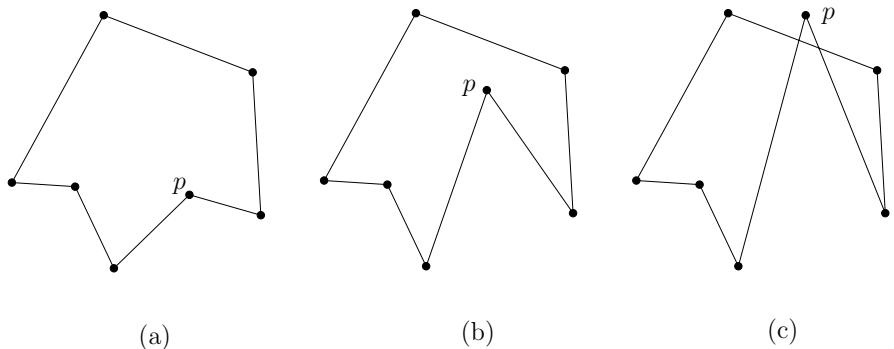
Let  $\{p_0, \dots, p_{n-1}\}$  be a set of points in the plane. A simple polygon  $P$  with vertex set  $\{p_0, \dots, p_{n-1}\}$  is a non-crossing piecewise linear curve formed by the union of the closed line segments  $\overline{p_ip_{i+1}}$  joining  $p_i$  to  $p_{i+1}$ ,  $i = 0, \dots, n - 1$ , addition taken mod  $n$ .

We will suppose without loss of generality that no vertical line passes through two vertices of  $P$ . A polygon  $P$  is *monotone* if any vertical line intersects  $P$  in at most two points. Let  $u$  and  $v$  be the leftmost and rightmost vertices of  $P$ . Clearly there are two edge-disjoint paths contained in  $P$  joining  $u$  to  $v$ , one above the other. The first will be called the *upper chain* of  $P$ , the second the *lower chain* of  $P$ .

A basic operation that we will use in this paper is that of moving the elements of  $\{p_0, \dots, p_{n-1}\}$  around the plane. As the vertices of  $P$  move, strictly

speaking the polygon  $P$  defined by its vertices changes, nevertheless, abusing our terminology a bit, we will always refer to it as  $P$ . In other words, as the elements of  $\{p_0, \dots, p_{n-1}\}$  move around the plane, they and  $P$  *keep their respective identities*. We will restrict our point moves to those that do not destroy the simplicity of  $P$ .

In Figure 1(b), we illustrate the change that the polygon shown in Figure 1(a) undergoes as point  $p$  moves. Moving  $p$  from its initial position in Figure 1(a) to its position in Figure 1(c) is not allowed.



**Fig. 1.** (a) A polygon  $P$ . (b) A valid vertex move on  $P$ . (c) An invalid vertex move (simplicity is lost).

If  $P$  is a polygon, then  $\text{CH}(P)$  denotes the convex hull of  $P$ . A vertex of  $P$  that lies on the boundary of  $\text{CH}(P)$  is called a *convex hull vertex*, otherwise it is called an *interior vertex*. Let  $V^\bullet(P)$  and  $V^\circ(P)$  respectively denote the set of convex hull and interior vertices of  $P$ . Thus  $V(P) = V^\circ(P) \cup V^\bullet(P)$ .

A simple polygon  $P$  divides  $\mathbb{R}^2 \setminus P$  into two regions, a bounded region called the *interior of  $P$* , and an unbounded region called the *exterior of  $P$* . We say the two vertices  $u$  and  $v$  of  $P$  are  $P$ -visible if the relative interior of the line segment  $\overline{uv}$  joining them is contained in the interior of  $P$ . We call  $\{u, v\}$  a *visibility pair*. Note that according to our definition, consecutive vertices of  $P$  are not visible. Let  $\mathcal{N}(P)$  be the set of pairs of vertices of  $P$  that are not  $P$ -visible. As consecutive vertices of  $P$  are not  $P$ -visible,  $|\mathcal{N}(P)| \geq n$ .

The visibility graph  $VG(P)$  is the graph with vertex set  $\{p_0, \dots, p_{n-1}\}$  in which two vertices are adjacent if they form a visibility pair. Note that if the vertices of  $P$  move, the set of visible pairs of  $P$  may change, and in turn  $VG(P)$  may also change.

In this paper, we will be mainly concerned with the changes that polygons undergo as we move their vertices along line segments. Most of the time, we move one vertex at a time, and along a line segment.

A *point-move* operation on  $P$  is the translation of a vertex of  $P$ , say  $p_i$ , (the moving point) from an initial position on the plane  $(x_i, y_i)$  to a final position  $(x'_i, y'_i)$  along the line segment joining  $(x_i, y_i)$  to  $(x'_i, y'_i)$ .

We say that a vertex move is *visibility-preserving*, if the following condition holds:

- If  $p_j$  and  $p_k$  were  $P$ -visible, they remain  $P$ -visible while  $p_i$  moves (this condition also applies for visibilities between  $p_i$  and other vertices of  $P$ ).

If in addition the following is satisfied:

- The number of edges of  $VG(P)$  increases,

we call it a *visibility-increasing* vertex move.

Our main results here are the following:

**Theorem 1.** *There are non-convex polygons that have no visibility-increasing vertex moves.*

and

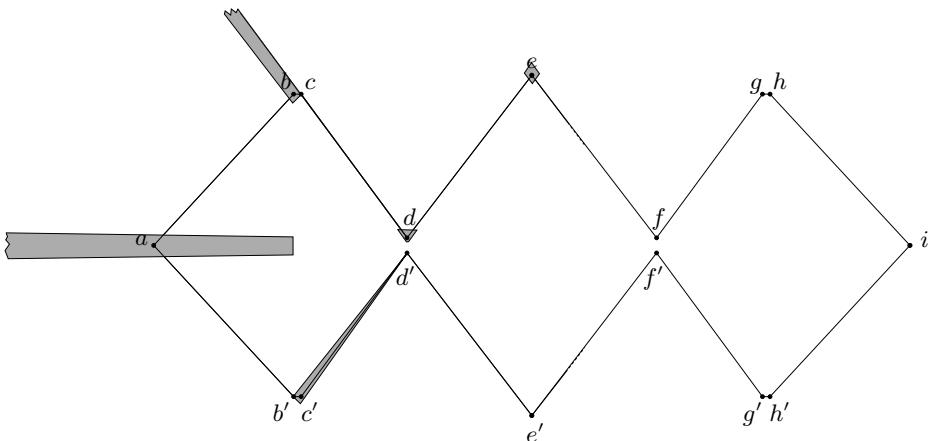
**Theorem 2.** *Every monotone polygon can be convexified with a sequence of visibility-preserving moves.*

## 2 A Counterexample to Question (1)

*Proof of Theorem 1.* Consider the monotone polygon  $P$  shown in Figure 2. The coordinates of the vertices of  $P$  are  $a = (-100, 0)$ ,  $b = (-63, 40)$ ,  $c = (-61, 40)$ ,  $d = (-33, 2)$ , and  $e = (0, 45)$ . The points  $\{f, g, h, i\}$  are obtained from the points  $\{a, b, c, d\}$  by reflecting them along the  $y$ -axis. Points  $b'$  to  $h'$  are obtained from the points  $b$  to  $h$  by a reflection along the  $x$ -axis.

To show that  $P$  does not admit any visibility-increasing vertex move, it is sufficient to consider the vertices of  $P$  in the set  $\{a, b, c', d, e\}$ . The remaining cases follow by symmetry. For each of these vertices, we show in Figure 2 the open shaded region into which any of these points can be moved without losing any visibility pairs in  $P$ .

For example, the region  $R$  into which vertex  $a$  can move without destroying any visibility pair of  $P$  is bounded by three lines; the one joining points  $d$  and  $f$ , the line joining  $d'$  to  $f'$ , and the line through points  $b$  and  $b'$ . If we move  $a$  above the line joining  $d$  to  $f$ , we lose visibility between  $a$  and  $f$ . If we move  $a$  to the right of the line joining  $b$  and  $b'$ , then the visibility between  $b$  and  $b'$  is lost. Lastly, if we move  $a$  below the line joining  $d'$  and  $f'$ , we again lose visibility between  $a$  and  $f'$ . Moreover, it is easy to see that when we move  $a$  around  $R$ , no additional internal visible pair of  $P$  can be gained. A similar case analysis is easily done for the other vertices of  $P$ . This completes the proof of Theorem 1.



**Fig. 2.** A monotone polygon without visibility-increasing vertex moves. Shaded areas indicate visibility-preserving regions.

### 3 Visibility-Preserving Vertex Moves

In this section we establish some basic properties of visibility-preserving vertex moves. They will be used in the next section to prove our second main result. For a point  $q \in \mathbb{R}^2$  and some  $\delta > 0$ , we denote by  $B_\delta(q)$  the closed disk with radius  $\delta$  with center at point  $q$ .

Let  $P = \{p_0, \dots, p_{n-1}\}$  be a set of points on the plane in general position (no three collinear). We say that  $\delta > 0$  is a *safe threshold* of  $P$  if there are no three elements  $p_i, p_j$ , and  $p_k$  of  $P$  such that  $B_\delta(p_i), B_\delta(p_j)$ , and  $B_\delta(p_k)$  are all intersected by a line. Equivalently, we can say that  $\delta$  is a safe threshold of  $P$  if there are no three points  $p_i, p_j, p_k \in P$  such that when we translate each of them to a point within  $\delta$  distance of them, they become aligned.

The following observations are evident, but useful.

**Observation 1.** *Every point set  $P$  in the plane in general position has a safe threshold  $\delta$ , for some  $\delta > 0$ .*

**Observation 2.** *Let  $P$  be a polygon such that  $V(P) = \{p_0, \dots, p_{n-1}\}$  is in general position. If a vertex move of a vertex  $p_i$  of  $P$  is not visibility-preserving, then at some point while  $p_i$  is moving, it becomes collinear with two other vertices of  $V(P)$ .*

The next proposition follows immediately from these two observations.

**Proposition 1.** *Let  $P$  be a polygon such that  $V(P)$  is in general position, and let  $\delta$  be a safe threshold of  $V(P)$ . Let  $S = \{p_{\sigma(1)}, \dots, p_{\sigma(m)}\}$  be any subset of vertices of  $P$ . Then if one at a time, each element  $p_{\sigma(i)}$  of  $S$  is moved from its original position to a new position within distance  $\delta$  along a line segment, then every vertex move is visibility-preserving.*

The following lemma will be useful to prove Theorem 2.

**Lemma 1.** *Let  $P$  be a monotone polygon. Then there is a sequence of visibility-preserving vertex moves of some vertices of  $P$  such that at the end of the sequence, the vertices of  $P$  are in general position,  $P$  remains monotone, and  $|V^\circ(P)| + |\mathcal{N}(P)|$  does not increase during the vertex movements.*

*Proof.* Recall that we are assuming that no vertical line passes through two vertices of  $P$ . We proceed by induction on the number of *collinear triples* in  $V(P)$ . Let  $\text{tri}(P)$  denote this number of collinear triples in  $V(P)$ . A vertex of  $P$  is called *weak* if it is part of a collinear triple. If  $\text{tri}(P) = 0$  there is nothing to prove. Observe now that for any  $p_i \in P$ , if we move it a small distance, say  $\epsilon_i$ , up or down along the vertical line through it, and in such a way that it does not cross any line joining two points of  $P$ , then:

- we do not create any new collinear triple, and
- if  $p_i$  belongs to a collinear triple  $p_i, p_j, p_k$ , these points cease to form a collinear triple.

Suppose then that  $\text{tri}(P) \geq 1$ , and that a vertex  $p_i$  of  $P$  is part of one such triple. Assume without loss of generality that  $p_i$  belongs to the upper chain of  $P$  and that it is the *leftmost vertex* of the chain that participates in a collinear triple.

If  $p_i$  is in  $V^\bullet(P)$ ; that is,  $p_i$  belongs to the convex hull of  $P$ , then move  $p_i$  vertically down by a distance  $\epsilon_i$ ,  $\epsilon_i$  as described above, otherwise move  $p_i$  vertically up by a distance  $\epsilon_i$ . In both cases  $\text{tri}(P)$  is reduced by at least one, and the sets  $V^\circ(P)$  and  $V^\bullet(P)$  remain unchanged. In the first case,  $|\mathcal{N}(P)|$  might be reduced by one. Our result follows by induction on  $\text{tri}(P)$ .

## 4 Convexifying Monotone Polygons

*Proof of Theorem 2.* By Lemma 1, we can assume that  $V(P)$  is in general position. We proceed by induction on the sum of the number of interior vertices plus the number of non-visible pairs. If the vertices of  $P$  are in convex position, there is nothing to prove. Observe that  $P$  is convex if  $|V^\circ(P)| + |\mathcal{N}(P)| = n$ . Suppose then that  $|V^\circ(P)| + |\mathcal{N}(P)| > n$  and assume that the theorem holds for all polygons  $Q$  with  $|V^\circ(Q)| + |\mathcal{N}(Q)| < |V^\circ(P)| + |\mathcal{N}(P)|$ .

Since  $P$  is not convex, it follows that  $V^\circ(P)$  is not empty. Suppose without loss of generality that there are  $k \geq 1$  interior vertices of  $P$  on its upper chain. Relabel them as  $v_1, v_2, \dots, v_k$ , in increasing order with respect to their  $x$ -coordinate.

Let  $\delta > 0$  be a safe threshold for the *initial position* of  $V(P)$ . Our algorithm will initially execute the following basic procedure **BP** once:

**BP:** *One at a time from left to right, move  $v_1, v_2, \dots, v_k$  upwards, by a distance  $\delta$ .*

Once  $v_1, v_2, \dots, v_k$  have all been moved, we execute **BP** repeatedly until one of the following occurs:

1. a vertex in  $\{v_1, v_2, \dots, v_k\}$  reaches the convex hull of  $P$ ,
2. a new visible pair occurs,
3. or the visibility-preserving property is lost.

If we stop because (1) or (2) occurs, then we are done, by our induction hypothesis. We will show now that (3) cannot happen, since before it happens there must be a visibility-increasing event. This will prove our result.

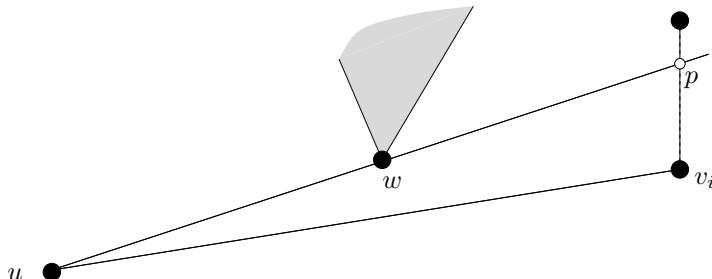
Observe that during the first iteration of **BP**, we are moving  $v_1, v_2, \dots, v_k$  upwards by a distance  $\delta$  which is the threshold of the initial configuration of  $P$ . Thus no collinear triple contained in  $\{v_1, v_2, \dots, v_k\}$  arises at any time during the first execution of **BP**. Observe also that at the end of each execution of **BP**, if we consider only the elements of  $\{v_1, v_2, \dots, v_k\}$ ,  $\delta$  is a safe threshold of them. In particular, this implies that if a visibility is lost during the execution of **BP**, it must involve one or two elements of  $P$  not in  $\{v_1, v_2, \dots, v_k\}$ . We summarize this in the following observation:

**Observation 3.** *During none of the next executions of **BP**, is the visibility between any pair of points in  $\{v_1, v_2, \dots, v_k\}$  blocked by a third element in  $\{v_1, v_2, \dots, v_k\}$ .*

Suppose then that **BP** stops because (3) occurs before any event of type (1) or (2). Suppose that the *critical move*, i.e. the vertex move in which a visibility loss occurs for the first time when vertex  $v_i$ ,  $i \in \{1, 2, \dots, k\}$ , is moved upwards during the  $r$ -th iteration of **BP**,  $r \geq 2$ .

Thus before the critical move takes place, each of  $v_1, v_2, \dots, v_{i-1}$  has been moved upwards vertically  $r \geq 2$  times, moving in total a distance  $r\delta$ . Each of  $v_{i+1}, v_{i+2}, \dots, v_k$  has been moved upwards vertically  $(r-1)$  times.

Since during the critical move when  $v_i$  moved from its second-to-last position to its last position a visibility of  $P$  was lost, at some point in time  $v_i$  passed over a point  $p$  collinear with two vertices of  $P$ ; call them  $u$  and  $w$ , such that one of them, say  $w$ , blocked the visibility between  $p$  and  $u$ ; see Figure 3.



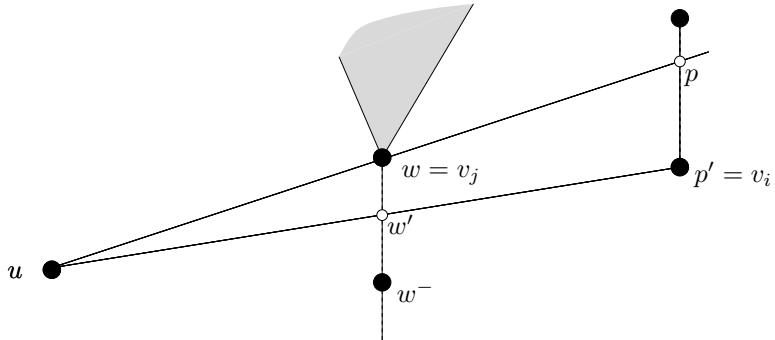
**Fig. 3.** When  $v_i$  passes over  $p$ ,  $w$  blocks the visibility between  $v_i$  and  $u$

By Observation 3, at least one of  $u$  or  $w$  belongs to the convex hull of  $P$ . We now show that  $w \in \{v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_{k-1}, v_k\}$ .

Obviously  $w$  can not be in the lower chain of  $P$ . It remains to show that  $w$  does not belong to  $\text{CH}(P)$ . Let  $p'$  be the point where  $v_i$  was located before the critical move was executed. Assume to the contrary that  $w$  is in  $\text{CH}(P)$ . Since before  $v_i$  reached  $p$ , no event of type (1) took place, any point on the line segment joining  $p$  to  $p'$  is in the interior of  $\text{CH}(P)$ .

It is easy to see now that not only  $w$ , but also  $p$  and  $u$  must be in  $\text{CH}(P)$ . Since  $w$  is contained in the straight segment joining  $p$  and  $u$ , and since  $p$  is also in  $\text{CH}(P)$ , it follows that there must be another vertex  $z$  of  $\text{CH}(P)$  such that  $p$  is contained in the straight line segment from  $w$  to  $z$ . Thus  $p, u, w$ , and  $z$  are all collinear. But  $u, w$ , and  $z$  are distinct vertices of  $\text{CH}(P)$ , contradicting that  $V(P)$  is in general position. Thus  $w = v_j$  for some  $1 \leq j \leq k$ ,  $j \neq i$  as claimed.

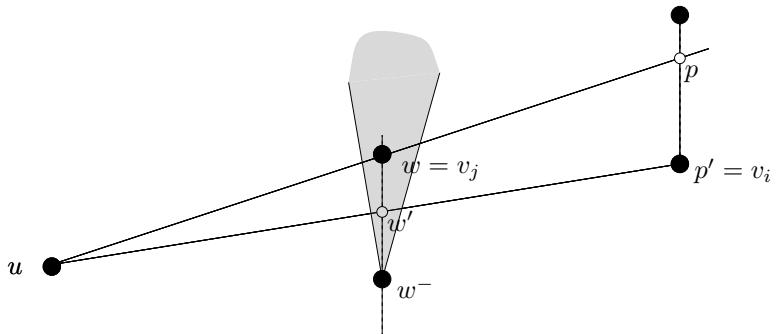
Since  $w$  and  $v_i$  belong to  $\{v_1, v_2, \dots, v_k\}$ , then by Observation 3,  $u$  cannot belong to this set. Thus  $u$  belongs to the lower chain of  $P$  or to  $\text{CH}(P)$ . To finish the proof it suffices to show that some vertex move performed before the so-called critical move was visibility-increasing.



**Fig. 4.** In the polygon  $P_\lambda$ ,  $u$  is visible from  $p' = v_i$

Consider the triangle whose vertices are  $p, u$ , and  $p'$ ; see Figure 4. Let  $L$  be the vertical line that passes through  $w$ , and let  $w'$  be the point where  $L$  crosses  $up'$ . Since  $pp'$  has length  $\leq \delta$ , it follows that the length of the segment  $ww'$  is strictly less than  $\delta$ . Therefore the point  $w^-$  that lies a distance  $\delta$  vertically below  $w$  lies below  $w'$ . Let  $\lambda$  denote the vertex move that took  $v_j$  from  $w^-$  to  $w$ , and let us denote by  $P_\lambda$  the position of polygon  $P$  before  $\lambda$  was executed. We remark that  $w^-$ , and hence  $\lambda$  and  $P_\lambda$ , are well defined, since every vertex was moved upwards at least once by a distance  $\delta$  before any collinearity arose.

The monotonicity of  $P_\lambda$  implies that the open vertical ray above  $w^-$  is in the exterior of  $P_\lambda$ . Since this open vertical ray intersects  $up'$  (namely at  $w'$ ),



**Fig. 5.** In the polygon  $P_\lambda$ ,  $u$  is not  $P_\lambda$ -visible from  $p' = v_i$ . Therefore one of the moves involved in taking  $P_\lambda$  into  $P$  must be visibility-increasing.

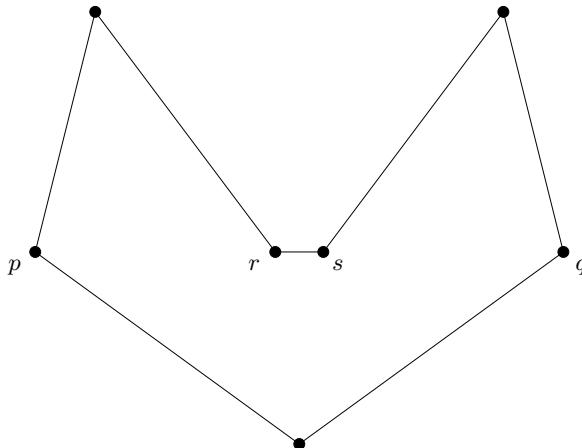
it follows that  $u$  is not  $P_\lambda$ -visible from  $p'$ ; see Figure 5. Since at the beginning of the critical move  $u$  is  $P$ -visible from  $p'$ , it follows that during one of the vertex moves executed between  $\lambda$ , call it  $\lambda'$ , and the critical event,  $v_i$  and  $u$  became visible, thus  $\lambda'$  is visibility-increasing. Our result follows.

## 5 Conclusion and Open Problems

Although the proof of Theorem 2 is essentially algorithmic, it does not yield an efficient algorithm to convexify monotone polygons. In fact, the number of vertical vertex moves it requires can be arbitrarily large. For example for the polygon  $P$  shown in Figure 6, the number of steps required to convexify it depends on the distance between the line passing through vertices  $r$  and  $s$ , and the line through  $p$  and  $q$ . The closer these lines are, the more vertex moves we will have to execute to convexify  $P$ . The main problem is the size of the safe thresholds involved at each iteration of the main procedure. On the other hand, it is not hard to see that if we allow horizontal and vertical vertex moves, then the polygon in Figure 6 can be convexified with a constant number of vertex moves. Can this problem be solved? We believe that this is not always possible. We conjecture:

*Conjecture 1.* There is a positive integer  $n_0$  with the following property. For each integer  $n \geq n_0$  and each  $M$ , there exists an  $n$ -vertex monotone polygon  $P$  such that the number of visibility-preserving vertex moves required to convexify  $P$  is at least  $M$ .

What happens if we allow more than one vertex to move at a time? In a  $k$ -vertex move, we allow up to  $k$  vertices to move simultaneously. It is worth mentioning that every monotone polygon on  $n$  vertices can be convexified, preserving visibility, with only one  $(n - 3)$ -vertex move. Indeed if all the vertices in the interior of  $P$  in the lower (respectively upper) chain of  $P$  move downwards (respectively



**Fig. 6.** The closer  $\overline{rs}$  is to  $\overline{pq}$ , the more vertex moves are needed to convexify  $P$

upwards) towards  $\text{CH}(P)$  (each vertex stopping as soon as it reaches  $\text{CH}(P)$ ), the resulting  $(n - 3)$ -vertex move convexifies  $P$ , and is visibility-preserving.

Thus the following question arises:

*Question 2.* Determine the complexity of the problem of convexifying a monotone polygon with  $m$ -vertex moves,  $1 \leq m \leq n - 3$ .

What happens to Conjecture 1 and Question 2 if instead of monotone polygons, we consider simple polygons? Devadoss *et al.* [9] showed how to convexify star-shaped polygons by moving all the vertices in the polygon simultaneously. We believe that our techniques should extend to star-shaped polygons, but so far we have not been able to do it.

To conclude we mention that we believe that question (2) has a positive answer:

*Conjecture 2.* Every simple polygon can be convexified by a sequence of visibility-preserving 1-vertex moves.

**Note:** During the final round of revisions of this paper, progress was made on the problem of convexifying simple (not necessarily monotone) polygons while maintaining internal visibility. In [1] it was shown that if there is a convexifying transformation of a simple polygon moving any number of vertices at a time, then there is a convexifying transformation moving one vertex at a time. Finally Conjecture 2 was shown to be true in [2].

**Acknowledgements.** We would like to thank Erik Demaine and Stefan Langerman for valuable discussions on the topic.

## References

1. Ábrego, B.M., Cetina, M., Leaños, J., Salazar, G.: Visibility-preserving convexifications using single-vertex moves. *Information Processing Letters* 112(5), 161–163 (2012)
2. Aichholzer, O., Aloupis, G., Demaine, E.D., Demaine, M.L., Dujmović, V., Hurtado, F., Lubiwi, A., Rote, G., Schulz, A., Souvaine, D.L., Winslow, A.: Convexifying Polygons Without Losing Visibilities. In: Proceedings of the 23rd Canadian Conference on Computational Geometry (CCCG 2011), pp. 229–234 (2011)
3. Aichholzer, O., Cortes, C., Demaine, E.D., Dujmović, V., Erickson, J., Meijer, H., Overmars, M., Palop, B., Ramaswami, S., Toussaint, G.: Flipturning polygons. *Discrete and Computational Geometry* 28, 231–253 (2002)
4. Aloupis, G., Ballinger, B., Bose, P., Damian, M., Demaine, E.D., Demaine, M.L., Flatland, R.Y., Hurtado, F., Langerman, S., O'Rourke, J., Taslakian, P., Toussaint, G.T.: Vertex Pops and Popturns. In: Proceedings of the 19th Canadian Conference on Computational Geometry (CCCG 2007), pp. 137–140 (2007)
5. Biedl, T.C., Demaine, E.D., Lazard, S., Robbins, S.M., Soss, M.A.: Convexifying Monotone Polygons. In: Aggarwal, A.K., Pandu Rangan, C. (eds.) ISAAC 1999. LNCS, vol. 1741, pp. 415–424. Springer, Heidelberg (1999)
6. Demaine, E.D., Gassend, B., O'Rourke, J., Toussaint, G.T.: All polygons flip finitely... right? In: Surveys on Discrete and Computational Geometry. Contemp. Math., vol. 453, pp. 231–255 (2008)
7. Demaine, E.D., Langerman, S.: Personal communication, Montreal (2008)
8. Demaine, E.D., O'Rourke, J.: Open Problems from CCCG 2008. In: Proceedings of the 21st Canadian Conference on Computational Geometry (CCCG 2009), pp. 75–78 (2009)
9. Devadoss, S.L., Shah, R., Shao, X., Winston, E.: Visibility graphs and deformations of associahedra, [arXiv:0903.2848](https://arxiv.org/abs/0903.2848) (March 2009)
10. Erdős, P.: Problem 3763. *Amer. Math. Monthly* 42, 627, 463–470 (1935)
11. Grünbaum, B.: How to convexify a polygon. *Geocombinatorics* 5, 24–30 (1995)
12. de Sz.-Nagy, B.: Solution of problem 3763. *Amer. Math. Monthly* 49, 176–177 (1939)

# On the Number of Radial Orderings of Colored Planar Point Sets

José Miguel Díaz-Báñez<sup>1,\*</sup>, Ruy Fabila-Monroy<sup>2,\*\*</sup>,  
and Pablo Pérez-Lantero<sup>3,\*\*\*</sup>

<sup>1</sup> Departamento Matemática Aplicada II, Universidad de Sevilla, Seville, Spain  
[dbanez@us.es](mailto:dbanez@us.es)

<sup>2</sup> Departamento de Matemáticas, Centro de Investigación y Estudios Avanzados del  
Instituto Politécnico Nacional, Mexico City, Mexico  
[ruyfabila@math.cinvestav.edu.mx](mailto:ruyfabila@math.cinvestav.edu.mx)

<sup>3</sup> Escuela de Ingeniería Civil en Informática, Universidad de Valparaíso,  
Valparaíso, Chile  
[pablo.perez@uv.cl](mailto:pablo.perez@uv.cl)

**Abstract.** Let  $n$  be an even natural number and let  $S$  be a set of  $n$  red and  $n$  blue points in general position in the plane. Let  $p \notin S$  be a point such that  $S \cup \{p\}$  is in general position. A *radial ordering* of  $S$  with respect to  $p$  is a circular ordering of the elements of  $S$  by angle around  $p$ . A *colored radial ordering* is a radial ordering of  $S$  in which only the colors of the points are considered. We show that: the number of distinct radial orderings of  $S$  is at most  $O(n^4)$  and at least  $\Omega(n^2)$ ; the number of colored radial orderings of  $S$  is at most  $O(n^4)$  and at least  $\Omega(n)$ ; there exists sets of points with  $\Theta(n^4)$  colored radial orderings and sets of points with only  $O(n^2)$  colored radial orderings.

**Keywords:** radial orderings, colored point sets, star polygonizations.

## 1 Introduction

Throughout this paper  $n$  is an even natural number and  $S$  is a set of  $n$  red and  $n$  blue points in general position in the plane. Let  $p$  be a point not in  $S$ , such that  $S \cup \{p\}$  is also in general position; we call  $p$  an *observation point*. A *radial ordering* of  $S$  with respect to  $p$  is a circular clockwise ordering of the elements of  $S$  by their angle around  $p$ . A *colored radial ordering* with respect to  $p$  is the circular list of the colors of the points in  $S$  in their radial ordering with respect to  $p$ . Thus permutations between points of the same color yield the same colored radial ordering. Unless otherwise noted, all point sets in this paper are in general position.

---

\* Partially supported by project MEC MTM2009-08652 and ESF EUROCORES programme EuroGIGA, CRP ComPoSe: grant EUI-EURC-2011-4306.

\*\* Partially supported by Conacyt of Mexico, grant 153984.

\*\*\* Partially supported by project MEC MTM2009-08652 and grant FONDECYT 11110069.

Let  $\rho(S)$  be the number of distinct radial orderings of  $S$  with respect to every observation point in the plane. Likewise let  $\tilde{\rho}(S)$  be the number of distinct colored radial orderings of  $S$  with respect to every observation point in the plane. We define the following functions:

$$\begin{aligned} f(n) &:= \max\{\rho(S) : |S| = 2n\} \\ \tilde{f}(n) &:= \max\{\tilde{\rho}(S) : |S| = 2n\} \\ g(n) &:= \min\{\rho(S) : |S| = 2n\} \\ \tilde{g}(n) &:= \min\{\tilde{\rho}(S) : |S| = 2n\} \end{aligned} \tag{1}$$

In this paper we show the following bounds.

$$\begin{aligned} f(n) &= \Theta(n^4) \\ \tilde{f}(n) &= \Theta(n^4) \\ \Omega(n^2) &\leq g(n) \leq O(n^4) \\ \Omega(n) &\leq \tilde{g}(n) \leq O(n^2) \end{aligned} \tag{2}$$

For the first equality, the fact that  $f(n)$  is  $O(n^4)$  has been noted before in the literature [5,10]. As far as we know all the other bounds are new.

A different problem but in the same setting has been studied recently in [10]. In that particular paper, the authors study what a robot can infer from its environment when all the information that is available is the cyclic positions of some landmarks as they appear from the robot's position. Other authors have considered problems of the same flavor, when a similar kind of information is available. See for example [6,7,9].

We point out that computing the radial ordering of  $S$  around every point in  $S$  is an unavoidable step in some geometric algorithms, as for example, doing a radial sweeping. Moreover, many optimization problems are solved by considering the arrangement of lines passing through two points in  $S$  and finding the optimum point inside each of the  $O(n^4)$  cells in the arrangement [8]. In many cases this is because the radial ordering of the points in  $S$  around every point  $p$  within a cell is the same. It could be interesting in this scenario to know how many cells induce the same radial ordering.

For a bi-colored point set, a radial sweeping algorithm also requires the ordering as an initial step, so it could be useful to know bounds on the number of different colored radial ordering of  $S$  from points in the plane. From the combinatorial point of view, this problem is related to partitioning bi-chromatic point sets by using *k-fans* [1,2]. A *k*-fan in the plane is a point  $p$  (called the center) and  $k$  rays emanating from  $p$ . This structure can be used to partition  $S$  into  $k$  monochromatic subsets and it depends only on the colored radial ordering of  $S$  with respect to  $p$ . The existence and non-existence of balanced *k*-fans for colored point sets have been studied in recent papers [4,3] but, as far as we know, the number of different monochromatic partitions induced by *k*-fans has not yet been considered. At a glance, the number of distinct colored radial orderings seems much smaller than the number of non-colored radial orderings. Surprisingly, in this paper we provide an example of a bi-colored set of  $2n$  points with  $\Theta(n^4)$  distinct colored radial orderings.

The assumption that  $S$  has the same number of red and blue points and that  $n$  is even, may seem arbitrary. However both of them are crucial hypothesis in our results. We elaborate on this in Section 3.

## 2 Bounds

We discretize the problem by partitioning the set of observation points into a finite number of sets so that two points in a same set induce the same radial ordering. This partition is made by half-lines, which if crossed by an observation point generate a transposition of two consecutive elements in the radial ordering. For every pair of points  $x_1, x_2 \in S$ , consider the line passing through them. Contained in this line we have two half-lines; one begins in  $x_1$  and does not contain  $x_2$  while the other begins in  $x_2$  and does not contain  $x_1$ . Two observation points are in the same element of the partition if they can be connected by a curve which does not intersect any of these half-lines. We call this partition the *order partition*. Since it induces a decomposition of the plane, we refer to its elements as *cells*. Note that if a point moves in a curve not crossing any half-line, the radial ordering with respect to this point is the same throughout the motion. Thus points in the same cell induce the same radial ordering.

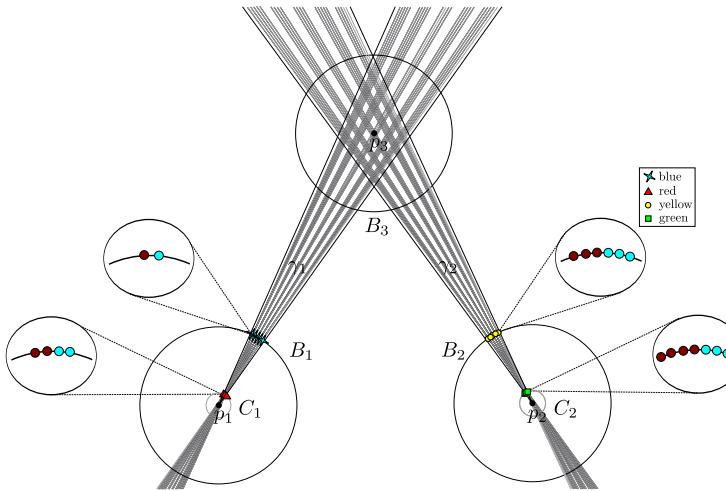
**Theorem 1.**  $f(n) \leq O(n^4)$  and  $\tilde{f}(n) \leq O(n^4)$ .

*Proof.* The first inequality follows from the fact that the number of cells in the order partition is  $O(n^4)$ . The second inequality follows from the first and from the observation that  $\tilde{\rho}(S) \leq \rho(S)$ .  $\square$

**Theorem 2.**  $g(n) \geq \Omega(n^2)$ .

*Proof.* Let  $\ell$  be a straight line having points of  $S$  both above and below. Let  $p$  and  $q$  be two points in  $\ell$ , such that when walking from  $p$  to  $q$  in a straight line a half-line is crossed. Let  $x_1, x_2 \in S$  be the points defining this half-line. We show that  $p$  and  $q$  induce different radial orderings. Note that  $x_1$  and  $x_2$  are both above or below  $\ell$ . Let  $x_3 \in S$  be on the side opposite to  $x_1$  and  $x_2$ . Assume without loss of generality that the radial ordering of  $\{x_1, x_2, x_3\}$  with respect to  $p$  is  $[x_1, x_2, x_3]$ . Since the half-line is crossed only once, the radial ordering of  $\{x_1, x_2, x_3\}$  with respect to  $q$  is  $[x_2, x_1, x_3]$ . Therefore, the radial ordering of  $S$  with respect to  $p$  is different from the radial ordering with respect to  $q$ . This implies that the number of different radial orderings with respect to points in  $\ell$  is equal to the number of half-lines intersecting  $\ell$  plus one. It remains to show that there is a choice for  $\ell$  crossing a quadratic number of half-lines. Choose  $\ell$  to be a line having only one point of  $S$  above and all the others below, not parallel to any half-line. Note that  $\ell$  intersects one of the two half-lines defined by every pair of points of  $S$  below  $\ell$ . Since there are  $2n - 1$  such points, they define  $\Theta(n^2)$  half-lines and the result follows.  $\square$

**Theorem 3.**  $f(n) \geq \Omega(n^4)$  and  $\tilde{f}(n) \geq \Omega(n^4)$ .



**Fig. 1.** A bi-colored set with  $\Omega(n^4)$  different colored radial orderings

*Proof.* We assume that  $n \geq 20$ . Given that  $\rho(S) \geq \tilde{\rho}(S)$ , it suffices to construct a set  $S$  of  $n$  red and  $n$  blue points such that  $\tilde{\rho}(S) = \Omega(n^4)$ . We start by constructing a four-colored set of points  $S'$  with  $\Omega(n^4)$  distinct colored radial orderings; afterwards we obtain  $S$  by replacing each point of a given color with a suitable “pattern” of red and blue points. These four patterns are chosen so that if they appear consecutively in a radial ordering, then any other equivalent radial ordering must match patterns of the same type. Since the patterns behave like the original four colors, the new set also has  $\Omega(n^4)$  colored radial orderings.

Let  $B_1$ ,  $B_2$  and  $B_3$  be three balls of radius  $1/4$ , whose centers  $p_1$ ,  $p_2$  and  $p_3$ , are the vertices of an equilateral triangle of side length equal to one. Let  $\varepsilon, \alpha > 0$ . Let  $C_1$  and  $C_2$  be circles of radius  $\varepsilon$  centered at  $p_1$  and  $p_2$ , respectively. Let  $\gamma_1$  and  $\gamma_2$  be infinite wedges of angle  $\alpha$ , with apices  $p_1$  and  $p_2$  respectively. Assume that  $\gamma_1$  is bisected by the line segment joining  $p_1$  and  $p_3$ , while  $\gamma_2$  is bisected by the line segment joining  $p_2$  and  $p_3$ . Refer to Figure 1. Let  $m$  and  $r$  be the only natural numbers such that  $n = 10m + r$  and  $10 \leq r \leq 19$ . Divide  $\gamma_1$  with  $m$  infinite rays emanating from  $p_1$  and intersecting  $B_3$ , such that the angle between two consecutive rays is  $\alpha/(m+1)$ . Do likewise for  $\gamma_2$ , with  $m$  infinite rays emanating from  $p_2$ . At every point of intersection of these rays with the boundary of  $B_1$ , place a blue point; at every point of intersection with  $C_1$  a red point; at every point of intersection with the boundary of  $B_2$  a yellow point; finally at every point of intersection with  $C_2$  a green point. Thus  $m$  points of each color are placed. This ends the construction of  $S'$ .

Let  $L$  be the set of lines passing through a red and a blue point. Let  $L'$  be the set of lines passing through a yellow and a green point. Choose  $\alpha$  and  $\varepsilon$  small enough so that the following conditions are met: (1) Every line in  $L \cup L'$  intersects the interior of  $B_3$  and these are the only lines passing through two

points of  $S'$  that do. (2) Neither two lines in  $L$  nor two lines in  $L'$  intersect at a point in the interior of  $B_3$ . (3) Every line in  $L$  intersects every line in  $L'$  at a point in  $B_3$ . By the previous conditions and the fact that  $|L| = m^2$  and  $|L'| = m^2$ ,  $L \cup L'$  splits  $B_3$  into precisely  $m^4$  cells. For each of these cells choose a point  $q_i$  in its interior. We show that the colored radial orderings of  $S'$  as seen from each of these points is different. Note that for each point in  $B_3$  there is a line separating the red and blue points from the green and yellow points. Thus we may assume, that the colored radial orderings as seen from points in  $B_3$  are written so that all the blue and red points appear before the green and yellow points. Consider the colored radial orderings that appear when walking from two distinct points  $q_i$  and  $q_j$  in a straight line. By conditions (1) and (2) the only transpositions that occur when the colored radial ordering changes is between a red and a blue point or between a yellow and a green point. This implies that the  $k$ -th red point will always be the same red point and that the number of blue points after the  $k$ -th red point is either increasing or decreasing monotonically; the same observation holds for the green and yellow points. Therefore in the walk once a line in  $L \cup L'$  is crossed all colored radial orderings afterwards will be distinct. Thus the number of different colored radial orderings of  $S'$  is at least  $m^4$ , which is  $\Omega(n^4)$ .

To construct  $S$ , we replace the points in  $S'$  by patterns of red and blue points, in such a way that the colored radial orderings at points  $q_i$  remain different. The points in the patterns replacing a point  $p \in S'$  are placed consecutively in the same circle containing  $p$ . If these points are placed close enough to  $p$ , then they will appear consecutively in the colored radial ordering with respect to every point  $q_i$ . The points of  $S'$  are replaced in the following way: every blue point with a pattern of one red and one blue point; every red point with a pattern of two red and two blue points; every yellow point with a pattern of three red and three blue points; and every green point with a pattern of four red and four blue points. Refer to Figure 1. Note that our choice of patterns implies that two equivalent radial orderings must match patterns of the same type. Also, if necessary, we may assume that  $\alpha$  is small enough so that this augmented set is in general position. So far  $10m$  red and  $10m$  blue points have been placed. The remaining  $2r$  points can be placed in such a way that in the radial ordering with respect to every point  $q_i$  these  $r$  red points appear consecutively followed by these  $r$  blue points. This final condition guarantees that the colored radial orderings at each  $q_i$  remain different.  $\square$

**Theorem 4.**  $\tilde{g}(n) \leq O(n^2)$ .

*Proof.* Recall that we are assuming that  $n$  is even. We employ a similar strategy as in the proof of Theorem 3. We start with a set  $S'$  of  $n/2$  points, placed almost evenly in the unit circle. All the points of  $S'$  have the same color and thus the colored radial orderings of  $S'$  are all equivalent. Afterwards, we replace each point of  $S'$  with a symmetric pattern of red and blue points. This is done in such a way that the new number of distinct colored radial orderings increases at most to  $O(n^2)$ .

Let then  $X$  be a set of  $n/2$  points placed evenly on the unit circle centered at the origin. Explicitly  $X := \{(\cos(\frac{4\pi}{n} \cdot i), \sin(\frac{4\pi}{n} \cdot i)) \mid i = 1, 2, \dots, n/2\}$ . To avoid degeneracies we choose  $S'$  to be a set of  $n/2$  points (of the same color) each arbitrarily close to a distinct point of  $X$ . Let  $L'$  be the set of lines passing through every pair of points of  $S'$  together with the lines passing through every point  $p$  of  $S'$  and tangent to the unit circle centered at the origin and passing through  $p$ . The degeneracy we wish to avoid is that not three of these lines intersect at a point not in  $S'$ . It is easy to see that such a set exists. Let  $L(S')$  be the line arrangement generated by  $S'$ .

Let  $\delta > 0$ . We form a new set  $S_\delta$ , by replacing each point  $p$  of  $S'$  with a pattern “red,blue,blue,red” of four points, placed clockwise consecutively at the same distance as  $p$  from the origin. The first (red) point is placed at  $p$ , the next three points are placed at distance  $\delta$  from the previous point. Note that for some small enough value  $\delta'$ , all line arrangements  $L(S_\delta)$  with  $\delta < \delta'$  are combinatorially equivalent. From now on assume that  $\delta < \delta'$  and let  $\delta$  tend to 0. The set of lines passing through each pair of points of  $S_\delta$  tends to  $L'$ . Likewise  $L(S_\delta)$  tends to  $L(S')$ . By this we mean that for each element (vertex, edge or cell)  $C$  of  $L(S_\delta)$  there is an element  $D$  of  $L(S')$ , such that for each open set  $\mathcal{O}$  containing  $D$  there is a small enough value for  $\delta$  so that  $C$  is contained in  $\mathcal{O}$ .

We calculate an upper bound on the number of different colored radial orderings of  $S_\delta$  by considering the colored radial orderings with respect to points in the interior of each cell of  $L(S')$ . Let  $C$  be a cell of  $L(S_\delta)$  and  $q$  be a point in the interior of  $C$ . There are three different cases according to the limit of  $C$  in  $L(S')$ :

- $C$  tends to a cell of  $L(S')$ .

In this case, every pattern replacing a point of  $S'$  will appear consecutively in the colored radial ordering around  $q$ . Moreover, by the symmetry of the patterns they are all “red,blue,blue,red”. Thus in this case there is only one possible colored radial ordering.

There are  $O(1)$  colored radial orderings in this case.

- $C$  tends to an edge of  $L(S')$ .

We distinguish two subcases; whether the edge is contained in one of the lines passing through two points  $p_i$  and  $p_j$  of  $S'$  or whether it is contained in one of the tangent lines passing through a point  $p_k$  of  $S'$ .

In the first subcase, the patterns at points different from  $p_i$  and  $p_j$  will appear consecutively in the colored radial ordering. However at  $p_i$  and  $p_j$ , the points in the patterns will appear together but intermixed. Since there are only 8 points involved, there is only a constant number of ways in which this can happen. The second subcase is similar.

There are  $O(1)$  colored radial orderings in both subcases.

- $C$  tends to a vertex of  $L(S')$ .

In this case,  $C$  may tend to a vertex that is a point of  $S'$  or the intersection of two lines  $\ell_1$  and  $\ell_2$  of  $L'$ .

Suppose that  $C$  tends to a point  $p_i$  of  $S'$ . Then  $C$  must be bounded by lines passing through points of the pattern at  $p_i$  and points of  $S_\delta$ . Note that

in total there are at most  $8n$  such lines and thus there are at most  $O(n^2)$  such cells. Moreover, since we may assume that  $S'$  is arbitrarily close to  $X$ , our analysis does not depend on the choice of  $p_i$ . Hence, the cells tending to any other point of  $S'$  will induce the same set of colored radial orderings.

In the second case the lines  $\ell_1$  and  $\ell_2$  may be defined by two points or one point of  $S'$ . In both cases we have that the patterns at the points of  $S'$  defining each line appear together but intermixed. The patterns at any other points will appear consecutively. Since there is at most four points defining  $\ell_1$  and  $\ell_2$ , the number of ways in which their respective pattern points can appear is at most a constant. The only thing left to consider is the  $O(n)$  number of ways in which the patterns at the points defining  $\ell_1$  can appear with respect to those of  $\ell_2$ .

There are  $O(n^2)$  colored radial orderings in the first case and  $O(n)$  in the second.

Note that the total number of distinct colored radial orderings of  $S_\delta$  is at most  $O(n^2)$ . Therefore by setting  $S := S_\delta$ , the result follows.  $\square$

We now give a linear lower bound on the number of colored radial orderings of  $S$ . First we show a combinatorial lemma on cyclic permutations.

**Lemma 1.** *Let  $\sigma$  be a cyclic permutation of symbols in the set  $\{r, b\}$ , in which an  $r$ -symbol appears  $n - 1$  times and a  $b$ -symbol appears  $n$  times. Then the  $n$  cyclic permutations that arise from inserting in  $\sigma$  an  $r$ -symbol after a  $b$ -symbol are all different.*

*Proof.* We refer to the  $r$ -symbols as being “red” and to the  $b$ -symbols as being “blue”. Assume that  $\sigma$  is written starting at some blue symbol. For  $i = 0, \dots, 2n - 2$ , denote with  $\sigma(i)$  the  $(i + 1)$ -th symbol of  $\sigma$ . Let  $\sigma_k$  be the cyclic permutation that arises from inserting a red symbol  $p$  after the  $k$ -th blue symbol of  $\sigma$  and let  $\sigma_l$  be the the cyclic permutation that arises from inserting  $p$  after the  $l$ -th blue symbol of  $\sigma$ . Assume that  $\sigma_k$  and  $\sigma_l$  are written starting at the same blue symbol as  $\sigma$ . We claim that if  $k \neq l$ , then  $\sigma_k$  and  $\sigma_l$  are different cyclic permutations.

Suppose to the contrary that they are equal. Then there exists a natural number  $j > 0$  such that, for all  $i = 0, \dots, 2n - 1$ , the color of the symbol  $\sigma_k(i)$  equals the color of the symbol  $\sigma_l(i + j)$  (where addition is taken modulo  $2n$ ).

We now define a directed graph that captures the relationship between  $\sigma_k$  and  $\sigma_l$ ; we employ the structure of this graph to conclude that  $k$  must equal  $l$ . Let  $G$  be the directed graph whose vertices are the symbols of  $\sigma$ , together with  $p$ . In which, for all  $i = 0, \dots, 2n - 1$ , there is an arc from  $\sigma_k(i)$  to  $\sigma_l(i + j)$ . Note that every vertex in  $G$  has indegree and outdegree equal to 1. Therefore  $G$  is the union of disjoint directed cycles of symbols of the same color. Let  $\Gamma$  be the cycle containing  $p$  and let  $V' := V(G) \setminus V(\Gamma)$ . Let  $\sigma'_k$  be the cyclic permutation that arises from deleting from  $\sigma_k$  the symbols in  $\Gamma$ . Let  $\sigma'_l$  be the cyclic permutation that arises from deleting from  $\sigma_l$  the symbols in  $\Gamma$ . Assume that  $\sigma'_k$  is written starting at  $\sigma_k(0)$ , whereas  $\sigma'_l$  is written starting at  $\sigma_l(2n - j)$ ; note that being blue, these symbols are not in  $\Gamma$ .

Let  $G'$  be the directed graph whose vertex set is  $V'$  and in which there is an arc from  $\sigma'_k(i)$  to  $\sigma'_l(i)$ . As before every edge in  $G'$  has indegree and outdegree equal to 1. Since  $V'$  comes from removing a cycle in  $G$ ,  $\sigma'_k(i)$  is of the same color as  $\sigma'_l(i)$ . Therefore  $G'$  is the union of disjoint directed cycles of vertices of the same color (in fact  $G'$  is the subgraph of  $G$  induced by  $V'$ ). Since  $p$  is not in  $\sigma'_k$  nor in  $\sigma'_l$ , there exists a natural number  $j' > 0$  such that, for all  $i = 0, \dots, |V'| - 1$ , the symbol  $\sigma'_k(i)$  comes from the same symbol in  $\sigma_k$  as  $\sigma'_l(i + j')$ . Thus  $\sigma'_l$  is just a “shift” of  $j'$  places to the right of  $\sigma'_k$ . Thus the cycles in  $G'$  have the same length  $m$ . Therefore, both the number of red and blue symbols in  $V'$  are multiples of  $m$ . This implies that the number of vertices in  $\Gamma$  is also a multiple of  $m$ .

Let  $r \cdot m$  be the length of  $\Gamma$ , since  $\Gamma$  is not empty, then  $r \geq 1$ . Assume that  $\Gamma$ , starting from  $p$  is given by  $(p = v_1, v_2, \dots, v_m, \dots, v_{2m}, \dots, v_{rm})$ . Let  $b_k$  be the  $k$ -th blue symbol in  $\sigma$  and  $\Gamma' := (b_k = u_1, \dots, u_m)$  be the cycle in  $G$  containing  $b_k$ . Consider the following sequence of pairs of vertices  $(u_1, v_1), (u_2, v_2), \dots, (u_m, v_m)$ . Note that in  $\sigma_k$ , the symbol  $v_1 = p$  is just after the symbol  $u_1 = b_k$  and there after, for  $2 \leq i \leq m$ , the symbol  $v_i$  is just after the symbol  $u_i$  in both  $\sigma_l$  and  $\sigma_k$ . This is because if  $\sigma_k(i'_i) = u_i$  for some  $i'$  and  $\sigma_k(i'_i + 1) = v_i$ , then  $\sigma_l(i'_i + j) = u_{i+1}$  and  $\sigma_l(i'_i + 1 + j) = v_{i+1}$ . Suppose that  $r > 1$ , then the symbol  $v_{m+1}$  is just after  $u_1$  in  $\sigma_l$  while in  $\sigma_k$  it is just after the symbol  $v_1$  (which is equal to  $p$ ). Consider now the following sequence of vertices  $(v_1, v_{m+1}), (v_2, v_{m+2}), \dots, (v_{(r-1)m+1}, v_{rm+1} = p)$ . For the same arguments as before for  $m+2 \leq i \leq rm$  the symbol  $v_i$  is just after the symbol  $v_{i-m}$  in both  $\sigma_k$  and  $\sigma_l$ . For  $i = rm+1$ , the symbol  $v_{rm+1}$  is just after the symbol  $v_{(r-1)m+1}$  in  $\sigma_l$ , but  $v_{rm+1} = p$  and  $v_{(r-1)m+1}$  is red, a contradiction since  $p$  is just after a blue symbol in  $\sigma_l$ . Thus  $r = 1$ . This implies that  $\sigma_k = \sigma_l$ , since  $v_1$  (which is equal to  $p$ ) is after  $u_1$  (which is equal to  $b_k$ ) in both  $\sigma_l$  and  $\sigma_k$ .  $\square$

**Theorem 5.**  $\tilde{g}(n) \geq n$ .

*Proof.* To obtain the claimed lower bound, we show a walk in which  $n$  distinct colored radial orderings are seen. First choose a red point  $p$  of  $S$  and let  $\mathcal{C}$  be a circle centered at  $p$ . Afterwards walk once clockwise around  $\mathcal{C}$ . Choose  $\mathcal{C}$  to be small enough so that the only half-lines crossed in the walk are those involving  $p$ . Consider the non-colored radial orderings seen in this walk. Note that since we do not cross any half-line defined by points of  $S \setminus \{p\}$ , the points of  $S \setminus \{p\}$  remain fixed in these non-colored radial orderings. The only point that changes position is  $p$ ; it moves counter clockwise, transposing an element of  $S \setminus \{p\}$  every time a half-line is crossed. Thus by Lemma 1 every time  $p$  transposes a blue point we obtain a different colored radial ordering.  $\square$

### 3 Conclusions

We proved an upper bound of  $O(n^4)$  and a lower bound of  $\Omega(n^2)$  on the number of radial orderings that every set of  $2n$  points in the plane must have. The upper bound was first given in [5]. As a corollary in the same paper it was noted that every set of  $n$  points in the plane contains  $O(n^4)$  different star-shaped

Polygonizations. In our  $\Omega(n^2)$  bound we did not use the assumption on the parity of  $n$  nor the fact that  $S$  has  $2n$  points. Thus our result implies that every set of  $n$  points in general position in the plane has  $\Omega(n^2)$  different star-shaped hamiltonian non-crossing polygonal paths. Where in this context star-shaped means that the whole path is visible from a point.

We leave the closing of the gap on the number of radial orderings that every set of  $2n$  points in the plane must have as an open problem.

*Conjecture 1.*  $g(n) = \Theta(n^4)$ .

For colored point sets the situation is far more intriguing, here we have been able to prove that such a gap exists. Mainly that there are bi-colored sets of  $2n$  points with  $\Theta(n^4)$  colored radial orderings and sets with only  $\Theta(n^2)$ . The best lower bound we have been able to provide is of  $\Omega(n)$ . We make the following conjecture.

*Conjecture 2.*  $\tilde{g}(n) = \Theta(n^2)$ .

Finally we remark that we used the assumption that  $n$  is even heavily in the proof of Theorem 4. In fact it can be shown that the number of colored radial orderings may increase to  $\Theta(n^3)$  if a red and a blue point are added to the construction in the proof of Theorem 4. Also in the proof of Theorem 5 we relied on the fact that the number of red points equals the number of blue points. It is possible to construct a set of  $n$  red and  $n - 1$  blue points such that a walk like the one described in Theorem 5 yields only one colored radial ordering. It is possible that the bounds given in Theorems 4 and 5 no longer hold when either of these two hypothesis is dropped.

**Acknowledgments.** The authors would like to thank Clemens Huemer, Merce Claverol and David Wood for helpful comments.

## References

1. Bárány, I., Matoušek, J.: Simultaneous partitions of measures by  $k$ -fans. *Discrete & Comput. Geom.* 25, 317–334 (2001)
2. Bárány, I., Matoušek, J.: Equipartition of two measures by a 4-fan. *Discrete & Comput. Geom.* 27, 293–301 (2002)
3. Bereg, S.: Equipartitions of measures by 2-fans. *Discrete & Comput. Geom.* 34, 87–96 (2005)
4. Bespamyatnikh, S., Kirkpatrick, D., Snoeyink, J.: Generalizing ham-sandwich cuts to equitable subdivisions. *Discrete & Comput. Geom.* 24, 605–622 (2000)
5. Deneen, L., Shute, B.: Polygonizations of point sets in the plane. *Discrete Comput. Geom.* 3(1), 77–87 (1988)
6. Gfeller, B., Mihalák, M., Suri, S., Vicari, E., Widmayer, P.: Counting Targets with Mobile Sensors in an Unknown Environment. In: Kutylowski, M., Cichoń, J., Kubiak, P. (eds.) ALGOSENSORS 2007. LNCS, vol. 4837, pp. 32–45. Springer, Heidelberg (2008)

7. Kannianen, O., Allho, T.M.R.: Minimalist navigation for a mobile robot based on a simple visibility sensor information, pp. 68–75 (2008)
8. Langerman, S., Steiger, W.: Optimization in Arrangements. In: Alt, H., Habib, M. (eds.) STACS 2003. LNCS, vol. 2607, pp. 50–61. Springer, Heidelberg (2003)
9. Suri, S., Vicari, E., Widmayer, P.: Simple robots with minimal sensing: From local visibility to global geometry. *Int. J. Rob. Res.* 27, 1055–1067 (2008)
10. Tovar, B., Freda, L., LaValle, S.M.: Using a robot to learn geometric information from permutations of landmarks. In: Topology and Robotics. Contemp. Math., vol. 438, pp. 33–45. Amer. Math. Soc., Providence (2007)

# Notes on the Twisted Graph\*

Elsa Omaña-Pulido and Eduardo Rivera-Campo

Departamento de Matemáticas,  
Universidad Autónoma Metropolitana - Iztapalapa,  
Av. San Rafael Atlixco 186, México D.F. 09340  
`{eop,erc}@xanum.uam.mx`

**Abstract.** The twisted graph  $T_n$  is a complete topological graph with  $n$  vertices  $v_1, v_2, \dots, v_n$  in which two edges  $v_i v_j$  ( $i < j$ ) and  $v_s v_t$  ( $s < t$ ) cross each other if and only if  $i < s < t < j$  or  $s < i < j < t$ . We study several properties concerning plane topological subgraphs of  $T_n$ .

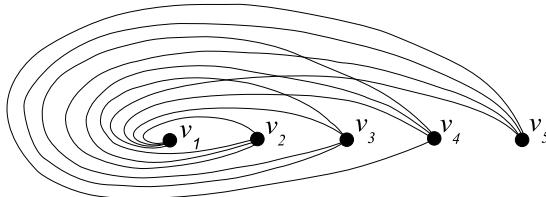
**Keywords:** Alternating Path, Tree Graph, Max Graph, Matching Graph.

## 1 Introduction

Let  $P$  be a set of points in the plane. A *topological graph* with vertex set  $P$  is a simple graph drawn in the plane with edges as Jordan curves in such a way that any two edges have at most one point in common. A *geometric graph* is a topological graph in which all edges are straight line segments.

Two topological graphs  $G$  and  $G'$  with vertex sets  $P$  and  $P'$ , respectively, are weakly isomorphic if there is a bijection  $\alpha : P \rightarrow P'$  such that  $\alpha(u)\alpha(v) \in E(G')$  if and only if  $uv \in E(G)$  and two edges  $\alpha(x)\alpha(y)$  and  $\alpha(u)\alpha(v)$  of  $G'$  intersect if and only if  $xy$  and  $uv$  intersect in  $G$ .

The *twisted graph*  $T_n$  is a complete topological graph with  $n$  vertices  $v_1, v_2, \dots, v_n$  in which two edges  $v_i v_j$  ( $i < j$ ) and  $v_s v_t$  ( $s < t$ ) cross each other if and only if  $i < s < t < j$  or  $s < i < j < t$ . See Fig. 1



**Fig. 1.** A twisted graph  $T_5$

The graphs  $T_n$ ,  $n \geq 5$  were introduced by Harborth and Mengersen [4] as examples of complete topological graphs containing no topological subgraphs

---

\* Partially supported by Conacyt, México, project 83856.

weakly isomorphic to the complete convex geometric graph  $C_5$  on 5 vertices. Later, Pach *et al.* [9] proved that every complete topological graph with  $n$  vertices contains a complete topological subgraph with  $m \geq c \log^{1/8} n$  vertices which is weakly isomorphic to either the complete convex graph  $C_m$  or to the twisted graph  $T_m$ . In this paper we study problems for the twisted graph  $T_m$  for which the corresponding problems for the complete geometric convex graph  $C_m$  have been studied by several authors, see [1–3], [5], [7], [8] and [10]. Due to the above result by Pach *et al.*, in most cases, a result for general topological graphs follows.

A colouring of a graph  $G$  is an assignment of colours to the vertices of  $G$ . Here, two adjacent vertices may have the same colour. A colouring of a graph  $G$  with two colours is *balanced* if the colour classes have the same size. Akiyama and Urrutia [1] gave an algorithm to decide whether there is a plane alternating spanning path for a given balanced colouring of the complete geometric graph  $C_n$  with  $n$  vertices in convex position. In Section 2 we show that for each balanced colouring of  $T_n$ , there is a plane spanning path of  $T_n$  whose vertices alternate colours. The best known lower bound for the largest plane alternating path for any balanced colouring of  $C_n$  is  $\frac{n}{2} + c\sqrt{n}$  [3].

Not all connected topological graphs contain plane spanning trees. Rivera-Campo [10] proved that if  $G$  is a geometric graph with  $n \geq 3$  vertices in convex position such that  $G - v$  has a plane spanning tree for each vertex  $v$  of  $G$ , then  $G$  also has a plane spanning tree. Also in Section 2, we prove the same result for topological subgraphs of the twisted graph  $T_n$ .

The *tree graph*  $G(T)$  of a topological graph  $T$  is the abstract graph whose vertices are the plane spanning trees of  $T$  in which two trees  $Q$  and  $R$  are adjacent if there are edges  $q$  and  $r$  of  $T$  such that  $R = (Q - q) + r$ . Avis and Fukuda [2] proved that the graph  $G(T)$  is connected whenever  $T$  is a complete geometric graph. In Section 3 we prove that  $G(T_n)$  is always connected.

Let  $F_n$  be a complete topological graph with  $n$  vertices. A plane topological subgraph  $G$  of  $F_n$  is a *maximal plane subgraph* of  $F_n$  if for each edge  $uv$  of  $F_n$ , not in  $G$ , there is an edge  $xy$  of  $G$  that intersects  $uv$ . The *max graph*  $MP(F_n)$  of  $F_n$  is the abstract graph whose vertices are the maximal plane topological subgraphs of  $F_n$  in which two graphs  $G$  and  $H$  are adjacent if there are edges  $g$  of  $G$  and  $h$  of  $H$  such that  $H = (G - g) + h$ . In the case where  $F_n$  is a complete geometric graph with vertex set  $P$ , the graph  $MP(F_n)$  is called the *graph of triangulations* of  $P$  which is well known to be connected [8]. In Section 4 we prove that  $MP(T_n)$  is also connected.

Hernando *et al.* [5] proved that given any two plane perfect matchings  $N$  and  $M$  of the complete convex geometric graph  $C_{2m}$ , there is a sequence  $N = N_0, N_1, \dots, N_t = M$  of plane perfect matchings of  $C_{2m}$  such that for  $i = 0, 1, \dots, t-1$ , the symmetric difference  $N_i \Delta N_{i+1}$  is a plane cycle with four edges. In Section 5 we prove the corresponding result for the twisted graph  $T_{2m}$ .

Throughout the paper we denote by  $v_1, v_2, \dots, v_n$  the vertices of  $T_n$  from left to right as in Fig. 1.

## 2 Alternating Paths and Plane Spanning Trees

Let  $c$  be a colouring of a topological graph  $G$  with two colours. An *alternating path* of  $G$  is a path in  $G$  whose vertices alternate colours.

**Theorem 1.** *For every balanced colouring of  $T_n$  with two colours, there is a plane alternating spanning path of  $T_n$ .*

*Proof.* Let  $c$  be a balanced colouring of  $T_n$ . Start a path at vertex  $v_{i_1} = v_1$  and for  $k = 1, 2, \dots, n - 1$  let the next vertex  $v_{i_{k+1}}$  be such that

$$i_{k+1} = \min \{j : v_j \notin \{v_{i_1}, v_{i_2}, \dots, v_{i_k}\} \text{ and } c(v_j) \neq c(v_{i_k})\}$$

Since  $c(v_{i_1}) \neq c(v_{i_2})$ , the path  $v_{i_1}, v_{i_2}$  is a plane alternating path  $R_1$  of  $T_n$  with length one. Let  $k > 1$  and assume the path  $R_{k-1} = v_{i_1}, v_{i_2}, \dots, v_{i_k}$  is a plane alternating path of  $T_n$ .

If the edge  $v_{i_k}v_{i_{k+1}}$  intersects an edge  $v_{i_j}v_{i_{j+1}}$  of  $R_{k-1}$ , then either  $i_k < i_j, i_{j+1} < i_{k+1}$  or  $i_{k+1} < i_j, i_{j+1} < i_k$  or  $i_j < i_k, i_{k+1} < i_{j+1}$  or  $i_{j+1} < i_k, i_{k+1} < i_j$ . Because the vertices in  $R_k$  alternate colours, then:  $c(v_{i_j}) = c(v_{i_k})$  and  $c(v_{i_{j+1}}) = c(v_{i_{k+1}})$ , or  $c(v_{i_{j+1}}) = c(v_{i_k})$  and  $c(v_{i_j}) = c(v_{i_{k+1}})$ .

In the case where  $i_k < i_j, i_{j+1} < i_{k+1}$  and  $c(v_{i_j}) = c(v_{i_k})$ , we reach a contradiction since  $j < k$  and  $i_k < i_j$ .

Analogous contradictions can be obtained in the remaining cases. Therefore  $R_k = v_{i_1}, v_{i_2}, \dots, v_{i_k}, v_{i_{k+1}}$  is a plane alternating path of  $T_n$  with length  $k$ . By induction,  $T_n$  contains a plane spanning alternating path  $R_{n-1}$ .  $\square$

The following result provides a sufficient condition for a topological subgraph  $G$  of  $T_n$  to contain a plane spanning tree.

**Theorem 2.** *Let  $n \geq 3$  and let  $G$  be a topological spanning subgraph of  $T_n$ . If  $G - v$  contains a plane spanning tree for each  $v \in V(G)$ , then  $G$  also contains a plane spanning tree.*

*Proof.* Let  $t = \min \{j : v_1v_j \in E(G)\}$  and let  $F_t$  be a plane spanning tree of  $G - v_t$ . Since  $F_t$  is connected, there is at least one edge  $v_1v_s$  of  $F_t$  incident with  $v_1$ . As no edge of  $F_t$  intersects the edge  $v_1v_s$ , then  $v_iv_j \notin E(F_t)$  for  $1 < i < j < s$ . By the choice of  $t$ ,  $s > t$  and therefore  $v_1v_t$  does not intersect any edge of  $F_t$ . This implies that  $F_t$  together with the edge  $v_1v_t$  is a plane spanning tree of  $G$ .  $\square$

## 3 Tree Graph

Let  $S_n$  be the plane spanning tree of  $T_n$  with edges  $v_1v_2, v_1v_3, \dots, v_1v_n$ . That is,  $S_n$  is the spanning star of  $T_n$  with center vertex  $v_1$ .

**Theorem 3.** *For each plane spanning tree  $R$  of  $T_n$  there is a path  $R_1, R_2, \dots, R_t$  in  $G(T_n)$  with  $R_1 = S_n$  and  $R_t = R$ .*

*Proof.* For each plane spanning tree  $R$  of  $T_n$  let  $k(R)$  be the largest integer  $j$  such that  $v_1v_2, v_1v_3, \dots, v_1v_j$  are edges of  $R$ . If  $n - k(R) = 0$ , then  $S_n = R$ . Assume  $n - k(R) = m + 1$  and that the result holds for each plane spanning tree  $R'$  of  $T_n$  for which  $n - k(R') = m$ .

Since  $R$  is a plane subgraph of  $T_n$  and  $v_1v_{k(R)}$  is an edge of  $R$ , for  $2 \leq i < j \leq k(R) - 1$  the edge  $v_iv_j$  of  $T_n$  cannot be an edge of  $R$ . This implies that  $v_1v_{k(R)+1}$  does not intersect any edge of  $R$  and therefore  $R + v_1v_{k(R)+1}$  is a plane subgraph of  $T_n$ . Clearly  $R + v_1v_{k(R)+1}$  contains an edge  $uv_{k(R)+1}$  with  $u \neq v_1$  such that  $R' = (R + v_1v_{k(R)+1}) - uv_{k(R)+1}$  is a tree. Then  $R'$  is a plane spanning tree of  $T_n$  with  $k(R') = k(R) + 1$  and  $n - k(R') = m$ . By induction there is a path  $R_1, R_2, \dots, R_t$  in  $G(T_n)$  with  $S_n = R_1$  and  $R_t = R'$ . Since  $R'$  and  $R$  are adjacent in  $G(T_n)$ ,  $R_1, R_2, \dots, R_t, R$  is also a path in  $G(T_n)$ .  $\square$

As an immediate consequence we obtain the following:

**Corollary 1.** *For each positive integer  $n$ , the tree graph  $G(T_n)$  is connected.*

## 4 Max Graph

A *vertex ordering* of a graph  $G$  is a numbering of the vertices  $v_1, v_2, \dots, v_n$  of  $G$ . Let  $G$  be a graph with a vertex ordering  $v_1, v_2, \dots, v_n$ . Two edges  $v_iv_j$  ( $i < j$ ) and  $v_sv_t$  ( $s < t$ ) of  $G$  are *nested* if  $i < s < t < j$  or  $s < i < j < t$ . A set  $X$  of edges of  $G$  is a *queue* if no two edges in  $X$  are nested. The sets of edges of plane topological subgraphs of  $T_n$  are queues of the complete graph  $K_n$  with the corresponding vertex ordering. Klazar [6] proved that if  $n \geq 2$ , then every maximal queue of  $K_n$  has exactly  $2n - 3$  edges and that the number of maximal queues of  $K_n$  is the Catalan number  $\frac{1}{n-1} \binom{2n-4}{n-2}$ . This implies that every maximal plane topological subgraph of  $T_n$  has exactly  $2n - 3$  edges and that the number of maximal plane topological subgraphs of  $T_n$  is  $\frac{1}{n-1} \binom{2n-4}{n-2}$  which is also the number of triangulations of any set  $P_n$  of  $n$  points in convex position.

