Lines in Graphs

Jozef Jirásek, Pavel Klavík

Faculty of Mathematics and Physics,
Charles University in Prague

Contents

Abstract

In this work we explore a particular structural property of graphs. Every graph defines a finite discrete metric space with its vertices as points and with metric being the distance of vertices. The notion of lines from Euclidean space is extended to this metric: a line defined by two points contains all the points satisfying the triangle equality. V. Chvátal extended a particular theorem from the euclidean plane to this metric. W. Klee and S. Wagon later conjectured, that a discrete metric space with n points either contains n different lines, or one line containing all of the points. We show several results about lines in graph metric spaces related to this conjecture.

First, we describe the structure of lines defined by a pair of adjanced and non-adjanced vertices. We show several results about lines containing all the points in a graph, so called "universal lines". These lines are strongly related to cuts and separators in a graph. We give formulas to calculate the number of different lines for several graph classes: cycles, trees, complete k-partite graphs. Then we use them to show several examples of graphs where every line is different. A complete graph is one such example.

Further, we prove that for any given integer k, there exist only finitely many graphs with exactly k different lines without a universal line. There exist infinitely many graphs with exactly k different lines, if and only if there exists at least one such a graph containing a bridge.

We also describe the problem of reconstructing a graph from a list of its lines. This is not always possible, as there exist different graphs with the same line structure. However, we can solve this problem for certain graphs, such as trees.

Last, we consider a generalization of the problem for weighted graphs, which are equivalent to finite discrete metric spaces. We will see that many of our results hold even in this general case.

Chapter 1

Introduction

1.1 Motivation in geometry

Let us introduce the problem with a short motivation in geometry and an open problem which inspired our research. In March 1893, J. J. Sylvester [1] proposed the following problem:

> Prove that it is not possible to arrange any finite number of real points so that a right line through every two of them shall pass through a third, unless they all lie in the same right line.

The problem was solved about forty years later by T. Gallai and it is known as the Sylvester–Gallai theorem. More information about this theorem can be found in Chvátal [2].

In the same article Chvátal considered a generalization of this theorem for finite discrete metric spaces. This generalization, also known as the Sylvester–Chvátal theorem, was proven by Chen [3] a year later. However, in this article we consider a more natural definition of a line for metric spaces induced by graphs. Using this definition the Sylvester–Chvátal theorem does not hold. We will later show that the circle C_5 is a simple counterexample.

A related theorem proven by de Bruijn and Erdős [4] states:

Theorem 1.1 (de Bruijn and Erdős, 1948)**.** *Let S be a set of n points in the euclidean plane. Then either*

(1) there exists a line that contains all points in S, or
(2) the points determine at least n different lines.

This is not unexpected. If all the points are collinear, they define exactly one line passing through all of them. Otherwise, we can find many different lines. If the points are in general position, every pair of points defines a unique line, $\binom{n}{2}$ different lines in total. The lower bound for the number of different lines if not all of the points are collinear can be achieved, as shown in the following figure.

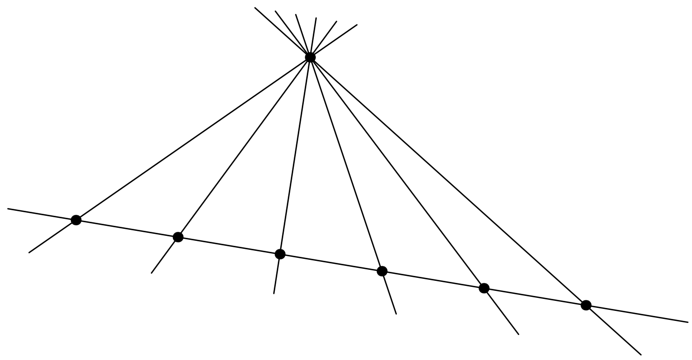

Figure 1.1: A positioning of n points defining exactly n different lines.

Can this theorem be extended to finite discrete metric spaces as well? This question was asked by Klee and Wagon in Section 6 (Points on Lines) of [5]. In this work, we consider the metric space defined by a simple graph. We show many interesting properties of lines in a graph and prove several results closely related to this problem.

1.2 Metric space and lines in a graph

Unless explicitly stated otherwise, throughout the work we consider only simple unoriented connected graphs. We assume knowledge of basic notions and notation of graph theory, such as the complete graph K_n or the circle C_n. Unless stated otherwise, G denotes a graph, V denotes the set of its vertices and E denotes the set of its edges.

We define the metric space induced by a graph $G = (V, E)$ as follows:

Definition 1.2 (Graph metric space)**.** Let G be a simple connected graph. A graph metric space for G is a metric space (V, d) such that

- $V = V(G)$, points are vertices of G.
- $d(u, v)$, where $u, v \in V$, is the distance of vertices u and v in G, defined as the number of edges on the shortest path from u to v.

It is easy to see that this indeed satisfies the definition of a metric space.

For a given metric space we define a *line* as follows:

Definition 1.3. Let (X, ρ) be a metric space. For each $a, b \in X$, $a \neq b$ we define the line $\overleftrightarrow{ab}$ as:

$$\overleftrightarrow{ab} = \{x \mid \rho(x,a) + \rho(a,b) = \rho(x,b) \vee \tag{1.1}$$

$$\vee\ \rho(a,x) + \rho(x,b) = \rho(a,b) \vee \tag{1.2}$$

$$\vee\ \rho(a,b) + \rho(b,x) = \rho(a,x)\} \tag{1.3}$$

In another words, points belonging to the line $\overleftrightarrow{ab}$ are points satisfying the triangle equality with a and b. Two lines are considered the same (equal, identical...) if they contain the same sets of points. Note that the points a and b always lie on the line $\overleftrightarrow{ab}$, and that the line $\overleftrightarrow{ab}$ is always equal to the line $\overleftrightarrow{ba}$.

A generalization of the de Bruijn-Erdős theorem for finite discrete metric spaces follows.

Conjecture 1.4 (V. Klee and S. Wagon [5]). *Let (S, ρ) be a finite discrete metric space, such that $|S| = n$. Then either*

(1) there exists a line containing all the points in S, or
(2) there exist at least n different lines defined by points in S.

If there is a line containing all the points in S, we will call such a line *universal line*.

Note the difference from Theorem 1.1 for the euclidean plane. For instance, in the euclidean plane any two points lying on a line $\overleftrightarrow{ab}$ define the same line as $\overleftrightarrow{ab}$. In a general metric space this does not need to hold.

1.3 A simple example

Let us illustrate the definitions with a simple example graph.

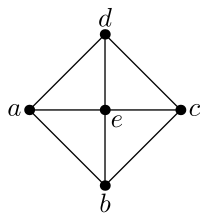

Figure 1.2: An example graph.

First, consider the line $\overleftrightarrow{ab}$. The point d lies on $\overleftrightarrow{ab}$, as it satisfies condition (1.1) from the definition. Similarly, the point c satisfies (1.3). However, the point e does not satisfy either of the conditions, therefore it does not lie on $\overleftrightarrow{ab}$. We can also see that a and b satisfy all of the conditions. It is obvious that a line always contains the two points defining it. Therefore, $\overleftrightarrow{ab} = \{a, b, c, d\}$.

In the same way one can deduce that $\overleftrightarrow{bc} = \overleftrightarrow{cd} = \overleftrightarrow{da} = \overleftrightarrow{ab}$, that $\overleftrightarrow{ae} = \overleftrightarrow{ec} = \{a, e, c\}$, similarly $\overleftrightarrow{be} = \overleftrightarrow{ed} = \{b, e, d\}$, and finally $\overleftrightarrow{ac} = \overleftrightarrow{bd} = \{a, b, c, d, e\}$. Lines $\overleftrightarrow{ac}$ and $\overleftrightarrow{bd}$ are universal, therefore the graph satisfies Conjecture 1.4.

We will soon show easier ways to determine whether a given point lies on a given line.

Chapter 2

Basic Results

In this chapter we introduce basic observations about the structure of different lines in graphs.

2.1 Line structure

A line $\overleftrightarrow{ab}$ induces four important subsets of V.

Definition 2.1 (Subsets induced by a line)**.** Let $\overleftrightarrow{ab}$ be a line in a graph. Then we define:

$$\begin{aligned} a\text{-part} &= \{x \mid d(x,a) + d(a,b) = d(x,b)\} \\ ab\text{-part} &= \{x \mid d(a,x) + d(x,b) = d(a,b)\} \\ b\text{-part} &= \{x \mid d(a,b) + d(b,x) = d(a,x)\} \\ \emptyset\text{-part} &= V \setminus \overleftrightarrow{ab} \end{aligned}$$

Each subset corresponds to a single condition defining the line, the $\emptyset$-part are vertices not belonging to the line. An example:

Note that the vertices a and b belong in both ab-part and a-part or b-part respectively. If we specifically want to exclude those points, we will talk about *proper* parts.

Definition 2.2 (Proper parts of a line)**.**

$$\begin{aligned} a^+\text{-part} &= a\text{-part} \setminus \{a\} \\ ab^+\text{-part} &= ab\text{-part} \setminus \{a,b\} \\ b^+\text{-part} &= b\text{-part} \setminus \{b\} \end{aligned}$$

We also say that a vertex c is an *inner point* of the line $\overleftrightarrow{ab}$, if $c \in ab$-part and that it is an *outer point* of the line, if $c \in a$-part or $c \in b$-part. Similarly,

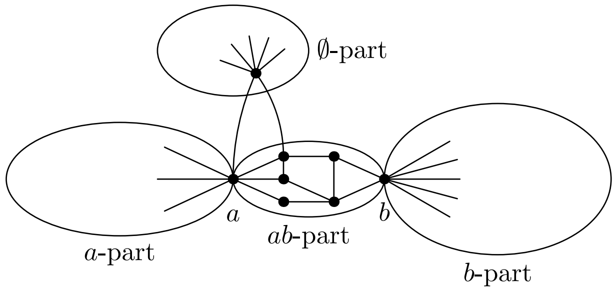

Figure 2.1: Four subsets ∅-part, a-part, b-part and ab-part.

proper inner points and *proper outer points* are inner (or outer) points other than the vertices a and b.

