DISS. ETH No. 14628, 2002

# Planar subdivisions

A dissertation submitted to the
Swiss Federal Institute of Technology, ETH Zürich
for the degree of Doctor of Technical Sciences

presented by
Csaba Dávid Tóth
M.Sc.Math. Eötvös Loránd University, Budapest
DEA MDFI. Université de la Méditerranée, Marseille
born in 1974, citizen of Hungary

accepted on the recommendation of
Prof. Dr. Emo Welzl, ETHZ, examiner
Prof. Dr. Günter Rote, FU Berlin, co-examiner

# Acknowledgment

I thank Frank Hoffmann, Gyula Y. Katona, Joseph O'Rourke, János Pach, Günter Rote, Bettina Speckmann, László Szabó, Géza Tóth, and Emo Welzl for pointing me to the problems discussed in this work, for many useful remarks and suggestions on the presentation of the results, and for their continuous encouragement.

Let me also thank all the colleagues at the Budapest University of Technology, at the ETH Zürich, and at the Free University Berlin for the friendly atmosphere and support they provided during the past few years.

I'm grateful to the joint Berlin-Zürich graduate program "Combinatorics, Geometry, and Computation", financed by the German Science Foundation (DFG) and the ETH Zürich, for the tranquil but stimulating conditions allowing me to complete this thesis.

# Summary

Subdivisions of the real Euclidean space and, specifically, planar subdivisions represent a fundamental tool in combinatorial and computational geometry. Many "divide and conquer" techniques exploit geometric structure through use of adequate subdivisions. Applications in range searching, ray tracing, robot motion planning, incremental algorithms, incidence problems, surface approximation algorithms, etc. reduce their running time and space complexity by working on a small number of cells of bounded complexity instead of the complete geometric structure.

This thesis does not intend to give an overview of planar subdivisions, nor of their internal hierarchy. Rather, it exhibits some recent progress on three different, though related, types of planar subdivisions, which are discussed separately in the three chapters to follow.

The three topics can be placed in a common framework of visibility. Two points $p$ and $q$ are mutually visible if the open line segment $pq$, connecting the two points, does not interfere (i.e., intersect or cross) any of the given obstacle.

Chapter 1 considers subdivisions of simple polygonal domains (or preferably, *art galleries*). Our goal is to cover any art gallery with a small number of polygonal domains such that each is visible from one point under a fixed angle $\alpha \in (0^\circ, 360^\circ]$ (called $\alpha$-guard, for obvious reasons) within the art gallery. In such a covering, we say that the guards monitor the art gallery. Guards may be required to be stationed at vertices of the polygon, at most one at each vertex. The optimization problem, that is, finding the minimal number of guards for a given art gallery is known to be NP-hard. There is, however, a minimal number of $\alpha$-guards who can monitor any possible art gallery with a fixed number $n$ of vertices. As a function of $n$, we investigate this minimum and show recent results and, sometimes, the optimal solutions.

In Chapter 2, we are given $n$ disjoint obstacles in the plane. Our task is to illuminate either all boundary points of the obstacles, or all points not covered by the obstacles using point-like light sources. Finding a minimal set of light sources, or even finding the minimal number of necessary light sources is NP-hard. The known reductions to the 3-SAT or the set-cover cover problem are analogous to those for art gallery problems. Here, again, we present a worst case analysis, showing that a certain number of light sources –as a function of the number $n$ of obstacles – suffice for any set of disjoint obstacles of one type (e.g. triangles, line segments). In certain cases, we are able to find the worst case optimal solution. Interestingly enough, this simple illumination problem splits into several slightly different versions once the obstacles in question are

line segments or curves in the plane. One version was considered and solved by J. O'Rourke, we address or show some progress on the others (and solve three of them).

Chapter 3 deals with a given set of disjoint line segments in the plane, too. The plane is partitioned by the *binary space partitioning*, or glass cutting algorithm along line segments. The partitioning segments are not necessarily *visibility segments*; i.e., they may cross some given segments but each such crossing costs one unit. We do not aim at finding the partitioning with minimal cost, but give new upper and lower bounds on the worst case cost of a partitioning as a function of the number $n$ of given line segments and the number $k$ of distinct orientations of the segments.

# Sommaire

Les subdivisions de l'espace réel Euclidien et en particulier les subdivisions du plan constituent des outils fondamentaux de la géométrie combinatoire et algorithmique. Plusieurs techniques de type "diviser et construire" exploitent la structure géométrique en utilisant une subdivision approprié. Les applications dans les domaines du lancé de rayon, de la planification de trajectoires pour un robot, des algorithmes incrémentaux, des problèmes d'incidences, des algorithmes pour l'approximation de surface etc. peuvent réduire leur temps d'exécution et l'espace du calcul en considèrant un nombre limité de regions de complexité bornée au lieu de la structure géométrique complète.

Ce rapport de thèse ne souhaite pas offrir un vue globale ni sur les subdivisions variée du plan ni sur leur hiérarchie interne. Par contre, il démontre quelques progrès récents sur trois types de subdivisions du plan differents mais liés, qui sont considerés un par un dans les trois chapitres à suivre.

Les trois sujets sont liés par le mot clé commun de visibilité. Deux points $p$ et $q$ sont mutuellement visibles si le segment ouvert $pq$, qui relie les deux points, n'interfère pas (soit il ne coupe pas, soit il ne traverse pas) avec des obstacles donnés.

Chapitre 1 considère des subdivisions de polygones simples (ou bien, de *galleries d'art*). Notre objectif est de couvrir toute gallerie d'art par un nombre limité de polygones tel que chaqu'un soit visible d'un point sous un angle fixé $\alpha \in (0°, 360°]$ (ce qui est nommé $\alpha$-garde pour des raisons évidents) dans la gallerie. Lorsqu'une telle couverture éxiste, on dit que les gardes monitorent la gallerie d'art. Les gardes peuvent être contrainetes à se placer aux sommets du polygone, au plus une garde par sommet. Le problème d'optimisation, i.e. detérminer le nombre minimal de gardes pour un gallerie d'art donnée, est connu d'être NP-difficile. Il éxiste, par contre, un nombre minimal de $\alpha$-gardes qui sont capable de monitorer n'importe quelle gallerie d'art avec un nombre $n$ fixé de sommets. En fonction de $n$, nous étudions ce minimum et présentons des resultats récents, parfois les solutions optimales.

Dans le Chapitre 2, $n$ obstacles disjoints sont donnés dans le plan. Notre objective est d'illuminer soit les points de la frontières des obstacles, soit le complèmentaire de l'union des obstacles en utilisant des points comme source de lumière. Il est NP-difficile de trouver un ensemble minimum de sources de lumière, de même, de déterminer le nombre minimal de lumières nécessaires. La réduction au 3-SAT par le problème de couverture des ensemble est similaire à celle appliquée pour les galleries d'art. De nouvau, nous présentons une analyse du pire cas : un certain nombre de sources de lumière (en fonction du

nombre d'obstacles) est suffisant pour n'importe quel ensemble d'obstacles d'un type spécifié (e.g. triangles ou segments). Dans quelques cas, on est capable de trouver la solution optimale. Étonnamment ce problème simple d'illumination se divise en plusieurs differentes variantes si les obstacles considérés sont des segments ou des courbes planes. Une variante a été étudiée et resolue par J. O'Rourke, nous attaquons les autres (et en résolvons trois).

Le sujet du Chapitre 3 est, aussi, des ensembles de segments disjoint dans le plan. Le plan est partitionné par l'algorithme de *partitionnement binaire d'espace* le long de segments. Les lignes de partitionnements ne sont pas nécessairement des segments de visibilité, elles peuvent couper les segments donnés mais chaque coupe coûte une unité. On ne souhaite pas trouver le partitionnement de coût minimal, mais on arrive à donner des bornes supérieures et inférieures pour le coût du partitionnement dans le pire cas en fonction du nombre $n$ de segments donnés et du nombre $k$ d'orientations des segments.

# Contents

**Art gallery problem for simple polygons** **1**

1.1 Introduction . . . . . 1

1.2 Dense polygons . . . . . 5

1.2.1 Dense decomposition . . . . . 6

1.2.2 Notation and star polygons . . . . . 7

1.3 Simple polygons with $45°$-guards . . . . . 8

1.3.1 Proof of Lemma 1.3 . . . . . 9

1.4 Simple polygons with point $\pi$-guards . . . . . 12

1.4.1 Proof of Theorem 1.6 . . . . . 13

1.4.2 Proof of Lemma 1.4 . . . . . 16

1.4.3 Proof of Lemma 1.5 . . . . . 18

1.4.4 An example . . . . . 26

1.5 Lower bounds on the number of $\alpha$-guards . . . . . 28

1.5.1 Lower bound for $\alpha < 180°$ . . . . . 28

1.5.2 Lower bound for $\alpha < 90°$ . . . . . 30

1.5.3 Lower bounds for $\alpha < 60°$ . . . . . 33

1.6 Simple polygons with vertex $\pi$-guards . . . . . 35

1.6.1 Basics on pseudo-triangles . . . . . 35

1.6.2 Edge-aligned $\pi$-guards . . . . . 38

1.6.3 $\pi$-guards in one pseudo-triangle . . . . . 41

1.6.4 Upper bound for the general model . . . . . 42

1.6.5 Proof of Lemma 1.8 . . . . . 45

**Illumination in the presence of obstacles in the plane** **51**
2.1 Introduction . . . . . 51
2.2 Convex partitionings and dual graphs . . . . . 55
2.2.1 Line segments in the plane . . . . . 55
2.2.2 Polygons with holes . . . . . 57
2.3 Illuminating line segments . . . . . 60
2.3.1 $P_3$-packing of graph $K$ . . . . . 60
2.3.2 Weak illumination of the free space . . . . . 63
2.3.3 Strong illumination of the free space and illuminating both sides of segments . . . . . 64
2.3.4 Illuminating segments . . . . . 65
2.4 The structure of graph $H$ and $K$ . . . . . 70
2.4.1 The dual graph $K$ has a $\{P_{\geq 3}\}$-factor . . . . . 70
2.4.2 Basic properties of $H$ and $K$ . . . . . 73
2.5 Illuminating triangles . . . . . 75
2.5.1 Three vertices on every cell . . . . . 77
2.5.2 Proof of Lemma 2.8 . . . . . 80

**Binary space partitions for line segments** **85**
3.1 Introduction . . . . . 85
3.2 Lower bound construction . . . . . 87
3.2.1 Auxiliary construction . . . . . 88
3.2.2 Main construction . . . . . 91
3.2.3 Autopartitions of $L^k$ . . . . . 94
3.2.4 Binary plane partitions of $L^k$ . . . . . 96
3.3 Binary space partition with few distinct directions . . . . . 101
3.3.1 Partition algorithm . . . . . 102
3.3.2 Cycles and convex cycles . . . . . 103
3.3.3 BSP for boundary segments . . . . . 107
3.3.4 Autopartition for boundary segments . . . . . 114

**Bibliography** **117**

**Curriculum Vitae** **125**

# Art gallery problem

## 1.1 Introduction

In a typical art gallery problem, an art gallery is represented by a simple closed polygonal domain (*polygon*, for short) $P$ and we must allocate points representing guards in $P$ under various conditions so that any point of $P$ be visible by at least one of the guards [33]. V. Chvátal [11] proved in 1975 the first result in this field.

**Theorem 1.1.** *[11] If $P$ is an art gallery with $n$ sides, then one can place $\lfloor \frac{n}{3} \rfloor$ guards at vertices of $P$ such that every point of $P$ is visible from at least one of them; this number is the best possible.*

A point $p \in P$ is *visible* from a point $q$ if the open line segment $pq$ lies completely in $P$. In 1987, S. Fisk [21] gave an elegant, "book" proof [2] of this theorem using only triangulations of simple polygons and vertex 3-colorings.

Ever since, many variants and generalizations of Chvátal's watchmen theorem popped up. Point guards, vertex guards, and edge guards are distinguished; moving guards and shortest paths of moving guards were considered; the art gallery can refer to an orthogonal polygon, a simple polygon or a polygon with holes. Note also that art gallery problems are often posed within the terminology of illumination where a set of *light sources* or *floodlights* should *illuminate* a certain region, instead of *guards* who *monitor* the gallery. See [58] or [47] for an overview on art gallery and illumination problems.

In this chapter, we concentrate on one generalization of Chvátal's problem, that of guards of uniform $\alpha \in (0°, 360°]$ range of vision. J. Urrutia posed the following question: What is the minimum number $f(n)$ of guards needed to monitor any art gallery of $n$ sides if the guards are to be stationed at fixed

points and their range of vision is $180°$. In 1992, H. Bunting and D. Larman [31] showed that $f(n) \leq \lfloor \frac{4}{9}(n+\frac{1}{4}) \rfloor$. Then Gy. Csizmadia and G. Tóth [12] proved $f(n) \leq \lceil \frac{2}{5}(n-3) \rceil$ for $n > 3$. I settled the problem by showing $f(n) \leq \lfloor \frac{n}{3} \rfloor$, which is the best possible upper bound [48].

A guard of range of vision $\alpha$ (or *$\alpha$-guard* as a shorthand) is a pair of a point $p \in P$ and an angular domain $H_p$ of angle $\alpha$ with its apex at $p$. A point $q \in P$ is monitored by $(p, H_p)$, if $pq \subset P \cap H_p$. The *art gallery* $P$ is *monitored* if any $q \in P$ is monitored by at least one of the guards. (Note that the term *$\alpha$-floodlight* is also used in the literature of illumination problems with the same content; see [8], [47], and [58] for floodlight illumination problems).

**Problem 1.1.** *Given $\alpha \in (0°, 360°]$, what is the minimal number of $\alpha$-guards required to monitor any art gallery with $n$ vertices.*

Denote by $f(n, \alpha)$ the maximal number of $\alpha$-guards required to monitor an art gallery with $n$ vertices for a fixed $\alpha \in (0°, 360°]$. If we are not interested in classes of $k$-gons for any specific $k$, we can formulate the problem in an asymptotic version. Let

$$g(\alpha) = \limsup_{n \to \infty} \frac{f(n, \alpha)}{n}.$$

**Problem 1.2.** *Determine $g(\alpha)$ for all $\alpha \in (0°, 360°]$.*

Clearly $g(\alpha)$ is monotone decreasing, and $\lim_{\alpha \to 0°} g(\alpha) = \infty$. In this chapter, we exhibit some recent improvements on Problem 1.2 for different values of $\alpha \in (0°, 360°]$. In the intervals $\alpha \in [45°, 60°)$ and $\alpha \in [90°, 360°]$, we determine the value of $g(\alpha)$; only $g(360°) = 1/3$ was known previously. In other intervals, we have non-matching upper and lower bounds.

The variant considering vertex guards goes back to Urrutia [19], too. What is the minimal number of vertex $180°$-guards that can collectively monitor any art gallery $P$ with $n$ vertices. A vertex $\alpha$-guard is an $\alpha$-guard whose apex is at a vertex of $P$. There may be at most one vertex $\alpha$-guard at each vertex.

The notion of vertex $\alpha$-guard makes sense for any $\alpha \in (0°, 360°]$. However the case $\alpha < 180°$ is answered under the condition that there may be at most one vertex guard at each vertex. Estivill-Castro et al. [19] proved that for any angle $\alpha < 180°$, there exist convex polygons $P_n$ with $n \geq n_\alpha$ vertices such that $n$ vertex $\alpha$-guards cannot illuminate $P_n$.

A $180°$-guards (or $\pi$-guard, for short) at a reflex vertex cannot monitor the complete angular domain bounded by the two incident edges. Restrictions on the orientation of the halfplane $H_v$, where $v$ is a reflex vertex, lead to three

variants of the problem. We may require that at every reflex vertex $v$, the possible guard $(v, H_v)$ be

- *inward-facing*, i.e. the two sides of $P$ incident to $v$ are disjoint from the interior of $H_v$;
- *edge-aligned*, i.e. inward-facing and the boundary of $H_v$ is collinear with one of the edges incident to $v$;
- *general*, i.e. we do not restrict the orientation of $H_v$.

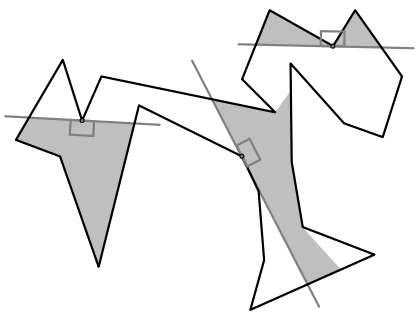

**Figure 1.1:** *A polygon with an inward-facing (left), an edge-aligned (middle), and an outward-facing (right) vertex $\pi$-guard.*

Urrutia [59] found a family of polygons that require $\lfloor 5(n-1)/8 \rfloor$ inward-facing or edge-aligned vertex $\pi$-guards, but showed only the sufficiency of $n-2$ vertex $\pi$-guards [19]. In the restricted models of *inward-facing* or *edge-aligned* $\pi$-guards, this is still the best previously known upper bound [35]. For *general* $\pi$-guards the upper bound was improved to $\lfloor (3n-5)/4 \rfloor$ using a so-called *dense decomposition* of the simple polygon [51]; later Brumberg et al. [9] claimed a simple proof for an upper bound of $\lfloor 5n/6 \rfloor$. The best known lower bound of $\lfloor (n-1)/2 \rfloor$ for the general case follows from a family of polygons called *monotone mountains* [34] (polygons with two chains of reflex vertices). We were able to improve the upper bounds for all three models.

**Theorem 1.2.** *$\lfloor n/2 \rfloor$ vertex $\pi$-guards can always monitor a simple polygon with $n$ vertices in the general model.*

**Theorem 1.3.** *$\lfloor (2n-k)/3 \rfloor$ edge-aligned vertex $\pi$-guards can always monitor a simple polygon with $n$ vertices, $k$ of which are convex.*

The latter bound is tight for the family of polygons of Urrutia, for which the number of convex vertices is $k = (n-1)/8+2$. Although our bound improves upon the best previously known upper bound of $n-2$, it is not always tight.

The results of this chapter are organized as follows. Section 1.2 introduces a new decomposition of simple polygons based on triangulations. This decomposition is then used in Section 1.3 to obtain a quadrilateral decomposition of simple polygons with an even number of vertices, which is further used to prove an upper bound on $g(45°)$. Section 1.4 proves that $g(180°) \leq 1/3$.

In Section 1.5, we determine lower bounds for $g(\alpha)$ for all $\alpha < 180°$. We expect that our lower bounds give the the exact values of $g(\alpha)$. In each case, the proof is based on explicit constructions of polygons that require at least a certain number of $\alpha$-guards. In the proofs of the theorems for $\alpha \in [90°, 180°)$ and $\alpha \in [60°, 90°)$, we use a combinatorial argument. The number of necessary guards is determined by placing points in the polygons so that no pair of points can be monitored by one $\alpha$-guard.

For all $\alpha < 60°$, our lower bounds are based on one family of polygon built of equilateral triangles. The main tool relies on a result of Zs. Páles [37] based upon metric theory and analysis.

This chapter leave many questions open. The exact value of $g(\alpha)$ for $\alpha \in (0°, 45°)$ and $\alpha \in [60°, 90°)$ is unknown, and we expect that the lower bounds given in Section 1.5 to be best possible.

For the study of $g(\alpha)$, we did not aim at maximizing the number of guards located at vertices. This could be a topic for further investigations. As a smallest example, there is a pentagon that cannot be monitored by $\lfloor \frac{5}{3} \rfloor = 1$ vertex $180°$-guard, but one guard at the boundary of the pentagon suffices (Figure 1.9(a)).

We show in Theorem 1.6 that $\lfloor \frac{n}{3} \rfloor$ $180°$-guards can illuminate an art gallery with $n$ vertices, but this theorem does not decide if $\lfloor \frac{n}{3} \rfloor$ $180°$-guards located exclusively on the boundary of the art gallery are sufficient or not. We conjecture that there is a positive answer to this question. But our method does place guards at interior points (see Subsection 1.4.4 for an example).

The proof of Theorem 1.6 can easily be turned into an $O(n^2)$ algorithm which locates the positions of at most $\lfloor \frac{n}{3} \rfloor$ $180°$-guards; but the optimal time complexity of such an algorithm is not known.

Determining the minimal number of vertex $180°$-guards necessary to monitor any art gallery with $n$ vertices remains an open problem. Finally, nothing is known about the version of vertex $\alpha$-guard placement if several guards may be located at a common vertex. In this relaxed model of vertex $\alpha$-guards only

the easy case of convex polygons is settled.

## 1.2 Dense polygons

The *dense decomposition* of simple polygons, introduced here, is a nice concept and will be used in to obtain a quadrilateral decomposition in Section 1.3. The notion of a *triangulation* of simple polygons and that of a *dual graph* will be fundamental in every part of this chapter. See also [45] for basic definitions.

Let $P$ be a simple polygon with $n$ vertices and let $T$ be a set of triangles in the plane. $T$ is the triangulation of $P$, if $P = \bigcup T$, the triangles of $T$ are pairwise non-overlapping, and the vertices of the triangles are vertices of $P$. It is known that every simple polygon has a triangulation. Every triangulation consists of exactly $n-2$ triangles, although the triangulation is not necessarily unique. We define the graph $G(T)$ on a triangulation $T$. The nodes of the graph correspond to the elements of $T$, two nodes are adjacent if and only if the corresponding triangles have a common side. ($G(T)$ is also known as the *dual graph* of the triangulation [45].) $G(T)$ is a tree on $n-2$ nodes, the maximal degree in $G(T)$ is three, since a triangle $t \in T$ may have a common side with at most three other triangles of $T$.

In our terminology, *polygons* have *vertices* and *sides*, while *graphs* have *nodes* and *edges*. If there is no danger of confusion, we will make no distinction between a *node* of $G(T)$ and the corresponding triangle of $T$. A *dense polygon* can be defined as follows.

**Definition 1.1.**

- *A tree $G$ is* dense*, if each node of $G$ has degree one or three.*
- *A simple polygon $S$ is* dense *if $G(T)$ is dense for every triangulation $T$ of $S$.*

**Proposition 1.1.** *Every dense tree $G$ has an even number of nodes. If a dense tree $G$ has $2\ell$ nodes then it has exactly $\ell + 1$ leaves.*

*Proof.* Suppose that $G$ has $s$ leaves and $t$ nodes of degree 3. The number of edges is $(s + 3t)/2 = (s + t)/2 + t$, hence $s + t$ is even. $G$ is a tree, so the number of edges equals to the number of nodes minus one: $(s + 3t)/2 = s + t - 1$, that is, $s = (s + t)/2 + 1$. □

### 1.2.1 Dense decomposition

The *dense decomposition* $D$ of a simple polygon $P$ is a set of pairwise non-overlapping dense polygons such that $P = \bigcup D$ and the vertices of the dense polygons are vertices of $P$. We can define the tree $G(D)$ on a dense decomposition $D$ just like $G(T)$. The nodes of $G(D)$ correspond to the dense polygons of $D$, two nodes are adjacent if and only if the corresponding polygons have a common side.

The union of the triangulations of the elements of a dense decomposition is a triangulation $T$ of $P$. So $T$ contains an even number of triangles. Clearly, this is impossible if $P$ has an odd number of vertices. We can prove, however, the following lemma.

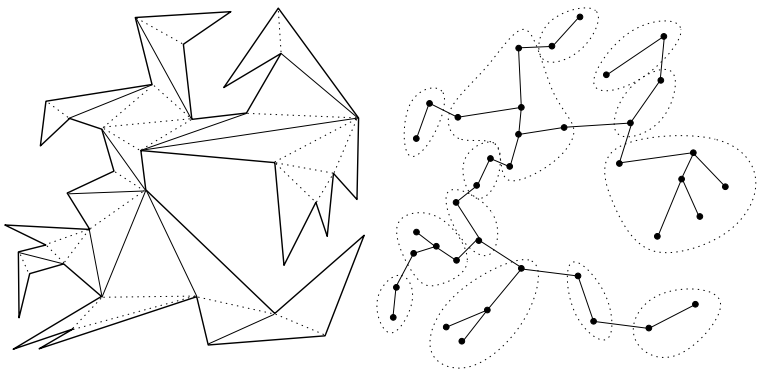

**Figure 1.2:** *Dense decomposition of a polygon $P$ and the dual graph $G(T)$.*

**Lemma 1.1.** *If $P$ is a simple polygon with an even number of vertices, then $P$ has a dense decomposition.*

*Proof.* We proceed by induction on the number of nodes of $G(T)$. Every quadrilateral is dense. If the polygon $P$ is dense, then the proof is complete.

If $P$ is not dense then there is a triangulation $T$ of $P$ such that there exists a node of degree two in $G(T)$. Consider the tree $G(T)$ as a rooted tree $(G(T), r)$ where an arbitrary leaf $r$ of $G(T)$ is chosen as the root. Let $v \in G(T)$ be a node of degree two such that no descendant of $v$ has degree two in $G(T)$. Let $S$ denote the subtree containing $v$ and all its descendants in $(G(T), r)$. According to Proposition 1.1, the subtree $S$ has an even number of nodes, hence

$G(T) \setminus S$ has an even number of nodes as well. The polygons corresponding to $S$ and $G(T) - S$ have dense decompositions by induction, and together they give a dense decomposition of $P$. □

### 1.2.2 Notation and star polygons

In the application of our dense decomposition, quadrilaterals always have special roles. We call *star-polygon* a dense polygon with at least 6 vertices. (Note that *star* polygons are not necessarily *star-shaped*, i.e., they cannot always be illuminated by one $2\pi$-guard.) Consider a star-polygon $S \in D$ and fix a triangulation $T_S$ in each $S \in D$. An *outer vertex* of $S$ is a vertex which belongs to exactly one triangle $t \in T_S$. All other vertices of a $S$ are called *inner*. Outer vertices are vertices of the leaves of $T_S$, so there are exactly $\ell + 1$ of them. Clearly, every second vertex of a dense polygon is outer and every outer vertex is convex.

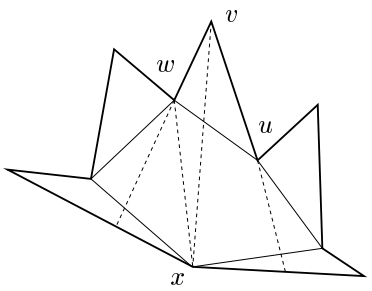


**Figure 1.3:** *This polygon is not dense, since $v$ is outer and $vx$ is a diagonal.*

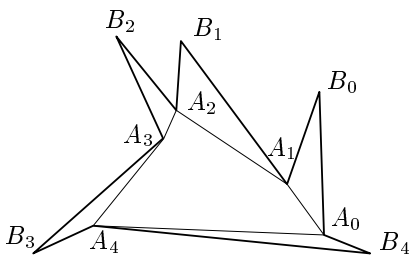


**Figure 1.4:** *Notation of vertices in a dense polygon.*

**Proposition 1.2.** *For two vertices $v$ and $x$ of a star-polygon $S$, $vx$ cannot be a diagonal of $S$ if $v$ is an outer vertex for a triangulation $T_S$.*

*Proof.* Let $u, v$, and $w$ be three consecutive vertices of $S$ such that $uvw$ is a triangle in a triangulation $T_S$ of $S$. Suppose that $vx$ is a diagonal of $S$. We may suppose w.l.o.g. that $ux$ is not a side of $S$, since $S$ has at least six vertices. First we state that there is a vertex $y$ such that both $uy$ and $vy$ are diagonals. If $y = x$ does not have this property, then let $y$ be the vertex in $uvx$ where the angle $\angle uvy$ is minimal. There is a triangulation $T'_S$ of $S$ such that $uvy \in T'_S$. In $G(T'_S)$, the node corresponding to $uvy$ has degree 2, a contradiction. □

**Corollary 1.4.** *If a vertex $v$ of a star polygon $S$ is outer for one triangulation of $S$, then it is outer for every triangulation of $S$. In other words: the $\ell + 1$ leaves of a star polygon $S$ with $\ell$ vertices correspond to the same $\ell+1$ triangles in any triangulation of $S$.*

Dense polygons have some more interesting properties which we do not use here. Among others, any dense polygon has at most one convex inner vertex.

## 1.3 Simple polygons with $45°$-guards

In this section, we consider exclusively guards of $45°$ range of vision. The upper bound $g(45°) \leq 1$ is given by the following theorem.

**Theorem 1.5.** *[56]*

$$f(n, 45°) \leq \begin{cases} n-1, & \text{if } n \text{ odd}, \\ n-2, & \text{if } n \text{ even}. \end{cases}$$

It is clear that a triangle can be illuminated by at most two guards. A triangle with an angle less than or equal to $45°$ can be illuminated by one guard. We give here the elementary proof for the case $n = 4$.

**Lemma 1.2.** *A quadrilateral can be illuminated by two $45°$-guards.*

*Proof.* Consider the quadrilateral $Q = ABCD$. If $Q$ is convex, then two guards placed at a vertex with angle less than or equal to $90°$ can illuminate $Q$.

$Q$ has at most one reflex angle. Assume w.l.o.g. that $\angle ABC > 180°$. If $\angle CDA \leq 90°$, then two guards at $D$ illuminates $Q$. If $\angle BCD \leq 45°$ and $\angle BAD \leq 45°$ then two guards at $C$ and $A$ illuminate the triangles $BCD$ and $BAD$ respectively.

Suppose that $\angle ABC > 180°$, $\angle CDA > 90°$, $\angle BCD > 45°$, and therefore $\angle BAD < 45°$. Let $E$ be the intersection of the ray $\overrightarrow{AB}$ and side $CD$. Now $\angle BAD + \angle CBE < 45°$ follows from the fact that $\angle ABC = 180° + \angle CBE$. Both $\angle BAD$ and $\angle CBE$ are smaller than $45°$. Two guards at $A$ and $B$ can illuminate the triangles $ADE$ and $CBE$ respectively. □

Our next lemma states a decomposition property of simple polygons.

**Lemma 1.3.** *Every simple polygon with $2(\ell + 1)$ vertices can be partitioned into $\ell$ quadrilaterals.*

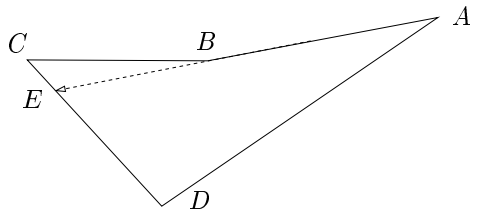


**Figure 1.5:** *A quadrilateral $ABCD$ where $\angle ABC > 180°$, $\angle CDA > 90°$, $\angle BCD > 45°$, and $\angle BAD < 45°$.*

The proof of Lemma 1.3 constitutes the subject of Subsection 1.3.1. We are ready now to establish Theorem 1.5.

*(of Theorem 1.5).* Consider a simple polygon $P$ with $n$ vertices. If $n$ is even, then by Lemma 1.3 it can be partitioned into $n/2 - 1$ quadrilaterals. Each quadrilateral can be illuminated, independently, by at most two guards. So $P$ can be illuminated by at most $n - 2$ guards.

If $n$ is odd then choose a triangle $t$ corresponding to a leaf of $G(T)$. Now $P \setminus t$, which has an even number of vertices, can be illuminated by $(n-1)-2$ guards and two guards can illuminate $t$. □

### 1.3.1 Proof of Lemma 1.3

Consider a simple polygon $P$ with $2(\ell + 1)$ vertices. The statement holds for $\ell = 1$, that is, if $P$ is a quadrilateral. We proceed by induction on $\ell$. Suppose that Lemma 1.3 holds for every $\ell' < \ell$.

Let $D$ be the dense decomposition of $P$. If $D$ has more than one element, then denote them by $S_1, S_2, \ldots, S_k \in D$, $k \geq 2$. If $S_i$ has $2(\ell_i + 1)$ vertices for every $i = 1, \ldots k$, then we know that $2(\ell + 1) = 2 + \sum_{i=1}^{k} 2\ell_i$, since the triangulation of each $S_i$ contains $2\ell_i$ triangles. By induction, each $S \in D$ can be partitioned into $\ell_i$ quadrilaterals. It follows that $P$ has a partition into $\sum_{i=1}^{k} \ell_i = \ell$ quadrilaterals.

Assume that $P = S$ is a star polygon $A_0B_0A_1B_1 \ldots A_\ell B_\ell$ in counter-clockwise order where the $B_i$'s are the outer vertices. Arithmetics on the indices of $A_0, \ldots, A_\ell$ and $B_0, \ldots, B_\ell$ is meant mod $\ell + 1$, that is, $\ell + 1$ stands for 0.

**Proposition 1.3.** *Assume that there is an index $i \in [0, \ell]$ such that $A_iO \subset P$ and $A_{i+2}O \subset P$ where $O$ is the intersection of the lines through $B_iA_i$ and $B_{i+1}A_{i+2}$, and either (i) $O \in \mathrm{int(P)}$ or (ii) $O \in B_iA_i$ or $O \in B_{i+1}A_{i+2}$. Then $P$ can be partitioned into $\ell$ quadrilaterals.*

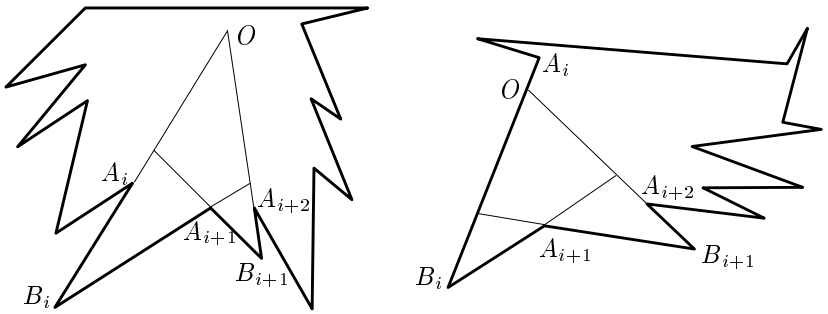


**Figure 1.6:** *The line segments $A_iO$ and $A_{i+2}O$ are in $S$. On the left $O \in \mathrm{int(P)}$, and on the right $O \in B_iA_i$.*

*Proof.* Polygon $A_{i+1}B_iOB_{i+1}$ is a quadrilateral. The simple polygon $S \setminus A_{i+1}B_iOB_{i+1}$ has $2\ell$ vertices, and by induction, it can be partitioned into $\ell - 1$ quadrilaterals (see Fig. 1.6.) □

Let the *core* of the star polygon $S$ be the simple $(\ell+1)$-gon $K = A_0A_1 \dots A_\ell$. We would like to determine which parts of the boundary of $K$ are visible from an outer vertex $B_i$. By Proposition 1.2, $A_jB_i$ is not a diagonal for any pair $i, j$. But the cone spanned by the rays $\overrightarrow{B_iA_i}$ and $\overrightarrow{B_iA_{i+1}}$ intersects $K$. More precisely, ray $\overrightarrow{B_iA_i}$ or ray $\overrightarrow{B_iA_{i+1}}$ enters into the interior of $K$. This implies that the only sections on the boundary of $K$ visible from $B_i$ are $A_iA_{i+1}$ and a part of $A_{j(i)}A_{j(i)+1}$ for some $j(i) \in \{0, .., \ell\} \setminus \{i\}$. Possibly, $A_{j(i)} = A_{i+1}$ or $A_{j(i)+1} = A_i$, but not both. Let $C_i$ (resp., $D_i$) be the point where ray $\overrightarrow{B_iA_i}$ (resp., $\overrightarrow{B_iA_{i+1}}$) hits $A_{j(i)}A_{j(i)+1}$, that is, $C_iD_i \subset A_{j(i)}A_{j(i)+1}$ is visible from $B_i$.

**Proposition 1.4.** *If Proposition 1.3 cannot be applied to $S$, then $j(i) = j(i+1)$ for every $i = 0, 1, \dots \ell$.*

*Proof.* By definition, $C_{i+1} \in A_{j(i+1)}A_{j(i+1)+1}$ and $D_i \in A_{j(i)}A_{j(i)+1}$. As the segments $A_{i+1}B_i$, $A_{i+1}B_{i+1}$, $A_{i+1}D_i$, and $A_{i+1}C_{i+1}$ are arranged around $A_{i+1}$ in this (counter-clockwise) order. The ray $\overrightarrow{B_{i+1}A_{i+1}}$ hits $C_{i+1}$

and $A_{i+1}C_{i+1}$ dissects $K$ into two parts. The ray $\overrightarrow{B_iA_{i+1}}$, which hits side $A_{j(i)}$, crosses the line through $A_{i+1}C_{1+1}$ at $A_{i+1}$. We deduce that the index $j(i)$ of the side of $K$ visible from $B_i$ is one of $i+1, i+2, \ldots, j(i+1)$.

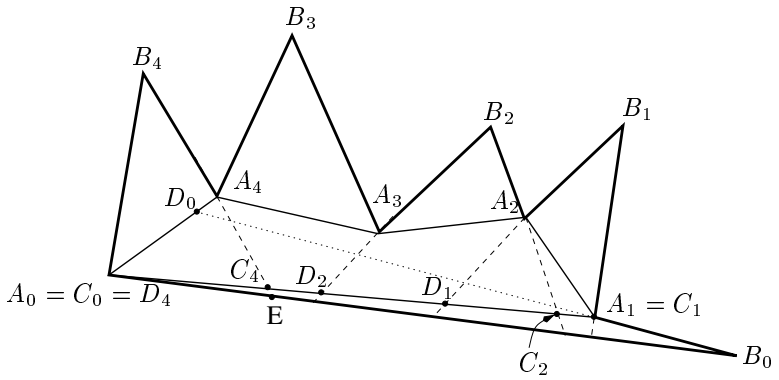


**Figure 1.7:** *Part of $A_0A_1$ is visible from $B_i$ for all $i \in \{1, .., \ell\}$.*

Similarly, $C_i \in A_{j(i)}A_{j(i)+1}$ and $D_{i+1} \in A_{j(i+1)}A_{j(i+1)+1}$. Let $O$ be the intersection of the lines $B_iA_i = B_iC_i$ and $B_{i+1}A_{i+2} = B_{i+1}D_{i+1}$. We may suppose that $A_iO, A_{i+2}O \not\subset S$ or $O \notin \text{int(P)}$, otherwise Proposition 1.3 could be applied. We deduce that $j(i) \neq i+2, i+3, \ldots, j(i+1)-1$.

If $j(i) = i+1$, then $D_i = A_{i+1}$ and $C_i \in A_{i+1}A_{i+2}$. (See Fig. 1.7 with $i = 4$ and $i+1 = 0$.) The ray $\overrightarrow{B_iA_i} = \overrightarrow{B_iC_i}$ intersects the triangle $A_{i+1}B_{i+1}A_{i+2}$, Denote by $E$ the intersection point of the ray $\overrightarrow{B_iA_i}$ and the polygonal curve $A_{i+1}B_{i+1} \cup B_{i+1}A_{i+2}$. If $E \in B_{i+1}A_{i+2}$ then Proposition 1.3 could be applied.

If $E \in B_{i+1}A_{i+1}$, then we argue that there is a vertex $v$ such that $B_{i+1}v$ is a diagonal, contradicting Proposition 1.2. If $A_i$ is visible from $B_{i+1}$, then $B_{i+1}A_i$ is a diagonal of $S$, and $v = A_i$. Otherwise, let $v$ be the vertex in $A_iEB_{i+1}$ for which the angle $\angle EB_{i+1}v$ is minimal. This is necessarily a reflex vertex of $S$, and therefore it is outer. $v \neq A_{i+2}$, because otherwise $\overrightarrow{B_iA_i}$ and $\overrightarrow{B_{i+1}A_{i+2}}$ would intersect in the interior of $S$ and, again, Proposition 1.3 could be applied (this is the case in Fig. 1.7 with $i = 4$ and $i+1 = 0$).

We conclude that $j(i) \neq i+1$, hence $j(i) = j(i+1)$. (See Fig. 1.7 with $i = 1$ and $i+1 = 2$.) □

We have shown that if $S$ cannot be decomposed by Proposition 1.3, then from every $B_i$, $i = 0, 1, \ldots \ell$, part of the same side $A_jA_{j+1}$ is visible. That implies

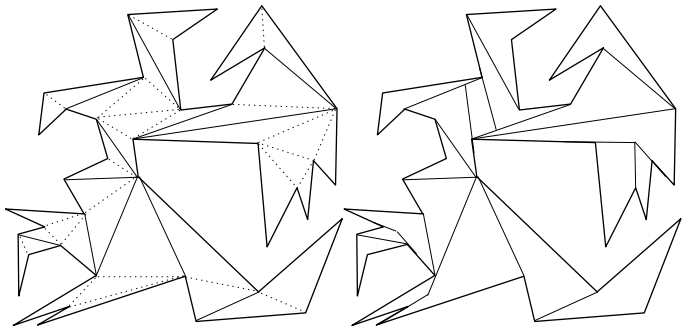

**Figure 1.8:** *Dense decomposition (on the left) of a simple polygon with* $40$ *vertices and its partition into into* $19$ *quadrilaterals (on the right).*

that from $B_j$ only one side of $K$ is visible, a contradiction. This completes the proof of Lemma 1.3.

## 1.4 Simple polygons with point $\pi$-guards

The main theorem of this section considers point guards of $180^\circ$ range of vision in simple polygons and establishes $g(\alpha) = \frac{1}{3}$ for all $\alpha \in [180^\circ, 360^\circ]$.

**Theorem 1.6.** *[48] Any art gallery with* $n$ *vertices can be monitored by at most* $\lfloor \frac{n}{3} \rfloor$ $180^\circ$*-guards.*

Throughout this section, the term *guard* refers to $180^\circ$-guard. Our proof is constructive, in terms that we specify the actual location of the guards. It is straightforward to transform our proof to an $O(n^2)$ algorithm, though it is possible that algorithms with smaller complexity exist even if based on our techniques.

The structure of this section is as follows. The Subsection 1.4.1 contains the proof of the main theorem. A crucial step of the argument, the proof of Lemma 1.4, is postponed to Subsection 1.4.2, and the proof of an important tool, Lemma 1.5, with fundamental technical details are left to Subsection 1.4.3. The last Subsection is devoted to the analysis of an example, a polygon of 23 sides. The kind reader is encouraged to consult the example while reading the proof.

**Definition 1.2.** *A polygonal path $c$ is a* cut *of $P$ if the two endpoints of $c$ are in $\partial P$ and* $\mathrm{int(c)} \subset \mathrm{int(P)}$.

*A cut $c$ decomposes $P$ into two simple closed polygonal domains $P_1$ and $P_2$, where $P = P_1 \cup P_2$, $c = P_1 \cap P_2$. Let $n_1$ and $n_2$ denote the number of sides of $P_1$ and $P_2$, resp.*

*A cut $c$ is called a* diagonal cut *if $c$ is a diagonal of $P$, i.e., a line segment connecting two non-adjacent vertices of $P$ and contained in $P$ [45].*

For every diagonal cut $c$ there is a triangulation $T$ such that $c$ corresponds to an edge of $G(T)$, and the graph $G(T_1)$ and $G(T_2)$ of triangulations of $P_1$ and $P_2$ can be obtained by deleting an edge of $G(T)$.

The induction steps of our proof is heavily rely upon dissection of an art gallery into two galleries along *cut*s. The cuts, we apply, are composed of one or two line segments, and are not restricted to diagonal cuts. In the proofs, we prefer using diagonal cuts, but in lack of appropriate diagonal cut we use freely the plausible cut that fits for our purposes. The reason why we cannot control the location of guards relative to the original art gallery $P$ is the method of consecutive dissections itself rather than the complexity of the cuts we use.