The following lemmas will be used in the proof of Theorem 4.

**Lemma 1.** *Let  $n \geq 3$ , the following properties are satisfied by any maximal plane topological subgraph  $F$  of  $T_n$ .*

1. *If  $v_1v_k$  is an edge of  $F$ , then  $v_1v_2, v_1v_3, \dots, v_1v_{k-1}$  are also edges of  $F$ .*
2. *If  $v_tv_n$  is an edge of  $F$ , then  $v_{t+1}v_n, v_{t+2}v_n, \dots, v_{n-1}v_n$  are also edges of  $F$ .*
3.  *$v_1v_2, v_1v_3, v_{n-2}v_n$  and  $v_{n-1}v_n$  are edges of  $F$ .*

*Proof.* To prove (1), let  $k$  be such that  $v_1v_k$  is an edge of  $F$  and assume  $v_1v_t$  is not an edge of  $F$  for some  $t < k$ . Since  $F$  is a maximal plane topological subgraph of  $T_n$ , the edge  $v_1v_t$  must intersect an edge  $v_iv_j$  of  $F$  which is only possible if  $1 < i, j < t$ . In this case  $v_1v_k$  must also intersect the edge  $v_iv_j$ , a contradiction. In an analogous way we can obtain a proof of (2). To show (3) just notice that the edges  $v_1v_2, v_1v_3, v_{n-2}v_n$  and  $v_{n-1}v_n$  do not intersect any edge of  $T_n$ .  $\square$

For any maximal plane topological subgraph  $F$  of  $T_n$ , let  $k(F)$  be the largest integer  $k$  such that  $v_1v_k$  is an edge of  $F$ . By Lemma 1 (3),  $k(F) \geq 3$ . Let  $MP_3(T_n)$  be the subgraph of  $MP(T_n)$  induced by the set of maximal plane topological subgraphs  $F$  of  $T_n$  for which  $k(F) = 3$ .

**Lemma 2.** *For  $n \geq 2$ ,  $MP_3(T_{n+1})$  and  $MP(T_n)$  are isomorphic graphs.*

*Proof.* For a graph  $G$  in  $MP_3(T_{n+1})$  let  $\phi(G)$  be the graph obtained from  $G$  by deleting the vertex  $v_1$ . Clearly  $\phi(G)$  is a maximal plane topological subgraph of  $T_{n+1} - v_1$  and that if  $\phi(G_1) = \phi(G_2)$  for  $G_1$  and  $G_2$  in  $MP_3(T_{n+1})$ , then  $G_1 = G_2$ . Let  $F$  be a graph in  $MP(T_{n+1} - v_1)$  and let  $G$  be the graph obtained from  $F$  by adding the vertex  $v_1$  together with the edges  $v_1v_2$  and  $v_1v_3$ . Since the edges  $v_1v_2$  and  $v_1v_3$  do not intersect any edges of  $T_{n+1}$ , the graph  $G$  is a plane topological subgraph of  $T_{n+1}$  and by the way  $\phi$  is defined,  $\phi(G) = F$ .

If  $G_1$  and  $G_2$  are graphs adjacent in  $MP_3(T_{n+1})$ , then there are edges  $v_iv_j$  and  $v_kv_l$  of  $T_{n+1}$  such that  $G_2 = (G_1 - v_iv_j) + v_kv_l$ . Since  $G_1$  and  $G_2$  are graphs in  $MP_3(T_{n+1})$ ,  $v_1v_2 \neq v_iv_j$ ,  $v_iv_j \neq v_kv_l$ ,  $v_1v_2 \neq v_kv_l$  and  $v_kv_l \neq v_1v_3$ . Therefore  $\phi(G_2) = (\phi(G_1) - v_iv_j) + v_kv_l$ . This implies that  $\phi(G_1)$  and  $\phi(G_2)$  are adjacent in  $MP(T_{n+1} - v_1)$ . Analogously, if  $G_1$  and  $G_2$  are such that  $\phi(G_1)$  and  $\phi(G_2)$  are adjacent in  $MP(T_{n+1} - v_1)$ , then  $G_1$  and  $G_2$  are adjacent in  $MP_3(T_{n+1})$ . Therefore  $MP_3(T_{n+1})$  is isomorphic to  $MP(T_{n+1} - v_1)$ . The lemma follows since  $T_{n+1} - v_1$  and  $T_n$  are weakly isomorphic.  $\square$

**Theorem 4.** *For each positive integer  $n$ , the graph  $MP(T_n)$  is connected.*

*Proof.* For the base of induction notice that the graphs  $MP(T_1)$ ,  $MP(T_2)$  and  $MP(T_3)$  contain exactly one vertex and therefore are connected. We proceed assuming  $n \geq 3$  and that  $MP(T_n)$  is a connected graph. We claim that for each maximal plane topological subgraph  $F$  of  $T_{n+1}$  with  $k(F) = k$ , there is a path  $F_k, F_{k-1}, \dots, F_3$  in  $MP(T_{n+1})$  with  $F = F_k$  and  $F_3$  in  $MP_3(T_{n+1})$ . By the induction hypothesis and by Lemma 2, the graph  $MP_3(T_{n+1})$  is connected. Therefore  $MP(T_{n+1})$  is also connected and the theorem follows by induction.

*Proof of claim.* If  $k(F) = 3$ , then  $F$  lies in  $MP_3(T_{n+1})$ . Assume the result follows for any graph  $F'$  in  $MP(T_{n+1})$  for which  $k(F') = t$  and let  $F$  be a maximal plane topological subgraph of  $T_{n+1}$  such that  $k(F) = t + 1$ . By Lemma 1 (1)  $v_1v_2, v_1v_3, \dots, v_1v_{t+1}$  are edges of  $F$ . Since  $F$  is a plane topological subgraph of  $T_{n+1}$ ,  $v_2v_t$  is not an edge of  $F$ ; let  $F' = (F - v_1v_{t+1}) + v_2v_t$ . If  $v_2v_t$  intersects an edge  $v_iv_j$  of  $F'$  with  $i < j$ , then either  $i < 2 < t < j$  or  $2 < i < j < t$ . In the former case  $i = 1$  and  $j \geq t + 1$  which is not possible since  $k(F) = t + 1$  and  $v_1v_{t+1}$  is not an edge of  $F'$ . In the latter case  $v_iv_j$  is an edge of the plane graph  $F$  which intersects the edge  $v_1v_{t+1}$  also in  $F$ , a contradiction. Therefore  $F'$  is a maximal topological subgraph of  $T_{n+1}$  with  $k(F') = t$ . By induction there is a path  $F_t, F_{t-1}, \dots, F_3$  in  $MP(T_{n+1})$  with  $F_3$  in  $MP_3(T_{n+1})$  and such that  $F' = F_t$ . The claim follows by induction since  $F$  and  $F'$  are adjacent in  $MP(T_{n+1})$ .  $\square$

## 5 Matching Graph

The *matching graph*  $\mathcal{M}(T_{2m})$  of  $T_{2m}$  is the abstract graph with vertex set given by the set of plane perfect matchings of  $T_{2m}$  where two matchings  $L$  and  $N$  are adjacent if the symmetric difference  $L \Delta N$  is a plane cycle with four edges. We prove that  $\mathcal{M}(T_{2m})$  is always connected.

For any plane perfect matching  $N$  of  $T_{2m}$ , let  $k = k(N)$  be such that  $v_1v_k \in N$  and let  $\mathcal{M}_2(T_{2m})$  be the subgraph of  $\mathcal{M}(T_{2m})$  induced by the matchings  $N$  for which  $k(N) = 2$ .

**Lemma 3.** *For  $m \geq 1$ ,  $\mathcal{M}_2(T_{2m+2})$  and  $\mathcal{M}(T_{2m})$  are isomorphic graphs.*

*Proof.* For a matching  $N$  in  $\mathcal{M}_2(T_{2m+2})$  let  $\psi(N) = N \setminus \{v_1v_2\}$ . Clearly  $\psi(N)$  is a plane perfect matching of  $T_{2m+2} - \{v_1, v_2\}$  and that if  $\psi(N_1) = \psi(N_2)$  for  $N_1$  and  $N_2$  in  $\mathcal{M}_2(T_{2m+2})$ , then  $N_1 = N_2$ . Let  $M$  be a plane perfect matching of  $T_{2m+2} - \{v_1, v_2\}$  and let  $N = M \cup \{v_1v_2\}$ . Since the edge  $v_1v_2$  intersects no edges of  $T_{2m+2}$ ,  $N$  is a plane perfect matching of  $T_{2m+2}$  and by the way  $\psi$  is defined,  $\psi(N) = M$ .

If  $\psi(N_1)$  and  $\psi(N_2)$  are matchings which are adjacent in  $\mathcal{M}_2(T_{2m+2})$ , then there are edges  $v_{i_1}v_{j_1}$ ,  $v_{i_2}v_{j_2}$  in  $N_1$  and  $v_{k_1}v_{l_1}$ ,  $v_{k_2}v_{l_2}$  in  $N_2$  such that  $N_1 = (N_2 - \{v_{i_1}v_{j_1}, v_{i_2}v_{j_2}\}) + \{v_{k_1}v_{l_1}, v_{k_2}v_{l_2}\}$ . Since  $v_1v_2 \in N_1 \cap N_2$ , it cannot happen that  $v_1v_2$  is one of the edges  $v_{i_1}v_{j_1}$ ,  $v_{i_2}v_{j_2}$ ,  $v_{k_1}v_{l_1}$  or  $v_{k_2}v_{l_2}$ . This implies that  $\psi(N_1)$  and  $\psi(N_2)$  are adjacent in the graph  $\mathcal{M}(T_{2m+2} - \{v_1, v_2\})$ . Analogously if  $N_1$  and  $N_2$  are matchings in  $\mathcal{M}_2(T_{2m+2})$  such that  $\psi(N_1)$  and  $\psi(N_2)$  are adjacent in  $\mathcal{M}(T_{2m+2} - \{v_1, v_2\})$ , then  $N_1$  and  $N_2$  are adjacent in  $\mathcal{M}_2(T_{2m+2})$ . Therefore  $\mathcal{M}_2(T_{2m+2})$  and  $\mathcal{M}(T_{2m+2} - \{v_1, v_2\})$  are isomorphic graphs. The Lemma follows since  $T_{2m}$  and  $T_{2m+2} - \{v_1, v_2\}$  are weakly isomorphic.  $\square$

**Theorem 5.** *For each positive integer  $m$ , the graph  $\mathcal{M}(T_{2m})$  is connected.*

*Proof.* Graph  $\mathcal{M}(T_2)$  is a graph with exactly one vertex and therefore connected. Assume  $m \geq 1$  and that  $\mathcal{M}(T_{2m})$  is a connected graph. We claim that for each plane matching  $N$  of  $T_{2m+2}$  with  $k(N) = k$ , there is a path  $N_k, N_{k-1}, \dots, N_2$  in  $\mathcal{M}(T_{2m+2})$  with  $N = N_k$  and  $N_2$  in  $\mathcal{M}_2(T_{2m+2})$ . By the induction hypothesis and by Lemma 3, the graph  $\mathcal{M}_2(T_{2m+2})$  is connected. Therefore  $\mathcal{M}(T_{2m+2})$  is also connected which completes the induction.

*Proof of claim.* If  $k(N) = 2$ , then  $N$  lies in  $\mathcal{M}_2(T_{2m+2})$ . Assume the result follows for any matching  $N'$  in  $\mathcal{M}(T_{2m+2})$  for which  $k(N') = t$  and let  $N$  be a plane matching of  $T_{2m+2}$  such that  $k(N) = t+1$ . Let  $s$  be such that  $v_tv_s \in N$ . Since  $N$  is a plane matching and  $v_1v_{t+1}$  is also an edge in  $N$ , it must be the case that  $t+1 < s$ . This implies that the edges  $v_1v_t$  and  $v_{t+1}v_s$  of  $T_{2m+2}$  intersect no edges in  $N$ . Let  $N'$  be the matching obtained from  $N$  by deleting the edges  $v_1v_{t+1}$  and  $v_tv_s$  and adding the edges  $v_1v_t$  and  $v_{t+1}v_s$ . Since  $v_1v_t$  and  $v_{t+1}v_s$  of  $T_{2m+2}$  intersect no edges in  $N$ ,  $N'$  is a plane perfect matching of  $T_{2m+2}$  with  $k(N') = t$  and by induction there is a path  $N_t, N_{t-1}, \dots, N_2$  in  $\mathcal{M}(T_{2m+2})$  with  $N_2$  in  $\mathcal{M}_2(T_{2m+2})$  and such that  $N' = N_t$ . Since  $N$  and  $N'$  are adjacent in  $\mathcal{M}(T_{2m+2})$ , the claim follows by induction.  $\square$

## References

1. Akiyama, J., Urrutia, J.: Simple alternating path problem. *Discrete Math.* 84(1), 101–103 (1990)
2. Avis, D., Fukuda, K.: Reverse search enumeration. *Discrete Applied Math.* 6, 21–46 (1996)
3. Hajnal, P., Mészáros, V.: Note on noncrossing alternating path in colored convex sets. Accepted in *Discrete Mathematics and Theoretical Computer Science*
4. Harborth, H., Mengersen, I.: Drawings of the complete graph with maximum number of crossings. In: *Proceedings of the Graph Theory and Computing*, Boca Raton, Fl. Congressus Numerantium, vol. 88, pp. 225–228. Utilitas Math., Winnipeg (1992)
5. Hernando, C., Hurtado, F., Noy, M.: Graphs of non-crossing perfect matchings. *Graphs and Combin.* 18, 517–532 (2002)
6. Klazar, M.: Counting pattern-free set partitions II. Noncrossing and other hypergraphs. *Electron. J. Combin.* 7, Research Paper 34, 25 pages (2000)
7. Kynčl, J., Pach, J., Tóth, G.: Long alternating paths in bicolored point sets. *Discrete Math.* 308(19), 4315–4321 (2008)
8. Lawson, C.L.: Software for  $C^1$ -interpolation. In: Rice, J. (ed.) *Mathematical Software III*, pp. 161–194. Academic Press, New York (1977)
9. Pach, J., Solymosi, J., Tóth, G.: Unavoidable configurations in complete topological graphs. *Discrete Comput. Geom.* 30, 311–320 (2003)
10. Rivera-Campo, E.: A Note on the Existence of Plane Spanning Trees of Geometric Graphs. In: Akiyama, J., Kano, M., Urabe, M. (eds.) *JCDCG 1998*. LNCS, vol. 1763, pp. 274–277. Springer, Heidelberg (2000)

# Locating a Service Facility and a Rapid Transit Line

José Miguel Díaz-Báñez<sup>1,\*</sup>, Matias Korman<sup>2,\*\*</sup>, Pablo Pérez-Lantero<sup>3,\*\*\*</sup>,  
and Inmaculada Ventura<sup>1,\*</sup>

<sup>1</sup> Departamento de Matemática Aplicada II, Universidad de Sevilla, Spain  
[{dbanez,iventura}@us.es](mailto:{dbanez,iventura}@us.es)

<sup>2</sup> Universitat Politècnica de Catalunya (UPC), Barcelona  
[matias.korman@upc.edu](mailto:matias.korman@upc.edu)

<sup>3</sup> Escuela de Ingeniería Civil en Informática, Universidad de Valparaíso, Chile  
[pablo.perez@uv.cl](mailto:pablo.perez@uv.cl)

**Abstract.** In this paper we study a facility location problem in the plane in which a single point (facility) and a rapid transit line (highway) are simultaneously located in order to minimize the total travel time of the clients to the facility, using the  $L_1$  or Manhattan metric. The rapid transit line is represented by a line segment with fixed length and arbitrary orientation. The highway is an alternative transportation system that can be used by the clients to reduce their travel time to the facility. This problem was introduced by Espejo and Rodríguez-Chía in [8]. They gave both a characterization of the optimal solutions and an algorithm running in  $O(n^3 \log n)$  time, where  $n$  represents the number of clients. In this paper we show that the Espejo and Rodríguez-Chía's algorithm does not always work correctly. At the same time, we provide a proper characterization of the solutions with a simpler proof and give an algorithm solving the problem in  $O(n^3)$  time.

**Keywords:** Geometric optimization, Facility location, Time distance.

## 1 Introduction

Suppose that we have a set of clients represented as a set of points in the plane, and a service facility represented as a point to which all clients have to move. Every client can reach the facility directly or by using an alternative rapid transit line or highway, represented by a straight line segment of fixed length and arbitrary orientation, in order to reduce the travel time. Whenever a client moves

---

\* J.M. D.-B., P. P.-L. and I. V. were partially supported by project FEDER MEC MTM2009-08652. J.M. D.-B., M.K and I. V. were partially supported by ESF EU-ROCORES programme EuroGIGA, CRP ComPoSe: grant EUI-EURC-2011-4306.

\*\* Partially supported by the support of the Secretary for Universities and Research of the Ministry of Economy and Knowledge of the Government of Catalonia and the European Union.

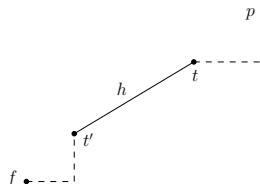
\*\*\* P. P.-L. was partially supported by grant FONDECYT 11110069.

directly to the facility, it moves at unit speed and the distance traveled is the Manhattan or  $L_1$  distance to the facility. In the case where a client uses the highway, it travels the  $L_1$  distance at unit speed to one endpoint of the highway, traverses the entire highway with a speed greater than one, and finally travels the  $L_1$  distance from the other endpoint to the facility at unit speed. All clients traverse the highway at the same speed. Given the set of points representing the clients, the facility location problem consists in determining at the same time the facility point and the highway in order to minimize the *total weighted travel time* from the clients to the facility. The weighted travel time of a client is its travel time multiplied by a weight representing the intensity of its demand. This problem was studied by Espejo and Chía [8]. We refer to [8] and references therein to review both the state of the art and applications of this problem.

Geometric problems related to transportation networks have been recently considered in computational geometry. Abellanas *et. al.* introduced the *time metric* model in [1]: Given an underlying metric, the user can travel at speed  $v$  when moving along a highway  $h$  or unit speed elsewhere. The particular case in which the underlying metric is the  $L_1$  metric and all highways are axis-parallel segments of the same speed, is called the *city metric* [3]. The optimal positioning of transportation systems that minimize the maximum travel time among a set of points has been investigated in detail in recent papers [2,6,4]. Other more general models are studied in [9]. The variant introduced by Espejo and Chía aims to minimize the sum of the travel times (transportation cost) from the demand points to the new facility service, which has to be located simultaneously with a highway. The highway is used by a demand point whenever it saves time to reach the facility.

Notation to formulate the problem is as follows. Let  $S$  be the set of  $n$  client points;  $f$  the service facility point;  $h$  the highway;  $\ell$  the length of  $h$ ;  $t$  and  $t'$  the endpoints of  $h$ ; and  $v > 1$  the speed in which the points move along  $h$ . Let  $w_p > 0$  be the weight (or demand) of a client point  $p$ . Given a point  $u$  of the plane, let  $x(u)$  and  $y(u)$  denote the  $x$  and  $y$  coordinates of  $u$ , respectively. The distance or travel time (see Figure 1), between a point  $p$  and the service facility  $f$  is given by the function

$$d_{t,t'}(p, f) := \min \left\{ \|p - f\|_1, \|p - t\|_1 + \frac{\ell}{v} + \|t' - f\|_1, \|p - t'\|_1 + \frac{\ell}{v} + \|t - f\|_1 \right\}$$



**Fig. 1.** The distance between a point  $p$  and the facility  $f$  using the highway

Then the problem can be formulated as follows:

**The Facility and Highway Location problem (FHL-problem):**

Given a set  $S$  of  $n$  points, a weight  $w_p > 0$  associated with each point  $p$  of  $S$ , a fixed highway length  $\ell > 0$ , and a fixed speed  $v > 1$ , locate a point (facility)  $f$  and a line segment (highway)  $h$  of length  $\ell$  with endpoints  $t$  and  $t'$  such that the function  $\sum_{p \in S} w_p \cdot d_{t,t'}(p, f)$  is minimized.

Espejo and Chía [8] studied the FHL-problem and gave the following characterization of the solutions. Consider the grid  $G$  defined by the set of all axis-parallel lines passing through the elements of  $S$ . They stated that there always exists an optimal highway having one endpoint at a vertex of  $G$ . Based on this, they proposed an  $O(n^3 \log n)$ -time algorithm to solve the problem. In this paper we show that the characterization given by Espejo and Chía is not true, hence their algorithm does not always give the optimal solution. Recently, there has been some claims on the originality and timing of our paper. Since this is a scientific publication, we decided to omit the topic. For more information on the controversy, the interested reader can access the on-line version of this paper [7].

In Section 2 we first provide a proper characterization of the solutions. Our proof uses geometric observations and is simpler than the proof given in [8]. After that we give a counterexample to the Espejo and Chía's characterization. We provide a set of five points, all having weight equal to one, and prove that no optimal highway has one endpoint in a vertex of  $G$ . In Section 3 we present an improved algorithm running in  $O(n^3)$  time that correctly solves the FHL-problem.

## 2 Properties of an Optimal Solution

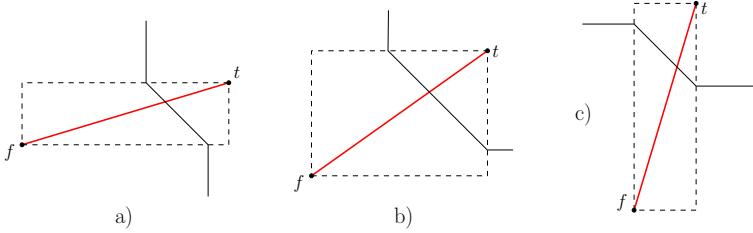
A primary observation (also stated in [8]) is that the service facility can be located at one of the endpoints of the rapid transit line. From now on, we assume throughout the paper that  $f = t'$ . This assumption simplifies the distance from a point  $p \in S$  to the facility to the following expression,

$$d_t(p, f) = \min \left\{ \|p - f\|_1, \|p - t\|_1 + \frac{\ell}{v} \right\}.$$

Using this observation, the expression of our objective function to minimize is  $\Phi(f, t) = \sum_{p \in S} w_p \cdot d_t(p, f)$ . We call this value the *total transportation cost* associated with  $f$  and  $t$  (or simply the *cost* of  $f$  and  $t$ ).

We say that a point  $p$  uses the highway if  $\|p - t\|_1 + \frac{\ell}{v} < \|p - f\|_1$ , and that  $p$  does not use it (or goes directly to the facility) otherwise. Given  $f$  and  $t$ , we say that the bisector of  $f$  and  $t$  is the set of points  $z$  such that  $\|z - f\|_1 = \|z - t\|_1 + \frac{\ell}{v}$ , see Figure 2. A geometrical description of such bisector can be found in [8].

**Lemma 1.** *There exists an optimal solution to the FHL-problem satisfying one of the next conditions:*



**Fig. 2.** The bisector of  $f$  and  $t$

- (a) One of the endpoints of the highway is a vertex of  $G$ .
- (b) One endpoint of the highway is on a horizontal line of  $G$ , and the other endpoint is on a vertical line of  $G$ .

*Proof.* Let  $f$  and  $t$  be the endpoints of an optimal highway  $h$  and assume that neither of the two conditions is satisfied. Using local perturbation we will transform this solution into one that satisfies one of the two conditions. Assume that neither  $f$  nor  $t$  is on a vertical line of  $G$ . Let  $\delta_1 > 0$  (resp.  $\delta_2 > 0$ ) be the smallest value such that if we translate  $h$  with vector  $(-\delta_1, 0)$  (resp.  $(\delta_2, 0)$ ) then either one endpoint of  $h$  touches a vertical line of  $G$  or a demand point hits the bisector of  $f$  and  $t$ . Given  $\varepsilon \in [-\delta_1, \delta_2]$ , let  $f_\varepsilon$ ,  $t_\varepsilon$ , and  $h_\varepsilon$  be  $f$ ,  $t$ , and  $h$  translated with vector  $(\varepsilon, 0)$ , respectively.

While doing the translation, the bisectors move accordingly. First notice that a point  $p$  of  $S$  can change the type of shortest path (from using the highway to walking, or *vice versa*) whenever  $p$  passes through a bisector. By choice of  $\delta_1$  and  $\delta_2$ , this cannot happen during this translation. Again, by construction of  $\delta_1$  and  $\delta_2$ , no grid line can pass through a highway endpoint. In particular, the shortest path topology from  $p$  to  $f$  is unaffected. Thus, by linearity of the  $L_1$  metric, we have  $|d_{t_\varepsilon}(p, f_\varepsilon) - d_t(p, f)| = \varepsilon$ .

Given a real number  $x$ , let  $sgn(x)$  denote the sign of  $x$ . We partition  $S$  into three sets  $S_1$ ,  $S_2$  and  $S_3$  as follows:

$$\begin{aligned} S_1 &= \{p \in S \mid sgn(d_{t_\varepsilon}(p, f_\varepsilon) - d_t(p, f)) = sgn(\varepsilon), \quad \forall \varepsilon \in [-\delta_1, \delta_2] \setminus \{0\}\} \\ S_2 &= \{p \in S \mid sgn(d_{t_\varepsilon}(p, f_\varepsilon) - d_t(p, f)) = -sgn(\varepsilon), \quad \forall \varepsilon \in [-\delta_1, \delta_2] \setminus \{0\}\} \\ S_3 &= \{p \in S \mid sgn(d_{t_\varepsilon}(p, f_\varepsilon) - d_t(p, f)) = -1, \quad \forall \varepsilon \in [-\delta_1, \delta_2] \setminus \{0\}\} \end{aligned}$$

Geometrically speaking, the elements of  $S_3$  belong to the bisector of  $f$  and  $t$ ,  $S_1$  contains the demand points that travel leftwards to reach  $f$ , and  $S_2$  contains the points that travel rightwards. Theoretically, one could consider the case in which a point belongs to set  $S_4 = \{p \in S \mid sgn(d_{t_\varepsilon}(p, f_\varepsilon) - d_t(p, f)) = 1, \quad \forall \varepsilon \in [-\delta_1, \delta_2] \setminus \{0\}\}$ . The points of this set are those that, when translating the highway in either direction, the distance between them and the entry point of the highway increases. This situation can only happen when the point is vertically aligned with the entry point: that is, point  $p \in S_4$  if and only if either (i)  $p$  uses the highway to reach the facility and it is vertically aligned with  $t$ , or (ii)  $p$  walks to the facility and it is vertically aligned with  $f$ . However,

by definition of  $\delta_1$  and  $\delta_2$ , no point of  $S$  can belong to (or enter)  $S_4$  during the whole translation.

Whenever we translate the highway  $\varepsilon$  units to the right (for some  $0 < \varepsilon \leq \delta_1$ ), the highway will be  $\varepsilon$  units closer for points in  $S_2 \cup S_3$ , but  $\varepsilon$  units further away for points of  $S_1$ . Analogously, the distance to the facility decreases for points in  $S_1 \cup S_3$  and increases for points of  $S_2$  when translating  $h$  leftwards. Given  $X \subseteq S$ , let  $w(X) = \sum_{p \in X} w_p$  be the sum of weights of the points in set  $X$ . Let  $N = w(S_1) - w(S_2)$  and  $k = w(S_3)$ . Thus, the change of the objective function when translating the highway with vector  $(\varepsilon, 0)$  (for  $\varepsilon \in [-\delta_1, \delta_2] \setminus \{0\}$ ) is,

$$\sum_{p \in S} w_p \cdot d_{t_\varepsilon}(p, f_\varepsilon) - \sum_{p \in S} w_p \cdot d_t(p, f) = w(S_1)\varepsilon - w(S_2)\varepsilon - w(S_3)|\varepsilon| = N\varepsilon - k|\varepsilon|.$$

We claim that  $k = 0$  and  $N = 0$ . In fact, since  $h$  is optimal  $N\varepsilon - k|\varepsilon| \geq 0$  for all  $\varepsilon \in [-\delta_1, \delta_2]$ . If  $\varepsilon > 0$  then  $N \geq k$ . Otherwise,  $\varepsilon < 0$  implies  $N \leq -k$ . Therefore, if  $k > 0$  then we have a contradiction, that is,  $N \leq -k$  and  $N \geq k$ . Thus  $k = 0$  implying  $N = 0$ . Therefore we can translate  $h$  either rightwards or leftwards in such a way the objective function keeps unchanged.

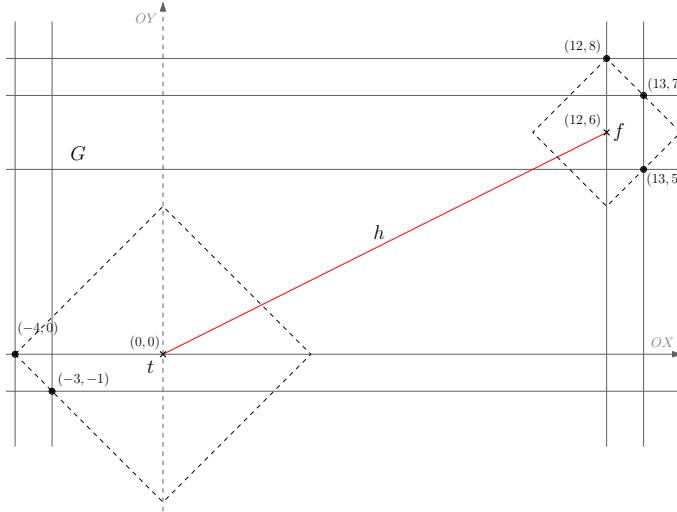
More importantly, observe that the value of  $k$  must remain 0 on the whole translation: if at some point it becomes positive, we can find a translation from that point that reduces the cost of the objective function. In particular, the set  $S_3$  must remain empty during the whole translation. Any point that changes from set  $S_1$  to  $S_2$  (or *vice versa*) must first enter  $S_3$ . Since the latter set remains empty during the whole translation, no point can change between sets  $S_1, S_2$ , or  $S_3$  until either  $f$  or  $t$  is vertically aligned with a point of  $S$ .

We repeat this operation until  $f$  or  $t$  is on a vertical line of  $G$  or a point of  $S$  reaches the bisector. In the latter case, we will have  $k > 0$ . In particular, we can find a translation that reduces the total cost of the solution. This contradicts with the assumption that the original highway location was optimal, so it cannot happen. Thus, we can translate both rightwards or leftwards until one of the two highway endpoints reaches a horizontal line of  $G$ . We repeat the same process for the  $y$ -coordinates, hence proving the Lemma.  $\square$

In [8] the authors stated that an optimal solution always exists satisfying Lemma 1 (a). Unfortunately, the above claim is false and their algorithm may miss some highway locations; indeed, it may miss the optimal location and thus fail. We provide here one counterexample and the following result.

**Lemma 2.** *There exists a set of unweighted points in which no optimal solution to the FHL-problem satisfies Lemma 1 (a).*

*Proof.* Consider the problem instance with five points whose coordinates are  $(-4, 0)$ ,  $(-3, -1)$ ,  $(12, 8)$ ,  $(13, 5)$ , and  $(13, 7)$ , respectively (see Figure 3). In the problem instance, we give unit weight to all points, and set the length  $h$  of the highway as  $\ell = \sqrt{180} \approx 13.5$ . For simplicity in the calculations, we also set  $v = \ell$ , but any other large number works as well. The cost associated with the highway of endpoints  $f = (12, 6)$  and  $t = (0, 0)$  is  $10 + 2\ell/v = 12$ . We claim



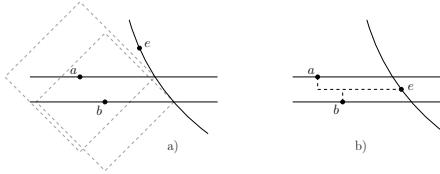
**Fig. 3.** A counterexample to the algorithm of Espejo and Rodríguez-Chía

that this location is better than any other solution with an endpoint at a vertex of  $G$ .

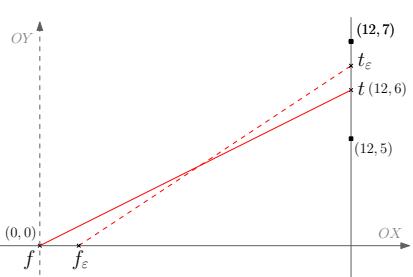
If one endpoint of  $h$  is a vertex of  $G$  in the line  $x = -3$ , then the other endpoint is located to the left of the line  $x = 11$  because  $-3 + \ell < 11$ . In that case we can translate  $h$  rightwards with vector  $(\frac{1}{2}, 0)$  improving the objective function. The same holds if one endpoint of  $h$  is a vertex in the line  $x = -4$ . Similarly, if one endpoint is a vertex in the line  $x = 13$ , then we can translate  $h$  leftwards with vector  $(-\frac{1}{2}, 0)$  and the objective function decreases.

Consider now locating one of the highway endpoints at coordinates  $(12, 0)$  or  $(12, -1)$ . Observe that the walking time (i.e., the traveling time when the highway is not used) from the points  $(-4, 0)$  and  $(-3, -1)$  takes at least 15 units of time, which is more than the cost associated with our solution. The same happens to the sum of the traveling times of the three other points. Hence, if  $f$  is located at one of the two vertices, the five points must use the highway (otherwise the travel time is higher than our solution). Analogously, if  $t$  is located at grid points  $(12, 0)$  or  $(12, -1)$ , no point of  $S$  will use the highway. In either case, the corresponding solution is at least as high as the sum of distances from all points of  $S$  to the geometric median, which is higher than the cost associated with our solution.

Consider now the cases in which one of the endpoints has coordinates  $(12, y_0)$  for some  $y_0 \in \{5, 7, 8\}$ . We start by showing that, in any of the three cases, the optimal position of the other endpoint of the highway (denoted by  $e$ ) must lie on the line  $y = 0$ . Since the highway's length is equal to  $\ell$ , the possible positions of  $e$  lie both in circle  $\sigma$  of radius  $\ell$  centered at  $(12, y_0)$  and to the left of line  $x = 12$ . Observe that the clients that walk to  $e$  are points  $a = (-4, 0)$  and  $b = (-3, -1)$ , located always to the left of  $e$ . Hence, we are interested in



**Fig. 4.**  $a = (-4, 0)$  and  $b = (-3, -1)$ . When one endpoint of the highway has coordinates  $(12, 8)$ ,  $(12, 7)$ , or  $(12, 5)$ , the optimal position of the other endpoint  $e$  is on the line  $y = 0$



**Fig. 5.** Definitions of  $f_\varepsilon$  and  $t_\varepsilon$

minimizing the expression  $\|a - e\|_1 + \|b - e\|_1$ . Let  $a', b' \in \sigma$  be the closest points to  $a$  and  $b$  with the  $L_1$  metric that verify  $y(a') = 0$  and  $y(b') = -1$ , respectively. Observe that if  $y(e) > 0$  then  $\|a - a'\|_1 < \|a - e\|_1$  and  $\|b - b'\|_1 < \|b - e\|_1$  implying

$$\|a - a'\|_1 + \|b - b'\|_1 < \|a - e\|_1 + \|b - e\|_1$$

(see Figure 4 a)). Similarly, if  $y(e) < -1$ , then

$$\|a - b'\|_1 + \|b - b'\|_1 < \|a - e\|_1 + \|b - e\|_1.$$

Therefore,  $e$  must satisfy  $-1 \leq y(e) \leq 0$  (see Figure 4 b)). In this case we have

$$\begin{aligned} \|a - e\|_1 + \|b - e\|_1 &= x(e) - x(a) + y(a) - y(e) + x(e) - x(b) + y(e) - y(b) \\ &= 2x(e) + 8 \end{aligned}$$

Then  $\|a - e\|_1 + \|b - e\|_1$  is minimized when  $x(e)$  is minimum, and it happens when  $y(e) = 0$ .

If  $y_0 = 8$ , then  $h$  can be translated downwards with vector  $(0, -\frac{1}{2})$  and the value of the objective function decreases. Thus point  $(12, 8)$  is discarded. It remains to show that there is a solution better than the one having an endpoint at either  $(12, 7)$  or  $(12, 5)$ , and the other endpoint on the line  $y = 0$ . Observe that if  $f$  and  $t$  belong to the lines  $y = 0$  and  $x = 12$ , respectively, then by exchanging  $f$  and  $t$  the value of the objective function reduces in  $\ell/v$ . Then consider the case where  $y(t) = 0$  and  $x(f) = 12$ .

Let  $t = (0, 0)$  and  $f = (12, 6)$ . Given a value  $\varepsilon$ , let  $t_\varepsilon$  be the point with coordinates  $(\varepsilon, 0)$  and  $f_\varepsilon$  be the point in the line  $x = 12$  such that  $y(f_\varepsilon) > 0$  and the Euclidean distance between  $f_\varepsilon$  and  $t_\varepsilon$  is equal to  $\ell$  (see Figure 5). Let  $[-\delta_1, \delta_2]$ ,  $\delta_1, \delta_2 > 0$ , be the maximal-length interval such that  $5 \leq y(f_\varepsilon) \leq 7$  for all  $\varepsilon \in [-\delta_1, \delta_2]$ . Note  $\delta_1 = \sqrt{155} - 12 < 1$  and  $\delta_2 = 12 - \sqrt{131} < 1$ . Then  $|\varepsilon| < 1$ .

The variation of the objective function's value when  $f$  and  $t$  are moved to  $f_\varepsilon$  and  $t_\varepsilon$ , respectively, is equal to

$$\begin{aligned} g(\varepsilon) &:= \Phi(f_\varepsilon, t_\varepsilon) - \Phi(f, t) \\ &= 2(x(t_\varepsilon) - x(t)) - (y(f_\varepsilon) - y(f)) \\ &= 2\varepsilon - \left( \sqrt{36 + 24\varepsilon - \varepsilon^2} - 6 \right). \end{aligned}$$

In the following we will show that  $\sqrt{36 + 24\varepsilon - \varepsilon^2} < 6 + 2\varepsilon$ , for all  $\varepsilon \in [-\delta_1, \delta_2] \setminus \{0\}$ . In particular, we will have  $g(\varepsilon) > 0$  (except when  $\varepsilon = 0$ ), implying that our highway location is optimal. First observe that  $4\varepsilon^2 + 24\varepsilon + 36 = (2\varepsilon + 6)^2 > 36 + 24\varepsilon - \varepsilon^2$ . Since  $|\varepsilon| < 1$  then  $2\varepsilon + 6 > 0$  and  $36 + 24\varepsilon - \varepsilon^2 > 0$ , which implies  $2\varepsilon + 6 > \sqrt{36 + 24\varepsilon - \varepsilon^2}$ . Thus  $g(\varepsilon) > 0$  and the highway with endpoints  $f$  and  $t$  gives a better solution than that having an endpoint at  $(12, 7)$  or  $(12, 5)$ . This completes the proof.  $\square$

From now on, we assume that there are no two points on a line having slope in the set  $\{-1, 0, 1, \infty\}$ .

### 3 The Algorithm

Lemma 1 can be used to find an optimal solution to the FHL-problem. Although the method is quite similar for both cases in Lemma 1, we address the two cases independently for the sake of clarity. By Vertex-FHL-problem we will denote the FHL-problem for the cases in which Lemma 1 a) holds, and by Edge-FHL-problem the FHL-problem for the cases in which Lemma 1 b) holds. In the next subsections we give an  $O(n^3)$ -time algorithm for each variant of the problem. In both of them we assume w.l.o.g. that the highway length  $\ell$  is equal to one.

In the following  $\theta$  will denote the positive angle of the highway with respect to the positive direction of the  $x$ -axis. For the sake of clarity, we will assume that  $\theta \in [0, \frac{\pi}{4}]$ . When  $\theta$  belongs to the interval  $[k\frac{\pi}{4}, (k+1)\frac{\pi}{4}]$ ,  $k = 1, \dots, 7$ , both the Vertex- and Edge-FHL-problem can be solved in a similar way.

Given a point  $u$  and an angle  $\theta$ , let  $u(\theta)$  be the point with coordinates  $(x(u) + \cos \theta, y(u) + \sin \theta)$ . There exists an angle  $\phi \in [0, \frac{\pi}{4}]$  such that the bisector of the endpoints  $f$  and  $t = f(\theta)$  has the shape in Figure 2 a) for all  $\theta \in [0, \phi]$ , and has the shape in Figure 2 b) for all  $\theta \in (\phi, \frac{\pi}{4}]$ . Such an angle  $\phi$  verifies  $\cos(\phi) - \sin(\phi) = \frac{1}{v}$ . Furthermore,  $\phi = \frac{1}{2} \arcsin(1 - \frac{1}{v^2})$  and  $\phi \neq \frac{\pi}{4}$  unless  $v$  is infinite.

Let  $\Pi_x$ ,  $\Pi_y$ , and  $\Pi_1$  denote set  $S$  sorted according to the  $x$ ,  $y$ , and  $(x+y)$  order, respectively.

#### 3.1 Solving the Vertex-FHL-Problem

For each vertex  $u$  of  $G$  we can solve the problem subject to  $f = u$  or  $t = u$ . We show how to obtain a solution if  $f = u$ . The case where  $t = u$  can be solved analogously. Suppose w.l.o.g. that the vertex  $f = u$  is the origin of the coordinate system and the highway angle is  $\theta$ , for  $\theta \in [0, \frac{\pi}{4}]$ . Thus, the distance between a point  $p \in S$  and the facility  $u$  has the expression  $c_1 + c_2 \cos \theta + c_3 \sin \theta$ , where  $c_1 > 0$  and either  $c_2, c_3 = \pm w_p$  ( $p$  uses the highway) or  $c_2 = c_3 = 0$  ( $p$  does not use the highway). When  $\theta$  goes from 0 to  $\frac{\pi}{4}$  this expression changes at the values of  $\theta$  such that:

- The point  $p$  switches from using the highway to going directly to the facility (or vice versa). We call these changes *bisector events*. A bisector event occurs

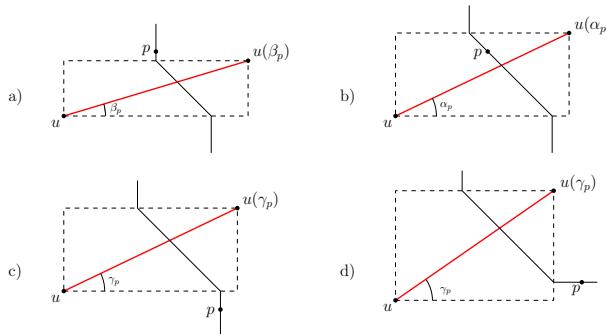
when the bisector between the highway's endpoints  $u$  and  $u(\theta)$ , contains  $p$ . At most two bisector events are obtained for each point  $p$ .

- The highway endpoint  $u(\theta)$  crosses the vertical or horizontal line passing through  $p$ . We call this event *grid event*. Again, each point of  $S$  generates at most two grid events.
- $\theta = \phi$ . This event is called  *$\phi$ -event*.

**Lemma 3.** *After an  $O(n \log n)$ -time preprocessing, the angular order of all the events associated with a given vertex of  $G$  can be obtained in linear time.*

*Proof.* The preprocessing consists in computing  $\Pi_x$ ,  $\Pi_y$ , and  $\Pi_1$ , which can be done in  $O(n \log n)$  time. Now, let  $u$  be a vertex of  $G$ . It is straightforward to see that there are  $O(n)$  grid events and that we can obtain their angular order in linear time by using both  $\Pi_x$  and  $\Pi_y$ . Let us show how to obtain the bisector events in  $O(n)$  time.

The bisector of  $u$  and  $u(\theta)$  consists of two axis-aligned half-lines and a line segment with slope -1 connecting their endpoints (see Figure 2 and [8] for further details). Given a point  $p$ , when  $\theta$  goes from 0 to  $\pi/4$  the bisector between  $u$  and  $u(\theta)$  passes through  $p$  at most twice, that is, when  $p$  belongs to one of the half-lines of the bisector and when  $p$  belongs to the line segment. If  $p$  belongs to the line segment of the bisector then the event is denoted by  $\alpha_p$  (see Figure 6 b)). If  $p$  belongs to the leftmost half-line of the bisector, which is always vertical, we denote that event by  $\beta_p$  (see Figure 6 a)). Otherwise, if  $p$  belongs to the rightmost half-line which can be either vertical or horizontal we denote that event by  $\gamma_p$  (see Figure 6 c) and d)). Observe that if the rightmost half-line is vertical then  $\gamma_p < \phi$ , otherwise  $\gamma_p > \phi$ .



**Fig. 6.** The bisector events of  $p$  when  $\theta \in [0, \frac{\pi}{4}]$ . a)  $p$  belongs to the leftmost half-line of the bisector of  $u$  and  $u(\theta)$ . b)  $p$  belongs to the segment. c,d)  $p$  belongs to the rightmost half-line of the bisector.

Let  $\Pi_1$  be the subsequence of  $\Pi_1$  containing all elements  $p$  such that  $\alpha_p \in [0, \frac{\pi}{4}]$ ,  $\Pi_2$  be the subsequence of  $\Pi_x$  containing all elements  $p$  such that  $\beta_p \in [0, \frac{\pi}{4}]$ , and  $\Pi_3$  be the subsequence of  $\Pi_x$  that contains all elements  $p$  such that

$y(p) < y(u)$  and  $\gamma_p \in [0, \frac{\pi}{4}]$ , concatenated with the subsequence of  $\Pi_y$  that contains all elements  $p$  such that  $x(p) > x(u)$  and  $\gamma_p \in [0, \frac{\pi}{4}]$ . Given a point  $p \in S$ , the corresponding events of  $p$  in  $[0, \frac{\pi}{4}]$  can be found in constant time, thus  $\Pi_1$ ,  $\Pi_2$ , and  $\Pi_3$  can be built in linear time. The following statements are true for any point  $p \in S$ :

- (a)  $x(p) + y(p) = \frac{1}{2}(\cos \alpha_p + \sin \alpha_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_1$ .
- (b)  $x(p) = \frac{1}{2}(\cos \beta_p - \sin \beta_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_2$ .
- (c)  $x(p) = \frac{1}{2}(\cos \gamma_p + \sin \gamma_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_3$  such that  $\gamma_p < \phi$ .
- (d)  $y(p) = \frac{1}{2}(-\cos \gamma_p + \sin \gamma_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_3$  such that  $\gamma_p > \phi$ .

Let  $\Gamma_1$  (resp.  $\Gamma_2$ ,  $\Gamma_3$ ) be the sequence obtained by replacing each element  $p$  in  $\Pi_1$  (resp.  $\Pi_2$ ,  $\Pi_3$ ) by  $\alpha_p$  (resp.  $\beta_p$ ,  $\gamma_p$ ). Therefore, from statements (a) – (d) and the monotonicity of the functions  $\cos \theta + \sin \theta$ ,  $\cos \theta - \sin \theta$ , and  $-\cos \theta + \sin \theta$  in the interval  $[0, \frac{\pi}{4}]$ , we obtain that  $\Gamma_1$ ,  $\Gamma_2$ , and  $\Gamma_3$  are sorted sequences. Using a standard method for merging sorted lists, we can merge in linear time  $\Gamma_1$ ,  $\Gamma_2$ ,  $\Gamma_3$ , the grid events, and the  $\phi$ -event. Therefore, the angular order of all events associated with a vertex  $u$  can be obtained in  $O(n)$  time and the result follows.  $\square$

**Theorem 1.** *The Vertex-FHL-problem can be solved in  $O(n^3)$  time.*

*Proof.* Let  $u$  be a vertex of  $G$ . Using Lemma 3, we obtain in linear time the angular order of the  $O(n)$  events associated with  $u$ . The events induce a partition of  $[0, \frac{\pi}{4}]$  into maximal intervals. For each of those intervals, the objective function takes the form  $g(\theta) := \sum_{p \in S} w_p \cdot d_t(p, u) = b_1 + b_2 \cos \theta + b_3 \sin \theta$ . This problem is of constant size in each subinterval and the minimum of  $g(\theta)$  can be found in  $O(1)$  time. Furthermore, the expression of  $g(\theta)$  can be updated in constant time when  $\theta$  crosses an event point distinct of  $\phi$  when going from 0 to  $\frac{\pi}{4}$ . In the case where  $\theta$  crosses  $\phi$ ,  $g(\theta)$  can be updated in at most  $O(n)$  time. Then the problem subject to  $f = u$  can be solved in linear time. The case in which  $t = u$  can be addressed in a similar way. It gives an overall  $O(n^3)$  time complexity because  $G$  has  $O(n^2)$  vertices.  $\square$

### 3.2 Solving the Edge-FHL-Problem

We now consider the case in which the optimal solution satisfies condition b) of Lemma 1. Namely, we consider a horizontal line  $e_h$  of  $G$  and each vertical line  $e_v$  of  $G$ . For every pair of such lines, we consider eight different sub-cases, depending on whether  $h$  is located above/below  $e_h$ , rightwards/leftwards of  $e_v$ , and  $f \in e_h$  and  $t \in e_v$  (or *vice versa*). For a fixed sub-case, we parameterize the location of the highway by the angle  $\theta$  that the highway forms with  $e_h$ . As in the Vertex-FHL case, we assume that  $f \in e_h$ ,  $t \in e_v$ , and  $\theta \in [0, \frac{\pi}{4}]$ .

We implicitly redefine the coordinate system so that  $e_h$  and  $e_v$  intersect at the origin  $o$ . Let  $\theta \in [0, \frac{\pi}{4}]$  be the positive angle of the highway with respect to the positive direction of the  $x$ -axis and  $f = x_\theta$ ,  $t = y_\theta$  be the highway endpoints.

First notice that, since we are again doing a continuous translation of  $h$ , the events that affect the value of the objective function are exactly the same as those that happen in the Vertex-FHL-problem: bisector-, grid- and  $\phi$ -events. We start by showing that the equivalent of Lemma 3 also holds:

**Lemma 4.** *After an  $O(n \log n)$ -time preprocessing, the angular order of all the events associated with a pair of perpendicular lines of  $G$  can be obtained in linear time.*

*Proof.* We can follow the arguments of Lemma 3. Firstly, we note that there are  $O(n)$  grid events and their angular order can be obtained in linear time by using both  $\Pi_x$  and  $\Pi_y$ .

Given a point  $p \in S$ , let the events  $\alpha_p$ ,  $\beta_p$ , and  $\gamma_p$  be defined as in the Vertex-FHL case. Refer to Figure 6. Let  $\Pi_1$  be the subsequence of  $\Pi_1$  containing all elements  $p$  such that  $\alpha_p \in [0, \frac{\pi}{4}]$ ,  $\Pi_2$  be the subsequence of  $\Pi_x$  containing all elements  $p$  such that  $\beta_p \in [0, \frac{\pi}{4}]$ , and  $\Pi_3$  be the subsequence of  $\Pi_x$  that contains all elements  $p$  such that  $y(p) < y(o)$  and  $\gamma_p \in [0, \phi]$ , concatenated with the subsequence of  $\Pi_y$  that contains all elements  $p$  such that  $x(p) > x(o)$  and  $\gamma_p \in [\phi, \frac{\pi}{4}]$ . Note that  $\Pi_1$ ,  $\Pi_2$ , and  $\Pi_3$  can be built in linear time.

Given a point  $p \in S$ , the following statements are true:

- (a)  $x(p) + y(p) = \frac{1}{2}(-\cos \alpha_p + \sin \alpha_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_1$ .
- (b)  $x(p) = \frac{1}{2}(-\cos \beta_p - \sin \beta_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_2$ .
- (c)  $x(p) = \frac{1}{2}(-\cos \gamma_p + \sin \gamma_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_3$  such that  $\gamma_p < \phi$ .
- (d)  $y(p) = \frac{1}{2}(-\cos \gamma_p + \sin \gamma_p + \frac{1}{v})$  for all points  $p$  in  $\Pi_3$  such that  $\gamma_p > \phi$ .

Let  $\Gamma_1$  (resp.  $\Gamma_2$ ,  $\Gamma_3$ ) be the sequence obtained by replacing each element  $p$  in  $\Pi_1$  (resp.  $\Pi_2$ ,  $\Pi_3$ ) by  $\alpha_p$  (resp.  $\beta_p$ ,  $\gamma_b$ ). Therefore, by using similar arguments to those used in Lemma 3 the angular order of all events can be obtained in  $O(n)$  time, once the lists  $\Pi_x$ ,  $\Pi_y$  and  $\Pi_{x+y}$  have been precomputed.  $\square$

Consider now a small interval  $[\theta_1, \theta_2]$  in which no event occurs. After the coordinate system redefinition, we have  $f = x_\theta = (-\cos \theta, 0)$ , and  $t = y_\theta = (0, \ell \sin \theta)$ . Let  $p \in S$  be a point that uses the highway to reach the facility; since only the  $y$ -coordinate of  $t$  changes, its distance to  $f$  can be expressed as  $c_1 \pm \sin \theta$  for some  $c_1 > 0$ . Analogously, if  $p$  walks to  $f$ , its distance is of the form  $c_1 \pm \cos \theta$  for some  $c_1 > 0$ . That is, the distance between a point of  $S$  and  $f$  in any interval is of the form  $c_1 + c_2 \sin \theta + c_3 \cos \theta$  for some constants  $c_1 > 0$  and  $c_2, c_3 \in \{-1, 0, 1\}$ .

**Theorem 2.** *The Edge-FHL-problem can be solved in  $O(n^3)$  time.*

*Proof.* We can use a method similar to the one used in the Vertex-FHL-problem. Let  $e_h$  be a horizontal line of  $G$  and  $e_v$  be a vertical line of  $G$ .

Using Lemma 4, we obtain in linear time the angular order of the  $O(n)$  events associated with  $e_h$  and  $e_v$ . The events induce a partition of  $[0, \frac{\pi}{4}]$  into maximal intervals. For each of those intervals the objective function has the form  $g(\theta) := \Phi(f, t) = \Phi(x_\theta, y_\theta) = b_1 + b_2 \cos \theta + b_3 \sin \theta$ , where  $b_1 > 0$ , and  $b_2, b_3 \in \mathbb{Z}$  are constants. This problem has constant size, hence the minimum of  $g(\theta)$  can

be found in  $O(1)$  time. Furthermore, the expression of  $g(\theta)$  can be updated in constant time when  $\theta$  crosses an event point distinct of  $\phi$  when it goes from 0 to  $\frac{\pi}{4}$ . In the case where  $\theta$  crosses  $\phi$ ,  $g(\theta)$  can be updated in at most  $O(n)$  time. Then the problem subject to  $f \in e_h$  and  $t \in e_v$  can be solved in linear time. It gives an overall  $O(n^3)$  time complexity because  $G$  has  $O(n^2)$  pairs consisting of a horizontal and a vertical line.  $\square$

## References

1. Abellanas, M., Hurtado, F., Icking, C., Klein, R., Langetepe, E., Ma, L., Palop, B., Sacristán, V.: Voronoi diagram for services neighboring a highway. *Information Processing Letters* 86, 283–288 (2003)
2. Ahn, H.-K., Alt, H., Asano, T., Bae, S.W., Brass, P., Cheong, O., Knauer, C., Na, H.-S., Shin, C.-S., Wolff, A.: Constructing optimal highways. In: Proceedings of the 13th Computing: The Australasian Theory Symposium (CATS 2007), pp. 7–14 (2007)
3. Aichholzer, O., Aurenhammer, F., Palop, B.: Quickest paths, straight skeletons, and the city Voronoi diagram. *Discrete & Computational Geometry* 31, 17–35 (2004)
4. Aloupis, G., Cardinal, J., Collette, S., Hurtado, F., Langerman, S., O'Rourke, J., Palop, B.: Highway hull revisited. *Computational Geometry: Theory and Applications* 43, 115–130 (2010)
5. Bae, S.W., Korman, M., Tokuyama, T.: All Farthest Neighbors in the Presence of Highways and Obstacles. In: Das, S., Uehara, R. (eds.) WALCOM 2009. LNCS, vol. 5431, pp. 71–82. Springer, Heidelberg (2009)
6. Cardinal, J., Collette, S., Hurtado, F., Langerman, S., Palop, B.: Optimal location of transportation devices. *Computational Geometry: Theory and Applications* 41, 219–229 (2008)
7. Díaz-Báñez, J.-M., Korman, M., Pérez-Lantero, P., Ventura, I.: Locating a service facility and a rapid transit line. CoRR, abs/1104.0753 (2011)
8. Espejo, I., Rodríguez-Chía, A.M.: Simultaneous location of a service facility and a rapid transit line. *Computers and Operations Research* 38, 525–538 (2011)
9. Korman, M., Tokuyama, T.: Optimal Insertion of a Segment Highway in a City Metric. In: Hu, X., Wang, J. (eds.) COCOON 2008. LNCS, vol. 5092, pp. 611–620. Springer, Heidelberg (2008)

# Simultaneously Flippable Edges in Triangulations\*

Diane L. Souvaine<sup>1</sup>, Csaba D. Tóth<sup>2</sup>, and Andrew Winslow<sup>1</sup>

<sup>1</sup> Tufts University, Medford MA 02155, USA

{dls,awinslow}@cs.tufts.edu

<sup>2</sup> University of Calgary, Calgary AB T2N 1N4, Canada

cdtoth@ucalgary.ca

**Abstract.** Given a straight-line triangulation  $T$ , an edge  $e$  in  $T$  is *flippable* if  $e$  is adjacent to two triangles that form a convex quadrilateral. A set of edges  $E$  in  $T$  is *simultaneously flippable* if each edge is flippable and no two edges are adjacent to a common triangle. Intuitively, an edge is flippable if it may be replaced with the other diagonal of its quadrilateral without creating edge-edge intersections, and a set of edges is simultaneously flippable if they may be all be flipped without interfering with each other. We show that every straight-line triangulation on  $n$  vertices contains at least  $(n - 4)/5$  simultaneously flippable edges. This bound is the best possible, and resolves an open problem by Galtier *et al.*.

**Keywords:** planar graph, graph transformation, geometry, combinatorics.

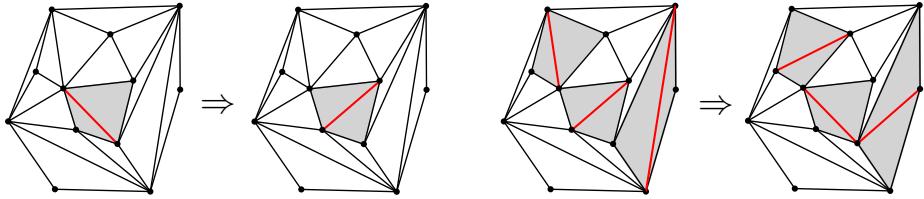
## 1 Introduction

A (*geometric*) *triangulation* of a point set  $P$  is a planar straight line graph with vertex set  $P$  such that every bounded face is a triangle, and the outer face is the exterior of the convex hull of  $P$ . An edge  $e$  of a triangulation is *flippable* if it is adjacent to two triangles whose union is a convex quadrilateral  $Q(e)$  (see Figure 1). An *edge flip* is performed by exchanging a flippable edge with the other diagonal of the convex quadrilateral it lies in. Hurtado *et al.* [5] proved that every triangulation on  $n$  vertices has at least  $(n - 4)/2$  flippable edges, and this bound cannot be improved in general.

A set  $E$  of edges in a triangulation are *simultaneously flippable* if each edge in  $E$  is flippable, and the quadrilaterals  $Q(e)$ ,  $e \in E$ , are pairwise interior disjoint. Note that this definition does indeed imply that simultaneously flippable edges can be flipped at the same time without interfering with each other. For a triangulation  $T_P$  of a point set  $P$ , let  $f_{\text{sim}}(T_P)$  denote the maximum number of simultaneously flippable edges in  $T_P$ , and let  $f_{\text{sim}}(n) = \min_{T_P: |P|=n} f_{\text{sim}}(T_P)$

---

\* Research of Diane L. Souvaine and Andrew Winslow supported in part by NSF grants CCF-0830734 and CBET-0941538. Research of Csaba D. Tóth supported in part by NSERC grant RGPIN 35586.



**Fig. 1.** Examples of an edge flip (left) and a simultaneous edge flip (right). In the simultaneous edge flip, the convex quadrilaterals containing each flippable edge (gray) are interior-disjoint.

be the minimum of  $f_{\text{sim}}(T_P)$  over all  $n$ -element point sets  $P$  in general position in the plane. The value of  $f_{\text{sim}}(n)$  played a key role in recent results on the number of various classes of planar straight line graph embedded on given point sets [2,4].

Galtier *et al.* [3] proved that  $f_{\text{sim}}(n) \geq (n-4)/6$  and that for arbitrarily large  $n$  there are triangulations  $T_P$  with  $|P| = n$  such that  $f_{\text{sim}}(T_P) \leq (n-4)/5$ . In this note we improve the lower bound to  $f_{\text{sim}}(n) \geq (n-4)/5$ , resolving an open problem posed in [3] and restated in [1]. We also describe a family of triangulations  $T_P$ , where  $|P| = n$ , for which  $f_{\text{sim}}(T_P) = (n-4)/5$ , and which contains the set given by Galter *et al.* as a special case.

## 2 Lower Bound

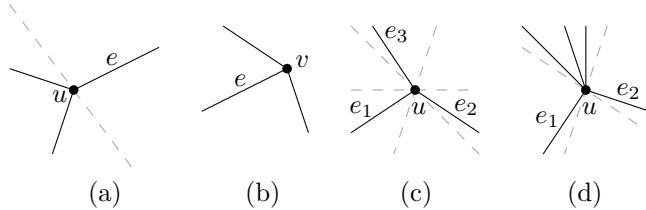
We obtain a lower bound of  $f_{\text{sim}}(n) \geq (n-4)/5$  using a three-part argument. First, we assign each non-flippable edge to an incident vertex and partition the vertices by degree and number of assigned non-flippable edges. Second, we use a coloring result by Galtier *et al.* [3] to reduce the problem to the case in which the number of flippable edges is relatively small. Third, we create a 2-colored augmented triangulation to develop an upper bound on the number of flippable edges that share faces with other flippable edges. Applying this upper bound to the situation where the number of flippable edges is small yields the desired lower bound of  $f_{\text{sim}}(n) \geq (n-4)/5$ .

### 2.1 Separable Edges

The bound  $f_{\text{sim}}(n) \geq (n-4)/5$  trivially holds for  $n \leq 4$ , so we assume  $n > 4$  for the remainder of the proof. Following the terminology in [4], we say that an edge  $e = uv$  of the triangulation is *separable* at vertex  $u$  iff there is a line  $\ell_u$  through  $u$  such that  $uv$  is the only edge incident to  $u$  on one side of  $\ell_u$ . Examples of separable and non-separable edges are shown in Figure 2. Note that an edge  $uv$  of  $T$  is flippable iff it is separable at *neither* endpoint, and that no edge is separable at *both* endpoints. Vertices on the convex hull are referred to as *hull vertices* and all other vertices as *interior vertices*. We use the following observations from [5]:

- If  $u$  is a hull vertex, then only the two incident hull edges are separable at  $u$ .
- If  $u$  is an interior vertex with degree 3, then all three incident edges are separable at  $u$ .
- If  $u$  is an interior vertex with degree 4 or higher, then at most two edges are separable at  $u$  and these edges are consecutive in the rotation of  $u$ .

Since no edge is separable at both endpoints, then no pair of interior degree 3 vertices is adjacent.



**Fig. 2.** Examples of edges separable and unseparable at a vertex  $u$ . The gray dashed lines denote separating lines. (a) An edge separable at a vertex (b) An edge not separable at a vertex (c) All three incident edges separable at an interior vertex (d) Two consecutive edges separable at an interior vertex.

Similarly to [5], we assign every non-flippable edge  $e$  to an incident vertex at which it is separable. If  $e$  lies on the boundary of the convex hull, assign  $e$  to its counterclockwise first hull vertex. If  $e$  is incident to an interior vertex of degree 3, then assign  $e$  to this vertex. Otherwise assign  $e$  to one of its endpoints at which it is separable, breaking ties arbitrarily. Recall that since no pair of interior degree 3 vertices is adjacent, each such vertex has all three incident edges assigned to it. See the left portion of Figure 4 for an example assignment.

Based on the above observations, we can now distinguish five types of vertices. Let  $h$  be the number of hull vertices (with  $h \geq 3$ ) and let  $n_3$  be the number of interior vertices of degree 3. Denote by  $n_{4,0}$ ,  $n_{4,1}$ , and  $n_{4,2}$  the number of interior vertices of degree 4 or higher, to which 0, 1, and 2 non-flippable edges, respectively are assigned. We have

$$n = h + n_{4,2} + n_{4,1} + n_{4,0} + n_3. \quad (1)$$

Using this notation, the number of non-flippable edges is exactly  $h + 3n_3 + 2n_{4,2} + n_{4,1}$ . By Euler's formula,  $T$  has a total of  $3n - h - 3$  edges. We use these facts to get that the total number of flippable edges in  $T$  is

$$\begin{aligned} f &= (3n - h - 3) - (h + 3n_3 + 2n_{4,2} + n_{4,1}) \\ &= (2h + 3n_3 + 3n_{4,2} + 3n_{4,1} + 3n_{4,0} - 3) - (h + 3n_3 + 2n_{4,2} + n_{4,1}) \\ &= h + n_{4,2} + 2n_{4,1} + 3n_{4,0} - 3. \end{aligned} \quad (2)$$

## 2.2 Coloring Argument

Galtier *et al.* [3] note that the edges of a geometric triangulation can be 3-colored such that the edges of each triangle have distinct colors. Any two flippable edges of the same color can be flipped simultaneously. Recall that there are at least  $(n - 4)/2$  flippable edges by the result of Hurtado *et al.* [5], and so the most popular color class contains at least  $(n - 4)/6$  simultaneously flippable edges.

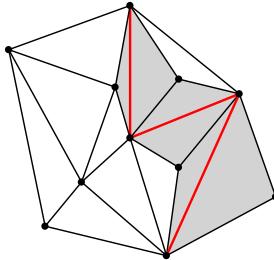
If  $f \geq 3\lceil(n - 4)/5\rceil - 2$ , then the above 3-coloring argument already implies that the largest color class of flippable edges contains at least  $\lceil(n - 4)/5\rceil$  simultaneously flippable edges, yielding the desired lower bound. Thus, in the remainder of the proof we assume that

$$f \leq 3 \left\lceil \frac{n - 4}{5} \right\rceil - 3 \leq \frac{3n}{5} - 3. \quad (3)$$

In this case we will show a slightly stronger bound, namely that  $f_{\text{sim}} > n/5 \geq \lceil(n - 4)/5\rceil$ . Since  $f \geq (n - 4)/2$ , the number of flippable edges must be in the range  $0.5n - 2 \leq f < 0.6n - 3$ . Combining (2) and (3), we have

$$\frac{3}{5}n \geq h + n_{4,2} + 2n_{4,1} + 3n_{4,0}. \quad (4)$$

We apply the 3-coloring result by Galtier *et al.* only for a subset of the flippable edges. We call a flippable edge  $e$  *isolated* if the convex quadrilateral  $Q(e)$  is bounded by 4 non-flippable edges (see Figure 3).



**Fig. 3.** The isolated edges of a triangulation. Each convex quadrilateral (gray) contains an isolated flippable edge (red), and none of its four boundary edges are flippable.

It is clear that an isolated flippable edge is simultaneously flippable with any other flippable edge. Let  $f_0$  and  $f_1$  denote the number of isolated and non-isolated flippable edges, respectively, with  $f = f_0 + f_1$ . Applying the 3-coloring argument to the non-isolated flippable edges only, the number of simultaneously flippable edges is bounded by

$$f_{\text{sim}} \geq f_0 + \frac{f_1}{3} = (f - f_1) + \frac{f_1}{3} = f - \frac{2}{3}f_1. \quad (5)$$

### 2.3 An Auxiliary Triangulation

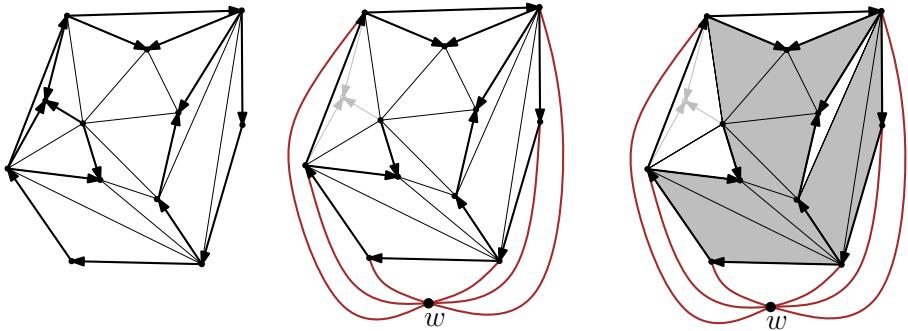
Similarly to Hurtado *et al.* [5] and Hoffmann *et al.* [4], we use an auxiliary triangulation  $\widehat{T}$ . We construct  $\widehat{T}$  from  $T$  as follows:

1. Add an auxiliary vertex  $w$  in the exterior of the convex hull, and connect it to all hull vertices.
2. Remove all interior vertices of degree 3 (and all incident edges).