First we give two observations about inner and outer points and their connection to shortest paths in the graph. We introduce the notion of *betweenness*:

Definition 2.3 (Betweenness)**.** For three vertices $a, b, c \in G$ we say that the vertex b lies between vertices a and c, and we write $[abc]$, if there exists a shortest path in G from the vertex a to the vertex c, such that this path contains the vertex b.

Similarly, we will write $[v_1 v_2 \dots v_n]$ if there exists a shortest path from v_1 to v_n such that this path contains the vertices $v_2, v_3, \dots v_{n-1}$ in this order. Naturally, this path may contain other vertices as well.

Observe that if $[v_1 v_2 \dots v_n]$, then for every i, j, k with $1 \le i \le j \le k \le n$, $[v_i v_j v_k]$ must hold. Second, observe that if $[abc]$ and $[acd]$, then $[abcd]$ must hold as well.

Observation 2.4 (Inner points)**.** *Point c is an inner point of a line $\overleftrightarrow{ab}$ iff $[acb]$.*

Proof. Any point c on the shortest path from a to b satisfies $d(a, c) + d(c, b) = d(a, b)$. To prove the other implication, let c be an inner point of $\overleftrightarrow{ab}$. Consider the shortest path from a to c, and the shortest path from c to b. Joining them together we create a path with length $l = d(a, c) + d(c, b) = d(a, b)$, therefore it is a shortest path from a to b. □

Observation 2.5 (Outer points)**.** *Point c is an outer point of the line $\overleftrightarrow{ab}$, iff $[cab]$ or $[abc]$.*

Proof. Analogous to the previous observation. □

Lemma 2.6 (Shortest path)**.** *Point c lies on the line $\overleftrightarrow{ab}$ iff $[cab]$ or $[acb]$ or $[abc]$.*

Proof. A direct consequence of the previous two observations. □

This lemma is very useful in determining whether a point lies on a line. We will refer to it as the *Shortest Path lemma.*

Corollary 2.7 (Triangle)**.** *Let a, b and c be three points. Then either*

(1) $a \in \overleftrightarrow{bc}$, $b \in \overleftrightarrow{ac}$ *and* $c \in \overleftrightarrow{ab}$
(2) or $a \notin \overleftrightarrow{bc}$, $b \notin \overleftrightarrow{ac}$ *and* $c \notin \overleftrightarrow{ab}$.

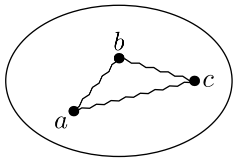

Figure 2.2: Vertices a, b and c create a triangle.

If the first case holds, we will call the points a, b, c *collinear.* Note that this relation does not need to be transitive. However, if three points are not collinear, they define three different lines.

Lemma 2.8 (Properties of line parts)**.** *For each line* $\overleftrightarrow{ab}$ *the following holds:*

(1) Each of the subgraphs induced by a-part, b-part and ab-part is connected.
(2) The subgraph induced by all vertices belonging to the line is also connected.
(3) No vertex of the graph, except for a and b, belongs to two different parts.

Proof. First, consider the a-part for example. Let c be a vertex in a-part. By Observation 2.5 $\overleftrightarrow{cab}$. However, this must also hold for any vertex d on the shortest path from c to a. Therefore, the entire path from c to a belongs to a-part. This holds for any c in the a-part, therefore the a-part is connected. The proof for other parts is similar.

The second statement is obvious, as each of the parts is connected, a connects a-part and ab-part and b connects ab-part and b-part. Note that the $\emptyset$-part does not have to be connected.

Last, recall the definition of line parts. Consider a vertex c on the line $\overleftrightarrow{ab}$, and the three distances $d(c, a)$, $d(a, b)$ and $d(b, c)$. If c belongs to the a-part, $d(c, b)$ is the maximum of the three. Moreover, if $c \neq a$, $d(c, a) \geq 1$ and $d(c, b)$ is a strict maximum. Similarly, if c is in ab-part and it is not a or b, the strict maximum is $d(a, b)$. And if c is in b-part, $c \neq b$, the strict maximum is $d(a, c)$. We can see that for no such c two of those can hold at the same time. □

2.2 Distinguishing of lines

Here we show several ways to show that two lines are different. We have already seen that if three points are not collinear, they define three different lines. We will show a similar result about so-called *distance subgraph.*

Definition 2.9. Graph H is a *distance subgraph* of a graph G, if H is a subgraph of G and for every $a, b \in V_H$ is $d_H(a,b) = d_G(a,b)$.

The metric defined by H is a restriction of the metric defined by G. Also, if a subgraph of G is a complete graph, it is also a distance subgraph—the distance between any two points is 1 in either graph.

Lemma 2.10 (Distance subgraphs)**.** *Let H be a distance subgraph of G. Then for each $a, b, c \in V_H$ $c \in \overleftrightarrow{ab}$ in H iff $c \in \overleftrightarrow{ab}$ in G.*

Proof. The distances between a, b and c are equal in H and in G, therefore if c satisfies one of the equalities in the line definition in H, it must satisfy it in G, and vice versa. □

Corollary 2.11. *Let H be a distance subgraph of G. Then G contains at least as many different lines as H.*

Proof. Every two lines different in H must be different also in G. Note that if two lines are equal in H, they do not need to be equal in G—there might be a vertex in G, but not in H, which distinguishes them. □

Corollary 2.12. *Let G be a triangle-free graph with a maximum degree of $\Delta(G) \geq 3$. Then G contains at least $\binom{\Delta(G)}{2} + 1$ different lines.*

Proof. Let a be a vertex with the maximal degree and let us denote its neighbors $b_1, b_2, \ldots, b_{\Delta(G)}$. If G is triangle-free, there is no edge between two vertices b_i and b_j. Therefore the subgraph H induced by $\{a, b_1, b_2, \ldots, b_{\Delta(G)}\}$ is a distance subgraph. Let us count the number of lines in H.

First, we can see that every line $\overleftrightarrow{ab_i}$ contains all vertices of the subgraph H since the shortest path from b_j to b_i is $b_j a b_i$. Second, every line $\overleftrightarrow{b_i b_j}$ contains only the points b_i, a and b_j. Therefore, the lines $\overleftrightarrow{b_i b_j}$ are pairwise different. By choosing i and j from 1 to $\Delta(G)$, we obtain $\binom{\Delta(G)}{2}$ different lines. The line $\overleftrightarrow{ab_i}$ is different from all of them. By the previous corollary, the graph G must have at least $\binom{\Delta(G)}{2} + 1$ lines. □

2.3 Edge lines

We will soon show that, compared to the general case, the line structure is much simpler if the two vertices defining a line are connected by an edge. We will call such lines *edge lines* and all other lines *long lines.* Edge lines do not contain any proper inner points.

Observation 2.13 (Edge Lines)**.** *Let $\overleftrightarrow{ab}$ be an edge line. Then $\overleftrightarrow{ab}$ contains exactly those points c for which $d(a,c) \neq d(b,c)$.*

Proof. Note that since ab is an edge, for any c holds that $|d(a,c) - d(b,c)| \leq 1$.

Let c be a point on the line $\overleftrightarrow{ab}$. It must be an outer point, therefore either $d(c,a) + d(a,b) = d(c,b)$ or $d(c,b) + d(b,a) = d(c,a)$. In either case $d(a,c) \neq d(b,c)$.

If $d(a,c) \neq d(b,c)$, then $d(b,c) = d(a,c) \pm 1$, and $d(a,b) = 1$, therefore either $d(c,b) = d(c,a) + d(a,b)$ or $d(c,b) + d(b,a) = d(c,a)$. In either case $c \in \overleftrightarrow{ab}$. □

Lemma 2.14. *Let $\overleftrightarrow{ab}$ be an edge non-universal line. Then there exists an odd cycle containing a and b.*

Proof. Consider a point c, such that $c \notin \overleftrightarrow{ab}$, and that $d(a,c)$ is minimal. Since $\overleftrightarrow{ab}$ is an edge line, $d(b,c) = d(a,c)$. Then the shortest paths ac and bc must be disjoint, otherwise there exists a vertex d on both of the paths. The distances $d(a,d)$ and $d(b,d)$ must be equal. However, that means $d \notin \overleftrightarrow{ab}$, which is in contradiction with the minimality of $d(a,c)$.

Therefore, the paths ac, cb and the edge ba form an odd cycle. □

Corollary 2.15. *Each edge line in a bipartite graph is a universal line.*

Proof. If the graph is bipartite, it does not contain any odd cycles. Therefore every edge defines a universal line. □

A result not unlike the converse of the previous Lemma 2.14 holds.

Proposition 2.16. *Let graph G contain an odd cycle $a_1a_2\dots a_k$. Then for each vertex c, there exists an edge a_ia_{i+1} of the cycle such that $c \notin \overleftrightarrow{a_ia_{i+1}}$.*

(Naturally, if $i = k$, we define a_{i+1} to be a_1.)

Proof. For a contradiction assume that there is a vertex c such that for each edge a_ia_{i+1} the vertex c lies on the line $\overleftrightarrow{a_ia_{i+1}}$.