### 1.4.1 Proof of Theorem 1.6

Without loss of generality, we may assume that the vertices of $P$ are in general position.

We apply induction on $n$. It is easy to check [12] that the theorem holds for $n = 3, 4$ and $5$. We prove the theorem for a polygon $P$ with $n > 5$ vertices, assuming that it holds for every polygon with fewer than $n$ vertices.

In what follows, we either decompose $P$ into two polygons $P_1$ and $P_2$, and apply the induction hypothesis, or give the location of the guards explicitly.

**Definition 1.3.** *A cut is called* good cut*, if it satisfies the following condition.*

$$\left\lfloor \frac{n_1}{3} \right\rfloor + \left\lfloor \frac{n_2}{3} \right\rfloor \leq \left\lfloor \frac{n}{3} \right\rfloor .$$

Once we have found a good cut in $P$, the proof can be completed by applying the induction hypotheses for both $P_1$ and $P_2$. Through the five Propositions of this section, we either show that any art gallery $P$ with $n$ vertices contains a *good cut*, or we can specify the location of at most $\lfloor \frac{n}{3} \rfloor$ guards who monitor $P$. In both cases, Theorem 1.6 holds for $P$.

In the Propositions of this section, we will always apply diagonal cuts to dissect $P$. Observe that diagonal cuts has the property that the total number of vertices of the two components $P_1$ and $P_2$ is $n_1 + n_2 = n + 2$.

Let $T$ be a fixed triangulation of $P$, and let $G(T)$ be the corresponding graph.

**Proposition 1.5.** *For $n = 3k$, any diagonal is a good cut.*

*Proof.* For any diagonal, the number of vertices of $P_1$ and $P_2$ is either $n_1 = 3k_1 + 1$ and $n_2 = 3k_2 + 1 \quad (k_1 + k_2 = k)$, or $n_1 = 3k_1 + 2$ and $n_2 = 3k_2 \quad (k_1 + k_2 = k)$. In both cases, $\lfloor \frac{n_1}{3} \rfloor + \lfloor \frac{n_2}{3} \rfloor = \lfloor \frac{n}{3} \rfloor$. □

**Proposition 1.6.** *For $n = 3k + 1$, there is a good diagonal cut.*

*Proof.* If there is a diagonal cut where the number of vertices of $P_1$ and $P_2$ is $n_1 = 3k_1 + 2$ and $n_2 = 3k_2 + 1 \quad (k_1 + k_2 = k)$, then we are done.

Suppose to the contrary that $n_1 = 3k_1$ and $n_2 = 3k_2 \quad (k_1 + k_2 = k + 1)$ for every diagonal cut. Consequently, deleting any edge of $G(T)$, the resulting two connected components have $3k_1 - 2$ and $3k_2 - 2$ nodes resp.

If there is a node $v \in G(T)$ of degree 2, then one of the edges incident to $v$ violates this condition.

Therefore every node of $G(T)$ has degree 1 or 3. That implies that there is a node $v \in G(T)$ adjacent to two leaves. The diagonal corresponding to the third edge of $v$ is a good diagonal cut with $n_1 = 5$ and $n_2 = n - 3$. ($G(T)$ has at least four nodes, as $n > 5$.) Contradiction. □

**Proposition 1.7.** *For $n = 3k + 2$, a cut is a good cut if the number of vertices of $P_1$ and $P_2$ is $n_1 = 3k_1 + 2$, and $n_2 = 3k_2 + 2 \quad (k_1 + k_2 = k)$ resp.* □

From now on, we suppose that $n = 3k + 2$ and every diagonal decomposes $P$ into polygons of $n_1 = 3k_1$ and $n_2 = 3k_2 + 1$ vertices $(k_1 + k_2 = k + 1)$.

Consequently, $G(T)$ has $3k$ nodes, and deleting any edge of $G(T)$, it falls into two subtrees of $3k_1 - 2$ and $3k_2 - 1$ nodes ($k = k_1 + k_2 + 1$).

This assumption means that there is no good diagonal cut. Figure 1.9(b), depicts a octagon with no good diagonal cut; however this octagon does have a good cut (dashed line in Figure 1.9(b)). The same octagon serves as an example for a polygon containing no one-segment good cut.

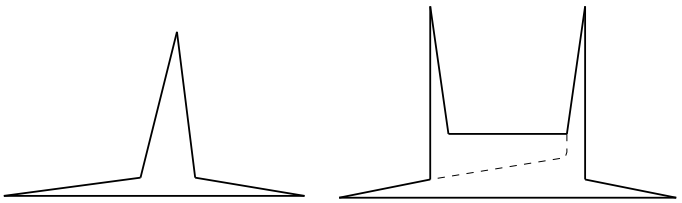


**Figure 1.9:** *A pentagon where one vertex $180°$-guard is not sufficient and an octagon with no good diagonal cut.*

**Proposition 1.8.** *$G(T)$ has exactly $k+1$ leaves.*

*Proof.* Introduce the following notation on subtrees of $G(T)$. Let $e$ be any edge of $G(T)$. Deleting $e$, we get two subtrees, say $G_1^e$ and $G_2^e$.

It is sufficient to establish the following statement which immediately implies Proposition 1.8:

*If $G_1^e$ has $3k_1 - 1$ or $3k_1 - 2$ nodes, then $G(T)$ has $k_1$ leaves in $G_1^e$.*

This implies Proposition 1.8: let $G_2^e$ be a leaf, and apply the statement to $G_1^e$ on $3k - 1$ nodes.

The proof of the statement is done by induction. Let $a$ be the node of $G_1^e$ incident to $e$. We distinguish two cases:

*Case 1:* $G_1^e$ has $3k_1 - 1$ nodes.

If $a$ has degree two, then let $f$ be the second edge incident to $a$. There are $3k_1 - 2$ nodes in $G_1^f$. The induction hypothesis implies that $G_1^f$ (and thus $G_1^e$) contains $k_1$ leaves of $G(T)$. If $a$ has degree three, then consider the two other edges $f$ and $g$. $G_1^f$ and $G_1^g$ have $3k_1' - 1$ and $3k_1'' - 1$ nodes, resp., where $k_1' + k_1'' = k_1$; the induction hypothesis completes the proof.

*Case 2:* $G_1^e$ has $3k_1 - 2$ nodes.

If $a$ has degree two, then let $f$ be the second edge incident to $a$. There are $3k_1 - 3$ nodes in $G_1^f$, so the diagonal cut corresponding to $f$ is a good diagonal cut. If $a$ has degree three, then consider the two other edges $f$ and $g$. $G_1^f$ and $G_1^g$ will have $3k_1' - 1$ and $3k_1'' - 2$ nodes, resp., where $k_1' + k_1'' = k_1$. By induction, $G(T)$ has $k_1' + k_1'' = k_1$ leaves in $G_1^e$. □

In any triangulation of $P$, every triangle corresponding to a leaf has one common angle with $P$. So $P$ has at least $k + 1$ convex angles. However, if we assume that there is no good cut, we can establish a much stronger assertion.

**Lemma 1.4.** *If $P$ has $n = 3k + 2$ vertices and contains no good cut, then $P$ has at most $k$ reflex angles.*

For the proof of Lemma 1.4, see Subsection 1.4.2. Lemma 1.4 allows to conclude the proof of Theorem 1.6.

**Proposition 1.9.** *If $P$ has at most $k$ reflex angles, then there are $k$ guards who monitor $P$.*

*Proof.* Observe that, if a polygon has at most one reflex angle, then it can be monitored by one guard.

If $P$ has at least two reflex angles, cut $P$ along the bisector of one of its reflex angles. We get two components, $P_1$ and $P_2$, and the total number of reflex angles is at most $k - 1$.

As long as any component has at least two reflex angles, cut it into two pieces along the bisector of one of its reflex angles. In each step, the total number of reflex angles decreases by one and the number of components increases by one.

Finally, we get at most $k$ components, each of which has at most one reflex angles. □

This completes the proof of Theorem 1.6. □

It is expedient to highlight the parts of the proof where guard location was explicitly given, which might be unclear from the structural propositions. Guard placement is carried out in Proposition 1.9 and at the very beginning of the proof for triangles, quadrilaterals and pentagons. In both cases, guards may be placed at sides of dissected components, i.e., also at interior points of the original art gallery. The following two sections discuss only decomposition methods except for Lemma 1.6 where one vertex guard is placed besides decomposition.

### 1.4.2 Proof of Lemma 1.4

We formulate two lemmas that add up to the proof of Lemma 1.4.

**Lemma 1.5.** *If $P$ has $n = 2k + 2$ vertices then one or the following two conditions is satisfied.*

*a, $P$ has a good cut,*

*b, for every $ABC \in T$ corresponding to a leaf in $G(T)$ (where $AC$ is a diagonal of $P$) either $\angle A < 180^\circ$ or $\angle C < 180^\circ$ in $P$.*

The proof of Lemma 1.5 is postponed to Section 4.

To every leaf of $G(T)$ we associate two convex angles of $P$ applying Lemma 1.5. For a leaf $ABC \in T$, $\angle B$ and $\angle A$, or $\angle B$ and $\angle C$ are convex angles of $P$. One convex angle $\angle B = \angle ABC$ is the common angle of $P$ and the triangle $ABC$, the other is located at another vertex of $ABC$.

Note that it may occur that we associate an angle, say $\angle BCD$, to both $ABC \in T$ and $CDE \in T$ (Figure 1.10). This case is treated in the following lemma.

**Lemma 1.6.** *If a convex angle of $P$ is associated to two leaves, then $\lfloor \frac{n}{3} \rfloor$ guards can monitor $P$.*

*Proof.* Let $\angle C = \angle BCD$ be the convex angle associated to two leaves of $G(T)$.

One guard at $C$ can monitor all the triangles adjacent to $C$. These triangles form a chain in $G(T)$. (See Figure 1.10 for an illustration.)

Deleting the nodes of this chain, $G(T)$ decomposes into $m$ smaller trees ($m \in \mathbf{N}$). Let us denote these trees by $L_1, L_2, \ldots, L_m$ along the chain. Each $L_i$ has $3k_i + 2$ or $3k_i + 1$ nodes.

Here we prove two short arithmetic propositions that allow to conclude the proof of lemma 1.6.

**Proposition 1.10.** *There is at least one node of degree two in the chain.*

*Proof.* Suppose that none of the nodes in the chain has degree 2. Clearly, the two extremal node of a chain has degree 1. If every $L_i$ has $3k_i + 2$ nodes, then there is altogether $2 + m + \sum_{i=1}^{m} 3k_i + 2 = 3\ell + 2$ nodes. But $G(T)$ has $3k$ nodes.

Consider the smallest index $i \in \mathbf{N}$ for which $L_i$ has $3k_i + 1$ nodes. Let $e$ be the edge that follows the node adjacent to $L_i$ in the chain. $e$ corresponds to a good cut, since deleting $e$, $G(T)$ decomposes to subtrees of $3\ell_1$ and $3(k - \ell_1)$ nodes, resp. □

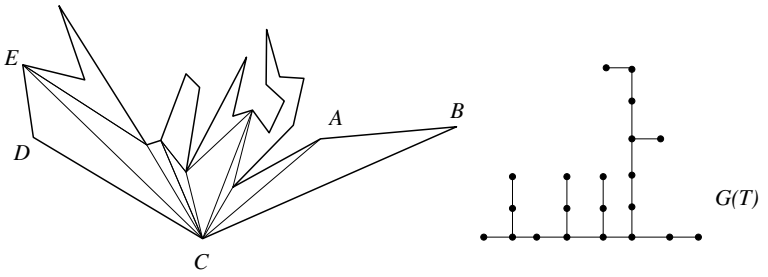


**Figure 1.10:** *Illustration for the proof of Lemma 1.6.*

**Proposition 1.11.** *There is at least one node of degree two in the chain between any pair of subtrees $L_i, L_j$ with $3k_i + 1$ and $3k_j + 1$ nodes.*

*Proof.* The same argument applies as in proof of Proposition 1.10. □

**Corollary 1.7.** *If there are $m'$ subtrees $L_i$ of $3k_i + 1$ nodes, then there are at least $m' + 1$ nodes of degree 2 in the chain.* □

The proof of Lemma 1.6 can now be completed as follows. By induction, $k_i + 1$ guards can monitor any $L_i$ with $3k_i + 2$ or $3k_i + 1$ nodes. Altogether we need $1 + \sum_1^m (k_i + 1)$ guards to monitor $P$, while the total number of nodes is at least $(2 + m + (m' + 1))$ in the chain plus $\sum_1^m (3k_i + 2) - m'$ in the $L_i$-s. That makes $3 + \sum_1^m (3k_i + 3)$. □

Assuming that $P$ does not have a good cut, Lemma 1.6 states that we have associated distinct convex angles to each of the $k + 1$ leaves. $P$ has at least $2k + 2$ convex vertices. This proves Lemma 1.4. □

### 1.4.3 Proof of Lemma 1.5

We distinguish two cases discussed separately.

**Definition 1.4.** *Consider a triangulation $T'$ of $P$. We call a leaf of $G(T')$ a* short leaf*, if it is adjacent to a node of degree* 3*. A* long leaf *is a leaf of $G(T')$ adjacent to a node of degree* 2*.*

**Remark.** Note that the node adjacent to a long leaf is also adjacent to a node of degree 3, otherwise one could find three consecutive nodes whose

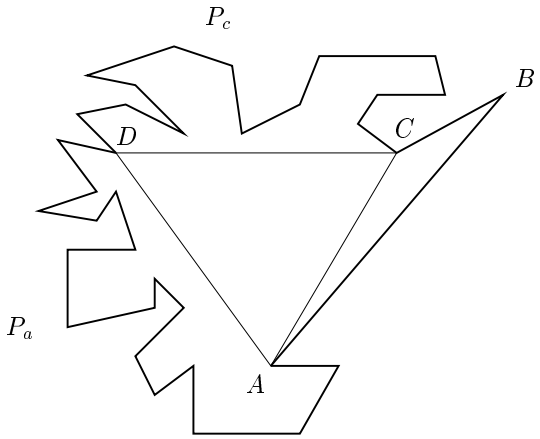


**Figure 1.11:** *Decomposition of $P$ if $ABC$ is a short leaf.*

corresponding triangles could form a pentagon, i.e., one component of a good diagonal cut.

### Short leaves

Let the triangle $ABC$ correspond to a short leaf in $G(T)$, where $AC$ is a diagonal of $P$. Suppose that $\angle A > 180°$ and $\angle C > 180°$ in $P$, and $P$ does not have a good cut. Let $ACD$ be the triangle corresponding to the node adjacent to the leaf $ABC$.

$AD$ and $CD$ decompose $P$ into polygons $P_a$, $ABCD$, and $P_c$, where $AD$ is a side of $P_a$ and $CD$ is a side of $P_c$ (see Figure 1.11).

**Claim 1.8.** *$P_a$ and $P_c$ have $3l_a + 1$ and $3l_c + 1$ sides, resp. (i.e., they have $3l_a$ and $3l_c$ common sides with $P$, where $n = 3l_a + 3l_c + 2$).*

*Proof.* If $P_a$ has $3l_a + 2$ vertices, then $AD$ is a good diagonal cut according to Proposition 1.7. If $P_a$ has $3l_a$ vertices, then $P_a \cup ABCD$ has $3l_a + 2$ sides, i.e., $CD$ is a good diagonal cut. □

**Claim 1.9.** *$ABCD$ is a non-convex quadrilateral.*

*Proof.* If $BD$ is a diagonal of $ABCD$, then $P_a \cup ABD$ has $3l_a + 2$ vertices, so $BD$ is a good diagonal cut. □

Thus, $ABCD$ has a reflex angle at $A$ or at $C$. We may suppose w.l.o.g. that it is at $C$.

Let $T'$ be a triangulation of $P$, where $ABC$ is a short leaf and the number of vertices of $P_a$ is minimal (i.e., $l_a$ minimal).

Denote the vertices of $P_a$ next to $A$ and $D$ by $A_0$ and $D_0$, respectively. The ray $\overrightarrow{BA}$ reaches the boundary of $P$ ($\partial P$, for short) at $A'$, the ray $\overrightarrow{CD}$ reaches $\partial P$ at $D'$, and the ray $\overrightarrow{BC}$ reaches $\partial P$ at $C'$. These points are all interior points of sides of $P$, since the vertices of $P$ are in general position.

The following useful definition describes a minimum search on the vertices of $P$ visible from a vertex $X$. It can also be interpreted as finding the first element of a *partial local sequence* of $X$ defined in [36].

**Definition 1.5.** *Let $X, Y$ and $Z$ be non-collinear points in the plane such that $X$ is a vertex of $P$, $Y \in \partial P$, and the line segment $XY$ is contained in $P$.*

*Then let $W = r(X, Y, Z)$ be the first vertex of $P$ visible from $X$ that is hit if we spin the ray $\overrightarrow{XY}$ around $X$. The orientation of the rotation be the same as the orientation of the triangle $XYW$.*

*When we say that we* rotate *$\overrightarrow{XY}$ towards $Z$, we determine $r(X, Y, Z)$. (Observe that such a vertex always exists, because the neighboring vertices of $X$ in $P$ are visible form $X$.)*

**Claim 1.10.** *$A'$ and $D'$ are points of the same side of $P$.*

*Proof.* Rotate $\overrightarrow{CA}$ towards $D'$. If the ray hits $D$, then $A'$ and $D'$ are points of the same side of $P$, so the claim holds.

If the ray hits $A_0$, then $\angle A = \angle BAA_0 < 180°$ and condition **b** of Lemma 1.5 holds.

Assume that the ray hits a vertex $O$ (see Figure 1.12). $AO$ and $CO$ are diagonals of $P$. There exists a triangulation of $P$ where $ACO$ is a triangle and $ABC$ is a short leaf.

Either $ABCO$ is a convex quadrilateral, contradicting Claim 1.9, or the reflex angle of $ABCO$ is $\angle BCO$ and the corresponding $P'_a$ is smaller than $P_a$, contradicting the minimality of $P_a$. □

**Claim 1.11.** *$\angle CDD_0 > 180°$ in $P$.*

*Proof.* If $\angle CDD_0 < 180°$, then $A'$ must be on $DD_0$ (so $C'$, too) according to Claim 1.10. Now $CC'$ is a good cut with two component of $3l_c + 2$ and $3l_c + 2$ vertices, resp., (see Figure 1.13).

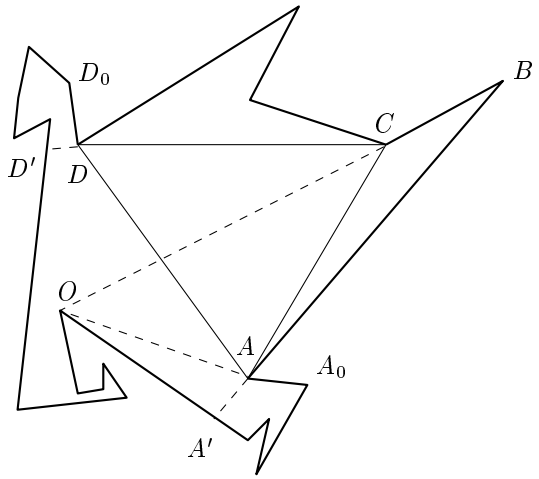


**Figure 1.12:** *Case where $r(C, A, D')$ equals $O$.*

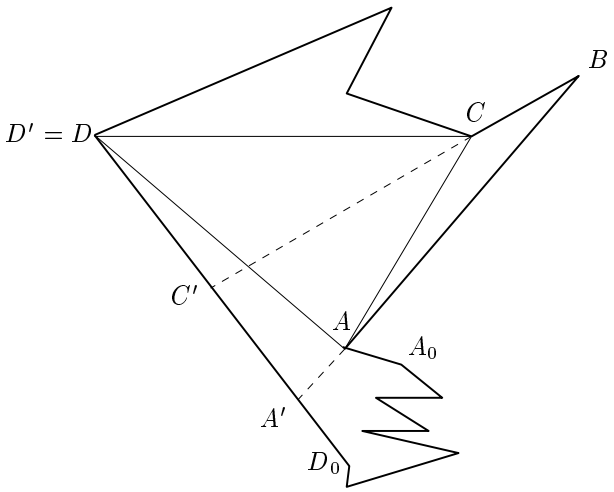


**Figure 1.13:** *Illustration for Proposition 1.11.*

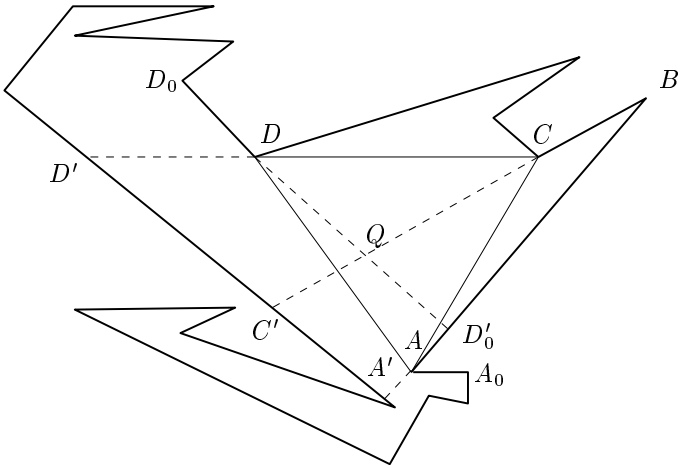


**Figure 1.14:** *$CC'$ and $D_0D'_0$ intersect inside $CAA'D'$.*

$\angle CDD_0 \neq 180°$, because the vertices are in general position. (But anyway, $\angle CDD_0 = 180°$ implies $D' = D_0$ and $CD$ would be a good cut with $3k_1 + 1, 3k_2 + 2$ vertices.) □

So far we have concluded that $A'$ and $D'$ are points of the same side of $P$, $\angle CDD_0$ is reflex angle, and $D_0$ is outside of $CAA'D'$.

**Claim 1.12.** *$\angle CDD_0 < 180°$ in $P$.*

*Proof.* Suppose that $\angle CDD_0 > 180°$ in $P$. Let $D'_0$ be the point where the ray $\overrightarrow{D_0D}$ reaches $\partial P$.

If the line segments $CC'$ and $D_0D'_0$ intersect inside $CAA'D'$ at $Q$, then $CQ \cup QD$ is a good cut with $P_c \cup CQD$ and $P_a \cup ABQD$ of $3l_a + 2$ and $3l_c + 2$ vertices, resp., (Figure 1.14).

If $CC'$ and $D_0D'_0$ do not intersect inside $CAA'D'$, then $D'_0$ is on the line segment $C'D'$ (Figure 1.15). In this case, one of $DD'_0, CC'$, and $AA'$ is a good cut.

To see this, count the number of edges in the component containing $D_0$. (Note that the total number of vertices increases by two at each cut.) Suppose that $D_0$ is in a component having $h$ vertices at the cut $DD'_0$. At the cut $CC'$, the number of vertices of the component containing $D_0$ is $h + 3l_c + 1$. Then at the

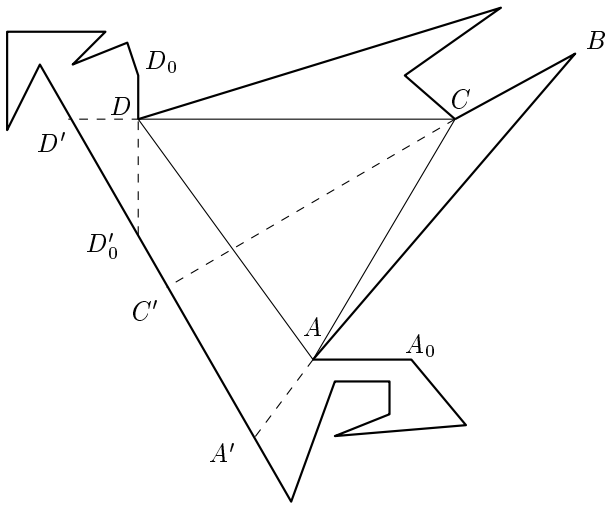


**Figure 1.15:** *$D'_0$ is on the line segment $C'D'$.*

cut $AA'$, the component of $D_0$ has one more vertex (i.e., along $BA'$). Either $h, h + 3l_c + 1$, or $h + 3l_c + 2$ is of the form $3k_1 + 2$. □

Contradiction, Claim 1.11 and Claim 1.12 cannot hold for the same $\angle CDD_0$. This completes the proof of Lemma 1.5. □

**Corollary 1.13.** *Suppose that a leaf $t$ of $G(T)$ corresponds to a triangle $ABC$ and there is a triangulation $T'$ where $ABC$ is a short leaf. Lemma 1.5 holds for the leaf $ABC$ even if it is a long leaf in our fixed triangulation $T$.*

### Long leaves

Let $ABC$ be a leaf of $G(T)$ which is not a short leaf in any triangulation of $P$ so that $AC$ is a diagonal of $P$. Assume that $\angle A > 180°$ and $\angle C > 180°$ in $P$, and $P$ does not have a good cut.

**Claim 1.14.** *In any triangulation $T'$ with $ABC \in T'$, the node adjacent to the node $ABC$ corresponds to the same triangle.*

*Proof.* Suppose that the node adjacent to the node $ABC \in T$ corresponds to two different triangles $ACD$ and $ACE$ in two different triangulations. Since

$ABC$ is not a short leaf in any triangulation, $D, A, B, C, E$ (or $E, A, B, C, D$) are consecutive vertices along $\partial P$.

Either $DE$ is a diagonal of $P$, and then $DE$ is a good diagonal cut since $DABCE$ is a pentagon; or there is a vertex of $P$ in the quadrilateral $DACE$. Let $O$ be the closest[1] vertex to the line $AC$ in $DACE$. Now $AO$ and $CO$ are diagonals, and there exists a triangulation $T'$ where $ABC$ is a short leaf adjacent to $ACO$. □

Let $ACD$ be the unique triangle adjacent to $ABC$. $\overrightarrow{CA}$ and $\overrightarrow{CD}$ reach $\partial P$ at $A'$ and $D'$, resp., (Figure 1.16).

**Claim 1.15.** *$A'$ and $D'$ are points of the same side of $P$.*

*Proof.* Rotate $\overrightarrow{CA'}$ towards $D'$. If the ray hits vertex $D$, then $A'$ and $D'$ are points of the same side of $P$, so the claim holds.

If the ray hits $A_0$, then $\angle BAA_0 < 180^\circ$ and condition **b** of Lemma 1.5 holds.

Suppose that the ray hits a vertex $O \neq D$. Then $ACO \subset P$, there is a triangulation of $P$ where $ACO$ is a triangle and $ABC$ is a short leaf. Contradiction. □

**Claim 1.16.** *$P$ has a convex angle at $A$, $C$ or $D$.*

*Proof.* Suppose that $P$ has reflex angles at $A, C$ and $D$ (Figure 1.16). $\overrightarrow{BC}$ and $\overrightarrow{BA}$ reach $\partial P$ at $C'$ and $B'$ resp. According to Claim 1.15, $C'$ and $B'$ are on the same side as $A'$ and $D'$.

A short counting argument shows that one of $DD'$, $CC'$, or $AB'$ is a good cut. Count the number of edges in the component containing $D_0$. Suppose that $D_0$ is in a component with $h$ vertices at the cut $DD'$. At the cut $CC'$, the number of vertices of the component containing $D_0$ is $h + 1$; at the cut $AB'$, the component of $D_0$ has one more vertex (i.e., along $BB'$). One of $h, h + 1$ or $h + 2$ is of the form $3k_1 + 2$. □

**Definition 1.6.** *Two sides of $P$, $XY$ and $ZW$, are equivalent with respect to vertex $V$, if the number of vertices of $P$ met when passing from the midpoint of $XY$ to the midpoint of $ZW$ along $\partial P$ on the arc not containing $V$ is a multiple of three.*

**Claim 1.17.** *$P$ cannot have a convex angle at $D$ and reflex angles at $A$ and $C$ at the same time.*

[1] Minimizing the point-line distance.

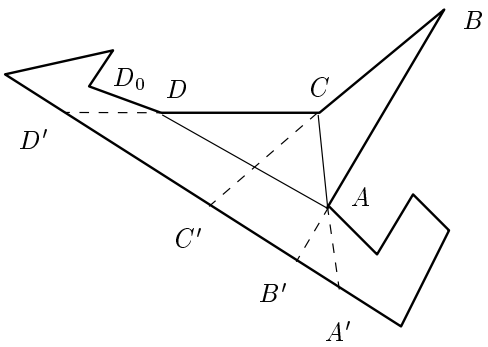


**Figure 1.16:** *Illustration for the proof of Claim 1.16.*

*Proof.* Suppose that $P$ has a convex angle at $D$ and reflex angles at $A$ and $C$. Claim 1.15 implies that $B'$ and $D'$ are on the side $DD_0$ (see Figure 1.17). Let $A'_0$ be the point where $\overrightarrow{A_0A}$ reaches $\partial P$. According to Claim 1.15, $A'_0$ may be on $CD$ or $DA'$. It cannot be on $DA'$, otherwise the cut $AA'_0$ were a good cut with a pentagonal component $ABCDA'_0$. We can deduce that at least an interval of the side $AA_0$ is visible from vertex $D$, this fact is used in the next paragraphs.

Label the vertices of $P$ from $A$ with numbers $A = 0, A_0 = 1$, and so on $2, 3, 4, \ldots$. Rotate $\overrightarrow{DA}$ towards $A_0$. The ray cannot hit $A_0$, since $DA_0$ would be a good diagonal cut (as $A_0ABCD$ is a pentagon). It hits a vertex $q \in \mathbf{N}$.

$\overrightarrow{(q-1)q}$ and $\overrightarrow{(q+1)q}$ reach $\partial P$ on either $CD$ or $A_0A$. Both rays reach $\partial P$ on sides equivalent to $AA_0$ with respect to $D$. $q$ is a multiple of 3, otherwise $\overrightarrow{(q-1)q}$ or $\overrightarrow{(q+1)q}$ is a good cut. If $q \equiv 1 \mod 3$, then the polygon $12 \ldots q(\overrightarrow{(q+1)q} \cap A_0A)$ or $CBA12 \ldots q(\overrightarrow{(q+1)q} \cap CD)$ has $2k' + 2$ vertices. If $q \equiv 2 \mod 3$, then the polygon $12 \ldots q(\overrightarrow{(q-1)q} \cap A_0A)$ or $CBA12 \ldots q(\overrightarrow{(q-1)q} \cap CD)$ has $3k' + 2$ vertices. Observe that for a modulo 3 counting, we do not have to know on exactly which side $\overrightarrow{(q-1)q}$ or $\overrightarrow{(q+1)q}$ reaches $\partial P$, it is enough to know that these sides are equivalent with respect to $D$ and check one possibility.

Continue rotating $\overrightarrow{Dq}$ towards $(q+1)$. Now again the ray hits $q'$, where $\overrightarrow{(q'-1)q'}$ and $\overrightarrow{(q'+1)q'}$ reach $\partial P$ on sides equivalent to $AA_0$ with respect to

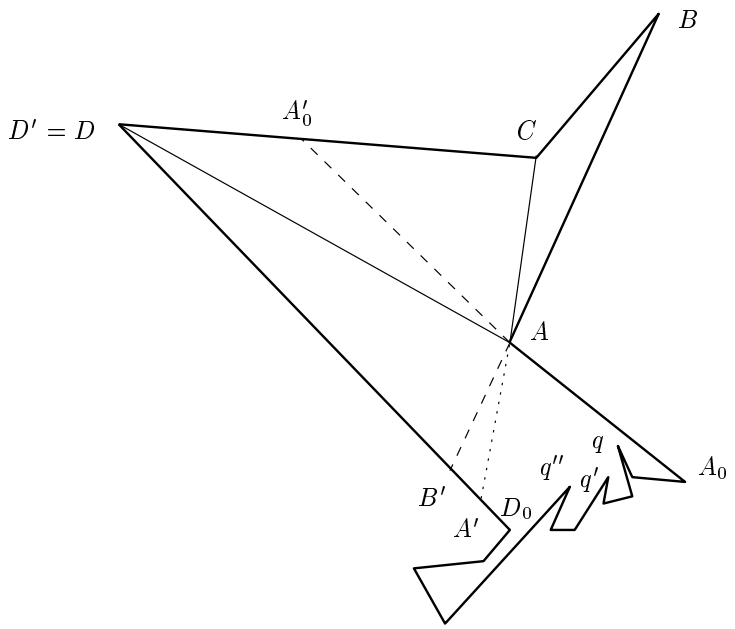


**Figure 1.17:** *Illustration for the proof of Claim 1.17.*

$D$. This implies that $q'$ is a multiple of 3.

Repeating this procedure, the ray spinning around $D$ must hit $D_0$ (Figure 1.17). We have just proved that the label of $D_0$ is a multiple of three.

Together with $D, C, B$ and $A$, $P$ would have $3\hat{k} + 1$ vertices. Contradiction, because $P$ has $3k + 2$ vertices. □

### 1.4.4 An example

In this subsection we demonstrate how the proof of Theorem 1.6 is applied to a specific art gallery. It facilitates the comprehension of the rather technical structural propositions, and also gives an example where our method places guards at an interior point of the art gallery.

The polygon $P$ in Figure 1.18 has $23 = 3 \cdot 7 + 2$ vertices. One possible triangulation $T$ of $P$ is drawn with dotted lines. The corresponding graph $G(T)$ is represented in Figure 1.19.

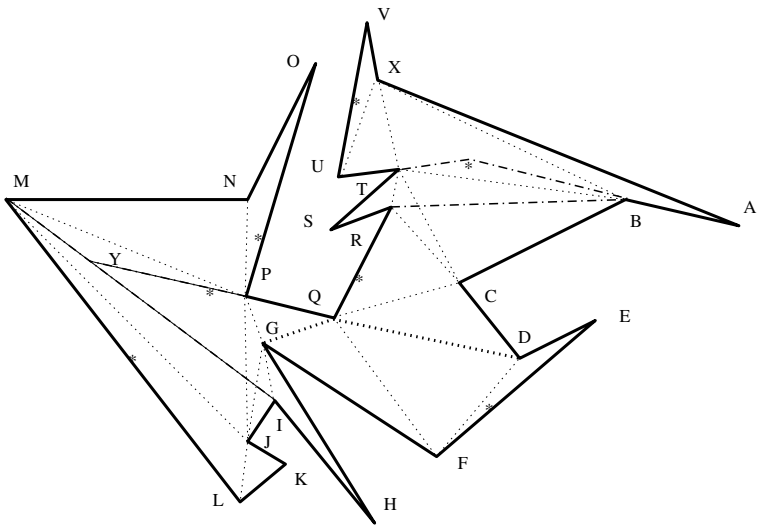


**Figure 1.18:** *Example for our* $180°$*-guard placement.*

Figure 1.19 also shows how $G(T)$ can be partitioned into three subtrees $G_1$, $G_2$, and $G_3$ with $3k_1$, $3k_2$, and $3k_3$ nodes respectively. The three corresponding polygons are $P_1 = ABCDQRSTUVX$, $P_2 = DEFGQ$, and $P_3 = GHIJKLMNOPQ$.

In $P_1$, there is a good diagonal cut $BR$ that dissects $P_1$ into pentagon $P_4 = BCDQR$ and octagon $P_5 = ABRSTUVX$. $P_5$ does not have any good diagonal cut. It has four reflex angles, so Proposition 1.9 cannot be applied to place guards in $P_5$. We consider the short leaf $ABX$. (Note that the triangulation of $P_5$ is not inherited from $P$, since $BR$ is a side of $P_5$.) Section 4.1 applies to the short leaf $ABX$, but Claim 1.12 does not hold for $\angle AXV$. Therefore Claim 1.12 shows a good (non-diagonal) cut in $P_5$. Thus $P_5$ is decomposed into two pentagons.

In $P_3$, there is a good diagonal cut $IM$ that dissects $P_3$ into pentagon $P_6 = IJKLM$ and octagon $P_7 = GHIMNOPQ$. $P_7$ does not have any good diagonal cut. It has four reflex angles, so Proposition 1.9 cannot be applied to $P_7$. Consider the long leaf $NOP$. It is not a short leaf in any triangulations of $P_7$. Section 4.2 applies to the long leaf $NOP$, but Claim 1.17 does not hold for $\angle OPQ$. Therefore Claim 1.17 (first paragraph) shows a good (non-diagonal) cut in $P_7$. Thus $P_7$ is decomposed into two pentagons.

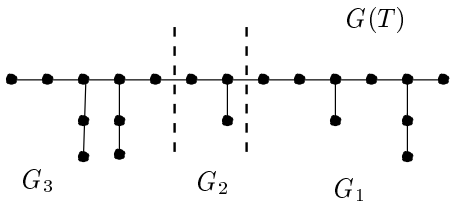


**Figure 1.19:** *Dual graph of the example polygon.*

The decomposition of $P$ into pentagons is drawn with dash-dotted lines. Each pentagon can easily be monitored by one guard. One possible allocation of guards (along sides of the pentagons) is marked by asterisks in Figure 1.18.

Observe that pentagon $GHIYQ$ can only be monitored with a guard in the interior of $P$.

## 1.5 Lower bounds on the number of $\alpha$-guards

### 1.5.1 Lower bound for $\alpha < 180°$

In this subsection, we show that $g(\alpha) = 2/3$ for all $\alpha \in [90°, 180°)$. We have shown in the previous section that $\lfloor n/3 \rfloor$ $180°$-guards can monitor a simple polygon with $n$ vertices. Therefore, $g(\alpha) \leq 2/3$ follows from an obvious guard allocation method: Consider an optimal $180°$-guard allocation and replace every $180°$-guard by two $90°$-guards. A lower bound of $g(\alpha) \geq 2/3$ is a straightforward consequence of the next theorem.

**Theorem 1.18.** *[53] For every $\ell \in \mathbb{N}$ and every $\alpha$, $0° < \alpha < 180°$, there exists an art gallery $P$ with $3\ell + 2$ vertices that cannot be monitored by less than $2\ell$ $\alpha$-guards.*

*Proof.* Fix an angle $\beta$, $\alpha < \beta < 180°$. We give the positions of all $3\ell + 2$ vertices of the polygon $P = B_0C_1B_1C_2 \dots B_{\ell-1}C_\ell B_\ell A_\ell A_{\ell-1} \dots A_0$ as follows. The vertices $A_i$ lie on a circle centered at $O$ of large radius such that $\mathrm{dist}(A_iA_{i+1})$ is constant. Let $C_i'$ be the middle point of the segment $A_{i-1}A_i$ for $i = 1, \dots, \ell$. Let $C_i$ be on ray $\overrightarrow{OC_i'}$ with $\mathrm{dist}(C_i', C_i) = \varepsilon$, where $\varepsilon > 0$ is a sufficiently small constant to be specified later. Let $B_i$, $i = 0, \dots, \ell$, be on

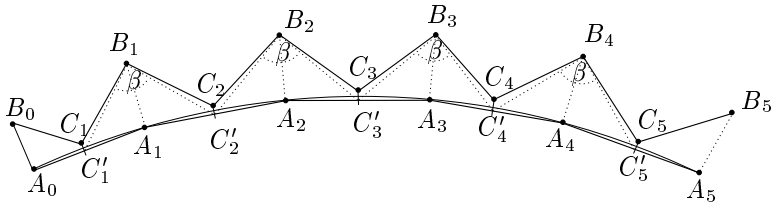


**Figure 1.20:** *Polygon $P$.*

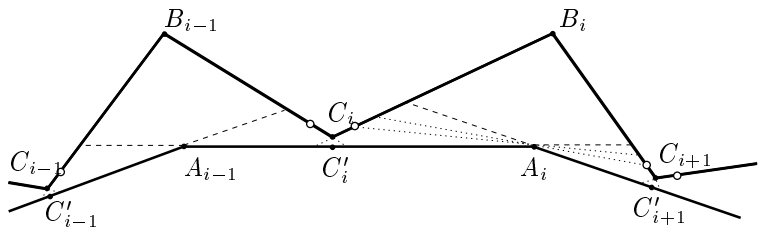


**Figure 1.21:** *Auxiliary lines and points.*

$\overrightarrow{OA_i}$ with constant $\mathrm{dist}\,(A_i, B_i)$ such that $\angle A_i B_i C'_{i+1} = \beta/2$ and $\angle c'_i B_i A_i = \beta/2$ (see Figure 1.20).

Observe that $\angle B_{i-1} C_i B_i > \angle B_{i-1} C'_i B_i$, because $B_{i-1} C_i B_i \subset B_{i-1} C'_i B_i$; and $\angle B_{i-1} C'_i B_i > C'_i B_i C'_{i+1} = \beta$ (that is, $\angle C_i C'_i B_i > \angle C'_i B_i A_i$), because points $B_i$ and $C_i$ are located on concentric circles. There is an $\varepsilon_0 > 0$ such that if $0 < \varepsilon \leq \varepsilon_0$ then $C_{i-1}$ and $C_i$ cannot see each other. Let $0 < \varepsilon \leq \varepsilon_0$.

We place $2\ell$ points $\mathcal{R} = \{R_1^-, R_1^+, R_2^-, R_2^+, R_3^-, \ldots, R_\ell^+\}$ in $P$ so that one $\alpha$-guard can see at most one of them at a time. This will prove that at least $2\ell$ guards are required to monitor $P$, which gives $f(3\ell + 2, \alpha) \geq 2\ell$.

Let $R_i^- \in C_i B_{i-1}$ and $R_i^+ \in C_i B_i$ such that the distance $\mathrm{dist}(C_i, R_i^-)$ and $\mathrm{dist}(C_i, R_i^+)$ are both equal to a constant $\eta > 0$. Let $\eta$ be so small that $R_i^+$ and $R_{i+1}^-$ cannot see each other for $\varepsilon = \varepsilon_0$. Now $R_i^+$ and $R_{i+1}^-$ cannot see each other for any $\varepsilon \leq \varepsilon_0$. The points $R_i$ are shown as small empty circles in Figure 1.21.

It is clear that two points $R_i^-$ and $R_j^-$ (resp., $R_i^+$ and $R_j^+$) are not visible from any point of $P$ if $j - i > 1$, since $C_i$ and $C_{i+1}$ do not see each other.

For every $i = 1, \ldots, \ell$, both $R_i^-$ and $R_i^+$ are visible from points of a triangle at $C_i$. Note that if there is a point $Q \in P$ such that $\angle R_i^- Q R_i^+ \leq \alpha$, then there is

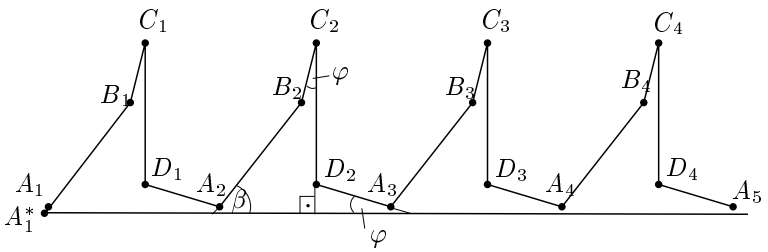


**Figure 1.22:** *Polygon $P$.*

$Q' \in A_{i-1}A_i$ with $\angle R_i^- Q' R_i^+ \leq \alpha$ ($Q'$ might be the perpendicular projection of $Q$ to $A_{i-1}A_i$). Since $Q'$ sees both $R_i^-$ and $R_i^+$, $Q$ is on the line segment $(A_{i-1}A_i \cap \overline{B_{i-1}C_i})(A_{i-1}A_i \cap \overline{B_iC_i})$. This segment is in an arbitrarily small neighborhood of $C_i'$ if $\varepsilon$ is small enough with respect to $\eta$. No $\alpha$-guard on this segment can monitor both $R_i^-$ and $R_i^+$, since $\angle B_{i-1}C_i'B_i > \angle B_{i-1}C_iB_i' > \beta > \alpha$ and $\angle R_i^- Q' R_i^+$ is arbitrarily close to $\angle B_{i-1}C_i'B_i$.