An example  $\widehat{T}$  is seen in the center portion of Figure 4. Notice that only non-flippable edges have been deleted from  $T$ . In the triangulation  $\widehat{T}$ , the number of vertices is  $n - n_3 + 1 = h + n_{4,0} + n_{4,1} + n_{4,2} + 1$  and all faces (including the unbounded face) are triangles. By Euler's formula, the number of faces in  $\widehat{T}$  is:

$$m = 2(n - n_3 + 1) - 4 = 2h + 2n_{4,2} + 2n_{4,1} + 2n_{4,0} - 2. \quad (6)$$

We 2-color the faces of  $\widehat{T}$  as follows: let all triangles incident to vertex  $w$  be white; let all triangles obtained by deleting a vertex of degree 3 be white; for each of the  $n_{4,2}$  vertices (which have degree 4 or higher in  $T$  and two assigned consecutive separable edges), let the triangle adjacent to both nonflippable edges be white; finally, color all remaining triangles of  $\widehat{T}$  gray. See the right portion of Figure 4 for an example coloring.



**Fig. 4.** Producing a 2-colored auxiliary triangulation  $\widehat{T}$  of a triangulation  $T$ . The original triangulation  $T$  (left) orients each separable edge towards a vertex at which it is separable. An auxiliary triangulation  $\widehat{T}$  (center) is produced from  $T$  by adding a vertex  $w$  and removing all interior vertices of degree 3. The faces of  $\widehat{T}$  are then 2-colored (right).

Under this coloring, the number of white faces is  $m_{\text{white}} = h + n_{4,2} + n_3$ . Using (6), the number of gray faces is

$$\begin{aligned} m_{\text{gray}} &= m - m_{\text{white}} \\ &= (2h + 2n_{4,2} + 2n_{4,1} + 2n_{4,0} - 2) - (h + n_{4,2} + n_3) \\ &= h + n_{4,2} + 2n_{4,1} + 2n_{4,0} - n_3 - 2. \end{aligned} \quad (7)$$

## 2.4 Putting It All Together

Observe that if a flippable edge  $e$  of  $T$  lies on the common boundary of two white triangles in the auxiliary graph  $\widehat{T}$ , then  $e$  is isolated. That is, if  $e$  is a non-isolated flippable edge in  $T$ , then it is on the boundary of a gray triangle in  $\widehat{T}$ . Since every gray triangle has three edges, the number of non-isolated flippable edges in  $T$  is at most  $3m_{\text{gray}}$ . Substituting this into our bound (5) on the number of simultaneously flippable edges, we have

$$\begin{aligned} f_{\text{sim}} &\geq f - \frac{2}{3}f_1 \\ &\geq f - 2m_{\text{gray}} \\ &= (h + n_{4,2} + 2n_{4,1} + 3n_{4,0} - 3) - 2(h + n_{4,2} + 2n_{4,1} + 2n_{4,0} - 2 - n_3) \\ &= 2n_3 - h - n_{4,2} - 2n_{4,1} - n_{4,0} + 1. \end{aligned} \tag{8}$$

Finally, combining twice (1) minus three times (4), we obtain

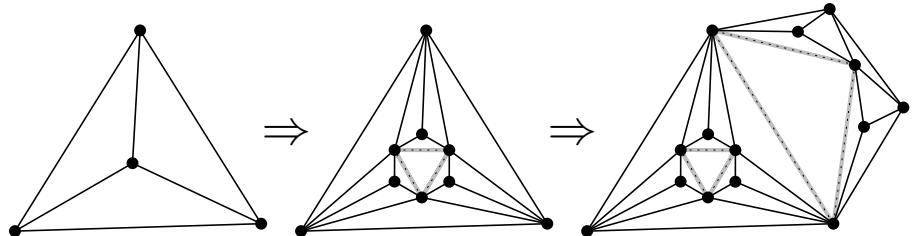
$$\begin{aligned} 2n - 3 \cdot \frac{3n}{5} &\leq 2(h + n_{4,2} + n_{4,1} + n_{4,0} + n_3) - 3(h + n_{4,2} + 2n_{4,1} + 3n_{4,0}) \\ \frac{n}{5} &\leq 2n_3 - h - n_{4,2} - 4n_{4,1} - 7n_{4,0} \\ &< 2n_3 - h - n_{4,2} - 2n_{4,1} - n_{4,0} + 1 \\ &\leq f_{\text{sim}}. \end{aligned} \tag{9}$$

This gives a lower bound of  $f_{\text{sim}} > n/5 \geq (n-4)/5$  under the condition in (3). Recall that if the condition does not hold, then a lower bound of  $f_{\text{sim}} \geq (n-4)/5$  is achieved by applying the 3-coloring argument by Galtier *et al.* to all flippable edges.

## 3 Upper Bound Constructions

In this section we construct an infinite family of geometric triangulations with at most  $(n-4)/5$  simultaneously flippable edges. This family includes all triangulations constructed by Galtier *et al.* [3]. First observe that a straight line drawing of  $K_4$  has no flippable edge. We introduce two operations that each increase the number of vertices by 5, and the maximum number of simultaneously flippable edges by one.

One operation replaces an interior vertex of degree 3 by a configuration of 6 vertices as shown at left in Fig. 5. The other operation adds 5 vertices in a close neighborhood of a hull edge as shown at right in Fig. 5. Note that both operations maintain the property of  $K_4$  that the triangles adjacent to the convex hull have no flippable edges. Each operation creates three new convex quadrilaterals formed by adjacent triangles. Because any pair of these quadrilaterals share a common triangle and none of the triangles were simultaneously flippable, the size of the largest disjoint subset of these quadrilaterals (and the number of newly-created simultaneously flippable edges) is 1. Each operation increases  $h + n_{4,2}$  by 3 and  $n_3$  by 2, as expected based on the previous section.

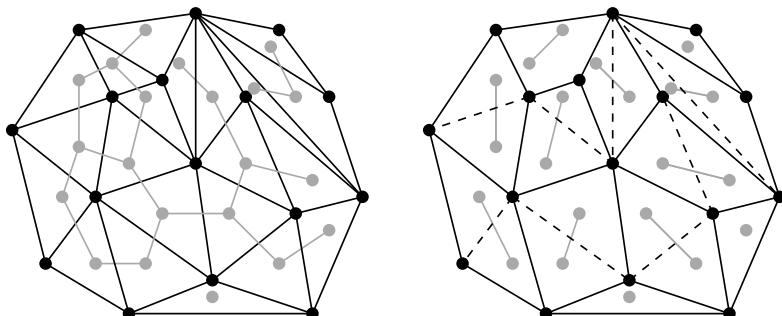


**Fig. 5.** Applying two successive operations to  $K_4$  to yield a triangulation with  $(n - 4)/5$  simultaneously flippable edges.

Let  $\mathcal{F}_{\text{sim}}$  denote the family of all geometric triangulations obtained from  $K_4$  via applying an arbitrary sequence of the two operations. Then every triangulation  $T \in \mathcal{F}_{\text{sim}}$  on  $n$  vertices has at most  $(n - 4)/5$  simultaneously flippable edges, attaining our lower bound for  $f_{\text{sim}}(n)$ . We note that all upper bound constructions by Galtier *et al.* [3] can be obtained by applying our 2nd operation successively to all sides, starting from  $K_4$ .

#### 4 Algorithmic Aspects

Finally, we consider computing a solution to the following problem: given a set of  $n$  points in the plane, and a triangulation  $T$  on these points, find a maximum set of simultaneously flippable edges of  $T$ . To solve this problem, we use the *flippable dual graph of  $T$* : a subgraph of the dual graph of  $T$  containing exactly the edges whose duals are flippable in  $T$  (see Figure 6). The following fact leads to a simple algorithm: a set of flippable edges  $E$  in  $T$  is simultaneously flippable if and only if the edges in  $E$  correspond to a matching in the flippable dual graph of  $T$ . This is true because a matching assigns at most one edge to each triangle, ensuring that no two edges are adjacent to a common triangle.



**Fig. 6.** Given the flippable dual graph (shown in gray at left) of a triangulation, a maximum matching on it corresponds to a maximum set of simultaneously flippable edges of the triangulation (shown in dashed lines at right).

Using this fact, a maximum set of simultaneously flippable edges can be computed by a maximum matching algorithm. Consider the subgraph of the triangulation consisting of flippable edges (both isolated and non-isolated) and find a maximum matching on this graph. A maximum matching can be computed in  $O(n^{1.5})$  deterministic time [7] or in  $O(n^{\omega/2}) = o(n^{1.19})$  expected time [6], where  $\omega$  is the time required to multiply two  $n \times n$  matrices. Since flippable edges can be identified in  $O(n)$  time, the algorithm runs in  $O(n^{1.5})$  deterministic time, or  $o(n^{1.19})$  expected time, depending upon which maximum matching algorithm is used.

**Acknowledgements.** We wish to acknowledge Ferran Hurtado for his influence and inspiring research on edge flippings. We also thank Oswin Aichholzer and Günter Rote for helpful comments, and anonymous reviewers for suggestions on improving presentation.

## References

1. Bose, P., Hurtado, F.: Flips in planar graphs. *Computational Geometry: Theory and Applications* 42, 60–80 (2009)
2. Dumitrescu, A., Schulz, A., Sheffer, A., Tóth, C.D.: Bounds on the maximum multiplicity of some common geometric graphs. In: 28th International Symposium on Theoretical Aspects of Computer Science, pp. 637–648. Dagstuhl Publishing, Germany (2011)
3. Galtier, J., Hurtado, F., Noy, M., Pérennes, S., Urrutia, J.: Simultaneous edge flipping in triangulations. *International Journal of Computational Geometry and Applications* 13, 113–133 (2003)
4. Hoffmann, M., Sharir, M., Sheffer, A., Tóth, C.D., Welzl, E.: Counting Plane Graphs: Flippability and Its Applications. In: Dehne, F., Iacono, J., Sack, J.-R. (eds.) WADS 2011. LNCS, vol. 6844, pp. 524–535. Springer, Heidelberg (2011)
5. Hurtado, F., Noy, M., Urrutia, J.: Flipping edges in triangulations. *Discrete and Computational Geometry* 22, 333–346 (1999)
6. Mucha, M., Sankowski, P.: Maximum matchings in planar graphs via Gaussian elimination. *Algorithmica* 45, 3–20 (2006)
7. Micali, S., Vazirani, V.V.: An  $O(\sqrt{|V|} \cdot |E|)$  algorithm for finding maximum matching in general graphs. In: 21st IEEE Symposium on Foundations of Computer Science, pp. 17–27. IEEE Press, New York (1980)
8. Urrutia, J.: Flipping edges in triangulations of point sets, polygons and maximal planar graphs. Invited talk, 1st Japanese Conference on Discrete and Computational Geometry, JCDCG 1997 (1997), [http://www.matem.unam.mx/~urrutia/online\\_papers/TrianSurv.pdf](http://www.matem.unam.mx/~urrutia/online_papers/TrianSurv.pdf)

# Spiral Serpentine Polygonization of a Planar Point Set

Justin Iwerks and Joseph S.B. Mitchell

Dept. of Applied Mathematics and Statistics  
Stony Brook University, Stony Brook, NY 11794-3600  
`{jiwerks,jsbm}@ams.stonybrook.edu`

**Abstract.** We introduce a simple algorithm for constructing a spiral serpentine polygonization of a set  $S$  of  $n \geq 3$  points in the plane. Our algorithm simultaneously gives a triangulation of the constructed polygon at no extra cost, runs in  $O(n \log n)$  time, and uses  $O(n)$  space.

**Keywords:** polygonization, triangulation, serpentine, point set, algorithm, computational geometry.

## 1 Introduction

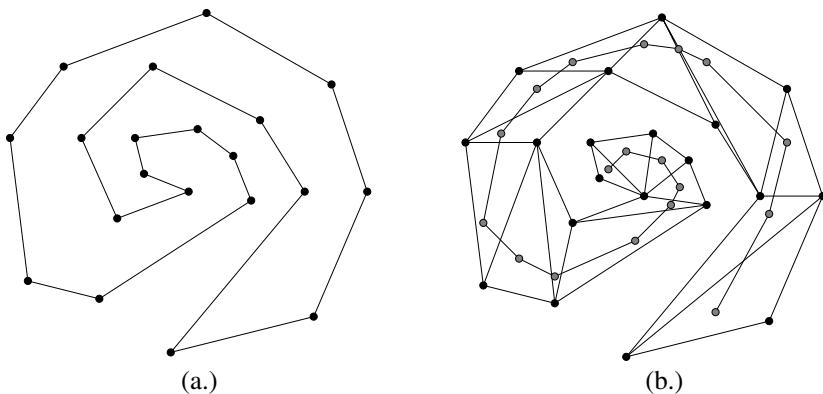
A polygonization of a planar point set  $S$  is a simple polygon having  $S$  as the set of its vertices. Different types of polygonizations have been investigated in settings where objects are being constructed from limited data, such as pattern recognition and image reconstruction [1], [5], [6]. The number of polygonizations for a given point set can be exponential in  $n$ , even when restricted to monotone or star-shaped polygonizations [11].

The existence of a polygonization for any point set  $S$  in general position was established by Steinhaus [14]. The Euclidean-TSP on  $S$  was shown to be simple by Quintas and Supnick, yielding the existence of a polygonization for any planar point set  $S$  [12]. Graham demonstrated that a *star-shaped* polygonization could be constructed explicitly in  $O(n \log n)$  time [7]. Later, Grünbaum showed implicitly that every point set  $S$ , where not all points are collinear, has a *monotone* polygonization and that it can be computed in  $O(n \log n)$  time [8]. Agarwal et al. have discussed the attractiveness of the subset of polygonizations that admit *thin* triangulations, which minimize the number of nodes of degree three in the dual, and in particular, *serpentine* triangulations, whose dual graph is a path. In [2], the authors gave an  $O(n \log n)$  algorithm for computing a monotone serpentine polygonization of a point set  $S$  of  $n$  points, not all of them collinear.

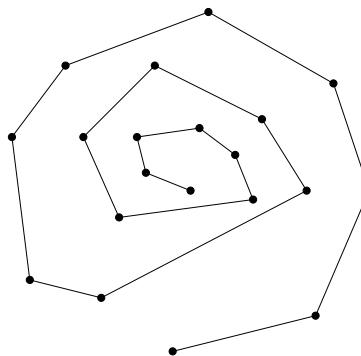
The *reflexivity* of a point set  $S$  was introduced and studied by Arkin et al. [3]; the reflexivity of  $S$  is the minimum number of reflex vertices possible in any polygonization of  $S$ . This paper motivated us to consider the problem of determining the *inflection number* of a point set  $S$ , which we define to be the minimum number of transitions between reflex and convex vertices possible in

any polygonization of  $S$ . We show that one can always polygonize  $S$  with zero or two transitions; the inflection number is zero if and only if  $S$  is in convex position.

In particular, we demonstrate that any point set  $S$  has a *spiral* serpentine polygonization. A spiral serpentine polygon is a simple polygon possessing at most one chain of reflex vertices, exactly one chain of convex vertices (see Figure 1) and admitting a serpentine triangulation. We note that it is trivial to find a spiraling polygonal simple *path* through a set of  $n$  points in the plane as in Figure 2. The task here is to construct a simple *cycle* through the points that possesses the spiral and serpentine properties.



**Fig. 1.** A spiral polygon is shown in (a.) and a serpentine triangulation of this polygon is shown in (b.), where the dual of the triangulation is the path depicted



**Fig. 2.** A spiraling polygonal simple path is easy to construct

We present a simple algorithm in Section 2 for constructing such a polygonization in  $O(n \log n)$  time, requiring  $O(n)$  space and explicitly giving a serpentine triangulation at no extra cost. Both the run-time and space complexities are optimal. To establish that the runtime is  $\Omega(n \log n)$ , we can use the same approach as in [13] where the input  $S$  contains  $n - 1$  collinear points with one additional point  $p$  not collinear with the rest. If we rotate  $S$  so that  $p$  has the lowest  $y$ -coordinate among all the points, then our algorithm will give a unique spiral serpentine polygonization of the point set that can be used to sort the  $n - 1$  collinear points.

In Section 3, we establish the correctness of the algorithm. Section 4 gives examples computed from an implementation of the algorithm. Section 5 provides a closing discussion and some related open questions.

## 2 The Algorithm

Here we introduce the algorithm `SpiralSerpentinePolygonize`, which produces a spiral serpentine polygonization of a planar set  $S$  of  $n \geq 3$  points. Refer to the pseudocode below.

In the first step of the algorithm, we compute the convex hull,  $H$ , of  $S$  and initialize a semi-dynamic convex hull data structure that allows only deletion operations [9]<sup>1</sup>. Three arrays are also initialized:  $C$ , storing the convex chain,  $R$ , storing the (possibly empty) reflex chain, and  $D$ , storing the set of diagonals for the triangulation.

After establishing the first two vertices,  $u$  and  $v$ , of the convex chain  $C$  in step 2.) and determining the next point  $w$  of  $S$  to consider in step 3.), the algorithm enters a **while** loop in step 4.) that processes one remaining point per iteration. In particular, at the start of each iteration of the loop we check whether or not the (possibly unbounded) wedge  $Q$  formed by rays emanating from  $w$  along the directions  $\overrightarrow{uw}$  and  $\overrightarrow{vw}$  contains any points of  $H$  (see Figure 3). This can be determined in  $O(\log n)$  time, using binary search in the data structure storing the points  $H$ .

If  $Q$  has a point of  $H$ , the first point  $q$  encountered by rotating counterclockwise from the ray centered at  $w$  in the direction  $\overrightarrow{uw}$  is located (in time  $O(\log n)$ ). (In determining  $q$ , ties are again broken by picking the point closest to  $w$ .) We then update  $u$  to be  $w$  and  $w$  to be  $q$ , append  $u$  to the reflex chain  $R$ , and append  $(u, v)$  to the set of diagonals  $D$ .

Otherwise, the first point  $q$  encountered by rotating counterclockwise the ray centered at  $u$  in the direction  $\overrightarrow{uw}$  is located. We then update  $v$  to be  $w$  and  $w$  to be  $q$ , append  $v$  to the convex chain  $C$ , and append  $(u, v)$  to the set of diagonals  $D$ .

---

<sup>1</sup> In a previous extended abstract of this paper [10], we give an alternative to using the dynamic convex hull data structure, based on computing the *convex layers* [4] of all points in  $S$  and exploiting special structure in the layers when our algorithm requires convex hull queries.

Once  $H$  is empty, we append  $w$  to  $C$  and let the polygonization  $P$  be stored as an array containing the concatenation of  $R$  in reverse order to the end of  $C$ . The polygonization can be constructed by outputting the edges  $(P[i], P[i + 1])$ ,  $(0 \leq i \leq n - 2)$ , along with  $(P[n - 1], P[0])$ . The triangulation  $T$  is constructed via the polygonization  $P$  and set of diagonals  $D$ .

**Algorithm:** SpiralSerpentinePolygonize( $S$ )

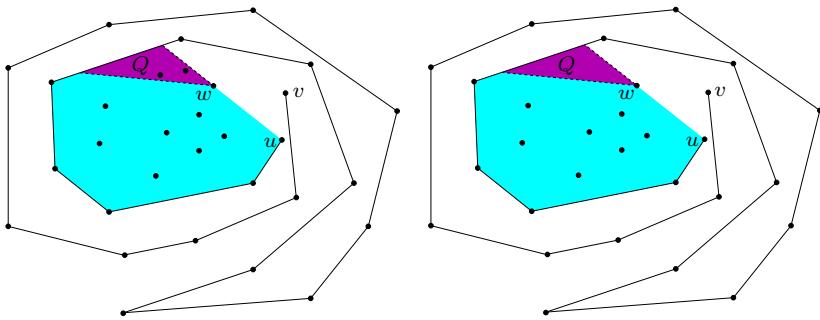
- 1.) Initialize  $H$  to be  $\mathcal{CH}(S)$ , stored in a dynamic (deletion only) convex hull data structure. Initialize the arrays  $C$  and  $R$  for the convex and reflex chains, respectively, and the array of vertex pairs  $D$  for the set of diagonals, to be empty:  $C = R = D = \emptyset$ .
- 2.) Determine  $u$ , the vertex of  $H$  with minimum  $y$ -coordinate (breaking ties by maximizing  $x$ -coordinate) and append it to  $C$ . Let  $v$  to be the next point on  $H$  in the counterclockwise direction and append it to  $C$ . Delete  $u$  and  $v$  from  $H$ .
- 3.) Let  $w$  be the first point encountered by rotating counterclockwise the ray emanating from  $u$  and passing through  $v$  (breaking ties by picking the point closest to  $u$ ). Delete  $w$  from  $H$ .
- 4.) **while**  $H$  is not empty
  - Let  $Q$  be the wedge formed by rays emanating from  $w$  along the directions  $\overrightarrow{uw}$  and  $\overrightarrow{vw}$ .
  - if**  $Q \cap H \neq \emptyset$ 
    - Find the next  $q \in H$  encountered by rotating ccw the ray centered at  $w$  in the direction  $\overrightarrow{uw}$ .
    - Set  $u = w$
    - Set  $w = q$
    - Append  $u$  to  $R$
    - Append  $(u, v)$  to  $D$
    - Delete  $q$  from  $H$
  - else**
    - Find the next  $q \in H$  encountered by rotating ccw the ray centered at  $u$  in the direction  $\overrightarrow{uw}$ .
    - Set  $v = w$
    - Set  $w = q$
    - Append  $v$  to  $C$
    - Append  $(u, v)$  to  $D$
    - Delete  $q$  from  $H$
- end while**
- 5.) Append  $w$  to  $C$ . Let  $P$  be the array obtained by concatenating  $R$  in reverse order to the end of  $C$ , giving a polygonization of  $S$ . The polygonization can be constructed by outputting the edges  $(P[i], P[i + 1])$ ,  $(0 \leq i \leq n - 2)$ , along with  $(P[n - 1], P[0])$ . The triangulation  $T$  is constructed via the polygon  $P$  and diagonals  $D$ .

We now summarize the important invariants of the algorithm.

- a.) Except for the final point appended in step 5.), each point  $w$  is either relabeled  $v$  and appended to the convex chain  $C$  or relabeled  $u$  and appended to the reflex chain  $R$ .

b.) In step 4.),  $w$  (once relabeled  $u$  or  $v$ ) always belongs to one single triangle  $t$  of the current triangulation,  $T_k$ . Therefore,  $t$  is always a leaf node in the dual graph (a path) of  $T_k$ .

c.)  $H$  is always contained in the intersection of left half-planes of edges along the reflex chain in the counterclockwise direction and the directed line  $\overrightarrow{uw}$ . This region is the union of the light shaded region and  $Q$ , as depicted in Figure 3.



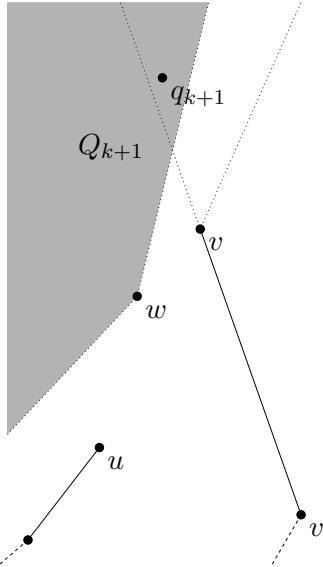
**Fig. 3.** The two possible cases that arise during each iteration of the algorithm’s **while** loop: When the dark shaded region  $Q$  contains at least one point of  $H$  (left) and when it contains no points of  $H$  (right)

### 3 Correctness

**Lemma 1.** *SpiralSerpentinePolygonize constructs a spiral polygon  $P$ .*

*Proof.* The proof is by induction on the iteration counter. After the first iteration of the **while** loop, we have just one triangle, which is trivially spiral. Assume the claim holds after  $k < n$  iterations of the **while** loop and consider the state after the  $(k+1)^{th}$  iteration. We remove the point  $w = q_{k+1}$  that was most recently discovered along with the two edges incident to  $q_k$ , the point discovered on the  $k^{th}$  iteration. Use labels  $u$ ,  $w$ ,  $v$  as assigned at the end of the  $k^{th}$  iteration. The resulting polygon is spiral by the induction hypothesis. We have two cases to consider: when  $q_{k+1}$  is in  $Q_{k+1}$  (the wedge  $Q$  during the  $(k+1)^{th}$  iteration) and when it is not.

We first assume that  $q_{k+1}$  is in  $Q_{k+1}$ . In this case we wish to show that when  $q_{k+1}$  is processed, (a)  $v$  remains a convex vertex, and (b)  $w$  becomes a reflex vertex. Let  $v'$  be the convex vertex adjacent to  $v$  in the clockwise direction. We observe that  $uvq_{k+1}$  must form a left turn (otherwise,  $q_{k+1}$  is not in  $Q_{k+1}$ ). Consider the wedge formed by rays centered at  $v$  along the directions  $\overrightarrow{uv}$  and  $\overrightarrow{v'v}$  (see Figure 4), which was previously examined after the vertex currently labeled  $v$  was inserted into  $C$ . The subsequent point discovered,  $w$ , was not



**Fig. 4.**  $v'vq_{k+1}$  cannot form a right turn

in this wedge and so the wedge is empty. It follows then that  $q_{k+1}$  must be to the left of the directed line  $v'v$  so that  $v$  remains a convex vertex. Since  $q_{k+1}$  is in  $Q_{k+1}$ ,  $uwq_{k+1}$  must form a left turn, which implies that  $w$  becomes a reflex vertex.

Next we suppose that  $q_{k+1}$  is not in  $Q_{k+1}$ . Here, we show that when  $q_{k+1}$  is processed, (a)  $u$  retains its convexity (convex or reflex), and (b)  $w$  becomes convex. Consider  $u$  being reflex in the  $k^{th}$  iteration of the loop. Then on the  $(k+1)^{th}$  iteration the ray  $\overrightarrow{uw}$  is rotated counterclockwise until it encounters  $q_{k+1}$  in  $H$ . Clearly,  $u$ 's interior angle cannot decrease so it remains reflex. Next we consider  $u$  being convex in the  $k^{th}$  iteration. Here,  $u$  represents the point first added to  $C$  in step 2.) of the algorithm. Since  $u \in \mathcal{CH}(S)$ ,  $u$  must remain convex after the  $(k+1)^{th}$  iteration. Finally, since  $q_{k+1}$  is not in  $Q_{k+1}$ ,  $vwq_{k+1}$  necessarily forms a left turn, indicating that  $w$  remains convex.

**Lemma 2.** **SpiralSerpentinePolygonize** constructs a triangulation of  $P$ .

*Proof.* In each iteration of the **while** loop of the algorithm, a triangle is effectively appended to an edge of the polygonization constructed so far. Specifically, a point  $w$  is relabeled as  $u$  or  $v$  and segments (one boundary edge and one diagonal) are attached from this vertex to the most recently added vertices along the convex and reflex chains,  $C$  and  $R$ , respectively. (If  $R$  is empty, the second segment connects to the first point selected in  $S$  in step 2.) By algorithm invariant c.),  $w$  is to the left of all previously constructed edges along the reflex chain in the counterclockwise direction and the directed line  $\overrightarrow{uw}$ . It follows that

the new segments appended do not intersect any other boundary points of the polygonization constructed so far, yielding a valid updated triangulation.

**Lemma 3.** *The triangulation constructed by `SpiralSerpentinePolygonize` is serpentine.*

*Proof.* After the first iteration we have a single triangle, which is trivially serpentine. During each subsequent iteration, the new triangle formed is adjacent to a leaf of the dual path of the current triangulation constructed so far. Hence, the updated triangulation remains serpentine.

Lemmas 1, 2 and 3 yield the desired result, stated in the following theorem:

**Theorem 1.** *`SpiralSerpentinePolygonize` constructs a spiral serpentine polygon.*

Finally, we examine the algorithm's runtime and space usage:

**Theorem 2.** *`SpiralSerpentinePolygonize` runs in  $O(n \log n)$  time and requires  $O(n)$  space.*

*Proof.* The semi-dynamic convex hull data structure requires  $O(n \log n)$  preprocessing time,  $O(\log n)$  amortized deletion time and  $O(n)$  space [9]. During each of the  $n - 3$  iterations of the **while** loop, a point of  $S$  is processed. For each iteration of the while loop, it takes  $O(\log n)$  time to determine if  $Q$  contains a point of  $H$ ,  $O(\log n)$  time to find the next vertex  $q$ , and constant time to update the arrays storing one of the polygonal chains ( $C$  or  $R$ ) and the array of diagonals ( $D$ ). This gives us the desired  $O(n \log n)$  run time, which is optimal. Since the semi-dynamic convex hull structure uses  $O(n)$  space and the input points and the outputted polygon and set of diagonals (for the triangulation) can be stored in arrays of linear size, we require just  $O(n)$  space.

## 4 Implementation and Examples

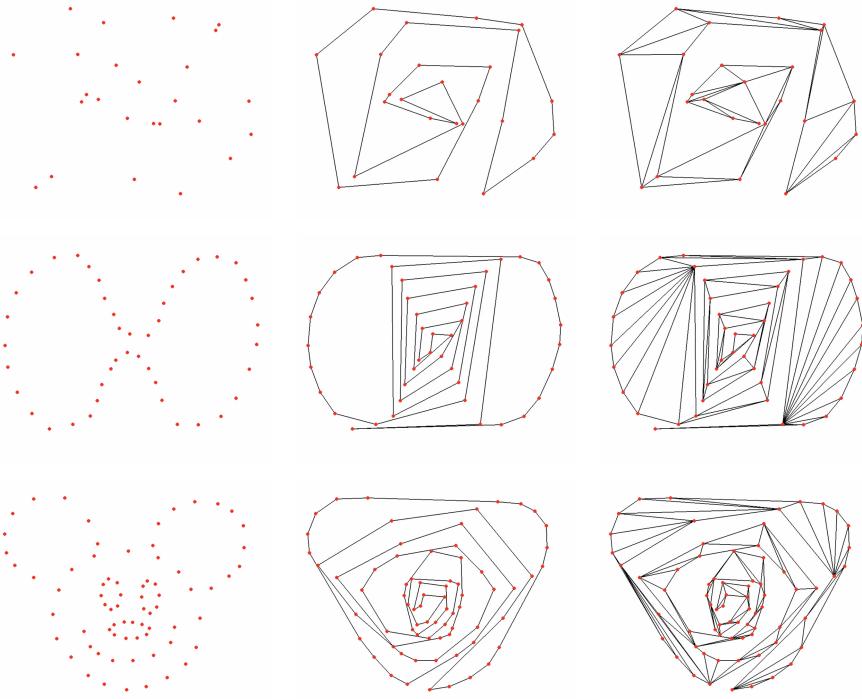
We have implemented `SpiralSerpentinePolygonize` in Java<sup>2</sup>. In Figure 5 we display three different input point sets, along with corresponding spiral polygonizations and serpentine triangulations produced by our algorithm.

In the first row of Figure 5, our input set is a cloud of 25 random points. The second row depicts a “dumbbell shaped” point set, and the last row resembles the face of Mickey Mouse. In each row, we show the point set (left), the spiral polygonization given by the algorithm (middle), and the serpentine triangulation of the spiral polygonization (right). Note that the spiral polygonization is not a good curve reconstruction polygonization; the spiral polygonizations we compute are not indicative of the “shape” of the point sets.

---

<sup>2</sup> An interactive Java applet can be accessed at

<http://www.ams.stonybrook.edu/~jsbm/SpiralSerpentinePolygonize/applet.html>



**Fig. 5.** Left to right in each row: an input point set  $S$ , a spiral polygonization of  $S$  produced, and a serpentine triangulation of the spiral polygonization

## 5 Conclusion

We have shown that every planar point set  $S$  with  $n \geq 3$  points admits a spiral serpentine polygonization, which can be computed with its accompanying triangulation in optimal  $O(n \log n)$  time and  $O(n)$  space.

This work suggests some open questions for further study. Deneen and Shute [5] have investigated the combinatorics of star-shaped polygonizations. An analogous question is how many spiral polygonizations exist among all sets of  $n$  points in the plane? Secondly, what is the complexity of finding a minimum length spiral serpentine polygonization or a minimum weight serpentine triangulation? Finally, in 3D we pose a surface reconstruction problem: what is the minimum number of transitions between *convex patches* and *reflex patches* based on dihedral angles?

**Acknowledgments.** We thank Ferran Hurtado for inspiring us to work on this polygonization problem and for inspiring an entire generation of computational geometers. We thank Estie Arkin, Sándor Fekete, Ferran Hurtado, Marc Noy, Vera Sacristán, and Saurabh Sethia, coauthors on [3], for helpful discussions on this work. We thank Diane Souvaine for suggesting the use of the dynamic

convex hull data structure, which helped simplify the presentation. We also thank an anonymous referee for providing several helpful suggestions. This work is partially supported by the National Science Foundation (CCF-1018388).

## References

1. Abellanas, M., García, J., Hernández-Peñalver, G., Hurtado, F., Serra, O., Urrutia, J.: Onion Polygonizations. *Information Processing Letters* 57, 165–173 (1996)
2. Agarwal, P.K., Hurtado, F., Toussaint, G.T., Trias, J.: On Polyhedra Induced by Point Sets in Space. *Discrete Applied Mathematics* 156, 42–54 (2008)
3. Arkin, E., Fekete, S.P., Hurtado, F., Mitchell, J.S.B., Noy, M., Sacristán, V., Sethia, S.: On the Reflexivity of Point Sets. In: *Discrete and Computational Geometry: The Goodman-Pollack Festschrift, Algorithms and Combinatorics*, vol. 25, pp. 139–156 (2003)
4. Chazelle, B.: On the Convex Layers of a Planar Set. *IEEE Transactions on Information Theory* 31, 509–517 (1985)
5. Deneen, L., Shute, G.: Polygonizations of Point Sets in the Plane. *Discrete and Computational Geometry* 3, 77–87 (1988)
6. Fekete, S.P.: On Simple Polygonizations with Optimal Area. *Discrete and Computational Geometry* 23, 73–110 (2000)
7. Graham, R.L.: An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set. *Information Processing Letters* 1, 132–133 (1972)
8. Grünbaum, B.: Hamiltonian Polygons and Polyhedra. *Geombinatorics* 3, 83–89 (1994)
9. Hershberger, J., Suri, S.: Applications of a Semi-Dynamic Convex Hull Algorithm. *BIT Numerical Mathematics* 32, 249–267 (1992)
10. Iwerks, J., Mitchell, J. S. B.: Spiral Serpentine Polygonization of a Planar Point Set. In: *Proceedings of XIV Spanish Meeting on Computational Geometry*, pp. 181–184 (2011)
11. Newborn, M., Moser, W.O.J.: Optimal Crossing-Free Hamiltonian Circuit Drawings of  $K_n$ . *Journal of Combinatorial Theory, Series B* 29, 13–26 (1980)
12. Quintas, L.V., Supnick, F.: On Some Properties of Shortest Hamiltonian Circuits. *American Mathematical Monthly* 72, 977–980 (1965)
13. Shamos, M.I., Hoey, D.: Closest-Point Problems. In: *Proceedings of the Sixteenth IEEE Symposium on Foundations of Computer Science*, pp. 151–162 (1975)
14. Steinhaus, H.: *One Hundred Problems in Elementary Mathematics*. Dover Publications, Inc., New York (1964)

# The 1-Center and 1-Highway Problem

José Miguel Díaz-Báñez<sup>1,\*</sup>, Matias Korman<sup>2,\*\*</sup>, Pablo Pérez-Lantero<sup>3,\*\*\*</sup>,  
and Inmaculada Ventura<sup>1,\*</sup>

<sup>1</sup> Departamento de Matemática Aplicada II, Universidad de Sevilla, Spain  
`{dbanez,iventura}@us.es`

<sup>2</sup> Universitat Politècnica de Catalunya (UPC), Barcelona, Spain  
`matias.korman@upc.edu`

<sup>3</sup> Escuela de Ingeniería Civil en Informática, Universidad de Valparaíso, Chile  
`pablo.perez@uv.cl`

**Abstract.** In this paper we extend the RECTILINEAR 1-CENTER problem as follows: Given a set  $S$  of  $n$  points in the plane, we are interested in locating a facility point  $f$  and a rapid transit line (highway)  $H$  that together minimize the expression  $\max_{p \in S} d_H(p, f)$ , where  $d_H(p, f)$  is the travel time between  $p$  and  $f$ . A point  $p \in S$  uses  $H$  to reach  $f$  if  $H$  saves time for  $p$ . We solve the problem in  $O(n^2)$  or  $O(n \log n)$  time, depending on whether or not the highway's length is fixed.

**Keywords:** Geometric optimization, Facility location, Time metric.

## 1 Introduction

The optimal location of a facility modeled as a geometric object is a well-studied problem in both operations research and computational geometry. Plastria [25] gives an overview on continuous location, and Díaz-Báñez et al. [15] survey extensive facility location models. Particularly, the geometrical nature of problems under the minmax criterion has led to a fruitful interaction between the above fields [17, 26, 27].

On the other hand, models dealing with alternative transportation systems have been suggested in location theory [7, 11, 24]. Although the metric given by a real urban transportation system is often quite complicated, simplified mathematical models have been widely studied in order to investigate basic geometric properties of urban transportation systems. Abellanas et al. [1] considered a geometric modeling of this environment: they represent highways as polygonal

---

\* J.M. D.-B., P. P.-L. and I. V. were partially supported by project FEDER MEC MTM2009-08652. J.M. D.-B., M.K and I. V. were partially supported by ESF EU-ROCORES programme EuroGIGA, CRP ComPoSe: grant EUI-EURC-2011-4306.

\*\* Partially supported by the support of the Secretary for Universities and Research of the Ministry of Economy and Knowledge of the Government of Catalonia and the European Union.

\*\*\* Partially supported by grant FONDECYT 11110069.

chains consisting of line segments in the plane, giving each line segment an associated speed. Then, the travel time between two points gives a metric called the *time distance*. Recently, there has been an interest in facility location problems derived from urban modeling. In many cases one is interested in locating a highway that optimizes some given function that depends on the distance between elements of a given point set [2–4, 10, 23].

In a recent paper, Espejo and Rodríguez-Chía [19] introduced a variant of this kind of problem in which one is given a set  $S$  of clients (represented by points) located in a city, and then is interested in locating a service facility (represented by a point) and a highway (represented by a straight segment of fixed length and any orientation) simultaneously in a way that the average supply time between the clients and the facility point is minimized. In this model clients enter and exit the highway at the endpoints only, and move through it with a given constant speed. They move at unit speed elsewhere under the  $L_1$ -metric. Unfortunately, there was an error in their approach, and their algorithm gives incorrect solutions in some cases (a correct algorithm was given in [12]). In this paper we study a variation of this problem in which, instead of the average travel time, we want to minimize the largest travel time between the clients and the facility.

Similarly to many other related problems [12, 19], in this paper we use  $L_1$  as the underlying metric. The Euclidean metric is in some cases hard to deal with [6], and using the  $L_1$ -metric instead is a logical simplification. Indeed, the  $L_1$ -metric is much simpler and gives constant-factor approximation when compared to the Euclidean metric. Moreover, it works very well in practice.

Since we are using  $L_1$  as the underlying metric, the natural objective could be to locate a horizontal or vertical highway. However, in many instances a highway of this type gives much less profit with respect to a highway with arbitrary orientation. This is a reality in most cities where, although most streets and avenues are either parallel or orthogonal, a few exceptions occur (like Broadway avenue in New York, or Diagonal avenue in Barcelona).

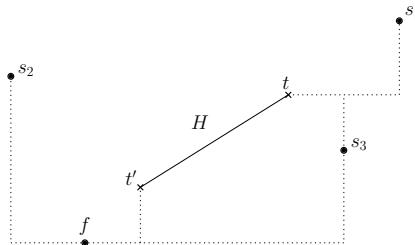
As we will show later, our problem is closely related to the well-known 2-CENTER problem: Given a set of  $n$  points in the plane representing customers, the goal is to locate two facilities that minimize the largest distance from a customer to its nearest facility. This problem received much attention in the 1990's; the currently best known algorithm for the EUCLIDEAN 2-CENTER problem is due to Eppstein [18]; Drezner [16] found a linear-time algorithm for the RECTILINEAR 2-CENTER problem; and the discrete version was considered by Bespamyatnikh and Segal [9]. However, the objective function in our problem is different, and it seems that we cannot directly apply these results to our case (see Section 4).

In our model we only allow entering and leaving the highway at only its endpoints (in other words this kind of highway is called *walkway* [10] or *turnpike* [5]). We note that there exists another model for highways (called *freeways* [5] or simply *highways* [1]) in which one is allowed to enter and leave at any point. Hence, a natural variant would be to consider the simultaneous location of a facility and a freeway (instead of a turnpike). The location of freeways under the min-sum

criterion was studied in [13] and under the min-max criterion can be found in the extended version of this paper [14].

## 2 Problem Definitions and Notation

Let  $S$  be the set of  $n$  client points,  $f$  be the service facility point, and  $H$  be the highway, represented by a straight segment with any orientation and endpoints  $t$  and  $t'$ . Given a point  $u$  of the plane, let  $x(u)$  and  $y(u)$  denote the  $x$ - and  $y$ -coordinate of  $u$ , respectively. For simplicity in the explanation, we assume that no two points of  $S$  share an  $x$ - or  $y$ -coordinate, but we note that our algorithms also work for the case in which points do not have this general position assumption. Let  $\ell$  be the length of  $H$  and let  $v > 1$  be its speed.



**Fig. 1.** The distance model; in the example  $s_1$  uses the highway from  $t$  to  $t'$  in order to reach  $f$  faster. The highway does not speed up transportation between  $s_2$  and  $f$ , hence is not used by  $s_2$ . Client point  $s_3$  however, can either walk or use the highway to reach  $f$ , and will need the same time in both cases. Observe that, since we are interested in paths that reach  $f$ , the highway will only be used in one direction.

If both the location of the facility and the endpoints of the highway are fixed, the distance from a client point  $p \in S$  to  $f$  is defined as  $d_H(p, f) := \min\{\|p - f\|_1, \|p - t\|_1 + \ell/v + \|t' - f\|_1, \|p - t'\|_1 + \ell/v + \|t - f\|_1\}$ , where  $\|\cdot\|_1$  represents the  $L_1$ -norm, see Figure 1. That is, the distance between  $p$  and  $f$  is the minimum time between either walking, using the highway in one direction or using it in the reverse direction, see Figure 1. Whenever  $d_H(p, f) \leq \|p - f\|_1$ , we say that  $p$  uses the highway to reach  $f$ . Otherwise, we say that  $p$  walks (or does not use  $H$ ) to reach  $f$ .

The problem that we study can be formulated as follows:

**The 1-CENTER AND 1-HIGHWAY problem (1C1H):** Given a set  $S$  of  $n$  points and a fixed speed  $v > 1$ , locate a point (facility)  $f$  and a line segment (highway)  $H$  with endpoints  $t$  and  $t'$  such that the function  $\max_{p \in S} d_H(p, f)$  is minimized.

The case in which the highway's length  $\ell = \|t - t'\|_2$  is fixed is called the FIXED LENGTH 1-CENTER AND 1-HIGHWAY problem (FL-1C1H for short). The case in which the highway can have any length is called the VARIABLE LENGTH 1-CENTER AND 1-HIGHWAY problem (or VL-1C1H for short).

It is easy to see that in either variant of the 1C1H problem, the highway will only be used in one direction. By using similar arguments to those given in [19, Lemma 2.1] we can prove that there always exists an optimal location in which one of the highway's endpoints coincides with the facility. Therefore, we assume throughout the paper that  $f = t'$ , thus the distance from a client point  $p \in S$  to  $f$  simplifies to  $d_H(p, f) := \min\{\|p - f\|_1, \|p - t\|_1 + \ell/v\}$ .

### 3 Solving the 1C1H Problem

In this section, we give a general algorithm to solve both variants of the 1C1H problem. We will then propose an improved method for the VL-1C1H problem. Using the standard transformation from the  $L_1$ -metric to the  $L_\infty$ -metric, we solve the problem using the  $L_\infty$ -metric instead.

Let  $(f^*, H^*, t^*)$  denote an optimal solution of a given problem instance, where  $f^*$  is the facility point,  $H^*$  is the highway, and  $t^*$  is the endpoint of  $H^*$  other than  $f^*$ . Let  $R^* = \max_{p \in S} d_{H^*}(p, f^*)$ . Let  $r_1$  be the maximum of  $d_{H^*}(p, f^*)$  among all points  $p \in S$  not using  $H^*$ , and  $r_2$  be the maximum of  $d_{H^*}(p, f^*) - \ell/v$  among all points  $p \in S$  that use  $H^*$ . Let  $B(u, r)$  denote the axis-parallel ball of radius  $r$  centered at  $u$ , that is, the ball of center  $u$  and radius  $r$  with respect to the  $L_\infty$ -metric. Note that  $S$  is covered by the balls  $B(f^*, r_1)$  and  $B(t^*, r_2)$ , and  $R^* = \max\{r_1, r_2 + \ell/v\}$  is satisfied. Furthermore, either  $r_1$  or  $r_2$  can be increased, without affecting the value  $R^*$  of the solution, so that  $r_1 = r_2 + \ell/v$ . From these observations, the following statement can be obtained:

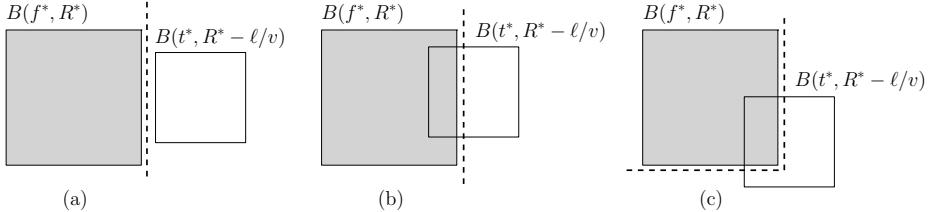
**Statement 1.** *The 1C1H problem is equivalent to finding two balls,  $B(f, R)$  and  $B(t, R - \ell/v)$ , such that  $R$  is minimum and  $B(f, R) \cup B(t, R - \ell/v)$  covers  $S$ .*

Consider again an optimal solution  $(f^*, H^*, t^*)$ . We partition  $S$  into two sets  $S_1^*$  and  $S_2^*$  as follows: set  $S_1^*$  contains the points whose  $L_\infty$ -distance to  $f^*$  is at most  $R^*$ ; and set  $S_2^* = S \setminus S_1^*$  contains the points that must use the highway to reach  $f^*$  in  $R^*$  or less units of time. Observe that we cannot have  $S_1^* = \emptyset$  since, by reversing the positions of  $f^*$  and  $t^*$ , we would obtain a better solution. Note that set  $S_2^*$  can be empty (for example, if we are forced to locate an extremely long highway). However, this case can be easily detected, since then  $f^*$  is the solution of the RECTILINEAR 1-CENTER problem (which can be computed in linear time [16]). Hence, from now on we assume that neither  $S_1^*$  nor  $S_2^*$  is empty for any optimal solution  $(f^*, H^*, t^*)$ .

We now state what we call the BASE problem: Given a partition  $\{S_1, S_2\}$  of  $S$ , find the coordinates of  $f$  and  $t$  and the smallest value  $R$  (called the *radius* of the partition) such that  $S_1 \subseteq B(f, R)$  and  $S_2 \subseteq B(t, R - \ell/v)$ . When considered the FL-1C1H problem, we add the constraint that  $\|f - t\|_2 = \ell$  is fixed.

Given that  $(f^*, H^*, t^*)$  is an optimal solution, it is easy to see that  $f^*$  and  $t^*$ , joint with  $R^*$ , form a solution of the BASE problem for the partition  $\{S_1^*, S_2^*\}$ . Moreover, the radius of any other partition of  $S$  will have radius greater than or equal to  $R^*$ .

Our algorithm works as follows: We consider different partitions of  $S$  and solve the BASE problem associated with each of them. We identify  $\{S_1^*, S_2^*\}$  as the partition whose radius is smallest. A naive method would be to guess the partition  $\{S_1^*, S_2^*\}$  among the  $O(2^n)$  candidates. In the following lemma we reduce the search space of partitions to one of polynomial size.



**Fig. 2.** Relative positions of balls  $B(f^*, R^*)$  and  $B(t^*, R^* - \ell/v)$ . For each of the cases, set  $S_1^*$  (the one covered by  $B(f^*, R^*)$ ) can be separated from  $S_2^*$  with either an axis-aligned line or an upper-left quadrant.

**Lemma 1.** *For any set  $S$  of  $n$  points, an optimal partition  $\{S_1^*, S_2^*\}$  of  $S$  can be found among  $O(n^2)$  candidates.*

*Proof.* Without loss of generality, we can assume that  $f^*$  is above and to the left of  $t^*$ . It is then easy to see that there are three possible relative positions of the two balls of Statement 1 (see Figure 2). In each of these cases the two sets can be separated by either an axis-aligned line or an upper-left quadrant (see dashed lines in Figure 2). Each possible partition is uniquely determined by the number of points above and/or to the left of the separating line/quadrant. In particular, there are  $O(n^2)$  different cases, hence the lemma follows.  $\square$

### 3.1 The General Approach via the BASE problem

Given a set  $P$  of points, let  $X(P) \subseteq P$  be the set containing the points with highest or lowest  $x$ - or  $y$ -coordinate of  $P$ . This set is called the set of *extreme points* of  $P$ . We define  $\delta(P)$  to be half of the largest  $L_\infty$ -distance between any two points of  $P$ . Note that  $\delta(P)$  is the minimum radius  $r$  needed such that all points of  $P$  can be enclosed in some  $L_\infty$ -ball of radius  $r$ . Observe that  $|X(P)| \leq 4$  and  $\delta(P) = \delta(X(P))$ . For any  $r \geq 0$ , let  $\text{Center}(P, r)$  be the locus of the points  $u$  such that  $P \subseteq B(u, r)$ , that is,  $\text{Center}(P, r) = \bigcap_{p \in P} B(p, r)$ . A similar definition has been used by Huang et al. [21].

**Lemma 2.** *For any set  $P$  of points the following properties hold:*

1.  $\text{Center}(P, r)$  can be empty, a point, an axis-parallel segment, or an axis-parallel rectangle.
2.  $\text{Center}(P, r) = \emptyset$  if and only if  $r < \delta(P)$ .

3.  $\text{Center}(P, r) = \text{Center}(X(P), r)$ . That is, the set  $\text{Center}(P, r)$  depends only on the coordinates of the extreme points of  $P$ .
4. For any  $\varepsilon > 0$  and  $r \geq \delta(P)$ ,  $\text{Center}(P, r) \subset \text{Center}(P, r + \varepsilon)$ . Moreover, the separation between the boundaries of  $\text{Center}(P, r)$  and  $\text{Center}(P, r + \varepsilon)$  is equal to  $\varepsilon$ .

All these properties are easy to prove from the definition of  $\text{Center}(P, r)$  and the linearity of the  $L_\infty$ -metric, thus we omit them. With these observations we can now solve the BASE problem efficiently.

**Lemma 3.** *Let  $\{S_1, S_2\}$  be a partition of  $S$ . If we are given the sets  $X(S_1)$  and  $X(S_2)$ , then the BASE problem can be solved in constant time; for both the FL-1C1H problem and the VL-1C1H problem.*

*Proof.* We start by giving an algorithm for the case in which we consider the FL-1C1H problem. Let  $\{S_1, S_2\}$  be a partition of  $S$  and  $R$  be its radius. Observe that we always have  $R \geq \max\{\delta(S_1), \delta(S_2) + \ell/v\}$  (otherwise an extreme point of either  $S_1$  or  $S_2$  will not be able to reach  $f$  in  $R$  units of time). If there exist two points  $f \in \text{Center}(S_1, \max\{\delta(S_1), \delta(S_2) + \ell/v\})$  and  $t \in \text{Center}(S_2, \max\{\delta(S_1), \delta(S_2) + \ell/v\})$  such that  $\|f - t\|_2 = \ell$ , then we are done.

Unfortunately, this does not always happen. In general, using Statement 1, we must find two values  $\varepsilon_1, \varepsilon_2 \geq 0$  such that: (i)  $\delta(S_1) + \varepsilon_1 = \delta(S_2) + \varepsilon_2 + \ell/v$ , (ii) there are points  $f \in \text{Center}(S_1, \delta(S_1) + \varepsilon_1)$  and  $t \in \text{Center}(S_2, \delta(S_2) + \varepsilon_2)$  satisfying  $\|f - t\|_2 = \ell$ , and (iii)  $\delta(S_1) + \varepsilon_1 = \delta(S_2) + \varepsilon_2 + \ell/v$  is minimized. The values  $\varepsilon_1$  and  $\varepsilon_2$  can be found in constant time as follows:

First, set  $\alpha_1 = \max\{0, \delta(S_2) + \ell/v - \delta(S_1)\}$  and  $\alpha_2 = \max\{0, \delta(S_1) - \delta(S_2) - \ell/v\}$ . That is, we increase either  $\delta(S_1)$  to  $\delta(S_2) + \ell/v$  in  $\alpha_1$  units or  $\delta(S_2)$  to  $\delta(S_1) - \ell/v$  in  $\alpha_2$  units. We now look for the smallest value  $x \geq 0$  such that there are points  $f \in \text{Center}(S_1, \delta(S_1) + \alpha_1 + x)$  and  $t \in \text{Center}(S_2, \delta(S_2) + \alpha_2 + x)$  satisfying  $\|f - t\|_2 = \ell$ . Note that this problem is of constant size and can be solved in  $O(1)$  time. Then,  $\varepsilon_1 := \alpha_1 + x$  and  $\varepsilon_2 := \alpha_2 + x$  satisfy conditions (i)–(iii).

We observe that this computation of  $\varepsilon_1$  and  $\varepsilon_2$  is correct. In fact, let  $\varepsilon_1, \varepsilon_2$  satisfy conditions (i)–(iii). Suppose  $\varepsilon_1 \geq \varepsilon_2$ . Then we have  $\delta(S_1) + (\varepsilon_1 - \varepsilon_2) = \delta(S_2) + \ell/v$ . By setting  $\alpha_1 := \varepsilon_1 - \varepsilon_2$ ,  $\alpha_2 := 0$ , and  $x := \varepsilon_2$  the correctness follows. The case where  $\varepsilon_1 < \varepsilon_2$  is analogous.

If we consider the FL-1C1H problem, then solving the BASE problem is slightly different. The main difference is that  $\varepsilon_1$  and  $\varepsilon_2$  must now minimize the expression  $\delta(S_1) + \varepsilon_1 = \delta(S_2) + \varepsilon_2 + g(\varepsilon_1, \varepsilon_2)/v$ , where  $g(\varepsilon_1, \varepsilon_2)$  denotes the smallest Euclidean distance between a point  $f \in \text{Center}(S_1, \delta(S_1) + \varepsilon_1)$  and a point  $t \in \text{Center}(S_2, \delta(S_2) + \varepsilon_2)$ . As before, this problem has constant size and thus can be solved in  $O(1)$  time.  $\square$

By combining the above results, we obtain a method to solve both problems:

**Theorem 2.** *Both variants of the 1C1H problem can be solved in  $O(n^2)$  time and  $O(n)$  space.*

*Proof.* Recall that, by Lemma 1, we can split set  $S$  into sets  $S_1^*$  and  $S_2^*$  by using either a vertical line or an upper-left quadrant. We start by considering first the case in which there is a vertical splitting line (cases (a) and (b) in Figure 2). Sort the points of  $S$  in increasing value of  $x$ -coordinates; let  $p_1, p_2, \dots, p_n$  be the obtained order. For any  $i \in [1..n - 1]$ , let  $S_i$  denote the set  $\{p_1, \dots, p_i\}$ . By sweeping  $S$  from left to right with a vertical, we can compute and store the set of extreme points  $X(S_i)$  for all  $i \in [1..n - 1]$  in  $O(n)$  time. Analogously, we can sweep  $S$  from right to left and compute the set  $X(S \setminus S_i)$  in linear time as well. Then, by using Lemma 3, we solve the BASE problem for each pair  $(S_i, S \setminus S_i)$ ,  $1 \leq i < n$ . The running time is dominated by the initial sorting of the points of  $S$ , in it is thus  $O(n \log n)$ .

In order to complete the proof, it remains to show how case (c) in Figure 2 can be solved in  $O(n^2)$  time and  $O(n)$  space. Let  $q_1, q_2, \dots, q_n$  be the points of  $S$  sorted in decreasing order of  $y$ -coordinates. For any  $i, j \in [1..n - 1]$ , let  $S_{i,j} := \{u \in S \mid y(u) \geq y(q_i) \wedge x(u) \leq x(p_j)\}$  and  $S'_{i,j} := \{u \in S \mid y(u) < y(q_i) \wedge x(u) \leq x(p_j)\}$ . For  $i = 1 \dots n - 1$ , we proceed as in the previous case: We first sweep  $S$  from left to right computing  $X(S_{i,j})$  and  $X(S'_{i,j})$  for all  $j \in [1..n - 1]$  in  $O(n)$  time. Using the sets  $X(S'_{i,j})$ ,  $j \in [1..n - 1]$ , previously computed, we sweep  $S$  from right to left computing  $X(S \setminus S_{i,j})$  for all  $j \in [1..n - 1]$  in  $O(n)$  time. Considering the above computation, and using Lemma 3, we then solve the  $O(n)$  instances  $(S_{i,j}, S \setminus S_{i,j})$ ,  $j \in [1..n - 1]$ , of the BASE problem in constant time each. We keep track, for each value of  $i$ , of a partition  $\{S_{i,j}, S \setminus S_{i,j}\}$  of smallest radius. Observe that  $O(n)$  space is used only.  $\square$

### 3.2 Locating a Highway of Variable Length

Using Theorem 2 we have an algorithm that runs in  $O(n^2)$  time for both the FL-1C1H problem and the VL-1C1H problem. The bottleneck of the algorithm is case (c) of Lemma 1. In the following we show how to treat this case more efficiently for the VL-1C1H problem.

Let  $p_N$  and  $p_S$  be the points with highest and lowest  $y$ -coordinate of  $S$ , respectively. Analogously, let  $p_E$  and  $p_W$  be defined with respect to the  $x$ -coordinates. Then  $X(S) = \{p_N, p_S, p_E, p_W\}$ . Without loss of generality, we assume  $\delta(S) = \frac{1}{2}(x(p_E) - x(p_W))$ . We further assume  $x(f^*) > x(t^*)$  and  $y(f^*) > y(t^*)$  (that is, the angle that the highway forms with the  $x$ -axis is between  $-\pi/2$  and  $0$ ).

**Lemma 4.** *If in every optimal solution of the VL-1C1H problem the two balls of Statement 1 satisfy that each of them contains a corner of the other one, then there exists an optimal solution  $(f^*, H^*, t^*)$  of the VL-1C1H problem in which the extreme points  $X(S)$  lie on the boundary of  $B(f^*, R^*) \cup B(t^*, R^* - \ell/v)$ .*

*Proof.* Observe that if any of the balls contains both  $p_N$  and  $p_S$  (or  $p_W$  and  $p_E$ ), the sets  $S_1^*$  and  $S_2^*$  can be separated by either a horizontal or a vertical line. Since we have assumed that this is not possible, the ball that contains  $p_N$  cannot contain  $p_S$  (analogously, the ball that contains  $p_W$  cannot contain  $p_E$ ). Without loss of generality, we assume that  $p_N, p_W \in B(f^*, R^*)$  and  $p_S, p_E \in$

$B(t^*, R^* - \ell/v)$  for any optimal solution  $(f^*, H^*, t^*)$ . Observe that this implies both  $x(f^*) < x(t^*)$  and  $y(f^*) > y(t^*)$ .

Assume that point  $p_N$  is not on the boundary of  $B(f^*, R^*)$ . In the following, we show how to perform a local perturbation to the solution so that  $p_N$  ends up on the boundary of  $B(f^*, R^*)$ . We translate  $f^*$  downwards continuously while keeping  $t^*$  unchanged until  $p_N$  reaches the top boundary of  $B(f^*, R^*)$ . Observe that, since initially  $f^*$  has a larger  $y$ -coordinate than  $t^*$ , the translation reduces the distance between  $f^*$  and  $t^*$  until both points have the same  $y$ -coordinate. Observe, however, that this cannot happen since otherwise we can split  $S$  into  $S_1^*$  and  $S_2^*$  with a vertical line. Moreover, no point of  $B(f^*, R^*)$  can leave the ball before  $p_N$  reaches the top boundary. Hence, optimality is preserved through this translation. Analogously, we can perform the same operation on the other extreme points and either obtain that all extreme points lie on the boundary of  $B(f^*, R^*) \cup B(t^*, R^* - \ell/v)$  or find a way to split  $S$  into  $S_1^*$  and  $S_2^*$  with an axis-aligned line.  $\square$

With this observation we can speed up the algorithm for the VL-1C1H problem:

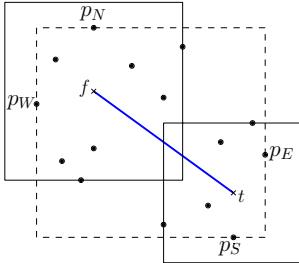
**Theorem 3.** *The VL-1C1H problem can be solved in  $O(n \log n)$  time and  $O(n)$  space.*

*Proof.* By Lemma 4 there is an optimal solution  $(f^*, H^*, t^*)$  such that either there exists an axis-parallel line that splits  $S$  into the sets  $S_1^*$  and  $S_2^*$  or all extreme points of  $S$  lie on the boundary of  $B(f^*, R^*) \cup B(t^*, R^* - \ell/v)$ . The former case can be treated in  $O(n \log n)$  time by using the same approach as in Theorem 2. Hence we focus on the latter case.

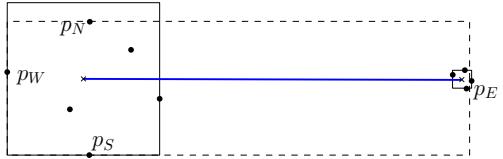
Without loss of generality, we assume that  $x(f^*) < x(t^*)$ , and  $y(f^*) > y(t^*)$ . In particular, this implies that  $p_N, p_W \in B(f^*, R^*)$  and  $p_E, p_S \in B(t^*, R^* - \ell/v)$ . Denote by  $u = (x(p_W), y(p_N))$  the top-left corner of the smallest enclosing axis-aligned rectangle of  $S$ . By Lemma 4, this point must also be the top-left corner of  $B(f^*, R^*)$ . Let  $r_1, r_2, \dots, r_n$  be the elements of  $S$  sorted in increasing order of  $L_\infty$ -distance from  $u$ . Now apply the same approach as in the linearly separable case of Theorem 2 using this new ordering instead. The result follows.  $\square$

## 4 Conclusions and Further Research

In this paper we have considered a facility location problem in which one is interested in locating a supply center point and a highway simultaneously. The highway (represented by a line segment with arbitrary orientation) is an alternative transportation system that can be used by the clients to reduce their travel time to the facility. Two variants were studied, depending on whether or not the length of the segment is fixed in advance. Observe that, in particular, our algorithms can also be used to solve the RECTILINEAR 2-CENTER problem subject to the additional constraint that the distance between the two centers is fixed. It suffices to run our algorithm setting the highway's speed as infinite. Hence, both balls of Statement 1 will have the same radius, and both problems become equivalent.



(a) The optimal solution for the FL-1C1H problem.



(b) The optimal solution for the VL-1C1H problem.

**Fig. 3.** (a) Neither of the two balls shares diametral vertices with the smallest enclosing axis-parallel rectangle of the client points (the bounding box), and no extreme point lies on the boundary of their union. (b) The two optimal balls do not contain two diametral vertices of the bounding box.

Our problem is closely related to the RECTILINEAR 2-CENTER problem. Since this one can be solved in linear time [8, 16], one could think that similar approaches should work for the 1C1H-problem. However, let us notice two observations. First, we recall that the crucial property for solving the RECTILINEAR 2-CENTER problem in linear time (Lemma 1 in [8]) is the fact that there always exist two optimal balls containing two diametral vertices of the bounding box of  $S$ . Unfortunately, this fact is not true for neither variant of the 1C1H problem (see Figure 3(a) and Figure 3(b) for the fixed and variable length case, respectively). Second, it is not difficult to show, by using an example similar to the one in Figure 3(a), that the problem is not *LP*-type. Thus, it seems unlikely that the usual approaches that solve the RECTILINEAR 2-CENTER problem can be adapted for the 1C1H-problem.

Even if linear were not possible, we would like to know if the FL-1C1H problem can be solved in  $o(n^2)$  time. The fast algorithm for the VL-1C1H problem is based on Lemma 4. Unfortunately, it is easy to give examples in which this result does not hold whenever the highway's length is fixed (see Figure 3(a)). Again, extending the same approach for the fixed-length case seems difficult. Finding lower bounds for either problem would also be worth obtaining.

A natural extension of the problems studied in this paper is to consider the general  $p$ -Center and  $k$ -Highway problem. Although it is easy to see that the problem is NP-Hard when either  $p$  or  $k$  is part of the input (even when the other parameter is set to zero) [20, 22], it would be interesting to find efficient algorithms for small values of  $p$  or  $k$ .

## References

1. Abellanas, M., Hurtado, F., Icking, C., Klein, R., Langetepe, E., Ma, L., Palop, B., Sacristán, V.: Voronoi diagram for services neighboring a highway. Information Processing Letters 86(5), 283–288 (2003)

2. Ahn, H.-K., Alt, H., Asano, T., Bae, S.W., Brass, P., Cheong, O., Knauer, C., Na, H.-S., Shin, C.-S., Wolff, A.: Constructing optimal highways. *Int. J. Found. Comput. Sci.* 20(1), 3–23 (2009)
3. Aichholzer, O., Aurenhammer, F., Palop, B.: Quickest paths, straight skeletons, and the city Voronoi diagram. *Discrete & Computational Geometry* 31, 17–35 (2004)
4. Aloupis, G., Cardinal, J., Collette, S., Hurtado, F., Langerman, S., O'Rourke, J., Palop, B.: Highway hull revisited. *Computational Geometry: Theory and Applications* 43, 115–130 (2010)
5. Bae, S.W., Korman, M., Tokuyama, T.: All Farthest Neighbors in the Presence of Highways and Obstacles. In: Das, S., Uehara, R. (eds.) WALCOM 2009. LNCS, vol. 5431, pp. 71–82. Springer, Heidelberg (2009)
6. Bajaj, C.: Proving geometric algorithm non-solvability: An application of factoring polynomials. *J. Symb. Comput.* 2(1), 99–102 (1986)
7. Battaudatta, R., Rajan, S.P.: Mixed planar/network facility location problems. *Computers & Operations Research* 15, 61–67 (1988)
8. Bespamyatnikh, S., Kirkpatrick, D.: Rectilinear 2-center Problems. In: Proc. 11th Canadian Conference on Computational Geometry, pp. 68–71 (1999)
9. Bespamyatnikh, S., Segal, M.: Rectilinear Static and Dynamic Discrete 2-Center Problems. In: Dehne, F., Gupta, A., Sack, J.-R., Tamassia, R. (eds.) WADS 1999. LNCS, vol. 1663, pp. 276–287. Springer, Heidelberg (1999)
10. Cardinal, J., Collette, S., Hurtado, F., Langerman, S., Palop, B.: Optimal location of transportation devices. *Computational Geometry: Theory and Applications* 41, 219–229 (2008)
11. Carrizosa, E., Rodríguez-Chía, A.M.: Weber problems with alternative transportation systems. *European Journal of Operational Research* 97(1), 87–93 (1997)
12. Díaz-Báñez, J.-M., Korman, M., Pérez-Lantero, P., Ventura, I.: Locating a service facility and a rapid transit line. *CoRR*, abs/1104.0753 (2011)
13. Díaz-Báñez, J.-M., Korman, M., Pérez-Lantero, P., Ventura, I.: Locating a single facility and a high-speed line. *CoRR*, abs/1205.1556 (2012)
14. Díaz-Báñez, J.-M., Korman, M., Pérez-Lantero, P., Ventura, I.: The 1-Center and 1-Highway problem. *CoRR*, abs/1205.1882 (2012)
15. Díaz-Báñez, J.M., Mesa, J.A., Schobel, A.: Continuous location of dimensional structures. *European Journal of Operational Research* 152, 22–44 (2002)
16. Drezner, Z.: On the rectangular  $p$ -center problem. *Naval Research Logistics (NRL)* 34(2), 229–234 (1987)
17. Drezner, Z., Hamacher, H.W.: Facility location – applications and theory. Springer (2002)
18. Eppstein, D.: Faster construction of planar two-centers. In: Proc. of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 1997, pp. 131–138 (1997)
19. Espejo, I., Rodríguez-Chía, A.M.: Simultaneous location of a service facility and a rapid transit line. *Computers and Operations Research* 38, 525–538 (2011)
20. Feder, T., Greene, D.: Optimal algorithms for approximate clustering. In: Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing, STOC 1988, pp. 434–444. ACM, New York (1988)
21. Huang, P.-H., Tsai, Y.T., Tang, C.Y.: A fast algorithm for the alpha-connected two-center decision problem. *Inf. Process. Lett.* 85(4), 205–210 (2003)
22. Korman, M.: Theory and Applications of Geometric Optimization Problems in Rectilinear Metric Spaces. PhD thesis, Tohoku University (2009); Committee: Nishizeki, T., Shinohara, A., Shioura, A., Tokuyama, T.

23. Korman, M., Tokuyama, T.: Optimal Insertion of a Segment Highway in a City Metric. In: Hu, X., Wang, J. (eds.) COCOON 2008. LNCS, vol. 5092, pp. 611–620. Springer, Heidelberg (2008)
24. Laporte, G., Mesa, J.A., Ortega, F.A.: Optimization methods for the planning of rapid transit systems. European Journal of Operational Research 122, 1–10 (2000)
25. Plastria, F.: Continuous location problems. In: Drezner, Z. (ed.) Facility Location: A Survey of Applications and Methods, pp. 225–262. Springer (1995)
26. Robert, J.-M., Toussaint, G.: Computational Geometry and Facility Location. In: Proc. Inter. Conference on Operations Research and Management Science, pp. 11–15 (1990)
27. Woeginger, G.J.: A comment on a minmax location problem. Operations Research Letters 23(1-2), 41–43 (1998)

# Compact Grid Representation of Graphs<sup>\*</sup>

José Cáceres<sup>1</sup>, Carmen Cortés<sup>2</sup>, Clara Isabel Grima<sup>2</sup>, Masahiro Hachimori<sup>3</sup>,  
Alberto Márquez<sup>2</sup>, Raiji Mukae<sup>4</sup>, Atsuhiro Nakamoto<sup>4</sup>, Seiya Negami<sup>4</sup>,  
Rafael Robles<sup>2</sup>, and Jesús Valenzuela<sup>2</sup>

<sup>1</sup> Universidad de Almería, Spain  
[jcaceres@ual.es](mailto:jcaceres@ual.es)

<sup>2</sup> Universidad de Sevilla, Spain  
[{ccortes,grima,almar,rafarob,jesusv}@us.es](mailto:{ccortes,grima,almar,rafarob,jesusv}@us.es)

<sup>3</sup> Tsukuba University, Japan  
[hachi@sk.tsukuba.ac.jp](mailto:hachi@sk.tsukuba.ac.jp)

<sup>4</sup> Yokohama National University, Japan  
[{mkerij,nakamoto,negamij}@edhs.ynu.ac.jp](mailto:{mkerij,nakamoto,negamij}@edhs.ynu.ac.jp)

**Abstract.** A graph  $G$  is said to be *grid locatable* if it admits a representation such that vertices are mapped to grid points and edges to line segments that avoid grid points but the extremes. Additionally  $G$  is said to be *properly embeddable in the grid* if it is grid locatable and the segments representing edges do not cross each other. We study the area needed to obtain those representations for some graph families.