Since $c \in \overleftrightarrow{a_ia_{i+1}}$, we have $d(c,a_i) \neq d(c,a_{i+1})$. This means that $d(c,a_{i+1}) = d(c,a_i)+1$ or $d(c,a_{i+1}) = d(c,a_i)-1$, that it, the parity of $d(c,a_{i+1})$ is different from the parity of $d(c,a_i)$. Thus the parity of distances of the vertex c from vertices $a_1, a_2, \dots, a_k$ alternates between odd and even. However, if this holds, then the circle must be even, which is a contradiction. □

Let us consider a complete graph K_n. Each line $\overleftrightarrow{ab}$ in this graph is an edge line. Moreover, for every $c \neq a,b$ we have $d(a,c) = d(b,c) = 1$, and so this line contains only the points a and b. Therefore, all the lines in the graph K_n are pairwise different, so K_n contains $\binom{n}{2}$ different lines.

Corollary 2.17. *Every graph G with a maximal complete subgraph of size $\omega(G)$ contains at least $\binom{\omega(G)}{2}$ different lines.*

Proof. Recall Lemma 2.10 on distance subgraphs. Each complete subgraph is also a distance subgraph, therefore the corollary holds. □

2.4 Long lines

Recall that we call a line $\overleftrightarrow{ab}$ *long line*, if ab is not an edge. The structure of long lines is somewhat different from edge lines. Consider a long line $\overleftrightarrow{ab}$. We will call the subgraph induced by its inner points *layer subgraph.* The vertices are divided into *layers* according to their distance from a. Note that we can interchangeably use distance from b; if two inner points have equal distance from a, they must also have equal distances from b.

We can see that edges can exist only between vertices in two consecutive layers, or between vertices in the same layer. We will call the latter *ladder edges.*

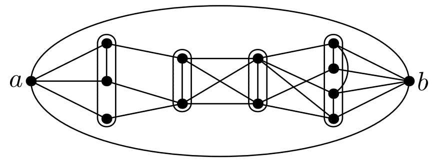

Figure 2.3: An example of a layer subgraph.

Lemma 2.18 (Unconnected parts). *Let $\overleftrightarrow{ab}$ be a long line. Then there exists no edge between a^+-part and ab^+-part, and no edge between a^+-part and b-part. Similarly for b^+-part.*

Proof. Let c be a vertex in a^+-part and let d be a vertex in ab^+-part. For a contradiction, assume that cd is an edge. From the definition of proper parts, we can see that $d(d,b) < d(a,b) < d(c,b)$. However, $d(c,d) = 1$, therefore $|d(d,b) - d(c,b)|$ is at most 1. This gives a contradiction.

If $c \in a^+$-part, then $d(c,b) = d(c,a) + d(a,b) > 1$. Therefore, there can be no edge between a vertex in a^+-part and the vertex b.

Now, let c be a vertex in a^+-part and let e be a vertex in b^+-part. For a contradiction, assume that ce is an edge. Without loss of generality, let $d(a,c) \leq d(b,e)$. From the definition of parts, and because ab is not an edge, we get:

$$d(a,e) = d(a,b) + d(b,e) > d(b,e) + 1$$

However, since ce is an edge, by the triangle inequality it also holds that:

$$d(a,e) \leq d(a,c) + d(c,e) \leq d(b,e) + 1$$

This again gives a contradiction. $\square$

Note that this lemma does not hold in general for edge lines.

Chapter 3

Cuts, Separators, and Universal Lines

Cuts and separators are natural structures in graph theory. In this chapter, we will prove some result on the existence of an universal line in a graph and we will show their close connection with cuts and separators.

A cut is in general a set of edges that divides a graph into two or more components. For a set of vertices A, with $\emptyset \neq A \neq V$, a *fundamental cut* C_A is uniquely defined by the set A, and consists of all edges that go between A and $V \setminus A$, that is, of all edges e, for which $|e \cap A| = 1$. All cuts considered in this chapter are fundamental cuts.

3.1 Bridges

Let us start with the simplest cut—a bridge. A bridge defines an edge line. We will show that each such line is universal.

Observation 3.1 (Bridge universality)**.** *Each edge line defined by a bridge is a universal line.*

Proof. Let ab be a bridge defined by a set of vertices A. Let $B = V \setminus A$. Without loss of generality, let $a \in A$, $b \in B$. Let c be a vertex in A. Note that ab is the only edge connecting A and B, therefore every path from c to b must go through a. By the Shortest Path lemma, $c \in \overleftrightarrow{ab}$.

Similarly, for each vertex d in B, the shortest path from d to a goes through b. $\square$

Another property of bridges is that subdividing a bridge does not significantly change the structure of the lines in a graph.

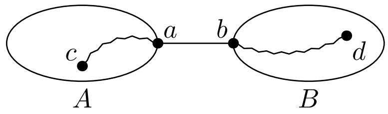

Figure 3.1: A bridge ab defines a universal line.

Theorem 3.2 (Bridge pumping). *Let ab be a bridge in the graph G defined by a set of vertices A. Let $B = V \setminus A$. Let G' be a graph obtained by subdividing the edge ab with a new vertex c. Then the following holds:*

(1) Each line $\overleftrightarrow{xy}$ in G contains the same set of points in G and in G', only in G' it may additionaly contain the point c.
(2) Every two lines that are equal in G are also equal in G'.
(3) Each line cx in G' is equal to some line defined only by vertices of G.

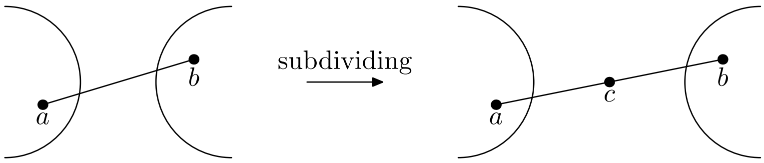

Figure 3.2: Subdividing of the bridge ab with the vertex c.

Proof. (*1*) First, observe that for every three vertices $x, y, z \in G$ the relation $[xyz]$ is not changed after subdividing a bridge. Together with the Shortest Path lemma this gives that $z \in \overleftrightarrow{xy}$ in G if and only if $z \in \overleftrightarrow{xy}$ in G'.

(*2*) If the line $\overleftrightarrow{xy}$ in G contains both vertices a and b, we now know that it contains them in G' as well. By Lemma 2.8, the line is connected, therefore it must contain c as well. If $\overleftrightarrow{xy}$ contains neither a nor b, we can similarly see that it can not contain c, otherwise the line would not be connected.

So let the line $\overleftrightarrow{xy}$ contain exactly one of the vertices a and b, without loss of generality, let it contain a. The line has to be connected, therefore both x and y must be in A. We can see that $[xcy]$ cannot hold, therefore c is not an inner point of $\overleftrightarrow{xy}$. Now assume that c is an outer point of the line, and let, for example, $[xyc]$. However, cb is a bridge, and therefore also $[xycb]$. But then b lies on the line $\overleftrightarrow{xy}$, which is a contradiction.

We see that c lies on the line $\overleftrightarrow{xy}$ in G' if and only if $\overleftrightarrow{xy}$ in G contains both a and b. Therefore, two equal lines $\overleftrightarrow{xy}$ and $\overleftrightarrow{uv}$ either both contain c in G' or none of them does. We have already seen that they can not differ in G' by any other points.

(*3*) Consider a line $\overleftrightarrow{xc}$ in G', without loss of generality let $x \in A$. We are going to show that $\overleftrightarrow{xc} = \overleftrightarrow{xb}$.

First, note that for each vertex $y \in B$, both $[xcy]$ and $[xby]$ hold. Therefore, both lines contain all of the vertices in B. Now let $z \in A$ be an inner point of $\overleftrightarrow{xc}$. This means that $[xzc]$, but because cb is a bridge, it also holds that $[xzcb]$, and therefore z lies on the line $\overleftrightarrow{xb}$ as well. Observe that the converse also holds: If z is an inner point of $\overleftrightarrow{xb}$, then $[xzb]$, and naturally also $[xzcb]$, and then $z \in \overleftrightarrow{xc}$.

Similarly, if z is an outer point in x-part, we can see that $[zxc]$ if and only if $[zxb]$. Therefore, every point of the line $\overleftrightarrow{xc}$ is a point of the line $\overleftrightarrow{xb}$, and vice versa. □

Corollary 3.3. *Subdividing a bridge does not change the number of different lines in a graph.* □

3.2 Cuts

We now show several results about general cuts and their connection with universal lines.

Lemma 3.4 (Cut union)**.** *Let C_A be a fundamental cut. Then*

$$\bigcup_{uv \in C_A} \overleftrightarrow{uv} = V.$$

Proof. We are going to show that for each vertex $a \in G$, there exists an edge uv of the cut C_A such that $a \in \overleftrightarrow{uv}$.

Without loss of generality, let $a \in A$. Consider the shortest path form a to any vertex not in A. This path must pass through at least one edge uv of the cut. But we see that $[auv]$, therefore $a \in \overleftrightarrow{uv}$. □

Corollary 3.5. *Let k be the size of the minimal cut in the graph G, that is, the cut with minimal number of edges. Then there exists a line in the graph G containing at least $\frac{|V|}{k}$ points.*

Proof. Assume the corollary does not hold. Then the cut defines k edge lines, each containing less than $\frac{|V|}{k}$ points. However, this is not enough for their union to contain all of the vertices in the graph G. □

Note that if the graph contains a bridge, $k = 1$ and there is a line with $\frac{|V|}{1} = |V|$ points—an universal line.