For every $i = 1, \ldots, \ell - 1$, both $R_i^+$ and $R_{i+1}^-$ are visible from points of a deltoid in $A_iC_{i+1}B_iC_i$. It is easy to see that among all points $Q \in P$ that see both $R_i^+$ and $R_{i+1}^-$, $\angle R_i^+ Q R_{i+1}^-$ is minimized by $Q = B_i$. But $\angle R_i^+ B_i R_{i+1}^- > \beta > \alpha$. □

**Corollary 1.19.** *For every $n \geq 3$ and $\alpha < 180°$, there exists an art gallery $P$ of $n$ sides that cannot be monitored by less than $2n/3 - 3$ $\alpha$-guards.*

*Proof.* For $n = 3, 4$ the assertion is void. If $n = 3\ell + 2$, then the theorem above gives the bound. If $n = 3\ell + 3$, or $3\ell + 3$, then we can insert one or two additional vertices with angle $(180 - \delta)°$, with a very small $\delta > 0$, to a polygon described in the proof of Theorem 1.18. □

### 1.5.2 Lower bound for $\alpha < 90°$

**Theorem 1.20.** *[53] For every $\ell \in \mathbb{N}$ and every $\alpha$, $0° < \alpha < 90°$, there exists an art gallery $P$ of $4\ell + 1$ vertices that cannot be monitored by less than $3\ell$ $\alpha$-guards.*

*Proof.* Fix an angle $\beta$, $\alpha < \beta < 180°$. We give the locations of all $4\ell + 1$ vertices of the polygon $P = A_1^* B_1 C_1 D_1 A_2 B_2 C_2 D_2 A_3 \ldots A_\ell B_\ell C_\ell D_\ell A_{\ell+1}^*$ as follows. Pentagons $A_iB_iC_iD_iA_{i+1}$ are translates of each other for $i =$

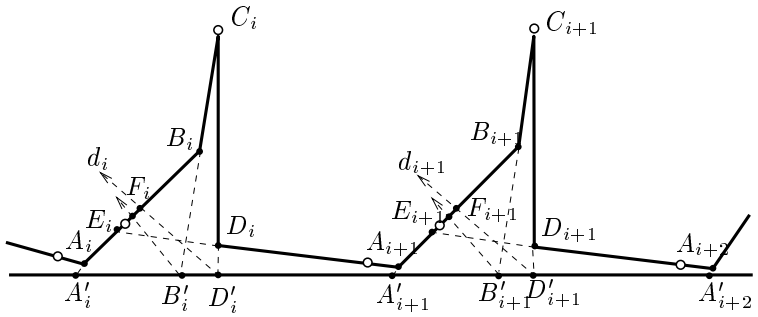


**Figure 1.23:** *Auxiliary lines and points.*

$1, \ldots, \ell$ . Vertices $A_1^*$ and $A_{\ell+1}^*$ are below the line of collinear vertices $A_i$, $i = 1, \ldots, \ell+1$, such that $A_1^* \in A_1B_1$, $A_{\ell+1}^* \in D_\ell A_{\ell+1}$, and $\operatorname{dist}(A_i, A_1A_{\ell+1}) = \varepsilon$ with a sufficiently small $\varepsilon > 0$. Side $C_iD_i$ is on the (perpendicular) bisector of $A_iA_{i+1}$ for all $i = 1, 2, \ldots, \ell$. The angles of $A_iB_iC_iD_iA_{i+1}$ are $\angle A_{i+1}A_iB_i = \beta$, $\angle B_iC_iD_i = \angle D_iA_{i+1}A_i = \varphi$ with a sufficiently small $0° < \varphi < 90° - \beta$ (see Figure 1.22).

Note that $\varepsilon$ can be scaled by moving the line $A_1^*A_{\ell+1}^*$ closer to line $A_1A_{\ell+1}$. Similarly, $\varphi$ can be scaled by moving the points $C_i$ and $D_i$ such that the distances $\operatorname{dist}(C_i, A_1A_{\ell+1})$ increase and the distances $\operatorname{dist}(D_i, A_1A_{\ell+1})$ decrease. The two parameters, $\varepsilon$ and $\varphi$, can be chosen independently.

We place $3\ell$ points $\mathcal{R} = \{R_1^b, R_1^c, R_1^d, R_2^b, R_2^c, R_2^d, R_3^b, \ldots, R_\ell^d\}$ in $P$ so that one $\alpha$-guard can see at most one of them at a time. This will prove that at least $3\ell$ guards are required to monitor $P$, which gives $f(4\ell + 1, \alpha) \geq 3\ell$.

Now we define auxiliary lines and points that will help placing the points of $\mathcal{R}$ for every $i = 1, \ldots \ell$.

- Let $A'_i = A_1^*A_\ell^* \cap \overline{B_iA_i}$.
- Let $B'_i = A_1^*A_{\ell+1}^* \cap \overline{C_iB_i}$.
- Let $D'_i = A_1^*A_{\ell+1}^* \cap \overline{C_iD_i}$.
- Let $E_i = \overline{A_{i+1}D_i} \cap A_iB_i$.
- Let $d_i$ be the ray emanating from $D'_i$ such that $\angle(\overrightarrow{D'_iC_i}, d_i) = \beta$ and $d_i$ is on the same side of the line $D'_iC_i$ as $A_i$.

- Let $s_i = \{X \in A_iB_i : \exists Y \in B_i'D_i', C_iY \subset P, \angle XYC_i = \beta\}$ be a segment on $A_iB_i$.

There is a constant $\varepsilon_0$ such that if $0 < \varepsilon \leq \varepsilon_0$ then ray $d_i$ hits the relative interior of $A_iB_i$, and $\alpha < \angle B_iA_i'A_{i+1} < \beta$. Let $F_i = A_iB_i \cap d_i$ and assume that $0 < \varepsilon \leq \varepsilon_0$.

We determine $\varphi$ for a fixed $\varepsilon = \varepsilon_0 > 0$. If $\varphi$ is small enough, then $A_iE_i$ is in an arbitrarily small neighborhood of $A_i$, and $B_i'D_i'$ is in an arbitrarily small neighborhood of $D_i'$, and in turn, segment $s_i$ is in an arbitrarily small neighborhood of $F_i$. Let $\varphi$ be so small that $A_iE_i$ and $s_i$ are disjoint. If $\varepsilon < \varepsilon_0$, then the gap between $A_iE_i$ and $s_i$ is increasing.

Let $R_i^b \in A_iB_i$ be a point between segments $A_iE_i \subset A_iB_i$ and $s_i \subset A_iB_i$. Let $R_i^c = C_i$. Let $R_i^d \in D_iA_{i+1}$ in a small neighborhood of $A_{i+1}$ such that $\alpha < \angle B_iA_i'R_i^c < \beta$. The location of the points of $\mathcal{R}$ are shown as empty dots in Figure 1.23. Note that the location of $\mathcal{R}$ is independent of $\varepsilon$, it depends only on $\varepsilon_0$ and $\varphi$.

First we show that an $\alpha$-guard monitoring $R_i^c$ cannot monitor any other point of $\mathcal{R}$. If there is a point $Q \in P$ such that $\angle R_i^cQR_i^b \leq \alpha$, then there is such a $Q$ on the line segment $B_i'D_i'$. But $\varphi$ was chosen so that $\angle R_i^cQR_i^b \geq \beta$ for all $Q \in B_i'D_i'$. Among all points $Q \in P$ that see both $R_i^c$ and $R_i^d$ the angle $\angle R_i^cQR_i^d$ is minimized by $Q = B_i'$. But $\angle R_i^cB_i'R_i^d = \angle B_iB_i'R_i^d < \angle B_iA_i'R_i^d$, and $\angle B - iA_i'R_i^d$ is arbitrarily close to $\alpha < \angle B_iA_i'A_{i+1}$ if $\varepsilon$ is suffitiently small.

Next, we consider other pairs of $\mathcal{R}$. Both $R_i^b$ and $R_{i+1}^b$ are visible from points of a small triangle at $A_{i+1}$, and $\angle R_i^bA_{i+1}R_{i+1}^b > 90°$. Both $R_i^d$ and $R_{i+1}^b$ are visible from points of another small triangle at $A_{i+1}$, and $\angle R_i^dA_{i+1}R_{i+1}^b > 90°$. Both $R_i^d$ and $R_{i+1}^d$ are visible from a third small triangle at $A_{i+1}$, and $\angle R_i^dA_{i+1}R_{i+1}^d > 90°$. If $\varepsilon < \varepsilon_0$ is small enough with respect to $\varphi$, then these triangles are in arbitrarily small neighborhood of $A_{i+1}$, and no $\alpha$-guard can see both of these pairs of points.

It is easy to see that among all points $Q \in P$ that see both $R_i^b$ and $R_i^d$, $\angle R_i^bQR_i^d$ is minimized by $Q = A_i'$. In this case, $\angle R_i^bA_i'R_i^d > \alpha$, if $\varepsilon$ is small enough, because $\angle B_iA_iR_i^d > \alpha$. □

**Corollary 1.21.** *For every $n \geq 3$ and every $\alpha < 90°$, there exists an art gallery $P$ of $n$ vertices that cannot be monitored by less than $3n/4 - 4$ $\alpha$-guards.*

*Proof.* For $n = 3, 4$ the assertion is void.

If $n = 4\ell+1$, then the theorem above gives the bound. If $n = 4\ell+2$, $4\ell+3$, or $4\ell+4$, then insert one, two, or three additional vertices with angle $(180-\delta)^\circ$, with a very small $\delta > 0$, to a polygon described in Theorem 1.20. □

### 1.5.3 Lower bounds for $\alpha < 60^\circ$

For every angle $\alpha < 60^\circ$, we employ one type of polygon in the argument. Still we expect that the bound deduced is best possible. The key tool is an analytic theorem due to Zsolt Páles.

**Theorem 1.22.** *[37] Let $D$ be a unit disk and let $C$ be a concentric disk of radius $r > 1$. If $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a set of $k$ angles such that $k$ guards of range of vision $\alpha_1, \alpha_2, \ldots, \alpha_k$, resp., located in $C$ monitor the disk $D$, then $\sum_{i=1}^{k} \alpha_i \geq 2 \cdot \arcsin(1/r)$.* □

**Corollary 1.23.** *For every $\alpha < 60^\circ$, the equilateral triangle cannot be monitored by less than $\left\lceil \frac{60^\circ}{\alpha} \right\rceil$ $\alpha$-guards.*

*Proof.* Consider an equilateral triangle $T$. Let $D$ be the maximal inscribed disk of $T$ and let $C$ be the circumscribed disk of $T$. Suppose that $k$ $\alpha$-guards can monitor $T$. Consider a placement of $k$ $\alpha$-guards monitoring $T$. These $k$ $\alpha$-guards are in $C$ and monitor $D$. Now $k \cdot \alpha \geq 60^\circ$ according to Theorem 1.22. □

**Theorem 1.24.** *[53] For every odd $n \geq 3$ and every $\alpha < 60^\circ$, there exists an art gallery $P$ with $n$ vertices that cannot be monitored by less than $\left\lceil \frac{60^\circ}{\alpha} \right\rceil \frac{n-1}{2}$ $\alpha$-guards.*

*Proof.* Consider the polygon $P = A_1^* B_1 A_2 B_2 A_3 \ldots A_\ell B_\ell A_{\ell+1}^*$ as shown in Figure 1.24. The triangles $A_i B_i A_{i+1}$ are translated copies of an equilateral triangle $A_1 B_1 A_2$ for $1 \leq i \leq \ell$. Vertices $A_1^*$ and $A_{\ell+1}^*$ are below the line $A_1 A_{\ell+1}$ at distance $\varepsilon$ with a sufficiently small $\varepsilon > 0$.

Consider the inscribed disks of the triangles $A_i B_i A_{i+1},\ i = 1, 2, \ldots, \ell$. Scale each disk from the vertex $B_i$ of triangle $A_i B_i C_i$ such that for the resulting disks $D_i$, $\operatorname{dist}(D_i, A_i A_{i+1}) = \eta$ with an $\eta > 0$ small enough.

Let $\eta$ be so small that the disk $D_i$ in $A_i B_i A_{i+1}$ can be seen from $A_i$ and $A_{i+1}$ under an angle $\frac{2}{3}60^\circ + \frac{1}{3}\left(\left\lceil \frac{60^\circ}{\alpha} \right\rceil - 1\right)\alpha < 60^\circ$. Note that parameter $\varepsilon$ is independent of $\eta$.

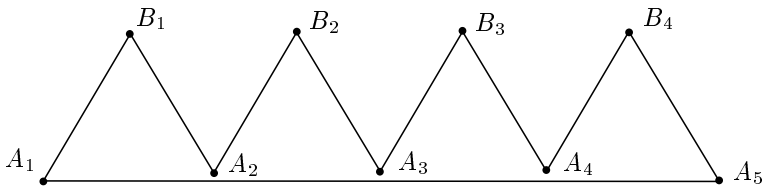


**Figure 1.24:** *Polygon $P$.*

For every $i = 2, 3, \ldots, \ell$, let $A_i^-$ (resp., $A_i^+$) be the point where the line tangent to $D_i$ (resp., to $D_{i-1}$) through $A_i$ intersects $A_1^* A_{\ell+1}^*$ (see Fig. 1.25).

Observe that some part of the disk $D_i$ is visible within $P$ from pentagon $A_i^- A_i B_i A_{i+1} A_{i+1}^+$. If $\varepsilon$ is sufficiently small, then $A_i^-$ and $A_i^+$ are in an arbitrarily small neighborhood of $A_i$. Let $\varepsilon > 0$ be so small that the disk $D_i \subset A_i B_i A_{i+1}$ is visible from points $A_i^-$ and $A_{i+1}^+$ (in the plane and not in polygon $P$) under an angle $\frac{1}{3}60^\circ + \frac{2}{3}\left(\left\lceil \frac{60^\circ}{\alpha} \right\rceil - 1\right)\alpha < 60^\circ$.

Both $D_i$ and $D_{i+1}$ are visible from triangle $A_{i+1}^- A_{i+1} A_{i+1}^+$. If $\varepsilon$ is small enough, then an $\alpha$-guard in $A_{i+1}^- A_{i+1} A_{i+1}^+$ cannot see both $D_i$ and $D_{i+1}$. Now every $\alpha$-guard monitors part of at most one disk $D_i$, it is enough to prove that the number of $\alpha$-guards required to monitor a disk $D_i$ is at least $\left\lceil \frac{60^\circ}{\alpha} \right\rceil$.

We can repeat the argument of the previous corollary. Suppose that $k$ $\alpha$-guards can monitor $D_i$. Consider and fix a placement of $k$ such $\alpha$-guards.

For the disk $D_i$, let $C_i$ be a concentric disk with $A_i^-$ and $A_{i+1}^+$ on its boundary. $C_i$ contains pentagon $A_i^- A_i B_i A_{i+1} A_{i+1}^+$. Now, according to Theorem 1.22,

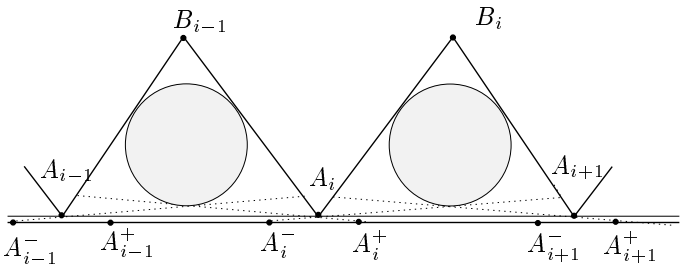


**Figure 1.25:** *Auxiliary disks, lines, and points.*

$k \cdot \alpha \geq \frac{1}{3}60^\circ + \frac{2}{3}\left(\left\lceil \frac{60^\circ}{\alpha} \right\rceil - 1\right)\alpha$. That is, $k \cdot \alpha > \left(\left\lceil \frac{60^\circ}{\alpha} \right\rceil - 1\right)\alpha$, and $k \geq \left\lceil \frac{60^\circ}{\alpha} \right\rceil$. □

## 1.6 Simple polygons with vertex $\pi$-guards

We describe the locations and the orientations of the vertex $\pi$-guards in a simple polygon $P$ establishing Theorems 1.2 and 1.3. Our allocation of vertex $\pi$-guards is based on a pseudo-triangulation [41, 40] or geodesic triangulation [57, 10] of the simple polygon $P$. A pseudo-triangulation decomposes $P$ along non-crossing diagonals into *pseudo-triangles*. It is easy to see (it is shown in Proposition 1.12) that a pseudo-triangle with $\ell$ vertices can always be monitored with at most $\lfloor (2\ell-3)/3 \rfloor$ vertex $\pi$-guards. Note though, that this does not immediately imply that an upper bound of $2n/3 + O(1)$ for the total number of guards required by $P$: Adjacent pseudo-triangles of the decomposition share two vertices and optimal guard allocations in two pseudo-triangles sharing a vertex $v$ can prescribe contradicting orientations for a guard $(v, H_v)$. As we will see in Section 1.6.2, this kind of conflict can be resolved by using a special type of pseudo-triangulation, namely a pulling pseudo-triangulation, where every vertex is either a convex vertex of $P$ or a reflex vertex of one of its adjacent pseudo-triangles.

The rest of this section is organized as follows. We first recall some definitions and give some basic facts concerning pseudo-triangles and pseudo-triangulations. We then proceed to prove Theorem 1.3 in Subsection 1.6.2. Subsection 1.6.3 illustrates how a single pseudo-triangle with $\ell$ vertices can be guarded with $\lfloor \ell/2 \rfloor$ vertex $\pi$-guards on its boundary. In Subsection 1.6.4, we prove Theorem 1.2 using a technical lemma whose proof is postponed to Subsection 1.6.5.

### 1.6.1 Basics on pseudo-triangles

A *pseudo-triangle* $T$ is a simple polygon with exactly three convex vertices, called *corner*s. If $a, b$, and $c$ are the corners of $T$ in counter-clockwise order, then we denote by $\overline{ab}$ (and similarly by $\overline{bc}$ and $\overline{ca}$) the chain of reflex vertices between $a$ and $b$ in counterclockwise direction on the boundary of $T$. For a simple polygon $P$, a *pseudo-triangulation* is a set $D$ of non-overlapping pseudo-triangles such that $P = \bigcup D$ and vertices of pseudo-triangles are vertices of $P$. Since triangles are also pseudo-triangles, any triangulation of a simple polygon is also a pseudo-triangulation.

We consider exclusively a special kind of pseudo-triangulations, namely minimum pseudo-triangulations. A *minimum pseudo-triangulation* is a pseudo-triangulation with the minimum number of pseudo-triangle faces. Equivalently, minimum pseudo-triangulations have the property [44], that every vertex of the polygon $P$ is either convex in $P$ or reflex in a pseudo-triangle $T \in D$. This property is also called *pointedness*. For the remainder of the paper we will use the term pseudo-triangulation to refer to a minimum pseudo-triangulation.

Every simple polygon has at least one pseudo-triangulation, but this pseudo-triangulation is not necessarily unique. One possible method to obtain a pseudo-triangulation of a simple polygon is to add non-crossing diagonals as long as they do not violate the pointedness property. Another approach is to dissect $P$ along non-crossing *geodesic paths* connecting convex vertices [57, 10]. Pseudo-triangulations are also referred to as *geodesic triangulation*s, since every pseudo-triangulation can be obtained in this way.

A pseudo-triangle in a pseudo-triangulation is uniquely determined by its three corners; hence we use the (non-abusive) notation $abc$ for a pseudo-triangle with corners $a, b$, and $c$. The number of pseudo-triangles in a pseudo-triangulation of a simple polygon $P$ is the number of convex vertices of $P$ minus two.

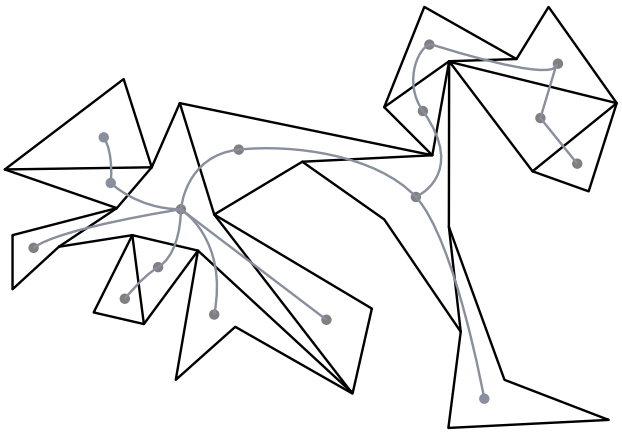

**Figure 1.26:** *A pseudo-triangulation and its dual graph for a simple polygon.*

The *dual graph* $G(D)$ of a pseudo-triangulation $D$ is defined by choosing $D$ as

the node set of $G(D)$ and connecting two nodes by an edge iff the corresponding pseudo-triangles have a common edge (a diagonal of $P$). $G(D)$ is always a tree if $D$ is a pseudo-triangulation of a simple polygon (see Fig. 1.26). Note that in our terminology, polygons have *vertices* and *sides*; graphs have *nodes* and *edges*.

For every pseudo-triangulation $D$ of $P$, there is a convex vertex $x \in V(P)$ such that $x$ is a vertex (a corner) of a unique pseudo-triangle $T \in D$. Let $y \in V(P)$ be a vertex adjacent to $x$. We denote by $G(D)_{xy}$ the rooted tree $G(D)$ where the root $R$ corresponds to the pseudo-triangle containing side $xy$. We write $S < T$ for two pseudo-triangles $S, T \in D_{xy}$ if $T$ is the ancestor of $S$ in $G(D)_{xy}$. For example, $S < R$ for all pseudo-triangles $S \neq R$. As a shorthand, we denote by $D_{xy}$ the pseudo-triangulation $D$ together with the partial order of ancestor relationships in the tree $G(D)_{xy}$.

For a reflex vertex $v$ of a pseudo-triangle we consider only vertex $\pi$-guards which are *inward-facing* (as defined above) or *outward-facing*, that is, the sides of $P$ incident to $v$ are in the closed halfplane $H_v$ (see also Fig. 1.1). For a corner $v$, we consider only two possibilities: either $(v, H_v)$ completely covers the angular domain, or it does not.

**Proposition 1.12.** *A pseudo-triangle $T$ with $\ell$ vertices can be monitored by at most $\lfloor (2\ell-3)/3 \rfloor$ vertex $\pi$-guards: one at a corner and at most $\lfloor 2(\ell-3)/3 \rfloor$ outward-facing guards at reflex vertices along two chains.*

*Proof.* Observe that one guard at a corner and outward-facing guards along the two adjacent chains can monitor the complete pseudo-triangle. For the three corners, this gives three possible guard allocations jointly using a total of $3+2(\ell-3) = 2\ell-3$ guards. This implies that at least one guard allocation uses at most $\lfloor (2\ell-3)/3 \rfloor$ guards (see Fig. 1.27). □

Using the same idea, we can deduce an upper bound of $\lfloor (\ell-1)/2 \rfloor$ for so-called *monotone mountains* – defined by J. O'Rourke [34]. A *mountain* is a pseudo-trianglewith at least two adjacent corners (i.e., one of the chains of reflex vertices is empty).

**Proposition 1.13.** *A pseudo-triangle $abc$ with $\ell$ vertices and $\overline{bc} = \emptyset$ can be monitored by at most $\lfloor (\ell-1)/2 \rfloor$ vertex $\pi$-guards: one at the corner $b$ or $c$ (but not at both) and at most $\lfloor (\ell-3)/2 \rfloor$ outward-facing guards at reflex vertices along a chain different from $\overline{bc}$.*

*Proof.* Consider the two guard allocations which put one $\pi$-guard at corner $b$ (resp., $c$) and outward-facing $\pi$-guards at vertices along the adjacent chain $\overline{ba}$

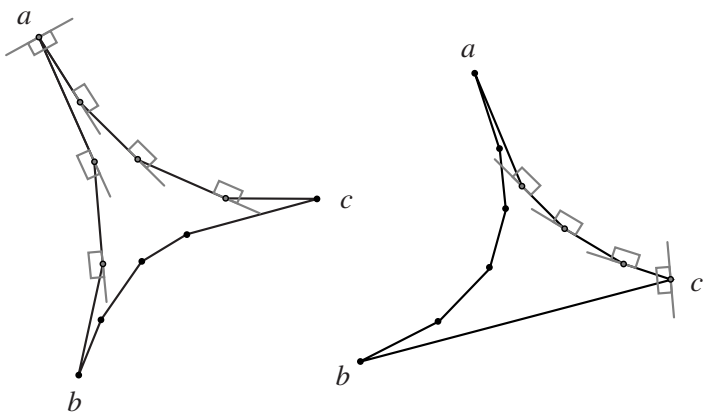


**Figure 1.27:** *Guards at a corner and along two adjacent chains monitor the complete pseudo-triangle.*

(resp., $\overline{ca}$). Both monitor $T$ and jointly use a total of $\ell - 1$ guards (see Fig. 1.27 (right)). □

### 1.6.2 Edge-aligned $\pi$-guards

We now proceed to prove Theorem 1.3.

*Proof (of Theorem 1.3).* Following, and generalizing the idea of Proposition 1.12, we describe three different guard allocations for $P$ that jointly use a total of $2n - k$ edge-aligned vertex $\pi$-guards. We define these allocations with resepct to an arbitrary but fixed pseudo-triangulation $D_{xy}$. In all three guard allocations and for each pseudo-triangle $T \in D_{xy}$, there will be vertex $\pi$-guards at one corner of $T$ and along the two chains adjacent to that corner such that these guards collectively monitor $T$. The three guard allocation schemes will jointly use all three corners (and thus different pairs of chains) for each $T \in D_{xy}$.

Consider the pseudo-triangle $T_1$ at the root of $G(D)_{xy}$. We can guard $T_1$ in three different ways by choosing any one of its three corners and placing guards at that corner (the *guarded corner* $gc(T_1)$ of $T_1$) and along its two adjacent chains (see Fig. 1.27). Every choice of guard allocation for $T_1$ induces a direction on its chains: The sides on the chains adjacent to $gc(T_1)$ are directed *away* from $gc(T_1)$, the sides on the remaining chain have *no* direction. Note

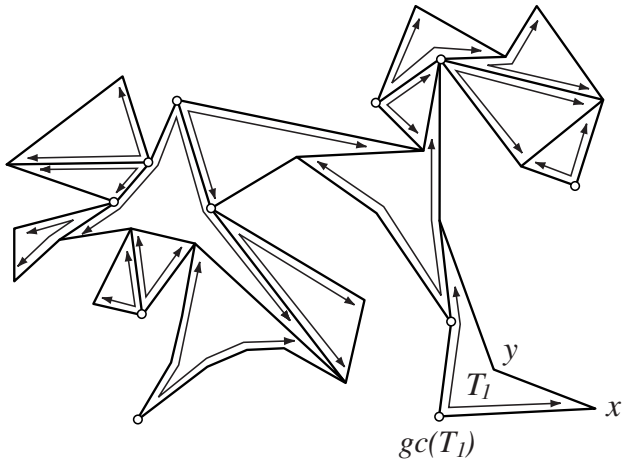


**Figure 1.28:** *Propagating the side directions along $G(D)_{xy}$.*

that we do indeed distinguish three directions for each side – the two standard directions and *no* direction. Fixing the directions for one chain of a pseudo-triangle uniquely determines the directions for the remaining ones. Furthermore, each side of the pseudo-triangulation can only have one direction, i.e., a pseudo-triangle inherits its directions from its parent via their joined side. Therefore, choosing a guarded vertex $gc(T_1)$ in $T_1$ not only induces a unique direction on the sides of $T_1$, but actually determines the direction of all sides of $D_{xy}$ (see Fig. 1.28). Furthermore, since each pseudo-triangle $T$ now has exactly one corner with two outgoing edges, we define this corner to be the guarded corner $gc(T)$. Finally, note that for each corner $v$ of every $T \in D_{xy}$ there is a choice of the guarded corner of $T_1$, such that the resulting orientation of $D_{xy}$ implies that $v$ is the guarded corner of $T$.

After choosing $gc(T_1)$ and propagating the directions along $G(D)_{xy}$ we place in a second step vertex $\pi$-guards along the directed edges of $D_{xy}$. More specifically, in each $T \in D_{xy}$ we place a vertex $\pi$-guard at the corner $gc(T)$ and *outward-facing* $\pi$-guards along the two directed chains adjacent to $gc(T)$ (see Fig. 1.29). By Property 1.12, these guards collectively monitor $T$. Note that it is always sufficient to place at most one $\pi$-guard at each vertex since every vertex is a reflex vertex of at most one pseudo-triangle of $D_{xy}$ and $\pi$-guards at reflex vertices are outward-facing. (This already gives an upper bound of $\lfloor (2n-k)/3 \rfloor$ in the general model.)

In the last step, we rotate every vertex $\pi$-guard $(v, H_v)$ at every reflex vertex $v$

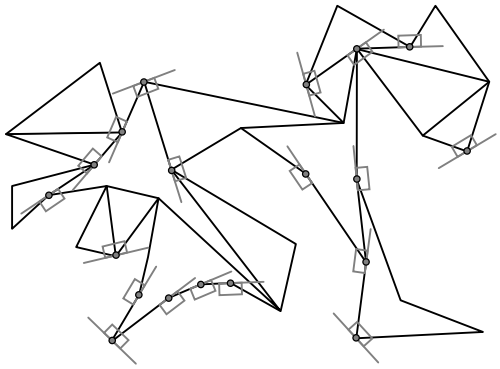

**Figure 1.29:** *Guard allocation according to the vertex assignment depicted in Fig. 1.28 after the second step.*

until it is edge-aligned. Recall that $v$ is a reflex vertex of exactly one $T_v \in D_{xy}$ and its guard monitors $T_v$ only if $v$ is on a chain adjacent to $gc(T_v)$. Rotate $(v, H_v)$ so that the area monitored in $T_v$ increases. The resulting set of vertex $\pi$-guards still monitors each $T \in D_{xy}$: The $\pi$-guards at reflex vertices of a $T \in D_{xy}$ monitor at least as much area as before, and the $\pi$-guards at $gc(T)$ still monitor the angular domain of $gc(T)$ in $T$, by the choice of $gc(T)$ (see Fig. 1.30).

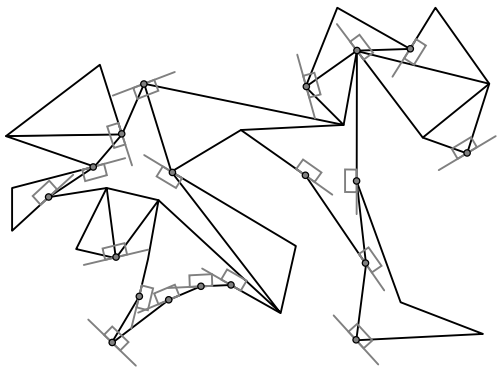

**Figure 1.30:** *Guard allocation according to the vertex assignment depicted in Fig. 1.28 after rotation.*

Finally, we show that the total number of guards is $2n - k$. We use $2\ell_1 - 3$ guards on the boundary of the pseudo-triangle $T_1$ corresponding to the root of

$D_{xy}$, where $\ell_1$ is the number of vertices of $T_1$. For all other $T_i \in D_{xy}$ with $\ell_i$ vertices, two guards out of the $2\ell_i - 3$ are already counted on the boundary of the ancestor of $T_i$. (This can be easily checked for all possible mutual positions of two adjacent pseudo-triangles: i.e., the common vertices are both corners, both reflex, or one is a corner and the other one is reflex.) Hence all three guard allocations use a total of $2 + \sum_{i=1}^{p}(2\ell_i - 5) = 2 - 5p + 2\sum_{i=0}^{p}\ell_i$ guards where $p$ denotes the number of of pseudo-triangles, i.e., $p = k - 2$. Using the fact that $n - 2 = \sum_{i=1}^{p}(\ell_i - 2) = (\sum_{i=1}^{p}\ell_i) - 2p$, we obtain $2n - 2 - p = 2n - k$ guards in total. □

### 1.6.3 $\pi$-guards in one pseudo-triangle

**Lemma 1.7.** *A pseudo-triangle with $\ell$ vertices can be monitored by $\lfloor \ell/2 \rfloor$ vertex $\pi$-guards such that there is at most one inward-facing $\pi$-guard, which lies on a prescribed chain, and all other $\pi$-guards are either located at a convex vertex or are outward-facing.*

*Proof.* We propose two possible guard allocations for a pseudo-triangle $abc$ with a total of at most $\ell$ guards, such that all $\pi$-guards at reflex vertices are outward-facing, except for at most one guard on the prescribed chain $\overline{bc}$. We consider two cases:

**Case 1:** There is a vertex $v \in \overline{bc}$ visible from $a$.

One guard allocation uses an inward-facing guard $(v, (va)^+)$ and outward-facing guards on the sub-chain $\overline{bv}$ monitoring the pseudo-triangle $abv$, combined with one guard at $a$ and outward-facing guards along $\overline{ac}$ monitoring $avc$. The other allocation uses an inward-facing guard $(v, (va)^-)$ and outward-facing guards on the sub-chain $\overline{vc}$ monitoring the pseudo-triangle $avc$, combined with one guard at $a$ and outward-facing guards along $\overline{ab}$ monitoring $abv$ (see Fig. 1.31). The two allocations jointly use a total of $\ell$ guards. Note that the guard at $v$ van be rotated inward facing w.r.t. $abc$.

**Case 2:** There is no vertex on $\overline{bc}$ visible from $a$.

Part of the relative interior of a side $uv$, $u, v \in \overline{bc} \cup \{b, c\}$, is visible from $a$. Denote by $g(u)$ (resp., $g(v)$) the first vertex on $\overline{ba}$ (resp., $\overline{ca}$) visited by the geodesic paths connecting $u$ to $a$ (resp., $v$ to $a$). The ray $\overrightarrow{ug(u)}$ (resp., $\overrightarrow{vg(v)}$) hits the polygonal chain connecting $a$ and $c$ at point $h(u)$ (resp., $h(v)$ on the polygonal chain connecting $a$ and $b$). The segment $uh(u)$ (resp., $vh(v)$) dissects $abc$ into three pseudo-triangles $T_u(a), T_u(b)$, and $T_u(c)$ containing the

vertices $a, b$, and $c$ respectively. Note that $T_u(b)$ (resp., $T_v(c)$) can be empty if $b = u$ (resp., $c = v$). One guard allocation uses an inward-facing guard $(u, (uv)^+)$, outward-facing guards along $\overline{uc}$ monitoring $T_u(c)$, and outward-facing guards along $\overline{ab}$ monitoring $T_u(a)$ and $T_u(b)$. The other guard allocation uses an inward-facing guard $(v, (uv)^+)$, outward-facing guards along $\overline{vb}$ monitoring $T_u(b)$, and outward-facing guards along $\overline{ac}$ monitoring $T_u(a)$ and $T_u(c)$ (see Fig. 1.32). The two allocations jointly use a total of $\ell - 1$ guards. □

Note the differences between Cases 1 and 2. In Case 1 there is always a $\pi$-guard at corner $a$, and no guard at $b$ nor at $c$; one allocation uses at most $\lfloor \ell/2 \rfloor$ guards. In Case 2 there is never a guard at $a$, and there is a guard at $b$ or at $c$ only if $b = u$ or $c = v$; one guard allocation uses at most $\lfloor (\ell - 1)/2 \rfloor$ guards. Furthermore, Case 2 also contains the proof of Proposition 1.13 with $b = u$ and $c = v$.

### 1.6.4 Upper bound for the general model

To establish the upper bound in the general model, we rely on a special pseudo-triangulation of $P$, which corresponds to the so-called *pulling triangulation* of convex polygons. A *pulling pseudo-triangulation* of a simple polygon is obtained by choosing (i.e., *pulling*) a convex vertex $x$ and connecting it to all other convex vertices of $P$ via geodesic paths (see Fig. 1.33). The pulling pseudo-triangulation has the special property that every pseudo-triangle $T \in D_{xy}$, with the exception of the root, has at least one corner adjacent to its parent, i.e., at least one of the two vertices which $T$ shares with its parent is a corner of $T$. We are now ready to present the proof of Theorem 1.2:

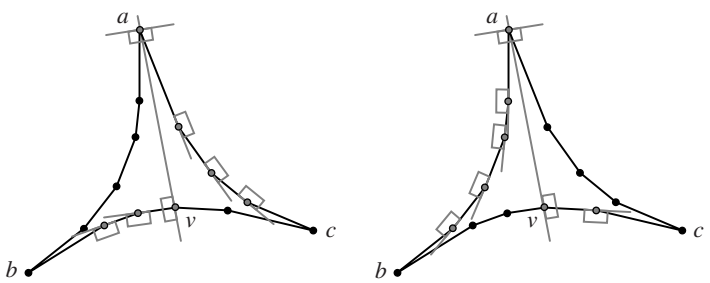


**Figure 1.31:** *The two guard allocations in Case 1.*

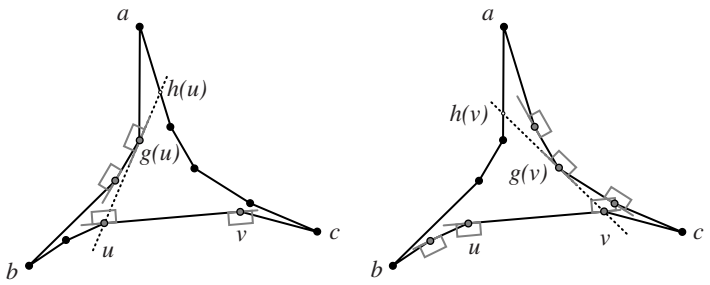


**Figure 1.32:** *The two guard allocations in Case 2.*

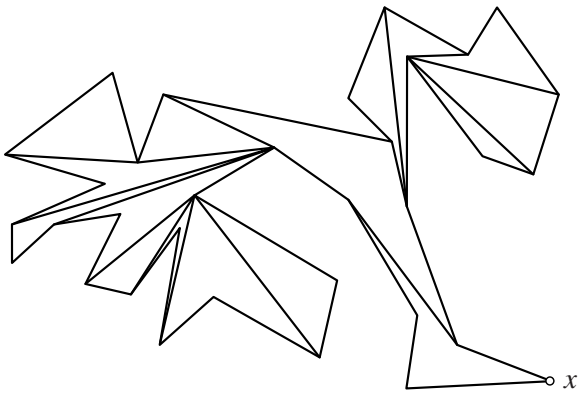


**Figure 1.33:** *A pulling pseudo-triangulation with respect to $x$.*

*Proof (of Theorem 1.2).* The basic idea of the proof is to construct a partition $\mathcal{S}$ of the vertices $V_P$ of $P$ and to place vertex $\pi$-guards at at most $|S|/2$ vertices for each $S \in \mathcal{S}$. It follows that the total number of vertex $\pi$-guards will be at most $\lfloor n/2 \rfloor$.

In order to define the partition $\mathcal{S}$, consider an ordered pulling pseudo-triangulation $D_{xy}$. Let $V_T$ denote the vertex set of a pseudo-triangle $T \in D_{xy}$. We assign a subset $S_T \subseteq V_T$ to each pseudo-triangle $T \in D_{xy}$ such that $\mathcal{S} := \{S_T : T \in D_{xy}\}$ is a partition of $V_P$. Then we place vertex $\pi$-guards at some vertices $R_T \subset S_T$.

The pseudo-triangles of $D_{xy}$ are processed from bottom to top, that is, $T \in D_{xy}$ is processed when all its descendants have already been processed. Pro-

cessing $T \in D_{xy}$ involves the following three tasks:

1. defining $S_T \subseteq V_T$
2. specifying $R_T \subset S_T \subseteq V_T$ with $|R_T| \leq \frac{1}{2}|S_T|$
3. verifying that
   (a) the vertex $\pi$-guards placed on the boundary of $T$ (not only those in $R_T$) can be oriented such that they collectively monitor $T$
   (b) for each vertex $v$, the orientation of a $\pi$-guard at $v$ is consistent in every $T' \in D_{xy}$ adjacent to $v$

The consistency of the orientation relies heavily on the pointedness of our pseudo-triangulation: a $\pi$-guard at a vertex $v$ can monitor all convex angular domains adjacent to $v$ and at the same time it can be an outward-facing $\pi$-guard with several possible orientations in a pseudo-triangle where $v$ is a reflex vertex. This means that if we include at one step a vertex $v$ in $R_T$ (andplace a guard at $v$) where $v$ is a corner in $T$ then we are free to choose the orientation of the $\pi$-guard at $v$ whenever we are processing an ancestor $T'$ of $T$ in which $v$ is a reflex vertex as long as the guard remains outward-facing in $T'$.

Now let $T = abc \in D_{xy}$ and suppose that all descendents $T' < T$ of $T$ have already been processed. Let $K_T = \{v \in V_T : \nexists\, T' < T, v \in S_{T'}\}$ be the set of vertices of $T$ which are not assigned to any descendent of $T$. Since $D_{xy}$ is a pulling pseudo-triangulation, we can assume that one of the corners of $T$, say $a$, is adjacent to its parent (or $a = x$ if $T$ is the root of $D_{xy}$). Let us assume further that $s \in \overline{ab} \cup \{b\}$ is the other vertex that $T$ shares with its parent (or $a = y$ if $T$ is the root of $D_{xy}$). Let $\tilde{K}_T = K_T \setminus \{a, s\}$.

Since none of the elements of $\tilde{K}_T$ is a vertex of any ancestor of $T$ it follows that $\tilde{K}_T \subseteq S_T$ has to hold. Furthermore, $S_T$ cannot contain vertices of $T$ that have already been assigned to any of the previously processed pseudo-triangles, that is $S_T \subseteq K_T$. This implies that the set $S_T \in \mathcal{S}$ of vertices assigned to $T$ has to be one of $\tilde{K}_T,\ \tilde{K}_T \cup \{a\},\ \tilde{K}_T \cup \{s\}$, and $K_T$.

In order to ensure that $T$ is guarded by the guards placed on its boundary, we establish the following invariant:

**Invariant 1.1.** *For every $v \in V_T \setminus K_T$ (i.e., for every vertex $v \in V_T$ that has previously been assigned to an ancestors of $T$), if $v \in R_{T'}$, $T' < T$ (i.e., there is guard placed at $v$), then the guard at $v$*

- *is outward-facing in $T$ if $v$ is a reflex vertex of $T$;*
- *covers the angular domain of $v$ in $T$ if $v$ is a corner of $T$.*

Invariant 1.1 clearly holds as long as no pseudo-triangle has been processed, i.e., as long as no guard has been placed. In order to maintain this invariant while processing a pseudo-triangle $T$ it is sufficient to respect the following two rules: $(i)$ if $a$ or $s$ is in $S_T$, then it is also in $R_T$, and $(ii)$ if $s \in S_T$ (which implies $s \in R_T$) and $s$ is reflex in $T$ then the guard at $s$ is outward-facing.