**Keywords:** graph drawing, grid locatable, grid embeddable, chromatic number.

## 1 Introduction

Graph drawing applies topology and geometry to derive suitable representations of graphs. Particularly, grid representations of graphs have attracted the attention of many researchers (see, for example [4,7,11]).

A *grid point* is a point of the plane having integer coordinates. A line segment  $s$  joining two grid points is said to be *primitive* if the only grid points in  $s$  are its extremes.

A graph  $G$  is said to be *grid locatable* (or *locatable in the grid*), if each vertex is represented by a grid point and each edge by a primitive segment joining its extremes.  $G$  is said to be *properly embeddable in the grid* (*p-embeddable* for short) if it is grid locatable and the segments representing edges do not cross each other. Observe, that the graphs so obtained are subgraphs of the visibility graph of the grid points, where two points are visible if there is no other grid point between them (this visibility graph was defined, for example, in [12]).

---

\* Partially supported by projects JA-FQM305, ESF EUROCORES programme EuroGIGA - ComPoSe IP04 - MICINN project EUI-EURC-2011-4306, MTM2008-05866-C03-01, and PAI-FQM141.

In this way, a graph is grid locatable (resp. p-embeddable) if it is isomorphic to a subgraph (resp. plane subgraph) of the visibility graph of the grid points.

The following characterization is shown in [9].

**Theorem 1.** [9] *A graph  $G$  is grid locatable if and only if  $G$  is 4-colorable.*

One of the reasons to study grid representations of graphs is to test how compact a given representation can be. Thus, a very important question in this kind of representation is to minimize the area needed to represent a given graph.

A *line* of the grid  $\mathbb{Z} \times \mathbb{Z}$  is a set  $\{(x, k)/x \in \mathbb{Z}\}$  for  $k \in \mathbb{Z}$  fixed. By grid locating (or p-embedding) a graph  $G$  in  $\ell$  lines we mean to locate (or embed)  $G$  in  $\mathbb{Z} \times \mathbb{Z}$  mapping the vertices to grid points in  $\ell$  consecutive lines (there is no unused line between two lines in our drawing).

In [9], an upper bound on the number of lines needed to grid locate a graph in the grid is given.

**Corollary 1.** [9] *A graph  $G$  with chromatic number  $\chi(G) \leq 4$  can be located in the grid in, at most,  $\chi(G)$  lines.*

Although the authors give an upper bound to the number of lines needed in order to grid locate a graph with  $\chi(G) \leq 4$ , in general, that upper bound is not tight. In this work, we deal with this problem. Namely, to find a tighter upper bound of the area (number of lines) needed to grid locate or p-embed a given graph in the grid.

A simple but useful observation is the following:

*Remark 1.* An edge joining two grid points  $(x_1, y_1)$  and  $(x_2, y_2)$  is primitive if and only if  $|x_1 - x_2|$  and  $|y_1 - y_2|$  are relatively prime to each other.

The structure of this paper is as follows. In Sect. 2, we show some results about p-embeddability of graphs in the grid. In Sect. 3, we provide our main results, giving a characterization of those graphs that are grid locatable in two and three lines. We shall conclude with some remarks and open problems.

## 2 P-Embedding a Graph in the Grid

Until very recently, it was not known whether any planar graph could be p-embedded in the plane or not, and there were only some partial results on the subject (see [9]). This problem has just been solved in the affirmative independently by M. Balko [1] and Flores-Peña et al. [8], although the grid size of the resulting representation is still excessively large. So it makes sense to find bounds on the size of the grid for some specific families of graphs, as we do below for bipartite graphs.

**Proposition 1.** *Any plane bipartite graph can be p-embedded in the grid.*

*Proof.* In [2] it is shown how to embed a quadrangulation in the grid. They prove firstly that any planar quadrangulation admits two different representations in two 2–books (in a 2–book representation, vertices are mapped to a line, called the *spine*, and edges are represented by crossing-free curves which do not cross the spine either). Then, they assign to each vertex the coordinates of its relative positions in the spines of both 2–books. A similar representation can be performed by forcing the differences of the  $x$ -coordinates to be relatively prime each other, guaranteeing the condition of primitiveness. This can be achieved by replacing the first  $n$ -positive integer  $x$ -coordinates of the graph by the first  $n$  primes.  $\square$

As a consequence of the proof of Proposition 1, and applying a direct implication of the well known Prime Number Theorem that establishes that the asymptotic expression for the  $n$ th prime number is  $n \log n$  (see [10]), we obtain a (not really tight) upper bound on the area needed to p-embed a bipartite graph in the grid.

**Corollary 2.** *Any plane bipartite graph with  $n$  vertices can be p-embedded in a rectangle of area  $O(n^2 \log n)$ .*

On the other hand, a complete characterization is given in [5] for those graphs that are p-embeddable in two lines.

**Proposition 2.** [5] *A graph  $G$  is p-embeddable in two lines if and only if  $G$  can be extended to a maximal outerplanar graph such that its dual (excluding the vertex representing the outerface) is a path.*

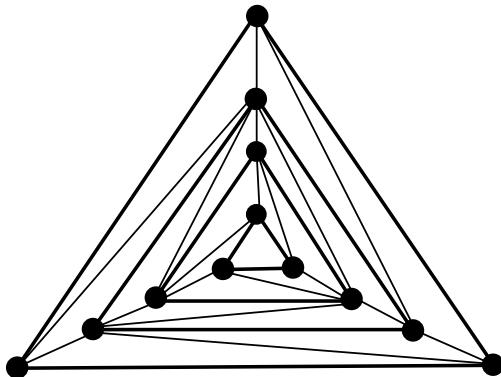
It is known that if a graph is grid locatable in the plane, then it is grid locatable in at most four lines, however, it is hopeless to try to obtain a similar result regarding p-embeddability. This is because some graphs with  $n$  vertices, as the *nested triangles graphs* (see Fig. 1), are shown to require a grid of size  $(2n/3 - 1) \times (2n/3 - 1)$  for a plane straight-line drawing [6]. Then, it seems to be worthwhile to improve the bounds on the area needed for certain families of graphs to be p-embeddable.

### 3 Locating a Graph in the Grid

As we said above, the upper bound provided in the number of lines required to grid locate a graph given in Theorem 1 is not always optimal. In this section we tackle the problem of minimizing such number of lines. In other words, given a concrete 4-colorable graph, how many lines are needed to represent it? Next result answers this question for outerplanar graphs.

**Proposition 3.** *Any outerplanar graph can be located in two lines.*

*Proof.* We will show that any maximal outerplanar graph  $G$  can be located in two lines (and then the result for any outerplanar graph is straightforward).



**Fig. 1.** The nested triangle

The proof can be obtained inductively, just starting by a triangle of the maximal outerplanar graph  $G$  and adding its adjacent triangles (at most three) taking into account some easy rules.

Assume that a subgraph  $G'$  (made of adjacent triangles) of  $G$  has already been located in two lines and that a new triangle  $uvw$  of  $G$  is going to be added and it shares the edge  $uv$  with a previously located triangle of the subgraph  $G'$ . Then, one of the following situations holds:

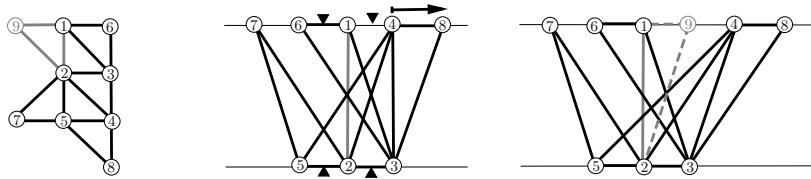
- Either  $u$  and  $v$  are mapped to contiguous vertices in the same line, or
- $u$  and  $v$  are mapped to vertices in different lines but at least one of them is not adjacent (in  $G'$ ) to one of its contiguous vertices in the same line.

In the first situation, we add the new vertex to the other line and in the second case, we add the new vertex to the place of the contiguous vertex that is not adjacent. In this last case, we possibly have to move some of the previous vertices in one of the lines (the one the new vertex is added to) in order to make room for the new vertex  $w$ . Figure 2 shows how to add a vertex in this last case. Triangle 129 is going to be added and it shares the edge 12 with the previously located triangle 123. Since 1 and 4 are not adjacent, then vertices 4 and 8 move to map the vertex  $w = 9$  (the tiny triangles in the central image mean that one of those points is in the outerface).  $\square$

We can describe the general structure of any graph that is grid locatable in two or three lines.

**Theorem 2.** *Let  $G = (V, E)$  be a graph such that  $\chi(G) \leq 4$ . Then,*

1.  *$G$  is grid locatable in two lines if and only if  $V$  can be partitioned into two subsets  $V_1, V_2$ , such that the subgraph of  $G$  induced by  $V_i$ ,  $i = 1, 2$ , is a disjoint union of paths.*



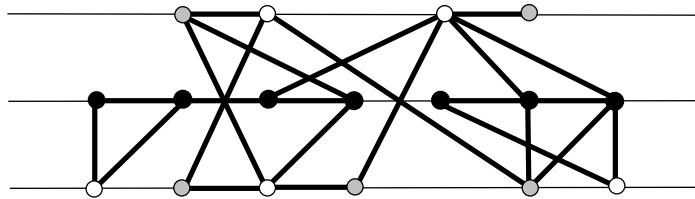
**Fig. 2.** Adding a vertex in the second case

2.  $G$  is grid locatable in three lines if and only if  $V$  can be partitioned into three subsets  $V_1, V_2, V_3$ , such that the subgraph of  $G$  induced by  $V_1$  is a disjoint union of paths, and  $V_2$  and  $V_3$  are independent sets.

*Proof.* The proof of (1) is trivial.

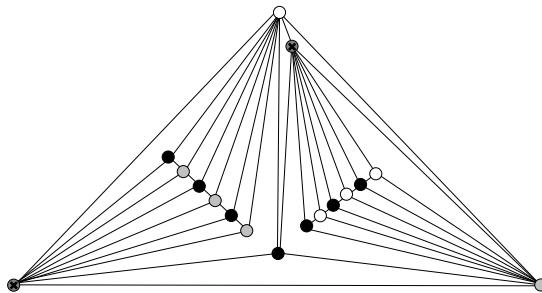
To prove (2), assume first that  $G$  is grid locatable in three lines and that those lines are numbered 1, 2 and 3 from bottom to top. We assign all the vertices in line 2 to  $V_1$ , vertices with odd abscissae in lines 1 and 3 to  $V_2$  and vertices with even abscissae in lines 1 and 3 to  $V_3$  (see Fig. 3).

To prove the other implication, once the set  $V$  is partitioned into sets  $V_1, V_2, V_3$ , we can draw the graph in three lines numbered 1, 2 and 3 from bottom to top as follows. Since the subgraph of  $G$  induced by  $V_1$  is a disjoint union of paths, we embed those paths in line 2. Then, vertices of  $V_2$  are mapped to vertices with odd abscissae of line 1, and vertices in  $V_3$  are mapped to vertices with even abscissae in line 3.  $\square$



**Fig. 3.** The coloring of the vertices of the graph located in three lines

Theorem 2 can be used to obtain examples of planar graphs which prove that the bound given in Corollary 1 is sometimes reached. For instance, consider the graph depicted in Fig. 4; it has chromatic number 4 and it requires exactly four lines to be located in the grid. This is because (up to permutation of the colors) there is only one 4-coloring and there does not exist any assignment of the colors avoiding vertices of degree 3 in the subgraph induced by two colors.



**Fig. 4.** Four lines are needed to locate this graph in the grid

Additionally, Theorem 2 is one of the main tools used to prove some results regarding grid locatability when the maximum degree of the graph is bounded.

From now on, for a graph  $G$ ,  $\delta(v)$  denotes the degree of a vertex  $v$  and  $\Delta(G)$  the maximum vertex degree in  $G$ . Similarly, for a directed graph  $G$ ,  $\delta^-(v)$  denotes the in-degree of a vertex  $v$  and  $\Delta^-(G)$  the maximum vertex in-degree in  $G$ .

**Lemma 1.** *Let  $G$  be a connected, directed, acyclic and not isomorphic to  $K_5$  graph, with  $\Delta(G) \leq 4$  and  $\Delta^-(G) \leq 3$ . Then  $G$  is grid locatable in two lines.*

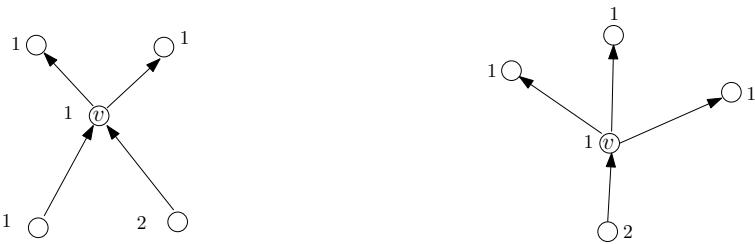
*Proof.* We are going to make use of the first statement of Theorem 2 by partitioning the vertices of  $G$  into two subsets,  $V_1$  and  $V_2$ , such that the subgraphs of  $G$  induced by them are disjoint union of paths.

Each directed acyclic graph induces a partial order  $\leq$  on its vertices, where  $u \leq v$  if there exists a directed path from  $u$  to  $v$ . Once the vertices are sorted, we sequentially assign them labels 1 or 2 as follows: Vertex with no incoming edges are marked with label 1, and the rest are assigned the less popular label among those used by their tails (for vertices with exactly two incoming edges with different colors in their tails, we choose any of them).

After this procedure, the subgraphs of  $G$  induced by each label class,  $G|_{V_1}$  and  $G|_{V_2}$ , may have vertices of degree three but this may only occur when a vertex  $v$  is the tail of two or three vertices of its same color class, as it is depicted in Fig. 5. This is avoided by sweeping the vertices of  $G$  in reverse order and changing the labels of such vertices. Note that to change the label of a vertex may produce a new degree three vertex, but always in a vertex that has not been visited yet.

It only rests to prove that the graphs  $G|_{V_1}$  and  $G|_{V_2}$ , regardless of the edges orientation, have no cycles.

First note that a vertex that have changed its label in the previous sweeping cannot take part of a cycle, since it becomes a vertex with degree one in its class. This is obvious in the right case in Fig. 5. In the left case,  $v$  could be a vertex of degree two in  $G|_{V_2}$  if its tail with label 1 would change to label 2 in a later step of the sweeping. But this can not happen since now its head,  $v$ , has a different label than it.



**Fig. 5.** Vertices with degree 3 in the subgraph induced by one of the color classes

Finally, if there is a monochrome cycle, there must be at least a vertex  $v$  with two incoming edges (since  $G$  is a directed acyclic graph). But if two incoming edges have the same color in their tails then  $v$  should have been previously assigned a different color and we have already shown that no vertex of the cycle has changed its color, what is a contradiction.  $\square$

**Theorem 3.** Let  $G$  be a connected graph, not isomorphic to  $K_5$ , with  $\Delta(G) \leq 4$ .

- a) If  $G$  is 4-regular, then  $G$  is grid locatable in three lines.
- b) If there exists a vertex  $v$  such that  $\delta(v) \leq 3$ , then  $G$  is grid locatable in two lines.

*Proof. Proof of a):*

Given that  $\Delta(G) \leq 4$ , since  $G$  is 4-regular, every vertex has -at least- two non-adjacent neighbors. Let  $x_n$  be one of its vertices, with two non-adjacent neighbors  $x_1$  and  $x_2$ .

By mimicking Brooks' greedy algorithm to color a graph [3], we assign colors  $\{1, 2, 3, 4\}$  to the vertices as we describe below.

Sort the vertices of  $G$  so that if  $i < j$  then  $d(x_n, x_i) \geq d(x_n, x_j)$  for  $i \geq 3$  (for vertices at the same distance from  $x_n$ , we freely choose how to sort them) and sequentially color them as follows:

1. Assign color 1 both to  $x_1$  and  $x_2$ .
2. For  $i \geq 3$ , assign preferentially colors 1 or 2 to  $x_i$  unless colors 3 or 4 are bound to be used.

First, we show that this assignment of colors provides a 4-coloring of  $G$ . This can be done by induction: Obviously, it is a proper coloring for  $x_1$  and  $x_2$  since they are not adjacent. It also works for  $x_n$  since both  $x_1$  and  $x_2$  has the same color. Now assume that, for  $3 \leq i < n$ ,  $x_i$  has already been well-colored and let us color  $x_{i+1}$ . There must exist a vertex  $x_k$ ,  $i+1 < k \leq n$  such that  $d(x_n, x_k) \leq d(x_n, x_{i+1})$ . Since  $\delta(x_k) \leq 4$ , at most three of its neighbors have been already colored and hence there is still a free color for  $x_{i+1}$ .

Next we prove that the subgraph induced by colors 3 and 4, denoted  $H_{34}$ , is a disjoint union of paths.

Observe that  $\delta_{H_{34}}(x_i) \leq 2$  for every  $x_i \in V(H_{34})$ ,  $i \neq n$ . On the other hand, if  $x_n \in V(H_{34})$  then  $\delta_{H_{34}}(x_n) = 1$ . This proves that  $H_{34}$  is the union of paths and cycles and, in the case of having cycles,  $x_n$  does not take part of none of them.

If  $H_{34}$  is supposed to include a cycle, consider the vertex  $x_j$  whose label  $j (< n)$  is the maximum among the vertices lying on the cycle. Since  $x_j$  has either color 1 or 2 and colors 3 and 4 are assigned preferentially,  $x_j$  has (at least) two neighbors previously colored 3 and 4 and still another neighbor to be colored later, what implies  $\delta(x_j) \geq 5$ , a contradiction.

The proof is concluded by virtue of Theorem 2.

**Proof of b):** Assume now the case of  $G$  having a vertex, say  $x_n$ , with degree at most three.

We sort the vertices of  $G$  so that if  $i < j$  then  $d(x_n, x_i) \geq d(x_n, x_j)$  for  $1 \leq i < n$  (again, for vertices at the same distance from  $x_n$ , we freely choose how to sort them).

This ordering induces an acyclic digraph in  $G$  with maximum in-degree 3. Now, applying Lemma 1, the result holds.  $\square$

## 4 Concluding Remarks and Open Questions

As it was said in the introduction, grid embeddings of graphs as subgraphs of the visibility graph of the grid points has been considered previously in the literature, mainly in [9].

But, being an important question that of the compactness of such representations, we are not aware of many results in this subject. This is the main gap we have tried to fill in this work: What is the minimum number of lines needed to represent (either to p-embed – without crossing of edges– or to grid locate – allowing edges crossing–) in the grid a given graph?

Regarding p-embeddability, although it is known that any planar graph can be p-embedded in the plane, the question of the minimal area needed to obtain such representation is still open, since we have provided only partial results concerning bipartite graphs.

On the other hand, regarding grid location, we have proved that a 4-regular graph (non-isomorphic to  $K_5$ ) can be located in three lines, and that if a graph is not regular and its maximal degree is 4, then it can be located in two lines. Other cases are still open.

## References

1. Balko, M.: Grid Representations and the Chromatic Number. In: 28th European Workshop on Computational Geometry, pp. 45–48 (2012)
2. Barrière, L., Huemer, C.: 4-Labelings and Grid Embeddings of Plane Quadrangulations. In: Eppstein, D., Gansner, E.R. (eds.) GD 2009. LNCS, vol. 5849, pp. 413–414. Springer, Heidelberg (2010)

3. Brooks, R.L.: On colouring the nodes of a network. Cambridge Philosophical Society, Math. Phys. Sci. 37, 194–197 (1941)
4. Chrobak, M., Nakano, S.: Minimum width grid drawings of plane graphs. Computational Geometry 11(1), 29–54 (1998)
5. Cornelsen, S., Schank, T., Wagner, D.: Drawing Graphs on Two and Three Lines. In: Goodrich, M.T., Kobourov, S.G. (eds.) GD 2002. LNCS, vol. 2528, pp. 31–41. Springer, Heidelberg (2002)
6. Dolve, D., Leighton, F.T., Trickey, H.: Planar embedding of planar graphs. Advances in Computing Research. VLSI Theory, vol. 2. JAI Press, Inc., Greenwich (1984)
7. Fraysseix, H., Pach, J., Pollack, R.: How to draw a planar graph on a grid. Combinatorica 10(1), 41–51 (1990)
8. Flores-Peñaloza, D., Santos, F., Zaragoza, F.J.: Encajes primitivos de gráficas aplanables. In: XXVII Coloquio Víctor Neumann (2012)
9. Flores-Peñaloza, D., Zaragoza, F.J.: Every four-colorable graph is isomorphic to a subgraph of the visibility graph of the integer lattice. In: 21st Canadian Conference on Computational Geometry, CCCG 2009, pp. 91–94 (2009)
10. Hardy, G.H., Wright, E.M.: An introduction to the theory of numbers. Oxford University Press (1979)
11. Kára, J., Pór, A., Wood, D.R.: On the chromatic number of the visibility graph of a set of points in the plane. Discrete and Computational Geometry 34(3), 497–506 (2005)
12. Por, A., Wood, D.R.: On visibility and blockers. J. Computational Geometry 1(1), 29–40 (2010)

# On the Heaviest Increasing or Decreasing Subsequence of a Permutation, and Paths and Matchings on Weighted Point Sets

Toshinori Sakai<sup>1</sup> and Jorge Urrutia<sup>2</sup>

<sup>1</sup> Tokai University, Shibuya-ku, Tokyo, Japan  
[sakai@tokai-u.jp](mailto:sakai@tokai-u.jp)

<sup>2</sup> Universidad Nacional Autónoma de México, México D.F., México  
[urrutia@matem.unam.mx](mailto:urrutia@matem.unam.mx)

**Abstract.** Let  $S = \{s(1), \dots, s(n)\}$  be a permutation of the integers  $\{1, \dots, n\}$ . A subsequence of  $S$  with elements  $\{s(i_1), \dots, s(i_k)\}$  is called an *increasing* subsequence if  $s(i_1) < \dots < s(i_k)$ ; It is called a *decreasing* subsequence if  $s(i_1) > \dots > s(i_k)$ . The weight of a subsequence of  $S$ , is the sum of its elements. In this paper, we prove that any permutation of  $\{1, \dots, n\}$  contains an increasing or a decreasing subsequence of weight greater than  $n\sqrt{2n}/3$ .

Our motivation to study the previous problem arises from the following problem: Let  $P$  be a set of  $n$  points on the plane in general position, labeled with the integers  $\{1, \dots, n\}$  in such a way that the labels of different points are different. A non-crossing path  $\Pi$  with vertices in  $P$  is an increasing path if when we travel along it, starting at one of its end-points, the labels of its vertices always increase. The weight of an increasing path, is the sum of the labels of its vertices. Determining lower bounds on the weight of the heaviest increasing path a point set always has.

We also study the problem of finding a non-crossing matching of the elements of  $P$  of maximum weight, where the weight of an edge with endpoints  $i, j \in P$  is  $\min\{i, j\}$ .

## 1 Introduction

Let  $n$  be a positive integer. Consider any permutation  $S$  of the integers  $\{1, \dots, n\}$ . A well known result of Erdős and Szekeres [4], asserts that  $S$  always contains an increasing or a decreasing subsequence with at least  $\lceil \sqrt{n} \rceil$  elements. The weight of a subsequence of a permutation  $S$  is the sum of its elements. For example if we consider the permutation:

$$5, 2, 8, 1, 7, 4, 3, 6,$$

the weight of the increasing subsequence 2, 4, 6 is equal to 12. In this paper we study the problem of finding the *heaviest* increasing, or decreasing subsequence of a permutation. For the permutation considered before, the heaviest increasing, or decreasing subsequence, consists of 8, 7, 4, 3 and has weight 22.

We prove that any permutation of  $\{1, \dots, n\}$  always has an increasing or a decreasing subsequence with weight greater than  $n\sqrt{n/3}$ , our bound is asymptotically tight. Our solution is somehow related to a well known problem of Tutte [8] involving the efficient packing of squares of different sizes into a rectangle, or a square of small area. The permutations giving the tight bound for our problem, produce efficient packings of squares with areas  $1^2, 2^2, \dots, n^2$  under some conditions.

A finite sequence  $a_1, a_2, \dots, a_k$  is said to be *unimodal* (resp. *anti-unimodal*) if there is an  $m$ ,  $1 \leq m \leq k$ , such that  $a_1 < a_2 < \dots < a_m$  and  $a_m > a_{m+1} > \dots > a_k$  (resp.  $a_1 > a_2 > \dots > a_m$  and  $a_m < a_{m+1} < \dots < a_k$ ). We also study the problem of finding a heavy unimodal or anti-unimodal subsequence of a permutation of  $\{1, \dots, n\}$ . We show that any permutation of  $\{1, \dots, n\}$ , always has a unimodal or anti-unimodal subsequence of weight greater than  $n\sqrt{2n/3}$ .

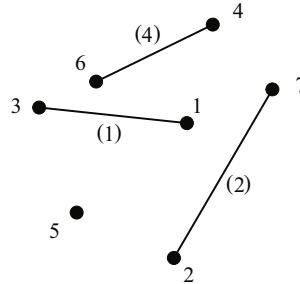
Let  $P$  be a set of  $n$  points on the plane in general position which are labeled with  $\{1, \dots, n\}$ . Our results are motivated by the study of longest or heaviest non-crossing increasing paths connecting elements of  $P$ . In particular, as we will see in Section 2, for labeled point sets in convex position, finding a non-crossing increasing path with maximum weight, is related to that of finding a heaviest unimodal or anti-unimodal subsequence in some permutation. The problem of finding non-crossing paths, cycles and trees with vertices on point sets on the plane optimizing some given functions, has been of interest to many computational geometers for some time now. Alon, Rajagopalan, and Suri [1], and Dumitrescu and Tóth [3] studied the problem of finding non-crossing paths, matchings, cycles, and trees of maximum length, where the length od a cycle, matching, and tree, is the sum of the lengths of its edges.

Károlyi, Pach, and Tóth [6] show that if the edges of a complete geometric graph on  $k^2 + 1$  points are colored red or blue, then there always exists a non-self intersecting red or blue path with  $k + 1$  edges.

The problem of finding non-crossing monotonic paths with many edges was first studied by Czyzowicz et al [5]. They proved that any labeled point set in convex position, contains a non-crossing monotonic path with at least  $\sqrt{2n}$  edges. This bound was improved recently by Sakai and Urrutia [7] to  $\sqrt{3n - 3/4} - 1/2$  by giving a simple proof for a result by Chung [2] concerning the length (number of elements) of a longest unimodal or anti-unimodal subsequences in a sequence.

For labeled point set  $P$ , we also study the problem of finding a non-crossing matching of  $P$ , not necessarily perfect, such that the sum of the weights of its edges is maximized, where the *weight* of an edge joining points  $i$  and  $j$  is the smallest of  $\{i, j\}$ , see Figure 1.

We show that  $P$  always has a matching of weight greater than  $n^2/6$  for  $n \geq 2$ . When the elements of  $P$  are in convex position, then  $P$  has a matching of weight at least  $\lfloor n^2/5 \rfloor$ . We observe that if we define the weight of the edge joining point  $i$  to point  $j$  to be the largest of  $\{i, j\}$ , and  $P$  has  $n = 2m$  elements, then we can always find a perfect matching with weight  $(m + 1) + \dots + 2m$ , which is the largest possible. To prove this, we recall a well known result that if a point set  $S$  has  $2m$  elements (in general position),  $m$  colored red, and  $m$  colored blue, then



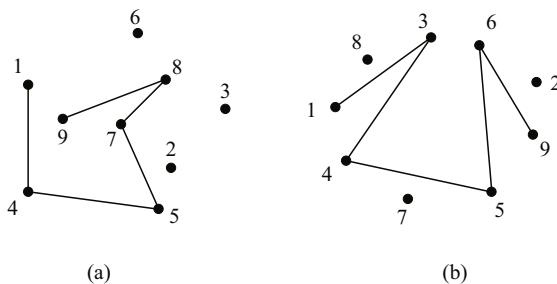
**Fig. 1.** A non-crossing matching with edges having weights 1, 2, 4

there always exists a non-crossing perfect matching of the elements of  $S$ , such that any edge in the matching has a blue and an red endpoint. Now color the elements of  $P$  as follows: those elements whose label is smaller than or equal to  $m$  are colored red, and the remaining points blue. Our result follows.

## 2 The Heaviest Monotonic Subsequences and Paths

In the rest of this paper,  $P$  will denote a set of  $n$  points in general position whose elements are labeled with the integers  $\{1, \dots, n\}$  in such a way that different points receive different labels. A point in  $P$  labeled  $i$  will be denoted as point  $i$ .

A path  $\Pi$  of  $P$  is a sequence of elements  $i_1, \dots, i_k$  of  $P$  together with the closed line segments  $[i_r, i_{r+1}]$  (called the *edges* of  $P$ ) joining  $i_r$  to  $i_{r+1}$ ,  $r = 1, \dots, k - 1$ . We say that  $\Pi$  is non-crossing if its edges do not cross each other. If in addition  $i_r < i_{r+1}$ ,  $1 \leq r \leq k - 1$ , we call  $\Pi$  a non-crossing increasing path of  $P$ . The *weight* of  $\Pi$  is equal to  $i_1 + \dots + i_k$ .



**Fig. 2.** (a) A non-crossing increasing path. (b) Finding the non-crossing increasing path with vertices 1, 3, 4, 5, 6, 9 is equivalent to finding the unimodal subsequence 3, 6, 9, 5, 4, 1 in the permutation {8, 3, 6, 2, 9, 5, 7, 4, 1} (or the anti-unimodal subsequence 9, 5, 4, 1, 3, 6 in the permutation {9, 5, 7, 4, 1, 8, 3, 6, 2}).

In this section we will study the following problem: Find the largest integer  $\mathcal{W}(n)$  such that any labeled point set with  $n$  elements has a non-crossing increasing path of weight at least  $\mathcal{W}(n)$ . In Figure 2(a), we show a non-crossing increasing path of weight 34. The path with vertices 1, 7, 8, 9 is increasing, but has a crossing.

Observe that if the elements of  $P$  are in convex position, that is, they are the vertices of a convex polygon, then the problem of finding a non-crossing increasing path of maximum weight, can be reduced to that of finding a unimodal or anti-unimodal subsequence of maximum weight in a permutation of  $\{1, \dots, n\}$  obtained from  $P$ , by reading its elements starting at some point of  $P$ , see Figure 2(b).

## 2.1 The Heaviest Monotonic Subsequence of a Permutation

Let  $S = \{s(1), \dots, s(n)\}$  be a permutation of  $\{1, \dots, n\}$ . A subsequence of  $S$  with elements  $\{s(i_1), \dots, s(i_k)\}$  is called an *increasing* subsequence if  $s(i_1) < \dots < s(i_k)$ , or a *decreasing* subsequence if  $s(i_1) > \dots > s(i_k)$ . The subsequence is also called a *monotonic* subsequence if it is increasing or decreasing. The *weight* of  $\{s(i_1), \dots, s(i_k)\}$  is  $s(i_1) + \dots + s(i_k)$ .

For each  $s(i)$  of  $S$ , let us associate to it the point  $(x_i, y_i)$  as follows:  $x_i$  is the weight of the heaviest increasing subsequence of  $S$  staring at  $s(i)$ , and  $y_i$  is the weight of the heaviest decreasing subsequence of  $S$  staring at  $s(i)$ . For example if  $S = \{4, 3, 7, 2, 5, 1, 6\}$ , then to  $s(2) = 3$ , we associate the point  $(x_2, y_2) = (3 + 5 + 6, 3 + 2 + 1) = (14, 6)$ . It is easy to see that for indices  $i < j$ , if  $s(i) < s(j)$ , then  $x_i \geq s(i) + x_j$ , and if  $s(i) > s(j)$ , then  $y_i \geq s(i) + y_j$ . We now prove:

**Theorem 1.** *Any permutation of  $\{1, \dots, n\}$  contains a monotonic subsequence of weight greater than  $n\sqrt{n/3}$ . Our bound is asymptotically tight.*

*Proof.* We now associate to each  $s(i)$  the square  $SQ(i)$  whose top-right vertex is the point  $(x_i, y_i)$ , and whose bottom-left vertex is the point  $(x_i - s(i), y_i - s(i))$ .

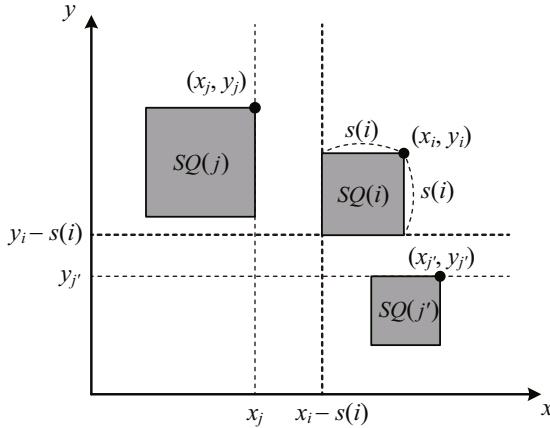
We now prove that if  $i \neq j$  then  $SQ(i)$  and  $SQ(j)$  have disjoint interiors. Suppose that  $i < j$ . Two cases arise. Suppose first that  $s(i) < s(j)$ . Then  $x_j \leq x_i - s(i)$ , and  $SQ(j)$  lies to the left of the vertical line with equation  $x = x_i - s(i)$ . If  $s(i) > s(j)$  then  $y_j \leq y_i - s(i)$ , and  $SQ(j)$  lies below the horizontal line  $y = y_i - s(i)$ . Thus in both cases  $SQ(i)$  and  $SQ(j)$  have disjoint interiors. See Figure 3.

Let  $R$  be a rectangle whose bottom-left vertex is the origin, and whose top-right vertex is the point  $(\alpha, \beta)$ , where  $\alpha$  is the maximum among all  $x_i$  and  $\beta$  is the maximum among all  $y_i$ . Since  $R$  contains all the squares  $SQ(i)$ ,  $i = 1, \dots, n$ ,

$$\text{the area of } R = \alpha\beta \geq 1^2 + 2^2 + \dots + n^2 > \frac{n^3}{3},$$

and thus

$$\max\{\alpha, \beta\} > n\sqrt{\frac{n}{3}}.$$



**Fig. 3.**  $SQ(i)$  and  $SQ(j)$ ,  $i \neq j$ , have disjoint interiors

Therefore for some  $i$ ,  $x_i$  or  $y_i$  is greater than  $n\sqrt{n/3}$ .

To see that our bound is asymptotically tight, we proceed as follows. Define  $k$  and  $m$  by  $k = \sqrt[4]{4n^3/3}$ , and  $m = \sqrt{3n}/2$ . Consider now the permutation  $\Pi$  generated as follows:

$$\begin{aligned} \lceil k \rceil, \lceil k \rceil - 1, \dots, 1, \lceil \sqrt{2}k \rceil, \lceil \sqrt{2}k \rceil - 1, \dots, \lceil k \rceil + 1, \\ \lceil \sqrt{3}k \rceil, \lceil \sqrt{3}k \rceil - 1, \dots, \lceil \sqrt{2}k \rceil + 1, \dots, n, n - 1, \dots, \lceil \sqrt{m-1}k \rceil + 1. \end{aligned}$$

The permutation  $\Pi$  consists of  $m$  blocks of decreasing integers, such that for each of them, the sum of its elements is  $n\sqrt{n/3} + O(n)$ . On the other hand, the heaviest increasing subsequence of  $\Pi$ , is the one containing the elements  $\lceil k \rceil, \lceil \sqrt{2}k \rceil, \lceil \sqrt{3}k \rceil, \dots, n$ , which again has weight  $n\sqrt{n/3} + O(n)$ . Our result follows.  $\square$

We can now prove:

**Theorem 2.** *Any labeled point set  $P$  with  $n$  elements has a non-crossing increasing path of weight greater than  $n\sqrt{n/3}$ .*

*Proof.* Assume that no two elements of  $P$  lie on a vertical line, otherwise rotate  $P$  slightly to achieve this condition. Project the elements of  $P$  on the  $x$ -axis, and let  $\Pi$  be the permutation of  $\{1, \dots, n\}$  obtained by reading the projections of the elements of  $P$  on the  $x$ -axis from left to right. Then  $\Pi$  has an increasing or decreasing subsequence of weight greater than  $n\sqrt{n/3}$ . This subsequence induces a non-crossing increasing path of  $P$  with weight greater than  $n\sqrt{n/3}$ .

We conclude this section with general result concerning heaviest subsequences and paths. By arguing as in the proof of Theorems 1 and 2, we obtain:

**Theorem 3.** (a) Any sequence  $T = \{a_1, a_2, \dots, a_n\}$  of  $n$  distinct positive numbers contains a monotonic subsequence of weight at least  $\sqrt{\sum_{i=1}^n a_i^2}$ .  
(b) Any set of points in general position, labeled with elements of  $T$ , has a non-crossing increasing path of weight at least  $\sqrt{\sum_{i=1}^n a_i^2}$ .

## 2.2 The Heaviest Increasing Paths of Point Sets in Convex Position

In this section we study the problem of finding a non-crossing increasing path of large weight for point sets in convex position. Our result will be based on the next result on unimodal or anti-unimodal subsequences of permutations.

**Theorem 4.** Any permutation  $S$  of  $\{1, \dots, n\}$  contains a unimodal or anti-unimodal subsequence of weight greater than  $n\sqrt{2n/3}$ .

*Proof.* The result follows immediately for  $n \leq 3$ . Thus assume  $n \geq 4$ . Let  $S = \{s(1), \dots, s(n)\}$  be a permutation, and let  $A = \sum_{i=1}^n s(i)^2$ .

**Lemma 1.** There exist  $j$ ,  $1 \leq j \leq n$ , and  $d$  with  $|d| \leq n^2/2$  such that

$$s(1)^2 + \dots + s(j)^2 = \frac{A}{2} - d \text{ and } s(j+1)^2 + \dots + s(n)^2 = \frac{A}{2} + d.$$

*Proof.* Take the maximum  $i$  such that  $s(1)^2 + \dots + s(i)^2 \leq A/2$ . If in addition  $(A - n^2)/2 \leq s(1)^2 + \dots + s(i)^2$ , then we have only to let  $j = i$ . Thus assume that  $s(1)^2 + \dots + s(i)^2 < (A - n^2)/2$ . In this case, we have  $A/2 < s(1)^2 + \dots + s(i)^2 + s(i+1)^2 < (A + n^2)/2$  by our choice of  $i$  and by  $0 < s(i+1) \leq n$ . Thus  $j = i+1$  satisfies the desired property.  $\square$

Let  $j$  and  $d$  be as in Lemma 1, and define the sequences  $S_1$  and  $S_2$  by

$$S_1 = \{s(1), \dots, s(j)\} \text{ and } S_2 = \{s(j+1), \dots, s(n)\}.$$

Denote by  $\alpha_1$  (resp.  $\beta_1$ ) the weight of the heaviest increasing (resp. decreasing) subsequence of  $S_1$ , and by  $\alpha_2$  (resp.  $\beta_2$ ) the weight of the heaviest increasing (resp. decreasing) subsequence of  $S_2$ . First note that  $\alpha_1 + \beta_2$  and  $\beta_1 + \alpha_2$  are weights of a unimodal subsequence and an anti-unimodal subsequence of  $S$ , respectively.

By arguing as in the proof of Theorem 1, we obtain

$$\begin{aligned} \alpha_1 \beta_1 &\geq s(1)^2 + \dots + s(j)^2 = \frac{A}{2} - d, \text{ and} \\ \alpha_2 \beta_2 &\geq s(j+1)^2 + \dots + s(n)^2 = \frac{A}{2} + d. \end{aligned}$$

From these and the well known inequality  $(x+y+z+w)/4 \geq \sqrt[4]{xyzw}$  for positive numbers  $x, y, z$  and  $w$ , it follows that

$$\begin{aligned} (\alpha_1 + \beta_2) + (\beta_1 + \alpha_2) &\geq 4\sqrt[4]{(\alpha_1 \beta_1)(\alpha_2 \beta_2)} \\ &\geq 4\sqrt[4]{(A/2)^2 - d^2} \\ &\geq 4\sqrt[4]{(A^2 - n^4)/4} \end{aligned}$$

Since  $A = n(n+1)(2n+1)/6 > (n^3 + n^2)/3$ , we have  $A^2 - n^4 > n^6/9$  for  $n \geq 4$ . Thus  $(\alpha_1 + \beta_2) + (\beta_1 + \alpha_2) > 2n\sqrt{2n/3}$ , and hence  $\max\{\alpha_1 + \beta_2, \beta_1 + \alpha_2\} > n\sqrt{2n/3}$ , as desired.  $\square$

At this point the best lower-bound we have for the weight of a unimodal or anti-unimodal subsequence of a permutation is approximately  $2n\sqrt{n/3}$ , and is given by the same permutation stated in the proof of Theorem 1. On the other hand, the best upper-bound we have for the weight of a non-intersecting increasing path of the set of  $n$  labeled points in convex position is approximately  $n\sqrt{2n}$ . This is given by the following (circular) permutation for  $n = 2k^2$ :

$$\begin{aligned} n-k+1, & n-2k+1, \dots, k+1, 1, \\ n-k+2, & n-2k+2, \dots, k+2, 2, \\ & \dots, \\ n=2k^2, & n-k, \dots, 2k, k. \end{aligned}$$

It is easy to see that among all the unimodal or anti-unimodal subsequences of all possible permutations, a heaviest one is:

$$n-k+1, n-k+2, \dots, n, n-k, \dots, 2k, k,$$

which has weight

$$\frac{(2n-k+1)k}{2} + \frac{n(2k-1)}{2} < 2nk = n\sqrt{2n}.$$

Thus we have:

**Theorem 5.** *Any set  $P$  of  $n$  labeled points in convex position has a non-crossing increasing path of weight greater than  $n\sqrt{2n/3}$ . Furthermore, there is a labeled point set in convex position such that any non-crossing increasing path has weight less than  $n\sqrt{2n}$ .*

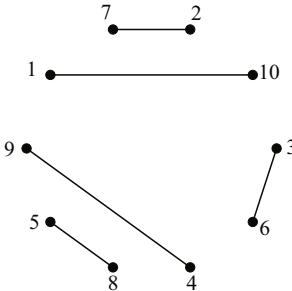
*Conjecture 1.* Any set  $P$  of  $n$  labeled points in convex position has a non-crossing increasing path of weight greater than  $n\sqrt{2n} - O(n)$ .

### 3 The Heaviest Matchings

In this section, we study the problem of finding non-crossing matchings of labeled point sets that maximize the sum of the weights of its edges. Recall that the weight of an edge joining point  $i$  to point  $j$ , is the minimum of  $\{i, j\}$ . The sum of the weights of all edges of a matching is called the *weight* of the matching.

#### 3.1 Point Sets in Convex Position

In this section, we consider the case where  $P$  is in convex position. We start by observing that even if  $P$  has  $2m$  elements, a non-crossing matching of maximum



**Fig. 4.** The weight of the perfect matching shown is 15. Moreover, any perfect matching of  $P$  has weight 15.

weight of  $P$  is not necessarily a perfect matching. For example for the point set shown in Figure 4, the weight of any non-crossing perfect matching is 15.

To prove this, observe that in any non-crossing perfect matching of the point set in Figure 4, any edge must leave an even number of points on each of its sides. Thus any such edge joins a point in  $\{1, 2, 3, 4, 5\}$  to a point in  $\{6, 7, 8, 9, 10\}$ . It follows now that the weights of the edges of any non-crossing perfect matching  $\mathcal{M}$  of  $P$  are precisely  $\{1, 2, 3, 4, 5\}$ , and thus the weight of  $\mathcal{M}$  is 15. On the other hand, a non-perfect matching with non-crossing edges  $\{[7, 2], [9, 10], [5, 6], [8, 4]\}$  has larger weight 20( $> 15$ ).

This example can be generalized as follows: Let  $P$  be an unlabeled point set with  $2m$  elements in convex position. We now color the elements of  $P$  red or blue, in such a way that when we traverse the boundary of the convex hull of  $P$  the colors of its elements alternate. Observe now that any edge of perfect matching  $\mathcal{M}$  of  $P$  joins a red to a blue point. Label the red and blue points of  $P$  with the integers  $\{1, \dots, m\}$  and  $\{m+1, \dots, 2m\}$ , respectively. Then the weight of any edge of  $\mathcal{M}$  belongs to  $\{1, \dots, m\}$ . Since different edges in  $\mathcal{M}$  have different weights, the weight of  $\mathcal{M}$  is precisely  $\binom{m+1}{2}$ .

The maximum weight a non-crossing perfect matching of  $P$  can have, is achieved when the elements of  $P$  are labeled  $1, 2, \dots, n$  in the clockwise order around the boundary of the convex hull of  $P$ . In this case, the weight of the heaviest matching of  $P$  is precisely  $1 + 3 + \dots + (2m-1) = m^2$ .

Since these bounds are valid for labeled point sets in general position, we have:

**Lemma 2.** *The weight of any non-crossing perfect matching of a labeled point set  $P$  in general position with  $2m$  elements is at least  $\binom{m+1}{2}$  and at most  $m^2$ . These bounds are tight.*

Using similar arguments, we obtain:

**Lemma 3.** *Let  $Q$  be a set of  $2m$  points in general position whose elements are labeled with the integers  $r+1, r+2, \dots, r+2m$ . Then the weight of any perfect matching of  $Q$  has the following bounds, and these bounds are tight:*

- at least  $(r+1) + (r+2) + \cdots + (r+m) = rm + \binom{m+1}{2}$ , and
- at most  $(r+1) + (r+3) + \cdots + [r+(2m-1)] = rm + m^2$ .

Next we consider matchings not necessarily perfect.

**Lemma 4.** *Let  $n_1 \geq 0$  and  $n_2 \geq 1$  be integers, and  $Q$  be a set of  $2(n_1 + n_2)$  points in convex position such that  $2n_1$  elements are colored red, and  $2n_2$  blue. Let  $p$  be a blue element of  $Q$ . Then there is a non-crossing matching  $\mathcal{M}$  of  $Q$  such that the endpoints of each edge have the same color, and  $\mathcal{M}$  matches all elements of  $Q$ , except at most two blue elements containing  $p$ .*

*Proof.* We proceed by induction on  $n = n_1 + n_2$ . The result follows immediately for  $n = 1$ . Thus assume  $n \geq 2$ . We label the elements of  $P$  as  $p_0, \dots, p_{2n-1}$  around the boundary of the convex hull of  $P$ , where  $p_0 = p$ . The indices are to be read modulo  $2n$ . If  $p_i$  and  $p_{i+1}$  have the same color for some  $i \not\equiv 0, -1$ , then the edge  $[p_i, p_{i+1}]$  together with edges of some matching  $\mathcal{M}'$  of  $Q \setminus \{p_i, p_{i+1}\}$  form a matching  $\mathcal{M}$  with the desired properties. Thus assume that the colors of  $p_1, p_2, \dots, p_{2n-1}$  alternate (so  $p_1$  and  $p_{2n-1}$  have a same color). Then  $\mathcal{M} = \{[p_1, p_{2n-1}], [p_2, p_{2n-2}], \dots, [p_{n-1}, p_{n+1}]\}$  has the desired properties (and  $p_n$  is colored blue).  $\square$

We are now ready to prove the main result of this section.

**Theorem 6.** *Let  $P$  be a labeled point set with  $n$  elements in convex position. Then the heaviest non-crossing matching of  $P$  has weight at least  $\lfloor n^2/5 \rfloor$ .*

*Proof.* To make our proof easy to understand, let us assume first that  $P$  has  $n = 5s$  elements.

To start, discard from  $P$  all the elements with labels in  $\{1, \dots, s-2\}$ . Now color with blue all the elements with labels in  $\{s-1, \dots, 3s\}$ , and with red all those with labels in  $\{3s+1, \dots, 5s\}$ . By Lemma 4, we can find matchings  $\mathcal{M}'$  and  $\mathcal{M}''$  of  $\{s-1, \dots, 3s\}$  and  $\{3s+1, \dots, 5s\}$ , respectively, such that  $\mathcal{M} = \mathcal{M}' \cup \mathcal{M}''$  is non-crossing and  $\mathcal{M}$  leave at most two elements of  $\{s-1, \dots, 3s\}$ , including  $s-1$ , unmatched.

First suppose that all the elements of  $\{s-1, \dots, 3s\}$  are matched. Then by Lemma 3 with  $r = s-2$  and  $m = s+1$ , the weight of  $\mathcal{M}'$  is at least

$$(s-1) + s + \cdots + (2s-2) + (2s-1) = \frac{(3s-2)(s+1)}{2}.$$

Similarly, by applying Lemma 3 with  $r = 3s$  and  $m = s$ , the weight of  $\mathcal{M}''$  is at least:

$$(3s+1) + (3s+2) + \cdots + (4s-1) + 4s = \frac{(7s+1)s}{2}.$$

Adding up these summations, we get that the weight of  $\mathcal{M} = \mathcal{M}' \cup \mathcal{M}''$  is at least:

$$5s^2 + s - 1 = \frac{n^2}{5} + \frac{n}{5} - 1.$$

Next consider the case where two elements of  $\{s-1, \dots, 3s\}$ , including  $s-1$ , are unmatched. Let  $t$  be the unmatched point other than  $s-1$ . Then the weight of  $\mathcal{M} \cup \{[s-1, t]\}$  (which might have crossings) is at least  $5s^2 + s - 1$  in this case as well. Since the weight of  $\mathcal{M}$  decreases by  $s-1$  from the weight of  $\mathcal{M} \cup \{[s-1, t]\}$ , the weight of  $\mathcal{M}$  is at least  $5s^2 = n^2/5$ .

We can verify that this bound is valid for  $n = 5s + k$  ( $1 \leq k \leq 4$ ) as follows: We discard from  $P$  the elements  $1, \dots, s+k-4$ , and color the elements  $s+k-3, \dots, 3s+k-2$  with blue, and the elements  $3s+k-1, \dots, 5s+k$  with red. Now we can argue as in the case where  $n = 5s$  to see that  $P$  has a non-crossing matching with weight at least  $5s^2 + 2ks + k - 1 \geq \lfloor (5s+k)^2/5 \rfloor = \lfloor n^2/5 \rfloor$ .  $\square$

### 3.2 Point Sets in General Position

For point sets in general position, we have:

**Theorem 7.** *Let  $P$  be a labeled point set with  $n \geq 2$  elements in general position. Then the heaviest non-crossing matching of  $P$  has weight at least  $\lfloor (n^2 + n)/6 \rfloor$ .*

*Proof.* First consider the case where  $n = 3s$ . In this case, we discard from  $P$  all the elements with labels  $1, \dots, s$ . Then by Lemma 3, any matching  $\mathcal{M}$  of the remaining points has weight at least

$$(s+1) + (s+2) + \dots + 2s = \frac{(3s+1)s}{2} = \frac{n(n+1)}{6}.$$

We can argue similarly for the case where  $n = 3s+1$  or  $3s+2$  by discarding from  $P$  the elements with labels  $1, \dots, s+1$  or  $1, \dots, s$ , respectively.  $\square$

## References

1. Alon, N., Rajagopalan, S., Suri, S.: Long non-crossing configurations in the plane. *Fundamenta Informaticae* 22, 385–394 (1995)
2. Chung, F.R.K.: On unimodal subsequences. *J. Combin. Theory Ser. A* 29, 267–279 (1980)
3. Dumitrescu, A., Tóth, C.: Long non-crossing configurations in the plane. *Discrete Comput. Geom.* 44, 727–752 (2010)
4. Erdős, P., Szekeres, G.: A combinatorial problem in geometry. *Compositio Math.* 2, 463–470 (1935)
5. Czyzowicz, J., Kranakis, E., Krizanc, D., Urrutia, J.: Maximal length common non-intersecting paths. In: Proc. Eighth Canadian Conference on Computational Geometry, Ottawa, pp. 180–189 (August 1996)
6. Károlyi, G., Pach, J., Tóth, G.: Ramsey-type results for geometric graphs I. *Discrete and Computational Geometry* 18, 247–255 (1997)
7. Sakai, T., Urrutia, J.: Monotonic Polygons and Paths in Weighted Point Sets. In: Akiyama, J., Bo, J., Kano, M., Tan, X. (eds.) CGGA 2010. LNCS, vol. 7033, pp. 164–175. Springer, Heidelberg (2011)
8. Tutte, W.T.: The quest of the perfect square. *Amer. Math. Monthly* 72, 29–35 (1965)

# A Generalization of the Source Unfolding of Convex Polyhedra<sup>\*</sup>

Erik D. Demaine<sup>1</sup> and Anna Lubiw<sup>2</sup>

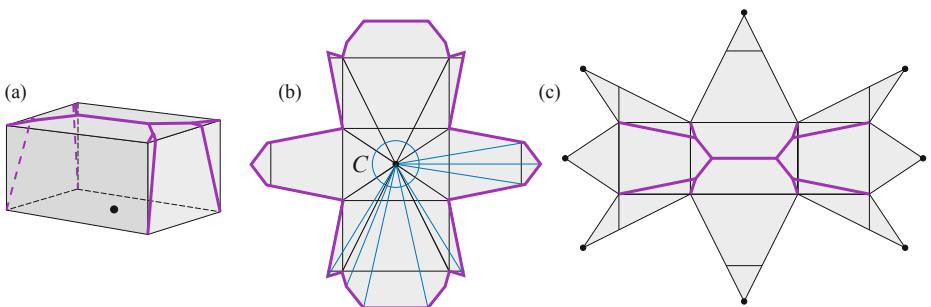
<sup>1</sup> MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, USA  
[edemaine@mit.edu](mailto:edemaine@mit.edu)

<sup>2</sup> David R. Cheriton School of Computer Science, University of Waterloo, Canada  
[alubiw@uwaterloo.ca](mailto:alubiw@uwaterloo.ca)

**Abstract.** We present a new method for unfolding a convex polyhedron into one piece without overlap, based on shortest paths to a convex curve on the polyhedron. Our “sun unfoldings” encompass source unfolding from a point, source unfolding from an open geodesic curve, and a variant of a recent method of Itoh, O’Rourke, and Vilcu.

## 1 Introduction

The easiest way to show that any convex polyhedron can be unfolded is via the *source unfolding* from a point  $s$ , where the polyhedron surface is cut at the *ridge tree* of points that have more than one shortest path to  $s$ , [10], or see [3]. The unfolding does not overlap because the shortest paths from  $s$  to every other point on the surface develop to straight lines radiating from  $s$ , forming a star-shaped unfolding. See Figure 1(b).



**Fig. 1.** [based on O’Rourke [9]] Unfolding a box from a point on the middle of the base: (b) source unfolding with some shortest paths shown. The source unfolding is the same as the sun unfolding relative to circle  $C$ . (c) star unfolding, with ridge tree shown.

Our main result is a generalized unfolding, called a *sun unfolding*, that preserves the property that shortest paths emanate in a radially monotone way,

\* Dedicated to Ferran Hurtado on the occasion of his 60th birthday.

although they no longer radiate from a point. We begin with an easy generalization where the point is replaced by a curve  $S$  that unfolds to a straight line segment (an *open geodesic curve*). Cutting at the ridge tree of points that have more than one shortest path to  $S$  produces an unfolding in which the shortest paths from  $S$  radiate from the unfolded  $S$ , so the unfolding does not overlap; see Figure 2.

For our general sun unfolding, the paths emanate radially, not from a point or a segment, but from a tree  $S$ , and the paths are not necessarily shortest paths from  $S$ . We define both  $S$  and the paths based on a convex curve  $C$  on the surface of the polyhedron. Let  $S$  be the ridge tree of  $C$  on the convex side and let  $R$  be the ridge tree of  $C$  on the other side. Let  $\mathcal{G}$  be the set of all shortest paths to  $C$ , where we glue together any paths that reach the same point of  $C$  from opposite sides. We prove that the paths emanate in a *radially monotone way* from the unfolded  $S$ , and hence that the polyhedron unfolds into a non-overlapping planar surface if we make the following cuts: cut  $R$  and, for every vertex  $v$  on the convex side of  $C$ , cut a shortest path from  $v$  to  $C$  and continue the cut across  $C$ , following a geodesic path, until reaching  $R$ . See Figure 3.

Our result generalizes source unfolding from a point or an open geodesic, by taking  $C$  to be the locus of points at distance  $\varepsilon$  from the source. See the curve  $C$  in Figure 1(b) and Figure 2(c). Our result is related to recent work of Itoh, O'Rourke, and Vilcu on “star unfolding via a quasigeodesic loop” [6]. In the remainder of this section we discuss the relationship between these results.

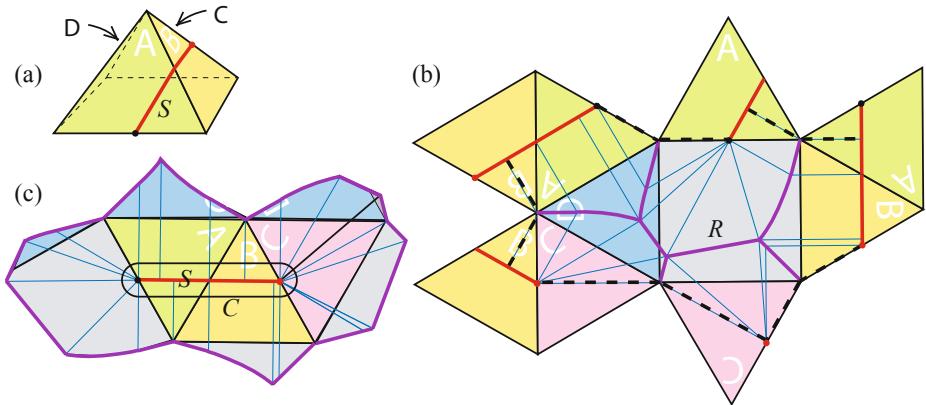
## Related Work

A *quasigeodesic loop* on the surface of a polyhedron is a closed polygonal curve such that every point on the curve except one has surface angle at most  $\pi$  on each side. Quasigeodesic loops are a special case of convex curves. Given a quasigeodesic loop  $Q$ , Itoh et al. [6] showed that cutting from every vertex to  $Q$  yields an unfolding of each half of the polyhedron, and furthermore that the two unfolded pieces can be joined together. See Figure 4. Our sun unfolding of this example is shown in Figure 4(c). The convex half unfolds the same way, but the non-convex half of the surface is cut into pieces and attached around the other half. Note that Itoh et al. call their unfolding a “star” unfolding from the quasigeodesic loop—they regard the quasigeodesic loop as the generalization of the point from which the star unfolding is defined. Our emphasis is different. We focus on the way that shortest paths emanate radially from the ridge tree  $S$ , so we regard the ridge tree as the generalization of the point from which the source unfolding is defined. Itoh et al. prove their unfolding result using Alexandrov’s work. We take a more self-contained approach of applying induction as the convex curve shrinks/expands on the surface of the polyhedron.

Itoh, O’Rourke, and Vilcu also have an interesting alternative unfolding where the convex curve  $C$  remains connected (developing as a path) while  $S$  and  $R$  are cut [5,7,8]. This is possible only for special convex polygonal curves.

## 2 Source Unfolding from an Open Geodesic

In this section we generalize source unfolding from a point  $s$  to source unfolding from an *open geodesic curve*—a path on the surface of  $P$  that does not self-intersect, starts and ends at distinct points that are not vertices, and is a locally shortest path between those endpoints. See Figure 2. Although this generalization is quite easy to prove, it seems to be a new observation.



**Fig. 2.** Source unfolding from an open geodesic: (a) a pyramid with an open geodesic curve  $S$  crossing two faces; (b) the ridge tree  $R$  lies in two faces, the base and face  $D$ . (The dashed lines, together with segments of  $S$  delimit a “dual” unfolding where the paths are attached to  $R$ .); (c) the source unfolding showing paths emanating radially from the open geodesic, and showing the convex curve  $C$  relevant to sun unfolding.

**Lemma 1.** *Let  $S$  be an open geodesic curve on the surface of a convex polyhedron  $P$ , and let  $R$  be the ridge tree of points on  $P$  that have more than one shortest path to  $S$ . Then cutting along  $R$  produces an unfolding of  $P$  that does not overlap.*

*Proof.* We will assume that  $R$  is a tree that includes all vertices of  $P$ —this follows from our general result below, or can be proved as Sharir and Schorr [10] do for the case where  $S$  is a point. Therefore  $R$  unfolds  $P$  to a planar surface. Because  $S$  is a geodesic, it unfolds to a straight line segment. Consider any point  $p$  on  $P$  that is not in  $R$ . Then there is a unique shortest path from  $p$  to  $S$ . We claim that this path reaches  $S$  in one of two ways: (1) it reaches an interior point of  $S$  and makes a right angle with  $S$ ; or (2) it reaches an endpoint of  $S$  and makes an angle between  $\pi/2$  and  $3\pi/2$  with  $S$ . The reason is that a path that reaches  $S$  in any other way can be shortened.

Observe that these shortest paths spread out in a radially monotone way from the unfolded  $S$ . See Figure 2. Thus we have an unfolding of  $P$  that does not overlap.  $\square$

### 3 Sun Unfolding

We define *sun unfolding* of a convex polyhedron  $P$  relative to a closed convex curve  $C$  on  $P$ . We will prove our unfolding result only for curves composed of a finite number of line segments and circular arcs, but will discuss more general convex curves. The curve  $C$  splits  $P$  into two “halves”, the convex or *interior* side  $C_I$ , and the *exterior* side  $C_E$ . For a point  $c$  on  $C$  (notated  $c \in C$ ), let  $\alpha_I(c)$  be the surface angle of  $C_I$  between the left and right tangents at  $c$ , and let  $\alpha_E(c)$  be the surface angle of  $C_E$  between those tangents. Then  $\alpha_I(c) + \alpha_E(c) \leq 2\pi$ , with equality unless  $c$  is a vertex of  $P$ . Also,  $\alpha_I(c) \leq \pi$ . A point  $c$  with  $\alpha_I(c) < \pi$  is called an *internal corner* of  $C$ . A point  $c \in C$  with  $\alpha_E(c) < \pi$  is called an *external convex corner* of  $C$ . If a point  $c \in C$  is a vertex then it is an internal corner or an external convex corner (or both); and if  $c$  is not a vertex, then it is not an external convex corner. See Figure 5.

The *ridge tree* (a.k.a. “cut locus”) in  $C_I$  [or  $C_E$ ] is the closure of the set of points that have more than one shortest path to  $C$ . Let  $S$  be the ridge tree of  $C$  in  $C_I$ , and let  $R$  be the ridge tree in  $C_E$ . Among all the shortest paths from points of  $C_I$  to  $C$ , let  $\mathcal{G}_I$  be the maximal ones. Among all the shortest paths from points of  $C_E$  to  $C$ , let  $\mathcal{G}_E$  be the maximal ones. If  $c \in C$  has  $\alpha_I(c) = \alpha_E(c) = \pi$ , then we concatenate together the unique paths of  $\mathcal{G}_I$  and  $\mathcal{G}_E$  that are incident to  $c$ . Let  $\mathcal{G}$  be the resulting set of paths, together with any leftover paths of  $\mathcal{G}_I$  and any leftover paths of  $\mathcal{G}_E$ . For example, a leftover path of  $\mathcal{G}_I$  reaches  $c_3$  in Figure 5; observe that  $c_3$  is in  $R$ . The figure also shows examples of leftover paths of  $\mathcal{G}_E$  reaching points  $c_2$  and  $c_4$ ; observe that  $c_2$  and  $c_4$  are in  $S$ .

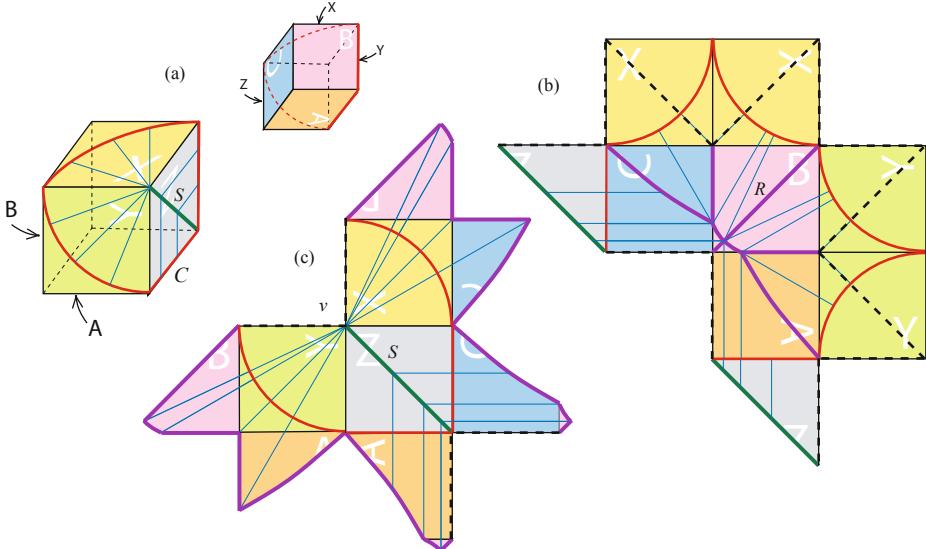
**Lemma 2.** *Both  $R$  and  $S$  are trees. Every vertex of  $P$  lies in  $R$  or  $S$  (or both). Every internal corner of  $C$  is a leaf of  $S$ . Every external convex corner of  $C$  is a leaf of  $R$ . Every path of  $\mathcal{G}$  goes from  $S$  to  $R$  and includes a point of  $C$ . The surface of  $P$  is covered by  $S$ ,  $R$ , and  $\mathcal{G}$ . Furthermore, any point not on  $S$  or  $R$  is in a unique path of  $\mathcal{G}$ .*

The proof of the lemma is in Section 4.

Let  $v$  be a vertex of  $P$ . If  $v$  is not in  $R$ , then it is in  $S$ , and we let  $\gamma(v)$  be a path of  $\mathcal{G}$  incident to  $v$ . The choice of  $\gamma(v)$  is not unique in general, but we fix one  $\gamma(v)$ . Observe that each  $\gamma(v)$  is a path from  $v$  to  $R$ , consisting of a shortest path from  $v$  to  $C$  possibly continued geodesically to  $R$ . We define *sun cuts* with respect to  $C$  to consist of  $R$  and the paths  $\gamma(v)$ , for  $v$  a vertex of  $P$  in  $C_I \cup C$ . Note that a vertex on  $C$  may be a leaf of  $R$ , in which case  $\gamma(v)$  has length 0.

**Theorem 1.** *Let  $C$  be a closed convex curve on the surface of a convex polyhedron  $P$ , such that  $C$  is composed of a finite number of line segments and circular arcs. Then sun cuts with respect to  $C$  unfold the surface of  $P$  into the plane without overlap.*

To prove the theorem, we first show that the sun cuts form a tree that reaches all vertices of  $P$ —hence the surface unfolds to the plane. To show that the unfolded

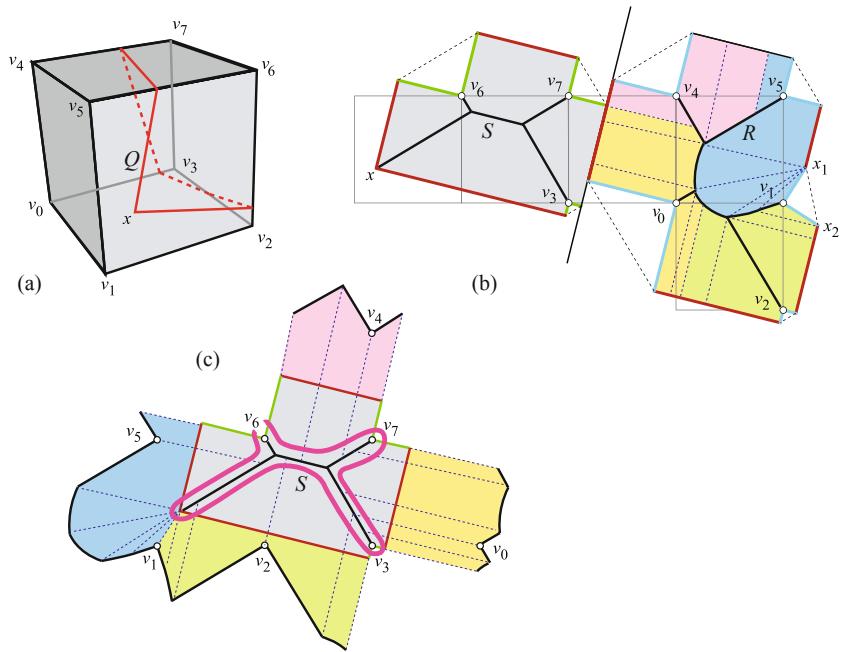


**Fig. 3.** Sun unfolding with respect to a convex curve: (a) a cube with a convex curve  $C$  and the ridge tree  $S$  on the convex side; (b) the ridge tree  $R$  on the non-convex side of  $C$ . (The dashed lines delimit a “dual” unfolding where the paths are attached to  $R$ .); (c) the sun unfolding showing paths emanating radially from  $S$ . Note that the two vertices on  $S$  require cuts  $\gamma(\cdot)$  to  $R$ , shown with dashed lines.

surface does not overlap, we prove, by shrinking  $C$  and applying induction, that  $S$  unfolds without overlap and that the paths of  $\mathcal{G}$  emanate from the unfolded  $S$  in a *radially monotone way*, defined as follows. Make a tour clockwise around the unfolded  $S$ , travelling in the plane an infinitesimal distance away from the unfolded  $S$ . See Figure 4(c). Parts of the tour are off the unfolded surface  $P$ . In particular, whenever a cut reaches  $S$ , there will be a gap in the unfolding. Apart from the gaps, at any point of the tour we are at a point  $p$  of  $P - (S \cup R)$ , so by Lemma 2 there is a unique path  $\gamma(p) \in \mathcal{G}$  containing  $p$  and extending to  $R$ . Extend this path to a ray, and let  $f(p)$  be the corresponding point on the circle at infinity. We say that  $\mathcal{G}$  emanates from the unfolded  $S$  in a *radially monotone way* if, as  $p$  tours clockwise around  $S$ ,  $f(p)$  progresses clockwise around the circle at infinity, i.e., if  $p'' > p' > p$  along the tour then  $f(p'') \geq f(p') \geq f(p)$  clockwise around the circle at infinity.