Let the line $\overleftrightarrow{ab}$ be an edge universal line. Consider the cut defined by a-part of the line. We will call this cut a *universal cut.*

Proposition 3.6 (Universal cuts)**.** *Let ab be an edge of the graph G. The edge line $\overleftrightarrow{ab}$ is universal if and only if there exists a cut C_A such that $ab \in C_A$ and for each edge $cd \in C_A$ is $d(a,c) = d(b,d)$. Moreover, for a given edge line $\overleftrightarrow{ab}$, this cut is defined uniquely by $A = a$-part.*

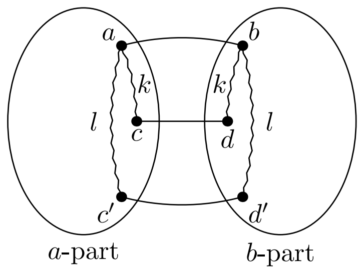

Figure 3.3: A universal cut defined by the edge universal line $\overleftrightarrow{ab}$.

Proof. Let the line $\overleftrightarrow{ab}$ be an edge universal line, and let C_A be a cut defined by $A = a$-part of the line. Let $B = V \setminus A$. The line $\overleftrightarrow{ab}$ is an edge universal line, therefore $\emptyset$-part $=$ ab-part $= \emptyset$, and $B = b$-part of $\overleftrightarrow{ab}$. Let cd be an edge of the cut C_A, such that $c \in A$ and $d \in B$.

By the triangle inequality, it holds that

$$d(a, d) \leq d(a, c) + d(c, d) = d(a, c) + 1.$$

The vertex d belongs to the b-part, therefore

$$d(a, d) = d(a, b) + d(b, d) = d(b, d) + 1.$$

Hence:

$$d(b, d) \leq d(a, c).$$

Similarly:

$$\begin{aligned} &d(b, c) \leq d(b, d) + d(d, c) = d(b, d) + 1, \\ &d(b, c) = d(b, a) + d(a, c) = d(a, c) + 1, \\ &d(a, c) \leq d(b, d). \end{aligned}$$

Therefore $d(a, c) = d(b, d)$.

Conversely, let C_A be a cut defined by a set of vertices A, and let ab be one of its edges, such that for each edge $cd \in C_A$, the equality $d(a, c) = d(b, d) = k$ holds.

Consider a vertex x, without loss of generality, let $x \in A$. Consider the shortest path xb. This path needs to cross one of the edges of C_A, let us denote the edge cd, $c \in A$, $d \notin A$. Consider the path $x-c-a-b$. Since $d(c, a) = d(d, b)$,

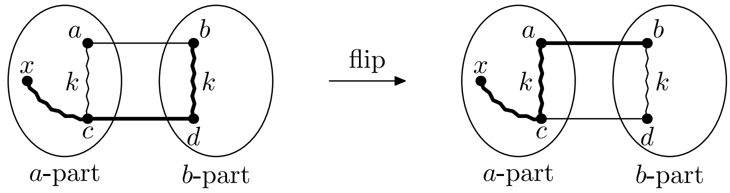

Figure 3.4: The paths $x - c - d - b$ and $x - c - a - b$ are equally long.

this path is as long as the path $x - c - d - b$, therefore it is a shortest path as well. This gives $[xab]$, and from that, $x \in \overleftrightarrow{ab}$. Similarly if $x \in V \setminus A$.

Finally, we prove that for a given edge universal line $\overleftrightarrow{ab}$, there exists only one cut with the desired properties. Let C_A be such a cut defined by a set of vertices A, and let $B = V \setminus A$. Exactly as above, we can show that for every $x \in A$, the relation $[xab]$ holds, therefore $x \in a$-part. Similarly, for every $y \in B$, $[aby]$ holds, and so $y \in b$-part. The line $\overleftrightarrow{ab}$ is a universal edge line, therefore the $\emptyset$-part and ab-part are empty. This gives $A = a$-part and $B = b$-part. □

3.3 Articulations

Lemma 3.7 (Articulation)**.** *Let the vertex c separate the graph G into two parts A and B. That is, let $A \cup B \cup \{c\} = V$, $A \cap B = \emptyset$ and no edge connects a vertex in A with a vertex in B. Let a and b be two different vertices in A. Then the following statements are equivalent:*

(1) The line $\overleftrightarrow{ab}$ contains all vertices in B.
(2) The line $\overleftrightarrow{ab}$ contains a vertex d, $d \in B$.
(3) The vertex c is an outer point of the line $\overleftrightarrow{ab}$.

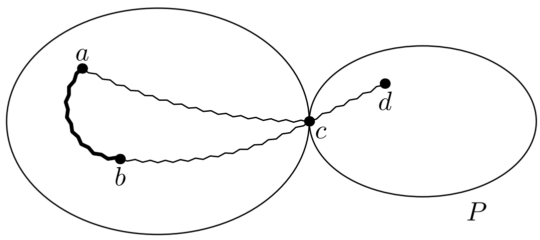

Figure 3.5: Articulation c separates A and B.

Proof. *(1)* →*(2)*: Trivial.

(2) →*(3)*: Obviously, d is not an inner point of the line $\overleftrightarrow{ab}$. Since each of the outer parts of the line is connected, the vertex c must be in the same part as d, therefore it is an outer point as well.

(3) →*(1)*: Without loss of generality, let $c \in b$-part. Then $[abc]$ holds. Since c is an articulation, for every $d \in B$ must hold $[acd]$. Therefore, $[abcd]$ holds, and $d \in \overleftrightarrow{ab}$. □

Lemma 3.8. *Let the vertex c separate the graph G as in the previous lemma. Let vertices a, b belong to A, and let d belong to B. Then the lines $\overleftrightarrow{ad}$ and $\overleftrightarrow{bd}$ are equal on vertices in B, that is, $\overleftrightarrow{ad} \cap B = \overleftrightarrow{bd} \cap B$.*

Proof. Let $x \in B$ be an inner point of the line $\overleftrightarrow{ad}$. Therefore it holds that $[acxd]$. But since c is an articulation, the shortest paths ad and bd must contain the same path between c and d. Then it must also hold that $[bcxd]$, and x is an inner point of $\overleftrightarrow{bd}$.

Now let $y \in B$ be an outer point of $\overleftrightarrow{ad}$, therefore $[acdy]$. Similarly, $[bcdy]$ holds as well and y is an outer point of $\overleftrightarrow{bd}$.

Analogously we prove that all vertices of the line $\overleftrightarrow{bd}$ in B belong also to the line $\overleftrightarrow{ad}$. □

Now we show several observations on the structure of lines in a graph with an articulation, based on the lines in each of its subgraphs separated by the articulation.

Definition 3.9 (Graph join)**.** Given two graphs G and H, we will call a graph F *join* of graphs G and H, denoted by $F = G \cdot H$, if

- F is separated by a vertex v into two parts, G' and H', as in the previous lemma.
- G is a subgraph of F induced by $G' \cup \{v\}$.
- H is a subgraph of F induced by $H' \cup \{v\}$.

Note that both G and H are distance subgraphs of $G \cdot H$, see Definition 2.10.

For the following observations, consider two graphs G and H, and their join $G \cdot H$ separated by a vertex c.

Observation 3.10. *Let the line $\overleftrightarrow{ab}$ be a universal line in G, such that c is its outer point. Then the line $\overleftrightarrow{ab}$ is universal even in $G \cdot H$.*

Proof. Line $\overleftrightarrow{ab}$ is universal in the distance subgraph G, therefore it contains all the vertices of G even in $G \cdot H$. The vertex c is its outer point, therefore by the previous lemma 3.7 the line contains all vertices of H as well. □

Observation 3.11. *Let a, b be different vertices in G, and let d, e be different vertices in H, such that the line $\overleftrightarrow{ab}$ is either not universal in G, or the vertex c is not its outer point.*

Then the lines $\overleftrightarrow{ab}$ and $\overleftrightarrow{de}$ are different in $G \cdot H$.

Proof. First, assume that the line $\overleftrightarrow{ab}$ is not universal in G. If the line $\overleftrightarrow{de}$ in $G \cdot H$ does not contain either of the vertices a or b, it is trivially different from the line $\overleftrightarrow{ab}$. If $\overleftrightarrow{de}$ contains both a and b, it must contain at least one point of G different from c. By the previous lemma 3.7, it contains the entire G. Since we assumed that $\overleftrightarrow{ab}$ is not universal in G, there is at least one vertex of G not in $\overleftrightarrow{ab}$ but in $\overleftrightarrow{de}$.

Second, assume that the vertex c is not an outer point of the line $\overleftrightarrow{ab}$. By the previous lemma, the line $\overleftrightarrow{ab}$ does not contain any point of H different from c. Therefore it can not contain both d and e, so it is clearly different from the line $\overleftrightarrow{de}$. □

Observation 3.12. *Let the graph G contain k different lines and let the graph H contain ℓ different lines. Then the graph $G \cdot H$ contains at least $k + \ell - 1$ different lines.*

Proof. Consider only lines defined by a pair of points in the same subgraph (G or H). Since G and H are distance subgraphs, two different lines defined by points in the same subgraph must be also different in $G \cdot H$. The lemma 3.7 shows that any line defined by points in G is different from any line defined by points in H, except for the case when both are universal lines with c as an outer point. However, each graph can have only one universal line, therefore at most one such pair exists.

To sum up, if we consider all the lines defined in G and all the lines defined in H, we get $k + \ell$ lines, with at most two of them equal, therefore the desired $k + \ell - 1$ different lines. □

3.4 Long universal lines

If the graph contains a long universal line, this line must have a very specific structure.