The following Lemma guarantees that we can process each pseudo-triangle $T \in D_{xy}$ while maintaining Invariant 1.1:

**Lemma 1.8.** *Assume we are given a pseudo-triangle $T = abc$ and a partition of its vertex set $V_T = K_T \cup^* L_T$. Denote by $s$ the vertex next to $a$ in $\overline{ab} \cup \{b\} \subset V_T$. Then $T$ can be monitored by vertex $\pi$-guards at a set of vertices $U_T \subset V_T$ such that*

- *$\pi$-guards at reflex vertices of $L_T \cup \{s\}$ are outward-facing,*
- *there is at most one inward-facing guard and it is located at $\overline{bc} \cap K_T$,*
- *at most $\lfloor |K_T|/2 \rfloor$ guards are placed at vertices of $K_T$, in particular if $|K_T|/2$ guards are placed, then there is a guard placed at both $a$ and $s$, and if $(|K_T| - 1)/2$ guards are placed, then there is a guard placed at at least one of $a$ and $s$.*

We postpone the proof of Lemma 1.8 to Subsection 1.6.5.

Now we are ready to process the pseudo-triangle $T$. Recall that at this point all descendants of $T$ have already been processed and Invariant 1.1 holds. We apply Lemma 1.8 with $L_T = V_T \setminus K_T$. This yields a guard allocation which places guards at the set of vertices $U_T \subset V_T$. Let $S_T = \tilde{K}_T \cup (\{a, s\} \cap U_T)$ and $R_T = K_T \cap U_T$. Lemma 1.8 and Invariant 1.1 imply that $T$ is now indeed guarded by consistently oriented $\pi$-guards placed on its boundary. Furthermore the orientations for the guard specified in Lemma 1.8 assure that Invariant 1.1 is maintained.

It remains to show that $|R_T| \leq |S_T|/2$. According to Lemma 1.8 there are at most $\lfloor |K_T|/2 \rfloor$ guards placed at vertices of $K_T$, that is $|R_T| = |U_T \cap K_T| \leq |K_T|/2$. If $|R_T| = |K_T|/2$ then $\{a, s\} \subseteq U_T$ and therefore $S_T = \tilde{K}_T \cup (\{a, s\} \cap U_T) = K_T$. If $|R_T| = (|K_T| - 1)/2$ then $a \in U_T$ or $s \in U_T$ and therefore $S_T = \tilde{K}_T \cup \{a\}$, $S_T = \tilde{K}_T \cup \{s\}$, or $S_T = \tilde{K}_T \cup \{a, s\}$, i.e., $|S_T| \geq |K_T| - 1$. Finally, if $|R_T| \leq (|K_T| - 2)/2$ then $|S_T| \geq |\tilde{K}_T| = |K_T| - 2$. □

### 1.6.5 Proof of Lemma 1.8

Before we begin the actual proof of Lemma 1.8 we formulate an additional Lemma which is a generalization of Proposition 1.13 as well, but slightly

weaker than Lemma 1.8 itself. It will be used as a building block during the rest of this section.

**Lemma 1.9.** *Assume we are given a mountain $T = xyz$ with $\overline{yz} = \emptyset$ and a partition of its vertex set $V_T = K_T \cup^* L_T$.*

**(1)** *If $\{x\} \cup \overline{xy} \cup \overline{xz} \subseteq L_T$ then $T$ can be monitored by vertex $\pi$-guards at a subset of $V_T$ such that*

- *$\pi$-guards at reflex vertices are outward-facing,*
- *we place no guard at vertices of $K_T$.*

**(2)** *If $x \in L_T$ and at least one of $\overline{xy} \cap K_T$ and $\overline{xz} \cap K_T$ is non-empty then $T$ can be monitored by either one of two guard allocations $A_2(y) \subset V_T$ and $A_2(z) \subset V_T$ such that*

- *$\pi$-guards at reflex vertices are outward-facing,*
- *$A_2(y)$ places a guard at $y$ but no guard at $z$, while $A_2(z)$ places a guard at $z$ but no guard at $y$,*
- *$A_2(y)$ and $A_2(z)$ jointly place $|K_T| - 1$ guards at vertices of $K_T$.*

**(3)** *If $x \in K_T$ then $T$ can be monitored by either one of two guard allocations $A_3(y) \subset V_T$ and $A_3(z) \subset V_T$ such that*

- *$\pi$-guards at reflex vertices are outward-facing,*
- *$A_3(y)$ places a guard at $y$ but no guard at $z$, while $A_3(z)$ places a guard at $z$ but no guard at $y$,*
- *$A_3(y)$ and $A_3(z)$ jointly place $|K_T| - 1$ guards at vertices of $K_T$.*

*Proof.* **(1)** If $\overline{xy} \subseteq L_T$ and $\overline{xz} \subseteq L_T$ (i.e., $K_T$ is one of $\{y\}$,$\{z\}$, and $\{y, z\}$) then one guard at $x$ and outward-facing guards along the chains $\overline{xy}$ and $\overline{xz}$ monitor $T$ using no guard at $K_T$.

**(2)** If $\overline{xy} \cap K_T$ or $\overline{xz} \cap K_T$ is non-empty then let $y_0 = y, y_1, y_2, \ldots, y_p$ and $z_0 = z, z_1, z_2, \ldots, z_q$ be the vertices on the chains $\{y\} \cup \overline{yx}$ and on $\{z\} \cup \overline{zx}$. If $\overline{yx} \cap K_T \neq \emptyset$ then let $y_i$ be the closest vertices to $x$ in $\overline{yx} \cap K_T$ and, equivalently, if $\overline{zx} \cap K_T \neq \emptyset$ then let $z_j$ be the closest vertices to $x$ in $\overline{zx} \cap K_T$. If $y_i$ is defined then let $e$ denote the intersection of the ray $\overrightarrow{y_{i-1}y_i}$ and the polygonal chain between $z$ and $x$. Similarly, if $z_j$ is defined let $f$ denote the intersection of the ray $\overrightarrow{z_{j-1}z_j}$ and the polygonal chain between $y$ and $x$. Note that at least one of $e$ and $f$ is defined.

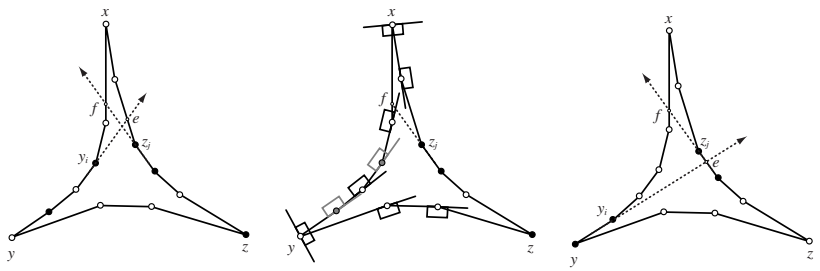


**Figure 1.34:** *Two possible partitions $K_T \cup^* L_T$ (left, right) and one possible guard allocation for the partition depicted on the left (middle).*

We may suppose w.l.o.g. that $f$ is defined. If $e$ is defined as well then we may also suppose that $f$ is between $y_i$ and $x$. The segment $z_j f$ dissects $T$ into two pseudo-triangles $xfz_j$ and $fyz$ as depicted in Figure 1.34.[2]

Both guard allocations $A_2(y)$ and $A_2(z)$ monitor the pseudo-triangle $xfz_j$ by one $\pi$-guard at $x$ and outward-facing $\pi$-guards along the two adjacent chains $\overline{xf}$ and $\overline{xz_j}$, that is, using no guards at vertices of $K_T$. $A_2(y)$ monitors the pseudo-triangle $fyz$ by one guard at corner $y$ and outward-facing $\pi$-guards at the adjacent chain $\overline{yf}$. Analogously $A_2(z)$ monitors $fyz$ by one guard at corner $z$ and outward-facing $\pi$-guards at the adjacent chain $\overline{zz_j}$.

Note that $z_j \in K_T$ is not a vertex of $fyz$ and so neither guard allocation places a guard at $z_j$. Since the two allocations use all other vertices of $K_T$ exactly once, they jointly use $|K_T| - 1$ guards at vertices of $K_T$.

**(3)** Following the allocation scheme described in Proposition 1.13, $A_3(y)$ places a guard at $y$ and outward-facing $\pi$-guards along the chain $\overline{yx}$. Analogously $A_3(z)$ places a guard at $z$ and outward-facing $\pi$-guards along the chain $\overline{zx}$. □

*Proof (of Lemma 1.8).* We distinguish three cases. Note that in all of the cases we assume that if $a \in L_T$ or $s \in L_T$ then there is a guard placed at $a$ or $s$ respectively.

**Case A**: $T$ is a mountain (this includes the case of $b = s$).

Since $T$ is a mountain there is at least one empty chain. Let $d$ denote the corner

[2] In Figure 1.34 and those that follow we depict the vertices of $K_T$ as hollow circles and the vertices of $K_T$ as filled circles. Similarly, guards placed at vertices of $K_T$ are drawn with black edges while guards placed at vertices of $K_T$ are drawn, as before, with grey edges.

opposite to such a chain. To guard $T$ we invoke Lemma 1.9 with $x = d$.

First let us note that all allocations of Lemma 1.9 use only outward-facing $\pi$-guards at reflex vertices.

Second, Lemma 1.8 clearly holds in the case of Lemma 1.9 (1). Otherwise one of the allocations uses at most $\lfloor(|K_T| - 1)/2\rfloor$ guards at vertices of $K_T$. If both allocations use exactly $(|K_T| - 1)/2$ guards, and both $a \in K_T$ and $s \in K_T$, then it remains to show that one allocation places a guard at $a$ or $s$: If $d = a$ then $A_3(y)$ places a guard at $s$. If $d \neq a$ then $a = y$ or $a = z$ and one of the allocations from Lemma 1.9 (2) or Lemma 1.9 (3) places a guard at $a$.

**Case B**: $T$ is not a mountain and there is a vertex $v \in K_T \cap \overline{bc}$ visible from $a$.

Note that this case corresponds to Case 1 of the proof of Lemma 1.7.

The segment $av$ dissects the pseudo-triangle $T = abc$ into two mountains $abv$ and $avc$. We guard $abv$ and $avc$ by invoking Lemma 1.9 with $x = b$ and $x = c$ respectively.

If $abv$ or $avc$ are handled by Lemma 1.9 (1), i.e., $\{b\} \cup \overline{bv} \cup \overline{ba} \subseteq L_T$ or $\{c\} \cup \overline{ca} \cup \overline{cv} \subseteq L_T$, then no guard from $K_T$ is necessary to monitor $abv$ or $avc$, respectively, and the guard allocation for $T \setminus abv$ or $T \setminus avc$ is therefore subsumed by Case A.

Otherwise we obtain the two guard allocations for $T$ by combining the two pairs of guard allocations for $abv$ and $avc$ described by Lemma 1.9 (2) or (3). Specifically, we combine allocation $A_i(a)$ for $abv$ and $A_j(v)$ for $avc$, $i, j \in \{2, 3\}$, and similarly allocation $A_i(v)$ for $abv$ and $A_j(a)$ for $avc$, $i, j \in \{2, 3\}$.

Both guard allocations place a $\pi$-guard at $a$ and at $v$. Lemma 1.9 (2) and (3) place only outward-facing $\pi$-guards at reflex vertices. Therefore, only the guard at $v$ is not outward-facing with respect to $T$ but it can always be oriented to be inward-facing.

The two guard allocations jointly use at most $|K_T \cap V_{abv}| - 1$ guards at vertices of $K_T \cap V_{abv}$ and at most $|K_T \cap V_{avc}| - 1$ guards at vertices of $K_T \cap V_{avc}$. Therefore, they use at most $|K_T \cap V_T| = |K_T|$ guards at vertices of $K_T$. It follows that one of them necessarily uses at most $|K_T|/2$ guards.

If both guard allocations for $T$ use exactly $|K_T|/2$ guards at vertices of $K_T$, then it remains to show that there always is a guard at $a$ and $s$. Recall that both allocations place a guard at $a$. Since $T$ is not a mountain, $s \neq b$ which implies that $s \in \overline{ab}$. If $b \in K_T$ then allocation $A_3(a)$ places a guard at $s$.

If $b \in L_T$ then it is possible that $s$ corresponds to $y_i$ or $z_j$ in the proof of Lemma 1.9 and so neither allocation places a guard at $s$. In this case $K_T \cap \overline{ab} =$

$\{s\}$. We can therefore replace $A_2(v)$ in our combination by an allocation that guards $abv$ by also using the guard at $a$ placed by the allocation $A_j(a)$ for $avc$, $j \in \{2, 3\}$, and outward-facing $\pi$-guards at $\overline{ab}$. This allocation uses at most a many guards as $A_2(v)$, since it does not place a guard at $v \in K_T$ but places a guard at $s$ instead.

**Case C**: $T$ is not a mountain and there is no vertex in $K_T \cap \overline{bc}$ visible from $a$.

Note that this case corresponds to Case 2 of the proof of Lemma 1.7.

Let $\overline{uv}$ be the largest subchain of $\overline{bc}$ such that $a$ sees all vertices of $\overline{uv}$. Necessarily $\overline{uv} \subseteq L_T$. Following the notation of Lemma 1.7, denote by $g(u)$ (resp., $g(v)$) the first vertex on $\overline{ba}$ (resp., $\overline{ca}$) visited by the geodesic path connecting $u$ to $a$ (resp., $v$ to $a$). Note that $s \in \overline{g(u)a} \cup \{g(u)\}$. The segments $ug(u)$ and $vg(v)$ dissect $abc$ into two mountains $bug(u)$ and $cg(v)v$, and a pseudo-triangle $auv$ (see Fig. 1.35). Note that $bug(u)$ (resp., $cg(v)v$) can be empty if $b = u$ (resp., $c = v$).

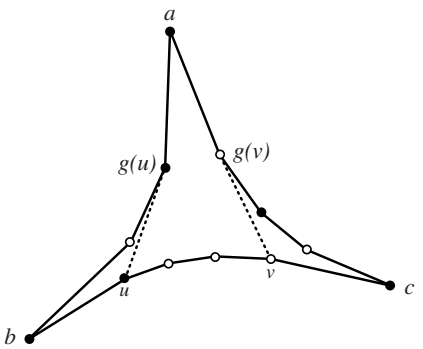


**Figure 1.35:** *The partition of $T$.*

If $bug(u)$ (resp., $cg(v)v$) is empty then we can simply remove it from all further consideration. If they are non-empty then we guard $bug(u)$ and $cg(v)v$ by invoking Lemma 1.9 with $x = b$ and $x = c$ respectively.

If $bug(u)$ or $cg(v)v$ are handled by Lemma 1.9 (1), i.e., $\{b\} \cup \overline{bu} \cup \overline{bg(u)} \subseteq L_T$ or $\{c\} \cup \overline{cg(v)} \cup \overline{cv} \subseteq L_T$, then no guard from $K_T$ is necessary to monitor $bug(u)$ or $cg(v)v$, respectively. In this case we can treat $T \setminus bug(u)$ (resp. $T \setminus vcg(v)$) as if $bug(u)$ (resp. $cg(v)v$) were empty.

Otherwise we define two guard allocations for $T$ that are a combination of the guard allocations for its parts. The first allocation guards $bug(u)$ by $A_i(u)$ and

$vcg(v)$ by $A_j(g(v))$, $i, j \in \{2, 3\}$, and it guards $auv$ by one guard at corner $v$ and outward facing $\pi$-guards at $\overline{vu} \subseteq L_T$ and $\overline{va}$. The second allocation guards $bug(u)$ by $A_i(g(u))$ and $vcg(v)$ by $A_j(v)$, $i, j \in \{2, 3\}$, and it guards $auv$ by one guard at corner $u$ and outward facing $\pi$-guards at $\overline{uv} \subseteq L_T$ and $\overline{ua}$.

Either allocation places guards at every vertex of $\overline{uv} \cup \{u, v\}$ and neither places a guard at $a$. Lemma 1.9 (2) and (3) place only outward-facing $\pi$-guards at reflex vertices. Observe that only the guard at either $v$ or $u$ (at one corner of $auv$) is not outward-facing with respect to $T$, but it can always be oriented to be inward-facing. The guards at $g(u)$ and $g(v)$ are always outwards-facing w.r.t. $T$.

The two guard allocations jointly use at most $|K_T \cap V_{bug(u)}| - 1$ guards at vertices of $K_T \cap V_{bug(u)}$ if $bug(u)$ is non-empty, and at most $|K_T \cap V_{cg(v)v}| - 1$ guards at vertices of $K_T \cap V_{cg(v)v}$ if $cg(v)v$ is non-empty. They jointly use $|(K_T \setminus \{a\}) \cap V_{auv}|$ guards at vertices of $(K_T \setminus \{a\}) \cap V_{auv}$. Therefore, they use at most $|(K_T \setminus \{a\}) \cap V_T| = |K_T \setminus \{a\}|$ guards at vertices of $K_T \setminus \{a\}$. It follows that one of them necessarily uses at most $|K_T|/2$ guards at vertices of $K_T$.

If both guard allocations for $T$ use exactly $|K_T \setminus \{a\}|/2$ guards at vertices of $K_T \setminus \{a\}$, then one of them places a $\pi$-guard at $s \in \overline{ag(u)} \cup \{g(u)\}$. So if both allocations use exactly $|K_T|/2$ guards at vertices of $K_T$, then necessarily $a \in L_T$, and we may suppose that there is a guard at $a$. □

# Obstacles in the plane

## 2.1 Introduction

Illumination of convex sets was first studied by L. Fejes Tóth [20] who proved that $4n - 7$ light sources are always enough and sometimes necessary to illuminate the boundaries of $n$, $n \geq 3$, disjoint compact sets in the plane. He also proved that $2n - 2$ light sources are always enough and sometimes necessary to illuminate the boundaries of $n$, $n \geq 3$, disjoint disks. Ever since, this worst case problem were studied for several subclasses of compact sets. The exact number of necessary light sources has not been known for any other subclass so far. For $n$ disjoint homothetic triangles, there are almost matching upper and lower bounds $n + 1$ and $n$ [14]. For $n$ disjoint rectilinear rectangles [25, 58], the maximum number of necessary light sources is between $n-1$ and $n + 1$ in the worst case. It is conjectured [58] that $n + c$ light sources always suffice to monitor $n$ disjoint convex quadrilaterals or $n$ disjoint triangles in the plane, $c$ is a constant.

Given a set $B$ of objects in the plane, a set $S$ of light sources collectively *illuminate* a point set $P$ if for every point $p \in P$ there is a light source $s \in S$ such that the open line segment $sp$ is disjoint from all objects in $B$. They *illuminate* the set $B$ of objects if they monitor the boundary of every object in $B$. This notion of illumination, was first studied by Valentine [60]. (We mention that Hadwiger [23] defined a different notion of illumination for a convex set, which we do not consider in this thesis. There, a light source $s$ illuminates a point $p \in \partial C$ on the boundary of a convex set $C \subset \mathbb{E}^d$ if the open line segment $sp$ is disjoint from the objects and the ray $\overrightarrow{sp}$ enters into the interior of $C$. It is a famous open problem to prove that $2^d$ light sources can always illuminate the boundary of one convex set. For a survey on Hadwidger's problem, we refer the reader to [32].)

Two problems can be distinguished: illuminating the obstacles and illuminating the *free space*, which is the complement of the union of the objects ($\mathbb{E}^2 \setminus \bigcup D$) in the presence of a set $D$ of disjoint obstacles. Interestingly, whenever the boundaries of every object in $D$ is a closed Jordan curve, then all known bounds equally apply to these two problems. This is true for the following theorem on triangular obstacles, too.

**Theorem 2.1.** *[54] Given a set $T$ of $n$ disjoint triangles in the plane ($\mathbb{E}^2$) in general position, one can place at most $\lfloor (5n+3)/4 \rfloor$ light sources in $\mathbb{E}^2 \setminus \bigcup T$ such that they collectively illuminate the every triangle and their free space.*

*General position* means that the triangles have no three collinear vertices. Illuminating the free space of polygonal objects is, in turn, related to a generalization of Chvátal's theorem to art galleries represented by polygons with holes [33, 58]. What is the minimal number of point guards (resp., vertex guards) that can collectively monitor any polygon with $h$ holes and a total of $n$ vertices? Hoffmann, Kaufmann, and Kriegel [26] and, independently, Bjorling-Sachs and Souvaine [7] gave two different algorithms to place $\lfloor (n + h)/3 \rfloor$ point guards so that they collectively monitor the gallery. There exist examples (a polygon with quadrilateral holes [58]) where at least $\lfloor (n+h)/3 \rfloor$ guards are necessary. The same lower bound applies to the cases where the guards are stationed at vertices, on the boundary, or at any point of the polygon.

The next theorem, which is intimately related to our Theorem 2.1, is the first non-trivial upper bound on the number of vertex guards in polygons with holes in an important special case [35, 58].

**Theorem 2.2.** *[54] A polygon $P$ with holes consisting of a triangular outer polygon and $k$ triangular holes can be monitored by at most $\lfloor (5k + 3)/4 \rfloor$ vertex guards.*

If the obstacles are open curves (and therefore their boundaries are not simple Jordan curves), then the illumination problem splits up to four distinct variants. We state them for a set $L$ of $n$ closed line segments in the plane and show progress on all variants. We suppose all through this section that the line segments are in general position: there are no three collinear segment endpoints and no two parallel segments.

**Problem 2.1.** *(Illuminating segments) Find a set $S \subset \mathbb{E}^2 \setminus \bigcup L$ of light sources such that they collectively illuminate $\bigcup L$.*

The maximum number of light sources required for any set of $n$ line segments is still not known. An asymptotics of $n/2 + O(1)$ was conjectured in[13]. The best known construction, due to Zaks [61], requires $4n/9 - 2$ light sources. We improve the best known upper bound $\lceil 2n/3 \rceil$ of Czyzowicz at al. [15].

**Theorem 2.3.** *[49] For Problem 2.1, $\lfloor (n+1)/2 \rfloor$ light sources are always sufficient.*

Another natural interpretation of illumination of segments was suggested by Welzl. We can require that every point of every given line segment is illuminated by at least two light sources, one on each side of the supporting line of the segment.

**Problem 2.2.** *(Illuminating segments from both sides) Find a set $S \subset \mathbb{E}^2 \setminus \bigcup L$ of light sources such that they collectively illuminate both sides of every $\ell \in L$. (A light source $s$ illuminates point $p \in \ell$ from one side, if $s$ is on one side of the line through $\ell$ and the open line segment $sp$ does not intersect any $\ell' \in L$.)*

We give exactly the minimum number of light sources required for Problem 2.2 for any set of pairwise disjoint line segments in general position.

**Theorem 2.4.** *[49, 50] For Problem 2.2, $\lfloor 4(n+1)/5 \rfloor$ light sources are always sufficient and sometimes necessary if $n \geq 2$.*

The problem of illumination of the *free space* of disjoint line segments can be formulated for two slightly different notion of illumination: For the definition we gave at the beginning of this section (strong illumination), and for a relaxation of the definition introduced by O'Rourke [33].

**Problem 2.3.** *(Illumination of the free space) Find a set $S \subset \mathbb{E}^2 \setminus \bigcup L$ of light sources such that they collectively illuminate $\mathbb{E}^2 \setminus \bigcup L$. A point $p \in \mathbb{E}^2 \setminus \bigcup L$ is illuminated by source $s$, if the open line segment $sp$*

*(a) does not intersect any $\ell' \in L$ (strong illumination);*

*(b) does not cross any $\ell' \in L$ (weak illumination).*

**Remark 2.1.** *O'Rourke [33] considered a variant of Problem 2.3 (b) where light sources may be placed on top of line segments. He showed that for $n > 5$, $\lfloor 2n/3 \rfloor$ light sources are always sufficient and sometimes necessary.*

Notice that if the free space is *weakly illuminated* (in the sense of Problem 2.3 (b)) then at most points on the lines through the segments are not *strongly illuminates* (in the sense of Problem 2.3 (a)). The fact that these problems are

really different is nicely indicated by the two results showing that they require different numbers of light sources in the worst case.

**Theorem 2.5.** *[49] For Problem 2.3 (a),* $\lfloor 4(n+1)/5 \rfloor$ *light sources are always sufficient and sometimes necessary if* $n \geq 2$ *.*

**Theorem 2.6.** *[55] For Problem 2.3 (b),* $\lfloor 2(n+1)/3 \rfloor$ *light sources are always sufficient and sometimes necessary.*

In our models, a line segment can be illuminated by several light sources, we do not require to have one light source $S$ for each line segment $\ell \in L$ illuminating every point of $\ell$. If such a constraint were imposed, then our theorems do not prevail. For this version of Problem 2.1, the best known upper bound due to Czyzowicz et al. [15] is $\lceil 2n/3 \rceil$, which is asymptotically tight.

A similar problem were considered by Jennings and Lenhart [27], where an arbitrary number of light sources can be placed to points of the line segments. The authors prove that there are always at most $\lfloor n/2 \rfloor$ segments from $L$ such that light sources placed at these segments can illuminate all the other segments, and this bound is tight.

The allocation of light sources in this chapter is based on the *matching technique* used in [33] for line segments where it yields tight bounds. We partition the *free space* $\mathbb{E}^2 \setminus \bigcup D$ of the set $D$ of disjoint objects into convex cells. We associate then a *dual graph* to the partition where the nodes represent convex cells, and two nodes are connected by an edge if a common light source can illuminate both cells, therefore the dual graph depend on the definition of illumination we use. We can show then that two or three adjacent cells can be illuminated by two or three light sources in different variants. A small covering set (or in many cases simply matching) of convex cells with common boundary points then gives a small set of light sources. This technique could possibly yield the best possible bound for $n$ disjoint triangles: In some cases, $\mathbb{E}^2 \setminus \bigcup T$ cannot be partitioned into less than $2n+1$ convex cells and no three cells have a common boundary. In this scenario $n+c$ light sources for disjoint triangle obstacles, the conjectured, would assume an almost perfect matching on the cells.

The key to our theorems about illuminating line segments is the fact that the *dual graph* of a convex partitioning of disjoint segments (defined below) can be decomposed into paths of 3, 4, or 5 nodes. Theorem 2.3 is based on a weaker property, i.e., the dual graph has a partition into paths of 2 or 3 nodes.

Kaneko [28] characterized $\{P_{\geq 3}\}$-factor graphs, that is, graphs which can be decomposed into paths of at least 3 nodes. Finding a maximum $P_3$ packing in a given graph is known to be NP-hard [30] even for planar graphs [6]. Recently, Kano, Katona, and Király [29] gave a simple proof to Kaneko's theorem, and Hartvigsen and Hell [24] presented a polynomial algorithm which finds an actual $\{P_{\geq 3}\}$-factor or states that none exists.

Section 2.2 discusses basic definitions and the convex partitions used in the proofs of our theorems. The theorems on disjoint line segments are shown in Section 2.3. An important lemma on the structure of auxiliary graphs is discussed in Section 2.4. Section 2.5 contains the proofs of Theorem 2.2. The proof of our Lemma 2.8, crucial for Theorem 2.2, and a modified convex partitioning of polygons with holes are presented in Section 2.5.1.

## 2.2 Convex partitionings and dual graphs

The proofs of Theorems 2.1 through 2.6 are all based on a decomposition of the free space of the obstacles into convex pieces. Convex decompositions are not always unique, we need to specify what partitioning we use. In this section we describe a simple method to partition the free space of disjoint segments and polygons with holes into convex pieces. This naive partitioning can directly be applied in the case of segments. As for polygons with holes, where there is more freedom in our algorithm, this partition will be the basis for another partition in Section 2.5.1.

We also define graphs on the convex partitions such that cells represented by adjacent nodes in the graph can be illuminated by one light source.

### 2.2.1 Line segments in the plane

Let $L$ be a set of $n$ disjoint line segments in the plane in general position. $\mathbb{E}^2 \setminus \bigcup L$ can be partitioned into convex closed cells by the following *convex partitioning* algorithm. Extend consecutively every line segment in both directions until it hits the relative interior of another line segment or a previously drawn extension (or to infinity if it does not hit anything). For a segment endpoint $v$, denote by $\beta(v)$ the extension of the segment beyond $v$. Starting with $n$ disjoint line segments, the plane is partitioned into $n + 1$ convex open regions [33].

The convex partitioning induced by $L$ is not unique. It depends on the order in which the segments are extended. However a simple algorithm may render it

unique by extending the line segments without generating intersection points. The *protruding algorithm* goes as follows.

Let $\delta > 0$ be a constant. Consider consecutively every endpoint $p$ of every segment $pq \in L$. If the ray $\overrightarrow{qp}$ hits another line segment (or a former extension) at point $r$ then extend $pq$ beyond $p$ by $\max(0, |pr| - \delta)$. If the ray $\overrightarrow{qp}$ does not hit any other (possibly extended) line segment, then extend $pq$ beyond $p$ by $1/\delta$. Denote the resulting set of extended line segments by $L^\delta$. Note that $L^\delta$ is not unique either, it depends, too, on the order in which the segment endpoints are considered.

A set $L$ of line segments is called *protruded*, if the convex partitioning induced by $L$ is unique. That is, no matter in which order the convex partitioning extends the segments, they are always extended until the same point or to infinity. Observe that if $L^\delta$ is protruded, then $L^{\delta'}$ is also protruded for any $0 < \delta' < \delta$. We show that for any $L$ there exists a $\delta$ and an $L^\delta$ such that $L^\delta$ is protruded set of segments in general position.

**Proposition 2.1.** *Let $P$ be the set of all the intersection points of lines induced by segments of $L$. If $\delta < \frac{1}{2} \cdot \min\{|pq| : p, q \in P\}$, and $1/\delta$ is bigger than the diameter of $P \cup L$, then the convex partitioning of any $L^\delta$ is unique.*

*Proof.* If $L$ is already protruded, then the proof is complete. If the convex partitioning for $L$ is not unique then there are two segments $p_1q_1$ and $p_2q_2$ such that both can be extended beyond the endpoints $q_1$ and $q_2$, resp., to a same point $x \in P$, $x \notin \bigcup L$, without crossing other line segments.

Consider the step of the protruding algorithm where $p_1q_1$ is extended beyond its endpoint $q_1$. Suppose that until this step no extended segment intersects $q_2x$. Denote by $w$ the last point of $P$ on the ray $\overrightarrow{p_1q_1}$. If $x = w$, then $|q_1x| < \operatorname{diam}(P \cup L)$, and the new endpoint $q_1'$ of the extension of $p_1q_1$ is beyond $x = w$. Therefore the partitioning algorithm extends $p_1q_1'$ beyond $q_1'$ to infinity. If $w \neq x$, then $\min\{|pq| : p, q \in P\} \leq |xw|$ and the new endpoint $q_1'$ of $p_1q_1$ is in the relative interior of the ray $\overrightarrow{xw}$, i.e., it is beyond $x$. Therefore the partitioning algorithm extends $p_1q_1'$ beyond $q_1'$ and beyond $x$, too. □

In our worst case analysis, we may suppose that $L$ is protruded. $\bigcup L \subset \bigcup L^\delta$, hence a light source can illuminate at most as many points in the presence of $L^\delta$ as in the presence of $L$.

The *graph $H$ of the convex partition* is defined as follows: Let the set $V(H)$ of nodes be the set of convex cells. For each endpoint $v$ of every segment $\ell$, we add an edge to the graph: we connect the two nodes corresponding to the two cells along the last portion of the extension of the segment beyond $v_1$.

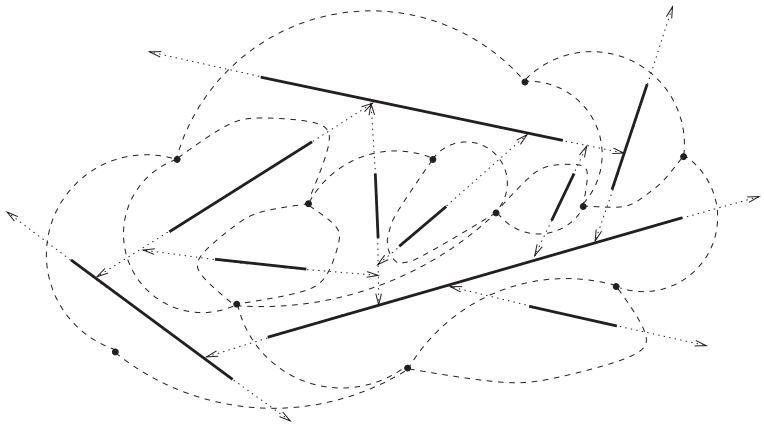

**Figure 2.1:** *A convex partitioning and the corresponding planar graph $H$.*

The graph $H$ has exactly $2n$ edges, since every segment endpoint gives rise to exactly one edge of $H$. $H$ may have double edges, if the extension of a line segment lies completely in the common boundary of two cells. Observe that $H$ is a planar graph without loops. Notice that if the partition algorithm and the protruding algorithm extends the segments of $L$ in the same order, then the same graph $H$ corresponds to the convex partition of $L$ and the unique convex partition of the protrudes $L^\delta$.

We obtain a simple graph $K$ from $H$ by replacing every double edge by a single edge. The minimal degree in $K$ is at least two. More on the structure of graphs $H$ and $K$ can be found in Section 2.4. For a node $v \in V(K)$, we denote by $R_v$ the closure of the cell in the convex partition corresponding to $v$. Similarly, for a set of nodes $D \subset V(K)$, $R_D = \bigcup\{R_v : v \in D\}$.

If the partitioning algorithm extends a segment $\ell = pq$ beyond its endpoint $p$ to a (possibly previously extended) segment $\ell'$ then we call the pair $(p, \ell')$ a *$T$-junction*. If it extends a segment $\ell = pq$ beyond its endpoint $p$ to infinity, then we call $(p, \infty)$ an *$I$-junction*. Every edge of $H$ corresponds to exactly one $T$-junction or $I$-junction.

### 2.2.2 Polygons with holes

We introduce some basic notation for polygons with holes.

**Polygon with $h$, $h \geq 0$, holes.** In this chapter, by polygon we mean a polygon

with holes, a polygon without holes will be called *simple polygon*. A polygon $P$ is a *closed* region in the plane whose boundary is covered by $h+1$ pairwise disjoint simple closed polygonal curves $P_0, P_1, P_2, \ldots, P_h$. The exterior $\mathbb{E}^2 \setminus P$ of $P$ has $h+1$ components, whose boundary is $P_0, P_1, P_2, \ldots, P_h$ respectively.

We may suppose that $P_0$ is the outer polygonal curve, i.e., the boundary of the unique unbounded component of $\mathbb{E}^2 \setminus P$. The *holes* are the simple polygons enclosed by $P_1, P_2, \ldots, P_h$. The line segments of $P_0, P_1, P_2, \ldots$, and $P_h$ are the *sides* of $P$, the segment endpoints form the set of *vertices* of $P$. Each vertex $v$ of $P$ is incident to exactly two sides of $P$, hence the angle $\angle v$ can be defined to be the angular domain formed by the two sides enclosing $P$. We distinguish reflex ($\angle v > \pi$) and convex ($\angle v < \pi$) vertices. We suppose w.l.o.g. that $P$ has no vertex with straight angle ($\angle v = \pi$).

**Convex partition** of $P$. A *convex partition* for a polygon $P$ is a finite set $\mathcal{C}$ of convex open non-overlapping polygonal domains in $P$ such that $\bigcup\{\mathrm{cl}(f) : f \in C\} = \mathrm{cl}(P)$. The elements in $C$ are called the *cells* of the partition.

convex non-overlapping simple polygons in $P$ such that $\bigcup\{\mathrm{cl}(f) : f \in C\} = P$.

**Dual graph** of a convex partition $C$ of $P$. The nodes of the *dual graph* $G$ correspond to the cells of $C$, two nodes are adjacent if and only if the corresponding cells have a common vertex of $P$ on their boundary.

For a node $v \in V(G)$, we denote by $R_v$ the closure of the cell in the convex partition corresponding to $v$. Similarly, for a set of nodes $D \subset V(G)$, $R_D = \bigcup\{R_v : v \in D\}$.

One convex partition of a polygon $P$ is produced by the following *convex partitioning* algorithm.

CONVPART: For every reflex vertex $v$ of $P$, let $\alpha(v)$ be a ray emanating from $v$ which partitions the reflex angle at $v$ into two convex or straight angles, and which does not pass trough any other reflex vertex of $P$. Draw consecutively a directed segment $\beta(v)$ from $v$ along $\alpha(v)$ until it hits a previously drawn segment $\beta(v')$ or the boundary of $P$.

The convex partition obtained in this way is not unique: it depends on the directions of the rays $\alpha(v)$ and on the order in which the segments $\beta(v)$ are drawn. For every reflex vertex $v$, we define recursively directed polygonal curves $\gamma(v)$, which we call *bar*. A bar $\gamma(v)$ starts from $v$ and ends at the boundary of the polygon $P$. Let $v_0 = v$ and $i = 0$. The bar $\gamma(v)$ starts at $v_0$ and it follows $\beta(v_i)$, $i = 0, 1, \ldots$, to its endpoint as long as the this endpoint

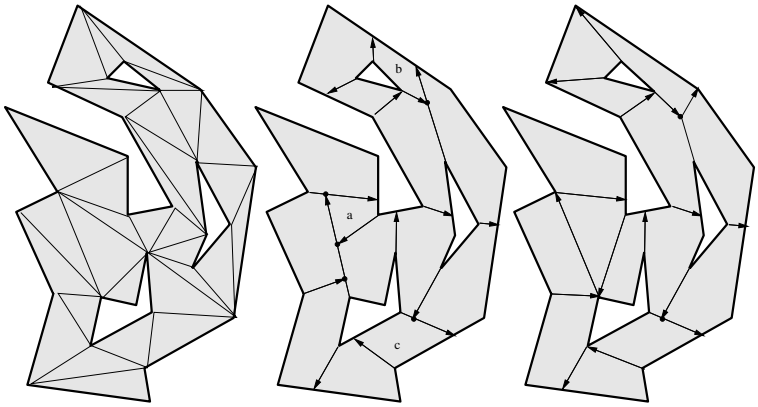


**Figure 2.2:** *A polygon with a triangulation (left), an output of* CONVPART *(middle), and the output of* AUGMENT *in Subsection 2.5.1 (right).*

lies in the relative interior of another segment $\beta(v_{i+1})$, for some $v_{i+1}$.

**Proposition 2.2.** *The convex partition obtained by* CONVPART *has always* $r - h + 1$ *cells, where* $r$ *denotes the number of reflex vertices of* $P$ *and* $h$ *is the number of holes.*

*Proof.* Consider the sets $P(j) := P \setminus \bigcup_{i=1}^{j} \beta(v_i)$ after each step where the algorithm CONVPART draws the segments $\beta(v_1), \beta(v_2), \ldots, \beta(v_r)$ (one for each reflex vertex of $P$). By adding one new segment $\beta(.)$, either $P(j+1)$ has one more connected components than $P(j)$, or the the total number of holes in all components of $P(j+1)$ drops by one. As $P(r)$ is a collection of disjoint convex sets without holes, there must be exactly $h$ steps for which the number of holes drops by one. Therefore $r - h$ steps increase the number of connected components by one each. □

**Proposition 2.3.** *For every cell* $f$ *of the convex partition returned by* CONVPART*, and for every point* $a \in \partial f \setminus \partial P$ *there is a reflex vertex* $v$ *of* $P$ *such that* $A \in \gamma(v)$ *and the portion of* $\gamma(v)$ *between* $v$ *and* $A$ *is on the boundary of* $f$*.*

*Proof.* If $P$ is a simple convex polygon, then there is only one cell. Otherwise on the boundary of every cell, there is part of a segment $\beta(.)$. There is no sequence $\beta(v_1), \beta(v_2), \ldots, \beta(v_t)$ where $\beta(v_i)$ hits $\beta(v_{i+1})$, $i = 1, \ldots, t-1$

and $\beta(v_t)$ hits $\beta(v_1)$, because the segments $\beta(v_i)$ are drawn consecutively. If we follow from $a$ the segments $\beta(v)$ on the boundary $\partial f$ in reverse orientation, the first common point with the boundary of $P$ is a starting point of a segment $\beta(v')$. Hence the reflex vertex $v'$ is on the boundary of $f$ and $a \in \gamma(v')$. □

**Corollary 2.7.** *In the convex partition returned by* CONVPART*, if $P$ has a reflex vertex, then there is at least one reflex vertex of $P$ on the boundary of every cell.*

It is possible, however, that a cell has only one or two vertices of $P$ on its boundary. See cells $a$, $b$, or $c$ in Fig. 2.2 (b).

## 2.3 Illuminating line segments

### 2.3.1 $P_3$-packing of graph $K$

Let $L$ be a set of $n$ disjoint line segments in the plane in general position. Consider a convex partition induced by $L$ and the graphs $H$ and $K$ described above. $n + 1$ light sources placed in the interior of all cells of the convex partitioning solve all four Problems. To beat this trivial upper bound, we need another planar subdivision based on the convex partition. We make the following crucial observation.

**Lemma 2.1.** *The graph $K$ has a $\{P_{\geq 3}\}$-factor if $|V(K)| \geq 3$.*

A $\{P_{\geq 3}\}$-factor is a partition of the graph into paths consisting of at least three nodes. The proof of Lemma 2.1 is postponed to Section 2.4. We apply two immediate corollaries in the next two Sections.

**Corollary 2.8.** *$K$ has a $P_3$-packing which covers at least $3\lceil |V(K)|/5 \rceil$ nodes, if $|V(K)| \geq 3$.*

**Corollary 2.9.** *The graph $K$ has a $\{P_2, P_3\}$-factor if $|V(K)| \geq 1$.*

In the remainder of this subsection, we explain why it is better to illuminate three adjacent cells forming a path $P_3$ than illuminating all convex cells independently. First, we define small regions corresponding to directed edges of $H$ and the notion of *bad* $P_3$. Then we proceed with Lemma 2.2 which states, that two light sources can basically illuminate $R_a$, $R_b$, and $R_c$ whenever $(a, b, c)$ forms a $P_3$ in $K$.

Fix a convex partition of a set $L$ of disjoint line segments. Let $e(a,b)$ be an edge of $H$ between the nodes $a$ and $b$. Assume that $e(a,b)$ corresponds to a $T$-junction $(v_1,\ell)$ and the extension of segment $v_1v_2$ hits the extended segment $\ell$ at point $w$. We associate a small region $r_\varepsilon(a,v_1)$ to the pair of the node $a$ and the edge $e(a,b)$, where $\varepsilon$ is a small positive constant. Let $r_\varepsilon(a,v_1)$ be the intersection of the closed cell $R_a$ and the closed disk centered at $w$ of radius $\varepsilon$.

Let $(a,b,c)$ be a $P_3$ in $K$. Assume that the lines through the line segments $v_1v_2$ and $v_3v_4$ form the common boundary of $R_a, R_b$, and $R_b, R_c$, resp., such that the segment endpoints $v_1$ and $v_3$ correspond to edges between $a,b$ and $b,c$, resp. We say that $(a,b,c)$ is a bad $P_3$ if $(v_3, v_1v_2)$ is a $T$-junction or symmetrically if $(v_2, v_3v_4)$ is a $T$-junction. The region $r_\varepsilon(b,v_2)$ or $r_\varepsilon(b,v_4)$, resp., is called the bad $\varepsilon$-corner of the bad $P_3$.

We define a *line fragment* to be a common portion of the boundary of a cell of the convex partition and a line segment in $L$.