## 4 Sun Unfolding: Proofs

We begin with a section about the local properties of shortest paths to  $C$ . Following that is a section on properties of the ridge trees, and in the final section we prove that the sun unfolding does not overlap.



**Fig. 4.** The sun unfolding with respect to a geodesic loop  $Q$  on a cube, based on an example from Fig. 1 of Itoh et al. [6]: (a) the cube and the quasigeodesic loop  $Q$ ; (b) the ridge trees  $S$  and  $R$  on the two sides of the curve  $Q$  (superimposed on the unfolding from [6]); (c) the sun unfolding with respect to  $Q$ , showing a tour around  $S$  and the paths emanating in a radially monotone way from  $S$

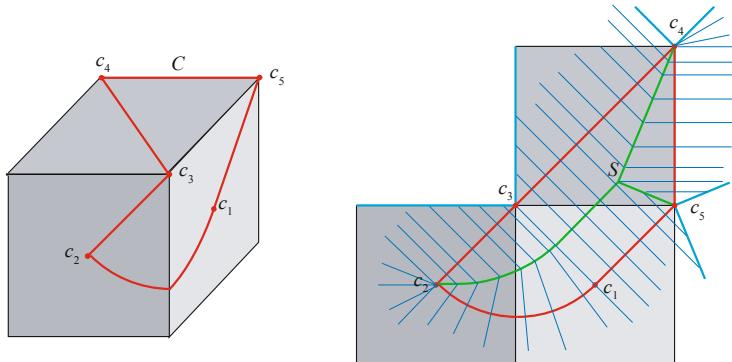
#### 4.1 Structure of Shortest Paths

*Claim.* Two shortest paths do not cross. [6, Lemma 2]. A shortest path reaches  $C$  at an angle greater than or equal to  $\pi/2$ .

We need the second statement only for the case where  $C$  is composed of a finite set of straight line segments and circular arcs, where it seems obvious (a path reaching  $C$  at an angle less than  $\pi/2$  can be locally shortened). A more general case is proved in [4].

As a consequence, we make the following observations about how shortest paths behave with respect to corners and reflex points of a convex curve  $C$ . See Figure 5. Let  $c$  be a point of  $C$ , and let  $s$  be a side, i.e.,  $s = I$  or  $E$ .

- If  $\alpha_s(c) = \pi$  then there is exactly one path in  $\mathcal{G}_s$  to point  $c$ .
- If  $\alpha_s(c) < \pi$  then there is no path of  $\mathcal{G}_s$  to point  $c$ .
- If  $\alpha_E(c) > \pi$  then there is a wedge of paths of  $\mathcal{G}_E$  to point  $c$ . In this case,  $\alpha_I(c)$  must be less than  $\pi$ .



**Fig. 5.** A convex curve  $C$  on the surface of a cube (*left*), and part of the sun unfolding with respect to  $C$  (*right*) showing the ridge tree  $S$  and some of the paths of  $\mathcal{G}$ . The points  $c_i$  show some possible vertex/corner/reflex point combinations.

## 4.2 Structure of the Ridge Trees

We believe that the sun unfolding works for any convex curve  $C$ , but our proof technique requires that the ridge trees of the curve  $C$  are trees, not only in the sense of being connected and acyclic, but also in the sense of having a finite number of nodes. Even for a closed curve in the plane, the ridge tree (which is then the *medial axis*) need not be finite: Choi, Choi and Moon [2] give an example of a smooth curve (in fact  $C^\infty$ ) whose medial axis has an infinite number of leaves. Their curve alternates infinitely often between convex and concave parts, but it is also possible to give an example of a convex (non-smooth) planar curve whose ridge tree has infinitely many leaves.

We will prove that the ridge trees are finite, and that the sun unfolding works, for any convex curve  $C$  composed of a finite number of line segments and circular arcs. We will discuss weaker conditions that probably suffice.

Our proofs use a “wavefront expansion” where we offset the curve  $C$  by distance  $d \geq 0$ , forming curve  $C_d$ . As  $d$  increases (and  $C$  “shrinks”) we stop at “events” where the curve changes in an essential way. The ridge tree of  $C$  consists of the ridge tree of  $C_d$  together with the portion of the ridge tree in the “band” between  $C$  and  $C_d$ , so we obtain our results by induction as long as the number of events is finite. This is why we need a finite ridge tree.

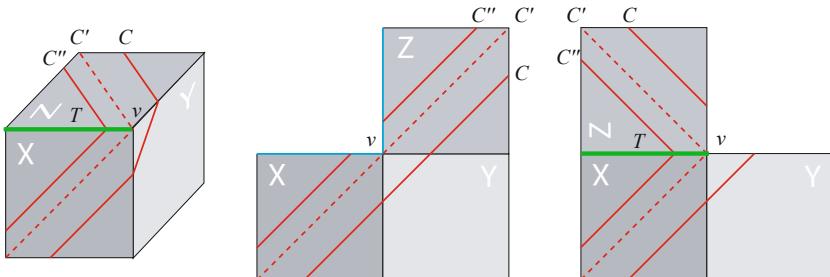
Shortest paths reach  $C$  along normals to the curve, which are maintained as  $C$  shrinks. At a point of  $C$  with surface angle less than  $\pi$  (a *corner*), the normals to the curve at either side of the point meet and form an edge of the ridge tree. As  $C$  shrinks, the corner traces out the edge of the ridge tree. At a point of  $C$  with surface angle greater than  $\pi$  (a “reflex” point), there is a wedge of normals reaching the point. For any  $d > 0$  the shrunken curve  $C_d$  has a circular arc in place of the point. Thus reflex points vanish immediately. We prove below that no reflex points arise during the shrinking process.

Therefore the general structure is that  $C$  is a curve with surface angle less than or equal to  $\pi$  at every point. The corners (where the surface angle is less than  $\pi$ ) trace out the edges of the ridge tree as  $C$  shrinks. For our application, we have a convex curve that we shrink both inside and outside. With respect to the inside, all points other than corners have positive curvature; with respect to the other side, for any  $d > 0$ , all points of  $C_d$  other than corners have negative curvature. Furthermore, for the special case of curves composed of a finite number of line segments and circular arcs, there are finitely many corners.

**Events.** As  $C$  shrinks, the possible events are as follows. Note that more than one event may happen simultaneously, but we claim that we can handle them one at a time.

**Closing Event.** The curve closes up and vanishes. It may vanish at a point, or more generally, at a tree, which forms part of the ridge tree. In case  $C$  is the inside or outside of a convex curve,  $C$  can only collapse to a point or a straight line segment.

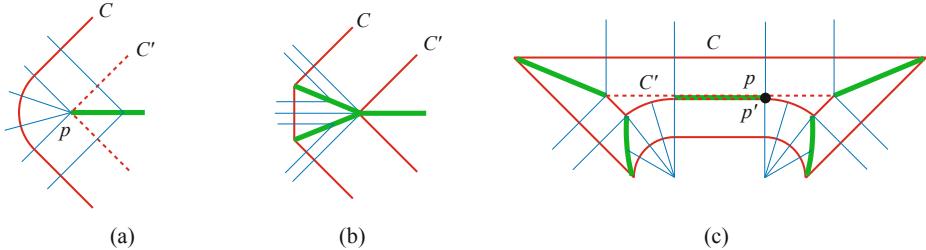
**Vertex Event.** The curve encounters a vertex  $v$  of  $P$ . See Figure 6. As the curve shrinks past  $v$  the surface angle of the curve at  $v$  decreases by the angular defect of  $v$ . The surface angle was at most  $\pi$  before hitting  $v$ , and is strictly less than  $\pi$  afterwards. Thus there is a corner at  $v$  after the event. It is a new corner except in the degenerate situation when an existing corner hits  $v$ . An edge of the ridge tree grows at the corner.



**Fig. 6.** A vertex event on a cube. The curve  $C$  shrinks to  $C'$ , forming a corner at vertex  $v$ , and then shrinks to  $C''$ , tracing out an edge of the ridge tree  $T$ .

**New Corner Event.** See Figure 7(a). In the general case a new corner appears at a convex point of the curve where the radius of the *osculating circle* becomes 0 (i.e. the curvature approaches positive infinity). In the special case of curves composed of line segments and circular arcs, a new corner appears when a circular arc shrinks to a point  $p$ . No new corners appear if  $C$  is the outside of a convex curve.

**Corner Merge Event.** See Figure 7(b). Two edges of the ridge tree meet and a new edge forms. In degenerate cases, it is possible that more than two edges of the ridge tree meet.



**Fig. 7.** Events: (a) a new corner event; (b) a corner merge event; (c) a pinch event. Thick lines indicate the ridge tree, and thin lines indicate paths of  $\mathcal{G}$ .

**Pinch Event.** The curve meets itself, but does not vanish. Figure 7(c) shows the curve meeting itself along a segment. Suppose point  $p$  becomes coincident with point  $p'$ , but no part of the curve from  $p$  to  $p'$  (traversed clockwise, say) becomes coincident.

**Lemma 3.** *In the closed curve from  $p$  to  $p'$ , point  $p = p'$  is a corner with surface angle 0 and turn  $\pi$ . Furthermore, there must be negative curvature arbitrarily close to  $p$  or  $p'$  (or both).*

*Proof.* Every point on the curve has surface angle at most  $\pi$ . Thus the surface angle of the original curve must be exactly  $\pi$  at  $p$  and at  $p'$  (because if one were less than  $\pi$  the other would have to be greater than  $\pi$ ). This implies that after the collapse, point  $p = p'$  has surface angle 0 and therefore turn  $\pi$ .

Now consider the curvature. If  $p$ , say, has positive curvature, then  $p'$  must have negative curvature, which completes the proof. So suppose some  $\varepsilon$ -interval of the curve from  $p$  has 0 curvature, and some  $\varepsilon'$ -interval of the curve from  $p'$  has 0 curvature. Then the curves are line segments in those intervals, and must be coincident within the smaller interval. Note that for the case of curves composed of a finite number of line segments and circular arcs, the neighbourhood argument is unnecessary— $p$  or  $p'$  must have negative curvature.  $\square$

Therefore, for a convex curve, a pinch event can only happen on the outside, and cannot involve three or more points of  $C$  meeting at the same point. The curve may become coincident along a segment, in which case the claim applies to both endpoints of the common segment in the two resulting closed curves.

**Ridge Trees.** We have a convex curve  $C$  that shrinks to the inside, forming ridge tree  $S$  in  $C_I$ , and to the outside, forming ridge tree  $R$  in  $C_E$ . On the outside, new corner events do not occur because there are no convex points apart from the corners. We claim that the ridge tree is therefore a [finite] tree—see Lemma 4 below. On the inside, Lemma 3 implies that pinch events do not occur, and, so long as new corner events happen only a finite number of times, we claim that the ridge tree is a [finite] tree—see Lemma 5 below. We will deal explicitly only

with the case where  $C$  is composed of a finite number of line segments and circular arcs, which guarantees a finite number of new corner events.

**Lemma 4.** *R is a [finite] tree. R contains all the vertices of P that lie in  $C_E$ , and R has a leaf at each external convex corner of C.*

We prove the lemma only for the case where  $C$  is a convex curve composed of a finite number of line segments and circular arcs, but we believe it holds for convex curves generally.

*Proof.* For the proof we will change our frame of reference with respect to inside/outside. We shrink  $C$  into  $C_E$ , and will now refer to that as “inside”. We will also say “corner” rather than “external convex corner.”

We shrink the curve and argue by induction on  $2n + c$  where  $n$  is the number of vertices lying inside  $C$ , and  $c$  is the number of corners of  $C$ . We use the property that  $C$  is composed of a finite number of line segments and circular arcs of negative curvature joined at points with surface angle at most  $\pi$ , and note that this holds after each event.

Shrink  $C$  to  $C'$  at the next event. As noted above in the general case, new corner events only occur when the curvature at some point approaches positive infinity. Because our curve has negative curvature except at the corners, there are no new corner events. We are left with four possible events:

**Closing Event.**  $C'$  is either a point or a line segment, since two circular arcs of negative curvature cannot collapse together.  $C'$  forms its own ridge tree, and the portions of the ridge tree edges between  $C$  and  $C'$  attach to it, which gives the desired result for  $C$ .

**Vertex Event.** As noted above, there is a corner at vertex  $v$  after the event, and it is a new corner except in the degenerate situation when an existing corner hits  $v$ . If the point of the curve that hits  $v$  is an internal point of a line segment or circular arc, then the segment or arc splits in two. The properties of the curve are maintained. Observe that  $n$  goes down by 1 and  $c$  goes up by at most 1. Therefore the quantity  $2n + c$  goes down, and by induction the ridge tree of  $C'$  is a tree that contains each vertex interior to  $C'$  and has a leaf at each corner of  $C'$ . The portions of the ridge tree edges between  $C$  and  $C'$  attach to the leaves of the ridge tree at the corners of  $C'$ , which gives the desired result for  $C$ .

**Corner Merge Event.** In this case  $c$  goes down by at least 1 and  $n$  stays the same. The quantity  $2n + c$  goes down and the result follows by induction.

**Pinch Event.** By Lemma 3,  $C'$  consists of two closed curves  $C'_1$  and  $C'_2$ . The vertices interior to  $C$  are partitioned between  $C'_1$  and  $C'_2$ , as are the corners of  $C$ . Also, by Lemma 3, each of  $C'_1$  and  $C'_2$  has one new corner, say at points  $p_1$  and  $p_2$ , respectively.

We first argue that we can apply induction to  $C'_1$  and  $C'_2$ . If  $C'_i$ ,  $i = 1$  or  $i = 2$  has fewer than  $n$  vertices in its interior, then the quantity  $2n + c$  goes

down and we can apply induction. Suppose  $C'_1$  has  $n$  vertices in its interior. Then  $C'_2$  has no vertices in its interior, which implies that it lies in the plane and therefore has total turn  $2\pi$ . Furthermore, by Lemma 3, the new corner  $p_2$  of  $C'_2$  has turn  $\pi$  and at least one part of the curve incident to  $p_2$  has negative curvature. Since all points except corners have non-positive curvature, the other corners of  $C'_2$  must have turn sum greater than  $\pi$ , and therefore there must be at least two other corners. This implies that  $C'_1$  has strictly fewer corners than  $C$ . Therefore we can apply induction to  $C'_1$ .

By induction, the ridge trees of  $C'_1$  and  $C'_2$  are trees and contain all the vertices of  $P$  that lie inside  $C$ . Furthermore, the ridge trees have leaves at  $p_1$  and  $p_2$ , respectively. If  $p_1 = p_2$  (i.e., the pinch event occurred at a single point) then the two ridge trees join to form the ridge tree of  $C$ . Otherwise, the pinch event occurred along a segment, and the segment plus the two ridge trees form the ridge tree of  $C$ .  $\square$

**Lemma 5.** *S is a [finite] tree. S contains all the vertices of P that lie in  $C_I$ , and S has a leaf at each internal corner of C.*

We prove the lemma only for the case where  $C$  is a convex curve composed of a finite number of line segments and circular arcs, but we believe that the lemma holds more generally for a convex curve so long as the number of leaves of  $S$  is finite, and that a natural sufficient condition for this is that there are a finite number of points of maximal curvature along  $C$ .

*Proof.* We are focused on the inside of  $C$ , so we will just say “corner” rather than “internal corner”. The proof is by induction on  $4n + 2a + c$  where  $n$  is the number of vertices lying in the interior of  $C$ ,  $c$  is the number of corners of  $C$ , and  $a$  is the number of circular arcs of  $C$ . We prove in addition that the curve is always convex and composed of line segments and circular arcs.

Shrink  $C$  to  $C'$  at the next event. By Lemma 3, a pinch event can only occur when the curve has a point of negative curvature, and therefore pinch events do not occur when  $C$  is convex. We consider each of the possible events:

**Closing Event.** The argument is the same as in Lemma 4.

**Vertex Event.** The argument about the ridge tree is the same as in Lemma 4. Convexity is maintained. Observe that  $n$  goes down by 1,  $c$  goes up by at most 1, and  $a$  goes up by at most 1. Therefore the quantity  $4n + 2a + c$  goes down, and by induction the ridge tree of  $C'$  is a tree that contains each vertex interior to  $C'$  and has a leaf at each corner of  $C'$ . The portions of the ridge tree edges between  $C$  and  $C'$  attach to the leaves of the ridge tree at the corners of  $C'$ , which gives the desired result for  $C$ .

**New Corner Event.** This occurs only when one of the circular arcs of  $C$  shrinks to a point. The point becomes a corner of  $C'$ . Convexity is maintained. Observe that  $c$  goes up by 1,  $a$  goes down by 1, and  $n$  is unchanged. Therefore the quantity  $4n + 2a + c$  goes down, and by induction the ridge tree of  $C'$  is a tree that contains

each vertex interior to  $C'$  and has a leaf at each corner of  $C'$ . Adding back the portions of ridge tree edges between  $C$  and  $C'$  gives the result for  $C$ .

**Corner Merge Event.** In this case  $c$  goes down by at least 1,  $a$  does not go up, and  $n$  stays the same. The quantity  $4n + 2a + c$  goes down and the result follows by induction.  $\square$

Having established the structure of  $R$  and  $S$ , we now wrap up the proof of Lemma 2.

*Proof (of Lemma 2).* By Lemma 5,  $S$  is a tree that includes every vertex of  $P$  that lies in  $C_I$  and every internal corner of  $C$ . By Lemma 4,  $R$  is a tree that includes every vertex of  $P$  that lies in  $C_E$  and every external convex corner of  $C$ . It remains to show that every vertex of  $P$  lies in  $R$  or  $S$ , or both. Vertices in  $C_I$  and  $C_E$  are in the respective trees. Consider a vertex  $v$  of  $P$  that lies on  $C$ . Since the surface angle at  $v$  is less than  $2\pi$ ,  $v$  must be an internal corner or an external convex corner. Therefore  $C$  is in  $S$  or  $R$ , or both.

To complete the proof, we must establish the properties of  $\mathcal{G}$ . By definition, every path  $\gamma$  of  $\mathcal{G}$  includes a point of  $C$ . If  $\gamma$  was formed by concatenating together a path of  $\mathcal{G}_I$  and  $\mathcal{G}_E$  then  $\gamma$  goes from  $S$  to  $R$ . If  $\gamma$  is a leftover path of  $\mathcal{G}_I$  then it reaches a point  $c \in C$  with  $\alpha_E(c) < \pi$ , so  $c$  is a convex external corner of  $C$ , and  $c$  is in  $R$ . If  $\gamma$  is a leftover path of  $\mathcal{G}_E$  then it reaches a point  $c \in C$  with  $\alpha_I(c) < \pi$ , so  $c$  is an internal corner of  $C$  and  $c$  is in  $S$ . Finally, it is clear that the surface of  $P$  is covered by  $S$ ,  $R$ , and  $\mathcal{G}$ , and that any point not on  $S$  or  $R$  is in a unique path of  $\mathcal{G}$ .  $\square$

### 4.3 Non-overlap of the Sun Unfolding

**Lemma 6.** *The sun cuts with respect to a closed convex curve  $C$  form a tree incident to every vertex of  $P$ .*

*Proof.* By definition, the sun cuts consist of  $R$  together with paths  $\gamma(v)$  for  $v$  a vertex of  $P$  in  $C_I \cup C$ . By Lemma 4,  $R$  is a tree. Adding the paths  $\gamma(v)$  still gives a tree, because every such path contains exactly one point of  $R$  and no two such paths intersect—this is because the portions of the paths in  $C_I$  do not intersect, the points where the paths reach  $C$  are distinct, and thus the portions of the paths in  $C_E$  do not intersect (except possibly at  $R$ ). Finally, by Lemma 4,  $R$  includes any vertex interior to  $C_E$  and, by construction, we add cuts reaching every vertex in  $C_I \cup C$ .  $\square$

As discussed in Section 3, we will prove that the sun cuts unfold  $P$  without overlap by proving that the paths  $\mathcal{G}$  emanate from the unfolded  $S$  in a radially monotone way. We will use an equivalent formulation of radial monotonicity. Let  $\bar{\mathcal{G}}$  be the infinite rays in the plane that extend paths of  $\mathcal{G}$  from their point of origin on the unfolded  $S$ . Radial monotonicity is equivalent to the property that two rays of  $\bar{\mathcal{G}}$  do not intersect except at a common point of origin. The main work, in Lemma 7 below, is to prove this for the interior of  $C_I$  and for the

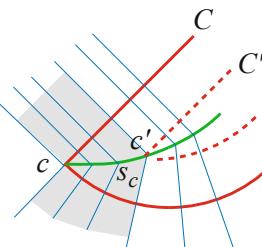
rays  $\bar{\mathcal{G}}_I$  that originate from points of  $S$  interior to  $C_I$ . Note that the sun cuts, restricted to  $C_I$ , consist of cuts from every vertex inside  $C_I$  along a shortest path to  $C$ . Itoh et al. [6] call this the “star unfolding” and prove that it develops without overlap for quasigeodesic loops. We need more general convex curves and we need the stronger result about radial monotonicity.

**Lemma 7.** *Two rays of  $\bar{\mathcal{G}}_I$  do not intersect except at a common point of origin. Consequently, the sun cuts unfold the interior of  $C_I$  without overlap.*

*Proof.* We prove the result by induction as  $C$  shrinks. We only treat the case where  $C$  consists of a finite number of line segments and circular arcs. The proof follows the same structure as the proof of Lemma 5. In particular, the induction is on the quantity  $4n + 2a + c$ , and we will not repeat the arguments about why we can apply induction at an event.

Recall that curve  $C$  shrinks to curve  $C'$  at the next event, and that there are no pinch events because  $C$  is convex. The ridge tree  $S$  consists of the ridge tree  $S'$  of  $C'$  together with the part of  $S$  that lies in the “band” between  $C$  and  $C'$ . Let  $\bar{\mathcal{G}}'$  denote the rays of  $\bar{\mathcal{G}}$  that are incident to  $S'$ . Let  $U'$  denote the unfolding of the interior of  $C'$  with the extended rays  $\bar{\mathcal{G}}'$ . Let  $U$  denote the analogous structure for  $C$ . By induction, we know that  $U'$  does not overlap itself. We want to prove the same for  $U$ .

To get  $U$  from  $U'$  we need to add the pieces of  $S$  that lie in the band between  $C$  and  $C'$  and the rays of  $\bar{\mathcal{G}}$  that originate from these new pieces of  $S$ . Let  $s_c$  be the portion of a ridge tree edge that is traced out by a corner  $c$  of  $C$  as it shrinks to corner  $c'$  of  $C'$ . See Figure 8. Let  $\bar{\mathcal{G}}_c$  be the rays of  $\bar{\mathcal{G}}$  that originate at points of  $s_c$ . The rays of  $\bar{\mathcal{G}}_c$  that originate from the point  $c'$  are the two normals to  $C'$  at  $c'$ . Because  $C$  is convex, the other rays of  $\bar{\mathcal{G}}_c$  unfold to lie in a wedge bounded by these normals. We need to show that this wedge fits into the unfolding  $U'$ . Note that  $c'$  is locally flat (we consider vertex events below). Also note that, because  $c'$  is a corner of  $C'$ , no shortest path from the interior of  $C'$  reaches  $C'$  at  $c'$ . This means that no sun cut reaches  $c'$ , and hence a neighbourhood of  $c'$  is intact in the unfolding  $U'$ . Therefore  $U'$  has a wedge-shaped gap at  $c'$  that exactly accommodates  $\bar{\mathcal{G}}_c$  (including the segment  $s_c$ ).



**Fig. 8.** A piece,  $s_c$ , of the ridge tree traced out by corner  $c$  of  $C$  as it shrinks to corner  $c'$  of  $C'$ . The shaded area indicates the rays of  $\bar{\mathcal{G}}_c$ .

The situations at a new corner event and a corner merge event are similar. Refer back to Figure 7(a) and (b). In both cases the event results in a corner  $c'$  of  $C'$ . There is a piece  $s_c$  of  $S - S'$  attached to  $c'$  (possibly just the point  $c'$ ), and some rays of  $\bar{\mathcal{G}}$  originate from points of  $s_c$ . By local planarity, these fit into the gap at  $c'$  in the unfolding  $U'$ .

At the closing event, the convex curve  $C$  collapses to a point or line segment, which becomes the ridge tree. The paths of  $\bar{\mathcal{G}}$  radiate outward from the ridge tree without intersections.

Finally, we consider a vertex event. Suppose  $C'$  goes through vertex  $v$  of  $P$ . Then the sun cuts include a shortest path from  $v$  to  $C$ . The unfolding  $U'$  has a gap at  $v$  with angle equal to the angular deficit at  $v$ . The unfolding  $U$  has no extra surface in the neighbourhood of  $v$ , and is therefore the same as  $U'$  in that neighbourhood.  $\square$

*Proof (of Theorem 1).* We prove that sun cuts unfold  $P$  without overlap by proving that no two rays of  $\bar{\mathcal{G}}$  intersect except at a common point of origin. Lemma 7 proves this for  $\bar{\mathcal{G}}_I$ . It remains to add the rays of  $\bar{\mathcal{G}}$  originating from points of  $S$  that lie on  $C$ .

Let  $c$  be a leaf of  $S$  lying on  $C$  that has paths of  $\mathcal{G}$  incident to it. Then  $c$  is a point with  $\alpha_I(c) < \pi$  and  $\alpha_E(c) > \pi$ . As in the proof of Lemma 7, a neighbourhood of  $c$  in  $C_I$  is intact in the unfolding, and has a wedge-shaped gap large enough for the wedge of rays of  $\bar{\mathcal{G}}$  originating at  $c$ . In case  $c$  is a vertex, one of the rays of  $\bar{\mathcal{G}}$  incident to  $c$  is in the set of sun cuts, so the rays of  $\bar{\mathcal{G}}$  originating at  $c$  are split into two wedges with a gap between them equal to the angular defect of the vertex.  $\square$

## 5 Unified Unfolding

The sun unfolding generalizes one of the basic unfolding methods for convex polyhedra, namely the source unfolding from a point. The other basic unfolding method is the *star unfolding* from a point  $s$ , where the polyhedron surface is cut along a shortest path from every vertex to  $s$  [1]. See Figure 1(c). This is dual to the source unfolding in that the shortest paths are attached in one case to  $s$  (for source unfolding) and in the other case to the ridge tree (for star unfolding). The dual of our sun unfolding would be to attach the paths of  $\mathcal{G}$  to the ridge tree  $R$  and cut the ridge tree  $S$  and paths of  $\mathcal{G}$  from vertices to  $S$ . See for example Figure 2(b) and Figure 3(b). We conjecture that this unfolds without overlap. In the general dual case the paths  $\mathcal{G}$  do not emanate in a radially monotone way from  $R$ , so a new proof technique will be needed. A first step would be to prove that the star unfolding from an open geodesic unfolds without overlap, i.e. to prove the dual version of Lemma 1.

**Acknowledgments.** We thank Joseph O'Rourke for helpful discussions.

## References

1. Aronov, B., O'Rourke, J.: Nonoverlap of the star unfolding. *Discrete and Computational Geometry* 8, 219–250 (1992)
2. Choi, H.I., Choi, S.W., Moon, H.P.: Mathematical theory of medial axis transform. *Pacific Journal of Mathematics* 181, 57–88 (1997)
3. Demaine, E.D., O'Rourke, J.: *Geometric Folding Algorithms: Linkages, Origami, Polyhedra*. Cambridge University Press (2007)
4. Ieiri, K., Itoh, J.-I., Vilcu, C.: Quasigeodesics and farthest points on convex surfaces. *Advances in Geometry* 11(4), 571–584 (2012)
5. Itoh, J.-I., O'Rourke, J., Vilcu, C.: Source unfoldings of convex polyhedra with respect to certain closed polygonal curves. In: Proc. 25th European Workshop Comput. Geom. (EuroCG), pp. 61–64 (2009)
6. Itoh, J.-I., O'Rourke, J., Vilcu, C.: Star unfolding convex polyhedra via quasi-geodesic loops. *Discrete and Computational Geometry* 44, 35–54 (2010)
7. O'Rourke, J., Vilcu, C.: Conical existence of closed curves on convex polyhedra. *CoRR*, abs/1102.0823 (2011)
8. O'Rourke, J., Vilcu, C.: Development of curves on polyhedra via conical existence. In: Proc. 23rd Canadian Conference on Computational Geometry (CCCG), pp. 71–76 (2011)
9. O'Rourke, J.: Unfolding polyhedra (July 2008)
10. Sharir, M., Schorr, A.: On shortest paths in polyhedral spaces. *SIAM Journal on Computing* 15(1), 193–215 (1986)

# Large Angle Crossing Drawings of Planar Graphs in Subquadratic Area<sup>\*</sup>

Patrizio Angelini<sup>1</sup>, Giuseppe Di Battista<sup>1</sup>, Walter Didimo<sup>2</sup>, Fabrizio Frati<sup>3</sup>,  
Seok-Hee Hong<sup>3</sup>, Michael Kaufmann<sup>4</sup>, Giuseppe Liotta<sup>2</sup>, and Anna Lubiw<sup>5</sup>

<sup>1</sup> Dipartimento di Informatica e Automazione - Università Roma Tre, Italy  
[{angelini,gdb}@dia.uniroma3.it](mailto:{angelini,gdb}@dia.uniroma3.it)

<sup>2</sup> Dip. di Ingegneria Elettronica e dell'Informazione - Perugia University, Italy  
[{walter.didimo,liotta}@diei.unipg.it](mailto:{walter.didimo,liotta}@diei.unipg.it)

<sup>3</sup> School of Information Technologies - University of Sydney, Australia  
[{brillo,shhong}@it.usyd.edu.au](mailto:{brillo,shhong}@it.usyd.edu.au)

<sup>4</sup> Wilhelm-Schickard-Institut für Informatik - Universität Tübingen, Germany  
[mk@informatik.uni-tuebingen.de](mailto:mk@informatik.uni-tuebingen.de)

<sup>5</sup> Cheriton School of Computer Science - University of Waterloo, Canada  
[alubiw@uwaterloo.ca](mailto:alubiw@uwaterloo.ca)

**Abstract.** This paper describes algorithms for computing non-planar drawings of planar graphs in subquadratic area such that: (i) edge crossings are allowed only if they create large angles; (ii) the maximum number of bends per edge is bounded by a (small) constant.

## 1 Introduction

Constructing planar drawings of planar graphs in small area is one of the most studied topics in graph drawing. Several drawing conventions have been investigated with this respect, e.g., straight-line drawings, poly-line drawings, orthogonal drawings, visibility representations, and many more [7, 17]. Quadratic upper bounds are known on the area requirements of planar graphs in each of such drawing conventions (see, e.g., [5, 6, 19–21]). However, despite the variety of considered drawing styles, no model ensuring high readability of the drawings has been so far defined in which planar graph drawings can be constructed in sub-quadratic area. Quadratic area lower bounds for planar drawings of graphs can be found, e.g., in [6, 14, 22]. Planar graphs and, more in general, graphs with constant chromatic number admit straight-line non-planar drawings in linear area [23].

In this paper we show subquadratic area bounds for drawing planar graphs with respect to two recently introduced drawing standards, called *Right Angle Crossing drawings (RAC drawings)* and *Large Angle Crossing drawings (LAC drawings)*.

A few years ago, cognitive experiments in graph visualization have been performed showing that the human understanding of a graph layout is negatively affected by edge

---

\* Work partially supported by the Italian Ministry of Research, project AlgoDEEP 2008TF-BWL4, and by the ESF project 10-EuroGIGA-OP-003 “Graph Drawings and Representations”. Work on these results began at the 5th Bertinoro Workshop on Graph drawing. Discussion with other participants is gratefully acknowledged.

crossings that form acute angles [15, 16]. As a consequence, a large body of papers has been recently devoted to the study of non-planar drawings where such “sharp angle crossings” are forbidden.

An *RAC drawing* is a drawing where edges cross only at right angles and an *LAC<sub>α</sub> drawing* is a drawing where edges cross only at angles larger than some constant  $\alpha$  between 0 and  $\pi/2$ . The *curve complexity* of a drawing is the maximum number of bends of any edge in the drawing; e.g., straight-line drawings have curve complexity 0. RAC and LAC drawings have been formally introduced in [9] (later also published in [10]) and in [8, 13], respectively.

Research on RAC and LAC drawings has so far been conducted in three main directions: (1) Determining the maximum number of edges that a graph admitting an RAC/LAC drawing can have; (2) determining the time complexity of deciding whether a graph admits an RAC/LAC drawing; (3) determining the area requirements of RAC and LAC drawings of several graph classes.

Concerning the extremal question of determining the maximum number of edges a graph admitting an RAC/LAC drawing can have, it is known from [9] that a graph admitting a straight-line RAC drawing has at most  $4n - 10$  edges (and this bound is tight); moreover, if a graph admits an RAC drawing with curve complexity 1 or 2, then it has  $O(n)$  edges [4]; furthermore, if a graph admits a straight-line LAC<sub>α</sub> drawing, then it has at most  $\pi(3n - 6)/\alpha$  edges [13], while if it admits an LAC drawing with curve complexity 1 or 2 in which each crossing angle is exactly  $\alpha$ , then it has  $O(n)$  edges [1]; finally, every graph admits an RAC drawing with curve complexity 3, hence there are graphs with  $O(n^2)$  edges admitting RAC drawings with curve complexity 3 [9].

Concerning the algorithmic question of deciding whether a graph admits an RAC drawing, NP-hardness results are known. This has been in fact first proved for straight-line upward RAC drawings of directed graphs [2] and later extended to straight-line RAC drawings of undirected graphs [3].

Concerning the area requirements of RAC/LAC drawings, it is known that every graph admits an RAC drawing with curve complexity 3 in  $O(n^4)$  area [9], and that every graph admits an RAC drawing with curve complexity 4 in  $O(n^3)$  area [8]. In [8] it is also proved that every graph admits an LAC<sub>α</sub> drawing with curve complexity 1 in  $O(n^2)$  area, for any fixed  $\alpha \in (0, \pi/2)$ , which is asymptotically optimal. Moreover, every graph with vertex degree at most 6 (at most 3) admits an RAC drawing with curve complexity 2 (curve complexity 1, resp.) in  $O(n^2)$  area [2]. In [2] it is also proved that straight-line RAC drawings of planar graphs may require  $\Omega(n^2)$  area.

In this paper we investigate RAC and LAC<sub>α</sub> drawings of graphs in small area and constant curve complexity. Our main contributions are the following:

- In Sect. 3 we describe an algorithm for constructing RAC drawings of graphs with  $n$  vertices,  $m$  edges, and cut-width  $c$  with curve complexity 4 in  $O(n + m) \times O(c)$  area. When the algorithm is applied to planar graphs with  $n$  vertices and degree  $d$ , the resulting RAC drawings have  $O(n\sqrt{dn})$  area. Note that, if  $d$  is bounded by any sublinear function of  $n$ , then such an area bound is subquadratic.
- In Sect. 3 we describe an algorithm for constructing LAC<sub>α</sub> drawings of graphs with  $n$  vertices and cut-width  $c$  with curve complexity 2 in  $O(n) \times O(c)$  area, for any fixed  $\alpha \in (0, \pi/2)$ . When such an algorithm is applied to planar graphs

with  $n$  vertices and degree  $d$ , the resulting  $\text{LAC}_\alpha$  drawings have  $O(n\sqrt{dn})$  area. Again, if  $d$  is bounded by any sublinear function of  $n$ , then such an area bound is subquadratic.

- In Sect. 4 we describe an algorithm for constructing  $\text{LAC}_\alpha$  drawings with curve complexity 2 of every graph  $G$  such that every subgraph  $J$  of  $G$  with  $n_J$  vertices and degree  $d_J$  has  $O(n_J)$  edges and cut-width at most  $\gamma(n_J, d_J)$  in  $O(\frac{n}{r} + \gamma(n, r)) \times O(n)$  area, for any  $1 \leq r \leq n$  and for any fixed  $\alpha \in (0, \pi/2)$ . When such an algorithm is applied to planar graphs with  $n$  vertices, the resulting  $\text{LAC}_\alpha$  drawings have  $O(n^{5/3})$  area.

The rest of the paper is organized as follows. In Sect. 2 we give some preliminaries; in Sect. 3 and in Sect. 4 we show algorithms for constructing RAC and  $\text{LAC}_\alpha$  graph drawings; finally, in Sect. 5 we conclude and suggest some open problems.

## 2 Preliminaries

A *poly-line grid drawing* of a graph is a mapping of each vertex to a distinct point of the plane with integer coordinates and of each edge to a sequence of rectilinear segments connecting the endpoints of the edge. A *bend* is a point in which an edge changes its slope, i.e., a point common to two consecutive segments in the sequence of segments representing the edge. The *curve complexity* of a drawing is the maximum number of bends of any edge in the drawing. In the following we always refer to poly-line grid drawings. The *bounding box* of a drawing  $\Gamma$  is the smallest rectangle with sides parallel to the axes that covers  $\Gamma$  completely. The *height* (*width*) of  $\Gamma$  is the height (resp. width) of its bounding box. The *area* of  $\Gamma$  is the height of  $\Gamma$  times its width.

The *cut-width* of a graph  $G$  is the smallest integer  $c$  such that the vertices of  $G$  can be arranged in a linear layout  $v_1, \dots, v_n$  in such a way that, for every  $i$ , at most  $c$  edges have one endpoint in  $v_1, \dots, v_i$  and the other in  $v_{i+1}, \dots, v_n$ . The following lemma has been proved by Diks et al. [12].

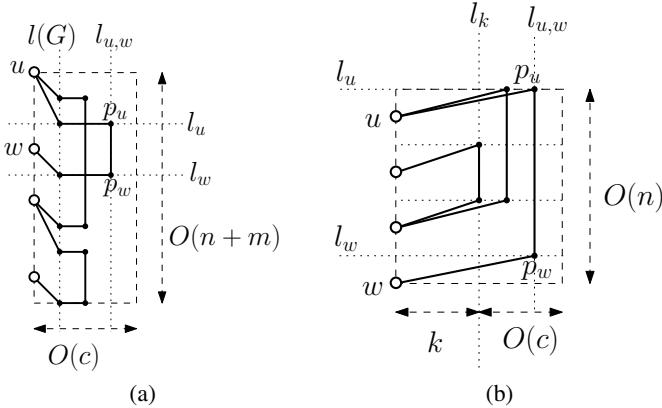
**Lemma 1.** *The cut-width of an  $n$ -vertex degree- $d$  planar graph is  $O(\sqrt{dn})$ .*

The proof of Diks et al. [12] is constructive. Namely, the authors present an algorithm for finding an  $O(\sqrt{dn})$  *edge-separator* for an  $n$ -vertex planar graph in  $O(n)$  time; they construct the linear layout with  $O(\sqrt{dn})$  cut-width by recursive applications of such an algorithm. A naive implementation of such an approach leads to an  $O(n^2)$ -time algorithm for laying out the vertices of a planar graph with  $O(\sqrt{dn})$  cut-width.

In the following, we will present algorithms for constructing RAC and  $\text{LAC}_\alpha$  drawings of graphs without explicitly mentioning their running time. However, the only non-trivial algorithmic steps will consist of constructing linear layouts of the vertices of the graphs with small cut-width. Hence, all the algorithms we will present can be implemented to run in polynomial time.

## 3 RAC and LAC Drawings of Bounded-Degree Planar Graphs

In this section we show two algorithms for constructing RAC drawings and  $\text{LAC}_\alpha$  drawings with curve complexity 4 and 2, respectively, of graphs with  $n$  vertices,  $m$  edges, and cut-width  $c$ . We start with RAC drawings.



**Fig. 1.** (a) Layout of Theorem 1. (b) Layout of Theorem 2.

**Theorem 1.** Every graph with  $n$  vertices,  $m$  edges, and cut-width  $c$  admits an RAC drawing with curve complexity 4 in  $O(n + m) \times O(c)$  area.

**Proof:** We compute an RAC drawing of the input graph  $G$  inside an axis parallel rectangle  $R(G)$  with height  $O(n + m)$  and width  $O(c)$ , where all the vertices of  $G$  are placed on the left side of  $R(G)$ .

First, compute the order  $u_1, \dots, u_n$  of the vertices of  $G$  corresponding to the optimal cut-width  $c$  of  $G$ . Then, place the vertices of  $G$  on the left side of  $R(G)$  in this order from top to bottom, starting by placing  $u_1$  on the top side of  $R(G)$ , and leaving  $d_i$  free horizontal lines between  $u_i$  and  $u_{i+1}$ , for each vertex  $u_i$ , where  $d_i$  is the degree of  $u_i$ . Let  $l(G)$  be the vertical line one unit to the right of the left side of  $R(G)$ .

We now show how to route the edges of  $G$ . Such edges are routed one by one in any order. Refer to Fig. 1(a). When an edge  $(u, w)$  has to be routed, consider the highest free horizontal line  $l_u$  ( $l_w$ ) below  $u$  (below  $w$ , resp.) that has not yet been used to route an edge incident to  $u$  (to  $w$ , resp.). Further, consider the leftmost vertical line  $l_{u,w}$  to the right of  $l(G)$  such that the segment having as endpoints the intersection points  $p_u$  and  $p_w$  of  $l_{u,w}$  with  $l_u$  and  $l_w$ , respectively, can be drawn without overlapping any edge of the current drawing. Route  $(u, w)$  by connecting  $u$  to the intersection point between  $l_u$  and  $l(G)$ , then to  $p_u$ , then to  $p_w$ , then to the intersection point between  $l_w$  and  $l(G)$ , and finally to  $w$ .

Observe that a crossing may only happen between a horizontal and a vertical segment, as all the edge-segments that are neither horizontal nor vertical are drawn without crossings in the vertical strip delimited by the left side of  $R(G)$  and by  $l(G)$ . By construction, no two edges overlap in the drawing. Hence, all the crossing angles are equal to  $\frac{\pi}{2}$ , and the drawing is an RAC drawing. As each edge bends four times, the curve complexity of the drawing is 4.

The height of the drawing is twice the number of edges plus the number of vertices, and hence it is  $O(n + m)$ . The width of the drawing is equal to the maximum number of edges crossing the same horizontal line plus a constant. Suppose that there are more than  $c$  edges intersecting the same horizontal line. Then, there would exist an

index  $i$  such that more than  $c$  edges connect vertices in  $u_1, u_2, \dots, u_i$  with vertices in  $u_{i+1}, u_{i+2}, \dots, u_n$ , thus contradicting the fact that the order  $u_1, \dots, u_n$  of the vertices of  $G$  is the order corresponding to the optimal cut-width  $c$  of  $G$ . Hence, the width of the drawing is  $O(c)$ , and the statement follows.  $\square$

The previous theorem, together with Lemma 1 and with the fact that planar graphs have  $O(n)$  edges, directly implies the following.

**Corollary 1.** *Every  $n$ -vertex degree- $d$  planar graph admits an RAC drawing with curve complexity 4 in  $O(n) \times O(\sqrt{dn})$  area.*

We now turn our attention to LAC $_\alpha$  drawings with curve complexity 2. The construction is analogous to the one of Theorem 1.

**Theorem 2.** *Every graph with  $n$  vertices and cut-width  $c$  admits an LAC $_\alpha$  drawing with curve complexity 2 in  $O(n) \times O(c)$  area, for any fixed  $\alpha \in (0, \pi/2)$ .*

**Proof:** We compute an LAC $_\alpha$  drawing of  $G$  inside an axis parallel rectangle  $R(G)$  with height  $2n$  and width  $O(c)$ , where all the vertices of  $G$  are placed on the left side of  $R(G)$ , as follows.

First, compute the order  $u_1, \dots, u_n$  of the vertices of  $G$  corresponding to the optimal cut-width  $c$  of  $G$ . Then, place the vertices of  $G$  on the left side of  $R(G)$  in this order from top to bottom, starting by placing  $u_1$  one unit below the top side of  $R(G)$ , and leaving a free horizontal line between any two consecutive vertices. For each vertex  $u$ , let  $l_u$  be the horizontal line one unit above  $u$ . Let  $l_k$  be the vertical line  $k = \frac{1}{\tan(\frac{\pi}{2} - \alpha)}$  units to the right of the left side of  $R(G)$ .

We now show how to route the edges of  $G$ . Such edges are routed one by one in any order. Refer to Fig. 1(b). When an edge  $(u, w)$  has to be routed, consider the leftmost vertical line  $l_{u,w}$  to the right of  $l_k$  such that the segment having as endpoints the intersection points  $p_u$  and  $p_w$  of  $l_{u,w}$  with  $l_u$  and  $l_w$ , respectively, can be drawn without overlapping any edge of the current drawing. Route  $(u, w)$  by connecting  $u$  to  $p_u$ , then to  $p_w$ , and finally to  $w$ .

Since all the edge-segments are either vertical or connecting two points with vertical distance 1 and horizontal distance at least  $k$ , and since by construction no two edges overlap, it follows that a crossing may only happen between a vertical and a non-vertical edge. Hence, any two crossing segments create an angle that is at least  $\frac{\pi}{2} - \arctan(\frac{1}{k}) = \frac{\pi}{2} - \arctan(\tan(\frac{\pi}{2} - \alpha)) = \alpha$ . As each edge has exactly two bends, the constructed drawing is an LAC $_\alpha$  drawing with curve complexity 2.

The height of the drawing is twice the number of vertices of  $G$ , and hence it is  $O(n)$ . The width of the drawing is given by  $k$  (which is a constant, for any fixed  $\alpha \in (0, \pi/2)$ ) plus the maximum number of edges crossing the same horizontal line which, as in the proof of Theorem 1, is  $O(c)$ , and the statement follows.  $\square$

The previous theorem, together with Lemma 1, directly implies the following.

**Corollary 2.** *Every  $n$ -vertex degree- $d$  planar graph  $G$  admits an LAC $_\alpha$  drawing with curve complexity 2 in  $O(n) \times O(\sqrt{dn})$  area, for any fixed  $\alpha \in (0, \pi/2)$ .*

## 4 LAC Drawings of General Planar Graphs

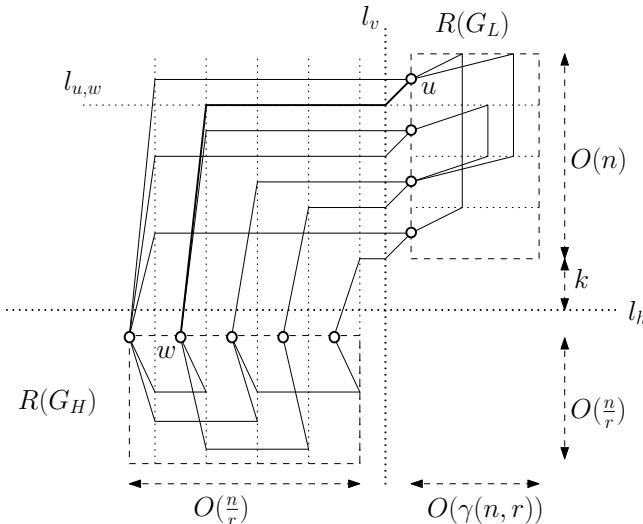
In this section we show our main result, namely that planar graphs admit  $\text{LAC}_\alpha$  drawings with curve complexity 2 in subquadratic area, for any fixed  $\alpha \in (0, \pi/2)$ . We in fact prove a stronger theorem, which directly implies the result on planar graphs as well as area bounds for  $\text{LAC}_\alpha$  drawings with curve complexity 2 of other graph classes.

**Theorem 3.** *Let  $G$  be a graph with  $n$  vertices, with  $O(n)$  edges, with degree  $d$ , and with cut-width  $\gamma(n, d)$ , for some function  $\gamma$  of  $n$  and  $d$ . Suppose that every subgraph  $J$  of  $G$  with  $n_J$  vertices and degree  $d_J$  has  $O(n_J)$  edges and cut-width at most  $\gamma(n_J, d_J)$ . Then, for every  $1 \leq r \leq n$  and for any fixed  $\alpha \in (0, \pi/2)$ ,  $G$  admits a  $\text{LAC}_\alpha$  drawings with curve complexity 2 in  $O(\frac{n}{r} + \gamma(n, r)) \times O(n)$  area.*

**Proof:** Let  $r$  be any value between 1 and  $n$ . First, we partition the vertices of  $G$  into heavy and light vertices. We say that a vertex is *heavy* if its degree is greater than or equal to  $r$  and it is *light* otherwise. Let  $H$  denote the set of heavy vertices and  $L$  denote the set of light vertices of  $G$ . Note that  $|H| \in O(\frac{n}{r})$ , given that every subgraph  $J$  of  $G$  with  $n_J$  vertices has  $O(n_J)$  edges, while  $|L| \in O(n)$ .

The  $\text{LAC}_\alpha$  drawing of  $G$  is constructed by drawing the two subgraphs of  $G$  induced by  $H$  and by  $L$ , respectively, and the edges connecting the vertices of such two sets inside three different regions, as follows. Refer to Fig 2.

Consider the subgraph  $G_H$  induced by the vertices in  $H$ . By Theorem 2,  $G_H$  admits an  $\text{LAC}_\alpha$  drawing with curve complexity 2 inside a rectangle  $R(G_H)$  such that: (a) all the vertices in  $H$  lie on the top side of  $R(G_H)$ ; (b) the width of  $R(G_H)$  is  $O(|H|) \in O(\frac{n}{r})$ ; and (c) the height of  $R(G_H)$  is  $O(\frac{n}{r})$ , given that every subgraph  $J$  of  $G$  with  $n_J$



**Fig. 2.** Layout of Theorem 3

vertices and degree  $d_J$  has  $O(n_J)$  edges and given that  $G_H$  has at most  $O(\frac{n}{r})$  vertices; since  $r \geq 1$ , we have that the height of  $R(G_H)$  is  $O(n)$ .

Consider the subgraph  $G_L$  induced by the vertices in  $L$ . By Theorem 2,  $G_L$  admits an LAC $_\alpha$  drawing with curve complexity 2 inside a rectangle  $R(G_L)$  such that: (a) all the vertices in  $L$  lie on the left side of  $R(G_L)$ ; (b) the height of  $R(G_L)$  is  $O(|L|) \in O(n)$ ; and (c) the width of  $R(G_L)$  is  $\gamma(n, r)$ , given that the degree of  $G_L$  is at most  $r$ . For the construction of such an LAC $_\alpha$  drawing we use the same algorithm of the proof of Theorem 2, except that we assume that two consecutive vertices  $u_i$  and  $u_{i+1}$  on the left side of  $R(G_L)$  have  $d_i - 1$  free horizontal lines separating them, where  $d_i$  is the number of heavy vertices adjacent to  $u_i$ . Observe that such a modification does not alter the asymptotic bounds on the height of the drawing, given that  $G$  has  $O(n)$  edges.

Consider a horizontal line  $l_h$  and a vertical line  $l_v$ . Place  $R(G_H)$  with its top side one unit below  $l_h$  and its right side one unit to the left of  $l_v$ . Place  $R(G_L)$  with its bottom side  $k$  units above  $l_h$ , where  $k = \frac{1}{\tan(\frac{\pi}{2} - \alpha)}$ , and its left side one unit to the right of  $l_v$ . Consider the vertices of  $L$  in the order they appear on the left side of  $R(G_L)$ . For each of such vertices, consider all the edges connecting it to the vertices of  $H$ . For each of such edges  $(u, w)$ , where  $u \in L$  and  $w \in H$ , consider the topmost horizontal line  $h_{u,w}$  that is not above the horizontal line through  $u$  and that can be drawn without overlapping any other edge. Route  $(u, w)$  by placing two bends on  $h_{u,w}$ , namely one in its intersection point with  $l_v$  and one in its intersection point with the free vertical line immediately to the right of  $w$ .

Since each vertex  $u_i$  in  $L$  has  $d_i$  free horizontal lines below it, then no two edges overlap. Since all the edge-segments are either horizontal or connecting two points with horizontal distance 1 and vertical distance at least  $k$ , and since a crossing may happen only between a horizontal and a non-horizontal edge, it follows that any two crossing segments create an angle that is at least  $\frac{\pi}{2} - \arctan(\frac{1}{k}) = \frac{\pi}{2} - \arctan(\tan(\frac{\pi}{2} - \alpha)) = \alpha$ . As each edge has exactly two bends, the constructed drawing is an LAC $_\alpha$  drawing with curve complexity 2. Note that the number of edges connecting a vertex of  $L$  to a vertex of  $H$  is  $O(n)$ , and hence all such edges are drawn inside the smallest rectangle  $R(G)$  containing both  $R(G_L)$  and  $R(G_H)$ , which implies that these edges do not increase the total area of the drawing.

We compute the area of  $R(G)$ . The height of  $R(G)$  is  $O(n)$ , as it is equal to the height of  $R(G_L)$ , which is  $O(n)$ , plus the height of  $R(G_H)$ , which is  $O(\frac{n}{r})$  and hence  $O(n)$ , plus a constant number of  $k + 1$  lines. The width of  $R(G)$  is  $O(\frac{n}{r} + \gamma(n, r))$ , as it is equal to the width of  $R(G_H)$ , which is  $O(\frac{n}{r})$ , plus the width of  $R(G_L)$ , which is  $\gamma(n, r)$ . This concludes the proof of the theorem.  $\square$

We get the following main result for planar graphs:

**Corollary 3.** *Every  $n$ -vertex planar graph  $G$  admits an LAC $_\alpha$  drawing with curve complexity 2 in  $O(n^{5/3})$  area, for any fixed  $\alpha \in (0, \pi/2)$ .*

**Proof:** Let  $G$  be any  $n$ -vertex planar graph. Let  $\gamma(n, d) = \sqrt{dn}$ . Since every subgraph of a planar graph is also planar, we have that every subgraph  $J$  of  $G$  with  $n_J$  vertices and degree  $d_J$  has  $O(n_J)$  edges and, by Lemma 1, cut-width at most  $\sqrt{d_J n_J}$ . Hence, by Theorem 3,  $G$  admits a LAC $_\alpha$  drawing in  $O(\frac{n}{r} + \gamma(n, r)) \times O(n)$  area. By setting

$r = n^{1/3}$ ,  $G$  admits a LAC $_{\alpha}$  drawing in  $O(n^{2/3} + \sqrt{n^{1/3}n}) \times O(n) = O(n^{2/3}) \times O(n)$  area, for any fixed  $\alpha \in (0, \pi/2)$ , thus proving the corollary.  $\square$

Also, we have that every graph  $G$  with  $o(\frac{n}{\log n})$  treewidth<sup>1</sup> and such that every subgraph of  $G$  has a linear number of edges admits LAC $_{\alpha}$  drawings in subquadratic area, for any fixed  $\alpha \in (0, \pi/2)$ .

**Corollary 4.** *Every  $n$ -vertex graph  $G$  with  $O(n)$  edges, with treewidth  $k$ , and such that every subgraph  $J$  of  $G$  with  $n_J$  vertices has  $O(n_J)$  edges admits an LAC $_{\alpha}$  drawing with curve complexity 2 in  $O(\sqrt{kn^3 \log n})$  area, for any fixed  $\alpha \in (0, \pi/2)$ .*

**Proof:** Korach and Solel [18] proved that any  $n$ -vertex graph with tree-width  $k$  and degree  $d$  has cut-width  $O(kd \log n)$ . Let  $G$  be any  $n$ -vertex graph with  $O(n)$  edges, with treewidth  $k$ , and such that every subgraph  $J$  of  $G$  with  $n_J$  vertices has  $O(n_J)$  edges. By Theorem 3, with  $r = \sqrt{\frac{n}{k \log n}}$ , we get that  $G$  admits a LAC $_{\alpha}$  drawing in  $O\left(\frac{n}{\sqrt{k \log n}} + k \sqrt{\frac{n}{k \log n}} \log n\right) \times O(n) = O(\sqrt{kn \log n}) \times O(n) = O(\sqrt{kn^3 \log n})$  area, for any fixed  $\alpha \in (0, \pi/2)$ , thus proving the corollary.  $\square$

Observe that the previous corollary implies subquadratic area bounds, e.g., for LAC drawings of graphs with bounded genus.

## 5 Conclusions and Open Problems

In this paper we have shown algorithms for constructing RAC and LAC drawings of graphs. While the asymptotic bounds provided by our algorithms do not outperform the previously known area bounds for general graphs [8, 9], they are very good for graphs with small cut-width. In particular, our algorithms lead to subquadratic area bounds for RAC and LAC drawings of bounded-degree planar graphs and general planar graphs, respectively.

**Table 1.** Area requirements for RAC and LAC drawings of planar graphs with small curve complexity

	$c = 0$	$c = 1$	$c \geq 2$
RAC drawings	$O(n^2)$ and $\Omega(n^2)$	$O(n^2)$ and $\Omega(n)$	$O(n^2)$ and $\Omega(n)$
LAC drawings	$O(n^2)$ and $\Omega(n)$	$O(n^2)$ and $\Omega(n)$	$O(n^{5/3})$ and $\Omega(n)$

Except for RAC drawings with curve complexity 0, no tight bounds are known for the area requirements of RAC/LAC drawings of planar graphs with curve complexity  $c$ , for any constant  $c$ . Table 1 summarizes the best known area upper and lower bounds for RAC and LAC drawings of planar graphs. Closing the gap between such bounds is an interesting open problem. We find particularly appealing the question whether RAC drawings with curve complexity  $c$  can be constructed in subquadratic area, for some constant  $c$ .

<sup>1</sup> For standard definitions about treewidth see, e.g., [11].

**Acknowledgments.** The authors are in debt with the reviewers for their comments that allowed us to significantly improve the results of this paper.

## References

1. Ackerman, E., Fulek, R., Tóth, C.D.: Graphs that admit polyline drawings with few crossing angles. *SIAM Journal on Discrete Mathematics* 26(1), 305–320 (2012)
2. Angelini, P., Cittadini, L., Di Battista, G., Didimo, W., Frati, F., Kaufmann, M., Symvonis, A.: On the perspectives opened by right angle crossing drawings. *Journal of Graph Algorithms and Applications* 15(1), 53–78 (2011)
3. Argyriou, E.N., Bekos, M.A., Symvonis, A.: The Straight-Line RAC Drawing Problem Is NP-Hard. In: Černá, I., Gyimóthy, T., Hromkovič, J., Jefferey, K., Králović, R., Vukolić, M., Wolf, S. (eds.) *SOFSEM 2011. LNCS*, vol. 6543, pp. 74–85. Springer, Heidelberg (2011)
4. Arikushi, K., Fulek, R., Keszegh, B., Moric, F., Tóth, C.D.: Graphs that admit right angle crossing drawings. *Computational Geometry: Theory & Applications* 45(4), 169–177 (2012)
5. Bonichon, N., Le Saëc, B., Mosbah, M.: Optimal Area Algorithm for Planar Polyline Drawings. In: Kučera, L. (ed.) *WG 2002. LNCS*, vol. 2573, pp. 35–46. Springer, Heidelberg (2002)
6. de Fraysseix, H., Pach, J., Pollack, R.: How to draw a planar graph on a grid. *Combinatorica* 10(1), 41–51 (1990)
7. Di Battista, G., Eades, P., Tamassia, R., Tollis, I.G.: *Graph Drawing*. Prentice Hall, Upper Saddle River (1999)
8. Di Giacomo, E., Didimo, W., Liotta, G., Meijer, H.: Area, curve complexity, and crossing resolution of non-planar graph drawings. *Theory of Computing Systems* 49(3), 565–575 (2011)
9. Didimo, W., Eades, P., Liotta, G.: Drawing Graphs with Right Angle Crossings. In: Dehne, F., Gavrilova, M., Sack, J.-R., Tóth, C.D. (eds.) *WADS 2009. LNCS*, vol. 5664, pp. 206–217. Springer, Heidelberg (2009)
10. Didimo, W., Eades, P., Liotta, G.: Drawing graphs with right angle crossings. *Theoretical Computer Science* 412(39), 5156–5166 (2011)
11. Diestel, R.: *Graph Theory*. Springer, Heidelberg (2010)
12. Diks, K., Djidjev, H., Sýkora, O., Vrto, I.: Edge separators of planar and outerplanar graphs with applications. *J. Algorithms* 14(2), 258–279 (1993)
13. Dujmović, V., Gudmundsson, J., Morin, P., Wolle, T.: Notes on large angle crossing graphs. *Chicago Journal of Theoretical Computer Science*, article 4, 1–14 (2011)
14. Frati, F., Patrignani, M.: A Note on Minimum-Area Straight-Line Drawings of Planar Graphs. In: Hong, S.-H., Nishizeki, T., Quan, W. (eds.) *GD 2007. LNCS*, vol. 4875, pp. 339–344. Springer, Heidelberg (2008)
15. Huang, W.: Using eye tracking to investigate graph layout effects. In: Hong, S.H., Ma, K.L. (eds.) *Asia-Pacific Symposium on Visualization (APVIS 2007)*, pp. 97–100. IEEE (2007)
16. Huang, W., Hong, S.H., Eades, P.: Effects of crossing angles. In: *Pacific Visualization (PacificVis 2008)*, pp. 41–46. IEEE (2008)
17. Kaufmann, M., Wagner, D. (eds.): *Drawing Graphs. LNCS*, vol. 2025. Springer, Heidelberg (2001)
18. Korach, E., Solel, N.: Tree-width, path-width, and cutwidth. *Discrete Applied Mathematics* 43(1), 97–101 (1993)

19. Schnyder, W.: Embedding planar graphs on the grid. In: Symposium on Discrete Algorithms (SODA 1990), pp. 138–148 (1990)
20. Tamassia, R., Tollis, I.G.: A unified approach to visibility representation of planar graphs. *Discrete & Computational Geometry* 1, 321–341 (1986)
21. Tamassia, R., Tollis, I.G.: Planar bookgrid embedding in linear time. *IEEE Transactions on Circuits and Systems* 36(9), 1230–1234 (1989)
22. Valiant, L.G.: Universality considerations in VLSI circuits. *IEEE Transactions on Computers* 30(2), 135–140 (1981)
23. Wood, D.R.: Grid drawings of  $k$ -colourable graphs. *Computational Geometry: Theory & Applications* 30(1), 25–28 (2005)

# Connecting Red Cells in a Bicolour Voronoi Diagram

Manuel Abellanas<sup>1</sup>, Antonio L. Bajuelos<sup>2</sup>, Santiago Canales<sup>3</sup>, Mercè Claverol<sup>4</sup>,  
Gregorio Hernández<sup>1</sup>, and Inês Matos<sup>2,4</sup>

<sup>1</sup> Dept. de Matemática Aplicada, Universidad Politécnica de Madrid, Spain  
[{mabellanas,gregorio}@fi.upm.es](mailto:{mabellanas,gregorio}@fi.upm.es)

<sup>2</sup> Dept. de Matemática & CIDMA, Universidade de Aveiro, Portugal  
[{leslie,ipmatos}@ua.pt](mailto:{leslie,ipmatos}@ua.pt)

<sup>3</sup> Esc. Téc. Superior de Ingeniería, Universidad Pontificia Comillas de Madrid, Spain  
[sccanales@dmc.icai.upcomillas.es](mailto:sccanales@dmc.icai.upcomillas.es)

<sup>4</sup> Dept. de Matemàtica Aplicada II & IV, Universitat Politècnica de Catalunya, Spain  
[merce@ma4.upc.edu](mailto:merce@ma4.upc.edu)

**Abstract.** Let  $S$  be a set of  $n + m$  sites, of which  $n$  are red and have weight  $w_R$ , and  $m$  are blue and weigh  $w_B$ . The objective of this paper is to calculate the minimum value of the red sites' weight such that the union of the red Voronoi cells in the weighted Voronoi diagram of  $S$  is a connected region. This problem is solved for the multiplicatively-weighted Voronoi diagram in  $\mathcal{O}((n+m)^2 \log(nm))$  time and for both the additively-weighted and power Voronoi diagram in  $\mathcal{O}(nm \log(nm))$  time.

**Keywords:** Weighted Voronoi diagrams, Bicolour Points.

## 1 Introduction

Red and blue points have been a popular topic of study in Geometry, and can be found in several branches of research such as subdivision problems, geometric graph problems, and so forth [8]. Suppose  $S = R \cup B$  is a set of  $n + m$  sites,  $n$  of which are red and  $m$  of which are blue; and let  $\text{VD}(S)$  denote the ordinary Voronoi diagram of  $S$  [4]. This diagram divides the plane into cells. Each cell is defined by a site  $s_i \in S$  and is composed of all the points on the plane that are closer to  $s_i$  than to any other site. A Voronoi cell of  $\text{VD}(S)$  is said to be red (resp. blue) if the corresponding generator site is red (blue). The objective of this paper is to connect all the red cells in order to allow one to travel within red regions following paths that do not cross blue regions. In other words, the union of red cells has to form a connected region. If that is not the case, then there are several ways to rectify this; for instance, one can add red sites or delete blue sites, or even move sites. Berg et al. [6] chose to remove the minimum number of blue sites in  $\mathbb{R}^d$  to achieve this goal. In so doing, they proved that finding such a solution is an NP-hard problem, unless there are only two red sites. In this instance, the problem is solved in  $\mathcal{O}(n \log n + n^{d-1})$  time, which is optimal for  $d = 2$ . In a related problem, Aichholzer et al. [1] studied the minimum number

of red sites needed to prevent the existence of an edge connecting two blue sites of the Delaunay triangulation of  $S$ . They showed that  $\frac{3n}{2}$  red sites are always sufficient to block a set of  $n$  blue sites if  $S$  is in general position. If the blue sites are in convex position, then  $\frac{5n}{4}$  red sites are always sufficient. Moreover, they proved that at least  $n - 1$  red sites are always necessary and there are examples requiring at least  $n$  red sites.

The approach taken in the following discussion differs from the previous ones because the sites are weighted, and therefore the weighted Voronoi diagram of  $S$  is used instead (see Okabe et al. [9] for detailed information). The weighted distances studied in this paper are the multiplicatively- and additively-weighted distances, as well as the power distance. All red sites are assigned the same weight  $w_R$ , as all blue sites are assigned  $w_B$ . Thus the goal is to study how these weights affect red cells' size and which weights allow them to become a connected region. As it is easy to realise, under these conditions the only relevant data is the relationship between  $w_R$  and  $w_B$ . Therefore, and assuming  $w_B$  is a constant such as 1, the problem can be restated as calculating the minimum weight of the red sites for which the union of red cells of the appropriate diagram is connected. Such weight will be denoted by  $w_R^*$ .

Let  $\text{VD}^w(S)$  denote the weighted Voronoi diagram of  $S$ . Bear in mind that  $\text{VD}^w(S) = \text{VD}(R)$  when  $w_R$  tends to infinity, which assures that our problem always has a feasible solution. It is clear that the structure of  $\text{VD}^w(S)$  depends on the weight of red sites and if  $w_R$  is small enough then  $\text{VD}^w(S)$  tends to  $\text{VD}(B)$ . Red cells' size varies in accordance with  $w_R$ , and therefore starting from a small weight that is slowly increasing, red cells will expand and eventually two of them will meet to form one connected red component of  $\text{VD}^w(S)$ . The point where two red cells meet is called a *critical point*. As red cells keep growing, more will connect at different weights and each of these weights will be defined by a sole critical point. The sought weight is the one that finally unites the last two disconnected red components of  $\text{VD}^w(S)$ . Consequently, finding these critical points is the key to solve the problem. Observe that it is possible for some critical points to exist “in the infinity” and therefore a solution involving them is not interesting to the present problem as a path visiting all red cells of  $\text{VD}^w(S)$  would be infinitely long. To solve this problem, it is assumed that all sites of  $S$  are enclosed in a polygon or bounding box. This case will be further explained in Section 5.

The remainder of this paper is divided in four sections, where the first three correspond to the three types of weighted distances in question: multiplicatively-weighted distance, additively-weighted distance and power distance. Section 5 concludes this paper and is followed by the acknowledgements.

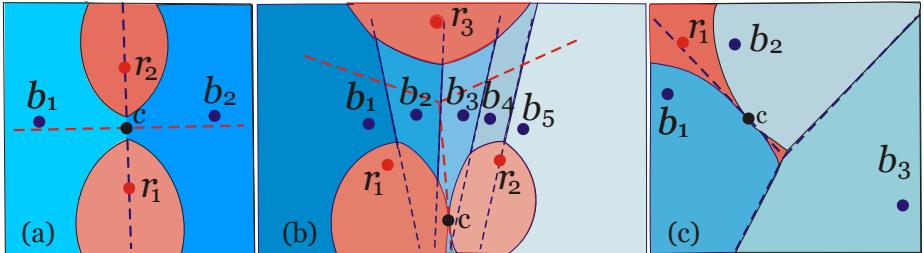
## 2 Multiplicatively-Weighted Distance

The weighted distance studied in this section is called the *multiplicatively-weighted distance*. Given a point  $p$  on the plane and site  $s_i \in S$ , said distance is defined by  $d_M(p, s_i) = \frac{1}{w} d_E(p, s_i)$ , where  $d_E$  is the Euclidean distance and  $w$  should

be replaced by the current weight of  $R$  if  $s_i$  is a red site or  $w_B$  if  $s_i$  is a blue site. The multiplicatively-weighted distance characterises the multiplicatively-weighted Voronoi diagram of  $S$  [3]. Let  $b(s_i, s_j) = \{p \in \mathbb{R}^2 : d_M(p, s_i) = d_M(p, s_j)\}$  denote the bisector between sites  $s_i$  and  $s_j$  of  $S$ . Note that each of these bisectors is part of an Apollonius circle. If all sites have the same weight, then the multiplicatively-weighted Voronoi diagram becomes the ordinary Voronoi diagram. And similarly to the latter, which can be seen as the lower envelope of a set of cones in 3D, the discs' variations of size induced by different weights can be seen as variations of the cones' angle.

As previously mentioned, critical points are the key to calculate the minimum weight that connects all red cells of  $\text{VD}^w(S)$ . In order to find them, we need to understand how red cells form clusters as they grow. To this end, the following definition characterises the points where red cells of  $\text{VD}^w(S)$  meet.

**Definition 1.** Suppose  $w_R$  is the exact weight of  $R$  for which two red cells of  $\text{VD}^w(S)$  meet at point  $c$ . In such a case,  $c$  is a type I critical point if there are red sites  $r_i$  and  $r_j$  and blue sites  $b_k$  and  $b_l$  such that  $\{c\} = b(r_i, b_k) \cap b(r_i, b_l) \cap b(r_j, b_k) \cap b(r_j, b_l)$ . Otherwise  $c$  is a type II critical point: if  $w_R < w_B$  and  $c$  belongs to the blue cell of  $\text{VD}^w(S)$  defined by  $b_k$  then  $\{c\} = b(r_i, b_k) \cap b(r_j, b_k)$ ; if  $w_R > w_B$  and  $c$  belongs to the red cell of  $\text{VD}^w(S)$  defined by  $r_i$  then  $\{c\} = b(r_i, b_k) \cap b(r_i, b_l)$ .