Observation 3.13 (Long universal line)**.** *Let the line $\overleftrightarrow{ab}$ be a long universal line. If its a^+-part is not empty, then the vertex a is an articulation separating the a^+-part from the rest of the graph. Similarly for b^+-part and the vertex b.*

Proof. In the section about long lines, we have proven that there are no edges between a^+-part and ab^+-part, nor between a^+-part and b-part. (see Lemma 2.18) We assume that the line is universal, so the $\emptyset$-part is empty. Therefore the vertex a separates the graph into two subgraphs, $G_1 = a^+$-part and $G_2 = ab^+$-part $\cup$ b-part.

Similarly for the b-part and the vertex b. □

Chapter 4

Line Structure in Special Graph Classes

In this chapter we consider several classes of graphs, for which we can describe their line structure and even accurately count the number of their lines.

4.1 Graphs with all lines universal

Theorem 4.1. *Each line of a graph G is universal if and only if the graph G is a cycle C_4, or a path P_n.*

Proof. By the Shortest Path lemma, each line in C_4 and in P_n is universal.

To prove that there are no other graphs, in which each line is universal, we will use the Long universal line lemma 3.13. Such graphs must be triangle-free, because a triangle is a distance subgraph and it defines three non-universal lines. If a graph G does not contain triangle, and it is not the trivial case P_1, it contains two vertices a and b with distance $d(a, b) = 2$. The line $\overleftrightarrow{ab}$ is a long universal line, therefore the vertices a and b separate the graph into three parts, with some of the vertices in a-part, some of them in b-part, and some in ab-part. The ab-part is a layer subgraph with three layers. (see Section 2.4)

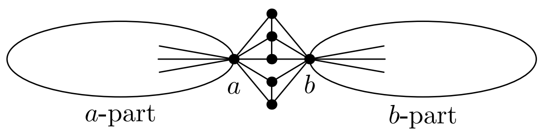

Figure 4.1: Line $\overleftrightarrow{ab}$ separates the graphs into three parts.

Using the same argument for the a-part, if it is not a single edge or a single vertex, there exist two points c and d in the a-part, such that $d(c, d) = 2$.

These vertices separate the a-part into smaller subgraphs. We can continue this process until every outer part of every considered line is an edge or a vertex.

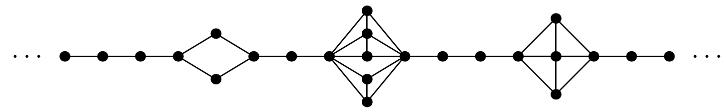

Figure 4.2: The graph after splitting of outer line parts.

Now we need to consider only the inner parts of the lines, such as the ab-part of the first considered line. Recall that for each considered line, the distance of its defining points is 2. Therefore, the inner parts contain exactly three layers. The first and last layers contain only the points defining the line. The middle layer does not contain any ladder edges, because the graph is triangle-free.

Assume that there are two different vertices u and v in the middle layer of a line $\overleftrightarrow{ab}$. Now we can see that if the graph contains any vertex x different from all of a, b, u, v, by the Shortest Path lemma the vertex x cannot belong to the line $\overleftrightarrow{uv}$ (examine the figure).

Therefore, either the graph is a C_4, or each of the considered lines contains only one inner point. In the latter case, the entire graph must be a single path. □

4.2 Graphs with all lines different

We have seen that n points in the euclidean plane in a general position define a different line for each pair of the points. There are graphs with a similar property—all the lines defined by vertices of the graph are pairwise different. In Section 2.13, we have already seen that the complete graph K_n is such a graph. We will show later in Theorem 6.2 that odd cycles C_{2k+1} have all lines different too. We are now going to show several more examples of such graphs.

Unlike points in the plane, there are graphs with all lines different containing a universal line. The following lemma shows one way of constructing such graphs.

Lemma 4.2. *Let all the lines in a graph G be different, and let G not have a universal line. Let the graph G have an articulation a separating it into subgraphs A and B. Construct a graph G' by adding a new vertex b and the edge ab to G. Then G' has a universal line and all the lines in G' are different.*

Proof. The edge ab is a bridge in G', therefore $\overleftrightarrow{ab}$ is a universal line. The original graph G is a distance subgraph of G', therefore every two lines defined only by vertices in G are also different in G'.

First, consider the structure of a line $\overleftrightarrow{bc}$, without loss of generality let $c \in A$. The vertex a is an inner point of the line $\overleftrightarrow{bc}$, therefore, by the Articulation lemma 3.7 no vertex in B lies on $\overleftrightarrow{bc}$.

First, we show that for all three vertices x, y, z of G, the line $\overleftrightarrow{bx}$ is different from the line $\overleftrightarrow{yz}$. For a contradiction, assume that the lines are equal. Without loss of generality, let $x \in A$. Then y and z must belong to A, otherwise they would not lie on the line $\overleftrightarrow{bx}$. The line $\overleftrightarrow{yz}$ contains b, therefore, by the Atrticulation lemma, a is its outer point. However, by the same lemma, $\overleftrightarrow{yz}$ must contain all the vertices in B as well. This makes it clearly different from $\overleftrightarrow{bx}$, which contains no vertex in B.

Now, we will show that for all two vertices x, y of G the lines $\overleftrightarrow{bx}$ and $\overleftrightarrow{by}$ are different. Again, assume that they are equal. Let x belong to A. If $y \in B$, then, as above, $y \notin \overleftrightarrow{bx}$. Therefore, both x and y must belong in the same part, without loss of generality, in A. Let c be a vertex in B. By the Lemma 3.8, the lines $\overleftrightarrow{xc}$ and $\overleftrightarrow{yc}$ contain the same points of B. But we assumed that they are different, therefore they must differ in A. Now observe that the vertex c separates the graph G' into components A and $C = B \cup \{b\}$. Since $b, c \in C$, by the same lemma, lines $\overleftrightarrow{bx}$ and $\overleftrightarrow{cx}$ contain the same vertices of A. Similarly, lines $\overleftrightarrow{by}$ and $\overleftrightarrow{cy}$ contain the same vertices of A. But since $\overleftrightarrow{cx}$ and $\overleftrightarrow{cy}$ differ in A, lines $\overleftrightarrow{bx}$ and $\overleftrightarrow{by}$ must differ there as well. $\square$

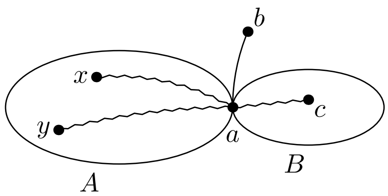

Figure 4.3: A graph with a universal line and all lines different.

We have already seen that graphs with a diameter of one, that is, complete graphs, have all lines different. In the following we will consider graphs with a diameter of two. We show a result connecting the line structure of such a graph with its complement. We will call the complement of a graph with diameter two an *antigraph.* If uv is an edge in the antigraph, we will call it an *antiedge.*

Lemma 4.3 (Antigraph lemma)**.** *Let G be a graph with diameter 2, and let H be its complement. Let $a, b \in V_G$, and let $N_H[v]$ be the neighborhood of the vertex v in H, that is, all vertices u for which $d_H(v, u) \leq 1$. Then for the line $\overleftrightarrow{ab}$ in G the following holds:*

(1) If ab is an antiedge, then $\overleftrightarrow{ab} = V_G \setminus (N_H[a] \cup N_H[b]) \cup \{a, b\}$.
(2) If ab is not an antiedge, then $\overleftrightarrow{ab} = N_H[a] \triangle N_H[b]$.

Proof. Note that since the diameter of G is 2, the distance of any two points of G is either 1 or 2.

(*1*) In the graph G, ab is not an edge, therefore $d(a, b) = 2$. By the definition of a line, the line $\overleftrightarrow{ab}$ contains, except for the vertices a and b, only such vertices c, for which $d(a, c) = d(b, c) = 1$. Those are exactly those vertices for which neither ac nor bc is an antiedge.

(*2*) Now ab is an edge, therefore $\overleftrightarrow{ab}$ contains the vertices a and b, and all vertices c, for which $d(a, c) = 1$ and $d(b, c) = 2$, or vice versa. Such vertices in H must be adjanced to exactly one of the vertices a and b. □

Also observe that if H is not connected, let C be a set of all the vertices c of H disconnected from both a and b. Then if ab is an antiedge, then the line $\overleftrightarrow{ab}$ in G contains all the vertices of C, and if ab is not an antiedge, then the line $\overleftrightarrow{ab}$ in G does not contain any vertex of C.

Observation 4.4. *Let G be a graph with diameter 2 with all the lines different, and let H be its complement. Then for all vertices a and b in the graph H, $N_H[a] \neq N_H[b]$.*

Proof. For a contradiction, assume that there are two vertices $a, b \in H$, such that $N_H[a] = N_H[b]$. Note that ab must be an antiedge, otherwise $a \notin N_H[b]$. Therefore, ab is not an edge in G, and since the diameter of G is 2, there exist a vertex c, such that ac and cb are edges in G. Observe that since ac and bc are not antiedges, and $N_H[a] = N_H[b]$, it must hold in G that $\overleftrightarrow{ac} = \overleftrightarrow{bc}$. This is a contradiction, as G has all its lines different. □

Now let us consider a graph G obtained by removing a cycle C_k from a complete graph K_n. The complement of G is a cycle, and optionally, several isolated vertices. Let us denote the vertices on the cycle $a_1, a_2, \dots a_k$ and the isolated vertices $b_1, b_2, \dots b_{n-k}$. For which values of n and k, all the lines in G are different?

For $k = 3$, the vertices a_1 and a_2 have the same neighborhoods in H, and by the previous observation, there are at least two equal lines in G. For $k = 6$, observe that the line $\overleftrightarrow{a_1a_2}$ is equal to the line $\overleftrightarrow{a_4a_5}$.