**Lemma 2.2.** *There is a positive constant $\varepsilon_0 = \varepsilon_0(L)$ such that for any $\varepsilon$, $0 < \varepsilon < \varepsilon_0$, and any $P_3$ $(a,b,c)$ of $K$, we can place two light sources in $\operatorname{int}(R_a) \cup \operatorname{int}(R_b) \cup \operatorname{int}(R_c)$ with the following properties.*

- *They illuminate all line fragments on the boundaries of $R_a$, $R_b$, and $R_c$ from at least one side.*
- *If $(a,b,c)$ is not bad, then they illuminate all line fragments on the boundaries of $R_a$, $R_b$, and $R_c$ from the side of $R_a$, $R_b$ and $R_c$ (from two sides on the common boundaries); and they strongly illuminate the interior of $R_a$, $R_b$, and $R_c$.*
- *If $(a,b,c)$ is bad, then they illuminate all line fragments on the boundaries of $R_a, R_b$, and $R_c$ from the side of $R_a$, $R_b$, and $R_c$ (from two sides on the common boundaries), except for the portions of line segments in the bad $\varepsilon$-corner from the side of $R_b$; and they strongly illuminate the interior of $R_a$, $R_b$, and $R_c$ except for the bad $\varepsilon$-corner.*
- *They (strongly) illuminate every $r_{2\varepsilon}(d,v)$ where $v \in V(H) \setminus \{a,b,c\}$ and $v$ corresponds to an edge of $H$ between $d$ and the $P_3$ $(a,b,c)$.*

*Proof.* Denote $v_1v_2$ and $v_3v_4$ the segments between the cells $a,b$ and $b,c$ as above. Let $w_1$ and $w_3$ be the endpoint of the extension of $v_1v_2$ beyond $v_1$ and that of $v_3v_4$ beyond $v_3$. Let $w_1'$ and $w_3'$ be two points on the common boundaries of $R_a, R_b$ and $R_b, R_c$ from distance $\eta > 0$ from points $w_1$ and $w_3$. Here $\eta$ is a fixed constant depending only on the set $L$ such that $w_1'$ and $w_3'$ are disjoint from the segments for all possible $(a,b,c)$ of $K$.

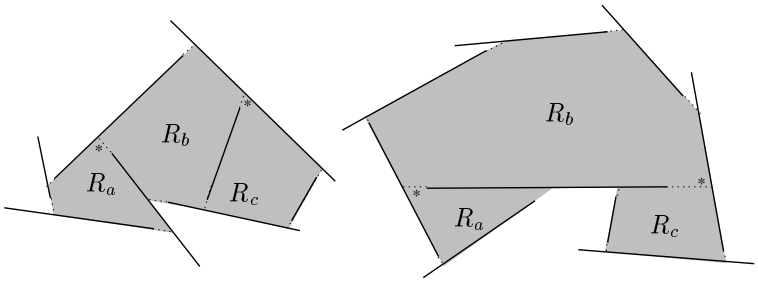


**Figure 2.3:** *Two examples where the two light sources are located in* $R_a \cup R_b \cup R_c$ *if* $(a, b, c)$ *is not a bad* $P_3$.

We place two light sources $s_1$ and $s_3$ into $\mathrm{int}(R_a) \cup \mathrm{int}(R_b) \cup \mathrm{int}(R_c)$ in the $\delta$-neighborhood of $w_1'$ and $w_3'$ but not on the common boundaries of cells, where $\delta > 0$ is a sufficiently small positive constant depending on $L$ and $\varepsilon$. If $v_1v_2 \neq v_3v_4$ then we place the two light sources in the interior of $R_a$ and $R_c$ (left of Fig. 2.3). If $v_1v_2 = v_3v_4$ then we place them on opposite sides of the line through $v_1v_2$, say, into the interior of $R_a$ and $R_b$ (right of Fig. 2.3).

Either light source illuminates the cell in which it is lying and the boundary of this cell. If $v_1v_2 = v_3v_4$, then the light source $s_3$ illuminates cell $R_c$, except for the cone between the rays $\overrightarrow{v_3v_4}$ and $\overrightarrow{s_3v_3}$ and the line fragment of $v_3v_4$ on the boundary of $R_c$. Both are illuminated by $s_1$ is $\delta$ is sufficiently small.

If $v_1v_2 \neq v_3v_4$, then the light source $s_1$ illuminates cell $R_b$; except for the cone between the rays $\overrightarrow{v_1v_2}$ and $\overrightarrow{s_1v_1}$ and the line fragment of $v_1v_2$ on the boundary of $R_b$. The light source $s_3$ illuminates cell $R_b$; except for the cone between the rays $\overrightarrow{v_3v_4}$ and $\overrightarrow{s_3v_3}$, and the line fragment of $v_3v_4$ on the boundary of $R_b$. They together illuminate $R_b$ and its boundary except if $(a, b, c)$ is bad. In that case the intersection of the non-illuminated pieces are all in the bad corner if $\delta$ is sufficiently small w.r.t. $\varepsilon$.

Finally $s_1$ and $s_2$ illuminate the regions $r_{2\varepsilon}(D, V)$ in adjacent cells if $\varepsilon$ is sufficiently small with respect to $\eta$. If $(a, b, c)$ is a bad $P_3$, then the $r_{2\varepsilon}(d, v_3)$ (or $r_{2\varepsilon}(d, v_4)$, resp.) opposite to the bad $\varepsilon$-corner is also illuminated by the light source $s_3$ (resp., $s_1$) (Fig. 2.3). Note that the upper bound on $\varepsilon$ depends only on $\eta$ and not on $\delta$. □

For the sake of completeness, we make the following observation about the bad corners.

**Proposition 2.4.** *There is a positive constant* $\varepsilon_1 = \varepsilon_0(L)$ *such that for any* $\varepsilon$,

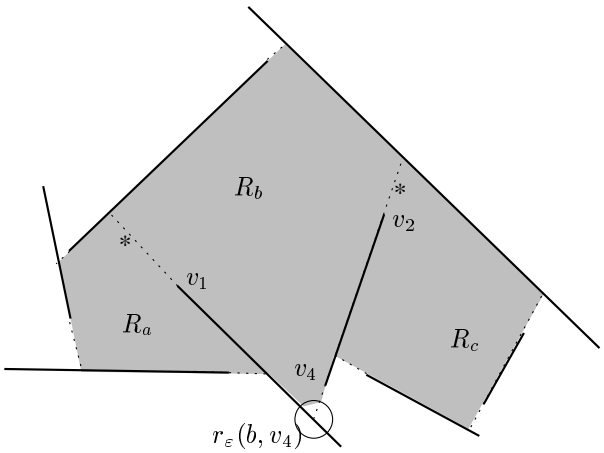


**Figure 2.4:** *Two light source do not illuminate a bad $\varepsilon$ corner of a bad $P_3$.*

$0 < \varepsilon < \varepsilon_1$*, and any* $a \in V(K)$*, we can place a light source in* $\mathrm{int}(R_a)$ *with the following properties.*

- *It illuminates all line fragments on the boundary of $R_a$ from the side of $R_a$; and it strongly illuminates the interior of $R_a$.*
- *It (strongly) illuminate every $r_{2\varepsilon}(d, e(a, d))$ where $v \in V(H) \setminus \{a\}$ and $e(a, d) \in E(H)$.*

### 2.3.2 Weak illumination of the free space

*Proof (of Theorem 2.6). Sufficiency.* According to Corollary 2.9, $K$ has a matching $M$ covering at least $2\lceil (n+1)/3 \rceil$ nodes. Put a light source on the common boundary of two cells corresponding to a pair of nodes in $M$ (but not on a segment of $L$), and put a light source in the interior of each cell not covered by $M$. The resulting set of light sources weakly illuminate $\mathbb{E}^2$. The total number of light sources used is

$$n + 1 - \left\lceil \frac{n+1}{3} \right\rceil = \left\lfloor \frac{2(n+1)}{3} \right\rfloor .$$

*Necessity.* Figure 2.5 depicts a construction of $n = 3k + 1$ line segments requiring $2k + 1 = \lfloor 2(n+1)/3 \rfloor$ light sources. Observe that each shaded region requires one distinct light source. The construction is composed of

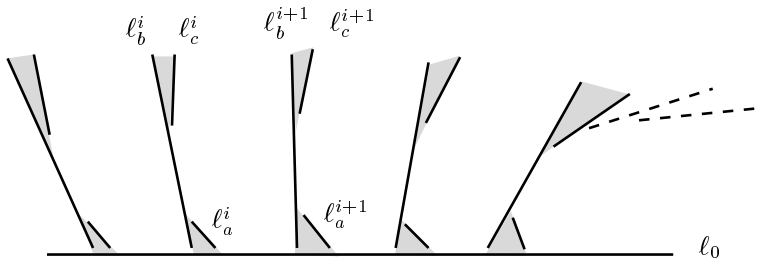


**Figure 2.5:** *A construction of* $n = 16$ *line segments requiring* 11 *light sources.*

a long horizontal line segment $\ell_0$ and $k$ triples $(\ell_a^i, \ell_b^i, \ell_c^i)$ of segments, $i = 1, 2, \ldots, k$. The directions of $\ell_b^1, \ell_c^1, \ell_b^2, \ell_c^2, \ldots, \ell_b^k, \ell_c^k$ follow in this order, and the two extensions of $\ell_a^i$ hit $\ell_0$ and $\ell_b^i$, resp. A construction on $n = 3k + 2$ or $3k + 3$ line segments can be obtained by adding one or two line segments shown as dashed segments in Figure 2.5. □

Our worst case lower bound construction gives also a tight lower bound for the theorem of O'Rourke mentioned in our Remark 2.1. It is, however, considerably simpler than the one presented by O'Rourke [33].

### 2.3.3 Strong illumination of the free space and illuminating both sides of segments

*Proof (of Theorems 2.4 and 2.5). Sufficiency.* According to Corollary 2.8, $K$ contains a $P_3$-packing which covers at least $3\lceil (n+1)/5 \rceil$ nodes. Place two light sources for every $P_3$ as described in the proof of Lemma 2.2 and one light source in the interior every cell not covered by a $P_3$ as described in Proposition 2.4. The two light sources in a $P_3$ (strongly) illuminate the interior and the boundaries of the three cell except for a bad corner in the case of bad corners. The bad corners are (strongly) illuminated from the opposite side: The other side is either another $P_3$ of our $P_3$-packing and the bad corner is illuminated by Lemma 2.2, or it is a single cell in our $P_3$-packing and the bad corner is illuminated by Proposition 2.4. The total number of light sources is

$$2\left\lceil \frac{n+1}{5} \right\rceil + \left( (n+1) - 3\left\lceil \frac{n+1}{5} \right\rceil \right) = \left\lfloor 4 \cdot \frac{n+1}{5} \right\rfloor .$$

*Necessity.* Figure 2.6 depicts a construction of $5k + 1$ line segments requiring $4k + 1 = \lfloor 4(n+1)/5 \rfloor$ light sources. Observe that each line segment in the

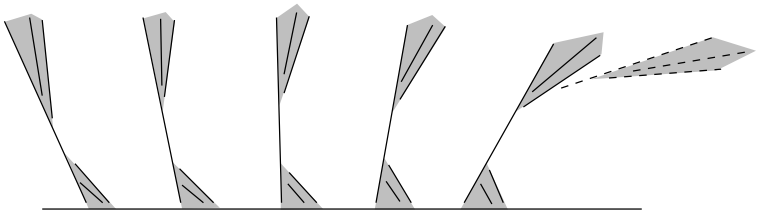

**Figure 2.6:** *A construction of* $n = 26$ *line segments requiring* 21 *light sources*

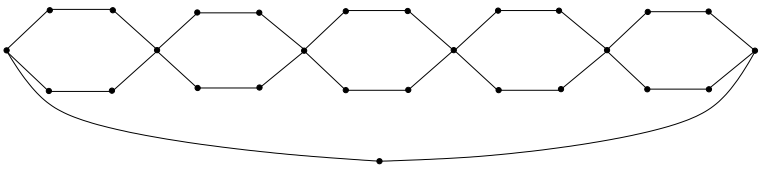

**Figure 2.7:** *The corresponding graph* $K$

interior of a shaded region requires two distinct light sources.

We obtain a construction on $n = 5k + 2, 5k + 3$, or $5k + 4$ line segments by adding one, two, or three line segments shown as dashed segments in Figure 2.6. □

Note that the maximum $P_3$-packing of the corresponding graph $K$ covers only $3k + 3$ nodes of the total $|V(K)| = 5k + 1$.

Theorem 2.4 says that the light source placement based on a maximum $P_3$-packing of $K$ is best possible in the worst case. Note, however, that there is no one-to-one correspondence between maximum $P_3$-packings of $K$ and optimal solutions to Problem 2.2.

### 2.3.4 Illuminating segments

We wish to find a set of light sources in the free space that illuminates every point of the line segments (from one side at least). We show first how an upper bound of $\lfloor 2(n+1)/3 \rfloor$ can be deduced from the fact that a maximum matching of $G$ covers at least $2\lceil (n+1)/3 \rceil$ nodes by Corollary 2.9.

Let $ab$ be a $P_2$ in $K$. We say that $ab$ is a *bad* $P_2$ if two edges of $H$ connects $a$ and $b$ (i.e., $ab$ is a double edge in $H$). Let $v_1v_2$ be the segment whose

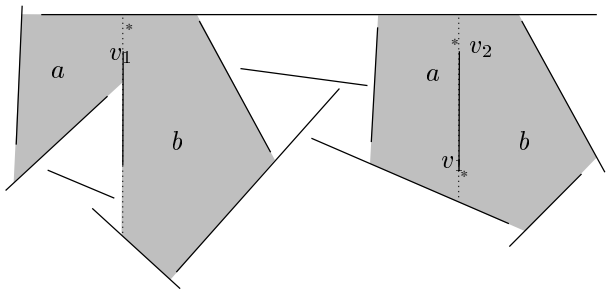


**Figure 2.8:** *A light source next to $v_1$ or $v_2$ cannot always illuminate the entire boundary of $R_{ab}$*

two endpoints correspond to the two edges of $H$ between $a$ and $b$. The two intersection points of the line through $v_1v_2$ and the boundary of $R_{ab}$ are the *bad points* of the bad $P_2$ $ab$.

**Lemma 2.3.** *There is a positive constant $\varepsilon_0 = \varepsilon_0(L)$ such that for any $\varepsilon$, $0 < \varepsilon < \varepsilon_0$, and any $P_2$ of $K$, we can place a light sources in $\operatorname{int}(R_a) \cup \operatorname{int}(R_b)$ with the following properties.*

- *If $ab$ is not bad, then it illuminates all line fragments on the boundaries of $R_a$ and $R_b$ from at least one side.*
- *If $ab$ is bad, then it illuminates all line fragments on the boundaries of $R_a$ and $R_b$, and $R_c$ from at least one side; except for the portion in one of four disjoint intervals on the boundary of $R_{ab}$. The four intervals has length $\varepsilon$ each and they lie in a $\varepsilon$ neighborhoods of the two bad points of $ab$ but do not include the bad points.*

*Proof.* Assume that the lines through the line segments $v_1v_2$ form the common boundary of $R_a$ and $R_b$ such that the segment endpoints $v_1$ correspond to an edge between $a$ and $b$.

If $ab$ is not bad, then we place a light source $s$ into $\operatorname{int}(R_a) \cup \operatorname{int}(R_b)$ in the $\delta$-neighborhood of $v_1$ but not on the common boundary of the cells, where $\delta > 0$ is a sufficiently small positive constant depending on $L$ only. $s$ illuminates the boundary of the cell in which it is lying. It also illuminates the boundary of the other cell, except for the cone between the rays $\overrightarrow{v_1v_2}$ and $\overrightarrow{sv_1}$. Either for $s \in \operatorname{int}(R_a)$ or for $s \in \operatorname{int}(R_b)$, this cone does not intersect any line fraction on the boundary of $R_{ab}$.

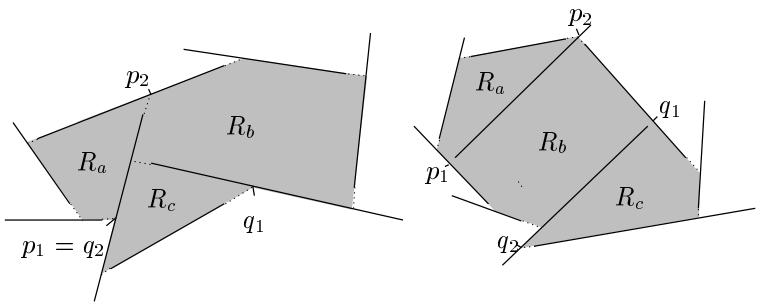


**Figure 2.9:** *Denote the two endpoints of* $\alpha = \partial R_{(a,b,c)} \cap \partial R_a$ *and* $\gamma = \partial R_{(a,b,c)} \cap \partial R_c$ *by* $p_1, p_2$ *and* $q_1, q_2$ *resp.*

If $ab$ is a bad $P_2$, then the segment endpoint $v_2$, too, corresponds to an edge of $H$ between $a$ and $b$. We place a two light source in the interior of $R_a$ or $R_c$, in the $\delta$-neighborhood of $v_1$ or $v_2$. Thus we have four disjoint regions for locating a light source $s$. There is only one portion of the boundaries of the two cells illuminated from neither side by $s$: The projection of $v_1v_2$ from $s$ to the boundary $\partial R_{ab}$. The non-illuminated portion of the boundary can be in four disjoint intervals around the two bad points depending on the four possible location of $s$. This portion can be smaller than $\varepsilon$, if $\delta$ is sufficiently small w.r.t. $\varepsilon$ (see Figure 2.8). □

Let $M$ be a maximum matching of $G$. Place one light source for every pair in the matching and one guard in the deficiency. For every bad pair $ab \in M$, there is a short interval which is not illuminated by the light source assigned to $ab$. All non-illuminated intervals fractions can be illuminated from the other side by choosing $\varepsilon$ sufficiently small in Lemma 2.3.

We can improve upon this bound by using a $\{P_2, P_3\}$-factor $F$ of $K$ which exists by Corollary 2.9. Consider a $P_3$ $(a, b, c)$ of $F$. The common boundaries $\alpha = \partial R_{(a,b,c)} \cap \partial R_a$ and $\gamma = \partial R_{(a,b,c)} \cap \partial R_c$ are connected polygonal curves. Denote the two endpoints of $\alpha$ and $\gamma$ by $p_1, p_2$ and $q_1, q_2$, resp., in clockwise order along $\partial R_{(a,b,c)}$. For every $P_3$ $(a, b, c)$ of $F$, we assign two polygonal paths $\kappa(a)$ and $\kappa(c)$ to the two extremal nodes $a$ and $c$, resp., such that $\kappa(a)$ is the polygonal path along $\partial R_{(a,b,c)}$ from $p_1$ to $q_1$, and $\kappa(c)$ is the polygonal path along $\partial R_{(a,b,c)}$ from $q_1$ to $p_1$ in clockwise order. As a result, $\partial R_{(a.b.c)} = \kappa(a) \cup \kappa(c)$, $\partial R_a \cap \partial R_{(a.b.c)} \subset \kappa(a)$, and $\partial R_c \cap \partial R_{(a.b.c)} \subset \kappa(c)$ (see Figure 2.9).

We define a graph $J_F$. Let the node set $V(J_F)$ be the set of all two extremal

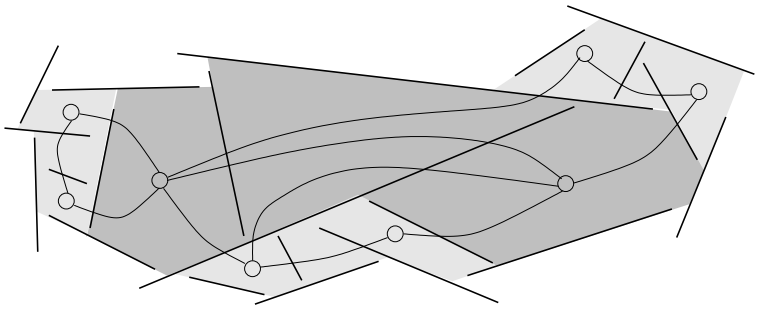

**Figure 2.10:** *Edges of graph $J_F$ defined on the extremal nodes of $P_3$'s of $F$.*

nodes from every $P_3 \in F$. Two nodes $u, v \in V(J_F)$ are connected by an edge if and only if at least one of the following conditions is satisfied (see Figure 2.10):

- $u$ and $v$ are in the same $P_3$ of $F$,
- $\kappa(u) \cap \kappa(v)$ contains a small segment (this includes the case where the cells $R_u$ and $R_v$ have common boundary points),
- $R_u$ and $R_v$ lie at the opposite side of two bad points of a bad $P_2$ in $F$.

Observe that $J_F$ is planar: We can place a point $p(a)$ in $\mathrm{int}(R_a)$ for every $a \in V(J_F)$, and connect the points corresponding to adjacent nodes of $J_F$ by curves without crossings. The two points, $p(a)$ and $p(b)$, for a $P_3$ $(a, b, c)$ of $F$ can be connected by a curve within $R_{(a,b,c)}$. If $\kappa(u_1) \cap \kappa(v_1)$ contains a segment then $p(u_1)$ and $p(v_1)$ can be connected to this segment within $R_{(u_1,u_2,u_3)}$ and $R_{(v_1,v_2,v_3)}$, resp., where $(u_1, u_2, u_3) \in F$ and $(v_1, v_2, v_3) \in F$. Finally, if $R_u$ and $R_v$ lie at bad points of a bad $P_2$ $ab$ of $F$, then $p(u)$ and $p(v)$ can be connected to the respctive bad point within $R_u$ and $R_v$, resp., and we can connect the two bad points by a curve in $R_{ab}$.

According to the four color theorem [3, 4], $J_F$ has an independent set $I_F \subset V(J_F)$ such that $|I_F| \geq \lceil |V(J_F)|/4 \rceil$. Denoting $T_F = V(J_F) \setminus I_F$, $|T_F| \leq \lfloor 3|V(J_F)|/4 \rfloor$.

In every $P_3$ of $F$ there are one or two nodes of $T_F$. Both extremal nodes of a $P_3$ of $F$ cannot be in $I_F$, because they are connected by an edge in $J_F$. Hence, at most half of the $P_3$'s of $F$ contain only one node of $T_F$, and at least half of them contain two.

Now decompose every $P_3$ of $F$ into a $P_2$ and a single node such that elements of $I_F$ are all single nodes. Together with the $P_2$'s of $F$, $V(K)$ is partitioned

into $q$ $P_2$'s and $s$ single nodes ($n + 1 = 2q + s$). Thus we obtain a new partition $F'$ of the graph $K$. Every single node is obtained from a $P_3$ of $F$, hence $s \leq q$, and at most half of the single nodes are in $T_F$.

Place one light source into each $P_2$ of $F$ as described in Lemma 2.3. (The proof of Lemma 2.3 suggests four possiblilities for the location of a light source for a bad pair. We use this freedom in the proof of Lemma 2.4.) If a $P_3$ $(a, b, c)$ has only one extremal node in $T_F$, say $a \in T_F$ and $c \notin T_F$, then place one light source into the $P_2$ $ab$ of $F'$ as described in Lemma 2.3. If a $P_3$ $(a, b, c)$ of $F$ has two extremal nodes in $T_F$, that is $a, c \in T_F$, then place two light sources in $R_{(a,b,c)}$ according to Lemma 2.2. We used one light source for every $P_2$ of $F$, one light source for at least half of the $P_3$'s of $F$, and two light sources for at most half of the $P_3$'s of $F$. All in all, we used at most $\lfloor (n + 1)/2 \rfloor$ light sources. Denote by $S = S(F, T_F)$ the set of these light sources.

**Lemma 2.4.** *The light sources of $S(F, T_F)$ can be arranged so that they collectively illuminate every point of $\bigcup L$.*

*Proof.* We prove that the line fragments on the boundary of the convex cells of the convex partition are all illuminated by some light source of $S(F, T_F)$.

If $S$ has two light sources in $R_{(a,b,c)}$ for a $P_3$ $(a, b, c)$ of $F$, then all line fragments on the boundaries of $R_a$, $R_b$, and $R_c$ are illuminated (from one side) by Lemma 2.2. If a $P_2$ $ab$ of $F'$ then all line on the boundaries of $R_a$ and $R_b$ are illuminated (from one side) by Lemma 2.3; except if $ab$ is a bad pair when one of four arbitrarily short segment in the neighborhood of one of the bad points is not illuminated from the side of $R_{ab}$. The boundary of $R_a$ where $a \in I_F$ is not illuminated from the side of $R_a$.

We distinguish two cases where a line fragment $f_1$ may be not completely illuminated.

(1) A line fragment $f_1$ is not illuminated from either side if $f_1$ is on the common boundary of two cells $R_a$ and $R_b$ belonging to $I_F$. But then $a, b \in V(J_F)$ and $ab \in E(J_F)$, therefore $a$ and $b$ cannot be both in the independent set $I_F$.

(2) Now suppose that $cd$ is a bad $P_2$ and the common boundaries of $R_{cd}$ with adjacent cell containing the two bad points of $cd$ are the line fragments $f_1$ and $f_2$. At least one of $f_1$ and $f_2$ should be illuminated from the other side. Neither $f_1$ nor $f_2$ is illuminated from either side if and only if there are two cells $R_a$ and $R_b$ lie on the other side of $f_1$ and $f_2$, resp., and $a$ and $b$ belong to $I_F$. Both $a$ and $b$ are in $I_F$ if and only if $a, b \in V(J_F)$. If $cd$ is a $P_2$ of $F$, then $ab \in E(J_F)$.

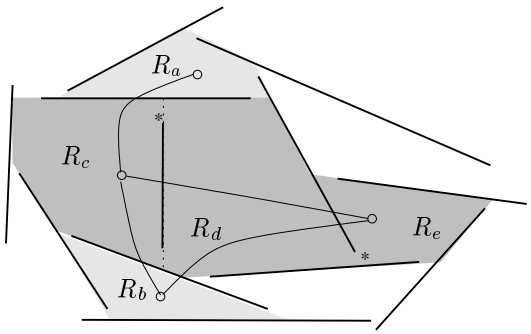


**Figure 2.11:** *A bad $P_2$ is part of a $P_3$ $(c, d, e)$ of $F$.*

Assume that $cd$ is a bad $P_2$ obtained from a $P_3$ $(c, d, e)$ of $F$. Then we have $a, b, c, e \in V(J_F)$. If $a = e$, then $ce, cb, eb \in E(J_F)$, and therefore at least one of $e = a$ and $b$ is not in $I_F$. If $e \neq a$ and $e \neq b$, then either $ac, ae \in E(J_F)$ or $bc, be \in E(J_F)$ (see Figure 2.11). If $a, b \in I_F$ then $c, e \notin I_F$. This implies, in turn, that $S$ has two light sources in $R_{(c,d,e)}$. According to Lemma 2.2, the two light sources in $R_{(c,d,e)}$ illuminate the line fragments $f_1$ and $f_2$. □

## 2.4 The structure of graph $H$ and $K$

In Subsection 2.4.1, we prove that the graph $K$ admits a $\{P_{\geq 3}\}$-factor; the rather technical proof of an important step, our Lemma 2.5, is postponed to Subsection 2.4.2. Relying on a $\{P_{\geq 3}\}$-factor of $H$.

### 2.4.1 The dual graph $K$ has a $\{P_{\geq 3}\}$-factor

A partition of a graph $G$ into paths $\{P_3, P_4, P_5\}$ is a $\{P_{\geq 3}\}$-factor of $G$. We will prove in Lemma 2.1 that the graph $H$ admits a $\{P_{\geq 3}\}$-factor. Here we cite Kaneko's theorem [28] on the characterization of $\{P_{\geq 3}\}$-factors.

**Theorem 2.10.** *[28] A simple graph $G$ has a $\{P_{\geq 3}\}$-factor if and only if*

$$sun(G - S) \leq 2|S|, \text{ for every } S \subseteq V(G). \tag{2.1}$$

□

The term $sun(G - S)$ requires further explanation.

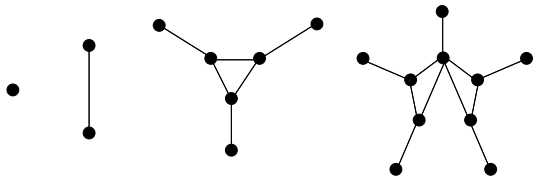

**Figure 2.12:** *Four sun graphs.*

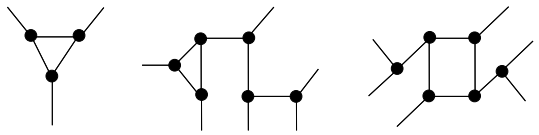

**Figure 2.13:** *Three wheels with all outgoing edges.*

**Definition 2.1.** *A graph $G$ is* factor-critical*, if for every $v \in V(G)$, $G - \{v\}$ contains a perfect matching (1-factor).*

*For a factor critical graph $G$ with $V(G) = \{v_1, v_2, \ldots, v_k\}$, add new vertices $u_1, u_2, \ldots, u_k$ and new edges $v_1u_1, v_2u_2, \ldots, v_ku_k$. The resulting graph is a* sun*. A single node is also considered as a* sun*.*

*For a graph $G$, $sun(G)$ is the number of its* sun components*.*

A factor-critical graph of $k > 1$ nodes has at least $k$ edges. To see this, observe that a factor-critical graph is connected, and it cannot contain leaves. Note that a single node is also factor-critical, the corresponding sun component is a path $P_2$. Consequently, a sun component of $2k$ nodes has at least $2k$ edges for every $k > 1$.

**Definition 2.2.** *For a graph $G$ and $S \subset V(G)$, denote the degree of $S$ in $G$ by $\deg_G(S)$ (i.e. the number of edges connecting nodes of $S$ to nodes of $V(G) \setminus S$).*

*Denote by $E_G(S) = E(G(S))$ the set of edges induced by $S$ in $G$. Let $G(S)$ be the subgraph of $G$ induced by $S$, and $G - S$ be the subgraph induced by $V(G) \setminus S$.*

**Definition 2.3.** *For a graph $G$, we call $S \subset V(G)$ a* wheel*, if $S$ induces a connected graph, $|E_G(S)| = |S|$, $\deg_G(S) = |S|$, and $\deg_G(v) > 2$ for every $v \in S$.*

The key to apply Kaneko's Theorem to $K$ is the following lemma.

**Lemma 2.5.** *Let $S \subset V(K)$. Then*

$$|E_H(S)| \quad \leq \quad 2|S| - p' - r', \tag{2.2}$$

*where $p'$ is the number of nodes in $S$ whose degree is two in both $K$ and $K(S)$, and $r'$ is the maximum number of disjoint subsets of $S$ which are wheels in both $K$ and $K(S)$.*

The rather technical proof of this Lemma is postponed to Section 2.4.2. Here we provide the proof of Lemma 2.1.

*Proof (of Lemma 2.1).* Suppose that (2.1) does not hold for the graph $K$ of the convex partitioning. Choose a subset of nodes $S \subset V(K)$, where $sun(K - S) - 2|S|$ is maximal.

If $S = \emptyset$ then, since $K$ is connected, $\mathrm{sun}(K) > 0$ implies that $|V(K)| \leq 2$.

Delete all components of $K - S$ which are not sun and denote the resulting graph by $K'$. We wish to apply Lemma 2.5, in particular Inequality (2.2), to $V(K')$. For this reason, we introduce parameters for the number of certain types of sun components of $K - S$. The $P_1$'s, $P_2$'s, and all other suns are considered separately.

Let $c_2$ and $c_3$ be the number of single node components $v \in V(K) \setminus S$ with $\deg_K(v) = 2$ and $\deg_K(v) \geq 3$ resp. (Note that here $\deg_{K'}(v) = \deg_K(v)$.)

Let $d_2$, $d_3$, and $d_4$ be the number of $P_2$ components $P \subset V(K) \setminus S$ such that $P \subset K'$ with $\deg_K(P) = 2$, $\deg_K(P) = 3$, and $\deg_K(P) \geq 4$ resp. Observe that $\deg_K(P) = 2$ (or $\deg_K(P) = 3$) means that $P$ contains two (resp., one) node(s) with $\deg_K(v) = 2$.

Consider the sun components of $K - S$ with more than 2 nodes. Let the total number of their nodes be $2k$. They have $k$ leaves. Let $f_2$ and $f_3$ be the number of leaves $v$ with $\deg_K(v) = 2$ and $\deg_K(v) > 2$ resp. Let $r_1$ and $r_2$ be the number sun components of $K - S$ with more than 2 nodes, whose factor critical subgraph is wheel and non-wheel resp. These components have at least $2k + r_2$ interior edges and at least $2k - f_2$ edges connect them to $S$.

Using these parameters, we can give a lower bound to the number of edges in $K'$ by $|E(K')| \geq \deg_{K'}(S) + (d_2 + d_3 + d_4) + (2k + r_2)$ where $\deg_{K'}(S) \geq 2c_2 + 3c_3 + 2d_2 + 3d_3 + 4d_4 + (2k - f_2)$. The number of nodes having degree 2 in both $K$ and $K'$ is at least $c_2 + 2d_2 + d_3 + f_2$.

Inequality (2.2) is written as

$$2c_2 + 3c_3 + 3d_2 + 4d_3 + 5d_4 + 4k + r_2 - f_2 \leq$$
$$\leq 2(|S| + c_2 + c_3 + 2d_2 + 2d_3 + 2d_4 + 2k) - (c_2 + 2d_2 + d_3 + f_2) - r_1,$$
$$c_2 + c_3 + d_2 + d_3 + d_4 + r_1 + r_2 \leq 2|S|.$$

The number of sun components is at most $2|S|$. According to Kaneko's theorem, $K$ has a $\{P_{\geq 3}\}$-factor. □

### 2.4.2 Basic properties of $H$ and $K$

Recall that $K$ is the graph received from $H$ by replacing double edges by simple edges. Notice that $V(K) = V(H)$ with $|V(K)| = n + 1$, $|E(K)| \leq |E(H)| = 2n$.

If $H$ has $p$ double edges, then $|E(K)| \leq 2n - p = 2|V(K)| - 2 - p$.

Denote by $q$ the number of line segments which are extended to a complete line by the convex partitioning algorithm. $q \in \{0, 1\}$ since the line segments are pairwise non-parallel. Denote this line by $\ell_0 \in L$ if exists. Observe that there is no cutting edge in $H$ nor in $K$.

**Lemma 2.6.** *The number of 2-edge cuts in $H$ is $q$.*

*Proof.* A proper subset $S \subset V(H)$ corresponds to a region $R_S$ in the plane, that is the union of the closed cells corresponding to the nodes in $S$. $R_S$ is connected if $S$ induces a connected subgraph. The boundary of $R_S$ is composed of pieces of the extended line segments of $L$. We distinguish two cases:

CASE 1: *$R_S$ or $\mathbb{E}^2 \setminus R_S$ is a bounded region.* Any bounded polygonal domain has at least three vertices. Every vertex corresponds to a $T$-junction. Every $T$-junction on the boundary of $R_S$ corresponds in turn to an edge of $H$ in $\deg_H(S)$, hence $\deg_H(S) \geq 3$.

CASE 2: *Both $R_S$ and $\mathbb{E}^2 \setminus R_S$ are unbounded regions.* There are at least two $I$-junctions on the boundary of $R_S$. If there are at least two line segments on the boundary of $R_S$, then there is at least one $T$-junction on the boundary, since the line segments are non-parallel, and therefore, $\deg_H(S) \geq 3$.

A set $S \subset V(H)$ for which both $R_S$ and $\mathbb{E}^2 \setminus R_S$ are unbounded can have $\deg_H(S) = 2$, if and only if there is exactly one line segment on the boundary of $R_S$. In this case, the unique line segment on the boundary of $R_S$ is extended to the infinity in both directions. □

**Corollary 2.11.** *For every node $v$ with $\deg_H(v) = 2$, $R_v$ is a half-plane of $\ell_0$.*

**Corollary 2.12.** *If $P \subset V(G)$ is a maximal path such that for every $v \in P$, $\deg_K(v) = 2$, $\deg_H(v) > 2$, and $\deg_H(P) = 2$, then $R_P$ is a half-plane of $\ell_0$.*

*Proof.* The two extremal edges of the path $P$ form a 2-edge cut. This is a 2-edge cut of both $H$ and $K$. □

**Lemma 2.7.** *Let $S \subset V(H)$, $S \neq V(H)$, such that $S$ induces a connected subgraph. Then*

$$2|S| \leq |E_H(S)| + \deg_H(S). \tag{2.3}$$

*Proof.* Consider a protruded set $L^\delta$ of segments obtained from $L$. The cells and the graphs $H$ and $K$ are the same for $L$ and for $L^\delta$. Let $a$ be the number of the line segments of $L^\delta$ in $\mathrm{int}(R_S)$. Note that both endpoints of a segment $v_1v_2 \subset \mathrm{int}(R_S)$ correspond to edges in $E_H(S)$. Let $b$ be the number of segment endpoints which corresponds to an edge in $E_H(S)$ but whose segments is not fully in $\mathrm{int}(R_S)$. Finally, let $c$ be the number of segments on the boundary of $R_S$ completely disjoint from $\mathrm{int}(R_S)$. Using these parameters, we estimate $|S|$, $|E_H(S)|$, and $\deg_H(S)$.

For estimating $|S|$, we use the argument of Proposition 2.2. We observe how $R_S$ is partitioned by our convex partitioning algorithm when the segment extensions $\beta(v_1), \beta(v_2), \ldots, \beta(v_{2n})$ are added one-by-one. Let $R(0) = R_S \setminus \bigcup L^\delta$. Denote by $h$ the number of holes of $R(0)$. In particular, every segment of $L^\delta$ in $\mathrm{int}(R_S)$ forms one connected component, hence $h \geq a$. Moreover if $\mathbb{E}^2 \setminus R_S$ is bounded then $h \geq a+1$. Put $R(j) := R_S \setminus \bigcup_{i=1}^{j} \beta(v_i)$, for $j = 1, 2, \ldots 2n$. By adding one new segment $\beta(.)$ intersecting $\mathrm{int}(R_S)$, either $R(j+1)$ has one more connected components than $R(j)$, or the the total number of holes in all components of $R(j+1)$ drops by one. As $R(2n)$ has no holes and at least $2a + b$ extensions intersect the interior of $R_S$, there must be $2a + b - h$ steps increasing the number of connected components by one each.

$|E_H(S)| = 2a + b$ follows from the definition of parameters $a$ and $b$. For $\deg_H(S)$, consider the consecutive line segments along every connected component of the boundary of $R_S$. Every pair of consecutive line segments on the boundary forms a $T$-junction and corresponds to an edge in $deg_H(S)$. The two endings of the last segments on an unbounded component of the boundary correspond to two $I$-junctions. This implies that $\deg_H(S) \geq c + b$ if $R_S$ or $\mathbb{E}^2 \setminus R_S$ is bounded, and $\deg_H(S) \geq c + b + 1$ if both are unbounded.

If the complement $\mathbb{E}^2 \setminus R_S$ is bounded, then our proof is complete: $|S| \leq a+b$ and $|E_H(S)|+\deg_H(S) \geq 2a+2b+c$. Otherwise we have $|S| \leq a+b+1$. If $R_S$ is bounded, then the convex hull $\mathrm{conv}(R_S)$ has at least three sides. On any curve on the boundary of $R_S$ between two endpoints of a side of $\mathrm{conv}(R_S)$, there is a segment of on the boundary of $R_S$ disjoint from $\mathrm{int}(R_S)$, hence $c \geq 3$. If both $R_S$ and $\mathbb{E}^2 \setminus R_S$ are unbounded, then $|E_H(S)| + \deg_H(S) \geq 2a + 2b + c + 1$ and $c \geq 1$ since $\mathrm{conv}(R_S)$ has at least one side. □

**Corollary 2.13.** *If $S \subset V(K)$ is a wheel and $|E_H(S)| = |S|$, then either $\mathbb{E}^2 \setminus R_S$ is bounded or $R_S$ is a half-plane of $\ell_0$.*

*Proof.* Suppose that $|S| = k$ and $S$ induces at most $k$ edges in $H$. Clearly, $S \neq V(H)$.

$$2k = |E_H(S)| + \deg_K(S) \leq |E_H(S)| + \deg_H(S) \leq |S| + k = 2k.$$

This means that (2.3) holds with equality. The argument of the previous proof implies that either $\mathbb{E}^2 \setminus R_S$ is bounded or both $R_S$ and $\mathbb{E}^2 \setminus R_S$ are unbounded and $c = 1$. □

*Proof (of Lemma 2.5).* We know that $|E(H)| = 2|V(H)| - 2$. Applying Lemma 2.7 to every component of $V(K) \setminus S$, we have $|E_H(S)| \leq 2|S| - 2$.

Let $S_1$ be the set of nodes $v \in S$ with $\deg_K(v) = \deg_{K(S)}(v) = 2$. Let $S_2$ be the union of maximal paths $P \subset K(S)$ such that $\deg_K(P) = \deg_{K(S)}(P) = 2$ and $\deg_H(v) > 2$ for every $v \in P$. Let $S_3$ be the union of disjoint vertex sets in $S$ which are wheels in both $K$ and $K(S)$. Observe that $S_1$, $S_2$, and $S_3$ are pairwise disjoint.

According to Corollaries 2.11, 2.12, and 2.13, there are at least $p' + r' - 2$ double edges in $H'$. Either $q = 1$ and the two halfplanes of $\ell_0$ do not give raise to a double edge in $H$, or $q = 0$ and there is a wheel component which gives raise to no double edge in $H$. □

## 2.5 Illuminating triangles

The following Proposition justifies the use of click covers of the dual graph.

**Proposition 2.5.** *Assume that the vertices of $P$ are in general position. For any collection of cells that share a common vertex of $P$ on their boundaries, the closures of the cells can be illuminated by one light source in* $\mathrm{int}(P)$.

*Proof.* Let $a_i, a_2, \ldots, a_k$ denote the cells sharing a common point of $W$ on their boundaries. Since the triangles are in general position, there are overlapping convex sets $A_1, A_2, \ldots, A_k$ in $P$ such that $\mathrm{cl}(a_i) \subset A_i$ for every $i = 1, 2, \ldots, k$. A light source in the interior of the intersection of these convex sets, $\mathrm{int}(\bigcap_{i=1}^{k} A_i)$ illuminates all $\mathrm{cl}(a_i)$, $i = 1, 2, \ldots, k$. □

Instead of click covers of the dual graph $G$, we consider matchings only. Berge's formula [5, 42] states that the *deficiency* $d(G)$ of $G$, the number of nodes not covered by a maximum matching, equals $\max_{S \subset V(G)} \mathrm{odd}(G-S) - |S|$ where $\mathrm{odd}(G-S)$ denotes the number of odd components of the induced subgraph $G - S$. Moreover, for any $S \subset V(G)$ for which $\mathrm{odd}(G-S) - |S|$ is maximal, there is a matching where every component of $G - S$ with $2k+1$ nodes contains $k$ matched pairs. We will consider a maximum matching of $G$, and count the number of vertices of $P$ on the boundaries of cells corresponding to connected components of $G-S$. There is one exceptional configuration we must consider.

**Definition 2.4.** *A 5-group in a convex partition is a set $Q$ of five cells which jointly have a set $U$ of six vertices of $P$ on their boundaries with the property that for any $a \in U$ there is a $b \in U$ such that the boundary of any cell of $Q$ contains $a$ or $b$.*

**Lemma 2.8.** *For a triangular polygon $P$ with $k$ triangular holes, there is a convex partition $C$ with $2k+1$ cells such that for any subset $D \subset V(G)$,*

- *$|W \cap R_D| \geq |D| + 2$, if $|D|$ is odd and $D$ does not form a 5-group,*
- *$|W \cap R_D| \geq |D| + 1$, if $|D|$ is even or $D$ forms a 5-group.*

The proof of Lemma 2.8 is postponed to Subsections 2.5.1 and 2.5.2.