**Fig. 1.**  $\text{VD}(R)$  is shown in a dashed red trace and  $\text{VD}(B)$  in a dark blue one. (a) Point  $c$  is a type I critical point. (b) Point  $c$  is a type II critical point for  $w_R < w_B$ . (c) Point  $c$  is a type II critical point for  $w_R > w_B$ .

In Figure 1(a) it is clear to see that a type I critical point can also be found as an intersection point between  $\text{VD}(R)$  and  $\text{VD}(B)$ . As the weight of red sites increases, red cells defined by  $r_1$  and  $r_2$  will meet at  $c$  and form one connected red component of  $\text{VD}^w(S)$ . Figures 1(b) and 1(c) illustrate two examples of critical points of type II. Figure 1(c) also shows that the regions of the multiplicatively-weighted Voronoi diagram of  $S$  may be disconnected. Hence a critical point is created when a region meets itself, as evinced by point  $c$  at the intersection of the two red cells that are part of the Voronoi region of  $r_1$ .

**Proposition 1.** Connected red components of  $\text{VD}^w(S)$  can only meet at critical points of type I or II.

*Proof.* Suppose all red cells of  $\text{VD}^w(S)$  are isolated. As  $w_R$  increases, two of these cells will eventually meet at a critical point. If  $w_R < w_B$  then Voronoi regions defined by red sites are convex, and therefore meet at bisectors between two red sites, that is, on edges of  $\text{VD}(R)$ . However, these events will only take place in the presence of blue sites, which means the study of critical points should be focused on the interaction between  $\text{VD}(R)$  and  $\text{VD}(B)$ . In this way, critical points found on edges of  $\text{VD}(R)$  can either coincide with edges of  $\text{VD}(B)$  or exist in the interior of cells of  $\text{VD}(B)$ . These intersections create type I critical points in the first case and type II critical points in the second. Otherwise if  $w_R > w_B$ , red sites may define disconnected Voronoi regions of  $\text{VD}^w(S)$ . In this instance, critical points can also be found when two red cells that are part of the same Voronoi region meet. Consequently, these points are located in the interior of cells of  $\text{VD}(R)$  that are crossed by edges of  $\text{VD}(B)$ . Hence a type II critical point.  $\square$

**Proposition 2.** *Let  $S$  be a set of  $n$  red sites and  $m$  blue sites. There are at most  $\mathcal{O}(n + m)$  points where red cells of  $\text{VD}^w(S)$  meet and decrease the number of connected red components.*

*Proof.* As  $w_R$  increases, red cells of  $\text{VD}^w(S)$  expand to form new connected red components. Since there are  $n$  isolated red cells to begin with, they can only meet  $n - 1$  times while  $w_R < w_B$  and the Voronoi regions of red sites remain convex. If  $w_R > w_B$ , in how many red cells can a Voronoi region of  $\text{VD}^w(S)$  be divided into? It is only possible to disconnect a Voronoi region of a red site if such region is surrounded by three blue cells of  $\text{VD}^w(S)$  (see Figure 1(c)). Observe that a red cell created this way corresponds to a vertex of  $\text{VD}(B)$ . Since  $\text{VD}(B)$  has size  $\mathcal{O}(m)$  [4], this event can only be reproduced  $\mathcal{O}(m)$  times. Note that connected red components can meet themselves, but these intersection points do not decrease the number of red components.  $\square$

A method to find candidates to critical points follows directly from Definition 1. As previously noted, candidates to type I critical points are easily found since they are the intersection points between the edges of  $\text{VD}(R)$  and  $\text{VD}(B)$ . If  $\{c\} = b(r_i, r_j) \cap b(b_k, b_l)$ , then the minimum weight  $w_R$  needed to reach  $c$  with red sites  $r_i$  and  $r_j$  is given by  $w_R = w_B \frac{d_E(c, r_i)}{d_E(c, b_k)}$ . Candidates to type II critical points can either be found on edges of  $\text{VD}(R)$  or on edges of  $\text{VD}(B)$ , depending on the relationship between  $w_R$  and  $w_B$ . Since the procedure to find them is analogous, only the first case will be described: candidates to critical points on edges of  $\text{VD}(R)$  that fall in the interior of cells of  $\text{VD}(B)$ . As an edge of  $\text{VD}(R)$  may cross several cells of  $\text{VD}(B)$  (see Figure 1(b)), one decides which is the blue site closest to  $c$  by computing all the intersection points between such edge of  $\text{VD}(R)$ ,  $b(r_i, r_j)$ , and  $\text{VD}(B)$ . For each intersection point  $p_k$ , which corresponds to a blue site  $b_k$ , calculate  $d_M(p_k, r_i) - d_M(p_k, b_k)$ . This distance will be zero at  $c$ , thus studying how this value alternates along  $b(r_i, r_j)$  will tell which is the cell of  $\text{VD}(B)$  where  $c$  lies. Once blue site  $b_k$  is found, one can infer the functions that describe  $b(r_i, b_k)$  and  $b(r_j, b_k)$ , both depending on the unknown weight of  $R$ . Finally,  $c$  is found at the minimum weight of  $R$  for which these bisectors are tangent.

**Proposition 3.** *Given a set  $S$  of  $n$  red sites and  $m$  blue sites, there are  $\mathcal{O}(nm)$  candidates to critical points.*

*Proof.* Candidates to type I critical points are the intersection points between the edges of  $\text{VD}(R)$  and  $\text{VD}(B)$ . Since each of these diagrams is linear in size [4], there are  $\mathcal{O}(nm)$  candidates to this type of event. Candidates to type II critical points are found on edges of  $\text{VD}(R)$  that cross cells of  $\text{VD}(B)$  or vice-versa. Each intersection point between an edge of  $\text{VD}(R)$  and  $\text{VD}(B)$  needs to be studied in order to find the closest blue site to the candidate associated with such edge. This search takes  $\mathcal{O}(m)$  time per edge of  $\text{VD}(R)$ , and therefore there are  $\mathcal{O}(nm)$  candidates to critical points.  $\square$

Note that this proposition evinces a gap between the actual number of critical points and the number of candidates to critical points. To conclude the algorithm above, a binary search will then locate the minimum range  $w_R^*$  amongst a list of weights obtained for each of these candidates to critical points. The following theorem states the complexity of this procedure.

**Theorem 1.** *Given a set  $S$  of  $n + m$  sites, the minimum weight  $w_R^*$  for which the union of red cells in the multiplicatively-weighted Voronoi diagram of  $S$  is a connected region can be found in  $\mathcal{O}((n + m)^2 \log(nm))$  time.*

*Proof.* The first task to find candidates to critical points is to construct and intersect  $\text{VD}(B)$  and  $\text{VD}(R)$ , which takes  $\mathcal{O}((n + m) \log(n + m) + nm)$  time [5]. According to Proposition 3, there are  $\mathcal{O}(nm)$  of these candidates, each corresponding to a different weight that can be calculated in constant time. These weights have to be sorted into a list in ascending order, which takes  $\mathcal{O}(nm \log(nm))$  time. A binary search will then be used to locate  $w_R^*$  in this list: for each listed weight, construct  $\text{VD}^w(S)$  in  $\mathcal{O}((n + m)^2)$  time [3]. Using this diagram, build graph  $G$  that has a node for each red cell and two nodes are connected if the respective cells are neighbours. To verify if  $G$  is connected, traverse it using the Depth-First Search algorithm, which also runs in  $\mathcal{O}((n + m)^2)$  time [7]. If  $G$  is indeed connected, then the search proceeds to lower weights on the list, otherwise it proceeds to higher weights. Finally, this step takes  $\mathcal{O}((n + m)^2)$  time for each weight and so it is concluded in  $\mathcal{O}((n + m)^2 \log(nm))$  time.  $\square$

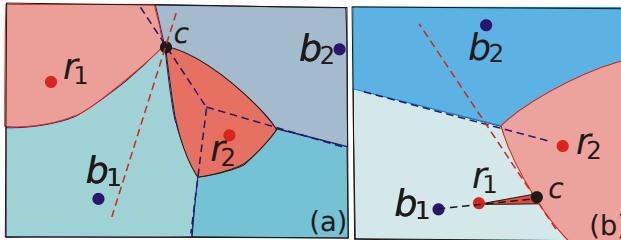
### 3 Additively-Weighted Distance

The weighted distance presented in this section is called the *additively-weighted distance* and is defined by  $d_A(p, s_i) = d_E(p, s_i) - w$ . As before,  $d_E$  is the Euclidean distance and  $w$  should be replaced by the current weight of  $R$  if  $s_i \in R$  or  $w_B$  if  $s_i \in B$ . Such distance characterises the additively-weighted Voronoi diagram of  $S$  [2], also called the *hyperbolic Dirichlet tessellation*. This diagram can present a peculiar structure since it is possible that a Voronoi region consists of a half-line. The edges of the additively-weighted Voronoi diagram are hyperbolic curves, hence its other name. Nonetheless, it also becomes the ordinary Voronoi diagram

when all sites have the same weight. The sites' different weights can be regarded as a variation of the cones' apex height, as opposed to the variation of the cones' angle that occurs in the multiplicatively-weighted Voronoi diagram. On the plane this diagram can also be interpreted as a set of discs that start growing at different times.

The method to find critical points on this diagram is similar to the one used in the last section. The main difference here is that the regions of the additively-weighted Voronoi diagram are always connected, if they exist. Let  $b(s_i, s_j) = \{p \in \mathbb{R}^2 : d_A(p, s_i) = d_A(p, s_j)\}$  represent the bisector between sites  $s_i$  and  $s_j$  of  $S$ .

**Definition 2.** Suppose  $w_R$  is the exact weight of  $R$  for which two red cells of  $\text{VD}^w(S)$  meet for the first time at point  $c$ . In that case,  $c$  is a type I critical point if there are red sites  $r_i$  and  $r_j$  and blue sites  $b_k$  and  $b_l$  such that  $\{c\} = b(r_i, b_k) \cap b(r_i, b_l) \cap b(r_j, b_k) \cap b(r_j, b_l)$ . Otherwise  $c$  is a type II critical point if  $b_k$  is the closest blue site to  $r_i$  and  $\{c\} = \overrightarrow{b_k r_i} \cap b(r_i, r_j)$ , where  $\overrightarrow{b_k r_i}$  is a ray from  $b_k$  to  $r_i$ .



**Fig. 2.**  $\text{VD}(R)$  is shown in a dashed red trace and  $\text{VD}(B)$  in a dark blue one. (a) Point  $c$  is a type I critical point. (b) Point  $c$  is a type II critical point.

As before, type I critical points are found at the intersections between edges of  $\text{VD}(R)$  and  $\text{VD}(B)$  (see Figure 2(a)). A type II critical point  $\{c\} = \overrightarrow{b_1 r_1} \cap b(r_1, r_2)$  is shown in Figure 2(b). In this instance, red site  $r_1$  does not have a Voronoi region until its weight wins over the weight of blue site  $b_1$ . At that exact moment, the Voronoi region of  $r_1$  is the line segment connecting  $r_1$  and  $c$ , and is therefore a degenerated region. The proofs of the following propositions are omitted as they coincide with the corresponding proofs of the previous section.

**Proposition 4.** Given a set  $S$  of  $n$  red sites and  $m$  blue sites, connected red components of  $\text{VD}^w(S)$  can only meet at critical points of type I or II. There are at most  $\mathcal{O}(n)$  points where red cells of  $\text{VD}^w(S)$  meet and decrease the number of connected red components.

Since the Voronoi regions of the additively-weighted Voronoi diagram cannot be disconnected, red cells of  $\text{VD}^w(S)$  can only meet and reduce the number of connected red components as many times as there are red cells on the diagram. Unfortunately this advantage does not help to lower the following complexity.

**Proposition 5.** Given a set  $S$  of  $n$  red sites and  $m$  blue sites, there are  $\mathcal{O}(nm)$  candidates to critical points.

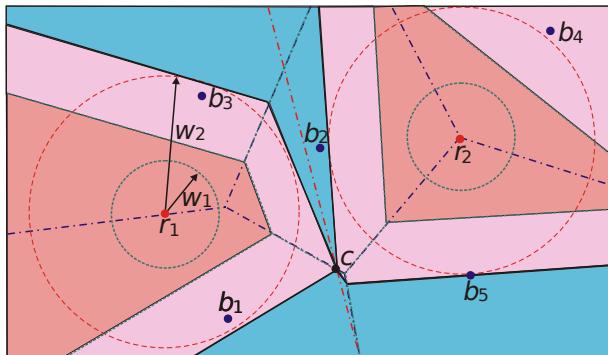
Once again there is a gap between the actual number of critical points and the number of candidates to critical points. As soon as the list of these candidates is compiled,  $w_R^*$  can be found by means of a binary search.

**Theorem 2.** Given a set  $S$  of  $n + m$  sites, the minimum weight  $w_R^*$  for which the union of red Voronoi cells in the additively-weighted Voronoi diagram of  $S$  is a connected region can be found in  $\mathcal{O}(nm \log(nm))$  time.

*Proof.* Since the method to find  $w_R^*$  is the same, this proof is similar to the one of Theorem 1. The main difference is that  $\text{VD}^w(S)$  is constructed in  $\mathcal{O}((n+m)\log(n+m))$  time [2], and therefore the binary search is concluded in  $\mathcal{O}((n+m)\log(n+m)\log(nm))$  time.  $\square$

## 4 Power Distance

The last weighted distance studied in this paper is the *power distance* [2]. It is also known as the *additively-weighted power distance* and is defined by  $d_P(p, s_i) = d_E(p, s_i)^2 - w$ , which characterises the power Voronoi diagram of  $S$  [2]. This diagram's structure is probably the closest to the usual Voronoi diagram of the three discussed in this paper. Observe that the edges of the power Voronoi diagram are line segments and its regions are convex as well. However, it is possible that a site does not have a Voronoi region, or that it does but does not belong to its own Voronoi region. Supposing all sites are the centre of discs of radius  $\sqrt{w}$ , then  $d_P(p, s_i)$  can be regarded as the distance between  $p$  and the intersection point of a line from  $p$  tangent to the disc centred at  $s_i$ .



**Fig. 3.**  $\text{VD}(R)$  is shown in a dashed red trace and  $\text{VD}(B)$  in dark blue. Power Voronoi diagram of  $S$  for  $w_R = w_1$  is shown in light green and for  $w_R = w_2$  is shown in black. Point  $c$  is a critical point.

As power Voronoi regions are convex, and thus simply-connected, there cannot be Voronoi regions within other Voronoi regions. Consequently, there is only one type of critical point (assuming  $S$  is in general position), which lies at the intersection of an edge of the red Voronoi diagram with an edge of the blue Voronoi diagram (see Figure 3). Since the power Voronoi diagram has linear size and can be built in the same amount of time as the additively-weighted Voronoi diagram [4], the proofs of the following results are omitted.

**Proposition 6.** *Given a set  $S$  of  $n$  red sites and  $m$  blue sites, there are  $\mathcal{O}(n)$  critical points.*

This proposition follows from the fact that red cells of  $\text{VD}^w(S)$  are convex and so they can only meet as many times as there are red cells on the diagram. Candidates to critical points are the intersection points between  $\text{VD}(R)$  and  $\text{VD}(B)$ , and therefore correspond to type I critical points of the previous two diagrams.

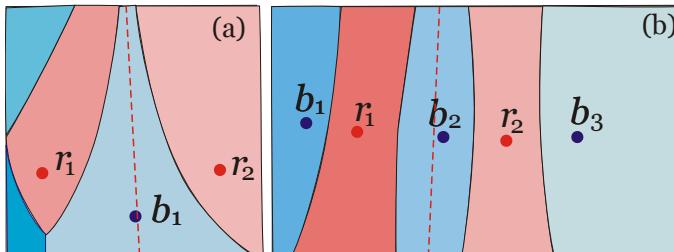
**Proposition 7.** *Given a set  $S$  of  $n$  red sites and  $m$  blue sites, there are  $\mathcal{O}(nm)$  candidates to critical points.*

Following the aforementioned method, candidates to critical points are found and the respective weights gathered into a list. Then a binary search will locate the weight  $w_R^*$  on such list.

**Theorem 3.** *Given a set  $S$  of  $n$  red sites and  $m$  blue sites, the minimum weight  $w_R^*$  for which the union of red Voronoi cells in the power Voronoi diagram of  $S$  is a connected region can be found in  $\mathcal{O}(nm \log(nm))$  time.*

## 5 Closing Remarks and Future Research

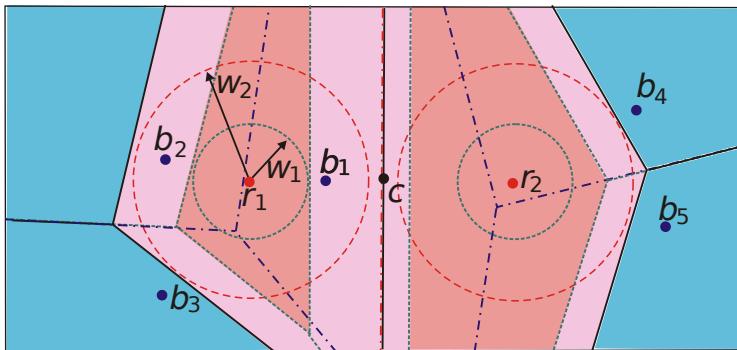
Given a set  $S$  of  $n$  red and  $m$  blue sites, this discussion was focused on minimising the weight of the red sites in order to guarantee that the union of red cells of



**Fig. 4.** (a) Bisectors  $b(r_1, b_1)$  and  $b(r_2, b_1)$  of the additively-weighted Voronoi diagram meet in the infinity. (b) Bisectors  $b(r_1, b_2)$  and  $b(r_2, b_2)$  of the multiplicatively-weighted Voronoi diagram meet twice in the infinity.

the weighted Voronoi diagram of  $S$  was connected. It was shown how to reach a solution in  $\mathcal{O}((n+m)^2 \log(nm))$  time for the multiplicatively-weighted distance and in  $\mathcal{O}(nm \log(nm))$  time for both the additively-weighted and the power distances. The first problem left open is how to close the gap or at least lessen the difference between the number of candidates to critical points and the number of actual critical points.

Observe that for the additively- and multiplicatively-weighted distances, it is possible that a type II critical point exists “in the infinity” as Figure 4 illustrates. However, such solution is not interesting to the present problem as a path visiting all red cells of  $\text{VD}^w(S)$  would be infinitely long. Therefore, we assume the existence of a bounding box or of a polygon containing all the sites of  $S$ . In this way, type II critical points will be the intersection points between the respective bisector of two red sites and the chosen bounding box.



**Fig. 5.** Power Voronoi diagram of  $S$  for  $w_R = w_1$  is shown in light green and for  $w_R = w_2$  is shown in black. Red cells generated by  $r_1$  and  $r_2$  meet at  $b(r_1, r_2)$ . Critical point  $c$  is the midpoint of  $\overline{r_1 r_2}$  and  $w_R^* = w_2$ .

There is also a similar glitch concerning power Voronoi diagrams. Assuming  $S$  is not in general position, three points might be collinear. In such a case, it is possible that a critical point is not an intersection between  $\text{VD}(R)$  and  $\text{VD}(B)$  but the midpoint of two red sites. Figure 5 shows an example of this case: sites  $r_1$ ,  $r_2$  and  $b_1$  are collinear and  $c$  is the critical point. The minimum weight  $w_R^* = d_E(r_1, c)^2 - d_E(b_1, c)^2 + w_B$  allows both red cells to meet at  $b(r_1, r_2)$  and leaves blue site  $b_1$  without a region on the power Voronoi diagram of  $S$ . In other words, since the three sites are collinear, bisectors  $b(r_1, b_1)$  and  $b(r_2, b_1)$  are parallel and so the only way to connect the red cells is to increase their weight until the blue cell of  $b_1$  no longer exists.

A variation of this problem can be attained by assigning each site a different weight or just assign different weights to red sites. Nevertheless, we suspect that this variation will result in an NP-hard problem. Furthermore, another optimisation problem associated with this last variation is associated with minimising the sum of weights of red sites.

**Acknowledgements.** This research began during the 8<sup>th</sup> Iberian Workshop on Computational Geometry held in Aveiro, Portugal, in 2011. Manuel Abellanas, Mercè Claverol and Gregorio Hernández were partially supported by the ESF EUROCORES programme EuroGIGA, CRP ComPoSe: MICINN Project EUI-EURC-2011-4306, for Spain. Mercè Claverol was also partially supported by projects MICINN MTM2009-07242 and Gen. Cat. DGR 2009SGR1040. Manuel Abellanas, Santiago Canales and Gregorio Hernández were partially supported by the ESF EUROCORES programme EuroGIGA, CRP Voronoi IP8, for Spain. Antonio L. Bajuelos and Inês Matos were supported by FCT through CIDMA of University of Aveiro. Inês Matos was also supported by grant SFRH/BPD/66572/2009.

## References

1. Aichholzer, O., Fabila-Monroy, R., Hackl, T., van Kreveld, M., Pilz, A., Ramos, P., Vogtenhuber, B.: Blocking delaunay triangulations. In: 22nd Canadian Conference on Computational Geometry, pp. 21–24 (2010)
2. Aurenhammer, F.: Voronoi diagrams a survey. Tech. rep., Institute for Information Processing Technical University of Graz (1988)
3. Aurenhammer, F., Edelsbrunner, H.: An optimal algorithm for constructing the weighted voronoi diagram in the plane. Pattern Recognition 17(2), 251–257 (1984)
4. Aurenhammer, F., Klein, R.: Voronoi diagrams. In: Sack, J.R., Urrutia, J. (eds.) Handbook of Computational Geometry, pp. 201–290. North-Holland, Amsterdam (2000)
5. Balaban, I.J.: An optimal algorithm for finding segments intersections. In: 11th Annual Symposium on Computational Geometry, SCG 1995, pp. 211–219. ACM, New York (1995)
6. de Berg, M., Gerrits, D., Khosravi, A., Rutter, I., Tsirogiannis, C., Wolff, A.: How alexander the great brought the greeks together while inflicting minimal damage to the barbarians. In: 26th European Workshop on Computational Geometry, pp. 73–77 (2010)
7. Golomb, S., Baumert, L.: Backtrack programming. Journal of ACM 12(4), 516–524 (1965)
8. Kaneko, A., Kano, M.: Discrete geometry on red and blue points in the plane – a survey. In: Discrete and Computational Geometry, The Goodman - Pollack Festschrift, pp. 551–570. Springer (2003)
9. Okabe, A., Boots, B., Sugihara, K., Chiu, S.N.: Spatial Tessellations - Concepts and Applications of Voronoi Diagrams, 2nd edn. Series in Probability and Statistics. John Wiley & Sons, Inc., New York (2000)

# Covering Islands in Plane Point Sets\*

Ruy Fabila-Monroy<sup>1</sup> and Clemens Huemer<sup>2</sup>

<sup>1</sup> Departamento de Matemáticas, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

[ruyfabila@math.cinvestav.edu.mx](mailto:ruyfabila@math.cinvestav.edu.mx)

<sup>2</sup> Departament de Matemàtica Aplicada IV, Universitat Politècnica de Catalunya  
[clemens.huemer@upc.edu](mailto:clemens.huemer@upc.edu)

**Abstract.** Let  $S$  be a set of  $n$  points in general position in the plane. A  $k$ -island  $I$  of  $S$  is a subset of  $k$  points of  $S$  such that  $\text{Conv}(I) \cap S = I$ . We show that, for an arbitrary but fixed number  $k \geq 2$ , the minimum number of  $k$ -islands among all sets  $S$  of  $n$  points is  $\Theta(n^2)$ . The following related counting problem is also studied: For  $l < k$ , an  $l$ -island covers a  $k$ -island if it is contained in the  $k$ -island. Let  $C_{k,l}(S)$  be the minimum number of  $l$ -islands needed to cover all the  $k$ -islands of  $S$  and let  $C_{k,l}(n)$  be the minimum of  $C_{k,l}(S)$  among all sets  $S$  of  $n$  points. We show asymptotic bounds for  $C_{k,l}(n)$ .

**Keywords:** planar point sets, islands, Horton sets.

## 1 Introduction

Let  $S$  be a set of  $n$  points in general position in the plane. An island  $I$  of  $S$  is a subset of  $S$  such that  $\text{Conv}(I)$ , the convex hull of  $I$ , does not contain points of  $S \setminus I$ . Problems related to islands have been studied recently in [4,5]. A  $k$ -island is an island  $I$  with  $|I| = k$ . The set of  $k$ -islands of  $S$  will be denoted as  $I_k(S)$ . Thus  $I_2(S)$  are the  $\binom{n}{2}$  segments connecting pairs of points of  $S$  and  $I_3(S)$  are the empty triangles of  $S$ . Throughout, we will use the  $O$ -,  $\Theta$ - and  $\Omega$ -notation to describe the asymptotic behavior. Katchalski and Meir [13] proved that the minimum number of empty triangles among all sets of  $n$  points is  $\Theta(n^2)$ . Lower and upper bounds for the exact leading constant of the quadratic term have then been improved in [2,3,8,16]. We show that also for  $k > 3$ , the minimum of  $|I_k(S)|$  among all sets  $S$  of  $n$  points is  $\Theta(n^2)$ . Note that for  $k > 3$ , a  $k$ -island might contain points in the interior of its convex hull. The special case when all the points of a  $k$ -island form a convex polygon, that is to say, they form an empty convex  $k$ -gon in the point set, has been studied extensively. For  $k = 4$ , the minimum number of empty convex  $k$ -gons is quadratic [3]. For  $k = 5$  and  $k = 6$  a linear lower bound is known [9,10,14,17]. A construction of  $n$  points

---

\* Research partially supported by Conacyt of Mexico, grant 153984 and MEC MTM2009-07242 and Gen.Cat. DGR 2009SGR1040 and ESF EUROCORES programme EuroGIGA - ComPoSe IP04 - MICINN Project EUI-EURC-2011-4306.

**Table 1.** The obtained bounds for  $C_{k,l}(n)$ 

value of $l$	$\leq \lfloor \frac{k+5}{6} \rfloor$	$\leq \lceil \frac{k}{2} \rceil$	$\geq \lceil \frac{k}{2} \rceil + 1$
$C_{k,l}(n)$	$\Theta(n)$	$O(n \log n)$	$\Omega(n \log n)$

without empty convex heptagons is due to Horton [11]. We will see that Horton sets also provide  $O(n^2)$ -examples for the minimum number of  $k$ -islands.

We then study coverings of  $k$ -islands. Some problems related to coverings were studied in [1,15]. For  $l < k$ , an  $l$ -island *covers* a  $k$ -island if it is contained in the  $k$ -island. Let  $C_{k,l}(S)$  be the minimum number of  $l$ -islands needed to cover  $I_k(S)$ . Let  $C_{k,l}(n)$  be the minimum of  $C_{k,l}(S)$  among all sets  $S$  of  $n$  points in general position in the plane. The problem to determine  $C_{k,l}(n)$  is closely related to counting empty convex polygons in point sets. For example, since every set of ten points contains an empty convex pentagon [10],  $C_{10,5}(n)$  is at most the minimum number of empty convex pentagons among all sets of  $n$  points. Also, the particular case  $C_{3,2}(n)$  has been studied recently in a different formulation in [6] by considering the question: What is the maximum number of edges that a geometric graph on a point set  $S$  can have such that it does not contain empty triangles? The results given there imply  $n - 2 + \lfloor \frac{n}{8} \rfloor \leq C_{3,2}(n) \leq O(n \log n)$ . The latter bound is achieved by the Horton set. We will see that also for each  $l \leq \lceil \frac{k}{2} \rceil$  the Horton set provides the upper bound  $O(n \log n)$  on  $C_{k,l}(n)$ . Table 1 summarizes our obtained bounds for  $C_{k,l}(n)$ .

Throughout the text,  $k \geq 2$  is an arbitrary but fixed natural number, and  $n$  is arbitrarily large with respect to  $k$ .

## 2 Islands of the Horton Set

A Horton set, see e.g. [2,3,7,11], is defined recursively as follows:  $H(1) = \{(1, 1)\}$  and  $H(2) = \{(1, 1), (2, 2)\}$ . When  $H(n)$  is defined, set

$$H(2n) = \{(2x - 1, y) \mid (x, y) \in H(n)\} \cup \{(2x, y + 3^n) \mid (x, y) \in H(n)\}. \quad (1)$$

We denote the subset of points of a Horton set  $H(n)$  with even x-coordinate as the upper set  $H^+(n)$ , and subset of  $H(n)$  with odd x-coordinate as the lower set  $H^-(n)$ . Horton sets have the following property: Any line connecting two points from  $H^+(n)$  leaves all points from  $H^-(n)$  below, and any line connecting two points from  $H^-(n)$  leaves all points from  $H^+(n)$  above.

We say that an island  $I$  of a point set  $S$  is *open from above* if the vertical stripe bounded by the leftmost point and rightmost point of  $I$  contains no points of  $S \setminus I$  above  $\text{Conv}(I)$ . Likewise  $I$  is *open from below* if there are no points of  $S \setminus I$  below  $\text{Conv}(I)$  in this stripe. Denote with  $I_k^+(S)$  and  $I_k^-(S)$  the set of  $k$ -islands of  $S$  open from above and open from below, respectively.

We estimate the number  $|I_k^+(H(n))|$  of  $k$ -islands of the Horton set  $H(n)$  open from above. The following lemma also applies to the number  $|I_k^-(H(n))|$  of  $k$ -islands open from below.

**Lemma 1.** *There exist positive constants  $c_k$  and  $c'_k$  that only depend on  $k$  such that  $c'_k n \leq |I_k^+(H(n))| \leq c_k n$ .*

*Proof.* The lower bound  $c'_k n \leq |I_k^+(H(n))|$  holds trivially, since any subset of  $k$  consecutive points in x-sorted order forms a  $k$ -island open from above. For the upper bound, consider the two Horton sets  $H^+(n)$  and  $H^-(n)$  that form  $H(n)$ . Each  $k$ -island open from above is either entirely contained in  $H^+(n)$  or it has points in  $H^+(n)$  and in  $H^-(n)$ . In the first case we can use recursion on  $H^+(n)$ . In the second case, assume  $i$  points of the  $k$ -island are in  $H^+(n)$  and  $k - i$  points are in  $H^-(n)$ . By the properties of a Horton set, these  $i$  points must be consecutive points in x-sorted order in  $H^+(n)$ . Hence, there are  $\frac{n}{2} - i + 1$  possible ways to choose these  $i$  points. Further, the  $k - i$  points of the  $k$ -island in the lower set have to be chosen among at most  $i + 1$  consecutive points of  $H^-(n)$ ;  $i - 1$  of them lie below the chosen points of  $H^+(n)$ , that is, within the vertical stripe defined by leftmost and rightmost point of the chosen points of  $H^+(n)$ . This also implies that  $i \geq \lfloor \frac{k}{2} \rfloor$ . The number of  $(k - i)$ -islands in this subset of  $i + 1$  points can be bounded by a constant  $d_k$ , because  $k$  is a constant. Hence, we have the following recurrence:

$$|I_k^+(H(n))| \leq |I_k^+(H(\frac{n}{2}))| + \sum_{i=\lfloor \frac{k}{2} \rfloor}^{k-1} (\frac{n}{2} - i + 1) d_k. \quad (2)$$

This further gives

$$|I_k^+(H(n))| \leq |I_k^+(H(\frac{n}{2}))| + d_{k'} \frac{n}{2}, \quad (3)$$

where  $d_{k'}$  is another constant. It then follows by induction on  $n$  that  $|I_k^+(H(n))| \leq c_k n$ , where  $c_k > d_{k'}$  is chosen big enough to also cover the base case of small values of  $n$ .  $\square$

**Lemma 2.** *The number of  $k$ -islands of the Horton set  $H(n)$  is  $\Theta(n^2)$ .*

*Proof.* The  $k$ -islands that are contained entirely in the upper set  $H(n)^+$  or in the lower set  $H(n)^-$  can be counted recursively. A  $k$ -island that has points in both of  $H(n)^+$  and  $H(n)^-$  consists of an  $i$ -island open from below in  $H(n)^+$  and a  $(k - i)$ -island open from above in  $H(n)^-$ , for some  $i \in \{1, \dots, k - 1\}$ . We thus obtain the recurrence for the number  $|I_k(H(n))|$  of  $k$ -islands of the Horton set:

$$|I_k(H(n))| = 2|I_k(H(\frac{n}{2}))| + \sum_{i=1}^{k-1} |I_i^-(H(\frac{n}{2}))| \cdot |I_{k-i}^+(H(\frac{n}{2}))|. \quad (4)$$

Since  $k$  is a constant,  $|I_i^-(H(\frac{n}{2}))|$  and  $|I_{k-i}^+(H(\frac{n}{2}))|$  are both in  $\Theta(n)$  by Lemma 1. Thus,

$$|I_k(H(n))| = 2|I_k(H(\frac{n}{2}))| + \Theta(n^2). \quad (5)$$

It follows that  $|I_k(H(n))|$  is  $\Theta(n^2)$ .  $\square$

**Lemma 3.** Let  $S$  be a set of  $n$  points in general position in the plane, and let  $k \geq 2$ . Then  $|I_k(S)| = \Omega(n^2)$ .

*Proof.* For a point  $p \in S$ , sort the points of  $S \setminus \{p\}$  cyclically around  $p$ . Divide these points into  $\Omega(n)$  groups of  $k - 1$  consecutive points in this order. Each group together with  $p$  forms a  $k$ -island. Repeating this for every point of  $S$ , we count  $\Omega(n^2)$   $k$ -islands; each one is counted at most  $k$  times.  $\square$

Lemmas 2 and 3 imply the following result:

**Theorem 1.** The minimum number of  $k$ -islands among all sets of  $n$  points in general position is  $\Theta(n^2)$ .

### 3 Covering Islands

We now show asymptotic bounds for  $C_{k,l}(n)$ .

**Theorem 2.** For  $l \leq \lfloor \frac{k+5}{6} \rfloor$ ,  $C_{k,l}(n)$  is  $\Theta(n)$ . For  $l \leq \lceil \frac{k}{2} \rceil$ ,  $C_{k,l}(n)$  is  $O(n \log n)$ . For  $l \geq \lceil \frac{k}{2} \rceil + 1$ ,  $C_{k,l}(n)$  is  $\Omega(n \log n)$ .

The proof is split into the following three cases.

- For  $l \leq \lfloor \frac{k+5}{6} \rfloor$ , there exist sets of  $n = ml$  points, for any  $m$ , such that their  $k$ -islands can be covered with  $m$   $l$ -islands.

*Proof.* The construction starts with a Horton set of  $m$  points; if  $m$  is not a power of 2 then take a larger Horton set and only consider its  $m$  leftmost points. Then  $l - 1$  additional points are placed in an  $\varepsilon$ -neighborhood of each point of the Horton set, where  $\varepsilon > 0$  is chosen small enough. It remains to provide a covering of all  $k$ -islands of this point set with  $m$   $l$ -islands: Choose each point of the given Horton set together with its  $l - 1$  additional nearest points as an  $l$ -island of the covering. Thus,  $m$   $l$ -islands are chosen. Assume the point set has a  $k$ -island  $I$  that is not covered by some chosen  $l$ -island. Note that  $k \geq 6(l - 1) + 1$ . Also note that  $I$  contains at most  $l - 1$  points of each chosen  $l$ -island. We now use the fact that Horton sets have no empty convex heptagons. Therefore,  $I$  cannot have points from more than six of the  $m$   $l$ -islands. But since  $k > 6(l - 1)$ , this gives a contradiction.  $\square$

- For  $l \leq \lceil \frac{k}{2} \rceil$ , the  $k$ -islands of the Horton set  $H(n)$  can be covered with  $O(n \log n)$   $l$ -islands.

*Proof.* To cover all the  $k$ -islands, choose all the  $l$ -islands open from below of  $H(n)^+$  and all the  $l$ -islands open from above of  $H(n)^-$ . By Lemma 1, the number of these  $l$ -islands is  $O(n)$ . Moreover, these  $l$ -islands cover all the  $k$ -islands that have points in both  $H(n)^+$  and  $H(n)^-$ . Recursively cover the  $k$ -islands of  $H(n)^+$  and  $H(n)^-$ . Let  $T(n)$  denote the number of  $l$ -islands in this covering. We obtain the recurrence:

$$T(n) \leq 2T\left(\frac{n}{2}\right) + O(n), \quad (6)$$

which gives  $O(n \log n)$  as an upper bound.  $\square$

- For  $l \geq \lceil \frac{k}{2} \rceil + 1$  and for every set  $S$  of  $n$  points, at least  $\Omega(n \log n)$   $l$ -islands are needed to cover all the  $k$ -islands of  $S$ .

*Proof.* Consider a halving line  $\ell_h$  for  $S$ . We want to find  $\Omega(n)$  pairwise disjoint  $k$ -islands crossing  $\ell_h$  such that at most  $\lceil \frac{k}{2} \rceil$  points of each island are on each side of  $\ell_h$ . But this follows easily as a special case of the Equitable Subdivision Theorem [12]: assume the points of  $S$  on one side of  $\ell_h$  are colored red; the remaining ones are colored blue. Possibly ignore the leftmost and rightmost points of  $S$  (up to  $\Theta(n)$  points) so that there are  $g \lceil \frac{k}{2} \rceil$  red points and  $g \lfloor \frac{k}{2} \rfloor$  blue points, where  $g$  is  $\Theta(n)$ . Then, there exists a subdivision of the plane into  $g$  disjoint convex polygons such that each of them contains exactly  $\lceil \frac{k}{2} \rceil$  red points and  $\lfloor \frac{k}{2} \rfloor$  blue points. Each such polygon gives a  $k$ -island. We thus have  $\Omega(n)$  pairwise disjoint  $k$ -islands crossing  $\ell_h$  and therefore need  $\Omega(n)$  pairwise disjoint  $l$ -islands to cover these  $k$ -islands. Since  $l \geq \lceil k/2 \rceil + 1$ , also each  $l$ -island crosses  $\ell_h$ . We iterate on both sides of  $\ell_h$  and obtain the following recurrence, which gives the claimed lower bound for  $C_{k,l}(n)$  of  $\Omega(n \log n)$ :

$$C_{k,l}(n) \geq 2C_{k,l}\left(\frac{n}{2}\right) + \Omega(n). \quad (7)$$

□

We finally determine  $C_{k,l}(H(n))$ , for  $l > \lceil \frac{k}{2} \rceil$ . Based on the following lemma, we believe that for  $l > \lceil \frac{k}{2} \rceil$ ,  $C_{k,l}(n)$  is  $\Theta(n^2)$ .

**Lemma 4.** *For  $l \geq \lceil \frac{k}{2} \rceil + 1$ , the number of  $l$ -islands needed to cover the  $k$ -islands of the Horton set  $H(n)$  is  $\Theta(n^2)$ .*

*Proof.* Divide the points of  $H(n)^+$  into  $\Theta(n)$  disjoint subsets of  $k-l+1$  consecutive points in x-sorted order. Also divide the points of  $H(n)^-$  into  $\Theta(n)$  disjoint subsets of  $l-1$  consecutive points. The union of any of these subsets of  $H(n)^+$  with any of these subsets of  $H(n)^-$  is a  $k$ -island. Note that an  $l$ -island covering one of these  $k$ -islands must have points in both  $H(n)^+$  and  $H(n)^-$ . Therefore, an  $l$ -island can only cover one of these  $k$ -islands and the result follows. □

## References

1. Aronov, B., Dulieu, M., Hurtado, F.: Witness (Delaunay) Graphs. Comput. Geom. 44, 329–344 (2011)
2. Bárány, I., Füredi, Z.: Empty simplices in Euclidean space. Canad. Math. Bull. 30, 436–445 (1987)
3. Bárány, I., Valtr, P.: Planar point sets with a small number of empty convex polygons. Stud. Sci. Math. Hung. 41, 243–269 (2004)
4. Bautista-Santiago, C., Cano, J., Fabila-Monroy, R., Flores-Péñaloza, D., González-Aguilar, H., Lara, D., Sarmiento, E., Urrutia, J.: On the Diameter of a Geometric Johnson Type Graph. In: 26th European Workshop on Computational Geometry, pp. 61–64 (2010)

5. Bautista-Santiago, C., Díaz-Báñez, J.M., Lara, D., Pérez-Lantero, P., Urrutia, J., Ventura, I.: Computing Maximal Islands. In: 25th European Workshop on Computational Geometry, pp. 333–336 (2009)
6. Bautista-Santiago, C., Heredia, M.A., Huemer, C., Ramírez-Vigueras, A., Seara, C., Urrutia, J.: On the number of edges in geometric graphs without empty triangles. Submitted to Journal (January 2011)
7. Devillers, O., Hurtado, F., Kàrolyi, G., Seara, C.: Chromatic variants of the Erdős-Szekeres Theorem. *Comput. Geom.* 26, 193–208 (2003)
8. Dumitrescu, A.: Planar sets with few empty convex polygons. *Stud. Sci. Math. Hung.* 36, 93–109 (2000)
9. Gerken, T.: Empty convex hexagons in planar point sets. *Discrete Comput. Geom.* 39, 239–272 (2008)
10. Harborth, H.: Konvexe Fünfecke in ebenen Punktmengen. *Elem. Math.* 33, 116–118 (1978)
11. Horton, J.D.: Sets with no empty convex 7-gons. *Canad. Math. Bull.* 26, 482–484 (1983)
12. Kaneko, A., Kano, M.: Discrete geometry on red and blue points in the plane — a survey. In: *Algorithms Combin.*, vol. 25, pp. 551–570. Springer (2003)
13. Katchalski, M., Meir, A.: On empty triangles determined by points in the plane. *Acta Math. Hung.* 51, 323–328 (1988)
14. Nicolás, C.M.: The empty hexagon theorem. *Discrete Comput. Geom.* 38, 389–397 (2007)
15. Sakai, T., Urrutia, J.: Covering the convex quadrilaterals of point sets. *Graphs Combin.* 23, 343–357 (2007)
16. Valtr, P.: On the minimum number of empty polygons in planar point sets. *Stud. Sci. Math. Hungar.* 30, 155–163 (1995)
17. Valtr, P.: On empty hexagons. In: *Surveys on Discrete and Computational Geometry, Twenty Years Later*, pp. 433–441. AMS (2008)

# Rectilinear Convex Hull with Minimum Area

Carlos Alegria-Galicia<sup>3</sup>, Tzolkin Garduño<sup>3</sup>, Areli Rosas-Navarrete<sup>3</sup>,  
Carlos Seara<sup>2,\*</sup>, and Jorge Urrutia<sup>1,\*\*</sup>

<sup>1</sup> Universidad Nacional Autónoma de México, Instituto de Matemáticas

<sup>2</sup> Universitat Politècnica de Catalunya

<sup>3</sup> Universidad Nacional Autónoma de México, Posgrado en Ciencia e  
Ingeniería de la Computación

{alegria\_c,garduno\_t,areli}@uxmcc2.iimas.unam.mx, carlos.seara@upc.edu,  
urrutia@matem.unam.mx

**Abstract.** Let  $P$  be a set of  $n$  points in the plane. We solve the problem of computing an orientation of the plane for which the rectilinear convex hull of  $P$  has minimum area in optimal  $\Theta(n \log n)$  time and  $O(n)$  space.

**Keywords:** rectilinear convex hull, ortho-convexity, optimization.

## 1 Introduction

The interest in the rectilinear convex hull of planar point sets arises from the study of the *ortho-convexity* [15], also known in the literature as *rectilinear*, *x-y*, or *orthogonal* convexity. This non-traditional notion of convexity has been widely studied since its formalization in the early eighties, and has found applications in theoretical research fields such as polyhedra reconstruction [4] and fixed point theory [9], as well as in the study of practical problems such as digital image shape analysis [5] and VLSI circuit layout design [16]. See also [7,8,11].

A set  $S \subset \mathbb{R}^2$  is said to be *ortho-convex* if the intersection of  $S$  with every horizontal or vertical line is connected, i.e., it is empty, a point, or a single interval. Unlike ordinary convex sets, an orthogonally convex set is not necessarily connected. The orthogonal convex hull of a set  $S$  is the intersection of all connected orthogonally convex supersets of  $S$ . So every convex set is orthogonally convex but not vice versa. For the same reason, the orthogonal convex hull itself is a subset of the convex hull of the same point set. A point  $p$  belongs to the orthogonal convex hull of  $S$  if and only if each of the closed axis-aligned quadrants having  $p$  as apex has a nonempty intersection with  $S$ . The orthogonal convex hull is also known as the rectilinear convex hull, or the *x-y* convex hull. As it is easy to see, an ortho-convex set might be non-convex or disconnected. The latter property hinders the definition of an ortho-convex closure that resembles

---

\* Supported by MTM2009-07242, Gen Cat DGR2009GR1040, and ESF EUROCORES programme EuroGIGA-ComPoSe IP04-MICINN Project EUI-EURC-2011-4306.

\*\* Partially supported by projects MTM2006-03909 (Spain), and SEP-CONACYT of México, Proyecto 80268.

the properties of the traditional convex hull and as a consequence, several definitions have been presented by different authors [13]. Throughout this paper, we will use the *maximal ortho-convex*, or *mr-convex* hull stated by Ottmann et al. [13] (see also Matousěk et al. [12]), which is defined as follows.

An *orthogonal wedge* is the intersection of two open half-planes whose supporting lines are orthogonal. Let  $X_\theta$  and  $Y_\theta$  denote the  $X$  and  $Y$  coordinate axes rotated counter-clockwise by  $\theta$  degrees, respectively. We call  $\theta$ -orientation to the coordinate system formed by  $X_\theta$  and  $Y_\theta$ . An *orthogonal  $\theta$ -wedge* is an orthogonal wedge whose supporting lines are parallel to the axes of the first quadrant of the  $\theta$ -orientation of the plane.

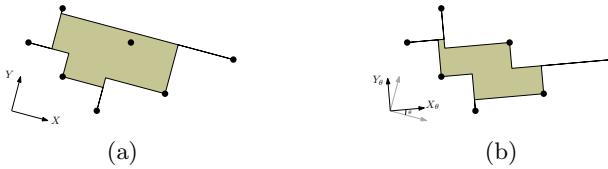
Let  $P \subset \mathbb{R}^2$  denote a set of  $n$  points in general position i.e., no three collinear points in  $P$ . An orthogonal  $\theta$ -wedge is called  $P$ -free if it does not contain points of  $P$ . Let  $W_\theta(P)$  be the set of all  $P$ -free orthogonal  $\theta$ -wedges, and let

$$\mathcal{W}_\theta(P) = W_\theta(P) \cup W_{\theta+\frac{\pi}{2}}(P) \cup W_{\theta+\pi}(P) \cup W_{\theta+\frac{3}{2}\pi}(P).$$

The *Rectilinear Convex Hull* of  $P$  with orientation  $\theta$  is the set

$$\mathcal{RH}_\theta(P) = \mathbb{R}^2 - \bigcup_{w \in \mathcal{W}_\theta(P)} w.$$

Figure 1(a) shows the rectilinear convex hull of a point set. The set  $\mathcal{W}_\theta(P)$  changes along with  $\theta$  and thus,  $\mathcal{RH}_\theta(P)$  is an orientation-dependent shape (Figure 1(b)). From this property, several optimization problems naturally arise for  $\mathcal{RH}_\theta(P)$  when considering the complete set of  $\theta$ -orientations of the plane.



**Fig. 1.**  $\mathcal{RH}_\theta(P)$  changes along with  $\theta$

In this paper we solve the problem of computing all the values of  $\theta \in [0, 2\pi]$  for which  $\mathcal{RH}_\theta(P)$  has minimum area in optimal  $\Theta(n \log n)$  time and  $O(n)$  space. We show that the set of orientations  $\{\theta : \theta \in [0, 2\pi]\}$  can be divided into a linear sequence,  $\{I_1, I_2, \dots, I_{O(n)}\}$ , of orientation intervals such that: (1) we can compute the orientation  $\theta \in I_1$  which minimizes the area of  $\mathcal{RH}_\theta(P)$  in  $O(n)$  time; (2) We perform the computation of the orientation of optimal area for the subsequent intervals in constant time. To obtain the set of orientation intervals we use an optimal  $\Theta(n \log n)$  time and  $O(n)$  space algorithm. Our work is based on techniques from Bae et al. [2], Avis et al. [1], and Díaz-Báñez et al. [6]. Our result improves the  $O(n^2)$  time complexity solution presented recently by Bae et al. [2].

## 2 The Un-oriented Rectilinear Convex Hull

In this section we present an algorithm to compute and maintain  $\mathcal{RH}_\theta(P)$  over all the  $\theta$ -orientations of the plane in  $\Theta(n \log n)$  time and  $O(n)$  space.

### 2.1 Preliminaries

In [10,14] the Set Maxima Problem for a point set  $P$  in the plane is considered. Concretely, given two points  $p_i, p_j \in P$ , the following dominance relation is established:  $p_i$  dominates  $p_j$  ( $p_j \prec p_i$ ), if  $x_j \leq x_i$  and  $y_j \leq y_i$ . The relation  $\prec$  is a partial order in  $P$ . A point  $p_i \in P$  is called *maximal* if there does not exist  $p_j \in P$  such that  $i \neq j$  and  $p_i \prec p_j$ . The maxima problem consists of finding all the maximal points of  $P$  under dominance. One can formulate maxima problems for each quadrant in the plane. Each set of maximal points has a *total ordering* and can be organized as a height balanced search tree. The maxima problem for  $P$  with respect to any of the four quadrants can be solved optimally in  $\Theta(n \log n)$  time and  $O(n)$  space [10,14]. The set of maxima points for  $P$  with respect to the four quadrants forms the rectilinear convex hull of  $P$ .

A *vertex* of  $\mathcal{RH}_\theta(P)$  is a point in  $P$  that lies on the boundary of  $\mathcal{RH}_\theta(P)$ . For any fixed  $\theta$ , the computation of  $\mathcal{RH}_\theta(P)$  has a close relation to the *Set Maxima Problem* [10,14], since the set of vertices of  $\mathcal{RH}_\theta(P)$  is equal to the set  $\mathcal{V}_\theta(P)$  of maximal points of  $P$  with respect to vector dominance in the four quadrants defined by  $X_\theta$  and  $Y_\theta$  [2,6,13].

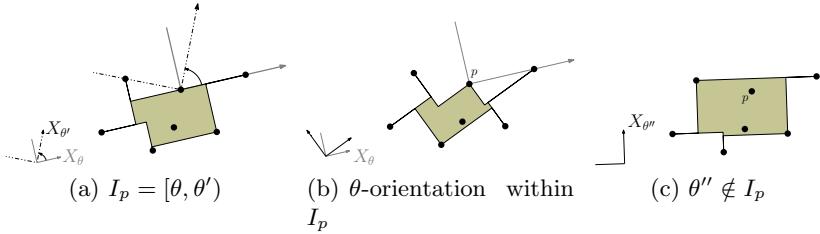
Since we can specify  $\mathcal{RH}_\theta(P)$  using  $\mathcal{V}_\theta(P)$ , it is possible to keep track of the changes in the set of vertices of  $\mathcal{RH}_\theta(P)$  over all  $\theta$ -orientations in the plane, by maintaining  $\mathcal{V}_\theta(P)$  as the coordinate axes rotates counter-clockwise. The maintenance of this set can be done using the results given in [1,6] in the manner we describe bellow.

We call *apex* of an orthogonal wedge to the intersection point of its supporting lines. Every point  $p \in \mathcal{V}_\theta(P)$  is the apex of a  $P$ -free orthogonal  $\theta$ -wedge contained in  $\mathcal{W}_\theta(P)$ . Figure 2 shows us a point  $p$  that is vertex of  $\mathcal{RH}_\theta(P)$  in the interval  $I_p = [\theta, \theta')$ , and the orthogonal  $\theta$  and  $\theta'$ -wedges with apex in  $p$ . The endpoints of  $I_p$  mark the *in-* and *out-*events of  $p$ , i.e., the values of  $\theta$  when  $p$  belongs and stops being in  $\mathcal{V}_\theta(P)$ . Moreover, a point can be the apex of an orthogonal  $\theta$ -wedge during at most three intervals of  $\theta$ -orientations. The next Theorem 1 appears in [6] not as a specific result but as a step of an  $\Theta(n \log n)$  time and  $O(n)$  space algorithm of a fitting problem studied by the authors.

**Theorem 1 (Díaz-Bañez et al. [6]).** *To compute and maintain  $\mathcal{RH}_\theta(P)$  over all  $\theta$ -orientations can be done in optimal  $\Theta(n \log n)$  time and  $O(n)$  space.*

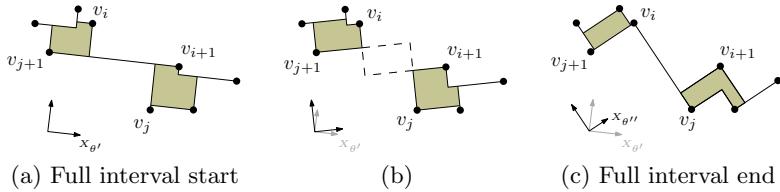
To obtain Theorem 1, the authors compute first the set of orientation intervals in which the elements of  $P$  are maximal with respect to the four quadrants defined by  $X_\theta$  and  $Y_\theta$  (using a result from Avis et al. [1]), and also the ordered set of in- and out-events. Both sets have a linear number of elements.

Let us denote by  $V_\theta(P)$  to the set of maximal points of  $P$  in the first quadrant defined by  $X_\theta$  and  $Y_\theta$ . Note that  $\mathcal{V}_\theta(P) = V_\theta(P) \cup V_{\theta+\frac{\pi}{2}}(P) \cup V_{\theta+\pi}(P) \cup V_{\theta+\frac{3\pi}{4}}(P)$ .



**Fig. 2.** A point  $p \in P$ , its orientation interval  $I_p$  and  $R H_{\theta''}(P)$

Since  $R H_\theta(P)$  is an ortho-convex set, it is monotone with respect to  $X_\theta$ . Thus, the points of  $P$  in  $V_\theta(P)$  can be relabelled as  $v_1, \dots, v_m$  in increasing order according to  $X_\theta$ . Let  $w_i(\theta)$  denote the  $P$ -free orthogonal  $\theta$ -wedge supported by two consecutive points  $v_i, v_{i+1} \in V_\theta(P)$ . We say that the wedges  $w_i(\theta)$  and  $w_j(\theta + \pi)$  are *opposite*. If the intersection of  $w_i(\theta)$  and  $w_j(\theta + \pi)$  is nonempty, then  $R H_\theta(P)$  is disconnected (Figure 3b). In such case we say that  $w_i(\theta)$  and  $w_j(\theta + \pi)$  *overlap*, and denote its intersection as  $O_\theta(i, j)$ .



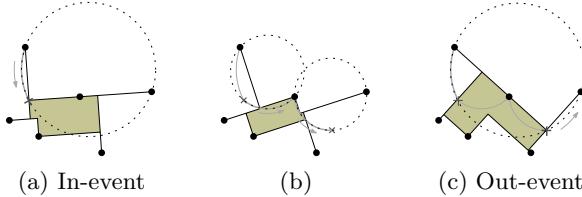
**Fig. 3.** An overlap disconnects  $R H_\theta(P)$ . The ends of the overlap full interval.

Let  $\mathcal{O}_\theta(P)$  be the set of all overlaps  $O_\theta(i, j)$  of  $R H_\theta(P)$  in the  $\theta$ -orientation.  $\mathcal{O}_\theta(P)$  can be computed from  $V_\theta(P)$  in  $O(n)$  time, and  $R H_\theta(P)$  can be computed in  $\Theta(n \log n)$  time and  $O(n)$  space [10,14] for a fixed value of  $\theta$ . To maintain the set  $\mathcal{O}_\theta(P)$  we will compute an ordered list of angles as the one computed for the in- and out- events. The overlap events do not necessarily match up with in- or out-events [2], thus they have to be computed independently.

## 2.2 Overlaps Characterization

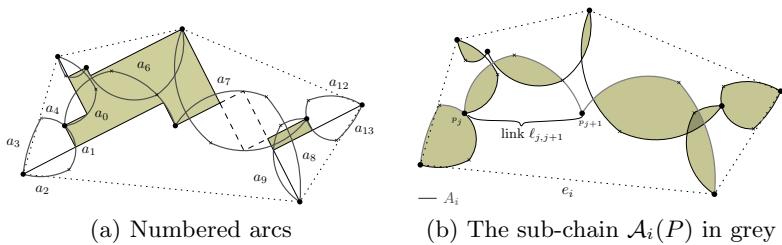
As the  $\theta$ -orientation rotates from 0 to  $2\pi$ , the apex of every  $P$ -free orthogonal  $\theta$ -wedge supported by at least two points in  $V_\theta(P)$  traces an oriented circular arc (Figure 4).

**Definition 1.** *The arc-chain  $\mathcal{A}(P)$  of  $P$  is the closed curve formed by the union of the set of arcs traced by the apices of the  $P$ -free orthogonal  $\theta$ -wedges supported by two consecutive elements of  $V_\theta(P)$  for some  $\theta \in [0, 2\pi)$  (Figure 5).*



**Fig. 4.** The apices path as the  $\theta$ -orientation changes

The arc-chain can be computed in  $O(n)$  time using the in- and out-events list: every arc is a segment of a circumference with diameter in consecutive points of the sets  $V_\theta(P)$ ,  $V_{\theta+\frac{\pi}{2}}(P)$ ,  $V_{\theta+\pi}(P)$ ,  $V_{\theta+\frac{3\pi}{4}}(P)$ ; and its ends correspond to the in- and out-events. Thus, each arc can be associated with an orientation interval in which, the wedge that traces the arc is part of the rectilinear convex hull.



**Fig. 5.** The arc-chain  $\mathcal{A}(P)$

Let  $\{e_1, \dots, e_h\}$  be the set of edges of the convex hull of  $P$ ,  $\mathcal{CH}(P)$ , in counter-clockwise order. A *sub-chain*  $A_i(P) \in \mathcal{A}(P)$  is a subsequence of consecutive elements of  $\mathcal{A}(P)$ , whose endpoints are the endpoints of  $e_i$ .  $A_i(P)$  is monotone in the direction determined by  $e_i$ . Thus, the orthogonal projection of  $A_i(P)$  on  $e_i$  defines a total order  $\prec_i$  on the set of endpoints of its arcs. Using the convex hull edges and  $\prec_i$ , we can label the arcs in  $\mathcal{A}(P)$  in counter-clockwise order, in such a way that  $\mathcal{A}(P) = \langle a_1, \dots, a_l \rangle$  (Figure 5a).

Assume that the ends of the arcs in any sub-chain  $A_i(P)$  are labelled as  $p_1, \dots, p_m$  such that if  $r < s$ , then  $p_r \prec_i p_s$ . A *link*  $\ell_{r,s}$  is a subsequence  $p_r, \dots, p_s$  of  $A_i(P)$  such that for  $r < t < s$ ,  $p_r, p_s \in P$ ,  $p_r \prec_i p_t \prec_i p_s$  and  $p_t \notin P$ .

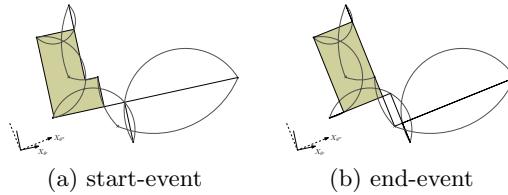
*Overlap characterization using the arc-chain.* Let  $v_i, v_{i+1} \in V_\theta(P)$  and let  $v_j, v_{j+1} \in V_{\theta+\pi}(P)$  that support two overlapping wedges  $w_i(\theta)$  and  $w_j(\theta + \pi)$ , respectively. If two opposite wedges in  $\mathcal{W}_\theta(P)$  overlap, since these wedges are  $P$ -free, the arcs traced by their apices belong to *links* that intersect (Figure 5a). The overlap between  $w_i(\theta)$  and  $w_j(\theta + \pi)$  is possible only in an interval  $[\theta', \theta'']$  where the  $\theta'$ -orientation is parallel to the segment  $\overline{v_{j+1}v_{i+1}}$ , and in the  $\theta''$ -orientation the  $Y$ -axis is parallel to the segment  $\overline{v_iv_j}$  (Figures 3a and 3c). We call *full overlap interval* to this orientation interval and we say that this interval is a proper interval when  $\theta' \leq \theta''$ .

Any arc  $a_j$  corresponds to a particular  $P$ -free orthogonal  $\theta$ -wedge  $w_j(\theta)$  for every  $\theta$  in some interval  $I_j = [\theta_{j1}, \theta_{j2}]$ . An overlap is possible only between opposite wedges. To determine if two given arcs  $a_j, a_k$  admit opposite wedges, we define the *facing interval* for them as the interval  $I_f = [\theta_{j1} + \pi, \theta_{j2} + \pi] \cap [\theta_{k1}, \theta_{k2}]$ .

Let  $a_j$  and  $a_k$  be arcs with their respective intervals  $I_j = [\theta_{j1}, \theta_{j2}]$  and  $I_k = [\theta_{k1}, \theta_{k2}]$ . Let  $w_j(\theta)$  denote the orthogonal  $\theta$ -wedge whose apex traces  $a_j$  for  $\theta \in I_j$  and let  $w_k(\theta)$  denote the respective wedge of  $a_k$  for any  $\theta \in I_k$ .

**Lemma 1.** (Overlap in terms of arcs) *The wedges  $w_j(\theta)$  and  $w_k(\theta)$  overlap if and only if the full overlap interval is proper, the facing interval is not empty, and the intersection between the full overlap interval and the facing interval is a proper interval.*

Thus, if there is an overlap between the wedges of given arcs, the overlap interval is the intersection between the full interval, the facing interval, and the intervals of both arcs. Let  $I_{i,j}$  be the interval of  $\theta$  for which  $w_i(\theta)$  and  $w_j(\theta + \pi)$  overlap. We call the ends of  $I_{i,j}$  the *start-* and *stop-* events of  $O_\theta(i, j)$ . The thick arc segments show the overlap interval in Figure 6.



**Fig. 6.** The orientation interval of an overlap  $I_{i,j} = (\theta', \theta'')$

If an overlap takes place, the links containing the arcs involved in the overlap had intersected. Thus, it is possible to compute the overlaps and the start- and end-events by computing the intersection points between links in the arc-chain. Next we show that the number of those intersection points is at most  $O(n)$ .

**Lemma 2.** *Let  $a, b, c \in A_i(P)$  such that  $a \prec_i b \prec_i c$ . Then,  $\frac{\pi}{2} \leq \angle abc < \pi$ .*

As a consequence of Lemma 2, any link  $\ell_{r,s}$  is contained in the disk that has the segment  $\overline{p_r p_s}$  as diameter, the link-disk  $D_{r,s}$ . Thus the number of links that intersect the link-disk is greater than or equal to the number of links that intersect  $\ell_{r,s}$ . We denote by  $d(r, s)$  to the *diameter* of  $\ell_{r,s}$ . Let  $\mathcal{D}(P) = \{D_0, D_1, \dots, D_h\}$  be the set of disks such that  $D_i$  has the convex hull edge  $e_i$  as a diameter and radius  $r_i$ . By Theorem 1 in [1] we have.

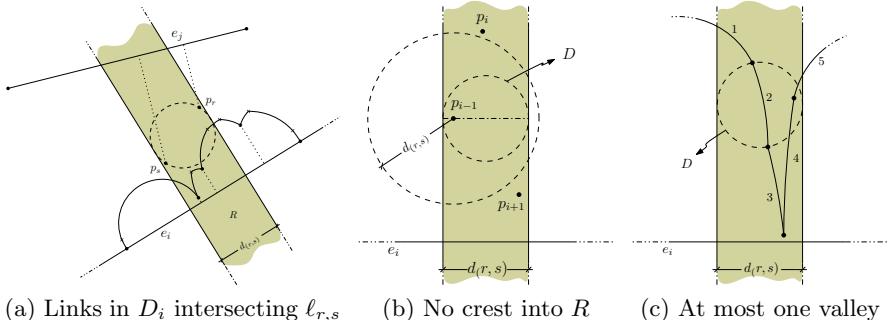
**Lemma 3.** *Given  $p, q \in P$ , the disk that has  $\overline{pq}$  as diameter and radius  $r$  is intersected by at most 28 disks  $D_i \in \mathcal{D}(P)$  such that  $r_i \geq r$ .*

For each of the disks in  $\mathcal{D}(P)$  that intersect a link-disk  $D_{r,s}$ , there is a constant number of links with diameter greater than  $d(r, s)$  that intersect  $\ell_{r,s}$ .

**Lemma 4.** Let  $D_i \in \mathcal{D}(P)$  be a disk of radius  $r_i$  that intersects some given link-disk  $D_{r,s}$  of radius  $d(r,s) < r_i$ . There are at most five links in  $D_i$  that intersect  $\ell_{r,s}$  with diameter greater than  $d(r,s)$ .

*Proof.* Let  $R$  be the strip region bounded by the lines that project  $D_{r,s}$  on the line containing  $e_i$  (Figure 7a). Since sub-chains are monotone, the links in  $A_i(P)$  that intersect  $D_{r,s}$  must have at least one endpoint in  $R$ .

Consider the links of  $A_i(P)$  with endpoints in  $R$ . Suppose that all these links have diameter greater than  $d(r,s)$ . In order to have many links into region  $R$ , the endpoints of these links should be placed as the configurations showed in Figure 7. The configuration in Figure 7b considers three points  $p_{i-1}, p_i, p_{i+1}$  such that  $p_i$  is a *peak* (i.e., the distances from the other points to  $e_i$  are smaller than the distance between  $p_i$  and  $e_i$ ). In this case, with the help of a circle with diameter parallel to  $e_i$  of length  $d(r,s)$  that passes through  $p_{i-1}$ , it is possible to show that: if  $p_i$  must be at distance greater than  $d(r,s)$  from  $p_{i-1}$  and from  $p_{i+1}$ , then  $p_i$  cannot be part of a link in  $A_i(P)$ , because the angle  $\angle p_{i-1}p_ip_{i+1}$  would be less than  $\frac{\pi}{2}$ , contradicting Lemma 2. Thus, since there are no crests into region  $R$ , there is at most one *valley* (Figure 7c). Therefore there are at most five links in  $A_i(P)$  that intersect  $\ell_{r,s}$ .



**Fig. 7.** A link  $\ell_{j,k}$  can be intersected by at most five links contained in some  $D_i \in \mathcal{D}(P)$

The open area bounded by  $A_i(P)$  and  $e_i$  is  $P$ -free since it is covered by  $P$ -free  $\theta$ -wedges. Thus, two intersecting links have at least two intersection points, and by Lemma 2 this number is tight, as none of the intersecting links can cross a line segment joining its intersection points. From lemmas 2, 3 and 4 we have.

**Lemma 5.** Any link  $\ell_{j,k}$  is intersected at most 280 times by links of diameter greater than or equal to  $d(j,k)$ .

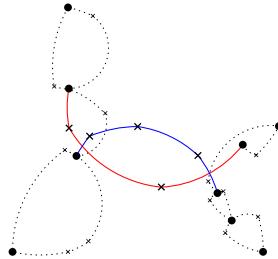
The following result is a central tool for computing the start- and stop-events list in  $O(n \log n)$  time.

**Theorem 2.** There are  $O(n)$  intersections between links in  $\mathcal{A}(P)$ .

### 2.3 Computing the Start- and Stop-Events

We color the arcs in  $\mathcal{A}(P)$  in the following way: red if the arc belongs to a sub-chain that corresponds to an edge in the upper chain of the convex hull, and the arc will be blue otherwise. Then we compute the bichromatic intersections using Bentley-Ottmann plane sweep algorithm [3] in  $O(n \log n)$  time and  $O(n)$  space.

Every intersection point will be associated with the respective links. After sorting these intersection points lexicographically according to the labels of the links involved in the intersection, we will be able to track corresponding pairs of links (Figure 8), since these links correspond to a couple of intersection points that involve the same links. We say that each of these couples form an *alternate link*. For each alternate link, let  $A_\alpha = \{\alpha_1, \dots, \alpha_v\}$  ( $A_\beta = \{\beta_1, \dots, \beta_w\}$ ) be the set of red (blue) arcs comprised in the alternate link. Let us recall that an arc  $\alpha_i$  is associated with an interval of orientations  $(\theta_{\alpha i_1}, \theta_{\alpha i_2})$ . We say that  $\theta_{\alpha i_1}$  is the *left endpoint* and  $\theta_{\alpha i_2}$  is the *right endpoint* of the interval. Using these intervals we can verify in linear time if the conditions of Lemma 1 are fulfilled.



**Fig. 8.** A couple of links forming an *alternate link*

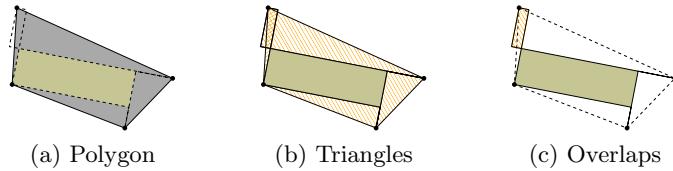
### 3 Minimum Area Algorithm

The list of event points obtained in the previous section generate a set of orientation intervals  $(\theta_1, \theta_2)$ , in which the set of vertices of  $\mathcal{RH}_\theta(P)$  and the set of overlaps in  $\mathcal{RH}_\theta(P)$  remains unchanged. For any  $\theta \in (\theta_1, \theta_2)$ , the area of  $\mathcal{RH}_\theta(P)$  is given by the following formula in [2]:

$$\text{area}(\mathcal{RH}_\theta(P)) = \text{area}(\mathcal{P}) - \sum_i \text{area}(\triangle_i(\theta)) + \sum_j \text{area}(\square_j(\theta)).$$

The formula has three main components: the first is a polygon constructed by joining counterclockwise consecutive vertices of  $\mathcal{RH}_\theta(P)$  (Figure 9a), the second is the sum of the areas of triangles formed by consecutive vertices of  $\mathcal{RH}_\theta(P)$  and the apex of the  $P$ -free  $\theta$ -wedge that they support (Figure 9b), finally the third one is the sum of the overlaps area (Figure 9c).