We leave it as an excercise to show that for every $k \geq 4$ not equal to 6 all the lines in G are different. Apply the Antigraph lemma 4.3 to determine which points belong to which line. Several examples of different lines for $n = 10, k = 7$ are shown in Figure 4.4.

We can also remove a forest with certain properties from a complete graph, and still obtain a graph with all lines different.

Theorem 4.5 (A forest antigraph)**.** *Let G be a graph with diameter at most 2, and let its complement H be a forest. Graph G has all the lines different, if and only if H does not contain any of the following four configurations:*

(1) One tree of H is an edge K_2.

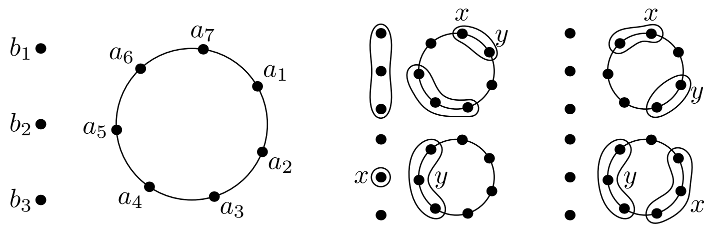

Figure 4.4: The graph H and several different lines of G, as described above.

(2) There exists a path $a_1a_2a_3a_4a_5$, with degrees of vertices $d_H(a_1) = d_H(a_5) = 1$, and $d_H(a_2) = d_H(a_4) = 2$. The degree of a_3 does not matter.

(3) The entire graph H is a single tree, and it contains a path $a_1a_2a_3a_4$ with degrees of vertices $d_H(a_2) = d_H(a_3) = 2$, and all neighbours of a_1 and a_4, except for a_2 and a_3, are leaves.

(4) The entire graph H consists of two stars, $K_{1,k}$ where $k \geq 1$, and $K_{1,\ell}$ where $\ell \geq 0$.

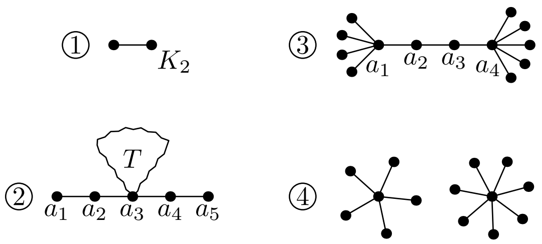

Figure 4.5: Forbidden configurations of the forest H.

Proof. Let G be a graph with a diameter of 2 and with all lines different. Let its complement H be a forest. We will show that if a line $\overleftrightarrow{ab}$ is equal to another line $\overleftrightarrow{cd}$, then the forest H has to contain one of the forbidden configurations.

The proof is a highly technical case analysis.

Throughout the proof, we will always discuss graph properties of the forest H, such as leaves or subtrees. On the contrary, the line $\overleftrightarrow{uv}$ will always denote a line in the graph G. If we assume that uv is an antiedge, by *neighbors* of v we will mean all neighbors of v, except for u. And last, by *distant neighbors* of v we will mean neighbors of neighbors of v, different from v. We hope that the reader will not be confused by this notation too much.

In the figures, empty discs and dashed lines will denote vertices or edges, which can not lie in H for a given condition to hold.

First, let both ab and cd be antiedges. If ab and cd belong to different trees, and one of a or b is not a leaf, then there exists a neighbor of a or b not contained

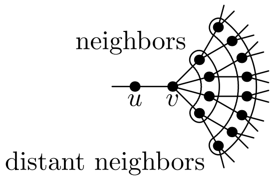

Figure 4.6: An antiedge uv, v's neighbors and distant neighbors.

in $\overleftrightarrow{ab}$. However, because cd is in a different component, the line $\overleftrightarrow{cd}$ must contain this neighbor. If both a and b are leaves, this gives the configuration (*1*).

Now, let both ab and cd belong to the same tree. We will show that at least one of the vertices a or b is a leaf. For a contradiction, assume that both a and b have neighbors. By the Antigraph lemma, these neighbors do not lie on the line $\overleftrightarrow{ab}$. The vertices c and d have to belong to the same subtree separated by ab, because cd is an antiedge. However, at least one of the neighbors of a or b is therefore not contained in a neighborhood of either c or d, as their distance is too large. Therefore, this neighbor belongs to $\overleftrightarrow{cd}$, but not to $\overleftrightarrow{ab}$, which is a contradiction. Similarly, we can prove that either c or d is a leaf. Without loss of generality, let a and d be leaves.

If either b or c is a leaf as well, we reach the configuration (*1*). So let both b and c be of degree at least two. None of b's neighbors is contained in $\overleftrightarrow{ab}$, therefore they can not be contained in $\overleftrightarrow{cd}$, therefore they have to be neighbors of either c or d. Since d is a leaf, they must be neighbors of c. This gives $d(b, c) = 2$, and we can also see that there can be only one such a neighbor, because H is a forest. This gives configuration (*2*), with $a_1 = a$, $a_2 = b$, $a_4 = c$ and $a_5 = d$.

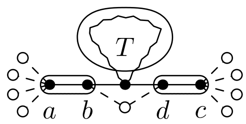

Figure 4.7: Two equal antiedge lines and their points.

Now, let ab be an antiedge and let cd not be an antiedge. Observe that the line $\overleftrightarrow{ab}$ contains a points in each of the trees of H. Similarly, the line $\overleftrightarrow{cd}$ can contain points in at most two different trees. Therefore, if $\overleftrightarrow{ab}$ and $\overleftrightarrow{cd}$ are to be equal, H has to contain at most two trees. Let us denote the one containing ab as A, and the other one (if it exists) B.

Assume now that c and d belong to two different trees. Without loss of generality, let c belong to A. The line $\overleftrightarrow{ab}$ contains all the points in B, therefore $\overleftrightarrow{cd}$ must contain them all too. The vertex c is not in B, so all vertices in B

must be in $N_H[d]$. This gives that B is a star $K_{1,\ell}$, where $\ell \geq 0$. Now note that vertices of A contained in $\overleftrightarrow{ab}$ must be the same as vertices in $N_H[c]$. This is possible only if A is a star $K_{1,k}$, where $k \geq 1$, where c is different from the central vertex of the star. This gives configuration (*4*).

Now assume that all of a, b, c, d belong to the same tree. Also observe that points lying on the line $\overleftrightarrow{cd}$ induce at most two connected components. If both a and b have a distant neighbor, note that the line $\overleftrightarrow{ab}$ contains at least three components—one induced by a and b, at least one induced by a's distant neighbors, and at least one induced by b's distant neighbors. These components are separated by the neighbors of a and b. Therefore, $\overleftrightarrow{ab}$ and $\overleftrightarrow{cd}$ can not be equal. If none of a and b has distant neighbors, the tree is too small to place c and d so that cd would not be an antiedge and $\overleftrightarrow{cd}$ would be equal to $\overleftrightarrow{ab}$.

This means that exactly one of the vertices a and b can have distant neighbors. Without loss of generality, let it be b. Therefore, all the neighbors of a are leaves.

Note that in order to belong to $\overleftrightarrow{cd}$, the vertex a must be in $N_H[c]$ or $N_H[d]$. Without loss of generality, let it belong to $N_H[c]$. However, c cannot be a neighbor of a, because these neighbors do not lie on $\overleftrightarrow{ab}$. If a has any neighbors, c can not be the same vertex as a, as $\overleftrightarrow{cd}$ would contain neighbors of a. Therefore c is equal to b. Now, in order not to lie on $\overleftrightarrow{cd}$, neighbors of b must be also neighbors of d. This means that there is at most one such neighbor. The vertex d can have neighbors, which belong to both lines. However, it cannot have distant neighbors, as those would belong to $\overleftrightarrow{ab}$ and not to $\overleftrightarrow{cd}$. Also, b cannot have distant neighbors except for d, as those would lie on $\overleftrightarrow{ab}$ and not on $\overleftrightarrow{cd}$. Therefore, we obtain the configuration (*3*).

If a does not have neighbors, c can be the same vertex as a. If b does not have distant neighbors, there is nowhere in the tree where d could be placed. Such distant neighbors of b must also be neighbors of d, then, similar to the previous paragraph, there is at most one such neighbor, and d can not have distant neighbors. Again, we obtain the configuration (*3*).

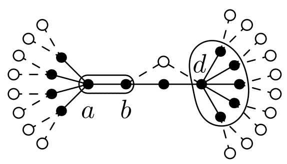

Figure 4.8: An antiedge line and an equal non-antiedge line.

Last, let neither of ab and cd be an antiedge. Let a belong to a tree A and let b belong to a different tree B. Note that neighborhoods of c and d have to cover the same vertices as neighborhoods of a and b, therefore c and d have to belong to the same two trees. Without loss of generality, let $c \in A$ and $d \in B$. Now, since A and B are not connected, in order for the lines to be equal, it must hold that $N_H[a] = N_H[c]$ and $N_H[b] = N_H[d]$. We know that either $a \neq c$ or $b \neq d$, so let $a \neq c$. But in a tree two vertices $a \neq c$ can have the same neighborhoods only if the tree is an edge, giving configuration (*1*).

The last situation says that both a and b belong to the same tree. In order for a to lie on $\overleftrightarrow{cd}$, it must belong to exactly one of the neighborhoods of c or d. Let a lie in $N_H[c]$. The vertex b cannot belong to $N_H[c]$ as well, because then c would not lie on $\overleftrightarrow{ab}$. Therefore b must belong to $N_H[d]$. Again, either $a \neq c$ or $b \neq d$. Let $a \neq c$.