*Proof (of Theorems 2.1 and 2.2).* Consider the convex partition $C$ assured by Lemma 2.8. Choose a set of nodes $S \subset V(G)$ for which $\mathrm{odd}(G-S) - |S|$ is maximal. Denote by $\mathrm{even}(G-S)$ and $\mathrm{odd}(G-S)$ the number of connected components of $G - S$ with even and odd number of nodes, respectively. Let $q$ denote the number of odd components which form 5-groups. If there is no edge between $D \subset V(G-S)$ and $E \subset V(G-S)$ then $W \cap R_D$ and $W \cap R_E$ are disjoint sets. Summing up the number of elements of $W$ in $R_D$ for every connected component $D$ of $G - S$, we have by Lemma 2.8

$$\mathrm{even}(G-S) + 2\,\mathrm{odd}(G-S) - q + (2k+1-|S|) \leq |W| = 3k+3, \quad (2.4)$$

$$2\,\mathrm{odd}(G-S) - |S| \leq k + q + 2 - \mathrm{even}(G-S). \quad (2.5)$$

If $S = \emptyset$, then $G$ is factor critical, and $k + 1$ light sources are sufficient. If $S$ is non-empty, then we have

$$-|S| \leq -1. \tag{2.6}$$

Combining Inequalities (2.5) and (2.6), we obtain $d(G) = \mathrm{odd}(G-S)-|S| \leq (k + q + 1 - \mathrm{even}(G - S))/2 \leq (k + q + 1)/2$. That is, the deficiency is at most $(k + q + 1)/2$. Moreover, the deficiency is always odd, because the total number of nodes is odd.

We illuminate all cells by placing one light source for every pair of cells in a maximum matching by Proposition 2.5, and one light source at a point of $W$ on the boundary of every cell in the deficiency. This set of light sources can be rearranged for every 5-group $Q$ containing two matched pair of cells: After a light source is placed on the boundary of the fifth cell (either it is in the deficiency or in a matching), the remaining four cells can be covered by one light source instead of two. The number of necessary light sources is therefore at most

$$\frac{2k + 1 - d(G)}{2} + d(G) - q = k - q + \frac{d(G) + 1}{2} \leq \left\lfloor \frac{5k + 3}{4} \right\rfloor.$$

This completes the proof of Theorem 2.1. For the proof of Theorem 2.2, we still need to take care about the boundaries of the open cells. Observe that the common boundary of two cells $R_a$ and $R_b$ is always illuminated by a light source on the boundary of one of the cells. It might happen that a segment $vz$ on the common boundary is collinear with a side $vw$ of a triangle on the boundary of $R_a$ and the light source assigned to $R_a$ is located at $w$. In this case, the light source at $w$ does not illuminate $vz$, but the light source assigned to $R_b$ does illuminate it, since $vz$ cannot be collinear with two sides of triangles. □

### 2.5.1 Three vertices on every cell

The main tool is a new convex decomposition of polygons with holes. The following theorem captures a crucial property of triangulations using a considerably smaller number of convex cells.

**Theorem 2.14.** *Any polygon $P$ with $h$ holes and with $r$ reflex vertices can be partitioned into $r-h+1$ convex cells such that there are at least three vertices on the boundary of every cell.*

*Proof.* Consider a convex partition $C$ returned by CONVPART. We make continuous local modifications on the partition until every cell has at least three vertices of $W$ on their boundary. This can been done independently one-by-one for each cell.

Let $f$ be a cell of $C$. By Proposition 2.3, the boundary of $f$ is composed of initial portions of bars $\gamma(v)$, and portions of the boundary of $P$. Every vertex $a$ of $f$ is either a vertex of $P$, or belongs to a bar $\gamma(v)$ such that the portion of $\gamma(v)$ between $v$ and $a$ is on the boundary of $f$. Our goal is to modify $f$ until

($*$) *every vertex $a$ of $f$ is either*

- *in $W$, or*
- *the endpoint of a side of $f$ containing a vertex $v$ of $P$ in its relative interior such that the directed segment $va$ is part of $\gamma(v)$.*

This implies that there are at least three points of $W$ on the boundary of $f$ because every convex cell has at least three vertices.

Let $v$ be a reflex vertex of $P$ on the boundary of $f$. It is possible that $v$ is in the relative interior of a side of $f$. Let $a_0, a_1, a_2, \ldots a_k$ be consecutive vertices of $f$ such that $a_0a_1 \not\subset \gamma(v)$, $a_1 \in \gamma(v)$, and $v \in \operatorname{relint}(a_0a_1) \cup \{a_1\}$. (We do not consider bars which are completely in the relative interior of a side of $f$.)

Observe that ($*$) holds for $a_1$. Suppose that ($*$) does not hold for $a_2$. We apply the following step on the boundary of $f$.

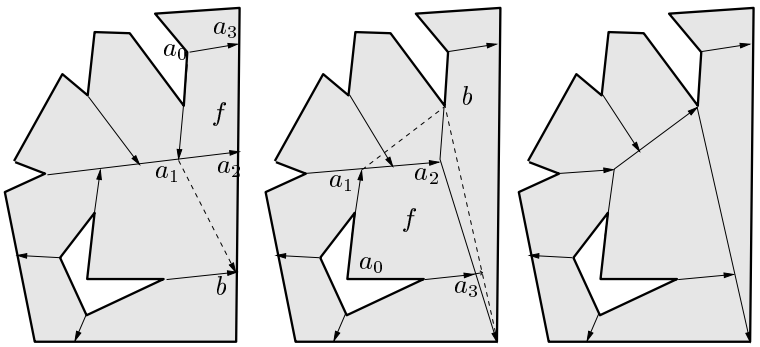


**Figure 2.14:** *Rotation in the two cases.*

AUGMENT($a_2$): Denote by $a_2a_3$ and $a_2b$, resp., the two sides incident to $a_2$ in the two cells of $C$ along side $a_1a_2$. We distinguish two cases.

A, $\angle a_1a_2a_3 + \angle a_1a_2b \leq \pi$: Rotate segment $a_1a_2$ to $a_1a'$ such that $a'$ moves along $a_2b$ until $a' = b$ or $\angle a_0a_1a' = \pi$. (See the left of Fig. 2.14.)

B, $\angle a_1a_2a_3 + \angle a_1a_2b > \pi$: Rotate segments $a_1a_2$ to $a_1a'$ and $a_2a_3$ to $a'a_3$ such that $a'$ moves along $a_2b$ until either $a' = b$, or $\angle a_0a_1a' = \pi$, or $\angle a'a_3a_4 = \pi$. (See the middle of Fig. 2.14.)

In the neighboring cells, shorten the sides whose endpoints were in the relative interior of the sides $a_1a_2$, and $a_1a_2$ or $a_3a_2$.

After a step AUGMENT, it can happen that $a_1$ or $a_3$ are not vertices of $f$ anymore, if $\angle a_0a_1a' = \pi$ or $\angle a'a_3a_4 = \pi$, respectively. Iterating step AUGMENT on the curve $\gamma(v)$, eventually $(*)$ holds for all vertices of $\gamma(v)$ on the boundary of $f$. Observe that the continuous modifications in process AUGMENT preserve the orientation of every portion of the bars. We can define the bar $\gamma(v)$ for every reflex vertex of $P$ at all times. If a new reflex vertex $w$ of $P$ appears on a bar $\gamma(v)$ after a rotation step, then necessarily $w$ will be the endpoint of $\gamma(v)$ and the remaining of the previous $\gamma(v)$ will belong to $\gamma(w)$.

We need to worry that a cell adjacent to $f$ is contracted to a point while we augment the area of $f$. This can only happen if $a'$ reaches $b$ and $\gamma(b)$ hits the side $va_1$. A cell might be contracted to a point only if it has only one vertex of $P$ on its boundary. We can avoid this if we first modify the cells until every cell has at least two vertices of $P$ on their boundary and then assure a third vertex for every cell.

We have a convex partition of $F$ at all times during the process. No vertex of $W$ is detached from $f$, nor from any other cell. We can assure, independently for every cell $f$, that there are at least three vertices of $P$ on the boundary of $f$. □

Theorem 2.14 can be generalized to the following form:

**Lemma 2.9.** *Any polygon $P$ with $h$ holes and with $r$ reflex vertices can be partitioned into $r - h + 1$ convex cells such that every closed Jordan curve along the boundaries of cells contains at least three vertices of $P$.*

*Proof.* Consider a convex partition returned by CONVPART. Let $\kappa$ be a closed Jordan curve along boundaries of cells. Curve $\kappa$ circumscribes a polygonal region $K$. There are at least three convex vertices on the boundary of $K$. We can modify the convex cells containing the convex vertices of $K$ until

$(**)$ *every convex vertex $a$ of $K$ is either*

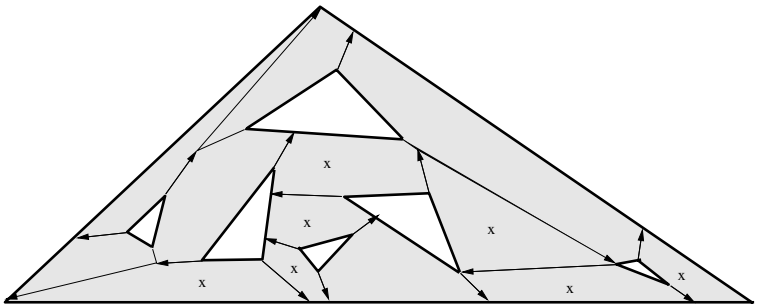

**Figure 2.15:** *A possible result of* CONVPART *and* AUGMENT *on an example polygon.*

- *in $W$, or*
- *the endpoint of a reflex curve along the boundary of $K$ containing a reflex vertex $v$ of $P$ in its relative interior such that the directed curve $va$ is part of $\gamma(v)$.*

Let $v$ be a reflex vertex of $P$ on the boundary of $f$. Let $a_0, a_1, a_2, \ldots a_k$ be consecutive vertices of $f$ such that $a_0a_1 \not\subset \gamma(v)$, $a_1 \in \gamma(v)$, and $v \in \text{relint}(a_0a_1) \cup \{a_1\}$. Suppose that $a_i$ and $a_j$, $0 < i < j$, are the first two convex vertices in this sequence. Note that $(**)$ holds for $a_i$. If $(**)$ does not hold for $a_j$, we can apply step AUGMENT($a_{i+1}$) on the boundary of $f$ until a new vertex of $F$ appears on the reflex chain $a_i, a_{i+1}, \ldots, a_j$ or until the angle at $a_i$ becomes a straight angle and $a_j$ becomes the first convex vertex on the chain between $a_0$ and $a_j$

It might be that as a result of an AUGMENT step, a curve $\kappa$ is not simple anymore. In this case we can remove this curve for forther consideration. No step AUGMENT creates new simple Jordan curves, nor detaches any point of $W$ from closed Jordan curves. Therefore, the statement of our lemma can be assured for every closed Jordan curve one by one. □

### 2.5.2 Proof of Lemma 2.8

It is enough to consider $|W \cap R_D|$ for the case where $R_D$ is a connected region. If $R_D$ is not connected then Lemma 2.8 follows by adding up the number of vertices of $P$ in all connected components of $R$.

For every $D \subseteq V(G)$, we construct a new partition $\Phi_D$ of $P$ into at most

$2k+1$, not necessarily convex, cells. The cells of the new partition correspond to the cells of $C$: For a node $a \in V(G)$, the closed region $R_a$ corresponds to the closed region $\varphi(R_a) \in \Phi_D$. Using the notation $\varphi(R_D) = \bigcup \varphi(R_a) : a \in D\}$, we construct $\Phi_D$ such that $|W \cap \varphi(R_D)| = |W \cap R_D|$ and for all other $D' \subseteq V(G)$, $|W \cap \varphi(R_{D'})| \geq |W \cap R_{D'}|$.

For every vertex $v$ of $P$, we draw a directed polygonal curve $\lambda(v)$ from $v$ to another vertex of $P$. Every curve $\lambda(v)$ will be constructed along the bars $\gamma(.)$ and the boundary of $P$. Unlike the curves $\gamma(.)$, the relative interiors of the curves $\lambda(.)$ will be disjoint.

Let $o_1o_2o_3$ be the outer triangle and let $\lambda(o_1)$, $\lambda(o_2)$, and $\lambda(o_3)$ be the sides of the outer triangle. We draw the other curves $\lambda(v)$ consecutively. $\lambda(v)$ follows $\gamma(v)$ until it reaches either a previously drawn $\lambda(v')$ at point $q$, or until the endpoint $w(v)$ of $\gamma(v)$. If $\lambda(v)$ reaches $w(v)$ at a vertex of $P$, then $\lambda(v)$ ends there. If $w(v)$ is at the relative interior of a side $ab$ of a triangle $abc \in T$ then $\lambda(v)$ follows $ab$ to either $a$ or to $b$. If $\lambda(v)$ reaches a previously drawn $\lambda(v')$ at point $q$ then it follows $\lambda(v')$ parallelly to its starting point or to its endpoint. In the last two cases, there are two possible directions to follow $ab$ or $\lambda(v')$ respectively. We choose the direction of $\lambda(v)$ as follows: If region $R_D$ is on one side of $w(v)$ (resp., $q$) and $F \setminus R_D$ is on the other side then $\lambda(v)$ follows the direction of $R_D$; otherwise $\lambda(v)$ follows either direction (see Fig. 2.16).

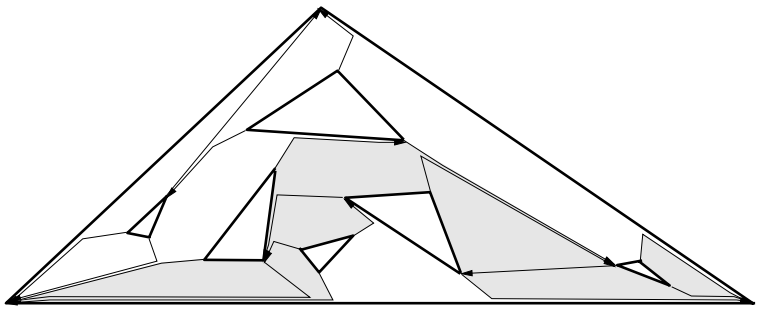

**Figure 2.16:** *The new partition where the shaded region is* $\varphi(R_D)$.

We have now a planar diagram where the nodes are the points of $W$, the edges are the curves $\lambda(.)$ and the sides of triangles, and the faces are cells of $\varphi(R_D)$ and triangles. Note that for every closed curve Lemma 2.9 holds.

It rests to show that $|W \cap \varphi(R_D)| \geq |D|+2$ if $|D|$ is odd, and $|W \cap \varphi(R_D)| \geq |D| + 1$ if $|D|$ is even. It clearly holds if $R_D = \mathrm{cl}(F)$ and $|D| = 2n +$

1. Otherwise consider the boundary of $\varphi(R_D)$. A *circumscribing curve* of $\varphi(R_D)$ is a weakly simple closed curve $\kappa$ in the boundary of $\varphi(R_D)$ such that $\varphi(R_D)$ is in the closed region circumscribed by $\kappa$ and the exterior of $\kappa$ is connected. (A closed curve is *weakly simple*, if it has no self-crossing but it is not necessarily simple.) Let $\kappa$ be a circumscribing curve of $\varphi(R_D)$ such that it contains the minimum number of faces. Notice that the boundary of $\varphi(F) \setminus \varphi(R_D)$ is not necessarily connected, and $\kappa$ forms the boundary of one component. We fill up these holes by regarding all components of $\varphi(F) \setminus \varphi(R_D)$ in the interior of $\kappa$ as new faces, and let $D'$ denote the union of $D$ and the new faces. $|D| \leq |D'|$ and we call $\varphi(R_{D'})$ the union of the closures of all faces in $D'$. If we consider $\varphi(R_{D'})$ and delete one-by-one the points of $W$ and the curves $\lambda(.)$ which are not contained in the curve $\kappa$, then we obtain the a partition into $|D'|$ faces.

We estimate $|D'|$: Denote by $b$ the number of curves $\lambda(.)$ along $\kappa$. Let $t_2$ and $t_3$ denote the number of triangles $\tau$ such that $\varphi(R_{D'}) \setminus \tau$ has 2 and 3 connected components, respectively. Let $T_2$ and $T_3$ be the set of these triangles. For every vertex $v \in \varphi(R_D) \cap W$, let $1 + h_v$ be the number of components of $\varphi(R_D) \setminus v$, and let $h = \sum_{v \in \varphi(R_D) \cap W} h_v$. We have

$$|D'| \leq 1 + (|W \cap \varphi(R_D)| - b) + (h + t_2 + 2t_3). \tag{2.7}$$

A triangle $\tau$, for which $\varphi(R_{D'}) \setminus \tau$ splits into two (resp., three) components, has two (resp., three) vertices on $\kappa$. From Lemma 2.9, we deduce that the boundary of every component of $\varphi(R_{D'}) \setminus (W \cup T_2 \cup T_3)$ contains at least two curves $\lambda(.)$. This implies

$$b \geq 2(h + 1) + 2t_2 + 4t_3. \tag{2.8}$$

Combining Inequalities (2.7) and (2.8), we obtain

$$|D'| + 1 + t_2 + 2t_3 + h \leq |W \cap \varphi(R_D)|.$$

This proves Lemma 2.8 if $t_2 + 2t_3 + h \geq 1$, $|D| < |D'|$, or at least one of the Inequalities (2.7) and (2.8) is not tight.

Suppose that $t_2 = t_3 = h = 0$, $|D| = |D'|$, $b = 2$, and Inequality (2.7) is tight. This latter assumption implies that there are exactly $|W \cap \varphi(R_D)| - 2$ curves $\lambda(.)$ which splits the interior of $\varphi(R_{D'})$ into two parts. Consequently, every triangle circumscribed by the curve $\kappa$ has at least one common vertex with $\kappa$. There are at most two such triangles, since $b = 2$. Denote them by $\tau_1$ and $\tau_2$.

If all elements of $W \cap \varphi(R_D)$ are on $\kappa$, then $W \cap \varphi(R_D)$ and the curves $\lambda(.)$ form a planar drawing of an outer-planar graph with $|W \cap \varphi(R_D)|$ nodes

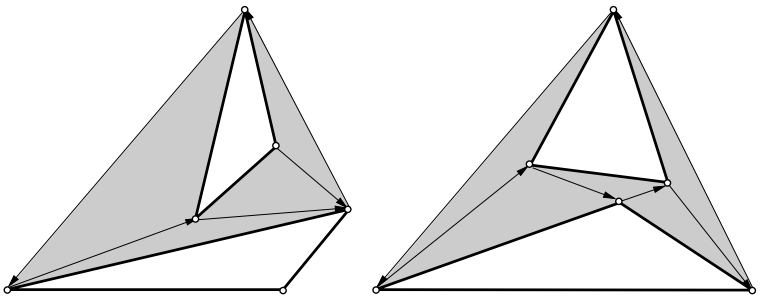

**Figure 2.17:** *The two exceptional diagrams, vertices of $P$ are depicted as small empty circles.*

and such that every face has at least three nodes on their boundary. In such a graph, the number of nodes exceed the number of faces by at least two.

Assume that not all elements of $W \cap \varphi(R_D)$ are on $\kappa$. If $\tau_1$ (or $\tau_2$) has a vertex in the interior of $\kappa$, then it necessarily has two vertices in the interior of $\kappa$ and one on $\kappa$, because $\kappa$ circumscribes a minimal number of faces in our diagram. If both $\tau_1$ and $\tau_2$ have only one-one vertex on $\kappa$ then $\kappa$ violates Lemma 2.9. Therefore, at most one of $\tau_1$ and $\tau_2$ can have one vertex on $\kappa$ and two vertices in the interior of $\kappa$. There are two possible configurations of the diagram $\varphi(R_D)$, shown in Figure 2.17, where $|W \cap \varphi(R_D)| < |D|+2$ and Lemma 2.9 is not violated. For every face $\varphi(R_v)$, $v \in D$, we have $|W \cap \varphi(R_v)| = 3$. Therefore $|W \cap R_v| = 3$ and $W \cap \varphi(R_v) = W \cap R_v$ holds, by Lemma 2.9, for every $v \in D$. One configuration has four faces, the other has five faces and forms a 5-group. This completes the proof of Lemma 2.8. □

# Binary space partitions

## 3.1 Introduction

The *binary space partition* tree (or BSP, as a shorthand) is a data structure introduced by Fuchs, Kedem, and Naylor [22], based on ideas of Schumacker et al. [43, 46]. Originally it was designed to give an efficient visibility algorithm for polygonal scenes using "painter's method". Beyond hidden surface removal, applications of BSP range through shadow generation, ray tracing, solid modeling, surface approximation, and robot motion planning.

A BSP $P$ for a set $L$ of $n$ disjoint $(d-1)$-dimensional objects in $\mathbb{E}^d$ is a rooted binary tree defined as follows. Each node $v \in P$ corresponds to a convex region $R_v$. The root of $P$ corresponds to $\mathbb{E}^d$. If $L_v = \{\ell \cap \mathrm{int}(R_v) : \ell \in L\}$ contains no object then $v$ is a leaf of $P$. Otherwise $R_v$ is partitioned into two regions $R_w$ and $R_u$ along a hyperplane and the two children $w$ and $u$ correspond to $R_w$ and $R_u$ resp.

A BSP for the set $L$ is called *autopartition*, if $R_v$ with is partitioned into two regions along the hyperplane spanned by an element of $L_v$ whenever there exists such an element. The convex regions corresponding to the leaves of $P$ give a partition of $\mathbb{E}^d$. The size of a BSP $P$ is the number of leaves of $P$.

Note that many different BSPs and autopartitions exist for the same set $L$. The size of a BSP depends largely on the choice of $\ell$ that splits $R_v$ at each node $v$. In fact, we consider a BSP for $L$ as a recursive *algorithm* of building the tree from the root, where we choose a hyperplane whenever there is a non-empty $L_v$ and split $R_v$ into two regions $R_w$ and $R_u$.

We denote by $p_d(L)$ the minimum size of a BSP for the set $L$ of disjoint $(d-1)$-dimensional objects in $\mathbb{E}^d$. Let $p_d(n) = \max\{p_L : |L| = n\}$. Denote by $r_d(L)$ the minimal size of an autopartition for $L$ and let $r_d(n) = \max\{r_L : |L| = n\}$. Clearly, $p_d(n) \leq r_d(n)$.

Paterson and Yao [39] showed in 1989 that $p_3(n) = O(n^2)$ matching the lower bound $p_3(n) = \Omega(n^2)$ described in [18]. In the same paper, they proved $r_2(n) = O(n \log n)$. The expected size of an autopartition, where the line segments are chosen in random order, is $O(n \log n)$. The authors conjectured $p_2(n) = O(n)$. I disproved this conjecture by showing the following theorem.

**Theorem 3.1.** *[52]*

$$p_2(n) = \Omega\left(n \cdot \frac{\log n}{\log \log n}\right).$$

The size of a BSP effects substantially the complexity of the algorithms using this data structure. Extensive efforts have been made to find effectively small BSPs for certain types of polygonal scenes. Linear bounds were found to the BSP for line segments which have only two different directions by Paterson and Yao [38], and for line segments where the ratio between the lengths of the longest and shortest segment is bounded by a constant by de Berg et al. [17]. Agarwal et al. [1] constructed a BSP of size $n2^{O(\sqrt{\log n})}$ for $n$ fat orthogonal rectangles in $\mathbb{E}^3$ in $n2^{O(\sqrt{\log n})}$-time.

Our Theorem 3.1 shows that we cannot expect linear BSPs for disjoint line segments in the plane. We may ask what is the smallest BSP if the segments have only few different orientations compared to the number of segments, which is the case in many applications. Paterson and Yao [39] proved an $O(n)$ upper bound on the smallest BSP for orthogonal line segments. Recently Dumitrescu et al. improved this bound to $2n - 1$, and gave a worst case lower bound of $2n - o(n)$. For more than two different directions (even for three distinct orientations) the best previously known bound was $O(n \log n)$, from Patersen and Yao. We have the following.

**Theorem 3.2.** *[51] If $L$ is a set of $n$ disjoint line segments in the plane with at most $k$ distinct directions, then there exists for $L$*

- $(i)$ *a BSP of size $O(n \log k)$, where every segment is dissected at most $O(\log k)$ times;*
- $(ii)$ *an autopartition of size $O(nk)$, where every segment is dissected at most $O(k)$ times.*

If $k$ is a constant, Theorem 3.2 yields a linear upper bound, which is, up to a constant factor, optimal. Interestingly enough, the bound of Theorem 3.2 $(i)$ is also asymptotically tight for the lower bound construction for our Theorem 3.1: There $n$ line segments have $\exp(\Theta(\log n / \log \log n))$ different directions and the minimum size BSP is as big as $\Theta(n \log n / \log \log n) =$

$\Theta(n \log k)$. For $k = n$, Theorem 3.2 $(i)$ returns the well known bound of $O(n \log n)$ due to Paterson and Yao [39].

Our proof is constructive. The procedures of constructing a BSP of size $O(n \log k)$ and an autopartition of size $O(nk)$ can easily be turned into algorithms using $O(n^2)$ and $O(n^2 k)$ time, respectively. Although, there could be algorithms with smaller complexity achieving the same bounds on the size of a BSP or an autopartition. For $k = 2$, d'Amore and Franciosa [16] gave a $O(n \log n)$ time algorithm to find an autopartition of size at most $4n$.

In Section 3.2, we describe a series of constructions of $n_k$ disjoint line segments with $n_k = 2^{\Theta(k \log k)}$ for all $k \in \mathbb{N}$. Subsection 3.2.3 contains the proof of Theorem 3.1. A crucial step of the proof, Lemma 3.1, is postponed to Subsection 3.2.4.

In Section 3.3, we describe a simple algorithm building an autopartition for a set $L$ of $n$ disjoint line segments with $k$ distinct directions. Instead of the number of leaves of the partition tree, we count the number of *cuts* on line segments, i.e., the events where a splitting line of an interior node $v \in P$ dissects some segment $\ell \in R_v$ into two pieces. Our Lemma 3.5 assures that every line segment is *cut* at most $O(\log k)$ and $O(k)$ times for BSP and for autopartitions respectively. The proof of Lemma 3.5 is in Sections 3.3.3 and 3.3.4. Section 3.3.2 discusses a key concept in our argument, that of *cycles* and *convex cycles*.

## 3.2 Lower bound construction

We start with a brief overview of the construction before we give the precise description. We define recursively $L^k$, a set of $n_k = 2^{\Theta(k \log k)}$ disjoint line segments, for every $k \in \mathbb{N}$. Every set $L^k$ is given together with a *bounding box* $B_k$. Every bounding box $B_k$ is partitioned by two horizontal lines into three parts. The upper part and the lower part are called *head-box* and *tail-box* of $L^k$ and denoted by $H_k$ and $T_k$, respectively.

The first construction, $L^1$, consists of four line segments as shown in Figure 3.1. Every subsequent construction $L^k$, $k > 1$, is composed of congruent copies of $L^{k-1}$ and some more line segments. The copies of $L^{k-1}$ are deformed before they are built into $L^k$ in two ways: On one hand, the bounding box $B_{k-1}$ is transformed by an affine transformation to be long and skinny; on the other hand, the head-box and the tail-box of each copy is very small compared to the height of the bounding box, which is achieved by setting a free parameter of $L^{k-1}$ appropriately.

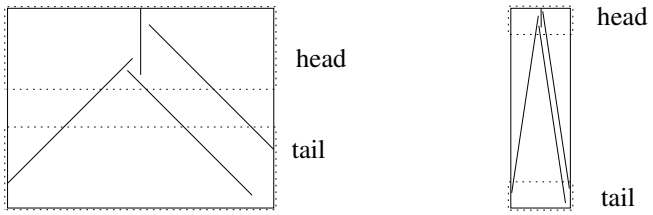


**Figure 3.1:** *Two copies of $L^1$ with their bounding boxes, head-boxes and tail-boxes. On the left: in the form as they appear in our descriptions, on the right in deformed form as they are used in the recursion.*

Most importantly, the intersection of any two lines spanned by segments in the same copy of $L^{k-1}$ intersect in the head-box of this copy of $L^{k-1}$, therefore a BSP for the segments clipped to the head-box of a copy of $L^{k-1}$ has the same size as a BSP for the full $L^{k-1}$. The head-box, however, is very small compared to the full $L^{k-1}$ and most of the line segments in $L^{k-1}$ connect the head-box and the tail-box. In the construction $L^k$, we can regard the copies of $L^{k-1}$ as a bunch of almost parallel line segments and, independently, a structure in the small head-box whose BSPs are just as complex as that of $L^{k-1}$. A splitting line, which is useful for a partitioning of one copy of $L^{k-1}$, will cut many other copies of $L^{k-1}$ between their head-boxes and tail-boxes. This will allow us to consider $k$ levels of $L^k$ separately, where a level $i$, $1 \leq i \leq k$, is formed by the head-boxes of all copies of $L^i$ within $L^k$. Our theorem is established, essentially, by showing that the size of the BSP is $O(n)$ in each level and the number of levels is $k = \Theta(\log n / \log\log n)$.

To describe how the line segments and the copies of $L^{k-1}$ are arranged in $L^k$, we use a common auxiliary construction $A(m, f, \delta)$. Here $m$ and $f$ denotes the number of line segment and copies of a previous construction (each represented by a directed line segment), $\delta$ denotes the ratio of the height of the head-box $H$ and its complement $B \setminus H$. The head-box and the tail-box will have the same height.

### 3.2.1 Auxiliary construction

We define the auxiliary construction $A(m, f, \delta)$ for all $m, f \in \mathbb{N}$, and $\delta$, $0 < \delta < 1$. $A(m, f, \delta)$ is composed of $m$ disjoint line segments and $3f$ disjoint directed line segments, which we call *arrows*. Let the bounding box of $A(m, f, \delta)$ be $B = [(-1, -1), (1, -1), (1, \delta), (-1, \delta)]$ where the head-box $H$

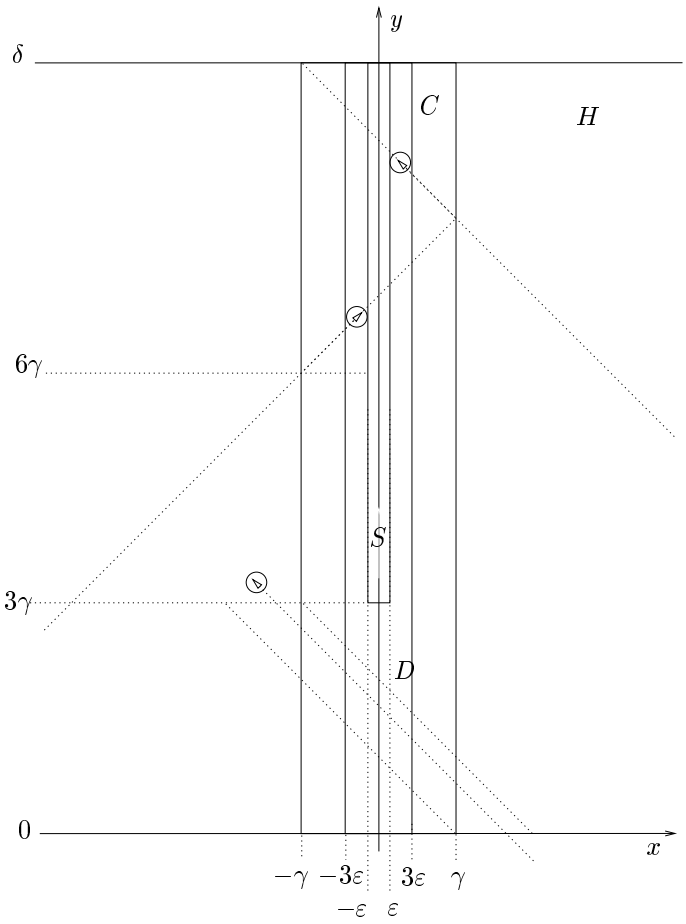


**Figure 3.2:** *The relative position of the boxes $H$, $C$, $D$, and $S$. The circles represent $\varepsilon$-disks around the heads of the three arrows of an $A(m, 1, \delta)$. (The condition on the choice of $\varepsilon$ is not enforced in this Figure.)*

is the part of $B$ above the $x$-axis, and the tail-box $T$ is the part of $B$ below the line $y = -1+\delta$. We introduce three new parameters: $\gamma = \delta/(4f+6)$, and two very small parameters $\varepsilon$, $0 < \varepsilon \leq \gamma/3$, specified later, and $\varphi$, $0 < \varphi \leq \gamma/(2f)$ chosen arbitrarily in this interval. We define three skinny rectangles in the head-box. The relative position of $B$, $H$, $C$, $D$, and $S$ is indicated in Figure 3.2. Let

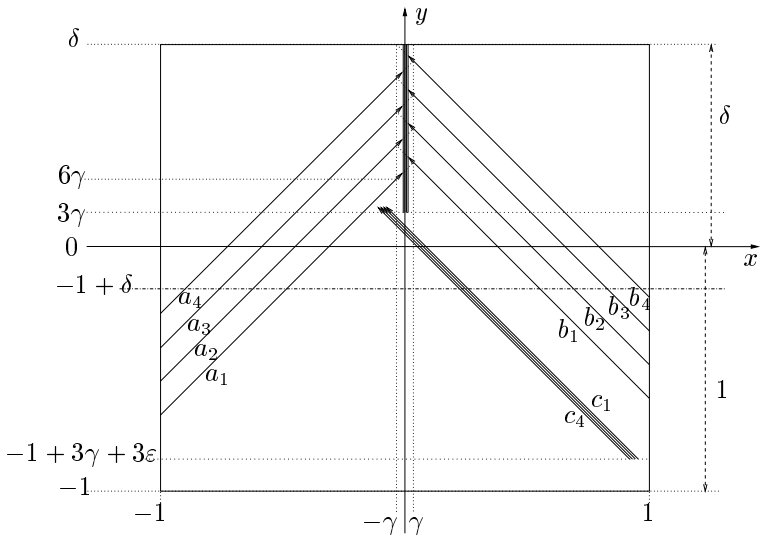


**Figure 3.3:** $A(4, 4, \delta)$ *with* $4$ *vertical line segments and* $3 \cdot 4 = 12$ *arrows.*

$$C = [(-\gamma, 0), (\gamma, 0), (\gamma, \delta), (-\gamma, \delta)],$$

$$D = [(-3\varepsilon, 0), (3\varepsilon, 0), (3\varepsilon, \delta), (-3\varepsilon, \delta)], \text{ and}$$

$$S = [(-\varepsilon, 3\gamma), (\varepsilon, 3\gamma), (\varepsilon, \delta), (-\varepsilon, \delta)].$$

Place $m$ vertical line segments of length $\delta - 3\gamma$ in $S$. Place three branches of $f$ parallel arrows as follows. Every arrow has slope $1$ or $-1$ and length $\sqrt{2}(1 - 2\varepsilon)$. For every $i$, $i = 1, \ldots, f$, let

$$a_i = [(-1, -1 + (4i + 3)\gamma), (-2\varepsilon, (4i + 3)\gamma - 2\varepsilon)],$$

$$b_i = [(1, -1 + (4i + 5)\gamma), (2\varepsilon, (4i + 5)\gamma - 2\varepsilon)], \text{ and}$$

$$c_i = [(1 - 3\gamma - i\varphi - 3\varepsilon, -1 + 3\gamma + 3\varepsilon), (-3\gamma - i\varphi - \varepsilon, 3\gamma + \varepsilon)].$$

See Figure 3.3 for an illustration.

W set two conditions on the size of $\varepsilon$, $0 < \varepsilon \leq \gamma/3$: Let $\varepsilon$ be so small that any line $\ell$ intersecting two $\varepsilon$-disks around heads of two arrows $a_i$ and $b_j$ must intersect both the top and bottom sides of $C$; and any line $\ell$ intersecting an $\varepsilon$-disk around heads of arrows $a_i$ or $b_j$ and the top or the bottom side of $D$ must

intersect both the top and the bottom sides of $C$. The choice of $\varphi$ is arbitrarily within the bounds $0 < \varphi \leq \gamma/(2f)$. The parameter $\varphi$ determines only the spacing between the arrows $c_i$, which is used only in Remark 3.2 below. This completes the construction of $A(m, f, \delta)$.

Observe that the arrows and line segments are disjoint; the vertical line segments are completely in the head-box, the head (resp., tail) of every arrow is in the head-box (resp., tail-box). Moreover, the head of every $c_i$ is above the line $y = 3\gamma$ (i.e., the bottom side of the box $S$) at length $\varepsilon$; and the head of every $a_i$ and $b_i$ enters $D$ at length $\sqrt{2}\varepsilon$.

Note also that the intersection of any two segments or arrows of $A(m, f, \delta)$ is in the head-box. In particular, the intersection point of lines through $a_i$ and $b_i$ is on the line $x = \gamma$ (the right side of the box $C$) for every $i$, $i = 0, \ldots, f$, and the intersection point of lines through $b_i$ and $a_{i+1}$ is on the line $x = -\gamma$ (the left side of the box $C$) for every $i$, $i = 0, \ldots, f-1$.

The following two remarks will be essential for the proof of Theorem 3.1.

**Remark 3.1.** Consider a construction $A(m, f, \delta)$. For a constant $d$, $\delta < d < 1$, consider the rectangle $R = [(-d, -d + 3\gamma + 3\varepsilon), (d, -d + 3\gamma + 3\varepsilon), (d, \delta), (-d, \delta)]$ (see Figure 3.4). Observe that the (portions of) line segments and arrows in $R$ form an $A(m, f, \delta/d)$ after appropriate scaling. This implies that after partitioning $A(m, f, \delta)$ by a horizontal line below $y = -\delta$, the upper part has essentially the same structure as $A(m, f, \delta)$.

**Remark 3.2.** Consider a construction $A(m, f, \delta)$. For an integer $g$, $0 < g \leq f$, consider the convex hexagon $R(g)$ defined by the following six points: $(-1, -1)$, $(1, -1)$, $(1, -1 + (6 + 4f - 4g)\gamma + \varepsilon)$, $(3\varepsilon, (6 + 4f - 4g)\gamma - 2\varepsilon)$, $(-3\varepsilon, (6 + 4f - 4g)\gamma - 2\varepsilon)$, and $(-1, -1 + (6 + 4f - 4g)\gamma + \varepsilon)$. $R(g)$ is shown in Figure 3.4 as a dashed hexagon. Now let $A_g(m, f, \delta)$ be the set of the (portions of) segments and arrows of $A(m, f, \delta)$ in $R(g)$ with the exception of $c_{g+1}, c_{g+2}, \ldots, c_f$. Observe that $A_g(m, f, \delta) = A(m, g, \delta)$. We note here that the only reason why we allow the parameter $\varphi$ to vary between 0 and $\gamma/(2f)$ is to establish this equality, thus the spacing between the arrows $c_i$ is not determined by the total number of arrows in the construction.

### 3.2.2 Main construction

In this subsection we describe recursively the constructions $L^k$ for every $k \in \mathbb{N}$. We leave a free parameter $\delta_k$ in the definition of $L^k$ which determines the ratio of the height of its head-box and the complement of its head-box. For any

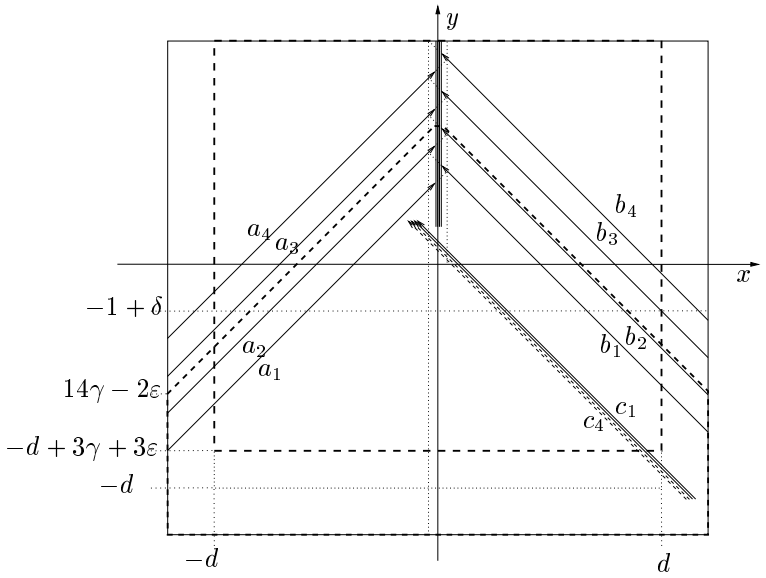


**Figure 3.4:** *The rectangle $R$ from Remark 3.1 and the hexagon $R(g)$ from Remark 3.2 are drawn with bold dashed lines.*

$\delta_1$, $0 < \delta_1 < 1$, the construction $L^1$ is obtained from $A(1, 1, \delta_1)$ by replacing each arrow by a line segment as depicted on the left of Figure 3.1.

Suppose that $L^{k-1}$ is defined for every $\delta_{k-1}$, $0 < \delta_{k-1} < 1$. We construct $L^k$ for any $\delta_k$, $0 < \delta_k < 1$. Let $m_{k-1}$ denote the number of line segments of $L^{k-1}$ which intersect both the head-box $H_{k-1}$ and the tail-box $T_{k-1}$. Clearly $m_1 = 3$.

Consider $A^k := A(m_{k-1}, k^2, \delta_k)$ together with its bounding box, head-box, and tail-box. Place a long and narrow box along the arrow of $A^k$ such that the boxes are congruent, pairwise disjoint, and disjoint from the line segments in $S$. Place a copy of $L^{k-1}$ in each box corresponding to an arrow such that the head-box of $L^{k-1}$ is in the $\varepsilon/3$-disk centered at the head of the arrow and the tail-box of $L^{k-1}$ is in the $\varepsilon/3$-disk centered at the tail of the arrow. This can be done, since a copy of $L^{k-1}$ is determined by two independent parameters: $\delta_{k-1}$ and the aspect ratio of the bounding box. $\delta_{k-1}$ can be arbitrarily small and, independently, the aspect ratio can be arbitrary by affine transformation. Every copy has the same ratio $\delta_{k-1}$ and the same aspect ratio. This completes the description of the construction of $L^k$.