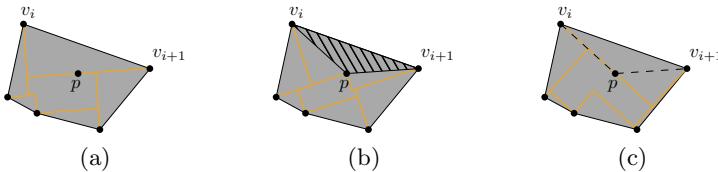
In [2] the authors showed that  $\text{area}(\triangle_i(\theta))$  and  $\text{area}(\square_j(\theta))$  of  $\mathcal{RH}_\theta(P)$  can be expressed as a function of  $\sin 2\theta$  and  $\cos 2\theta$ ; and used the derivatives to compute the value  $\theta \in (\theta_1, \theta_2)$  for which the area of  $\mathcal{RH}_\theta(P)$  is minimized:

**Fig. 9.** Three main components

$$\sum \text{area}'(\triangle_i(\theta)) = -[\sum A_i] \sin 2\theta + [\sum B_i] \cos 2\theta, \quad (1)$$

$$\sum \text{area}'(\square_j(\theta)) = [\sum C_i] \cos 2\theta - [\sum D_i] \sin 2\theta. \quad (2)$$

It takes  $O(n)$  time to compute the area of the polygon for the first orientation interval. This polygon changes with each in- and out-event. A point  $p$  is getting in the  $\mathcal{RH}_\theta(P)$  between points  $v_i, v_{i+1}$  (Figure 10a). The area of the polygon that includes  $p$  can be computed in constant time from the area of the previous polygon by subtracting the area of the triangle  $\triangle(v_i, p, v_{i+1})$  (Figure 10b). In the out-event of  $p$ , the area can be computed adding the area of the same triangle (Figure 10c).

**Fig. 10.** Updating the area of the polygon

Using the in- and out-event list, it is possible to compute the set of triangles that conform the second element of the formula. This list has a linear number of changes, since on each in-event two triangles are added to the list and a triangle is deleted from it and the same amount of triangles changes on each out-event. The areas of the overlaps can be computed in a way similar, but using the ordered list of the linear number of start- and stop-events. For each interval without changes, we update the formula in constant time by subtracting or adding new constant values. If there are more than one  $\theta$ -orientation in which  $\mathcal{RH}_\theta(P)$  has minimum area, our algorithm reports all of them. From the discussion above and since  $\mathcal{CH}(P)$  can be computed from  $\mathcal{RH}_\theta(P)$  in  $O(n)$  time, we obtain:

**Theorem 3.** *Computing the orientations for which the rectilinear convex hull of  $P$  has minimum area can be done in optimal  $\Theta(n \log n)$  time and  $O(n)$  space.*

## References

1. Avis, D., Beresford-Smith, B., Devroye, L., Elgindy, H., Guévremont, E., Hurtado, F., Zhu, B.: Unoriented  $\Theta$ -maxima in the plane: complexity and algorithms. *SIAM J. Comput.* 28(1), 278–296 (1999)
2. Bae, S.W., Lee, C., Ahn, H.-K., Choi, S., Chwa, K.-Y.: Computing minimum-area rectilinear convex hull and L-shape. *Computational Geometry: Theory and Applications* 42(3), 903–912 (2009)
3. Bentley, J., Ottmann, T.: Algorithms for reporting and counting geometric intersections. *IEEE Trans. Computers* C-28, 643–647 (1979)
4. Biedl, T., Genc, B.: Reconstructing orthogonal polyhedra from putative vertex sets. *Computational Geometry: Theory and Applications* 44(8), 409–417 (2011)
5. Biswas, A., Bhowmick, P., Sarkar, M., Bhattacharya, B.B.: Finding the Orthogonal Hull of a Digital Object: A Combinatorial Approach. In: Brimkov, V.E., Barneva, R.P., Hauptman, H.A. (eds.) IWCIA 2008. LNCS, vol. 4958, pp. 124–135. Springer, Heidelberg (2008)
6. Díaz-Báñez, J.M., López, M.A., Mora, M., Seara, C., Ventura, I.: Fitting a two-joint orthogonal chain to a point set. *Computational Geometry: Theory and Applications* 44(3), 135–147 (2011)
7. Fink, E., Wood, D.: Strong restricted-orientation convexity. *Geometriae Dedicata* 69, 35–51 (1998)
8. Franěk, V., Matousěk, J.: Computing D-convex hulls in the plane. *Computational Geometry: Theory and Applications* 42, 81–89 (2009)
9. Khamsi, M.A., Kirk, W.A.: An introduction to metric spaces and fixed point theory. Wiley-Interscience (2001)
10. Kung, H.T., Luccio, F., Preparata, F.P.: On finding the maxima of a set of vectors. *Journal of the ACM* 22, 469–476 (1975)
11. Martynchik, V., Metelski, N., Wood, D.: O-convexity: computing hulls, approximations, and orientation sets. In: Canadian Conference on Computational Geometry, pp. 2–7 (1996)
12. Matoušek, J., Plecháč, P.: On functional separately convex hulls. *Discrete and Computational Geometry* 19, 105–130 (1998)
13. Ottman, T., Soisalon-Soininen, E., Wood, D.: On the definition and computation of rectilinear convex hulls. *Information Sciences* 33, 157–171 (1984)
14. Preparata, F.P., Shamos, M.I.: *Computational Geometry: An Introduction*. Springer (1985)
15. Rawlins, G.J.E., Wood, D.: Ortho-convexity and its generalizations. In: Computational Morphology: A Computational Geometric Approach to the Analysis of Form, pp. 137–152. Elsevier Science Publishers B.V., North-Holland (1988)
16. Uchoa, E., De Aragão, M.P., Ribeiro, C.C.: Preprocessing Steiner problems from VLSI layout. *Networks* 40(1), 38–50 (2002)

# Separated Matchings and Small Discrepancy Colorings<sup>\*,\*\*</sup>

Viola Mészáros

Institute for Mathematics, TU Berlin,  
Sekretariat MA 6-1, Strasse des 17.  
Juni 136, 10623 Berlin, Germany  
[meszaros@math.tu-berlin.de](mailto:meszaros@math.tu-berlin.de)  
University of Szeged, Bolyai Institute,  
Aradi vértanúk tere 1, 6720 Szeged,  
Hungary  
[viola@math.u-szeged.hu](mailto:viola@math.u-szeged.hu)

**Abstract.** Consider  $2n$  points in the plane in convex position,  $n$  points are red and  $n$  points are blue. Edges are straight line segments connecting points of different color. A separated matching is a geometrically non-crossing matching where all edges can be crossed by a line. Separated matchings are closely related to non-crossing, alternating paths. Abellanas et al. and independently Kynčl et al. constructed convex point sets allowing at most  $\frac{4}{3}n + O(\sqrt{n})$  points on any non-crossing, alternating path. We give a class of configurations that contains at most  $\frac{4}{3}n + O(\sqrt{n})$  points in any separated matching. We also present a coloring with constant discrepancy parameter where the number of points in the maximum separated matching is very close to  $\frac{4}{3}n$ . When the discrepancy is at most three we show that there are at least  $\frac{4}{3}n$  points in the maximum separated matching.

## 1 Introduction

Consider a  $2n$ -element point set with a balanced coloring ( $n$  points red and  $n$  points blue) in the plane. Edges will be straight line segments connecting points of different color. Erdős posed the following problem: How many points are there on the longest non-crossing, alternating path in an arbitrary balanced  $2n$ -element convex point set in the plane? Without loss of generality we may assume that the points are on a circle  $C$ .

Erdős constructed a convex point set that allows at most  $\frac{3n}{2} + 2$  points on the longest non-crossing, alternating path. He conjectured that his configuration was asymptotically extremal. Erdős' conjecture was disproved. Kynčl, Pach and Tóth gave a single construction in 2008 [7] showing the  $\frac{4}{3}n + O(\sqrt{n})$  upper bound. In the same paper they proved the  $n + \Omega(\sqrt{n}/\log n)$  lower bound as well. Abellanas

---

\* Dedicated to Ferran Hurtado on the occasion of his 60th birthday.

\*\* Work on this paper was partially supported by ESF EuroGiga project ComPoSe (IP03), by OTKA Grant K76099 and by OTKA Grant 102029.

et al. found a similar construction independently of the previously mentioned researchers [2]. The presented upper bound is conjectured to be asymptotically tight. Hajnal and Mészáros improved the lower bound to  $n + \Omega(\sqrt{n})$ . They also constructed a class of configurations showing the  $\frac{4}{3}n + O(\sqrt{n})$  upper bound [5].

In the non-convex version of the problem you may find results in the following papers [1], [3], [4] and [6]. There is not much known about the general position. More researchers investigated some special positions of points. One of the early results was proved by Abellanas et al. stating that in case the color classes are separated by a line there is a non-crossing, alternating Hamiltonian path [1]. Another result was gained on a specific non-convex point set called *double-chain*. The double-chain consists of a convex and a concave chain, respectively. The reader may imagine the convex and concave chain as flipped semi-circles. All points are placed on the two chains. Cibulka et al. showed if at least one fifth of the points (regardless of their color) is contained on each of the chains, there is a non-crossing, alternating Hamiltonian path. On the other hand, if one of the chains contains at most  $\approx 1/29$  of all the points, there is no such path [3,4].

In convex position the proof techniques introduced the notion of *separated matchings*, that is, geometrically non-crossing matchings where all edges can be crossed by a single line. Separated matchings form a building element to alternating paths as every separated matching can be easily completed to a non-crossing, alternating path on a convex point set.

Recently new configurations were shown by Mészáros where any separated matching contains at most  $\frac{4}{3}n + O(\sqrt{n})$  points. A type of coloring was presented such that among these colorings in the optimal one any separated matching contains at most  $\frac{4}{3}n + O(\sqrt{n})$  points. These results may be found in [8,9].

Regarding the lower bound of the number of points in a maximal separated matching in a convex point set with balanced coloring there is a lot of room for improvement. The trivial lower bound is  $n$ . This bound is immediate. It is gained by a very simple idea of taking a halving line of the convex point set. Then at least half of the blue points will be on one side of the line and at least half of the red points will be on the other side. The mentioned points may be matched into a separated matching yielding the lower bound of  $n$  points.

An advantage of separated matchings is that we may consider point sets with small discrepancy. We say that the *discrepancy* is  $d$  if on any interval on the circle  $C$  the difference between the cardinality of color classes is at most  $d$ .

Small discrepancy colorings draw attention to the separated matching conjecture [7] that is formulated as follows. Let  $2k$  denote the number of alternations between the two colors in a  $2n$ -element point set on  $C$ . Then for any fixed  $k$  and large  $n$ , any configuration admits a separated matching that contains at least  $\frac{2k-1}{3k-2}2n + o(n)$  points.

So far no one was concerned with the discrepancy parameter since small discrepancy means many alternations among the two colors and that alone guarantees a long non-crossing, alternating path. However, when we consider separated matchings, it is reasonable to investigate this case. We believe it might shed light on the difficulties of the original Erdős problem.

Here we construct a new class of configurations allowing at most  $\frac{4}{3}n + O(\sqrt{n})$  points in any separated matching. This class significantly differs from the previous constructions for the reason that it contains arbitrarily many intervals consisting of short intervals that alternate in color. We will also present a coloring where the discrepancy parameter is constant and the number of points in the maximum separated matching is very close to the conjectured value. Furthermore, if we restrict the discrepancy, we obtain an interesting result. For discrepancies two and three we show that there are at least  $\frac{4}{3}n$  points in the maximum separated matching. This result suggests that the order of magnitude in the separated matching conjecture is feasible. A four-page extended abstract contains a brief summary of these results [10].

When the discrepancy is relatively small, the truth might be much closer to  $2n$  than  $\frac{4}{3}n$  at a balanced coloring. Although the case of small discrepancy looks very promising unfortunately already the analysis of discrepancy three is rather long by the current techniques. New ideas could yield further interesting results on small discrepancy colorings.

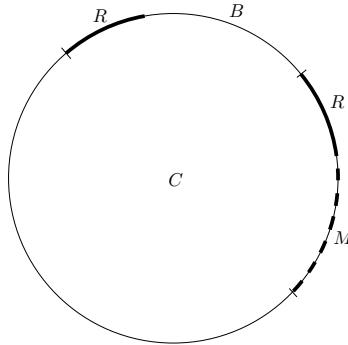
## 2 Coloring

First we introduce some necessary definitions to describe our coloring on  $C$ . Our  $2n$ -element convex point set with a balanced coloring we denote by  $P$ . Let an *arc* be an interval of points on  $C \cap P$ . We define the *size* of an arc to be the number of its elements. The points are ordered in an arc in the clockwise direction. A maximal set of consecutive points on  $C$  of the same color we call a *run*. The *length* of a run is the number of its elements.

The previous configurations contained long runs colored red or blue and at most two arcs consisting of alternating short runs of the two colors. We will present a coloring with arbitrary many arcs of alternating short runs. The idea originates from the Kynčl-Pach-Tóth construction which is described as follows. There are  $2k$  runs on  $C$ . The odd runs are colored red and the even ones blue. Suppose  $n$  is divisible by  $3k - 2$  (otherwise the construction is slightly different). Let the lengths of the runs be [7]:  $k\frac{n}{3k-2}, (2k-1)\frac{n}{3k-2}, k\frac{n}{3k-2}, \frac{n}{3k-2}, \dots, \frac{n}{3k-2}$ , and set  $k = \lfloor \sqrt{n} \rfloor$ . We cut the above construction into two pieces and we perform a small modification in the coloring. Then we repeat the two pieces we get in an arbitrary order and an equal number of times along the circle. The pieces themselves will not have a balanced coloring but the union of the two types of pieces will.

We describe two special arcs called *blocks*. They will be the building elements of our configuration. The *bluish* block will consist of a red run of length  $s$  and a blue run of length  $2s$ . We denote the bluish block by  $(s, 2s)$  block. The *reddish* block will consist of a red run of length  $s$  followed by a *mixed* arc  $M$ . The mixed arc  $M$  consists of  $2s$  points alternating in color. A blue and a consecutive red point in  $M$  will be referred to as a *period*. Note that the reddish block contains  $2s$  red and  $s$  blue points. We denote the reddish block by  $(s, s(1, 1))$  block.

If needed we introduce notations  $R$  and  $B$  for red runs of size  $s$ , and for blue runs of size  $2s$ , respectively. We call  $R$ ,  $B$  and  $M$  *subblocks* as they are the main building elements of blocks, see Figure 1.



**Fig. 1.** The blocks

The construction is a class of coloring  $\mathcal{C}(s, t)$ : Take  $t$  many  $(s, 2s)$  blocks and  $t$  many  $(s, s(1, 1))$  blocks and place them in an arbitrary order along  $C$ . In other words, the same number of bluish and reddish blocks are put along the circle and ordered arbitrarily. Observe that  $\mathcal{C}(s, t)$  is a balanced coloring.

### 3 Results

In this section we describe our results. First we give a general theorem about our class of coloring  $\mathcal{C}(s, t)$ . Later we take such a coloring from the previously given class where the discrepancy parameter is a large constant. Then we will estimate the number of points in the maximum separated matching in that coloring. In the end of this section we give another result on small discrepancy colorings. It will be a lower bound on the number of points in the maximum separated matching when the discrepancy is at most three.

Before we proceed to our claims we need to introduce another definition. Let the *size* of a separated matching be the number of points in the matching. Thus, it is twice the number of edges in the matching.

Now we are ready to make our statements.

**Theorem 1.** *Let  $C_0$  be any coloring from  $\mathcal{C}(s, t)$ . Then the size of any separated matching in  $C_0$  is at most  $\frac{4}{3}n + O(s + t)$ .*

In the theorem above we have  $O(s + t)$  as the remainder term. Parameters  $s$  and  $t$  can be set in a way that  $s, t = O(\sqrt{n})$ . Hence, we can achieve that the order of magnitude of  $O(s + t)$  is negligible.

**Observation 1** Let  $C_1$  be that coloring from  $\mathcal{C}(1000, t)$  where the reddish and bluish blocks alternate. Then the size of any separated matching in  $C_1$  is at most  $1.34n$ .

Observation 1 is a special case of this general theorem described above. We choose a setting where  $s$  is a large constant and  $t$  is  $\epsilon \cdot n$ . So  $O(s + t)$  is very small. The reason for choosing such a setting is that in  $C_1$  the discrepancy of the coloring is constant (2000). At the same time the size of the optimal matching is very close to the conjectured value.

Now we will consider colorings with small discrepancy. When the discrepancy parameter is bounded by three we can prove the following statement.

**Theorem 2.** For any two-coloring with discrepancy at most three of the  $2n$  points there is a separated matching of size at least  $\frac{4n}{3}$ .

In the case of the subsequent small integers at the discrepancy parameter, the  $\frac{4n}{3}$  lower bound is also feasible but the used techniques make the analysis rather long.

## 4 Proof of the Upper Bound

First we give the proofs of the upper bounds stated in our theorems. Here we will prove Theorem 1 and Observation 1.

We say that an arc  $A$  *faces* arcs  $A_1, A_2, \dots, A_n$  if all vertices of  $A$  that participate in the separated matching  $S$  have their pair in  $S$  on one of the arcs  $A_1, A_2, \dots, A_n$ .

Now we proceed to the proofs.

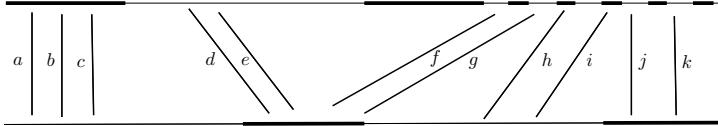
### Proof of Theorem 1:

Take any separated matching  $S$  in a coloring of  $C$  from  $\mathcal{C}(s, t)$ . Let line  $l$  be the axe, that is, a line that crosses all members of  $S$ . We think of  $l$  as a horizontal line deviding  $C$  into an upper and lower part. We can assume that both the upper and lower part of  $C$  consist of whole blocks by disregarding at most  $O(s)$  edges of  $S$ .

The edges of the matching can be ordered according to their intersection with  $l$ . We can partition the edges of  $S$  into classes in such a way that on each side (upper, lower) the endpoints belong to one subblock, see Figure 2. The previous partition determines  $O(t)$  pairs of arcs facing each other on  $C$ . Furthermore, if an arc  $A$  determined by the partition belongs to a mixed subblock, then we achieve that  $A$  contains complete periods. This can be done by removing at most  $O(t)$  many edges.

Let  $S_0$  be the remainder of  $S$ . We call it *the normalized matching*. Without loss of generality we may assume that we work with an arbitrary normalized matching.

Take a normalized separated matching  $S_0$  on  $C_0$  (an arbitrary member of  $\mathcal{C}(s, t)$ ). We will show that the size of  $S_0$  is at most  $\frac{4}{3}n$ .



**Fig. 2.** The classes:  $\{a, b, c\}$ ,  $\{d, e\}$ ,  $\{f, g\}$ ,  $\{i\}$ ,  $\{j, k\}$ . The edge  $h$  we delete to keep periods complete.

In each block the ratio between the color classes is 2 : 1. Hence, in each block we may call points of the major and the minor color class *major* and *minor* points, respectively. The number of minor points on  $C_0$  is  $\frac{2n}{3}$ . Our proof will be an assignment: to each edge of  $S_0$  we injectively assign a minor point. When we assign the point  $p$  to edge  $e$ , we say *mark*  $p$  for  $e$ .

If one endpoint of an edge  $e$  is minor point and the other endpoint is major point, then mark the minor point for  $e$ . If both endpoints of  $e$  are minor points, then mark the blue endpoint for  $e$ . Note, in this case the blue endpoint is in a mixed subblock  $M$ . Non-marked minor points give the set of *free* vertices. Observe, the set of free vertices is changing during the procedure of marking new points.

If both endpoints of  $e$  are major points, we distinguish two cases. If the major red point is in a mixed subblock, then we mark for  $e$  the free blue point that is in the same period with the red endpoint of  $e$ . We can freely do so as we achieved it in the normalization procedure that periods are complete. Otherwise, we call  $e$  a *bad* edge. Note, in this case one endpoint of  $e$  is in a blue subblock  $B$  and the other endpoint of  $e$  is in a red subblock  $R$  which is in an  $(s, s(1, 1))$  block. Bad edges are grouped according to blue subblocks. Take a blue subblock  $B$  and consider the bad edges incident to it. We can assume that  $B$  is on the upper side of  $l$ . Let  $R$  be the red subblock pair of  $B$ , that is,  $R$  and  $B$  form a block together. We distinguish two main cases.

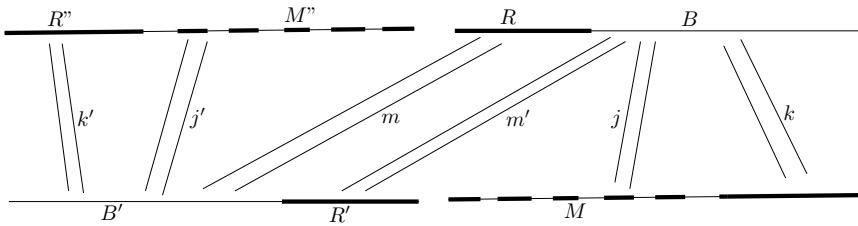
*Case 1: Subblock  $R$  contains only free vertices.* Let  $k$  denote the number of bad edges incident to  $B$ . If the red endpoints of the  $k$  bad edges are in the same red subblock, then  $k \leq s$  and we can injectively assign elements of  $R$  to each of the bad edges incident to  $B$ .

If the red endpoints come from different subblocks (of reddish blocks), then consider the mixed subblock  $M$  in the block of the rightmost bad edge's lower endpoint. Let  $j$  be the number of vertices matched in  $M$  to  $B$ . Consequently, we have at least  $s - j$  free vertices in  $M$  and  $s$  free vertices in  $R$ , altogether at least  $2s - j$  free vertices. As  $k + j \leq 2s$ , we get  $k \leq 2s - j$  as desired. Hence, we can mark free vertices injectively for each bad edge in  $B$ .

*Case 2: There is a non-free vertex in  $R$ .* Let the rightmost non-free vertex in  $R$  be incident to edge  $e$ . We can assume there is such an edge because the vertices of  $R$  will get marked only during the assignment of free vertices to bad edges incident to  $B$ . Let  $B'$  be the blue subblock of the lower endpoint of  $e$  and  $R'$  the red subblock pair of  $B'$ . Let  $k$  and  $M$  be defined as previously.

Suppose the red endpoints of the  $k$  bad edges come from different subblocks. Let  $M'$  be the subsequent mixed subblock to  $M$  on  $C$ . Let  $j$  be the number of vertices matched in  $M$  and  $M'$  to  $B$ . If there are vertices matched in  $M'$ , they have their pair in  $B$  and/or in  $R$ . The possible edges between  $M'$  and  $R$  leave their red endpoint free, so the number of free vertices is at least  $2s - j$  in  $M$ ,  $M'$  and in  $R$  altogether. As  $k + j \leq 2s$ , we are done.

Therefore we can assume that bad edges incident to  $B$  come from only one red subblock, see Figure 3. Let  $j$  be the number of vertices matched in  $M$  to  $B$ . Consider the bad edges incident to  $B'$ . Let their amount be denoted by  $k'$ . Let  $R''$  be the subblock of the rightmost bad edge's red endpoint and let  $M''$  be the subblock pair of  $R''$ . Let  $j'$  be the number of edges between  $M''$  and  $B'$ . Let  $m'$  be the number of edges incident to  $R'$ . Let  $m$  be the number of edges incident to  $R$  which have their blue endpoint in  $B'$ .



**Fig. 3.** The bad edges incident to  $B$  come from one red subblock

Now we injectively assign free vertices to all bad edges incident to  $B$  and  $B'$  at the same time. We will distinguish five subcases as follows.

i)  $k' = 0$  and  $m' = 0$ . The number of free vertices is  $s$  in  $R'$ . As  $k \leq s$ , we use the free vertices in  $R'$  to mark for the bad edges incident to  $B$  and  $B'$  (no bad edges in  $B'$ ).

ii)  $k' = 0$  and  $m' > 0$ . The number of free vertices is:  $s - m'$  in  $R'$ ;  $s - j$  in  $M$ . Together it is  $2s - j - m'$  free vertices. As  $k + j + m' \leq 2s$ , there are sufficiently many free vertices to be marked for all bad edges incident to  $B$  and  $B'$  (no bad edges in  $B'$ ).

iii)  $k' > 0$ ,  $m' = 0$  and there are no edges connecting  $R$  and  $M$ . The number of free vertices is:  $s - j'$  in  $M''$ ;  $s - m$  in  $R$ ;  $s$  in  $R'$ ;  $s - j$  in  $M$ . Altogether it makes  $4s - j - j' - m$  free vertices. As  $k' + j' + m \leq 2s$  and  $k + j \leq 2s$ , we get that  $k + k' \leq 4s - j - j' - m$  as desired.

iv)  $k' > 0$ ,  $m' = 0$  and there are  $l$  edges connecting  $R$  and  $M$ . The number of free vertices is:  $s - j'$  in  $M''$ ;  $s - m$  in  $R$ ;  $s$  in  $R'$ ;  $s - j - l$  in  $M$ . That gives  $4s - j - j' - l - m$  free vertices altogether. Also we know that  $k' + j' + m \leq 2s$  and  $k + j + l \leq 2s$ , and hence  $k + k' \leq 4s - j - j' - l - m$  as desired.

v)  $k' > 0$  and  $m' > 0$ . The number of free vertices is:  $s - j'$  in  $M''$ ;  $s - m$  in  $R$ ;  $s - m'$  in  $R'$ ;  $s - j$  in  $M$ . That is, we have  $4s - j - j' - m - m'$  free vertices altogether. As  $k' + j' + m \leq 2s$  and  $k + j + m' \leq 2s$  hence  $k + k' \leq 4s - j - j' - m - m'$ . So we can injectively assign free vertices to all bad edges incident to  $B$  and  $B'$ .

It remains to show that the assignment was injective. The blue subblocks are processed alone or in pairs. In the former case for bad edges of each blue subblock  $B$  we marked free vertices in the red subblock pair  $R$  of  $B$  or in a mixed subblock  $M$  (and  $M'$ ) facing  $B$  (and  $R$ ). Therefore, the assigned sets of free vertices are pairwise disjoint for different blue subblocks. In the latter case we considered two blue subblocks  $B$  and  $B'$  simultaneously and showed there are sufficiently many free vertices to mark in their red subblock pairs  $R$  and  $R'$  and in the mixed subblocks  $M$  and  $M'$  facing the block of  $B$  and  $B'$ , respectively. Hence, the assigned sets of free vertices are pairwise disjoint for the subblock pairs. We only need to argue that the pairs give a matching and there is no blue subblock  $B''$  that would make a pair with  $B$  or  $B'$ . Observe, if  $B''$  was in pair with  $B$  or  $B'$ , then no edges would be incident to  $B''$ .

We conclude that for each blue subblock the set of free vertices was well defined and disjoint from the set of free vertices for any other blue subblock. This completes the proof of Theorem 1.  $\square$

**Proof of Observation 1:** This is a special case of Theorem 1. The constants in front of  $s$  and  $t$  are small as a consequence of the number of disregarded edges in the normalization procedure. The claim of the theorem is immediate.  $\square$

## 5 Small Discrepancy

Now we will present our results on small discrepancy colorings.

Trivially, if we have a coloring with discrepancy one, then all points participate in the maximal separated matching. Moreover, regardless of it where we start we get all points covered by the separated matching if we make the matching in the most efficient way. This is obvious but at other colorings it is not the case by far.

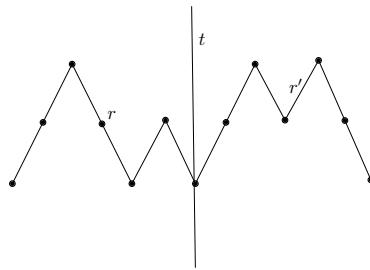
Instead of considering colorings with discrepancy two we will take colorings with runs of length at most two. Note that actually we consider a larger set of cases than colorings with discrepancy two. We will prove a lower bound on the size of separated matchings on this larger set. Consequently, our statement below is stronger than if we made the claim only for discrepancy two.

**Theorem 3.** *For any two-coloring with runs of length at most two of the  $2n$  points there is a separated matching of size at least  $\frac{4n}{3}$ .*

*Proof.* The colored point set can be viewed as follows: for each red point take a unit *up* line segment and for each blue point a unit *down* line segment. (When the discrepancy is one, then these up and down segments alternate.)

When the length of any run is at most two, we will not choose a good axe that divides our point set. We can be given any axe that halves the number of runs and we will construct a separated matching of the desired size.

There are two types of runs regarding their length: runs of length 1 and runs of length 2. Each run contains at most two up and at most two down segments, see Figure 4. Let us take a drawing for any coloring with runs of length at most



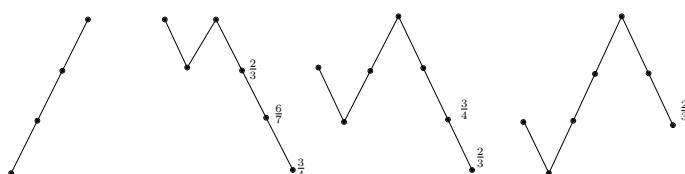
**Fig. 4.** When the length of any run is at most two, then at most  $\frac{1}{3}$  of the points will not participate in the constructed separated matching.

two and halve the number of runs by taking an axe  $t$ . Then we pair up all the runs. The run  $r$  will have pair run  $r'$  if  $r$  and  $r'$  are on different sides of  $t$  but for the same distance to  $t$  regarding the number of runs. We make the separated matching  $S$  so that each run will face only its pair in the matching. Therefore, all runs of length 1 will be fully covered in  $S$ . Now consider the runs of length 2. If a run  $r$  of length 2 faces a run  $r'$  of length 1, then  $\frac{2}{3}$  of all the vertices of  $r$  and  $r'$  will be in  $S$ . Otherwise, the run  $r$  is also fully covered in  $S$ . Hence, it follows that for colorings with runs of length at most two, there exists a separated matching of size at least  $\frac{4n}{3}$ .  $\square$

Note that the coloring did not have to be balanced in the theorem above. At balanced colorings an improvement is feasible. New ideas should be combined with the current methods to gain better results.

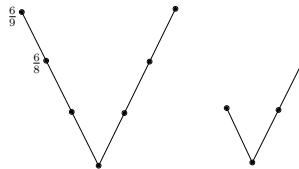
### 5.1 Colorings with Discrepancy Three

The case of discrepancy three uses similar ideas but it includes a more sophisticated pairing of the runs. We do not give the full case analysis here as it is quite extensive. We only give the pictures of the cases and a sketch of the proof: the left and right hand side of the picture correspond to intervals of the colored point set that will be faced to each other. We always face one specific side to the other side where we mark the ratio of points that can be matched on the

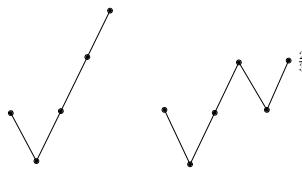


**Fig. 5.** The arc of three red points on the left faces the three arcs on the right, respectively. The number at a vertex on the right side shows the ratio of matched points in case the corresponding point is an endpoint of its run.

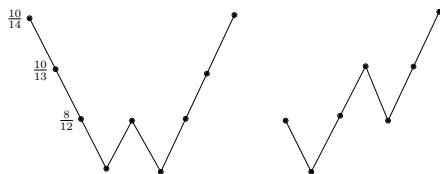
considered intervals. Some of the endpoints can vary on the sides and in that case we put the corresponding ratios to the alternate endpoints. On Figure 5 we merge more cases. There the left side contains a single interval while the right side contains three intervals. We face the left side to the intervals on the right, respectively. The three intervals on the right correspond to the three main cases. All forthcoming figures will be subcases of these.



**Fig. 6.** The right arc faces the left arc and there are two alternate endpoints on the left



**Fig. 7.** The left arc faces the right arc

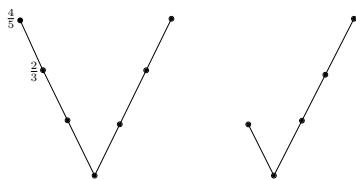


**Fig. 8.** The right arc faces the left arc and there are three alternate endpoints on the left

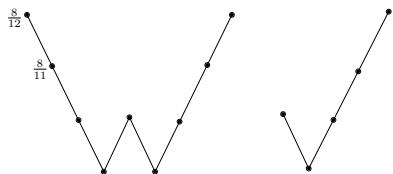
Observe, if we face two runs and one of the runs is of length two, then at least  $\frac{2}{3}$  of the points will be matched in these runs. Also it is easy to argue if the runs to be faced are of the same length. Therefore, we may suppose one of the runs to be faced is of length three and the other is of length one. Now we proceed on the circle and check what the lengths of the succeeding runs are that we did not process yet in the pairing algorithm. This gives the shape of the

intervals depicted on the figures. Note that the number of points altogether on two intervals that face each other in the case analysis is at most 14.

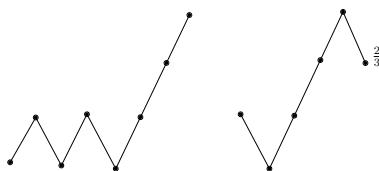
The first of the three main cases is complete on Figure 5. In each subcase indicated by the alternate endpoints a simple counting shows there are at least  $\frac{2}{3}$  of the points matched. The subcases of the second main case are given on Figure 6, Figure 7 and Figure 8. The rest of the figures belong to the third main case, see Figure 9, Figure 10, Figure 11 and Figure 12. The proof is a calculation in each of the cases. Checking that the given pictures cover all possible cases concludes Theorem 2.



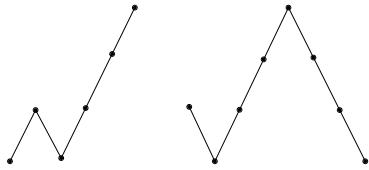
**Fig. 9.** The right arc faces the left arc



**Fig. 10.** The right arc faces the left arc and there are two alternate endpoints on the left



**Fig. 11.** The left arc faces the right arc



**Fig. 12.** The left arc faces the right arc

## 6 Closing Thoughts

Separated matchings is a new line of research in the area of long non-crossing, alternating paths. Recently more results were gained on the size of the maximal separated matching in a convex point set with balanced coloring. These are classes of configurations and types of colorings showing the  $\frac{4n}{3} + O(\sqrt{n})$  upper bound. All classes significantly differ from the previously known constructions and from each other, too. The mentioned results may be found in [5,8,9].

Small discrepancy colorings have an importance in trying to achieve a better lower bound for the problem of long non-crossing, alternating paths. Although at separated matchings the truth might be much closer to  $2n$  when the discrepancy parameter is small and the coloring is balanced, by our current methods it is not easy to reveal it. For discrepancies at most three a pairing algorithm of intervals yields that there are at least  $\frac{4n}{3}$  points in the maximum separated matching. It would be feasible to improve this result by new ideas. Also the case of the subsequent relatively small discrepancies seems promising. Just by our current methods it gets rather extensive.

Small discrepancy colorings also suggest the separated matching conjecture stated in the Introduction. A modification of the Kynčl-Pach-Tóth construction gives a constant in the remainder term. Roughly the original construction is the following:  $\frac{n}{3}$  blue,  $\frac{2n}{3}$  red,  $\frac{n}{3}$  blue points on  $C$  and an arc containing the rest of the points. The remaining  $\frac{n}{3}$  red and  $\frac{n}{3}$  blue points alternate in runs of length  $c\sqrt{n}$  where  $c$  is a positive constant. If we exchange the arc containing the short alternating runs of length  $c\sqrt{n}$  with an arc of points alternating in color, then the size of the largest separated matching is  $\frac{4}{3}n + 1$ . Note that the length of the longest non-crossing, alternating path is  $2n$  after performing this modification.

*Conjecture 4.* [5] Every balanced coloring of  $2n$  points on  $C$  admits a separated matching of size  $\frac{4}{3}n$ .

It would be an interesting result to settle this conjecture in the affirmative. That would also prove the conjecture related to the upper bound for non-crossing, alternating paths.

**Acknowledgement.** I would like to thank my advisor Péter Hajnal for his fruitful suggestions and for all the help during this research.

## References

1. Abellanas, M., García, J., Hernández, G., Noy, M., Ramos, P.: Bipartite embeddings of trees in the plane. *Discrete Appl. Math.* 93, 141–148 (1999)
2. Abellanas, M., García, A., Hurtado, H., Tejel, J.: Caminos Alternantes. In: Proc. X Encuentros de Geometría Computacional, Sevilla, pp. 7–12 (June 2003) (in Spanish)
3. Cibulka, J., Kynčl, J., Mészáros, V., Stolař, R., Valtr, P.: Hamiltonian Alternating Paths on Bicolored Double-Chains. In: Tollis, I.G., Patrignani, M. (eds.) GD 2008. LNCS, vol. 5417, pp. 181–192. Springer, Heidelberg (2009)
4. Cibulka, J., Kynčl, J., Mészáros, V., Stolař, R., Valtr, P.: Universal Sets for Straight-Line Embeddings of Bicolored-Graphs. In: Pach, J. (ed.) Thirty Essays on Geometric Graph Theory. Springer (to appear)
5. Hajnal, P., Mészáros, V.: Note on noncrossing alternating path in colored convex sets. Accepted to Discrete Mathematics and Theoretical Computer Science
6. Kaneko, A., Kano, M.: Discrete geometry on red and blue points in the plane — a survey. In: Aronov, B., et al. (eds.) Discrete and Computational Geometry, Goodman-Pollack Festschrift. Algorithms Comb., vol. 25, pp. 551–570. Springer (2003)
7. Kynčl, J., Pach, J., Tóth, G.: Long alternating paths in bicolored point sets. *Discrete Mathematics* 308(19), 4315–4321 (2008)
8. Mészáros, V.: An upper bound on the size of separated matchings. *Electronic Notes in Discrete Mathematics*, 633–638 (2011)
9. Mészáros, V.: Separated matchings in colored convex sets (manuscript)
10. Mészáros, V.: Separated matchings in convex point sets with small discrepancy, CRM Documents 8, Centre de Recerca Matemática, Bellaterra, Barcelona, pp. 217–220 (2011)

# A Note on the Number of Empty Triangles

Alfredo García\*

Departamento de Métodos Estadísticos and IUMA, Universidad de Zaragoza,  
Zaragoza, 50009, Spain

**Abstract.** Let  $P$  be a set of  $n$  points on the plane, in general position,  $H$  of them placed on the boundary of the convex hull of  $P$ . In this note we prove that there is a well defined family of empty triangles, the family of empty triangles not generated by an empty convex pentagon, containing exactly  $n^2 - 5n + H + 4$  empty triangles. This result immediately implies a slight improvement on the lower bound on the number of empty triangles that every set of  $n$  points in the plane must determine.

**Keywords:** Empty triangle, empty convex pentagon, convex hull, points in general position.

## 1 Introduction

In a classic paper, Barany and Füredi [3] proved that given a set  $P$  of  $n$  points in the plane in general position, the number of empty triangles determined by  $P$  is at least  $n^2 - O(n \lg n)$ . In the same paper, they proved that the expected number of empty triangles for a set of random points is  $2n^2 + O(n \lg n)$ , this last result implying that there are sets of points having  $\leq 2n^2 + O(n \lg n)$  empty triangles. In fact, Barany and Valtr [2] constructed sets of  $n$  points in general position containing  $\leq 1.6196n^2 + o(n^2)$  empty triangles.

The lower bound on the number of empty triangles was improved by Dehnhardt [4] in his doctoral dissertation to

$$n^2 - 5n + 10$$

(if  $n \geq 12$ ). However, this result in the thesis of Dehnhardt (in German) has been practically unnoticed in the community of Computational Geometry, and in fact it is not cited even in surveys on this subject like [1], [6] or [7].

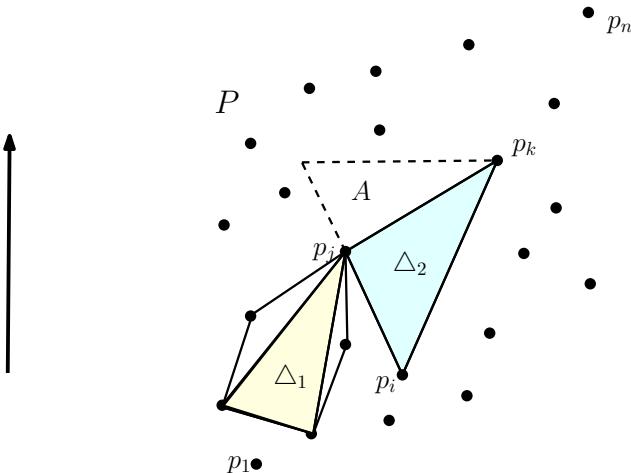
In this note, we slightly improve the lower bounds given in [3] and [4] proving that any set of points  $P$  contains at least

$$n^2 - 5n + H + 4 + 3 \cdot \left\lfloor \frac{n-4}{8} \right\rfloor$$

empty triangles, where  $H$  is the number of points on the boundary of the convex hull of  $P$ .

---

\* Partially supported by research grants MTM2009-07242 and E58-DGA.



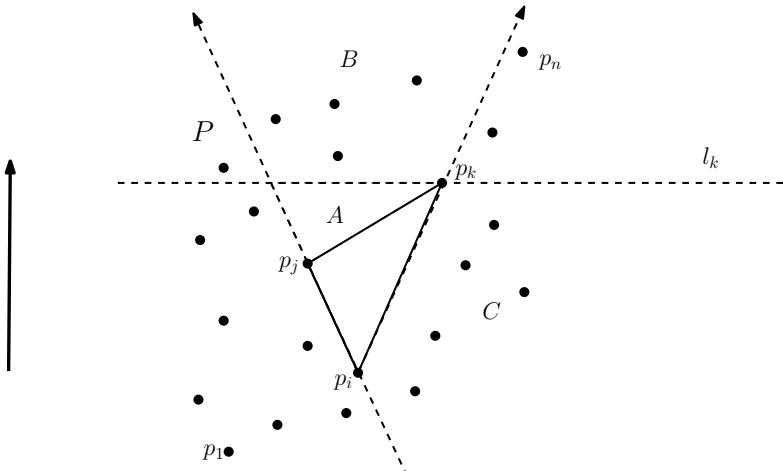
**Fig. 1.** A triangle generated by an empty convex pentagon ( $\Delta_1$ ), and a triangle not generable by an empty convex pentagon ( $\Delta_2$ ).

Probably more interesting than this new lower bound is the result that there is a well defined family of empty triangles, the family of "triangles not generated by an empty convex pentagon", containing an invariant number of empty triangles, exactly  $n^2 - 5n + 4 + H$  triangles.

In this paper we are going to use the definitions and notation used in Pinchasi et al. [7]. Let  $P = \{p_1, \dots, p_n\}$  be a set of  $n$  points in the plane in general position, sorted in increasing order of the ordinate  $y$ . That is,  $p_1$  is at the bottom and  $p_n$  is at the top of the point set. Let  $K$  be an empty convex pentagon spanned by  $P$  and let  $p$  be the top vertex of  $K$ . We name a triangle "generated by  $K$ " if it is spanned by  $p$  and the two vertices of  $K$  that are not adjacent (on the boundary of  $K$ ) to  $p$ . Consequently, the family of "triangles not generated by an empty convex pentagon" is the set of triangles that cannot be obtained in that way. See Fig. 1. Further, we denote by  $X_k(P)$  the number of empty convex  $k$ -gons determined by  $P$ .

## 2 Counting Empty Triangles

Consider an empty triangle  $\Delta = p_i p_j p_k$  ( $i < j < k$ ) of  $P$  and denote the horizontal line through  $p_k$  with  $l_k$ . The (bounded) region of the plane below  $l_k$  and between  $\Delta$  and the ray  $p_i p_j$  we call region  $A$  of  $\Delta$ . See Fig. 2. In the same way we name as region  $B$  of  $\Delta$  the (unbounded) region of the plane above  $l_k$  and between the two rays  $p_i p_j$  and  $p_i p_k$ . Furthermore, region  $C$  of  $\Delta$  is defined to be the (unbounded) region of the plane below  $l_k$  and between  $\Delta$  and the ray  $p_j p_i$ . We say that a region ( $A$ ,  $B$ , or  $C$ ) of a triangle is empty if it does not contain points of  $P$ .



**Fig. 2.** Regions  $A$ ,  $B$  and  $C$  of an empty triangle

Let us denote by  $F_e$  the family of empty triangles such that both region  $A$  and region  $C$  contain points of  $P$ . Notice that given an empty convex pentagon  $K$ , the triangle generated by  $K$  contains points in both regions  $A$  and  $C$ , so, it belongs to the family  $F_e$ . Reciprocally, a triangle  $\Delta = p_i p_j p_k$  of  $F_e$  is generated by the empty convex pentagon  $K$  formed by these three vertices of  $\Delta$ , plus the closest point to  $\Delta$  in region  $A$  and the closest point to  $\Delta$  in region  $C$ . That is to say,  $F_e$  coincides with the family of empty triangles generated by an empty convex pentagon. The complementary family, the family of empty triangles such that region  $A$  or region  $C$  (or both) is empty, let us denote it by  $F_o$ , is the one with an invariant number of triangles.

Now, the main result.

**Theorem 1.** *Let  $P$  be a set of  $n$  points in general position,  $H$  of them placed on the boundary of the convex hull of  $P$ . Then, the number of empty triangles of  $P$  such that region  $A$  or region  $C$  (or both) is empty equals*

$$n^2 - 5n + H + 4$$

*Proof.* First of all let us denote by  $T$  the triangulation obtained by the greedy method, (exploring the points in its natural order  $p_1, p_2, \dots, p_n$ ). Initially  $T$  contains only the triangle  $p_1, p_2, p_3$ , then we add to  $T$  the triangles (placed outside of the convex hull of the previous points) with top point  $p_4$ , next the triangles with top point  $p_5$  and so on. In this process, when we are adding triangles with top vertex  $p_k$ , a triangle  $\Delta = p_i p_j p_k$  is added to  $T$ , if and only if,  $p_k$  lies in one half-plane of the line  $p_i p_j$  and all the other points  $p_h$  ( $h < k$ ) lie in the other half-plane. Therefore,  $\Delta$  is added to  $T$  if and only if both region  $A$  of  $\Delta$  and region  $C$  of  $\Delta$  are empty. In other words, the  $2n - H - 2$  triangles of the greedy triangulation  $T$  are precisely the empty triangles of  $P$  with empty regions  $A$  and  $C$ .

Now we are going to define three mutually disjoint families ( $F_1$ ,  $F_2$ , and  $F_3$ ) of empty triangles of  $P$ . First, for each point  $p_i$  ( $1 \leq i \leq n - 2$ ) of  $P$  consider the points of  $P$  placed above  $p_i$ . Order these  $n - i$  points clockwise around  $p_i$  (starting with the leftmost) and denote them by  $p_{i_1}, \dots, p_{i_{n-i}}$  in that order. Consider the rays emanating from  $p_i$  and passing through  $p_{i_1}, \dots, p_{i_{n-i}}$ . Clearly, since the space between consecutive rays does not contain points of  $P$ , both region  $A$  and region  $B$  of each empty triangle  $p_i p_{i_j} p_{i_{j+1}}$  ( $1 \leq j \leq n - i - 1$ ) are empty. Thus, the first family of empty triangles  $F_1 = \{\text{empty triangle } | A = \emptyset \wedge B = \emptyset\}$ . Notice that  $|F_1| = (n - 1)(n - 2)/2$ .

Suppose now that, from the above configuration of rays emanating from  $p_i$ , we remove  $p_n$ , the top point of  $P$ . If the ray  $p_i p_n$  is neither the first one nor the last one of the above configuration of rays, that is  $p_n = p_{i_j}$ ,  $j \neq 1, j \neq n - i$ , then we can form a new empty triangle  $p_i p_{i_{j-1}} p_{i_{j+1}}$ . We can repeat this process removing the following top point  $p_{n-1}$ , then point  $p_{n-2}$ , and so on, until removing point  $p_{i+2}$ . After removing each point, if the corresponding ray is neither the first nor the last ray, we obtain a new triangle. Let us call  $F_2$  the family of triangles obtained by this method (this family can be empty). By construction, region  $A$  of each triangle of  $F_2$  is empty, and region  $B$  is nonempty. Reciprocally, given an empty triangle  $p_i p_j p_k$  without points in region  $A$  and with points in region  $B$ , this triangle belongs to the family  $F_2$ , because it is obtained by the previous process when we explore the rays emanating from  $p_i$  and remove the bottom point of region  $B$ . Therefore,  $F_2 = \{\text{Empty Triangles } | A = \emptyset \wedge B \neq \emptyset\}$ . As the process of removing top points is done  $n - i - 1$  times for each point  $p_i$ , we have

$$|F_2| + |\text{ExtremalRays}| = (n - 1)(n - 2)/2$$

where  $|\text{ExtremalRays}|$  is the number of times that an extremal ray (first or last ray) is found. Notice that  $|\text{ExtremalRays}| \geq n - 2$ , because when removing  $p_{i+2}$ , for each point  $p_i$ , an extremal ray ( $p_i p_{i+2}$ ) is found. Further observe, that the triangles of  $T$  are fully contained in the set  $F_1 \cup F_2$ .

Now, consider the case when we have to remove vertex  $p_k$ ,  $k > i + 1$  and ray  $r_k = p_i p_k$  is an extremal ray. Without loss of generality suppose that ray  $r_k$  is the last one. Keep in mind that at this moment of the process, we are only considering points in the range  $p_i, p_{i+1}, \dots, p_k$ , and we are supposing that ray  $r_k$  is the last one in this subset of rays. Let  $p_i p_j$  be the ray placed just before ray  $r_k$ , and consider the triangle  $\Delta = p_i p_j p_k$ . If region  $C$  of this triangle is nonempty, then rotating clockwise the ray  $p_i p_k$  we first reach a point  $p_h$  ( $h < i$ ) in that region, (the ray  $p_i p_h$  is pointing downwards). In this case, we can form a new empty triangle  $p_h p_i p_k$ . Doing this construction, when possible, we obtain a new family of triangles  $F_3$  (possibly empty). By construction, region  $C$  of these new triangles is empty. However, region  $A$  is nonempty because point  $p_j$  is in that region. Reciprocally, given an empty triangle  $p_h p_i p_k$  such that their region  $A$  is nonempty, and their region  $C$  is empty, this triangle belongs to the family  $F_3$ , because it is obtained by the previous process when we explore the extremal ray  $p_i p_k$ . It follows that  $F_3 = \{\text{Empty Triangles } | A \neq \emptyset \wedge C = \emptyset\}$ . Again, by

construction, we have

$$|ExtremalRays| = |F_3| + |CasesWith(C = \emptyset)|$$

where  $|CasesWith(C = \emptyset)|$  is the number of extremal rays  $r_k$  such that region  $C$  of the corresponding triangle  $p_i p_j p_k$  is empty.

Finally, observe that in the cases of an extremal ray from  $CasesWith(C = \emptyset)$  both region  $A$  and region  $C$  of the corresponding triangle are empty. Recall that this are exactly the  $2n - H - 2$  triangles of the greedy triangulation  $T$ .

Summarizing, we have

$$|F_1| + |F_2| + |F_3| + 2n - H - 2 = (n - 1)(n - 2)$$

On the other hand  $F_o$  consists of the triangles with  $A = \emptyset$  (these of  $F_1 \cup F_2$ ) plus the triangles with  $C = \emptyset$  and  $A \neq \emptyset$  (these of  $F_3$ ), therefore

$$|F_o| = |F_1 \cup F_2 \cup F_3| = |F_1| + |F_2| + |F_3| = n^2 - 5n + H + 4$$

□

### 3 Some Implications

1.- Notice that, for a given set  $P$  of points, the families  $F_o = F_1 \cup F_2 \cup F_3$  and  $F_e$  of empty triangles depend on the chosen direction  $\mathbf{d}$  (upwards, in our case), in which the points of  $P$  are ordered. But the numbers  $|F_o|$  and  $|F_e|$  are independent of  $\mathbf{d}$ .

2.- Recall that, given a triangle  $\Delta = p_i p_j p_k$  of  $F_e$  we can associate to it the empty convex pentagon  $K$  formed by the three vertices of  $\Delta$ , plus the closest point to  $\Delta$  in region  $A$  and the closest point to  $\Delta$  in region  $C$ . It is easy to see that different triangles of  $F_e$  have associated different pentagons, hence the number of empty convex pentagons of  $P$  is at least  $|F_e|$ , or in other words

$$X_5(P) \geq X_3(P) - (n^2 - 5n + H + 4)$$

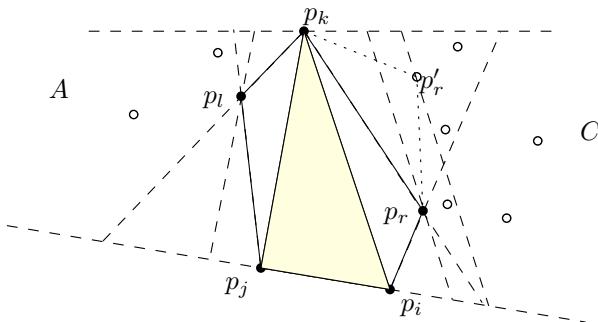
3.- Since  $X_3(P) = |F_o| + |F_e|$  we have

$$X_3(P) = n^2 - 5n + H + 4 + |F_e|$$

Notice that given an empty convex pentagon  $K$ , there is a unique empty triangle generated by  $K$ . However, different pentagons can generate the same triangle of  $F_e$ .

For points in convex position,  $|F_o| = n^2 - 5n + n + 4 = (n - 2)^2$ . Therefore, in this case there are exactly  $\binom{n}{3} - (n - 2)^2 = \binom{n-2}{3}$  triangles generated by empty convex pentagons.

In any case, since for  $n \geq 12$ ,  $P$  must contain at least 3 empty convex pentagons (see [4]), we can deduce the following bound for  $F_e$ .



**Fig. 3.** A triangle generated by two empty convex pentagons

**Theorem 2.** Let  $P$  be a set of  $n$  points in general position,  $H$  of them placed on the boundary of the convex hull of  $P$ . Then,

$$X_3(P) \geq n^2 - 5n + H + 4 + 3 \cdot \left\lfloor \frac{n-4}{8} \right\rfloor$$

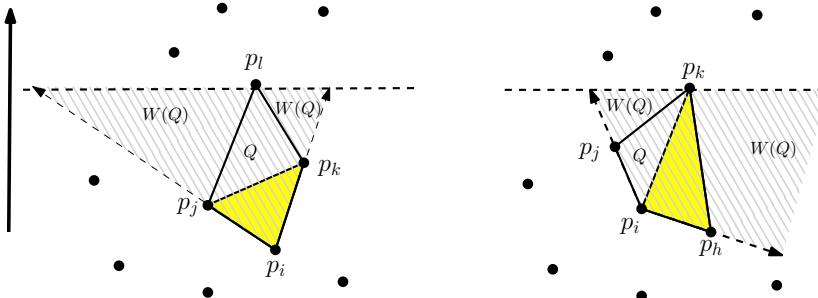
*Proof.* Take the 4 lowest points of  $P$ . Add the next 8 points of  $P$  to get the first set of 12 points. Take the topmost 4 points of this set and add the next 8 points of  $P$  to create a second set of 12 points. And so on. This results in  $\lfloor \frac{n-4}{8} \rfloor$  sets of 12 points, such that a pentagon completely contained in one set is not completely contained in any other. Furthermore, the top points of pentagons differ in different sets of 12 points.

Every set of 12 points contains at least 3 empty convex pentagons [4] (dissertation of Knut Dehnhardt). Consider a set of 12 points. If each of these 3 empty convex pentagons generates a different empty triangle, then we get 3 different empty triangles generated by empty convex pentagons in that set of 12 points. Otherwise, at least 2 empty convex pentagons generate the same empty triangle. We know that an empty triangle  $\triangle p_i p_j p_k$  of  $F_e$  is generated by the empty convex pentagon spanned by  $p_i, p_j, p_l, p_k, p_r$ , where  $p_l$  is the closest point to  $\triangle$  in region  $A$  and  $p_r$  is the closest point to  $\triangle$  in region  $C$ . See Fig. 3. If  $\triangle$  is generated by another empty convex pentagon, then there must be at least one additional point  $p'_l$  in region  $A$  such that triangle  $p'_l p_k p_j$  is empty or at least one additional point  $p'_r$  in region  $C$  such that triangle  $p'_r p_i p_k$  is empty. W.l.o.g. assume that there exists at least one additional point  $p'_r$  in  $C$  such that  $p'_r p_i p_k$  is empty and, among these additional points, let  $p'_r$  be the closest one to  $\triangle$ . Then, the points  $p_i, p_j, p_l, p_k, p'_r, p_r$  span an empty convex hexagon (as shown in Fig. 3). And this empty convex hexagon contains  $\binom{6-2}{3} = 4$  different empty triangles generated by empty convex pentagons.

Therefore, there exist at least 3 different empty triangles generated by empty convex pentagons in each of the  $\lfloor \frac{n-4}{8} \rfloor$  sets of 12 points. And thus  $|F_e| \geq 3 \cdot \lfloor \frac{n-4}{8} \rfloor$ , because the top points of the empty triangles differ for each of the sets of 12 points.  $\square$

4.- Given an empty convex pentagon  $K$  spanned by  $P$ , we name an empty convex quadrilateral  $p_i p_j p_k p_h$  generated by  $K$  if it is spanned by the vertices of  $K$  with the exception of one of the points adjacent (on the boundary of  $K$ ) to the top point of  $K$ . By construction, each empty convex pentagon generates two empty convex quadrilaterals. Let  $C_e$  be the family of quadrilaterals generated by empty convex pentagons.

Consider an empty convex quadrilateral  $Q = p_i p_j p_k p_h$ . Suppose that  $p_k$  is the top point of  $Q$  and  $p_i$  is the opposite vertex to  $p_k$  in  $Q$ . Let us denote by  $W(Q)$  the convex region bounded by the horizontal line passing through the top point  $p_k$  and the rays  $p_i p_h$  and  $p_i p_j$ . See Fig. 4(right). Observe that for quadrilaterals  $Q$  in  $C_e$ , region  $W(Q)$  is nonempty and, reciprocally, quadrilaterals with at least one point in  $W(Q)$  are generated by an empty convex pentagon. Therefore, if  $C_o$  is the family of quadrilaterals not generated by empty convex pentagons then  $C_o$  coincides with the family of quadrilaterals such that region  $W(Q)$  is empty. Let us prove that  $C_o$  contains an invariant number of quadrilaterals.



**Fig. 4.** Regions  $W(Q)$  of two empty convex quadrilaterals. (Left) A quadrilateral of  $C_2$ . (Right) A quadrilateral of  $C_3$ . The triangles of  $F_2$  and  $F_3$  are highlighted.

**Theorem 3.** *Let  $P$  be a set of  $n$  points in general position,  $H$  of them placed on the boundary of the convex hull of  $P$ . Then, the number of empty convex quadrilaterals such that region  $W(Q)$  is empty equals*

$$(n^2 - 7n + 6)/2 + H$$

*As a consequence,*

$$X_4(P) = |C_o| + |C_e| = (n^2 - 7n + 6)/2 + H + |C_e| \geq (n^2 - 7n + 6)/2 + H + 6 \cdot \left\lfloor \frac{n-4}{8} \right\rfloor$$

*Proof.* Observe that each triangle  $p_i p_j p_k$  belonging to the family  $F_2$  has been obtained after removing a top point  $p_l$ . Therefore, points  $p_i p_j p_l p_k$  form an empty convex quadrilateral  $Q$ , and its region  $W(Q)$  is empty. So, this quadrilateral belongs to  $C_o$ . Besides, the top point  $p_l$  of  $Q$  and the bottom point  $p_i$  are opposite vertices in  $Q$ . See Fig. 4(left). Let  $C_2$  be the family of quadrilaterals obtained in this way.

Similarly, from each triangle  $p_h p_i p_k$  of the family  $F_3$  we can form the empty convex quadrilateral  $p_i p_j p_k p_h$ . See Fig. 4(right). Let  $C_3$  be the family of quadrilaterals obtained in this way. Again, by construction, for a quadrilateral  $Q$  of  $C_3$  region  $W(Q)$  is empty. So, this quadrilateral belongs to  $C_o$ . But now,  $p_k$  (the top point of  $Q$ ) and  $p_h$  (the bottom point) are adjacent vertices in  $Q$ . Hence, the quadrilaterals of  $C_2$  are different from the quadrilaterals of  $C_3$ .

Now, let  $Q$  be a quadrilateral of  $C_o$ , with clockwise vertices  $p_i p_j p_k p_h$  and top point  $p_k$ . If the opposite vertex to  $p_k$  in  $Q$ , vertex  $p_i$ , is the bottom point of  $Q$ , then the triangle  $p_i p_j p_k$  belongs to  $F_2$  and, therefore,  $Q$  belongs to  $C_2$ . If  $p_i$  is not the bottom point of  $Q$ , then either the triangle  $p_h p_i p_k$  (if  $p_h$  is the bottom point of  $Q$ ) or the triangle  $p_i p_j p_k$  (if the bottom point of  $Q$  is  $p_j$ ) has to belong to  $F_3$  and, therefore,  $Q$  belongs to  $C_3$ . We conclude that  $C_o = C_2 \cup C_3$  and  $|C_o| = |F_2 \cup F_3|$ . Since  $F_2 \cup F_3$  contains  $(n-1)(n-2)/2 - (2n-H-2) = (n^2-7n+6)/2 + H$  triangles, we have proved the first statement of the theorem.

Observe that for points in convex position,  $|C_o| = (n-2)(n-3)/2$ . So, in this case,  $|C_e| = \binom{n}{4} - (n-2)(n-3)/2 = ((n+3)/4)\binom{n-2}{3}$ .

To complete the proof we only need to prove that

$$C_e \geq 6 \cdot \left\lfloor \frac{n-4}{8} \right\rfloor$$

and this result is proved using the same reasonings as in the proof of Theorem 2. Now, in each set of 12 points, there either are  $2 \cdot 3$  different empty convex quadrilaterals generated by empty convex pentagons, or at least one empty convex quadrilateral is generated by two different empty convex pentagons. In the latter case there has to exist an empty convex hexagon, this hexagon containing  $((6+3)/4)\binom{6-2}{3} = 9$  quadrilaterals of  $C_e$ . Therefore, there exist at least 6 different empty convex quadrilaterals of  $C_e$  in each of the  $\lfloor (n-4)/8 \rfloor$  sets of 12 points.

□

**Final Remark.** I first knew the results obtained by K. Dehnhardt in [4] after presenting a preliminary version of this paper in the XIV Spanish Meeting on Computational Geometry. I'm grateful to C. Huemer for informing me about the existence of these results and for translating some of the main results of that paper.

I'm specially grateful to Thomas Hackl that helped me to improve substantially the presentation of this paper, and sent me the bounds on  $|F_e|$  and  $|C_e|$  given in theorems 2 and 3.

## References

1. Aichholzer, O.: [Empty] [colored]  $k$ -gons - Recent results on some Erdős-Szekeres type problems. In: Proc. XIII Encuentros de Geometría Computacional, pp. 43–52. Prensas universitarias de Zaragoza, Spain (2009)

2. Bärany, I., Valtr, P.: Planar point sets with a small number of empty convex polygons. *Stud., Sci. Math. Hung.* 41(2), 243–269 (2004)
3. Bärany, I., Füredi, Z.: Empty simplices in Euclidean space. *Canadian Math. Bull.* 30, 436–445 (1987)
4. Dehnhardt, K.: Leere konvexe Vielecke in ebenen Punktmengen, Dissertation, TU Braunschweig (1987)
5. Harborth, H.: Konvexe Fünfecke in ebenen Punktmengen. *Elem. Math.* 33, 116–118 (1978)
6. Morris, W., Soltan, V.: The Erdős-Szekeres problem on points in convex position - a survey. *Bulletin (new series) of the American Mathematical Society* 37(4), 437–458 (2000)
7. Pinchasi, R., Radoicic, R., Sharir, M.: On empty convex polygons in a planar point set. *J. Comb. Theory, Ser. A* 113(3), 385–419 (2006)

# Meshes Preserving Minimum Feature Size

Greg Aloupis<sup>1,\*</sup>, Erik D. Demaine<sup>2</sup>, Martin L. Demaine<sup>2</sup>,  
Vida Dujmović<sup>3</sup>, and John Iacono<sup>4,\*\*</sup>

<sup>1</sup> Université Libre de Bruxelles

aloupis.greg@gmail.com

<sup>2</sup> Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

{edemaine, mdemaine}@mit.edu

<sup>3</sup> Carleton University, Ottawa, Ontario, Canada

vida@cs.mcgill.ca

<sup>4</sup> Polytechnic Institute of New York University, Brooklyn, New York, USA

jiacono@poly.edu

**Abstract.** The *minimum feature size* of a planar straight-line graph is the minimum distance between a vertex and a nonincident edge. When such a graph is partitioned into a mesh, the *degradation* is the ratio of original to final minimum feature size. For an  $n$ -vertex input, we give a triangulation (meshing) algorithm that limits degradation to only a constant factor, as long as Steiner points are allowed on the sides of triangles. If such Steiner points are not allowed, our algorithm realizes  $O(\lg n)$  degradation. This addresses a 14-year-old open problem by Bern, Dobkin, and Eppstein.

## 1 Introduction

Meshing is a field frequently studied in the context of computational geometry; see [BE95, She04] for surveys. In two dimensions, the typical forms of input are point sets, polygons, and most generally, planar straight-line graphs (PSLGs). The typical desired output is a decomposition into triangles or quadrangles, usually with Steiner points allowed (though usually aiming to minimize their number). A wide variety of quality measures dictate the desired decomposition. Often, decompositions are constructed so that there are no large angles, or instead no small angles, short edges, or short triangle heights. Most of these problems have been solved in the best sense possible. This paper highlights one problem that has not been fully solved.

*Problem statement.* Our goal is to mesh a polygon  $P$  into a triangulation  $G$  while avoiding the introduction of a small (Euclidean) distance between a vertex and a nonincident edge in  $G$ , compared to distances already existing in  $P$ . The minimum such distance

---

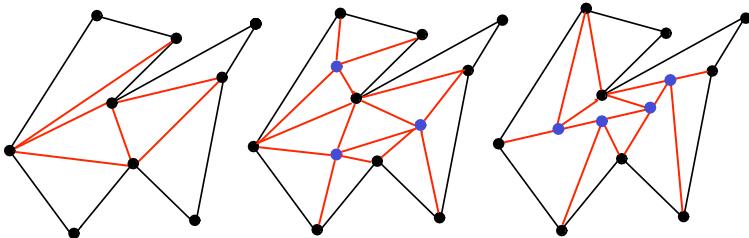
\* Chargé de Recherches du FNRS.

\*\* Research supported by NSF grants CCF-1018370, CCF-0430849 and by an Alfred P. Sloan fellowship. Research partially completed as a visiting professor at MADALGO (Center for Massive Data Algorithmics, a Center of the Danish National Research Foundation), Department of Computer Science, Aarhus University, IT Parken, Åbogade 34, DK-8200 Århus N, Denmark.

in  $G$  (or more generally in any PSLG) is called the *minimum feature size*, denoted by  $\text{mfs}(G)$ . See [Rup93, Dey07, HMP06, Eri03]. We call the ratio  $\frac{\text{mfs}(P)}{\text{mfs}(G)}$  the *degradation* of the decomposition of  $P$  into  $G$ .

Minimum feature size effectively describes the global resolution needed to visually distinguish elements in a mesh. For example, it measures the maximum (uniform) thickness that the edges in a mesh can be drawn without obscuring the individual components. Also,  $\text{mfs}$  measures the (maximum) error allowed in the placement of vertices while still guaranteeing preservation of the topology of the mesh. We were originally motivated to study minimum feature size as a way to obtain pseudopolynomial bounds on algorithms (specifically, geometric dissection) that start with a triangulation step; see Section 5 for details.

One important issue is the type of desired triangulation, which we show has a significant effect on the results that can be achieved. Refer to Figure 1. The most common decomposition of a polygon is the *classic triangulation*, which adds noncrossing chords between pairs of vertices of  $P$ , until the interior of  $P$  is partitioned into triangles. A *nonproper triangulation* allows the addition of Steiner points (extra vertices), and non-crossing edges between pairs of vertices (original or added), until each interior face has the geometric shape of a triangle. A *proper triangulation* has the additional property that any two edges that lie on the same interior face and are incident to a common vertex are not collinear; in other words, each interior face consists of only three vertices.



**Fig. 1.** Triangulation types: classic, proper, nonproper. Steiner points are blue

*Related results.* Steiner points are necessary to obtain a degradation smaller than linear. Mitchell constructed two illustrative examples, described in [BDE95]. The first example is a regular  $n$ -gon: all classic triangulations have a degradation of  $\Omega(n)$  (consider an ear). The second example is an  $r \times 1$  rectangle with two additional vertices approximately midway along one long edge, spaced at unit distance from each other: all classic triangulations have a degradation of  $\Omega(r)$ . Here  $r$  is the ratio of the polygon's diameter divided by its minimum feature size, often called the *spread*. These lower bounds extend to quadrangulations or any decomposition with constant-complexity faces; in the latter example, we simply add more vertices midway along the long edge.

When studying this problem, Bern, Dobkin, and Eppstein [BDE95] applied the notion of *internal feature size*  $\text{ifs}(P)$ , which is the minimum distance *inside*  $P$  between a vertex and a nonincident edge<sup>1</sup>. They proved that every polygon  $P$  (possibly with

<sup>1</sup> Note that “internal feature size” is called “minimum feature size” in [BDE95].

**Table 1.** Results on degradation, by triangulation type, when meshing a worst-case polygon or PSLG with  $n$  vertices and spread  $r$ 

Type of triangulation	Degradation of	
	minimum feature size	internal feature size
Classic	$\Omega(n + r)$ [BDE95]	$\Omega(n + r)$ [BDE95]
Proper	$\Omega\left(\frac{\log n}{\log \log n}\right)$ [ADD <sup>+</sup> 11], $\mathcal{O}(\log n)$ [§3]	$\Theta(1)$ [BDE95] [§4]
Nonproper	$\Theta(1)$ [§2]	$\Theta(1)$

holes) has a proper triangulation  $G$  in which every triangle has height  $\Omega(\text{ifs}(P))$ , and thus  $\text{ifs}(G) = \Omega(\text{ifs}(P))$ . In other words, they achieve  $\mathcal{O}(1)$  degradation for the interior of  $P$ . However, this process can substantially reduce the minimum feature size (externally). Consequently, the first open problem the authors list is whether their result can be generalized to planar straight-line graphs.

In fact, Ruppert’s Delaunay mesh refinement algorithm had already claimed constant degradation for proper triangulation of planar straight-line graphs [Rup93, Theorem 1],<sup>2</sup> but the “constant” actually depends on the minimum angle of the input graph (as well as the minimum triangle angle guaranteed by the algorithm).