If neither a nor c are leaves, the neighborhoods of b and c can not simultaneously cover a's and c's neighbors, so the lines differ. So let a be a leaf. If c is a leaf as well, we obtain configuration (*1*). Otherwise, for the lines to be equal, the neighbors of c must be covered by either d or b, but not both. As before, there can be only one such neighbor, let it be a vertex x. Also, b and d must be different vertices. Similarly, one of them has to be a leaf, and the other one is connected to x. We obtain configuration (*2*).

If you do not understand the proof, go through it again and draw the situations on a piece of paper. $\square$

4.3 Special graph classes

To conclude this chapter, we will show several results about the structure of lines in some specific graph classes.

First, consider cycles. By an application of Theorem 6.2 for simple graphs, for even cycles C_{2k} we get universal lines defined by each edge and each pair of opposite points. Any other two lines are different. This gives $\binom{2k}{2} - 3k + 1$ different lines. Similarly, for odd cycles C_{2k+1} we get $\binom{2k+1}{2}$ different lines, since each line is different.

Now, consider trees. Observe that the number of lines in a general tree can vary greatly. We have already shown that a path P_n defines only one line, and on the other hand a star $K_{1,n-1}$ defines $\binom{n-1}{2} + 1$ different lines.

The following observation tells us more about the structure of a line in a tree.

Observation 4.6. *Let T be a tree, and let a and b be its vertices. The line $\overleftrightarrow{ab}$ contains the shortest path between a and b, and subtrees separated by a and b (see Figure 4.9 below).*

Proof. Recall the Articulation lemma 3.7. Each inner vertex of a tree is an articulation. The inner points on the path ab are inner points of the line $\overleftrightarrow{ab}$, therefore by the lemma $\overleftrightarrow{ab}$ does not contain subtrees separated by these points

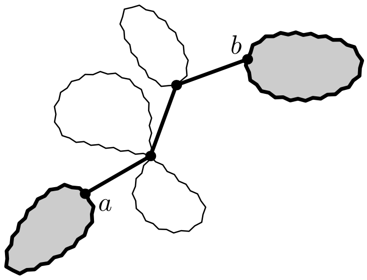

Figure 4.9: Points on a line in a tree.

(white in the figure). The points a and b are outer points of $\overleftrightarrow{ab}$, so the line $\overleftrightarrow{ab}$ contains the subtrees separated by a and b (grey in the figure). $\square$

Observation 4.7. *Let T be a tree and let $\overleftrightarrow{ab}$ be a line in T. Then:*

- *If c is a outer point of $\overleftrightarrow{ab}$, then $\overleftrightarrow{ab}$ contains all the neighbors of c.*
- *If d is a proper inner point of $\overleftrightarrow{ab}$, then $\overleftrightarrow{ab}$ contains exactly two neighbors of d.*

Proof. Obvious, observe the figure 4.9 above. $\square$

If a tree contains a vertex c of degree two, let us denote its neighbors a and b. Observe that c is a subdivision of the bridge ab. By the Bridge pumping theorem 3.2, contracting the edge ac does not change the structure of lines, nor the number of different lines. We can repeat this operation while there is a vertex of degree two in the tree. If there is no such vertex, we will call this tree a *normalized tree.*

Lemma 4.8. *Let T be a normalized tree. Then any two long lines $\overleftrightarrow{ab}$ and $\overleftrightarrow{cd}$ are different.*

Proof. The tree is normalized, therefore every vertex, which is not a leaf, has a degree of at least 3. Now, consider a subgraph S of T induced by the points of the line $\overleftrightarrow{ab}$. By the previous observation, a vertex in S has a degree of two if and only if it belongs to the ab^+-part of $\overleftrightarrow{ab}$. The line $\overleftrightarrow{ab}$ is a long line, hence S contains at least one such vertex. The ab^+-part is connected, therefore these vertices form a path P in S. The ends of this path are of degree 2, so they have a neighbor in S. By Lemma 2.18, there is no edge between ab^+-part and the other proper parts, therefore these neighbors must be a and b.

Consequently, we can distinguish the vertices a and b in a tree just by knowing the set of the points of the line $\overleftrightarrow{ab}$.

Note: this does not hold in general when the graph is not a tree.

If the paths $[ab]$ and $[cd]$ were equal, they would define the same set of points. Therefore, the vertices a and b have to be equal to vertices c and d. $\square$

Corollary 4.9. *Let T be a tree with n vertices, out of which k vertices have a degree of 2. Then T contains $\binom{n-k-1}{2}+1$ different lines.*

Proof. First, we normalize T to obtain a normalized tree T'. Note that the normalization operation does not change the degrees of the vertices, therefore after normalizing k initial vertices of degree 2, we obtain T' with $n-k$ vertices.

Each edge in a tree is a bridge, thus every edge line is universal. By the previous lemma, any pair of two nonadjancent vertices defines a unique, non-universal line. This gives $\binom{n-k}{2}-(n-k-1)=\binom{n-k-1}{2}$ long lines. Together with the universal line we get $\binom{n-k-1}{2}+1$ different lines. □

Last, we describe the line structure of complete k-partite graphs.

Observation 4.10. *Let G be a complete k-partite graph with parts $P_1, P_2, \ldots, P_k$ of sizes $a_1, a_2, \ldots, a_k$, where for all $i : a_i \geq 3$. Then G contains*

$$\binom{k}{2}+\sum_{i=1}^{k}\binom{P_i}{2}$$

different lines.

Proof. The graph G contains two distinct types of lines. First, consider two vertices a and b, $a \in P_i$, $b \in P_j$. We see that:

$$\begin{aligned}
&d(a,b)=1\\
\forall x \in P_i, x \neq a: \quad & d(a,x)=2, \quad d(b,x)=1, \quad x \in \overleftrightarrow{ab}\\
\forall y \in P_j, x \neq b: \quad & d(a,y)=1, \quad d(b,y)=2, \quad y \in \overleftrightarrow{ab}\\
\forall z \in P_k, k \neq i,j: \quad & d(a,z)=1, \quad d(b,z)=1, \quad z \notin \overleftrightarrow{ab}
\end{aligned}$$

Hence, the line $\overleftrightarrow{ab}$ contains all the vertices in $P_i \cup P_j$. By choosing i and j we obtain $\binom{k}{2}$ pairwise different lines.

Now let $c, d \in P_i$. Observe that:

$$\begin{aligned}
&d(c,d)=2\\
\forall x \in P_i, x \neq c,d: \quad & d(c,x)=2, \quad d(d,x)=2, \quad x \notin \overleftrightarrow{cd}\\
\forall y \in P_j, j \neq i: \quad & d(c,y)=1, \quad d(d,y)=1, \quad y \in \overleftrightarrow{cd}
\end{aligned}$$

Therefore, the line $\overleftrightarrow{cd}$ contains all the vertices in parts different from P_i, and also the vertices c and d. Again, every two such lines are different, and also different from any lines obtained in the previous paragraph. By choosing the vertices c and d, for every part P_i we get $\binom{a_i}{2}$ different lines. In total, we get

$$\binom{k}{2}+\sum_{i=1}^{k}\binom{a_i}{2}$$

different lines in G, as stated by the observation. □

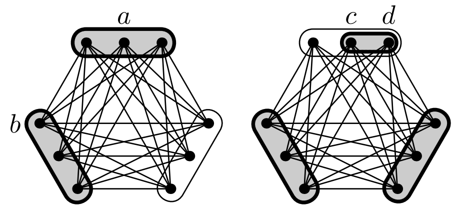

Figure 4.10: Lines in a k-partite graph.

For example, for the complete bipartite graph $K_{m,n}$ we obtain $1 + \binom{m}{2} + \binom{n}{2}$ different lines.

Note that if every $a_i \leq 3$, the graph does not contain a universal line. We will try to estimate a lower bound for the number of different lines. It can be seen that for a given n and k, we get a minimal number of lines if all a_i are equal. Let the graph contain k parts, each with $\frac{n}{k}$ vertices. Therefore, the number of lines is approximately $\frac{k}{2}^2 + k \cdot \frac{(k/n)}{2}^2 = \frac{1}{2}(k^2 + \frac{n}{k}^2)$. In order to minimize this number, let $k = n^{2/3}$, giving approximately $n^{4/3}$ lines.

Moreover, we conjecture that the minimal number of lines in any graph without a universal line is very close to $n^{4/3}$.

Chapter 5

Graph classes with k different lines

In this chapter we consider the following question: for a given integer k, which graphs contain exactly k different lines?

Definition 5.1. We say that a graph G is k-linear, if the vertices of G define exactly k pairwise different lines.

5.1 Graphs with small number of lines

Let us consider the cases where k is a small integer.

Proposition 5.2.

- *The only* 1*-linear graphs are the cycle* C_4 *and any path* P_n*.*
- *There is no* 2*-linear graph.*
- *The only* 3*-linear graph is the triangle* K_3*.*

Proof. Note that if a graph has a non-universal line $\overleftrightarrow{ab}$, there exist a vertex c, such that $c \notin \overleftrightarrow{ab}$. By the Triangle lemma 2.7, then both $a \notin \overleftrightarrow{bc}$ and $b \notin \overleftrightarrow{ac}$, giving at least three different non-universal lines. Therefore, if a graph is 1-linear, the line must be a universal line. We have already shown in Theorem 4.1, that these graphs are C_4 and any P_n.

If a graph is 2-linear, it contains at least one non-universal line. Then it has to contain at least three different lines, which is a contradiction.