We show by induction that $m_k = 3^k(k!)^2$ for every $k \in \mathbb{N}$. We know that $m_1 = 3^1(1!)^2 = 3$. By construction,

$$m_k = 3k^2 \cdot m_{k-1} = 3k^2 \cdot 3^{k-1}((k-1)!)^2 = 3^k(k!)^2.$$

The total number of line segments in $L^k$ is

$$n_k = m_{k-1} + 3k^2 \cdot n_{k-1}.$$

$n_k - m_k$ denotes the number of line segments fully in the head-box $H_k$. There are $m_{k-1}$ such vertical line segments in the box $S$, and $n_{k-1} - m_{k-1}$ line segment in the head-box of each of the $3k^2$ copies of $L^{k-1}$. This gives the recursion formula

$$n_k - m_k = m_{k-1} + 3k^2 \cdot (n_{k-1} - m_{k-1}).$$

An induction with initial value $n_1 - m_1 = 1$ establishes

$$n_k - m_k = \frac{m_k}{3} \cdot \sum_{i=1}^{k} \frac{1}{i^2} < m_k.$$

That is, $m_k < n_k < 2m_k$. Using the Stirling-formula, we have

$$\Theta(n_k) = m_k = 3^k \cdot (k!)^2 =$$

$$= 2\pi \left(\frac{3}{e^2}\right)^k \cdot k^{2k+1} \left(1 + O\left(\frac{1}{k}\right)\right)^2 = e^{\Theta(k)+(2k+1)\log k},$$

$$k = \Theta\left(\frac{\log n_k}{\log\log n_k}\right).$$

**Remark 3.3.** Let $t_k$ denote the number of distinct directions of line segments in $L^k$. Clearly $t_1 = 3$ (see Figure 3.1). By construction, we have the recursion formula $t_k = 2t_{k-1} + 1$. Therefore,

$$t_k = \sum_{i=0}^{k} 2^i = 2^{k+1} - 1 = 2^{\Theta\left(\frac{\log n_k}{\log\log n_k}\right)} = \exp\left(\Theta\left(\frac{\log n_k}{\log\log n_k}\right)\right).$$

### 3.2.3 Autopartitions of $L^k$

In this subsection, we prove Theorem 3.1. A crucial step of the proof, Lemma 3.1, is postponed to the next subsection. The machinery introduced here is, however, sufficient for the analysis of the autopartitions of the construction $L^k$. In our analysis, we consider the number of occasions where a splitting line or line segments of $P$ cuts a segment in $L$, and denote this quantity by $|P|$. The size of a BSP $P$ for a set $L$ of line segments is $|L| + |P|$, therefore a superlinear lower bound on $|P|$ implies a the same lower bound on the size of the BSP.

We study first the BSPs of the auxiliary construction. Fix an $A(m, f, \delta)$ and a BSP $P_0$. We say that a line (or line segment) $\ell_0$ chosen by $P_0$ *cut*s a line segment $\ell$ if $\ell_0$ intersects $\ell$. We have to define very carefully what we call a cut of an arrow $a$, because later we wish to say that a cut of $a$ is equivalent to $m_{k'}$ cuts of line segments in the copy of $L^{k'}$ placed along $a$. Let $N(a)$ be the $\varepsilon/3$-neighborhood of the arrow $w$ (i.e., $N(a) = \{p \in \mathbb{R}^2 : \mathrm{dist}(p, w) < \varepsilon/3\}$) where $\varepsilon$ is the parameter $0 < \varepsilon < \delta/(3(4f + 6))$ used in the description of $A(m, f, \delta)$. A cut of an arrow $a$ is a polygonal path $h$ along lines or segments chosen by $P_0$ such that $h$ traverses $N(a)$ but $h$ does not intersect either of the two $\varepsilon$-disks assigned to the endpoints of $a$. We consider two cuts of an arrow $a$ *distinct* if they correspond to non-intersecting polygonal paths.

For every line or line segment $\ell$ chosen by $P_0$ during the partitioning algorithm, let $w(\ell)$ (resp., $s(\ell)$) denote the maximum number of cuts on arrows (resp., on segments) of $A(m, f, \delta)$ until and including the step where $P_0$ chooses $\ell$. Denote by $w(P_0)$ (resp., $s(P_0)$) the maximum number of cuts on arrows (resp., on segments) of $A(m, f, \delta)$ at the end of the partitioning algorithm $P_0$.

**Lemma 3.1.** *Given an $A(m, f, \delta)$ and a BSP $P_0$ for the $m$ vertical segments of $A(m, f, \delta)$, the following inequality holds:*

$$w(P_0) \cdot m + s(P_0) \geq f \cdot m. \tag{3.1}$$

For the proof of Lemma 3.1, see the next subsection. We are now ready to prove Theorem 3.1.

**Lemma 3.2.** *For any BSP $P$ for $L^k$, we have $|P| \geq k \cdot m_k/3$.*

*Proof.* We precede by induction on $k$. Let $P$ be a BSP for $L^k$. In the case $k = 1$, clearly, $|P| \geq 1$ and $1 \cdot m_1/3 = 1/3$. As $P$ is also a BSP for the vertical line segments of the auxiliary construction $A^k$, we may use Lemma 3.1. A cut

of an arrow in $A^k$ corresponds to exactly $m_{k-1}$ cuts of line segments in $L^k$. By Lemma 3.1, $w(P)m_{k-1} + s(P) \geq k^2 m_{k-1} = m_k/3$, where $w(P)$ and $s(P)$ denote the number of cuts of arrows and segments of $A^k$ respectively. Therefore, $A^k$ contributes at least $m_k/3$ cuts of line segments of $L^k$.

By the definition of a cut of an arrow, we do not count the number of cuts in the $\varepsilon$-disks around the heads of arrows. That is, we do not count the number of cuts within the head-boxes of the copies of $L^{k-1}$ placed along the arrows in $A^k$. Using Remark 3.1, the $\varepsilon$-disk around the head of every arrow contains a copy of $L^{k-1}$. $P$ is a BSP for each of them, and each contributes at least $(k-1) \cdot m_{k-1}/3$ cuts by induction. In total, we have $3k^2 \cdot (k-1) \cdot m_{k-1}/3 + m_k/3 = k \cdot m_k/3$ distinct cuts on line segments of $L^k$. □

Lemma 3.2 immediately implies Theorem 3.1, since $m_k = \Theta(n_k)$ and $k = \Theta(\log n_k / \log\log n_k)$. We prove here an analogue of Lemma 3.1 for autopartitions.

**Lemma 3.3.** *For any* autopartition $P_0$ *for* $A(m, f, \delta)$*, we have*

$$w(P_0) \cdot m + s(P_0) \geq f \cdot m. \tag{3.2}$$

*Proof.* We proceed by induction on $f$. Fix a construction $A(m, f, \delta)$ and an autopartition $P_0$ for it. Let $\ell$ be the first segment or arrow chosen by $P_0$.

If $\ell$ is one of the vertical segments in $S$ or if $\ell = c_i$ for some $i$, $1 \leq i \leq f$, then $\ell$ cuts exactly $f$ arrows. Therefore $w(P_0) \geq w(\ell) = f$ and Inequality (3.2) follows.

If $\ell = a_i$ for some $i$, $1 \leq i \leq f$, then $\ell$ cuts all $m$ line segments in $S$ and the arrows $b_i, b_{i+1}, \ldots, b_f$. If $\ell = b_i$ for some $i$, $1 \leq i \leq f$, then $\ell$ cuts all $m$ line segments in $S$ and the arrows $a_{i+1}, a_{i+2}, \ldots, a_f$. In both cases, $w(\ell) \cdot m + s(\ell) \geq (f - i + 1) \cdot m$.

This implies Inequality (3.2) for $i = 1$. If $i > 1$, then observe that $\ell$ does not cut any line segment nor arrow in the hexagon $R(i-1)$. In particular, $\ell$ does not cut the construction $A_{i-1}(m, f, \delta)$. By Remark 3.2, $A_{i-1}(m, f, \delta) = A(m, i-1, \delta)$ and we can use the induction hypothesis for $A(m, i-1, \delta)$. We obtain $w(P_0) \cdot m + s(P_0) \geq (f - i + 1) \cdot m + (i-1) \cdot m \geq f \cdot m$. □

If we notice that the line spanned by a line segment within a copy of $L^{k-1}$ placed along an arrow $a$ of $A(m, f, \delta)$ *cut*s the same segments and arrows of $A(m, f, \delta)$ as the line through $a$, then we immediately conclude the following.

**Corollary 3.3.** *Given an* autopartition $P$ *for* $L^k$*, we have*

$$w(P) \cdot m_{k-1} + s(P) \geq k^2 \cdot m_{k-1} = m_k/3,$$

*where* $w(P)$ *and* $s(P)$ *denotes the maximum number of cuts on the segments and arrows of the auxiliary construction* $A^k = A(m_{k-1}, k^2, \delta_k)$*.*

**Corollary 3.4.**

$$r_2(n) = \Omega\left(n \cdot \frac{\log n}{\log\log n}\right).$$

**Remark 3.4.** It is not difficult to see that $r_2(L^k) = O(n_k \log n_k / \log\log n_k)$. One possible autopartition $P^k$ for $L^k$ within this bound can be the described recursively: First choose the leftmost vertical line segment in $S$. It cuts $k^2$ copies of $L^{k-1}$ placed along $c_1, c_2, \ldots, c_{k^2}$, that is, it cuts $m_{k-1} = m_k/3$ line segments. Second, if $k = 1$ then choose the segment along $a_1$ and $c_1$. Otherwise apply the autopartition $P^{k-1}$ (defined recursively) to the copies of $L^{k-1}$ along $a_1, a_2, \ldots, a_{k^2}$, then to the copies along $c_1, c_2, \ldots, c_{k^2}$ in this order. Third, choose consecutively the remaining $m_{k-1} - 1$ vertical line segments in $S$. Finally if $k = 1$ then choose $b_1$; otherwise apply the autopartition $P^{k-1}$ to the copies of $L^{k-1}$ along $b_1, b_2, \ldots, b_{k^2}$. In this way, line segments in copies of $L^{k-1}$ do not cut any line segment in other copies of $L^{k-1}$ nor line segments in $S$. The total number of cuts is the sum of $m_{k'}/3$ for all copies $L^{k'} \subset L^k$, $1 \leq k' \leq k$, which totals to $k \cdot m_k/3$.

### 3.2.4 Binary plane partitions of $L^k$

The proof of Lemma 3.1 follows the lines of that of Lemma 3.3. When partitions along arbitrary lines are allowed, we need to define and use polygonal curves along those lines, since any single splitting line might fail to cut arrows in the sense described in the beginning of the previous subsection.

The box $D$ divides the head-box $H$ into three rectangles $D^-$, $D$, and $D^+$. Fix a BSP $P_0$ for $L^k$. Consider a splitting line (or line segment) $\ell$ chosen by $P_0$ which intersects the box $D$. We define an $x$-monotone polygonal curve $\alpha(\ell)$ within the box $D$ and two $y$-monotone polygonal curves $\beta^-(\ell)$ and $\beta^+(\ell)$ within the boxes $D^-$ and $D^+$, respectively. (A plane curve is *$x$-monotone* (*$y$-monotone*) if its intersection with every line perpendicular to the $x$-axis ($y$-axis) is connected.)

**Definition 3.1.** *Set a starting point* $p \in \ell$*, put* $j = 0$*, and* $\ell_0 = \ell$*. We define separately the curves* $\alpha^-(\ell)$ *and* $\alpha^+(\ell)$ *on the left and on the right of p. The*

*full curve is $\alpha(\ell) := \alpha^-(\ell) \cup \alpha^+(\ell)$. For every $j \geq 0$, $\alpha^{\pm}(\ell)$ follows $\ell_j$ until it reaches another line or segment $\ell_{j+1}$ which was chosen by $P_0$ prior to $\ell_j$, or until it reaches the boundary of $D$. If it reaches $\ell_{j+1}$ then $\alpha^-(\ell)$ (resp., $\alpha^+(\ell)$) continues along $\ell_{j+1}$ in right-to-left (resp., left-to-right) direction when $\ell_{j+1}$ is not vertical, and in ascending (descending) direction when $\ell_{j+1}$ is vertical. (See Figure 3.6.)*

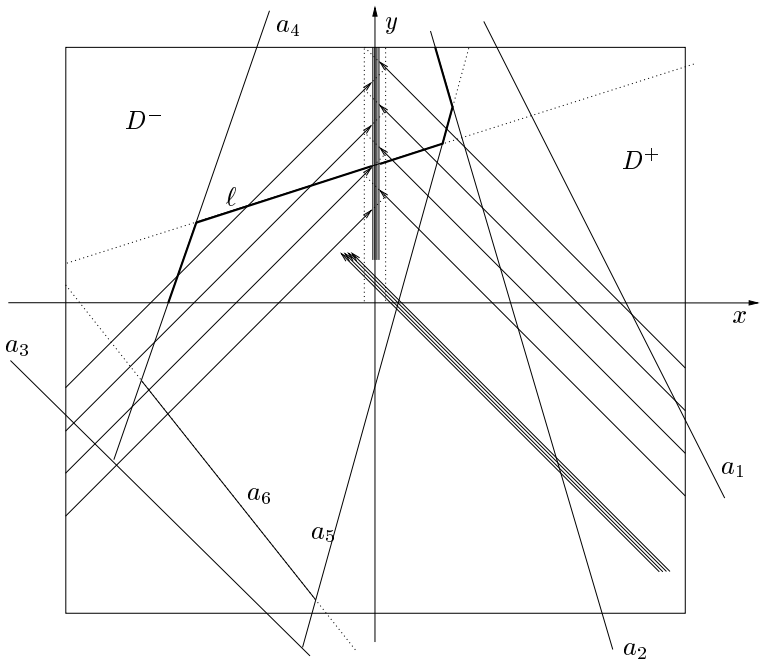


**Figure 3.5:** *A splitting line $\ell$ with positive slope, lines $a_1, a_2, \ldots, a_6$ chosen prior to $\ell$, and the curves $\beta^-(\ell)$ and $\beta^+(\ell)$.*

**Definition 3.2.** *If $\alpha(\ell)$ intersect the left (resp., right) side of $D$, then $\beta^-(\ell)$ (resp., $\beta^+(\ell)$) is defined in $D^-$ (resp., $D^+$) as follows.*

*Put $j = 0$ and let $\ell_0$ be the segment of $\alpha(\ell)$ which intersects the left (resp., right) side of $D$ at a point $p^-$ (resp., $p^+$). Set the starting point of $\beta^-(\ell)$ (resp., $\beta^+(\ell)$) to be $p^-$ (resp., $p^+$). For every $j \geq 0$, $\beta(\ell)$ follows $\ell_j$ until it reaches another line $\ell_{j+1}$ which was chosen by $P_0$ prior to $\ell_j$, or until it reaches the boundary of $D^-$ (resp., $D^+$). If $\beta^{\pm}(\ell)$ followed $\ell_j$ in ascending*

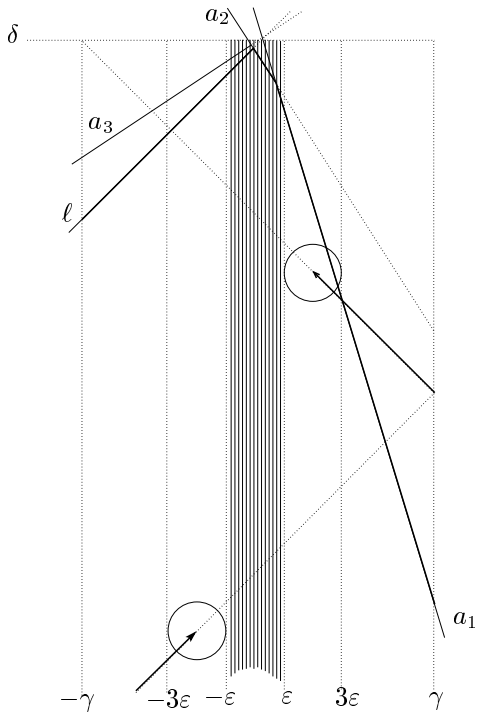


**Figure 3.6:** *$\ell$ is of Type 3, $\ell$ does not intersect all $m$ segments, but $\alpha(\ell)$ does.*

*(resp., descending) direction, then it continues along $\ell_{j+1}$ in ascending (resp., descending) direction. (See Figure 3.5 for an illustration.)*

The $x$-monotone and $y$-monotone curves are well defined, because all faces are convex in every step of the partitioning algorithm $P_0$. Note that if $\alpha(\ell)$ intersects both the top and the bottom sides of $D$, then $\beta(\ell)$ is not defined. We distinguish five cases depending on how $\ell$ intersects $D$.

Type 1, the line through $\ell$ intersects the lower and upper side of $C$.

Type 2, $\ell$ is not of Type 1 and $\ell$ intersects $D$ below the line $y = 3\gamma$.

Type 3, $\ell$ is not of Type 1 nor of Type 2; and $\alpha(\ell)$ intersects the left and the right side of $D$.

Type 4, $\ell$ is not of Type 1, 2, nor 3; but $\ell$ intersects $D$.

Type 5, $\ell$ does not intersect $D$.

Let us rephrase the properties of a line $\ell$ of Type 4: the line through $\ell$ intersects the left or the right side of $C$, $\ell$ intersects the box $D$ only above the line $y = 3\gamma$, and $\alpha(\ell)$ does not intersect both the left and the right side of $D$. Finally, we introduce one more important parameter, $\eta = \delta - \gamma - 2\varepsilon$, which is the $y$-coordinate of the head of the arrow $b_f$. Observe that $\varepsilon$ was chosen such that the line spanned by the segment $[(-3\varepsilon, \eta), (3\varepsilon, \delta)] \subset D$ intersects the top and the bottom sides of $C$.

**Lemma 3.4.** *Suppose that a BSP chooses lines and segments of Type 4 and 5 until (but not necessarily including) the step when it chooses $\ell$.*

*(A) If $\ell$ is of Type 4, then it intersects $D$ only above the line $y = \eta$.*

*(B) If $\alpha(\ell)$ intersects the left (resp., right) side of $D$ below the line $y = \eta$, then the region between the $y$-axis and $\beta^-(\ell)$ ($\beta^+(\ell)$) is convex below the line $y = \eta$.*

*Proof. (A)* As long as $P_0$ chooses lines and segments of Type 4 and 5, no $x$-monotone path traverses $D$ between its left and right sides. The curves $\alpha(\ell)$ and $\beta^{\pm}(\ell)$, if defined, are composed of $\ell$ and portions of lines and segments of Type 4 or 5. This implies that $\alpha(\ell)$ cannot intersect $D$ below the line $y = 3\gamma$. If $\ell$ is of Type 4, then the $x$-monotone path $\alpha(\ell)$ must end at the top side of $D$.

By induction, we may assume that no segment chosen by $P_0$ prior to $\ell$ intersects $D$ below $y = \eta$. Supposing that $\ell$ intersects $D$ below $y = \eta$, $\alpha(\ell)$ must coincide with $\ell$ below $y = \eta$. We may assume w.l.o.g. that $\ell$ has positive slope, therefore it intersects the left side of $D$ below $y = \eta$. The slope of $\ell$ is bigger than the slope of the line segment $[(-3\varepsilon, \eta), (3\varepsilon, \delta)]$. Therefore, the line spanned by $\ell$ intersects the top and bottom side of $C$, a contradiction. This completes the proof of part *(A)*.

*(B)* It suffices to show that, at every vertex of $\beta^-(\ell)$ (resp., $\beta^+(\ell)$) below $y = \eta$, the angle facing towards the $y$-axis is convex. Suppose w.l.o.g. that $\beta^+(\ell)$ is $y$-monotone increasing and $p$ is its first vertex below $y = \eta$ where the angle facing towards the $y$-axis is reflex. That means that $\beta^+(\ell)$ arrived to $p$ along some $\ell_q$ in $y$-monotone increasing direction and it continues along $\ell_{q+1}$ chosen by $P_0$ prior to $\ell_q$. We define an $y$-monotone decreasing curve $\beta^*$ staring from $p$: Set $u_0 = \ell_{q+1}$ and $j = 0$; for every $j \geq 0$, $\beta^*$ follows $u_j$ until it reaches another line $u_{j+1}$ which was chosen by $P_0$ prior to $u_j$, or until it reaches the boundary of $D^+$. Notice that $\beta^*$ cannot intersects the curve $\beta^+(\ell)$, because an intersection point $p'$ should be below $p$ and, by the definition, $\beta^+(\ell)$ should have followed $\beta^*$ from $p'$. Therefore, $\beta^+(p)$ must intersect the right side of $D$ below $y = \eta$ contradicting part *(A)* of our Lemma 3.4. □

*Proof (of Lemma 3.1).* Let $P_0$ be a BSP for the vertical segments in the box $S$ of $A(m, f, \delta)$. From Lemma 3.4 (A), we conclude that lines and segments of Type 4 and 5 cannot give a BSP for the vertical line segments in $S$, hence $P_0$ necessarily chooses a line or a segment of Type 1, 2, or 3. Let $\ell$ be the first line or segment chosen by $P_0$ of Type 1, 2, or 3. We distinguish three cases according to the type of $\ell$:

CASE 1: $\ell$ **is of Type 1.** In this case, we show that $\ell \cup \beta(\ell)$ cuts every $c_i$, $i = 1, 2, \ldots, f$. This implies $w(\ell) \geq f$ and hence Inequality (3.1). It is enough to show that $\ell$ or $\beta(\ell)$ traverses the box $E = [(-\gamma, 0), (\gamma, 0), (\gamma, 3\gamma), (-\gamma, 3\gamma)]$ between its top and bottom sides, because every $c_i$, $i = 1, 2, \ldots, f$, traverses $E$ between its left and right sides. Assuming, w.l.o.g., that the slope of $\ell$ is non-negative. $\alpha(\ell) \cap E = \ell \cap E$, since by Lemma 3.4 (A) the lines chosen by $P_0$ prior to $\ell$ intersect $D$ only above the top side of $E$. If $\ell$ itself does not intersects the top or bottom side of $E \cap D$, then it intersects the left or right (or both) side of $E \cap D$ and the curve $\beta^{\pm}(\ell)$ is defined. The $y$-monotone curve $\beta^{-}(\ell)$ (resp., $\beta^{+}(\ell)$) intersects the line $y = 0$ (resp., $y = 3\gamma$) and, by Lemma 3.4 (B), the intersection is closer to the $y$-axis than the intersection with the line spanned by $\ell$. Since the line spanned by $\ell$ traverses $E$ between its top and bottom sides, $\ell \cup \beta(\ell)$ traverses $E$ between its top and bottom sides, too.

CASE 2: $\ell$ **is of Type 2.** We show that $\beta(\ell)$ cuts every $a_i$ or every $b_i$, $i = 1, 2, \ldots, f$. This implies that $w(\ell) \geq f$ and hence Inequality (3.1). Assume first that the slope of $\ell$ is non-negative. $\ell$ intersects the right side of $D$, because, by the choice of $\varepsilon$ and Lemma 3.4 (A), it cannot intersect any line nor segment chosen by $P_0$ prior to $\ell$. Therefore the curve $\beta^{+}(\ell)$ is defined in $D^{+}$ and it is $y$-monotone increasing. The endpoint of $\beta^{+}(\ell)$ can be on the top side of $D^{+}$ or on the right side of $d^{+}$ above the $x$-axis; it can also be on the right side of $D^{+}$, but then by Lemma 3.4 (A) it must be above the head of the highest arrow $b_f$. By all means, $\beta^{+}(\ell)$ cuts all the arrows $b_1, b_2, \ldots, b_f$. Similarly, if the slope of $\ell$ is negative, $\beta^{-}(\ell)$ cuts all the arrows $a_1, a_2, \ldots, a_f$.

CASE 3: $\ell$ **is of Type 3.** In this case, $\alpha(\ell)$ cuts all $m$ line segments of $S$.

First, assume that the slope of $\ell$ is non-negative. Suppose that $\alpha(\ell)$ intersects the left side of $D$ at a height between $(4i + 3)\gamma - 2\varepsilon$ and $(4i + 7)\gamma - 2\varepsilon$ (i.e., the heights of the heads of $a_i$ and $a_{i+1}$). If $i = f$, then this implies $w(\ell)b + s(\ell) \geq b$. If $i < f$, then we follow the curves $\alpha(\ell)$ and $\beta^{+}(\ell)$. By the choice of $\varepsilon$ and Lemma 3.4 (A), $\alpha(\ell) = \ell$ (this does not necessarily hold for $i = f$). Hence the curve $\beta^{+}(\ell)$ is $y$-monotone increasing, too. Possibly $\alpha(\ell)$ intersects an $\varepsilon$-disk around the head of $b_i$. Similarly to the argument in our case 2, the curve $\beta^{+}(\ell)$ must cut all of $b_{i+2}, b_{i+3}, \ldots, b_f$ in $D^{+}$. Therefore,

$w(\ell)b + s(\ell) \geq (f - i - 1)b + b = (f - i)b$.

The argument is only slightly different in the case that the slope of $\ell$ is negative: Suppose that $\alpha(\ell)$ intersects the right side of $D$ at a height between $(4i+5)\gamma - 2\varepsilon$ and $(4i+9)\gamma - 2\varepsilon$ (i.e., the heights of the heads of $b_i$ and $b_{i+1}$). If $i \geq f - 1$, then this implies $w(\ell)b + s(\ell) \geq b$. If $i < f - 1$, then we follow the curves $\alpha(\ell)$ and $\beta^-(\ell)$. By the choice of $\varepsilon$ and Lemma 3.4 (A), $\alpha(\ell) = \ell$ (this does not necessarily hold for $i = f$ nor for $i = f - 1$). Hence the curve $\beta^-(\ell)$ is $y$-monotone ascending, too. Possibly $\alpha(\ell)$ intersect an $\varepsilon$-disk around the head of $a_{i+1}$, but the curve $\beta^-(\ell)$ must cut all of $a_{i+2}, a_{i+3}, \ldots, a_f$ in $D^-$. Therefore, $w(\ell)b + s(\ell) \geq (f - i - 1)b + b = (f - i)b$.

This complete the proof if $i = 0$. Suppose that $i > 1$. Observe that $\alpha(\ell)$ and $\beta^{\pm}(\ell)$, whose cuts on segments and arrows were counted so far, does not intersect the hexagon $R(i)$. By Remark 3.2, the hexagon $R(i)$ contains $A_i(m, f, \delta) = A(m, i, \delta)$. Moreover, $R(i)$ contains the $\varepsilon/3$-neighborhoods of every arrow in $A_i(m, f, \delta)$ except for the $\varepsilon$-disk around the head of $b_i$. Applying induction to $A(m, i, \delta)$, we obtain $w(P_0)b + s(P_0) \geq (f-i)b + ib = fb$. □

## 3.3 Binary space partition with few distinct directions

In order to prove Theorem 3.2, we introduce some notation. In particular, we distinguish several classes of segments within a convex region. Fix a set $L$ of disjoint line segments in the plane. A *cell* is defined as a bounded closed convex region in the plane. We avoid considering unbounded convex regions by associating the root of every BSP tree to the convex hull $\mathrm{conv}(\bigcup L)$ of segments (instead of $\mathcal{E}^2$). Observe that every node $v$ of a BSP corresponds to a cell $R_v$.

For a cell $C$, let $L(C) = \{\ell \cap \mathrm{int}(C) : \ell \in L\} \setminus \{\emptyset\}$. We use a simple observation throughout this section: If there is a segment $\ell \in L(R_v)$ for some interior node $v$ of the BSP such that $\ell$ connects two points on the boundary of the cell $R_v$, then we can split $R_v$ along $\ell$ without creating any new *cuts*. We apply such *free cuts* whenever possible. If a segment $\ell \in L$ is dissected into many pieces by consecutive cuts, we only have to care about two pieces, the two *endings* of $\ell$, which are two disjoint line segments in two non-overlapping regions.

We can assume that no segment of $L(C)$ has two endpoints on the boundary

of $C$, since otherwise a free cut could split $C$ into two cells. We partition $L(C)$ into two disjoint subsets: Let $B(C)$ be the set of segments from $L(C)$ which have exactly one endpoint on the boundary of $C$ (and the other endpoint in $\mathrm{int}(C)$) and let $I(C)$ be the segments from $L(C)$ which lie entirely in the interior of $C$.

Let $c$ be a line disjoint from $C$ and which is not parallel to any segment of $L$. Let $B_c(C) \subset B(C)$ be the set of all segments $\ell \in B(C)$ such that the extension of $\ell$ beyond the boundary $\partial C$ intersects $c$.

The following lemma is the key to the proof of Theorem 3.2 $(i)$ and $(ii)$. The proof of this lemma is postponed to the last two subsections.

**Lemma 3.5.** *Assume that we are given a set $L$ of disjoint line segments in the plane with at most $k$ distinct directions, a cell $C$, and a line $c$ such that $c \cap \mathrm{int}(C) = \emptyset$.*

- $(i)$ *There exists a BSP $P_C$ for $B_c(C)$ where every line segment of $B(C)$ is cut at most $7\lfloor \log_2 k \rfloor + 6$ times and every line segment of $I(C)$ is cut at most $14\lfloor \log_2 k \rfloor + 7$ times.*

- $(ii)$ *There exists a BSP $P_C$ for $B_c(C)$ where every line segment of $B(C)$ is cut at most $2k + 4$ times and every line segment of $I(C)$ is cut at most $2k + 9$ times. Moreover, every region $R_v$ corresponding to an internal node $v \in P_C$ is dissected along a segment of $L(R_v)$.*

### 3.3.1 Partition algorithm

We show here how Lemma 3.5 establishes our theorem. A BSP $P_C$ is constructed by the following recursive algorithm:

**Algorithm 3.1.** *Let $a$ and $b$ be two parallel lines in the plane such that the strip between $a$ and $b$ entirely contains all segments of $L$, and they are not parallel to any segment of $L$. Initialize $i := 1$ and $\mathcal{R}_1 := \{\mathrm{conv}(\bigcup L)\}$. As long as there is a $C \in \mathcal{R}_i$ such that $L(C)$ is non-empty, do the following.*

*Let $c = a$, when $i$ is odd, and let $c = b$ when $i$ is even. Consider one-by-one each cell $C \in \mathcal{R}_i$. If $B_c(C)$ is non-empty then, by Lemma 3.5, there is a BSP $P_C$ for $B_c(C)$ which dissects $C$ into some cells. If $B_c(C)$ is empty but $L(C)$ is non-empty, then dissect $C$ by an arbitrary $\ell \in L(C)$ into two cells. Let $\mathcal{R}^0_{i+1}$ be the set of all sub-cells for all $C \in \mathcal{R}_i$. Do all possible* free cuts *in each cell of $\mathcal{R}^0_{i+1}$ and let $\mathcal{R}_{i+1}$ be the resulting set of cells. Put $i := i + 1$.*

*Proof (of Theorem 3.2).* Algorithm 3.1 produces a BSP of $L$. If every cell is dissected along a segment in the cell, then it creates an autopartition. From the point of view of one line segment $\ell \in L$, the algorithm looks as follows. $\ell$ is cut into pieces at some step $i = q_\ell$. (There is the possibility of cutting along $\ell$ before it is cut otherwise. In this case $\ell$ is not cut at all by the procedure.) In this same step, $\ell$ may be cut several times, and its two endings are in two distinct cells $D_1, D_2 \in \mathcal{R}_{q_\ell+1}$. In step $i = q_\ell + 1$, the two endings are in sets $B_a(D_1)$ and $B_b(D_2)$. Assume that $q_\ell$ is odd. In step $q_\ell + 1$, BSPs for $B_b(D_1)$ and $B_b(D_2)$ are considered. In this step, both endings may also be cut several times. $D_1$ may be subdivided and the ending of the portion of $\ell$ within $B_a(D_1)$ belongs to $B_a(E_1)$ for a cell $E_1 \in \mathcal{R}_{q_\ell+2}$. In step $q_\ell + 2$, a BSP for $B_a(E_1)$ is considered but $\ell$ may still be cut several times. Analogously, if $q_\ell$ is even, both endings may be cut at step $q_\ell + 1$, but only one of them can be cut in step $q_\ell + 2$. Thus, $\ell$ may be cut in at most three consecutive steps of Algorithm 3.1.

For part (i) of Theorem 3.2, we determine the maximal number of cuts during the "life cycle" of a segment $\ell$ from Lemma 3.5 (i). In step $q_\ell$, segment $\ell$ is cut at most $14\lfloor \log_2 k \rfloor + 7$ times. In step $q_\ell + 1$, its either ending is cut at most $7\lfloor \log_2 k \rfloor + 6$ times. Finally in step $q_\ell + 2$, its surviving one ending might be cut at most $7\lfloor \log_2 k \rfloor + 6$ times. All in all, $\ell$ is cut at most $35\lfloor \log_2 k \rfloor + 25$ times during the algorithm.

For the proof of Theorem 3.2 (ii), we use Lemma 3.5 (ii). In step $q_\ell$, segment $\ell$ is cut at most $2k + 9$ times. In step $q_\ell + 1$, its either ending is cut at most $2k + 4$ times. Finally in step $q_\ell + 2$, its surviving one ending can be cut at most $2k + 4$ times. All in all, $\ell$ is cut at most $(2k + 9) + 3(2k + 4) = 8k + 21$ times during the algorithm. □

**Remark 3.5.** *The constants in the upper bound on the total number of cuts per segments can be slightly improved by a more careful analysis, namely by applying the detailed versions of Lemma 3.5 (i) and (ii) as they appear in Subsections 3.3.3 and 3.3.4.*

### 3.3.2 Cycles and convex cycles

Given $C$ and $L$ as above, fix some subset $X$ of $B(C)$ for this subsection. Recall that for every segment $\ell \in X$, one endpoint of $\ell$ is in $\mathrm{int}(C)$. Denote this common point by $p(\ell) = \ell \cap \partial C$. Extend $\ell$ to the interior of $C$ until it hits another segment of $X$ or the boundary of $C$ at some point $g(\ell)$. The extension of $\ell$ may cross several segments of $L(C) \setminus X$ before hitting some other segment $\ell' \in X$ or the boundary $\partial C$.

**Definition 3.3.** *A cycle of $X$ is a $t$-tuple $(\ell_1, \ell_2, \ldots, \ell_t)$, $t \in \mathbf{N}$, such that*

- $\ell_1, \ell_2, \ldots, \ell_t \in X$,
- $g(\ell_i) \in \ell_{i+1}$ *for* $i = 1, 2, \ldots t-1$, *and*
- $g(\ell_t) \in \ell_1$.

Figure 3.7 exhibits three cycles. Segments of cycles are depicted by bold lines. Segments of $L(C) \setminus X$ are not indicated. Note that a segment may appear in at most one cycle. The concept of *convex cycle*, defined below, is not a subclass of cycles, it requires a little less and a little more.

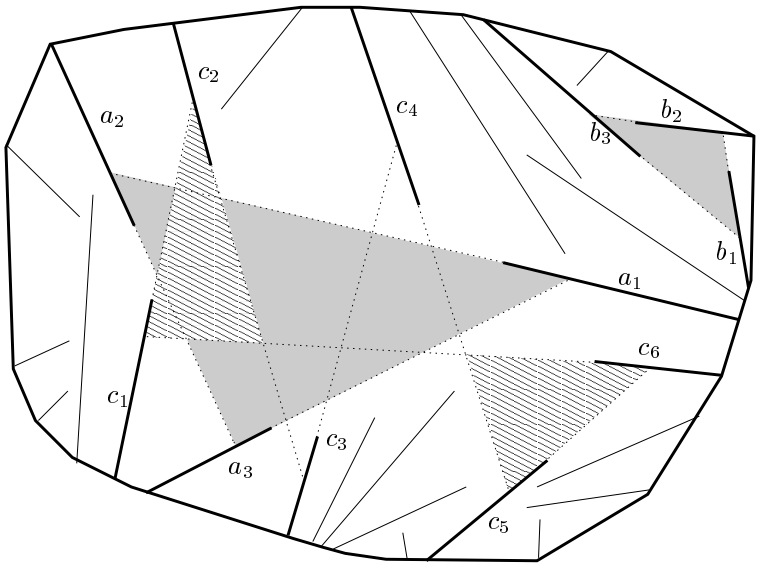


**Figure 3.7:** *Three cycles, and four convex cycles.*

**Definition 3.4.** *A convex cycle of $X$ is a $t$-tuple $(\ell_1, \ell_2, \ldots, \ell_t)$, $t \in \mathbf{N}$, such that*

- $\ell_1, \ell_2, \ldots, \ell_t \in X$,
- $g(\ell_i) \in \ell_{i+1}$ *for* $i = 1, 2, \ldots t-1$,
- *the extension of $\ell_t$ hits the line through $\ell_1$ at a point $a \in C$ but does not cross any segment of $X$ until it reaches $a$,*
- $g(\ell_1)g(\ell_2)\ldots g(\ell_{t-1})a$ *forms a convex $t$-gon.*

In Figure 3.7, two cycles are *convex cycles*. The third cycle is not convex but contains two convex cycles, which are indicated by their interior convex regions striped.

**Proposition 3.1.** *For any cycle $S = (\ell_1, \ell_2, \ldots, \ell_t)$, there is $q \in [1, t-2]$ and $r \in [q+2, t]$ such that $(\ell_q, \ell_{q+1}, \ldots, \ell_r)$ is a convex cycle.*

*Proof.* Consider the polygonal curve $\gamma = (p(\ell_1), g(\ell_1), g(\ell_2), \ldots, g(\ell_t))$. If $\gamma$ has no self-crossing, then $S$ forms a convex cycle.

Let $(p(\ell_1), g(\ell_1), g(\ell_2), \ldots, g(\ell_{r-1}), w) \subset \gamma$ be the longest prefix curve without self-crossing. That is, $w$ denotes the common point of two segments, say, $g(\ell_q)g(\ell_{q+1})$ and $g(\ell_{r-1})g(\ell_r)$, $q+1 < r-1$. We show next that $(\ell_q, \ell_{q+1}, \ldots, \ell_r)$ is a convex cycle. $g(\ell_i) \in \ell_{i+1}$ for $i = q, q+1, \ldots r-1$; the extension of $\ell_r$ hits the line through $\ell_q$ at point $w$; but it does not hit any segment of $X$ before. Extend the every segment $\ell_i$, for $i = 1, 2, \ldots, r-1$, to $G(\ell_i)$, and $\ell_r$ to $w$. We obtain a convex partition of $C$ where one face of the partition is $(g(\ell_q), g(\ell_{q+1}), \ldots, g(\ell_{r-1}), w)$, therefore it is a convex. □

We observe here a useful property of set $B_c(C)$ with respect to cycles.

**Proposition 3.2.** *$B_c(C)$ contains no cycle.*

*Proof.* Proposition 3.1 implies that it is enough to show that $B_c(C)$ contains no convex cycle. Suppose, to the contrary, that $S$ is a convex cycle in $B_c(C)$. Let $w$ be the furthest away vertex of the polygon $g(\ell_1)g(\ell_2)\ldots g(\ell_{t-1})a$ from the line $c$. Point $w$ is the intersection of two lines through, say, $\ell_i$ and $\ell_{i+1}$. The extension of $\ell_i$ or that of $\ell_{i+1}$ over the boundary $\partial C$ does not intersect $c$, a contradiction. □

Consider a convex cycle $S = (\ell_1, \ell_2, \ell_3, \ldots, \ell_t)$ of $X$. Extend every $\ell_i$, $i = 1, 2, \ldots, t-1$ to $g(\ell_i)$, and $\ell_t$ to $a$. The extended segments of $S$ partition the cell $C$ into $t+1$ convex cells $S^*, S_1, \ldots, S_t$. One of them, $S^* = g(\ell_1)g(\ell_2)\ldots g(\ell_{t-1})a$, lies completely in the interior of $C$. Clearly, $\mathrm{int}(S^*)$ is disjoint from all segments of $X$. Every $S_i$, $i = 1, \ldots t$, has a common boundary with $C$. We may choose the notation of $S_i$, $i = 1, 2, \ldots, t$, such that $S_i$ is bounded by portions of the extension of segments $\ell_i$ and $\ell_{i+1}$, and by $\partial C$.[1]

For the segments of a convex cycle $S = (\ell_1, \ell_2, \ell_3, \ldots, \ell_t)$, a BSP $P_S^0(i)$, $i = 1, \ldots, t$, can be obtained in the following way: We proceed in $t$ steps; in

[1] Here and in what follows, arithmetics on the indices of $\ell_1, \ell_2, \ldots, \ell_t$ and $S_1, S_2, \ldots, S_t$ is meant mod $t$, i.e., $t+1$ stands for 1.

step $j$, $j = 0, 1, 2 \ldots, t-1$, we dissect all sub-regions $R_v \subset C$ by the line through $\ell_{i-j}$ if $\ell_{i-j} \cap \operatorname{int}(R_v) \neq \emptyset$. Notice that $P_S^0(i)$ is an autopartition for $S$. We define another BSP $P_S^1(i)$, $i \in \{3, \ldots, t\}$, for $S$ if the line through $\ell_i$ crosses the extension of $\ell_1$ to $g(\ell_1)$ but does not intersect $\ell_1$: We proceed in $i$ steps; in step $j$, $j = 0, 1, 2 \ldots, i-1$, we dissect the sub-region $R_v \subset C$ containing $\ell_{i-j}$ by the line through $\ell_{i-j}$. Note that in $P_S^1(i)$, we do not cut along $\ell_{i+1}, \ell_{i+2}, \ldots, \ell_t$. These segments all intersect the line through $\ell_i$, because $p(\ell_{i+1}), p(\ell_{i+2}), \ldots, p(\ell_t)$, and $p(\ell_1)$ are all on the same side of $\ell_i$, and therefore they give raise to free cuts. Let the BSP $P_S(i) = P_S^1(i)$ if the line through $\ell_i$ intersects the extension of $\ell_1$ to $g(\ell_1)$ but does not intersect $\ell_1$, and let $P_S(i) = P_S^0(i)$ otherwise.

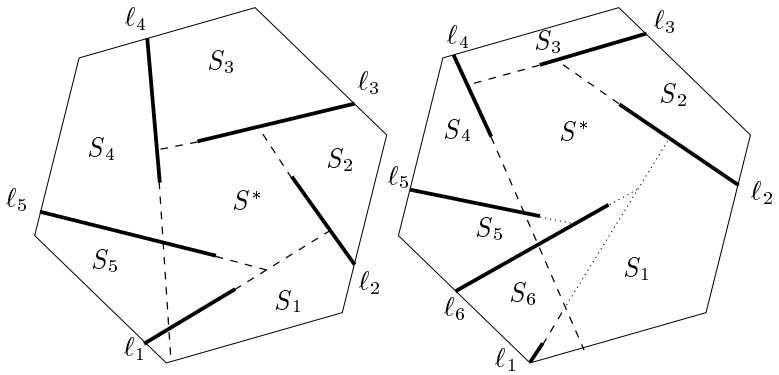


**Figure 3.8:** $P_S(4) = P_S^0(4)$ *on the left, and* $P_S(4) = P_S^1(4)$ *on the right.*

Notice that the cuts along $\ell_q$, $q \neq i$, in $P_S(i)$ can only cut segments which lie on the extension of the segment $\ell_q$ to $g(\ell_q)$. The first cut along $\ell_i$ may cut any segment, and it may also dissect some of the regions $S_{i+1}, S_{i+2}, \ldots, S_{i-2}$ into two pieces, but it does not cut regions $S^*$, $S_{i-1}$, and $S_i$. If $P_S(i) = P_S^0(i)$, then the cuts along the subsequent segments of the cycle always cut along all boundaries of the regions $S^*, S_1, \ldots, S_t$; and the partition contains these cells, some of them split into two by the line through $\ell_i$. If $P_S(i) = P_S^1(i)$, then we do not cut along $\ell_j$, $j = i+1, i+2, \ldots, t$, therefore the parts of $S_j$ (split by the line through $\ell_i$) which do not have a common boundary with $C$ form a common cell with $S^*$ and part of $S_1$. Denote by $\mathcal{P}(S, i)$ the set of cells in the partition of $C$ obtained by $P_S(i)$. In particular, $S_{i-1}$ and $S_i$ are always cells of $\mathcal{P}(S, i)$. The BSP $P_S(i)$ will be a basic building block of our partition algorithms in the next subsections.

**Proposition 3.3.** *In $P_S(i)$, every segment of $X$ is cut at most once, and every segment of $L(C) \setminus X$ is cut at most four times.*

*Proof.* A segment of $X$ can only be cut by the line through $\ell_i$. A segment $\ell \in L(C) \setminus X$ may be cut at most four times, because we cut along sides of a convex polygon and at most two other lines (namely the lines through $\ell_1$ and $\ell_i$). □

If $X$ has no cycle, it could still contain the following simple structure.