*Our results.* We address the open problem of [BDE95] by showing that PSLGs have proper triangulations with  $\mathcal{O}(\lg n)$  degradation. This is the first triangulation algorithm to achieve a reasonable bound on degradation; even for polygons, the only previous results bound internal feature size, not minimum feature size. Our algorithm uses  $\mathcal{O}(n)$  Steiner points and hence  $\mathcal{O}(n)$  triangles, and runs in  $\mathcal{O}(n)$  time.

In [ADD<sup>+</sup>11], we argue that  $\Omega(\lg n / \lg \lg n)$  degradation is in fact necessary for minimum feature size, even in polygons. This implies that our upper bound is nearly tight, and resolves the open problem mentioned above. This lower bound applies even to quadrangulation or any decomposition into constant-complexity faces. We are currently working on the details of improving this result to  $\Omega(\log n)$ , which would completely settle this question.

We also show a separation between proper and nonproper triangulations. Specifically, by allowing Steiner points along the sides of triangles, we show that PSLGs have nonproper triangulations with  $\mathcal{O}(1)$  degradation, which is clearly optimal up to constant factors. We actually present this nonproper upper bound first, in Section 2, before describing the simple modifications needed to obtain  $\mathcal{O}(\lg n)$  degradation for proper triangulations in Section 3. Our method can also be used to re-obtain the  $\mathcal{O}(1)$  internal feature size degradation result of [BDE95]; see Section 4.

In all of our upper bounds, we focus on the case of triangulating a single polygon using small degradation. It is trivial to extend to polygons with holes. Because our triangulations do not add Steiner points to the boundary of the polygon, and because they approximately preserve minimum feature size instead of just internal feature size, they can be applied separately to each face of a PSLG to obtain the claimed results.

Table 1 summarizes the best results on degradation for each type of triangulation.

---

<sup>2</sup> Incidentally, this is also the paper that first introduced the notion of feature size.

## 2 Nonproper Triangulation

In this section, we show how to construct a nonproper triangulation of any polygon  $P$  with a degradation of  $\Theta(1)$ . We use  $\Theta(n)$  Steiner points, and the construction can be computed in linear time.

We begin by explaining how to triangulate parallelograms and trapezoids. Trivially any rectangle can be triangulated by placing a Steiner vertex at its center. The mfs will degrade by a factor of 2. Suppose instead that we are given a parallelogram  $P$  with top and bottom edges horizontal, and tilted toward the right. A segment with one endpoint at the lower-right vertex determines  $\text{mfs}(P)$ . Its direction depends on the height of  $P$  and the length of the horizontal edges. The segment is either vertical representing the height, or orthogonal to the left edge of  $P$ . Either way, placing the Steiner vertex at the center yields a degradation of 2, as can be easily verified by examining similar parallelograms. Specifically the new mfs will be determined by a segment parallel to the original one, from the Steiner vertex to the boundary of  $P$ .

**Lemma 1.** *Any trapezoid  $H$  can be triangulated with a degradation  $d_{trap} \leq 2$ .*

*Proof.* Let  $L$  be the shorter of the parallel edges on  $H$ , with length  $\ell$ , and without loss of generality, at the bottom of  $H$ . Let  $U$  be the top edge of  $H$ , and  $h$  be the height of  $H$ . Consider the rectangle  $R$  obtained by projecting  $L$  vertically upward onto the line through  $U$ . Suppose that  $R$  is contained in  $H$ . Then  $\text{mfs}(H) = \min\{h, \ell\}$  (i.e., it is determined by the dimensions of  $R$ ). We place a Steiner vertex  $s$  at the middle of  $R$ . The distance from  $s$  to  $L$  or  $U$  is  $\frac{h}{2}$ . The distance from  $s$  to the sides of  $H$  is greater than  $\ell/2$ . So if  $h \leq \ell$ , the degradation of the resulting triangulation is 2. Otherwise it is even less.

Now suppose that  $R$  is not contained in  $H$ , in which case we know that both side edges of  $H$  are slanted in the same direction, and without loss of generality, toward the right. Then  $\text{mfs}(H)$  is determined by a segment with one endpoint on the lower-right vertex of  $H$ . Consider the parallelogram  $P$  obtained by sweeping a horizontal segment of length  $\ell$  from  $L$  to  $U$ , while keeping its left endpoint on the left side of  $H$ . Then  $H$  and  $P$  have the same minimum feature size, determined by the same segment. The center of  $P$  is suitable for  $s$ . By construction,  $s$  separates  $P$  into four similar parallelograms. By preceding arguments described for parallelograms, the degradation of the resulting triangulation is 2.  $\square$

**Lemma 2 (Perturbation Lemma).** *Moving all the vertices of a PSLG  $G$  by at most  $\alpha \text{mfs}(G)$ , for  $\alpha < \frac{1}{2}$ , results in a PSLG  $G'$  with degradation at most  $\frac{1}{1-2\alpha}$  relative to  $G$ . The drawings of  $G$  and  $G'$  are combinatorially equivalent.*

*Proof.* Any distance determined by a point and a nonincident edge can be shortened by at most  $\alpha \text{mfs}(G)$  at each end, and thus  $2\alpha \text{mfs}(G)$  total. These distances were at least  $\text{mfs}(G)$  to begin with. Combinatorial equivalence follows from the fact that no vertex is allowed to move enough to cross a nonincident edge.  $\square$

The next lemma is essentially the most critical element of the main theorem that will follow.

**Lemma 3.** Let  $R$  be a rectangle with  $\text{mfs}(R) = w$  and height  $h > w$ . Let  $E$  be either the top or bottom edge of  $R$ , with length  $c \cdot w$  for some integer  $c$ . Let  $Z$  be the set of positions on  $E$ , at distances  $j \cdot w$  from one endpoint, for integers  $0 \leq j \leq c$ . If  $R$  has no Steiner vertices on its boundary, except possibly for positions in the set  $Z$ , then  $R$  can be nonproperly triangulated, without placing any additional Steiner vertices on its boundary, with constant degradation. Furthermore the number of additional Steiner vertices (inside  $R$ ) is  $O(c)$ .

*Proof.* The proof is by construction, specifically the triangulation  $G$  shown in Figure 2 (C), where as a worst-case scenario we have placed Steiner vertices at every multiple of  $w$  on the bottom edge of  $R$ .

The main construction has a set of edges anchored at one corner of  $R$  (upper-left in the figure). Starting from the shortest (and most clockwise), each such edge  $e_i$  reaches to a horizontal coordinate twice as large as the previous one, and to a vertical coordinate  $\frac{h}{3}$  from the bottom, where it meets the midpoint of a vertical edge  $g_i$  of length  $\frac{2h}{3}$ . Also, between every two such successive edges  $g_i$ , there is a region below one of the anchored edges, that has a sawtooth pattern matching the Steiner vertices. Specifically, for  $e_i$ , the sawtooth region is bounded by  $e_i$ ,  $g_i$ ,  $g_{i-1}$ , and the bottom edge of  $R$ .

Among the newly constructed Steiner vertices on  $e_i$  (i.e., on the top side of its corresponding sawtooth), the leftmost,  $s_0$ , is closest to edge  $e_{i+1}$ . This vertex happens to be the intersection of  $e_i$  with  $g_{i-1}$ . Because  $e_{i+1}$  reaches twice as far as  $e_i$ , but to the same vertical coordinate, the vertical separation between  $e_{i+1}$  and  $s_0$  is  $\frac{h}{6}$ . The horizontal separation between  $e_{i+1}$  and  $s_0$  is at least  $w$ . We conclude that the distance between  $e_{i+1}$  and  $s_0$  is at least  $w \frac{h}{6} \frac{1}{\sqrt{(\frac{h}{6})^2 + w^2}}$ . This distance lower bounds the feature size of all triangles emanating from the top-left of  $R$ .

Each sawtooth consists of the bottom horizontal edge, two vertical edges (the left twice the length of the right), and a tilted top, where the tilt angle gets closer to horizontal as the sawtooths move further to the right. The minimum distance created within any sawtooth occurs at its right hand side and is determined by the angle of the internal diagonals and by the tilt of the top. The distance is minimized as the tilt of the top increases and as the diagonal becomes less vertical. So, the minimum distance overall is to be found at the leftmost sawtooth, in its rightmost component (triangle). See Figure 3. The new distance introduced is at least  $\frac{h}{3} \frac{w}{2} \frac{1}{\sqrt{(\frac{h}{3})^2 + (\frac{w}{2})^2}}$ . Notice that both terms calculated imply constant degradation.  $\square$

Next we provide an intuitive lemma that will be of use to start the proof of our main theorem.

**Lemma 4.** Let  $C$  be the curve that is the locus of points inside polygon  $P$ , at distance  $\frac{1}{4}$  from the boundary of  $P$ . If  $\text{mfs}(P) = 1$  then every straight edge of  $C$  has length at least  $\frac{1}{2}$ .

*Proof.*  $C$  is obtained from  $P$  using the grassfire transform, commonly used as a visualization of the construction of the medial axis. Note that each edge  $e$  on  $P$  transforms to an edge  $e'$  on  $C$  continuously as the grassfire progresses. The shape of  $e'$  depends on local conditions; specifically the angles of vertices at the endpoints of  $e$ . Let  $e$  be



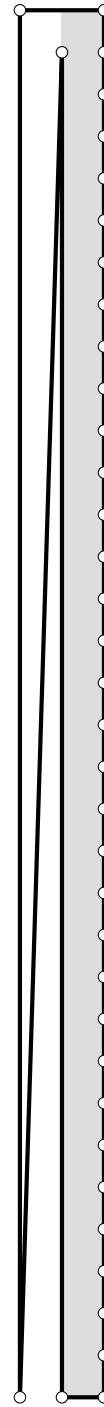
(A) Classic triangulation of  $R$  with all 66 vertices on the bottom edge. Observe that the interiors of many of the triangles are not discernible.



(B) Proper triangulation of  $R$  with all 66 vertices present, using a recursive construction. Here all triangles are discernible, but as the aspect ratio of  $R$  increases, the minimum feature size will slowly degrade.

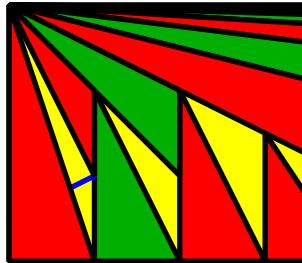


(C) Nonproper triangulation of  $R$ , using a novel construction. The triangulation is much easier to see than the previous two, and generalizing it to longer rectangles will not change the minimum feature size beyond that of this figure. The top figure shows the construction when all 66 vertices are present. The bottom figure illustrates how it can be adapted when vertices are not placed at all positions in  $Z$ .



(D) This polygon illustrates that Steiner vertices cannot be placed on a significant fraction of the boundary close to the reflex vertex. A construction like one of the three above can be used to triangulate the gray shaded area.

**Fig. 2.** Triangulations of a long rectangle  $R$ , with the properties stated in Lemma 3. In this figure,  $R$  has a set  $Z$  of 66 evenly spaced positions on the bottom edge. Vertices occupy some (or all) positions in  $Z$ , and the triangulations do not add more vertices on the boundary. Also illustrated is an example of why the ability to triangulate  $R$  in such a way is important.



**Fig. 3.** Closeup of Figure 2 (C); a short blue edge highlights the distance that defines the minimum feature size for the construction of  $R$  in Lemma 3

positioned horizontally, between vertices  $p_1, p_2$ . Suppose that the interior of  $P$  is below  $e$ . The edge  $e'$  must reside on the horizontal line at a distance  $\frac{1}{4}$  below  $e$ .

If both endpoints of  $e$  are reflex vertices on  $P$ , then  $e'$  will have the same length. If one of the vertices is reflex (without loss of generality, the left,  $p_1$ ), then the left endpoint  $v_1$  of  $e'$  will be located vertically below  $p_1$ , at a distance of  $\frac{1}{4}$ . Note that  $e'$  is just a subset of a longer edge on  $C$ . Follow a ray to the right of  $v_1$  for a distance of  $\frac{1}{2}$ , to construct a point,  $x$ . Let  $J_{p_1}$  be the unit quarter-circle in the lower-right quadrant of  $p_1$ . Then  $x$  is inside  $J_{p_1}$  and at a distance greater than  $\frac{1}{4}$  from the arc of  $J_{p_1}$ . Thus  $x$  cannot be a vertex on  $C$ , because it is not at a distance  $\frac{1}{4}$  from any point on  $P$  (excluding  $e$  itself). This implies that  $e'$  has length greater than  $\frac{1}{2}$ .

Finally, there is the case where both endpoints of  $e$  are convex vertices. Note that  $e'$  can have a length of  $\frac{1}{2}$  if  $e$  has length 1 and both convex angles are  $90^\circ$ . Then, the endpoints of  $e'$  are directed inward at angles of  $45^\circ$ , relative to the endpoints of  $e$ . If  $|e| = 1$ , then both convex angles must be at least  $90^\circ$ , so  $|e'| \geq \frac{1}{2}$ .

We can make  $e$  larger to allow for smaller convex angles at its endpoints. If we do so, the worst scenario is one where the edges adjacent to  $e$  in  $P$  are angled in a way that they eventually have a distance of 1 with each other (i.e., we close the angles as much as possible without violating feature size of  $P$ ). So, without loss of generality, assume that the angle at  $p_1$  is less than  $90^\circ$ . Follow the edge  $u$  neighboring  $e$  to the left until hitting a horizontal line at a distance of 1 from  $e$ . Note that this intersection point,  $y$ , must exist, otherwise we contradict the mfs assumption about  $P$  (the endpoint of  $u$  would be too close to  $e$ ). In other words  $y$  belongs to  $u$ . Construct the point  $z$  at a distance 1 vertically above  $y$ . This must be part of  $e$ , by construction. No part of the boundary of  $P$  intersects the triangle  $yzp_1$ . The left endpoint  $t_1$  of  $e'$  cannot be more than  $\frac{1}{4}$  to the right of the segment  $yz$ ; this happens if the angle at  $p_1$  is  $90^\circ$ , and as this angle decreases  $t_1$  moves to the left relative to  $yz$ . To visualize this, it is convenient to consider  $yz$  fixed, and move  $p_1$  to the left.

Consider the unit quarter-circle  $J_y$  centered at  $y$ , in its top-right quadrant. No part of  $P$  can intersect  $J_y$ ; if an edge not adjacent to  $y$  does so, it will be too close to  $y$ , and if an adjacent edge does so (if  $y$  were a real vertex), it will be too close to  $e$ . Now consider the vertex  $p_2$ , common to  $e$  and its edge  $e_r$  to the right. Wherever  $p_2$  is,  $|e'|$  will be minimized if we minimize the angle at  $p_2$ . If this angle is less than  $90^\circ$ , then because  $e_r$  must miss  $J_y$ , we follow the same reasoning as above to easily conclude that  $e'$  has

length greater than 1. That is, the right end  $t_2$  of  $e'$  will not be more than  $\frac{1}{4}$  to the left of some vertical segment  $y'z'$  analogous to  $yz$ .

So we are left with the case where  $p_2$  has angle greater than  $90^\circ$  and is located within one unit of  $z$ , i.e., its vertical projection intersects  $J_y$ . As mentioned,  $|e'|$  will be minimized if the angle at  $p_2$  is minimized, which is to say that  $e_r$  should be rotated as clockwise as possible, until it becomes tangent to  $J_y$ . This means that the bisector at  $p_2$  intersects  $y$ . So  $p_2$  should be placed as close to  $z$  as possible, to minimize  $|e'|$ . We will now work within the triangle  $p_1yp_2$ .

Let  $q$  be the intersection of  $yz$  with the line through  $e'$ . We know that  $t_2$  will be to the right of  $q$ . On the other hand,  $t_1$  can be arbitrarily to the left, or slightly to the right (specifically no more than  $\frac{1}{4}$ ). Let  $t_2 - z$  be  $e'_2$ , and  $z - t_1$  be  $e'_1$ . Then  $|e'|$  is equal to  $e'_1 + e'_2$ .

There is another constraint on the location of  $p_2$ : it cannot be too close to  $u$ . Let  $\omega$  be the angle at  $p_1$ . Then  $\sin \omega = \frac{1}{|e|}$ , meaning  $p_2$  is placed so that its distance to  $u$  is 1. Let  $z - p_1$  be  $e_1$ . From triangle  $yzp_1$  we have  $\tan \omega = \frac{1}{e_1}$ .

Let  $e_2$  be  $p_2 - z$ . Triangle  $yzp_2$  is similar to triangle  $yqt_2$ , so  $e'_2 = \frac{3e_2}{4}$ . From the triangle formed by  $p_1, t_1$ , and the projection of  $t_1$  onto  $e$ , we have  $\tan \frac{\omega}{2} = \frac{0.25}{e_1 - e'_1}$ ; ( $t_1$  is  $\frac{1}{4}$  below  $e$ , and on the bisector of  $\omega$ ).

Reordering and combining from above, we have:

$$e'_1 = e_1 - \frac{1}{4 \tan \frac{\omega}{2}} = \frac{1}{\tan \omega} - \frac{1}{4 \tan \frac{\omega}{2}}. \text{ Because } e_2 = \frac{1}{\sin \omega} - e_1, \text{ we have}$$

$$e'_2 = \frac{3}{4} \left( \frac{1}{\sin \omega} - e_1 \right) = \frac{3}{4} \left( \frac{1}{\sin \omega} - \frac{1}{\tan \omega} \right).$$

Notice that when  $\omega$  increases to  $90^\circ$  (i.e.,  $p_1$  approaches  $z$  from the left), we have  $e'_1 = -\frac{1}{4}$  and  $e'_2 = \frac{3}{4}$ , so  $|e'| = \frac{1}{2}$ . For  $\omega \leq 90^\circ$ ,  $e'$  is longer.  $\square$

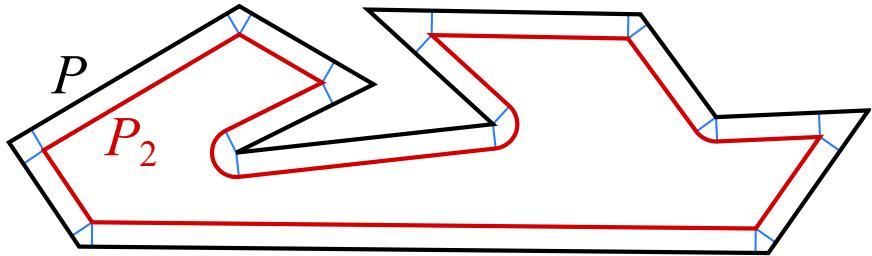
Now we prove the main theorem of this section:

**Theorem 1.** *Every  $n$ -gon has a nonproper triangulation with constant degradation, using  $O(n)$  Steiner vertices.*

*Proof.* Let  $P$  be an  $n$ -gon with minimum feature size 1. Let the curve  $P_2$  be the locus of points inside  $P$  that have minimum distance  $\frac{1}{4}$  from  $\partial P$ . This is obtained from the well-known grassfire transformation.  $P_2$  is a closed curve consisting of  $n$  line segments (one per segment of  $P$ ) as well as one circular arc corresponding to every reflex vertex of  $P$ . Each such arc spans less than  $180^\circ$ . See Figure 4 for an illustration of this process. Because  $P$  has minimum feature size 1, each line segment in  $P_2$  has length at least  $\frac{1}{2}$ , by Lemma 4.

$P_2$  splits the interior of  $P$  into two regions, which we call the *Interior* and the *Tube*.  $P_2$  itself belongs to both regions. We will modify and refine this boundary a few times, and then triangulate each region separately. Any operations in the interior of one region will not affect degradation in the other.

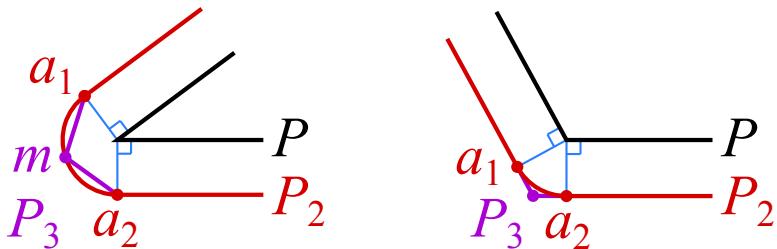
*Refinement of  $P_2$ :* Refer to Fig. 5. We create a polygon  $P_3$  by replacing all circular arcs on  $P_2$  with polylines. Let  $O$  be a given circular arc with endpoints  $a_1$  and  $a_2$ . If  $O$  spans more than  $90^\circ$ , we can replace it with segments  $a_1m$  and  $ma_2$ , where  $m$  is the midpoint of the arc. The minimum distance between  $P_3$  and  $P$  is  $\frac{1}{4\sqrt{2}}$  (achieved when  $O$  approaches  $180^\circ$ ). So far, this distance lower bounds the feature size of the



**Fig. 4.** A polygon  $P$  and the closed curve  $P_2$  created by the grassfire transformation

Tube, because (when  $O = 90^\circ$ ) the segments on  $P_3$  can have length as short as the side of a regular octagon of diameter  $\frac{1}{4}$ . That is,  $\frac{1}{4}2 \sin \frac{\pi}{8} \approx 0.19$ . We say that the Tube degradation is at most  $d_{\text{Tube}3} = 4\sqrt{2}$ . On the other hand, the feature size  $\frac{1}{d_3}$  of the Interior is  $\frac{1}{4}2 \sin \frac{\pi}{8}$ , i.e., the feature size is smallest on its boundary  $P_3$ . Distances through the interior of  $P_3$  are still at least  $\frac{1}{2}$ .

If instead  $O$  spans less than  $90^\circ$ , we extend its adjacent edges on  $P_2$ , through  $a_1$  and  $a_2$  respectively, until they meet. This extension remains at a distance greater than  $\frac{1}{4}$  from  $P$ , so the Tube degradation is unaffected. The extension also remains at most  $\frac{\sqrt{2}}{4}$  from the vertex on  $P$  that generated the arc (the max is achieved when the angle is  $90^\circ$ ). All other points on  $P_3$  are even closer to the boundary of  $P$ . Thus no two points from different edges on  $P_3$  will get closer than  $1 - \frac{\sqrt{2}}{2}$  to each other (roughly 0.29, not enough to reduce our bound on the feature size of the Interior).



**Fig. 5.** How to transform  $P_2$  to  $P_3$ . Left: arc spans more than  $90^\circ$ ; Right: arc spans less than  $90^\circ$ .

Let  $P_4$  be formed by snapping the vertices of  $P_3$  vertically to a horizontal grid of granularity  $g = \frac{1}{2d_3}$ . Any point can snap at most a distance of  $\frac{1}{4d_3}$ ; half the grid size. A pair of points on  $P_3$  at a *nearly* co-vertical position a distance of  $\frac{1}{d_3}$  from each other may snap toward each other, so snapping can degrade the feature size of  $P_3$ , and thus also the Interior, by a factor of 2. The effect on the minimum distance between  $P$  and  $P_3$  is smaller, because vertices of  $P$  remain fixed. This distance can drop to  $\frac{1}{d_{\text{Tube}3}} - \frac{1}{4d_3}$ . At this point though, the minimum distance in both regions is to be found on their common boundary, and the value is  $\frac{1}{2d_3}$ .

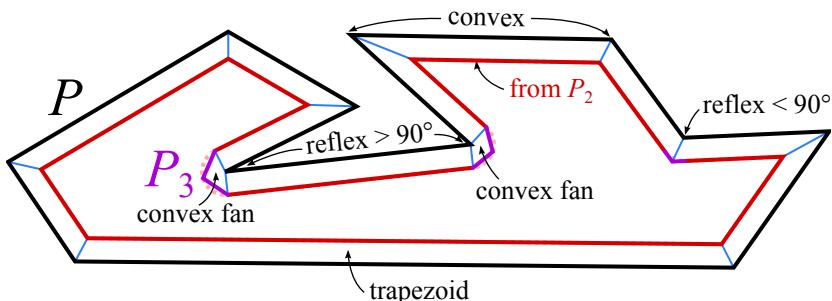
*Triangulation of Interior:* Let  $P_5$  consist of  $P_4$  and the horizontal trapezoidation of the interior of  $P_4$ . All vertices on  $P_4$  lie on the grid and thus the feature size is preserved. Let  $P_6$  consist of the triangulation of  $P_5$  obtained by placing a vertex in each trapezoid according to the method presented in Lemma 1. Thus, the degradation in this step is  $d_6 = d_{trap}$ .

The total degradation of the Interior is therefore  $2d_3d_6$ . The feature size of the Interior is  $\frac{1}{8} \sin \frac{\pi}{8} \approx 0.047$ .

*Triangulation of Tube:* Obtaining a triangulation of the Tube is done without adding more Steiner vertices to its boundary, and thus any degradation in this process will not amplify the degradation of Interior, or affect feature size via the exterior of  $P$ .

Before proceeding to the algorithm, which has four main steps, we require one definition: A *quasi-trapz* is a quadrilateral that can be transformed into a trapezoid by perturbing its vertices by an amount small enough so that the Perturbation Lemma can be applied.

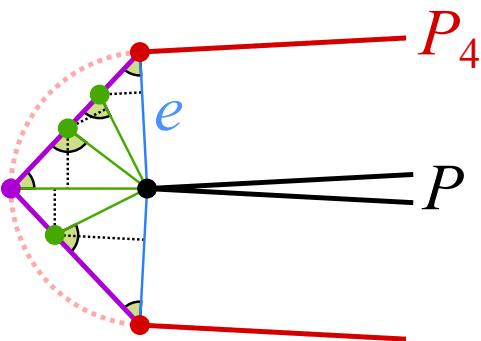
1. **Subdivision of the Tube into triangles and quasi-trapz.** Consider all convex polygonal chains that were created in  $P_3$  as replacements of circular arcs spanning more than  $90^\circ$  on  $P_2$ . Recall that such chains consist of two segments, which by now can also contain Steiner vertices from the trapezoidation of  $P_4$ . For each chain, we connect the endpoints to the unique reflex vertex  $v_r$  of  $P$  that generated the corresponding (replaced) arc via the grassfire transform. Similarly, we connect every convex vertex of  $P$  to its corresponding convex vertex on  $P_4$ , and we connect reflex vertices of  $P$  (with arcs spanning less than  $90^\circ$ ) to their unique corresponding reflex vertex on  $P_4$ . This subdivides the Tube into quasi-trapz and *convex fans* (i.e., a vertex visible from a convex chain). In fact any such fan is just a quadrilateral, because the chain opposite  $v_r$  had only two edges on  $P_3$  (with Steiner vertices added later on). See Figure 6.



**Fig. 6.** Phase 1 of Tube triangulation: Subdividing Tube into fans and quasi-trapz. Angles indicated correspond to arc spans.

The only degradation caused can be due to a newly created edge,  $e$ , and some nonincident vertex  $p$  on  $P_4$ . What matters is the angle that  $e$  makes with  $P_4$ , and

the proximity of  $p$  to the endpoint of  $e$  on  $P_4$ . The latter is at least  $\frac{1}{2d_3}$ . Without taking snapping into account, the aforementioned angle would be no less than  $45^\circ$ . See Figure 7. So, adding these edges would only degrade feature size from  $\frac{1}{2d_3}$  by a factor of  $\sqrt{2}$ . This means that the feature size of the fans could drop to roughly 0.067. This is not close to the smaller feature size in other areas that will be created, and snapping the vertices of the fan will not have any significant effect. The snapping would have to be so extreme that the angle mentioned would drop from  $45^\circ$  to under  $14^\circ$ . This is not possible when two vertices of a triangle move by less than a quarter of the shortest length.



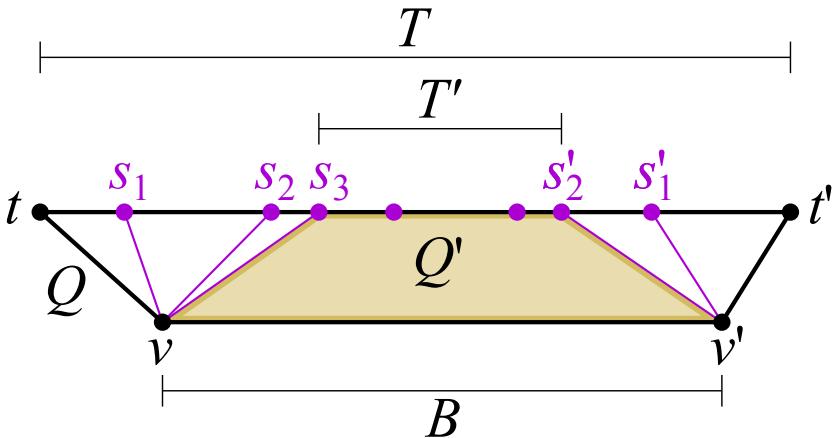
**Fig. 7.** Triangulating a convex fan: the angle between  $P_4$  (purple) and edge  $e$  (blue) is at least  $45^\circ$ . The same holds for other edges (green) from  $P$  to Steiner points on the fan (which may have been introduced when constructing  $P_6$  from  $P_4$ ). Dotted black segments are minimum distances in newly created triangles.

We continue to triangulate each fan by adding diagonals from  $v_r$  to all remaining vertices within. The preceding analysis follows verbatim. Structurally, the end result is that any nontriangulated region of the Tube is a quasi-trapz that has an edge of  $P$  (the *bottom*) and an edge of  $P_4$  (the *top*) on its boundary. Were it not for the snapping, the top and bottom would be parallel, and the quasi-trapz would be a trapezoid. Note that the top can contain Steiner vertices, generated during the trapezoidation of  $P_4$ .

2. **Subdivision of quasi-trapz with Steiner vertices on the top.** We will subdivide quasi-trapz so that the only remaining nontriangulated regions will be quasi-trapz without Steiner points on their boundary, and rectangles possibly with such Steiner points.

Let  $Q$  be a quasi-trapz to be subdivided. By assumption,  $Q$  has at least one Steiner vertex on its top,  $T$ . We will start adding diagonals from the bottom,  $B$ , to Steiner vertices on  $T$ . This will progressively cut off triangles, leaving a smaller quasi-trapz  $Q'$ . As we do this, we will be shortening  $T$ , so that it either has no Steiner vertices on it, or the internal angles of  $Q'$  at  $T$  are at least  $135^\circ$ . This will also imply that  $T$  comfortably projects onto the bottom,  $B$ , in a direction orthogonal to  $T$ .

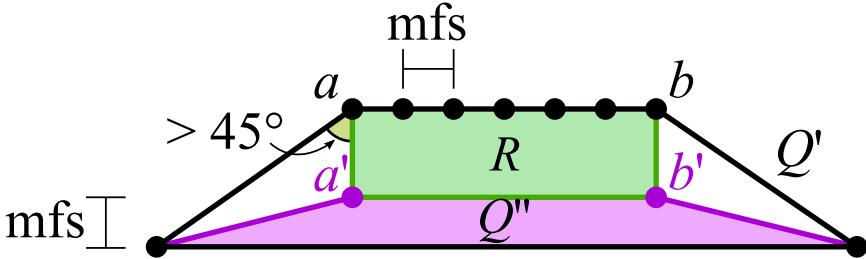
For each endpoint  $t$  of  $T$  with internal angle smaller than  $135^\circ$  (note that the angle is at least  $45^\circ$  to start with), do the following. Let  $v$  be the neighbor of  $t$  on  $Q$ , on edge  $B$ . Traverse  $T$  from  $t$  until a Steiner vertex  $s$  is found, and join  $s$  to  $v$ . While the angle condition is not met, keep forming such an edge to  $v$  for each successive  $s$ . This repeatedly cuts off triangles from  $Q$ , until it is either a triangle or Steiner-free quasi-trapz (if no more Steiner vertices remain), or until the angle condition is satisfied. The triangles cut off on each side form a convex fan triangulation (with  $v$  as the apex), identical in nature to those described in phase 1. Thus the degradation caused so far can be absorbed into the preceding analysis. See Figure 8.



**Fig. 8.** Phase 2 of Tube triangulation: decomposing  $Q$  into  $Q'$  and triangulated fans. Either the internal angles of  $Q'$  at the top are greater than  $135^\circ$ , or  $Q'$  has no Steiner vertices.

Let  $T'$  be the subset of  $T$  left over from this process (see Figure 8). Now  $T'$  and  $B$  form the parallel edges of  $Q'$ , which is a subset of  $Q$ . Consider the shape of  $Q'$  as it existed before snapping. Recall that before snapping,  $T'$  and  $B$  were parallel, at distance  $\frac{1}{4}$ . The minimum feature size of  $Q'$  is lower bounded by the separation of Steiner vertices on its boundary, i.e.,  $\frac{1}{2d_3}$ . Let  $a$  and  $b$  be the endpoints of  $T'$ . Let  $a'$  and  $b'$  be vertices placed at a distance  $\frac{1}{4} - \frac{1}{2d_3}$  away from  $a$  and  $b$ , respectively, so that  $ba'a'b'$  is a rectangle  $R$  inside  $Q'$ . Notice that  $a'$  and  $b'$  are  $\frac{1}{2d_3}$  from  $B$ , and even further from the sides of  $Q'$ . So this placement doesn't affect the feature size. Now, connect  $a'$  and  $b'$  to the vertices on  $Q'$ . Inside  $Q'$ , we are left with a Steiner-free trapezoid  $Q''$  (it is below  $a'b'$  and will again become a quasi-trapz when we account for snapping), the rectangle  $R$  with Steiner points on the edge  $ab$ , and two triangles. See Figure 9.

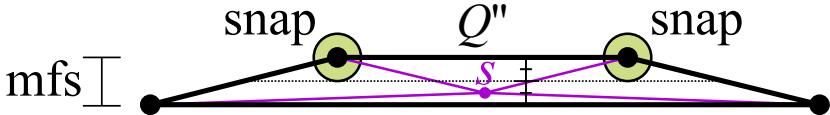
Finally we must reinstate the snapping of the segment  $ab$ . The positions of  $a'$  and  $b'$  will follow so that  $R$  moves rigidly. We now examine the effect of the motion of  $a'$  on the feature size of the components of  $Q'$ . Recall that  $a'$  is snapped by at most  $\frac{1}{4d_3} = \frac{1}{16\sqrt{2}}$  (roughly 0.044). So it can reduce the feature size of  $Q''$  to  $\frac{1}{4d_3}$  (by moving half way to the fixed edge  $B$  on  $Q''$ ). The effect of  $a'$  is even smaller on the



**Fig. 9.** Phase 2 of Tube triangulation: decomposing  $Q'$  into  $R$  and  $Q''$

feature size of the triangles in  $Q'$ , because its distance to their nonincident edges is greater than its distance to  $B$ . Of course, there is no effect on  $R$ .

3. **Triangulate all remaining Steiner free quasi-trapz.** Here we triangulate any quasi-trapz  $Q''$  constructed in the previous phase. One Steiner vertex  $s$  suffices, as with any trapezoid. There is a placement for  $s$  in the corresponding unsnapped trapezoid  $Q$  so that degradation is no more than 2. Because  $Q$  has height  $(\frac{1}{2d_3})$ ,  $s$  would normally be placed between the parallel edges of  $Q$ , i.e.,  $\frac{1}{4d_3}$  from  $B$ . However the edge  $a'b'$  might snap by this much, and this would create an arbitrarily small distance to  $s$ . So, instead we will place  $s$  at a distance  $\frac{1}{8d_3}$  from  $B$ , because  $B$  will remain fixed. The effect is that the feature size of  $Q''$  can be reduced to  $\frac{1}{8d_3}$ , but no less. Currently, this quantity lower bounds the overall feature size, at roughly 0.022. See Figure 10.



**Fig. 10.** Phase 3 of Tube triangulation: handling the last remaining quasi-trapz

4. **Triangulate rectangles generated in phase 2.** For each rectangle  $R = abb'a'$ , we use the construction presented in Lemma 3.  $R$  has height  $h = \frac{1}{4} - \frac{1}{2d_3}$ , and Steiner vertices are placed on one of its longer sides, at distances of  $\frac{1}{2d_3}$ . Then the formula in Lemma 3 yields a value of greater than 0.024 for  $\text{mfs}(R)$ .

The important conclusion is that each step described incurs a constant degradation, therefore the aggregate is also constant. Most of the steps described probably have tighter bounds. Furthermore, these steps could be optimized to work more harmonically, or replaced with more efficient constructions. For the record, we claim here that the degradation is under 45 (the inverse of 0.022 calculated in phase 3 in the Tube).

The construction of  $P_3$  adds a linear number of Steiner vertices, since a constant number of them is associated with each vertex of  $P$ . No Steiner vertices are added when we form  $P_4$ , since this only involves snapping.  $P_5$  is a trapezoidation formed by extending a line from every vertex. This creates a linear number of Steiner vertices (and trapezoids). As we construct  $P_6$ , we add one Steiner vertex per trapezoid. Thus the

boundary and interior of  $P_2$  contain a linear number of Steiner vertices. Finally, when we triangulate the Tube, we only add Steiner vertices to the interior. Step 1 adds no vertices. Step 2 adds two vertices to form  $Q''$  and  $R$ , in each quasi-trapz. Step 3 adds one vertex per quasi-trapz. There is one quasi-trapz per edge of  $P$ , so all of these steps add a linear number of Steiner vertices. Step 4 adds a number of vertices proportional to the number of vertices on the boundary of the rectangle  $R$  (by Lemma 3). The total number of vertices on all such rectangle boundaries is  $\mathcal{O}(n)$ , because they are formed when processing the interior of  $P_2$ .  $\square$

### 3 Proper Triangulations

Next we describe the few modifications necessary for proper triangulations, which lose a logarithmic factor in minimum feature size:

**Theorem 2.** *Every  $n$ -vertex polygon has a proper triangulation with  $\mathcal{O}(\log n)$  degradation.*

*Proof.* Use the algorithm of Section 2, with the exception of the nonproper triangulation of the rectangle  $R = abb'a'$ . That triangulation is instead done using the construction of Figure 2 (B). This is a simple recursive decomposition of a rectangle into proper triangles. For  $n$  vertices on the bottom of  $R$ , there are  $\mathcal{O}(\log n)$  horizontal layers in the construction. The separation between those vertices defines the minimum feature size of the input. The height of  $R$  could also equal this value (any more only helps to preserve feature size). Therefore the construction of these layers can cause a degradation of  $\mathcal{O}(\log n)$ . Triangulating the layers only affects the multiplicative constant.  $\square$

### 4 Internal Feature Size

For internal feature size in a polygon, we can avoid losing constant degradation or properness. This result is already known [BDE95], but can also be obtained by our methods:

**Theorem 3.** *Every  $n$ -vertex polygon has a proper triangulation with  $\mathcal{O}(1)$  internal feature size degradation.*

*Proof.* The only component where our triangulation from Section 2 is nonproper is within the quasi-trapz in the Tube, or more specifically, the rectangular regions. Instead, if we are not concerned with external feature size, we can triangulate the quasi-trapz by creating Steiner vertices on the boundary of  $P$ , to match those on the inner boundary of the Tube.  $\square$

### 5 Pseudopolynomial Dissection

Our original motivation for finding meshes that approximately preserve minimum feature size came from the classic problem of *geometric dissection*. The nearly

200-year-old algorithm of Lowry [Low14, Fre97] *dissects* any two given polygons  $P^1$ ,  $P^2$  of equal area into polygonal pieces such that the pieces of  $P^1$  can be translated and rotated to make up the pieces of  $P^2$ . Unfortunately, the number of pieces it uses can be extremely large.

How many pieces does polygon dissection need? In particular, do a *pseudopolynomial* number of pieces suffice? In computational geometry, a *pseudopolynomial* bound is polynomial in the number  $n$  of input coordinates and the size of a grid needed to express those coordinates. The latter bound is typically approximated by the *spread*  $r = D/w$ , where  $D$  is the diameter and  $w$  is the minimum feature size.<sup>3</sup> Some dependence beyond  $n$  is necessary for dissection: for example, dissecting a square into an  $r \times 1$  rectangle requires  $\Omega(\sqrt{r})$  pieces by a simple diameter argument.

We prove here that any two polygons  $P^1$  and  $P^2$  have a dissection using  $(n(P^1) + n(P^2)) \cdot (r(P^1) + r(P^2))$  pieces. This result follows by combining Lowry's original algorithm with an ifs-preserving triangulation algorithm, such as ours or the one in [BDE95]. More generally, for  $k$  polygons  $P^1, \dots, P^k$ , we obtain a dissection using  $k(n(P^1) + \dots + n(P^k)) \cdot (r(P^1) + \dots + r(P^k))$  pieces.

Lowry's algorithm starts by triangulating the polygon  $P^1$ , and then uses a dissection of Montucla from 1778 (see [Fre97]) to convert each triangle into a rectangle with a common height  $\varepsilon$ . The largest suitable  $\varepsilon$  is half the minimum height of all triangles, i.e., half the internal feature size of the triangulation. The common-height rectangles can then be assembled into one  $A/\varepsilon \times \varepsilon$  rectangle, where  $A$  is the area of  $P^1$ . The resulting number of pieces is  $\mathcal{O}(nA/\varepsilon)$ ; furthermore, each piece is bounded by a constant number of cuts, and the number of cuts hit by any vertical line is  $\mathcal{O}(1)$ . Finally the algorithm repeats this process for  $P^2$ , and overlays the two dissections of the  $A/\varepsilon \times \varepsilon$  rectangle. By the properties above, this overlay increases the number of cuts and thus pieces by only a constant factor, resulting in  $\mathcal{O}((n(P^1) + n(P^2))A/\varepsilon)$  pieces. More generally, for  $k$  polygons, each cut can be divided  $k$  times, so we obtain a piece bound of  $O(k(n(P^1) + \dots + n(P^k))A/\varepsilon)$ .

Lowry's algorithm does not specify a triangulation, so cannot efficiently bound  $\varepsilon$ . While a classic triangulation was originally intended, any nonproper triangulation suffices. Using an ifs-preserving triangulation, we obtain a triangulation with  $\varepsilon = \Theta(\text{ifs}(P^1))$ . Rescaling to make  $A = 1$  does not affect the algorithm, and uses a scale factor no smaller than  $\mathcal{O}(1/D)$ , where  $D$  is the diameter of any polygon. Thus we obtain  $\varepsilon = \Omega(\text{ifs}(P^1)/D) = \Omega(1/r)$  and  $A = 1$ . Plugging these bounds into the piece bound above, we obtain the desired result.

## 6 Discussion

Although several steps in our construction require subtle constructions and details to keep things at constant distance, we believe that the essential hurdle was how to triangulate rectangles with many Steiner points on a side, without adding new Steiner points on the boundary. This was the breakthrough needed to solve the problem at hand.

Preserving minimum feature size is by no means the only priority in meshing, but it is still a meaningful (and well-studied) measure of mesh quality. We leave to future

---

<sup>3</sup> This parameter also arises in many meshing results; we saw one in Table 1.

work the possibility of simultaneously attaining small feature-size degradation with other important mesh properties, such as maximum angle bounded away from  $180^\circ$ . This goal may be attainable by simply combining algorithms in a careful way.

Another direction for further research would be to extend our results to 3D. Our grassfire approach should work just as well. The central challenge, as in 2D, would seem to be the proper triangulation (tetrahedralization) of a box that is very thin in one or two of its dimensions.

**Acknowledgments.** This research was initiated at the 24th Annual Winter Workshop on Computational Geometry, co-organized by Erik Demaine and Godfried Toussaint, and held at the Bellairs Research Institute of McGill University in February 2009. We thank the other participants of the workshop for their helpful comments and for creating an environment conducive to creative thought: Zachary Abel, Brad Ballinger, Nadia Benbernou, Prosenjit Bose, Jean Cardinal, Sébastien Collette, Mirela Damian, Robin Flatland, Ferran Hurtado, Scott Kominers, Stefan Langerman, Robbie Schweller, David Wood, and Stefanie Wuhrer.

## References

- ADD<sup>+</sup>11. Aloupis, G., Demaine, E., Demaine, M., Dujmovic, V., Iacono, J.: Meshes preserving minimum feature size. Technical report arXiv:0908:2493 (2011)
- BDE95. Bern, M., Dobkin, D., Eppstein, D.: Triangulating polygons without large angles. *International Journal of Computational Geometry & Applications* 5(1-2), 171–192 (1995)
- BE95. Bern, M., Eppstein, D.: Mesh generation and optimal triangulation. In: Du, D.-Z., Hwang, F.K.-M. (eds.) *Computing in Euclidean Geometry*, 2nd edn. Lecture Notes Series on Computing, vol. 4, pp. 47–123. World Scientific (1995)
- Dey07. Dey, T.K.: Delaunay mesh generation of three dimensional domains. Technical Report OSU-CISRC-09/07-TR64, Ohio State University (October 2007)
- Eri03. Erickson, J.: Nice point sets can have nasty delaunay triangulations. *Discrete & Computational Geometry* 30(1) (July 2003)
- Fre97. Frederickson, G.N.: *Dissections: Plane and Fancy*. Cambridge University Press (November 1997)
- HMP06. Hudson, B., Miller, G., Phillips, T.: Sparse Voronoi Refinement. In: *Proceedings of the 15th International Meshing Roundtable*, Birmingham, Alabama, pp. 339–356. Springer (September 2006)
- Low14. Lowry: Solution to question 269, [proposed] by Wallace, W. In: Leybourn, T. (ed.) *Mathematical Repository*, vol. 3, part 1, pp. 44–46. W. Glendinning, London (1814)
- Rup93. Ruppert, J.: A new and simple algorithm for quality 2-dimensional mesh generation. In: *Proceedings of the 4th Annual ACM-SIAM Symposium on Discrete Algorithms*, Austin, Texas, pp. 83–92 (1993)
- She04. Shewchuk, J.R.: Theoretically guaranteed Delaunay mesh generation—in practice. Short course at 13th International Meshing Roundtable (2004), <http://www.cs.berkeley.edu/~jrs/papers/imrtalk.pdf>

# Geometric Graphs in the Plane Lattice

Mikio Kano<sup>1</sup> and Kazuhiro Suzuki<sup>2,\*</sup>

<sup>1</sup> Ibaraki University, Hitachi, Ibaraki, Japan

[kano@mx.ibaraki.ac.jp](mailto:kano@mx.ibaraki.ac.jp)

<http://gorogoro.cis.ibaraki.ac.jp/>

<sup>2</sup> Kochi University, Kochi, Kochi, Japan

[kazuhiro@tutetuti.jp](mailto:kazuhiro@tutetuti.jp)

**Abstract.** An *L*-line segment in the plane consists of a vertical line segment and a horizontal line segment having a common end-point. In this paper, we consider some problems on non-crossing geometric embeddings of graphs in the plane lattice, whose vertices are given points of the plane lattice in general position and whose edges are suitable *L*-line segments.

## 1 Introduction

A geometric graph in the plane is a graph drawn in the plane whose edges are straight line segments. There is a large amount of research on geometric graph drawings or geometric graph embeddings in the plane without crossings (for example, see [6], [3] and [5]). In this paper we consider graph drawings or graph embeddings in the plane lattice without crossings, whose edges are *L*-line segments defined below instead of straight line segments.

For a point  $x$  in the plane lattice, an *L*-shaped line consisting of a vertical ray and a horizontal ray emanating from  $x$  is called an *L-line*. Similarly, an *L*-shaped line segment consisting of a vertical line segment and a horizontal line segment with a common end-point is called an *L-line segment* (see (1) of Fig. 1). We consider some problems on geometric graphs in the plane lattice, whose edges are *L*-line segments. A set  $X$  of points in the plane is *in general position* if no three points of  $X$  lie on the same line. On the other hand, a set  $S$  of points in the plane lattice is said to be *in general position* if every vertical line and every horizontal line passes through at most one point of  $S$ . It is shown that *L*-line segments play a similar role in the plane lattice as line segments in the plane ([2]).

In this paper we consider some problems on geometric graph drawings and embeddings in the plane lattice, in which *L*-line segments plays a similar role as line segments in geometric graph drawings and embeddings in the plane. For example, for a 3-tree  $T$  of order  $|T|$  and for any given set  $S$  of  $|T|$  points in the plane lattice in general position, we consider a non-crossing geometric embedding of  $T$  on  $S$ , where a 3-tree is a tree with maximum degree at most three. In [7]

---

\* Partially supported by research grant by Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C).

and [4], some problems of non-crossing geometric embeddings or drawings of matchings and cycles are considered, and it is shown that they are  $\mathcal{NP}$ -hard. Some other results in the plane lattice related to this paper can be found in [1] and [6].

## 2 Geometric Hamilton Cycles

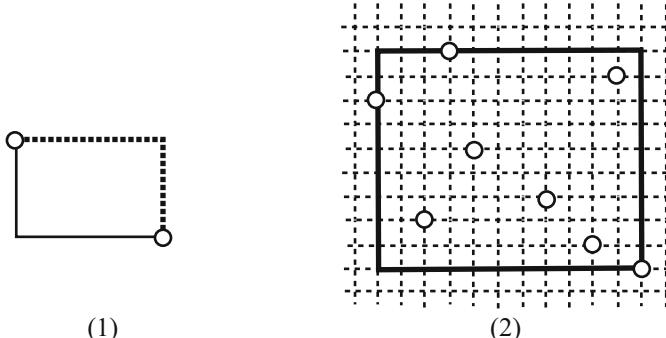
For a set  $X$  of points in the plane lattice in general position, a cycle (path) which passes through all the points of  $X$  and whose edges are  $L$ -line segments is called a *geometric Hamilton cycle (path) on  $X$* . The *rectangular hull* of  $X$ , denoted by  $rect(X)$ , is the smallest closed rectangle enclosing  $X$ , each of whose edges is a vertical or horizontal line segment ((2) of Fig. 1). In particular, every edge of  $rect(X)$  contains exactly one point of  $X$ . The cardinality of  $X$  is denoted by  $|X|$ .

We begin with the following result.

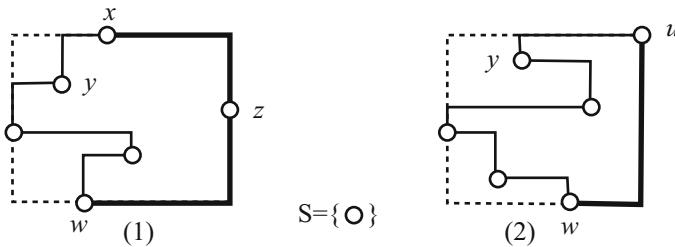
**Theorem 1.** *Let  $S$  be a set of points in the plane lattice in general position containing at least two points. Then there exists a non-crossing geometric Hamilton cycle on  $S$ .*

*Proof.* We may assume that  $|S| \geq 3$ . We first consider the case where the top edge, the left edge and the right edge of the rectangular hull  $rect(S)$  contain one distinct point of  $S$  each. Let  $x$  be the point of  $S$  lying on the top edge of  $rect(S)$ . By symmetry, we may assume that the second top point  $y$  of  $S$  is to the left of  $x$ . Take the point  $z$  of  $S$  lying on the right edge of  $rect(S)$ . Then there exists a non-crossing geometric Hamilton path on  $S - \{z\}$  starting at  $x$  and ending at the point  $w$  lying on the bottom edge of  $rect(S)$ . By adding two  $L$ -line segments joining  $w$  to  $z$  and  $z$  to  $x$  to this Hamilton path, we obtain the desired non-crossing geometric Hamilton cycle on  $S$  (see (1) of Fig. 2).

We next consider the case where the top-right corner of  $rect(S)$  contains a point of  $S$ , say  $u$ . There exists a non-crossing geometric Hamilton path on  $S$ ,



**Fig. 1.** (1) Two  $L$ -line segments joining two points; (2) A rectangular hull  $rect(X)$  of a set  $X$  of points in the plan lattice in general position



**Fig. 2.** (1),(2) Non-crossing geometric Hamilton cycles on  $S$

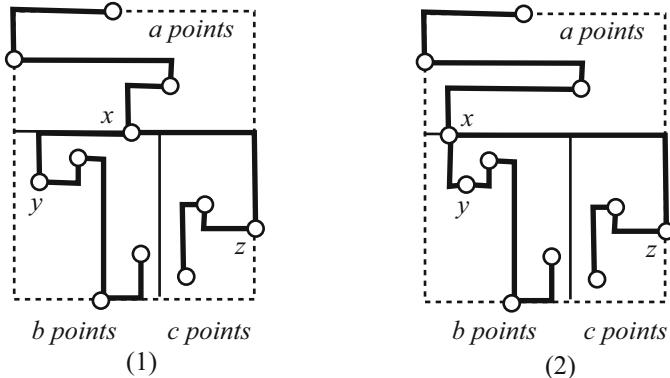
which starts at  $u$  and ends at the point  $w$  lying on the bottom edge of  $\text{rect}(S)$  (see (2) of Fig. 2). Then by adding an  $L$ -line segment connecting  $u$  and  $w$ , we obtain the desired non-crossing geometric Hamilton cycle on  $S$ .  $\square$

### 3 Geometric Embeddings of Spiders

Let  $G$  be a graph with vertex set  $V(G)$  and edge set  $E(G)$ . The degree of a vertex  $v$  of  $G$  is denoted by  $\deg_G(v)$ . The number of vertices of  $G$  is called the *order* of  $G$ , and denoted by  $|G|$ . Let  $S$  be a set of  $|G|$  points in the plane lattice in general position. Then we say that  $G$  is *geometrically embedded* on  $S$  if there exist a bijection  $\psi : V(G) \rightarrow S$  and a set of  $L$ -line segments joining two points of  $S$  such that two points  $\psi(x)$  and  $\psi(y)$  of  $S$  are joined by an  $L$ -line segment if and only if  $x$  and  $y$  are joined by an edge of  $G$ . We often call the above embedding a *geometric embedding*  $\psi : G \rightarrow S$ . If no two  $L$ -line segments intersect, such an geometric embedding is said to be *non-crossing*. A *spider* is a tree that has exactly one vertex with degree at least three, and the vertex of degree at least three is called its *trunk*.

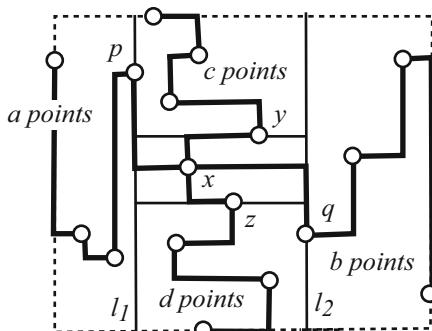
**Theorem 2.** *Let  $T$  be a spider such that the trunk has degree three or four, and let  $S$  be a set of  $|T|$  points in the plane lattice in general position. Then  $T$  can be geometrically embedded on  $S$  without crossings.*

*Proof.* Let  $u$  be the trunk of the spider  $T$ . We first consider the case where  $u$  has degree three. Then  $T - u$  consists of three paths, and their orders (i.e., the numbers of their vertices) are denoted by  $a, b, c$ , respectively. Sweep the points of  $S$  from top to bottom by horizontal line, and denote the  $(a+1)$ -th point by  $x$  (see (1) of Fig. 3). We remove these  $a+1$  points above  $x$  from  $S$ , and divide the rectangle containing the remaining points of  $S$  into two rectangles by a vertical line so that the left rectangle contains  $b$  points and the right rectangle contains  $c$  points. Let  $y$  be the left most point in the left rectangle and  $z$  be the right most point in the right rectangle (see (1) of Fig. 3). Then there exist three non-crossing geometric Hamilton paths that start at  $x, y$  and  $z$ , respectively, and have order  $a+1, b$  and  $c$ , respectively. Moreover there are two  $L$ -line segments joining  $x$  to  $y$  and  $x$  to  $z$ , respectively as shown in (1) or (2) of Fig. 3 according to the situation of  $x$  and  $y$ . Hence the spider  $T$  is embedded on  $S$  without crossings.



**Fig. 3.** (1) A non-crossing geometric embedding of a spider  $T$  with trunk degree 3 on  $S$ , where  $y$  is to the left of  $x$ . (2) A non-crossing geometric embedding of a spider  $T$  on  $S$ , where  $y$  is to the right of  $x$ .

We next consider the case where the trunk  $u$  has degree four. Then  $T - u$  consists of four paths, and their orders are denoted by  $a, b, c, d$ , respectively. Sweep the points of  $S$  from left to right by vertical line, and denote the  $a$ -th point by  $p$  and the vertical line passing  $p$  by  $l_1$ . Next sweep the points of  $S$  from right to left by vertical line, and denote the  $b$ -th point by  $q$  and the vertical line passing  $q$  by  $l_2$  (see Fig. 4). We remove these  $a + b$  points from  $S$ , and sweep the remaining points of  $S$  between  $l_1$  and  $l_2$  from top to bottom by horizontal line, and denote the  $c$ -th point, the  $(c + 1)$ -th point and the  $(c + 2)$ -th point by  $y, x$ , and  $z$ , respectively (see Fig. 4). We divide the points between  $l_1$  and  $l_2$  into two sets and one point  $\{x\}$  by two horizontal lines passing through  $y$  and  $z$ , respectively. Then each of the four regions contains a geometric Hamilton path starting at  $p, q, y$  or  $z$ . By joining  $x$  to  $p, q, y$  and  $z$  by  $L$ -line segments, we can obtain the desired non-crossing geometric embedding of  $T$ .  $\square$



**Fig. 4.** A non-crossing geometric embedding of a spider  $T$  with trunk degree 4

## 4 Geometric Embedding of 3-Trees

In this section we give a partial solution to the following our conjecture.

*Conjecture 1.* Let  $T$  be a 3-tree and  $S$  be a set of  $|T|$  points in the plane lattice in general position. Then  $T$  can be geometrically embedded on  $S$  without crossings.

Our result is the following. Namely, we shall prove that if the vertices of  $T$  with degree 3 is contained in a path of  $T$ , then the conjecture holds.

**Theorem 3.** *Let  $T$  be a 3-tree such that all the vertices of degree 3 is contained in a path of  $T$ , and let  $S$  be a set of  $|T|$  points in the plane lattice in general position. Then  $T$  can be geometrically embedded on  $S$  without crossings.*

In order to prove Theorem 3, we need some definitions and notation. We denote by  $V_3(T)$  the set of vertices of  $T$  with degree 3. Let  $T$  be a 3-tree such that  $V_3(T)$  contains at least two vertices and is contained in a path of  $T$ , and let  $Q$  be the shortest path of  $T$  that passes through all the vertices of  $V_3(T)$ . Let  $w_1$  and  $w_2$  be the two end-vertices of  $Q$ , (see Fig. 5). A vertex  $v$  of  $T$  is said to be *good* if  $v \notin V(Q)$  and  $\{v\} \cup V_3(G)$  is included in a path of  $T$ , in particular,  $\deg_T(v) \leq 2$  (see Fig. 5). If  $T$  has exactly one vertex with degree 3, then all the other vertices are good.

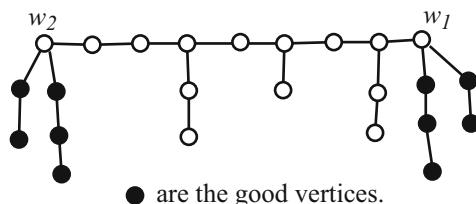
Theorem 3 follows from Theorem 1 if  $V_3(T)$  is empty, and otherwise it follows from the following proposition.

**Proposition 1.** *Let  $T$  be a 3-tree such that  $V_3(T)$  is non-empty and is contained in a path of  $T$ , and let  $S$  be a set of  $|T|$  points in the plane lattice in general position. Then for every good vertex  $v$  of  $T$  and for every point  $x$  of  $S$  on the edge of  $\text{rect}(S)$ , there exists a non-crossing geometric embedding  $\psi : T \rightarrow S$  that satisfies  $\psi(v) = x$ .*

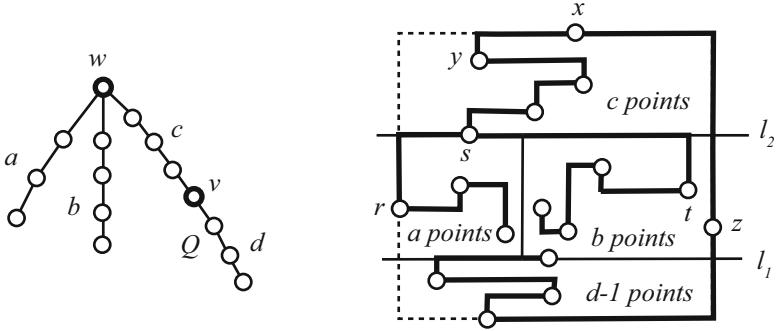
*Proof.* We shall prove the proposition by induction on  $|T|$ . By symmetry, we may assume that the point  $x$  lies on the top edge of  $\text{rect}(S)$ . We consider some cases.

*Case 1.*  $|V_3(T)| = 1$ .

Let  $w$  be the unique vertex of  $T$  with degree 3, and denote the orders of paths of  $T - w$  by  $a, b$  and  $c + d + 1$ , where the path  $Q$  of  $T - w$  containing  $v$  has order



**Fig. 5.** A 3-tree of Theorem 3 and its good vertices



**Fig. 6.** A 3-tree having one vertex of degree 3, and its non-crossing geometric embeddings

$c + d + 1$ , and the two paths of  $Q - v$  have orders  $c$  and  $d$ , respectively. Here we assume that  $x$  is not a corner point of  $\text{rect}(S)$  and  $c, d \geq 1$ . Notice that if  $x$  lies on the corner of  $\text{rect}(S)$ ,  $c = 0$  or  $d = 0$ , then by the almost the same argument given below, we can prove the the proposition.

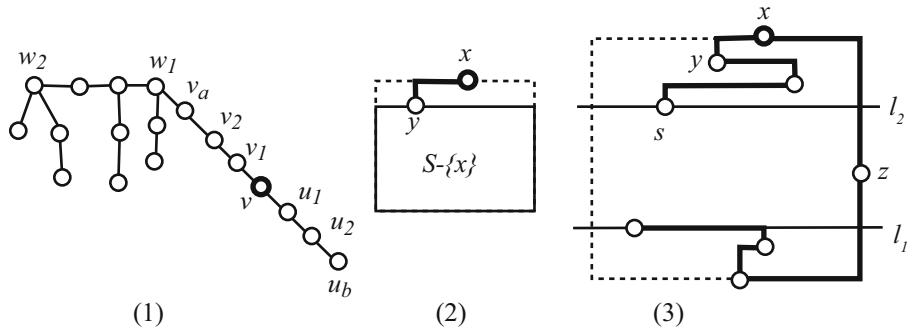
Let  $y$  be the highest point of  $S - x$ . By symmetry, we may assume that  $y$  is to the left of  $x$  (see Fig. 6). Let  $z$  be the point of  $S$  on the right edge of  $\text{rect}(S)$ , and let  $p$  be the point of  $S$  on the bottom edge of  $\text{rect}(S)$ . Sweep  $S - \{z\}$  from bottom to top by horizontal line, and take  $d - 1$  points and denote the horizontal line passing through the  $(d - 1)$ -th point by  $l_1$ . Sweep  $S - \{z\}$  from top to bottom by horizontal line, and take the  $(c + 1)$ -th point, say  $s$ , and denote the horizontal line passing through  $s$  by  $l_2$ . Divide the open rectangle between  $l_1$  and  $l_2$  into two open rectangles by vertical line segment so that the left open rectangle and the right open rectangle except  $z$  contain  $a$  and  $b$  points, respectively. Let us denote the left most point in the left rectangle by  $r$  and the right most point in the right rectangle by  $t$ . Then there are three Hamilton paths starting at  $r$ ,  $t$  and  $s$  as shown in Fig. 6, where a Hamilton path starting at  $s$  passes through  $y$ ,  $x$ ,  $z$ , and  $d - 1$  points below  $l_1$  including  $p$ . By joining  $s$  to  $r$  and  $t$  by  $L$ -line segments, we can obtain the desired non-crossing geometric embedding of  $T$  on  $S$ .

*Case 2.*  $|V_3(T)| \geq 2$  and  $\deg_T(v) = 1$ .

Let  $v_1$  be the vertex adjacent to  $v$ , and let  $y$  be the highest point of  $S - x$ . By induction,  $T - v$  can be embedded on  $S - x$  so that  $\psi(v_1) = y$  since  $v_1$  is a good vertex of  $T - v$ . Note that even if  $v_1 = w_i$  for some  $i \in \{1, 2\}$ ,  $v_1$  becomes a good vertex of  $T - v$ . By adding an  $L$ -line segment joining  $y$  to  $x$  to the embedding  $\psi$  of  $T - v$  on  $S - x$ , we obtain the desired embedding of  $T$  on  $S$  (see (2) if Fig. 7).

*Case 3.*  $|V_3(T)| \geq 2$  and  $\deg_T(v) = 2$ .

By symmetry, we may assume that there is a path that connects  $w_1$  and  $v$  and does not pass through  $w_2$ . Let  $(w_1, v_{a-1}, \dots, v_1, v, u_1, u_2, \dots, u_b)$  be the unique path in  $T$  that starts at  $w_1$ , passes through  $v$  and ends at end-vertex  $u_b$  of  $T$



**Fig. 7.** (1) A tree  $T$ ; (2) A geometric embedding of  $T$  in the case of  $\deg_T(v) = 1$ ; (3) A geometric embedding of  $T$  in the case where  $\deg_G(v) = 2$  and  $x$  is not a corner point.

(see (1) of Fig 7). Let  $y_1$  be the highest point of  $S - x$ . By symmetry we may assume that  $y_1$  is to the left of  $x$ .

We first assume that  $x$  is not a right-top corner point. Let  $z$  be the point of  $S$  on the right edge of  $\text{rect}(S)$ . Sweep  $S - \{z\}$  from bottom to top by horizontal line, and take  $b - 1$  points and denote the horizontal line passing through the  $(b - 1)$ -th point by  $l_1$ . Sweep  $S - \{z\}$  from top to bottom by horizontal line, and take the  $(a + 1)$ -th point, say  $s$ , and denote the horizontal line passing through  $s$  by  $l_2$ .

We embed the path  $(w_1, v_{a-1}, \dots, v_1, v, u_1, u_2, \dots, u_b)$  on the set of points of  $S$  that are above  $l_2$ , below  $l_1$  or  $z$  (see (3) of Fig 7). By induction and by the fact that  $w_1$  is a good vertex of a tree  $T^* = T - \{v_{a-1}, \dots, v_1, v, u_1, u_2, \dots, u_b\}$ ,  $T^*$  can be embedded on the set of points of  $S$  between  $l_2$  and  $l_1$  except  $z$  so that  $w_1$  is mapped to  $s$ . By combining these two embeddings, we can obtain the desired embedding of  $T$  on  $S$ .

We next assume that  $x$  is a right-top corner point. In this case, the point of  $S$  on the right edge of  $\text{rect}(S)$  is  $x$ . Then, sweep  $S$  from bottom to top by horizontal line, and take  $b$  points and denote the horizontal line passing through the  $b$ -th point by  $l_1$ . Latter, by applying the same argument give above, we can obtain the desired geometric embedding of  $T$  on  $S$ .

## References

1. Kaneko, A., Kano, M.: Discrete geometry on red and blue points in the plane — A survey. In: Discrete and Computational Geometry. Algorithms and Combinatorics, vol. 25, pp. 551–570. Springer (2003)
2. Kano, M., Suzuki, K.: Discrete geometry on red and blue points in the plane lattice (preprint)
3. Battista, G.D., Eardes, P., Tamassia, R., Tollis, I.G.: Graph drawing. Printice-Hall (1999)

4. Katz, B., Krug, M., Rutter, I., Wolff, A.: Manhattan-Geodesic Embedding of Planar Graphs. In: Eppstein, D., Gansner, E.R. (eds.) GD 2009. LNCS, vol. 5849, pp. 207–218. Springer, Heidelberg (2010)
5. Nishizeki, T., Rahman, M.S.: Planar graph drawing. World Scientific (2004)
6. Pach, J. (ed.): Towards a theory of geometric graphs. Contemporary Mathematics, vol. 342. AMS (2004)
7. Raghavan, R., Cohoon, J., Sahni, S.: Single bend wiring. *J. of Algorithms* 7, 232–257 (1986)

# Author Index

- Abellanas, Manuel 210  
Aichholzer, Oswin 1, 98  
Akiyama, Jin 14  
Alegría-Galicia, Carlos 226  
Aloupis, Greg 258  
Angelini, Patrizio 200  
Bajuelos, Antonio L. 210  
Bose, Prosenjit 29  
Cáceres, José 166  
Canales, Santiago 210  
Cetina, Mario 98  
Claverol, Mercè 210  
Cortés, Carmen 166  
Demaine, Erik D. 185, 258  
Demaine, Martin L. 258  
Díaz-Báñez, José Miguel 109, 126, 155  
Di Battista, Giuseppe 200  
Didimo, Walter 200  
Dujmović, Vida 258  
Fabila-Monroy, Ruy 79, 98, 109, 220  
Frati, Fabrizio 200  
García, Alfredo 249  
Garduño, Tzolkín 226  
Grima, Clara Isabel 166  
Hachimori, Masahiro 166  
Hackl, Thomas 1  
Hernández, Gregorio 210  
Hong, Seok-Hee 200  
Huemer, Clemens 220  
Iacono, John 258  
Itoh, Jin-ichi 85  
Iwerks, Justin 146  
Kano, Mikio 274  
Kaufmann, Michael 200  
Korman, Matias 126, 155  
Leaños, Jesús 98  
Liotta, Giuseppe 200  
Lubiw, Anna 185, 200  
Márquez, Alberto 166  
Matos, Inês 210  
Mészáros, Viola 236  
Mitchell, Joseph S.B. 146  
Mukae, Raiji 166  
Nakamoto, Atsuhiro 166  
Nara, Chie 85  
Negami, Seiya 166  
Omaña-Pulido, Elsa 119  
O'Rourke, Joseph 65  
Pach, János 45  
Pérez-Lantero, Pablo 109, 126, 155  
Radoičić, Radoš 45  
Rivera-Campo, Eduardo 119  
Robles, Rafael 166  
Rosas-Navarrete, Areli 226  
Sakai, Toshinori 175  
Salazar, Gelasio 98  
Sato, Ikuro 14  
Seara, Carlos 226  
Seong, Hyunwoo 14  
Souvaine, Diane L. 138  
Suzuki, Kazuhiro 274  
Tóth, Csaba D. 54, 138  
Tóth, Géza 45  
Toussaint, Godfried T. 54  
Urrutia, Jorge 98, 175, 226  
Valenzuela, Jesús 166  
Ventura, Inmaculada 126, 155  
Verdonschot, Sander 29  
Vilcu, Costin 85  
Vogtenhuber, Birgit 1  
Winslow, Andrew 54, 138  
Wood, David R. 79