**Definition 3.5.** *A series of $X$ is a $t$-tuple $(\ell_1, \ell_2, \ldots, \ell_t)$, $t \in \mathbf{N}$, such that*

- $\ell_1, \ell_2, \ldots, \ell_t \in X$,
- $g(\ell_i) \in \ell_{i+1}$ *for* $i = 1, 2, \ldots t-1$, *and*
- $g(\ell_t)$ *is on the boundary of* $C$.

*A series $(\ell_1, \ell_2, \ldots, \ell_t)$ is* convex, *if the point $p(\ell_1)$ and the points $g(\ell_i)$, $i = 1, 2, \ldots t$, are together in convex position.*

For the segments of a series $(\ell_1, \ell_2, \ell_3, \ldots, \ell_t)$, we obtain a BSP, by consecutively cutting along the segments of the series in reverse order (starting from $\ell_t$ and proceeding to $\ell_1$). In this BSP, segments of $X$ are not cut at all. If the series if convex, then a segment $\ell \in L(C) \setminus X$ may be cut at most twice because we cut along sides of a convex polygon.

### 3.3.3 BSP for boundary segments

For any $F \subset B_c(C)$, we define $F^*$ as follows. Let $q_c$ be the furthest vertex of $C$ from the line $c$. Now $\partial_F C \subset \partial C$ is defined to be the shortest polygonal path along the boundary $\partial C$ of $C$ containing all endpoints of $F$ on $\partial C$, but not containing $q_c$. Let $F^*$ be the set of segments in $B(C) \setminus F$ which have one endpoint in $\partial_F C$ (and one endpoint in $\operatorname{int}(C)$). (See Figure 3.9 for an illustration.)

**Proposition 3.4.** *If $X = B(C) \setminus F^*$ contains no cycle, then either there is a segment $\ell \in F$ such that $g(\ell) \notin \partial_F C$, or there is a series $(\ell_1, \ell_2, \ldots, \ell_t)$, $t \in \mathbb{N}$, of $X$ such that $\ell_1 \in B(C) \setminus (F \cup F^*)$ and $\ell_2, \ell_3, \ldots, \ell_t \in F$.*

*Proof.* Let $p(\ell_1)$ and $p(\ell_1')$ be the two extremal points of the curve $\partial_F C$ (with possibly, $p(\ell_1) = p(\ell_1')$ if $|F| = 1$). Suppose w.l.o.g. that $\operatorname{dist}(p(\ell_1), c) \geq \operatorname{dist}(p(\ell_1'), c)$. Consider the series $(\ell_1, \ell_2, \ldots, \ell_t)$ starting with $\ell_1$. Notice that if $\ell_i \in F$ for $i = 1, 2, \ldots j$, then $\operatorname{dist}(p(\ell_1), c) < \operatorname{dist}(g(\ell_1), c) < \operatorname{dist}(g(\ell_2), c) < \ldots < \operatorname{dist}(g(\ell_j), c)$. Therefore, if $g(\ell)$ is not in $\partial C \setminus \partial_F C$

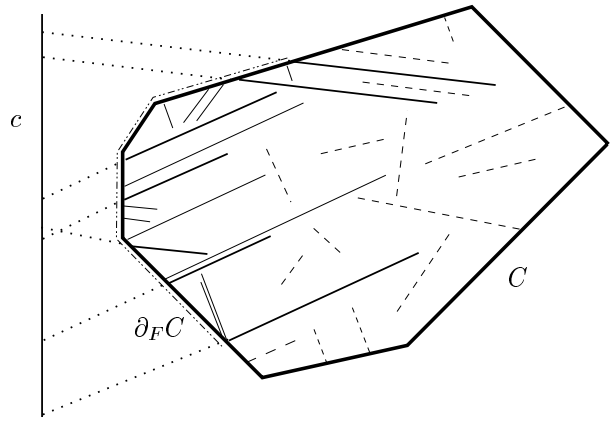


**Figure 3.9:** *Elements of $F$ are bold segments, elements of $F^* \setminus F$ are thin segments, and elements of $L(C) \setminus F^*$ are dashed.*

for any $\ell \in F$, then the series starting with $\ell_1$ must contain a segment of $B(C) \setminus (F \cup F^*)$. The post-fix $(\ell_{q+1}, \ell_{q+2}, \ldots, \ell_t)$ of this series is a series such that $\ell_q \in B(C) \setminus (F \cup F^*)$ and $\ell_{q+2}, \ldots, \ell_t \in F$. □

**Proposition 3.5.** *If $X = B(C) \setminus F^*$ contains no cycle, then every series $(\ell_1, \ell_2, \ldots, \ell_t)$, of $X$ such that $\ell_1 \in B(C) \setminus (F \cup F^*)$ and $\ell_2, \ell_3, \ldots, \ell_t \in F$ is convex.*

*Proof.* For every segment $\ell \in B(C)$, we distinguish a left and a right side such that the closed curve on the boundary of $C$ oriented in counter-clockwise order traverses $\ell$ from left to right. We may suppose w.l.o.g. that the extension of $\ell_1$ hits $\ell_2$ on the left side. If the extension of every $\ell_i$, $i = 2, 3, \ldots, t-1$, hits $\ell_{i+1}$ on the left side then the series is convex. Suppose, to the contrary, that $j \in \{2, 3, \ldots, t-1\}$ is the first index such that $\ell_j$ hits $\ell_{j+1}$ on the right side. This leads to a contradiction, since both $\ell_j$ and $\ell_{j+1}$ are in $F$: The polygonal curve $p(\ell_j)g(\ell_j)p(\ell_{j+1})$ dissects $C$ into two parts $C_1$ and $C_2$. One part, say $C_1$, contains $\ell_1 \in B(C) \setminus (F \cup F^*)$ and the same part have a convex angle at $g(\ell_j)$. Therefore the common boundary of $C$ and $C_1$ belongs to $\partial_F C$, and $\ell_1$ must be in $F$. □

We demonstrate Lemma 3.5 $(i)$ in the following refined formulation, which is necessary for the induction step in our proof. We use induction on the number of segments and on the number of consecutive directions, defined below.

**Definition 3.6.** *We say that the directions of a subset of segments $F \subset B_c(C)$ are* consecutive *if the directions are consecutive in the linear order (by slope) of all $k$ directions starting from, but not including, the direction of $c$ (which is non-parallel to any segment of $L$).*

**Lemma 3.6.** *Let $L$, $C$, and $c$ be as in Lemma 3.5. Consider a set $F \subset B_c(C)$, $F \neq \emptyset$, such that the directions in $F$ belong to $f$, $f \leq k$, consecutive directions. There is a BSP $P$ for $F$ such that*

| | | | |
|---|---|---|---|
| *1. every $\ell \in$* | $F$ | *is cut at most* | $2\lfloor \log_2 f \rfloor$ *times,* |
| *2. every $\ell \in$* | $B(C) \setminus (F \cup F^*)$ | *is cut at most* | $7\lfloor \log_2 f \rfloor + 1$ *times,* |
| *3. every $\ell \in$* | $F^*$ | *is cut at most* | $7\lfloor \log_2 f \rfloor + 6$ *times,* |
| *4. every $\ell \in$* | $I(C)$ | *is cut at most* | $14\lfloor \log_2 f \rfloor + 7$ *times.* |

We prove Lemma 3.6 by induction on $f$ and $|L(C)|$. We distinguish five cases. In each case, we describe a partitioning of $C$ and apply our lemma by induction to the portions of the segments in $F$ within each sub-cell $D$ of $C$. In each case $|L(D)| < |L(C)|$ will hold for every sub-cell of $C$. We apply induction to the number of consecutive directions only in our case 5. Throughout the following proof, we use the shorthand notation $F \cap D = \{\ell \cap \operatorname{int}(D) : \ell \in F\} \setminus \{\emptyset\}$ for the portions of the segments of $F$ in a sub-cell $D$.

*Proof (of Lemma 3.6).* Our lemma clearly holds for $|L(C)| = 0$. Note that $f = 0$ means that $F$ is empty and nothing needs to be dissected. Suppose that it holds for all $f' < f$ and for all $|L(C)'| < |L(C)|$. Put $X = B(C) \setminus F^*$. We distinguish five cases. The first four cases are similar in the sense that we use induction with the same parameter $f$, the fifth case uses induction with parameter $f$ and some $f' < f$, too.

In the first four cases, we dissect $C$ either along a series of $X$, or along a convex cycle of $X$, as described below. This guarantees that segments of $X = B(C) \setminus F^*$ are cut at most once and segments of $I(C) \cup F^*$ are cut at most four times. The resulting partition of $C$ into a set $\mathcal{P}$ of cells has, as we show below for each case, the following three properties:

- $(\alpha)$ No segment $\ell \in F$ is cut.
- $(\beta)$ If a segment $\ell \in B(C) \setminus (F \cup F^*)$ is cut and its ending is in $D \in \mathcal{P}$, then $D \cap F$ is empty.
- $(\gamma)$ If a segment $\ell \in I(C) \cup F^*$ is cut, then its ending(s) are in sets $B(D) \setminus (F \cap D)^*$ in the corresponding cell(s) $D \in \mathcal{P}$.

In the last property and hereafter, $(F \cap D)^*$ is defined for $(F \cap D) \subset B_c(D)$ with respect to the same line $c$ in each sub-cell $D \in \mathcal{P}$ . In all four cases, we can apply our Lemma by induction to $F \cap D$ in each cell $D \in \mathcal{P}$ where it is non-empty. Then induction completes the proof. Indeed: If a segment of $X$ is cut then, by $(\alpha)$ and $(\beta)$, its ending is cut no more in the induction step. If a segment of $F^* \cup I(C)$ is cut, then its ending(s) belong(s) to $B(D) \setminus (F \cap D)^*$ in the subproblem(s), and by $(\gamma)$ each ending is cut at most once at the induction step. It remains to describe the case distinction and establish properties $(\alpha)$ and $(\beta)$ in cases 1, 2, 3, and 4. Case 5 will be treated separately.

CASE 1: *There is no cycle in* $X$. If thee is a segment $\ell \in F$ whose extension does not hit any segment in $X$ and for which $g(\ell)$ is not in $\partial_F C$, then dissect $C$ along the line through $\ell$ into $C_1$ and $C_2$. Only segments of $L(C) \setminus X = F^* \cup I(C)$ may be cut. Otherwise there is a series $(\ell_1, \ell_2, \ldots, \ell_t)$ in $X$ where $\ell_1 \in B(C) \setminus (F \cup F^*)$ and $\ell_2, \ell_3, \ldots, \ell_t \in F$ by Proposition 3.4. According to Proposition 3.5, this is a convex series. Dissect $C$ along the segments of this series in reverse order.

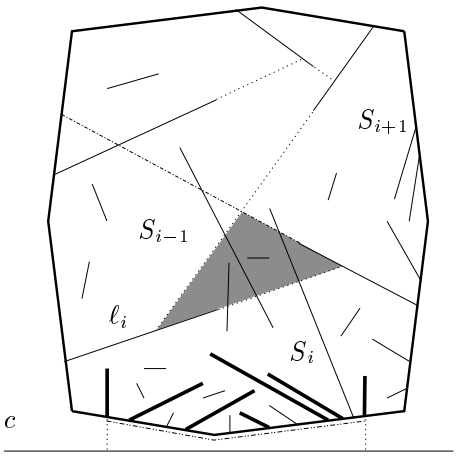


**Figure 3.10:** *Case 2 for Lemma 3.7. Segments of* $F$ *are bold.*

In the remaining four cases we assume that there is a cycle in $X$. By Proposition 3.1, there exists a convex cycle $S = (\ell_1, \ell_2, \ldots, \ell_t)$, which partitions $C$ into $S^*, S_1, \ldots, S_t$.

CASE 2: *$F$ and $S$ have no common element.* In this case $\partial_F C$ is on the boundary of one region, say, $S_i$. (See Figure 3.10.) Apply BSP $P_S(i)$, which cuts $C$ first along $\ell_i$. Every segment of $F$ is in cell $S_i$ of $\mathcal{P}(S,i)$, but the line through $\ell_i$ does not cut $S_i$. Hence, no segment of $F$ is cut. The ending of any $\ell' \in B(C) \setminus F^*$ cut by the line through $\ell_i$ is not in $S_i$, because $S$ is a convex cycle of $X$.

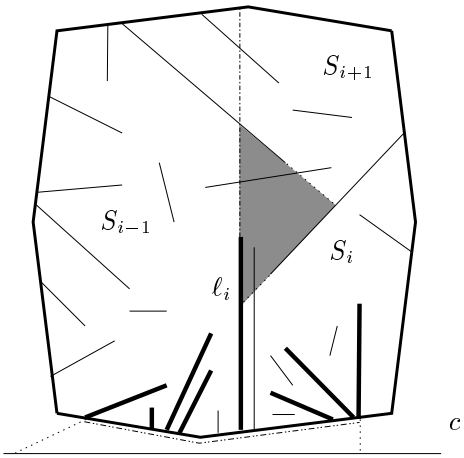


**Figure 3.11:** *Case 3 for Lemma 3.7. Segments of $F$ are marked bold.*

CASE 3: *$F$ and $S$ have exactly one common element.* Denote the common element of $F$ and $S$ by $\ell_i$. (See Figure 3.11.) Apply $P_S(i)$, which cuts $C$ first along $\ell_i$. Every segment of $F \setminus S$ is in $S_{i-1}$ or in $S_i$. The line through $\ell_i$ does not cut neither $S_{i-1}$ nor $S_i$. Similarly, the ending of any $\ell' \in B(C) \setminus F^*$ cut by the line through $\ell_i$ is not in $S_{i-1}$ nor in $S_i$.

From now on, we suppose that there are at least two segments of $F$ in the convex cycle $S$. Let us assume that the common segments are $\ell_i, \ell_{i+1}, \ldots,$ and $\ell_j$ appearing in this (cyclic) order in $S$.

CASE 4: *The supporting line of $\ell_j$ does not intersect $p(\ell_{i-1})g(\ell_{i-1})$.* (See Figure 3.12.) Apply $P_S(j)$, which cuts $C$ first along the line through $\ell_j$. The supporting line of $\ell_j$ may cut only the regions $S_{j+1}, S_{j+2}, \ldots, S_{i-2}$, Segments of $F$ are all in regions $S_{i-1}, S_i, \ldots, S_j$, which are cells of $\mathcal{P}(S,j)$.

CASE 5: *The supporting line of $\ell_j$ intersects $p(\ell_{i-1})g(\ell_{i-1})$.* In this case, the

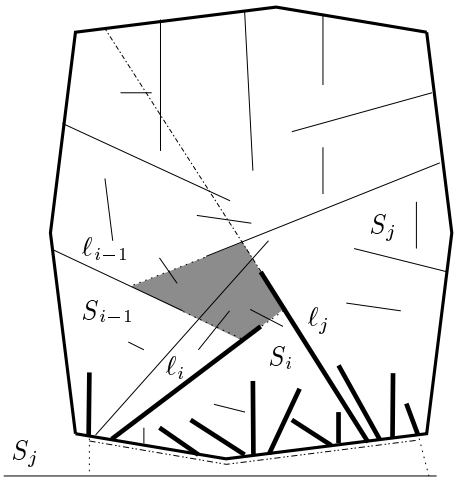


**Figure 3.12:** *Case 4 for Lemma 3.7. Segments of $F$ are marked bold.*

direction of $\ell_{i-1}$ belongs to the consecutive directions of $F$ between (but not including) $\ell_i$ and $\ell_j$. In this case, we cannot rely on properties $(\alpha)$-$(\gamma)$. We define and apply a slightly modified version of the family of BSPs $P_S(.)$ using the convex cycle $S$ and then we apply induction in the sub-cells.

We can define two new BSPs $Q^i_S(F)$ and $Q^j_S(F)$ for $S = (\ell_1, \ell_2, \ell_3, \ldots, \ell_t)$: $Q^i_S(F)$ cuts $C$ first along the line $p(\ell_i)p(\ell_{i-1})$, then it calls the BSP $P_S(j)$ as defined above. $Q^j_S(F)$ cuts $C$ first along the line $p(\ell_j)p(\ell_{i-1})$, then it calls the BSP $P_S(i)$.

The BSP $Q^i_S(F)$ (and $Q^j_S(F)$) has some useful properties (similarly to those of $P_S(.)$): Any segment of $I(C) \cup F^*$ is cut at most five times: at most once by the lines of $p(\ell_i)p(\ell_{i-1})$ (or $p(\ell_j)p(\ell_{i-1})$), $\ell_j$ (or $\ell_i$), and $\ell_1$, and at most twice along the boundary of the convex region $S^*$. Any segment of $X = B(C) \setminus F^*$ is cut at most twice: at most once by the line $p(\ell_i)p(\ell_{i-1})$ (or $p(\ell_j)p(\ell_{i-1})$) and at most once by the line through $\ell_j$ (or $\ell_i$).

The line $p(\ell_i)p(\ell_{i-1})$ cuts the region $S_{i-1}$ and the line $p(\ell_j)p(\ell_{i-1})$ cuts the regions $S_j, S_{j+1}, \ldots, S_{i-2}$. The two lines lie along or in sub-cells of the partition whose boundary have at most one common point with $\partial_F C$ (namely $p(\ell_i)$ or $p(\ell_j)$. Either line may cut segments of $X$ such that the properties $(\alpha)$-$(\gamma)$ do not hold for $Q^i_S(F)$ nor for $Q^j_S(F)$. Observe that the sets of directions of segments of $F$ which cross the segments $p(\ell_i)p(\ell_{i-1})$ and $p(\ell_j)p(\ell_{i-1})$,

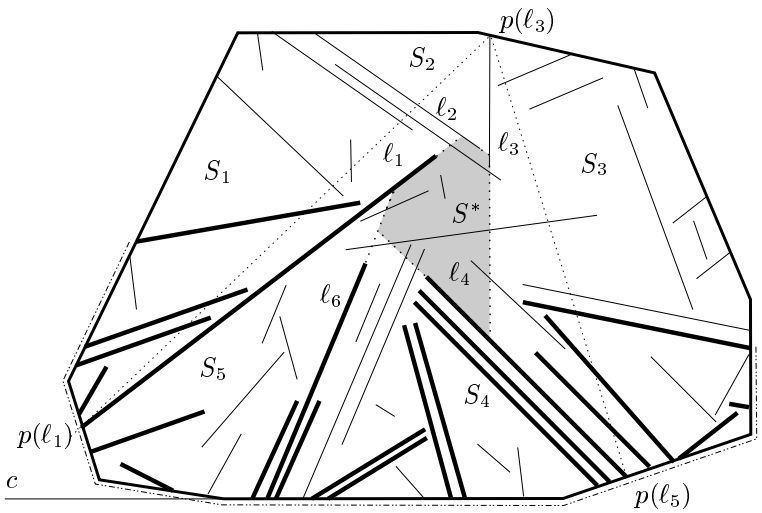


**Figure 3.13:** *Example for case 5 with a convex cycle highlighted and with* $i = 4$ *and* $j = 1$.

respectively, are disjoint. Moreover, segments of $F$ with direction $\ell_{i-1}$ can cross neither $p(\ell_i)p(\ell_{i-1})$ nor $p(\ell_j)p(\ell_{i-1})$.

If the segments of $F$ crossed by $p(\ell_i)p(\ell_{i-1})$ belong to at most $\lfloor f/2 \rfloor$ consecutive directions, then apply the BSP $Q_S^i(F)$, otherwise (i.e., if the segments of $F$ crossed by $p(\ell_j)p(\ell_{i-1})$ belong to at most $\lfloor f/2 \rfloor$ consecutive directions) apply $Q_S^j(F)$. We show that induction for $F \cap D$ in every $D$ resulting sub-cell establishes Lemma 3.6.

Suppose w.l.o.g. that we apply $Q_S^i(F)$ and it partitions $C$ into a set $\mathcal{Q}$ of sub-cells. Instead of properties $(\alpha)$-$(\gamma)$, we have the following two properties:

$(\alpha, \beta)'$ If a segment of $X$ is cut (at most twice) by $Q_S^i(F)$, then it is either part of $S$, or its ending is in sub-cells $D$ in which $F \cap D$ belongs to at most $\lfloor f/2 \rfloor$ consecutive directions.

$(\gamma)'$ A segment of $F^* \cup I(C)$ can be cut by $Q_S^i(F)$ in two possible ways:
Either it is cut at most five times and its each ending is in some $B(D) \setminus (F \cap D)^*$, $D \in \mathcal{Q}$;
or it is cut once and one ending is in $(F \cap D)^*$, $D \in \mathcal{Q}$, where segments of $F \cap D$ belong to at most $\lfloor f/2 \rfloor$ consecutive directions, and a second ending (for segments of $I(C)$) is in $B(D') \setminus (F \cap D')^*$, $D' \in \mathcal{Q}$.

□

### 3.3.4 Autopartition for boundary segments

In this section we formulate and prove a stronger version of Lemma 3.5 $(ii)$ which allows to apply induction on the number of distinct directions. The proof follows the outline of the previous section.

**Lemma 3.7.** *Let $C$ and $c$ be as in Lemma 3.5. Consider a set $F \subset B_c(C)$ such that the directions in $F$ belong to $f$ consecutive directions. If $f \geq 1$, then there is a BSP $P$ for $F$ such that*

1. *every segment of* $F$ *is cut at most* $2f$ *times,*
2. *every segment of* $F^*$ *is cut at most* $2f + 4$ *times,*
3. *every segment of* $I(C)$ *is cut at most* $2f + 9$ *times,*
4. *every segment of* $B(C) \setminus (F \cup F^*)$ *is cut at most once.*
5. *In particular, if an ending of a segment $\ell$ of $F^*$ or $I(C)$ is in a common sub-cell with an ending of segment of $B(C) \setminus (F \cup F^*)$, then $\ell$ is cut at most 5 or 7 times respectively.*
6. *Every region corresponding to an internal node $v \in P$ is dissected along a segment of $L(R_v)$.*

The proof of Lemma 3.7 is based on the same case distinction as the proof of Lemma 3.6. The main difference is in the argument for case 5 where we are bound to use a partitioning along segments of $C$. For Lemma 3.6, we applied induction to sets $F \cap D$ in the sub-cells of $C$. When we applied induction to $F \cap D$ where these segments were already cut, then we assured that they belong to at most $\lfloor f/2 \rfloor$ consecutive directions. For autopartitions we cannot half the number of consecutive directions, we can only assure that if segments $F \cap D$ are all cut then the number of their consecutive directions decreases by a constant.

*Proof (of Lemma 3.7).* Our lemma clearly holds for $f = 0$ or for $|L(C)| = 0$. Suppose that it holds for all $f' < f$ or $|L(C)'| < |L(C)|$. Put $X = B(C) \setminus F^*$. We distinguish five cases just like in the proof of Lemma 3.6. In cases 1, 2, 3, and 4, we can repeat *verbatim* the argument of Lemma 3.6. We only need to argue that properties *5* and *6* of Lemma 3.7 holds. They hold automatically, by induction.

Let us recall the setting in case 5: We assume that we are given a cell $C$, a line $c$ with $c \cap \mathrm{int}(\mathrm{C}) = \emptyset$, a set of segments $L$, and a subset $F \subset B_c(C)$ of the boundary segments. We assume, furthermore, that there is a convex cycle

$S \subset X = B(C) \setminus F^*$ and there are at least two common segments of $S$ and $F$. We denote the common segments by $\ell_i, \ell_{i+1}, \ldots, \ell_j$ appearing in this (cyclic) order in $S$. The supporting line of $\ell_j$ intersects $p(\ell_{i-1})g(\ell_{i-1})$, which implies that the direction of $\ell_{i-1}$ is between the consecutive directions of $F$ between $\ell_i$ and $\ell_j$.

Apply $P_S(i-1)$, which cuts $C$ first along $\ell_{i-1}$. We denote by $\mathcal{P}(S, i-1)$ the set of the resulting sub-cells and let $s$ denote the line through $\ell_{i-1}$. Line $s$ may cut the regions $S_i, S_{i+1}, \ldots, S_{j-1}$ into two pieces (at least one of them but possibly all). No segment of $B(C) \setminus (F^* \cup S)$ is cut.

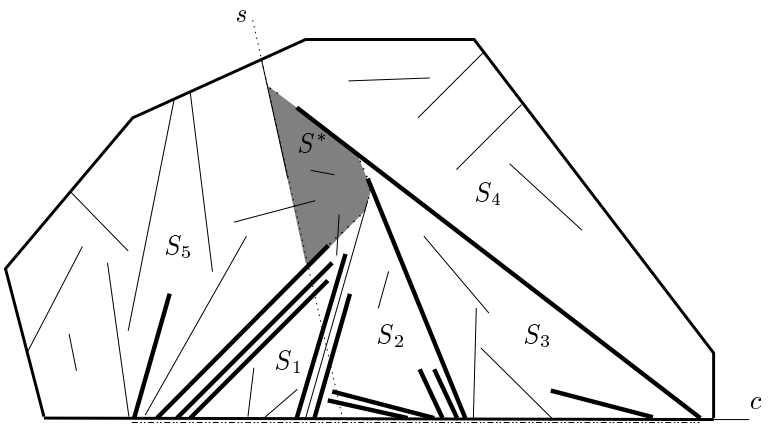


**Figure 3.14:** *Case 5. Segments of $F$ are marked as bold segments.*

In each sub-cell $D \in \mathcal{P}(S, i-1)$, we partition the endings of the segments in $F(C) \cap D$ into two sets: Let $F_s(D)$ be the subset of endings of segments cut by $s$, and let $F_c(D)$ be the subset of segments not cut by $s$. Observe that segments in $F_s(D)$ have at most $f-1$ consecutive directions: The direction of $s$ cannot occur in $F_s(D)$.

Observe, furthermore, that $F_s(D)^*$ and $F_c(D)^*$ (defined w.r.t. to line $c$, as $F_s(D), F_c(D) \subseteq B_c(D)$) are disjoint in each sub-cell $D \in \mathcal{P}(S, i-1)$. Elements of $F_c(D)^*$ all belong to $F^*$. Elements of $F_s(D)^*$ are endings of segments from $F^*$ or $I(C)$ cut by line $s$.

Next we apply, by induction, our Lemma to $F_c(D)$ in each sub-cell $D \in \mathcal{P}(S, i-1)$. We obtain a BSP $P_D$ in each $D$, which jointly partition $C$ into a collection $\mathcal{E}$ of sub-cells. Finally, we apply, by induction, our Lemma to $F_s(D) \cap E$ (or for short, $F_s \cap E$) in each sub-cell $E \in \mathcal{E}$, $D \in \mathcal{P}(S, i-1)$.

It remains to see that the cuttings implied by the three schemes, namely $P_S(i-1)$ and the two induction steps for $F_c(D)$ and $F_s(D) \cap E$, meets the bounds of Lemma 3.7. For shortness, we refer to the the three cutting schemes as $P_S(i-1)$, $P_c$ and $P_s$ (the last two can be carried out parallelly in the sub-cells). We check one-by-one the statements 1–5 of our lemma.

*1.* If a segment $\ell^1 \in F$ is not cut by $s$, then it belongs to $F_c(D)$, $D \in \mathcal{P}(S, i-1)$, and is cut by at most $2f$ times by induction. If $\ell^1 \in F$ is cut by $s$ then $P_c$ (by induction to $F_c(D)$ in the sub-cell $D \in \mathcal{P}(S, i-1)$ containing the ending of $\ell^1$) cuts it at most once more, and $P_s$ (by induction to $F_s(E)$, $E \subset D$) cuts it at most $2(f-1)$ times.

*2.* If a segment $\ell^2 \in F^*$ is cut by $P_S(i-1)$ (at most four times), then its ending is in $B(D) \setminus (F_c(D) \cup F_c(D)^*)$ for some $D \in \mathcal{P}(S, i-1)$, and then in $B(E) \setminus ((F_s \cap E) \cup (F_s \cap E)^*)$, $E \subseteq D$. Hence is cut at most once by $P_c$ and at most once by $P_s$, and altogether at most $6 \leq 2f+4$ times. If the ending of $\ell^2$ is in $F_c(D)^*$, $D \in \mathcal{P}(S, i-1)$, then it is not cut by $P_S(i-1)$ and it is cut at most $2f+4$ times by $P_c$. It can be cut once more by $P_s$, but then by Lemma 3.7 (5) it was cut at most 5 times by $P_c$; this totals to $6 \leq 2f+4$ cuts. If the ending of $\ell^2$ is in $F_s(D)^*$, $D \in \mathcal{P}(S, i-1)$, then it is cut exactly once by $P_S(i-1)$, at most once by $P_c$ and at most $2(f-1)+4$ times by $P_s$. Or else, $\ell^2$ is cut at most four times by $P_S(i-1)$, at most once by $P_c$, and at most once by $P_s$.

*3.* A segment $\ell^3 \in I(C)$ is cut at most four times by $P_S(i-1)$. Note that if $\ell^3$ is cut by $s$ then at most one of its endings may be in $F_s(D)^*$, $D \in \mathcal{P}(S, i-1)$. If one ending is actually in $F_s(D)^*$, then this ending is cut at most once by $P_c$ and at most $2(f-1)+4$ times by $P_s$; while the other ending is cut at most once by both $P_c$ and $P_s$. This implies at most $2f+9$ cuts on $\ell^3$. If neither ending is in $F_s(D)^*$, $D \in \mathcal{P}(S, i-1)$, then $P_c$ and $P_s$ cuts either endings at most once each; which makes a total of at most 8 cuts on $\ell^3$. If $\ell^3$ is not cut by $\mathcal{P}_S(i-1)$ and, therefore, it belongs to $I(D)$, $D \in \mathcal{P}(S, i-1)$, then $P_c$ cuts it at most $2f+9$ times. The subsequent BSP $P_s$ can cut its either ending at most once more, but in that case, by Lemma 3.7 (5), it was cut at most 7 times by $P_c$; thus we obtain a total of at most 9 cuts.

*4.* A segment $\ell^4 \in B(C) \setminus (F \cup F^*)$ is never cut by $s$. Moreover it cannot be in any of the sub-cells $S_m$, $m \in [i, i+1, \ldots, j-1]$, which are possibly cut by $s$. Therefore $P_c$ cuts $\ell^4$ at most once, but $P_s$ cannot cut the ending of $\ell^4$.

*5.* Observe that the BSP $P_s$ for $F_s \cap E$ was only applied in sub-cells $E \in \mathcal{E}$ where $F_s \cap E$ is non-empty, that is, only in sub-cells $\mathcal{P}(S, i-1)$ which lie along the extension of $\ell_{i-1}$ beyond $g(\ell_{i-1})$. If a segment $\ell^5 \in F^* \cup I(C)$ is cut by $P_S(i-1)$ (at most 4 times) and one of its endings belongs to a cell

$D_1 \in \mathcal{P}(S, i-1)$ which has no common boundary with (the relative interior of) $\partial_F C$, then this ending is necessarily in $B(D_1) \setminus (F_c(D_1) \cup F_c(D_1)^*)$ and is cut at most once by $P_c$. $P_s$ does not cut this ending, since $F \cap E = \emptyset$ for all such $E \subseteq D_1$. If furthermore $\ell^5 \in I(C)$ and its other ending belongs to some $D_2 \in \mathcal{P}(S, i-1)$ then it is in $B(D_2) \setminus (F_c(D_2) \cup F_c(D_2)^*)$ and then in $B(E) \setminus ((F_s \cap E) \cup (F_s \cap E)^*)$, $E \subseteq D_2$; hence is cut at most once by $P_c$ and at most once by $P_s$. The total number of cuts therefore is 5 and 7 if $\ell \in F^*$ and $\ell \in I(C)$, respectively. □

# Bibliography


[1] P. K. Agarwal, E. F. Grove, T. M. Murali, and J. S. Vitter, Binary space partitions for fat rectangles, *SIAM J. Comput.* **29** (2000), 1422–1448.

[2] M. Aigner and G. M. Ziegler, Proofs from the book, 2nd ed., Springer, Berlin, 2001, ch. 28, pp. 179–182.

[3] K. Appel and W. Haken, Every planar map is four colorable. part i. discharging, *Illinois J. Math.* **21** (1977), 429–490.

[4] K. Appel, W. Haken, and J. Koch, Every planar map is four colorable. part ii. reducibility, *Illinois J. Math.* **21** (1977), 491–567.

[5] C. Berge, Sur le couplage maximum d'un graphe, *C. R. Acad. Sciences, Paris* **247** (1958), 258–259.

[6] F. Berman, D. Johnson, T. Leighton, P. W. Shor, and L. Snyder, Generalized planar matching, *J. Algorithms* **11** (1990), 153–184.

[7] I. Bjorling-Sachs and D. L. Souvaine, An efficient algorithm for guard placement in polygons with holes, *Discrete and Comput. Geom.* **13** (1995), 77–109.

[8] P. Bose, L. Guibas, A. Lubiw, M. Overmars, D. Souvaine, and J. Urrutia, The floodlight problem, *Internat. J. Comput. Geom. Appl.* **7** (1997), 153–163.

[9] V. Brumberg, S. Ramaswami, and D. Souvaine, Experimental results on upper bounds for vertex pi-lights, in: *Proc. 11th Ann. Fall Workshop on Comput. Geom. (Brooklyn, NY, 2001)*.

[10] B. Chazelle, H. Edelsbrunner, M. Gringi, L. J. Guibas, J. Hershberger, M. Sharir, and J. Snoeyink, Ray shooting in polygons using geodesic triangulations, *Algorithmica* **12** (1994), 54–68.

[11] V. Chvátal, A combinatorial theorem in plane geometry, *J. Combin. Theory Ser. B* **18** (1975), 39–41.

[12] Gy. Csizmadia and G. Tóth, Note on an art gallery problem, *Comput. Geom. Theory Appl.* **10** (1998), 47–55.

[13] J. Czyzowicz, I. Rival, and J. Urrutia, Galleries, light matchings and visibility graphs, in: *Proc. Algorithms and Data Structures (Ottawa, ON, 1989)*, Lecture Notes in Comput. Sci., vol. 382, Springer, Berlin, 316–324.

[14] J. Czyzowicz, E. Rivera-Campo, and J. Urrutia, Illuminating rectangles and triangles on the plane, *J. Combin. Theory Ser. B* **57** (1993), 1–17.

[15] J. Czyzowicz, E. Rivera-Campo, J. Urrutia, and J. Zaks, On illuminating line segments in the plane, *Discrete Math.* **137** (1995), 147–153.

[16] F. d'Amore and P. G. Franciosa, On the optimal binary plane partition for sets of isothetic rectangles, *Inform. Proc. Lett.* **44** (5) (1992), 255–259.

[17] M. de Berg, M. de Groot, and M. Overmars, New results on binary space partitions in the plane, *Comput. Geom. Theory Appl.* **8** (1997), 317–333.

[18] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf, Computational geometry - algorithms and applications, 2nd ed., Springer, Berlin, 2000.

[19] V. Estivill-Castro, J. O'Rourke, J. Urrutia, and D. Xu, Illumination of polygons with vertex guards, *Inform. Process. Lett.* **56** (1995), 9–13.

[20] L. Fejes Tóth, Illumination of convex disks, *Acta Math. Acad. Sci. Hungar.* **29** (1977), 355–360.

[21] S. Fisk, A short proof of Chvátal's watchman theorem, *J. Combin. Theory Ser. B* **24** (1978), 374.

[22] H. Fuchs, Z. M. Kedem, and B. Naylor, On visible surface generation by a priory tree structures, in: *Proc. ACM Intern. Conf. on Comp. Graphics and Interactive Techniques (Seattle, WA, 1980)*, ACM-Press, 124–133.

[23] H. Hadwiger, Ungelöste probleme – nr. 38, *Elem. Math.* **15** (1960), 130–131.

[24] D. Hartvigsen and P. Hell, The $k$-piece packing problem, presented at the *17th Math. Prog. Symp. (Atlanta, GA)*, 2000.

[25] F. Hoffmann, On the rectilinear art gallery problem, in: *Proc. 17th Internat. Coll. Automata, Lang. and Prog. (Warwick, 1990)*, Lecture Notes in Comp. Sci., vol. 443, Springer, Berlin.

[26] F. Hoffmann, M. Kaufmann, and K. Kriegel, The art gallery theorem for polygons with holes, in: *Proc. 32nd Ann. Symp. on Foundations of Comp. Sci. (San Juan, PR, 1991)*, IEEE Comput. Soc. Press, 39–48.

[27] G. Jennings and W. Lenhart, An art gallery theorem for line segments in the plane, *Pattern Recognition Letters* **14** (1993), 727–732.

[28] A. Kaneko, A necessary and sufficient condition for the existence of a path factor every component of which is a path of length at least two, submitted, 2000.

[29] M. Kano, G.Y. Katona, and Z. Király, On path-factors and subfactors of graphs without paths of length one, submitted, 2000.

[30] D. G. Kirkcatrick and P. Hell, On the completeness of a generalized matching problem, in: *Proc. 10th ACM Symp. on Theory of Computing (San Diego, CA, 1978)*, ACM Press, 240–245.

[31] D. Larman. Lecture presented at the Oberwolfach conference on discrete geometry, 1994.

[32] H. Martini and V. Soltan, Combinatorial problems on the illumination of convex bodies, *Aequationes Mathematicae* **57** (3) (1999), 121–152.

[33] J. O'Rourke, Art gallery theorems and algorithms, Oxford University Press, New York, 1987.

[34] J. O'Rourke, Vertex $\pi$-lights for monotone mountains, in: *Proc. 9th Canadian Conf. Comput. Geom. (Kingston, ON, 1997)*, 1–5.

[35] J. O'Rourke, Open problems in the combinatorics of visibility and illumination, in: *Advances in Discrete and Computational Geometry (B. Chazelle, J. E. Goodman, and R. Pollack, eds.)*, Contemporary Mathematics, AMS, Providence, RI, 1998, pp. 237–243.

[36] J. O'Rourke and I. Streinu, The vertex-edge visibility graph of a polygon, *Comput. Geom. Theory Appl.* **10** (1998), 105–120.

[37] Zs. Páles, On a generalization of the plank problem, in: *Proc. 6th Int. Conf. on General Inequalities (Oberwolfach, 1990)*, Internat. Ser. Numer. Math., vol. 103, Birkhäuser, Basel, 473–476.

[38] M. S. Paterson and F. F. Yao, Binary space partitions for orthogonal objects, in: *Proc. 1st ACM-SIAM Symp on Discrete Algorithms (San Francisco, CA, 1990)*, 100–106.

[39] M. S. Paterson and F. F. Yao, Efficient binary space partitions for hidden-surface removal and solid modeling, *Discrete Comput. Geom.* **5** (1990), 485–503.

[40] M. Pocchiola and G. Vegter, opologically sweeping visibility complexes via pseudo-triangulations, *Discrete Comput. Geom.* **16** (1996), 419–453.

[41] M. Pocchiola and G. Vegter, Pseudo-triangulations: Theory and applications, in: *Proc. 12th ACM Symp. Comput. Geom (Philadelphia, PE, 1996)*, ACM-Press, 291–300.

[42] W. R. Pulleyblank, Matchings and extensions, in: *Handbook of combinatorics (R. Graham, M. Grötschel, and L. Lovász, eds.)*, MIT Press, Cambridge, MA, 1995, pp. 179–232.

[43] R. A. Schumacker, R. Brand, M. Gilliland, and W. Sharp, Study for applying computer-generated images to visual simulation, Tech. rep., US Air Force Human Resources Laboratory, Brooks Air Force Base, San Antonio, TX, 1969.

[44] I. Streinu, A combinatorial approach to planar non-colliding robot arm motion planning, in: *Proc. 41st Symp. Foundations of Comp. Sci. (Redondo Beach, CA, 2000)*, IEEE Comp. Soc., 443–453.

[45] S. Suri, Polygons, in: *Handbook of Discrete and Computational Geometry (J. Goodman and J. O'Rourke, eds.)*, CRC Press, Boca Raton, FL, 1997, pp. 429–444.

[46] I. E. Sutherland, R. F. Sproull, and R. A. Schumacker, A characterization of ten hidden-surface algorithms, *ACM Computing Survey* **9** (1984), 1–55.

[47] L. Szabó, Recent results on illumination problems, in: *Intuitive Geometry (I. Bárány and K. Böröczky, eds.)*, Bolyai Society Mathematical Studies, vol. 6, J. Bolyai Math. Soc., Budapest, 1997, pp. 207–221.

[48] Cs. D. Tóth, Art gallery problem with guards whose range of vision is $180^\circ$, *Comput. Geom. Theory Appl.* **17** (2000), 121–134.

[49] Cs. D. Tóth, Illuminating disjoint line segments in the plane, submitted, 2000.

[50] Cs. D. Tóth, Illuminating both sides of line segments, in: *Discrete and computational geometry (J. Akiyama, M. Kano, and M. Urabe, eds.)*, Lecture Notes in Comp. Sci., vol. 2098, Springer, Berlin, 2001, pp. 370–380.

[51] Cs. D. Tóth, Illuminating polygons with vertex $\pi$-floodlights, in: *Proc. Int. Conf. on Comput. Sci. (San Francisco, CA, 2001)*, Lecture Notes in Comp. Sci., vol. 2073 (V. N. Alexandrov et al., eds.), Springer, Berlin.

[52] Cs. D. Tóth, A note on binary plane partitions, in: *Proc. 17th ACM Symp. on Comput. Geom. (Medford, MA, 2001)*, ACM Press, 151–156.

[53] Cs. D. Tóth, Art galleries with guards of uniform range of vision, *Comput. Geom. Theory Appl.* **21** (2002), 185–192.

[54] Cs. D. Tóth, Convex decomposition for polygons with holes and a new bound on vertex guards, submitted, 2002.

[55] Cs. D. Tóth, Illumination in the presence of opaque line segments in the plane, *Comput. Geom. Theory Appl.* **21** (2002), 193–204.

[56] Cs. D. Tóth, Illumination of polygons by $45^\circ$-floodlights, *Discrete Maths.* (2002), to appear.

[57] G. T. Toussaint, Computing geodesic properties inside a simple polygon, *Revue D'Intelligence Artificielle* **3** (1989), 9–42.

[58] J. Urrutia, Art gallery and illumination problems, in: *Handbook on Computational Geometry (J. R. Sack and J. Urrutia, eds.)*, Elsevier, Amsterdam, 2000, pp. 973–1027.

[59] J. Urrutia, personal communication, 2001.

[60] F. A. Valentine, Visible shorelines, *Amer. Math. Monthly* **77** (1970), 146–152.

[61] J. Zaks, A note on illuminating line segments in the plane, manuscript, 1993.

# Curriculum Vitae

**Csaba Dávid Tóth**

born in 1974 in Szolnok, Hungary

| | |
|---|---|
| 2000–2002 | **Swiss Federal Institute of Technology Zürich (ETHZ)**<br>Institute for Theoretical Computer Science<br>Graduate student (CGC Program) |
| 1998–2000 | **Budapest University of Technology, Budapest (BME)**<br>Department of Geometry, Institute of Mathematics<br>Graduate student |
| 1998 spring | **Institut National de Rechereche en Informatique et en Autimatique, Sophia Antipolis (INRIA)**<br>Project SLOOP, practice |
| 1997–1998 | **Université de la Méditerranée, Marseille**<br>Insitut de Mathématiques de Luminy<br>DEA de Math. Discrètes et Fondements de l'Informatique |
| 1992–1997 | **Eötvös Loránd University, Budapest (ELTE)**<br>M.Sc. in Mathematics |
| 1988–1992 | **Verseghy Ferenc Gimnázium, Szolnok**<br>Baccalauréat |