Now consider 3-linear graphs. If such a graph contains a universal line, similarly to above, we find at least three different non-universal lines, four different lines in total. Therefore, a 3-linear graph can not contain a universal line. The only graph with at most 3 vertices with three different lines is K_3. Now consider a graph G with at least 4 vertices. Any edge in G is not universal, therefore, by Lemma 2.14, G contains an odd cycle. Observe that the shortest odd cycle in G is a distance subgraph: if there was a shorter path between two vertices on the cycle, it would induce two shorter cycles, at least one of which has to be odd. If the length of the shortest odd cycle is at least 5, by Theorem 6.2 the cycle contains at least $\binom{5}{2} = 10$ different lines, and by Lemma 2.10, since the cycle is a distance subgraph, these lines differ in G as well.

We now see that G has to contain a triangle abc. The three lines $\overleftrightarrow{ab}$, $\overleftrightarrow{bc}$, and $\overleftrightarrow{ca}$ are pairwise different. We assumed that G has more than three vertices, therefore at least one of the vertices of the triangle has to have a neighbor. Without loss of generality let a have a neighbor d. Now consider the distance subgraph induced by vertices a, b, c, d. Observe that

- If neither db nor dc are edges, the line $\overleftrightarrow{da}$ contains all of a, b, c, therefore it defines a fourth line.
- If db is an edge and dc is not, the line $\overleftrightarrow{dc}$ contains all of a, b, c, therefore it defines a fourth line. Similarly if dc is an edge and db is not.
- If both db and dc are edges, the four vertices induce a complete subgraph K_4, which defines $\binom{4}{2} = 6$ different lines.

In each of the cases, the vertices define at least 4 different lines. $\square$

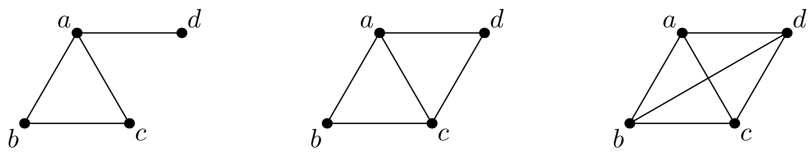

Figure 5.1: Three possible distance subgraphs induced by vertices a, b, c, d.

5.2 Finite and infinite classes of k-linear graphs

Observe that the set of all k-linear graphs can be empty (for $k = 2$), non-empty but finite (for $k = 3$), or infinite (for $k = 1$). For a given k, we ask which of the cases holds.

Observation 5.3. *There are infinitely many integers k, such that the set of all k-linear graphs is infinite.*

Proof. For a given integer $\ell \geq 3$, consider the star $K_{1,\ell}$. Let c be its center vertex, and let $a_1, a_2, \ldots a_\ell$ be the endpoints. Observe that each of the lines $\overleftrightarrow{a_i c}$ is universal, and each line $\overleftrightarrow{a_i a_j} = \{a_i b a_j\}$. Every two such lines are therefore different, and by choosing i and j we obtain $\binom{\ell}{2} + 1$ different lines. Now, by the

Bridge pumping theorem 3.2, we can subdivide any of the edges and we obtain a different graph with the same number of lines. We can repeat the subdivision to obtain any number of different graphs.

Therefore, for every integer $\ell \geq 3$, if we choose $k = \binom{\ell}{2} + 1$, the set of all k-linear graphs is infinite. $\square$

We have discovered many more formulas in the form $k = a\ell^2 + b\ell + c$, such that for every large enough ℓ, the set of all k-linear graphs is infinite. However, as of the time of writing, we do not know any k other than 2 or 3, for which the set of k-linear graphs is empty or finite. We conjecture that there are no 5-linear graphs, and the only 6-linear graphs are those pictured in Figure 5.2.

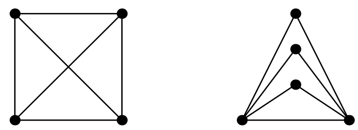

Figure 5.2: Two 6-linear graphs.

Now, for certain integers k, we can obtain an infinite class of k-linear graphs. However, every graph in the class we have shown contains a universal line. If there were infinitely many k-linear graphs without a universal line, this class would necessarily contain a graph with $k + 1$ vertices, which would be a counterexample against the Klee-Wagon conjecture 1.4. We will soon show that this is not possible.

In this and the following proof we will use one result of the Ramsey theory, which states:

Theorem 5.4 (Ramsey)**.** *Given any two integers k and ℓ, there exists an integer n, such that the following holds:*

Let G be a complete graph with at least n vertices. Let its edges be colored by ℓ different colors, that is, let there be a function $c : E \to \{1..\ell\}$. Then G contains a complete subgraph with at least k vertices, such that all edges of the subgraph are of the same color, that is, there is a set $W \subseteq V$, such that $|W| \geq k$, and $\forall u, v, x, y \in W, u \neq v, x \neq y : c(uv) = c(xy)$.

We will denote such n by $R(k, \ell)$.

Theorem 5.5. *For any given k, consider the set of all graphs G, such that G is k-linear and does not have a universal line. This set must be finite.*

Proof. We are going to show that there exists an integer n, such that every graph with at least n vertices contains either a universal line, or more than k different lines. There are only finitely many graphs with less than n vertices, so the theorem holds. This means that every "large enough" graph has either a universal line, or "enough" different lines.

Let $n = R(k, k)$, and let G be a graf with at least n veritces. For a contradiction, assume that G contains at most k different lines, none of which is universal.

Now, number the different lines in G by the numbers $1, 2, \ldots, k$. Construct a complete graph L with $V_L = V_G$. Let $c(uv)$ be the number of the line $\overleftrightarrow{uv}$ in G. By Ramsey's theorem 5.4, the graph L contains a complete subgraph with at least k vertices. Let us denote these vertices $v_1, v_2, \ldots v_k$. Note that for each pair of vertices, the edges between the vertices in L have the same color, therefore they define the same line in G. Let us denote this line $\boldsymbol{\mathcal{L}}$.

We assumed that the graph G does not have a universal line, hence there is a vertex $w \notin \boldsymbol{\mathcal{L}}$. Now let $1 \leq i < j \leq k$. The line $\overleftrightarrow{v_i v_j}$ is $\boldsymbol{\mathcal{L}}$, therefore it does not contain w. By the Triangle lemma 2.7, it holds that also $v_i \notin \overleftrightarrow{v_j w}$ and $v_j \notin \overleftrightarrow{v_i w}$. Therefore, every two lines $\overleftrightarrow{v_i w}$ and $\overleftrightarrow{v_j w}$ are different, and also different from $\boldsymbol{\mathcal{L}}$. However, this gives k different lines $\overleftrightarrow{v_i w}$, and the line $\boldsymbol{\mathcal{L}}$ different from each of them. In total, we have at least $k+1$ different lines. This is a contradiction. □

Finally, observe that we have generated an infinite set of k-linear graphs by finding one such graph containing a bridge and subdividing the bridge to obtain an arbitrary number of k-linear graphs. We are going to show that this is the only way we can obtain such an infinite set. Similarly to Lemma 4.8, we can contract an edge incident with an articulation of degree 2. This does not change the number of lines. As in Lemma 4.8, we obtain a normalized graph, which has no articulation with a degree of 2. We disallow bridge pumping by considering only normalized graphs.

Theorem 5.6. *For a given integer k, the class of all k-linear normalized graphs is finite.*

Similarly as above, we will show that for a given k, there exists an integer n, such that a normalized graph with at least n vertices contains more than k lines. We will prove this by a series of observations, each giving us a more specific structure which must be contained in the graph. To give a bound on n,

- let k_K be the smallest integer such that $\binom{k_K}{2} > k$,
- let k_C be the smallest even integer such that $\binom{k_C}{2} - \frac{3}{2}k_C + 1 > k$,
- and let k_T be the smallest integer such that $\binom{k_T - 1}{2} + 1 > k$.

Note that this is chosen so that the graphs K_{k_K}, C_{k_C} and a tree with k_T vertices with a degree different from 2 contain more than k different lines. Each of the calculations is given in Section 4.3.

Lemma 5.7. *For a given integer k_1, there exists an integer n, such that every normalized graph with at least n vertices and at most k different lines contains a 2-edge-connected subgraph with at least k_1 vertices.*

Proof. Let G be such a graph. Observe the structure of the graph. It contains several 2-edge-connected components separated by bridges and *isolated vertices*. Note that since G is normalized, the degrees of isolated vertices are either 1 or at least 3. Consult the figure 5.3.

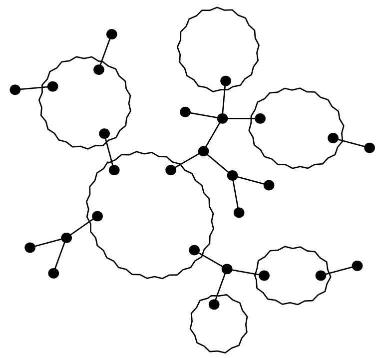

Figure 5.3: Structure of a normalized graph.

First, let us bound the number of isolated vertices. Consider a tree T obtained from the graph G by contracting all the 2-edge-connected components. Observe that the relation of betweenness among isolated vertices in G is preserved in T. Therefore, all lines different in T are also different in G. All isolated vertices in G have in T diameter different from 2. Therefore, by Corollary 4.9, if G contains at least k_T isolated vertices, they define more than k different lines.

Now, we will bound the number of 2-edge-connected components. If G contains at least two such components, for each component C we can find an edge ab going into C, that is $a \notin C$ and $b \in C$. Since C is 2-connected, b must have at least two different neighbors. Let us denote them c and d. Now consider the line $\overleftrightarrow{ac}$. We can see that regardless of whether cd is an edge, the vertex d does not lie on $\overleftrightarrow{ac}$. Also, observe the shortest paths between vertices a and c, and vertices in a component other than C. Observe that for every component, the line $\overleftrightarrow{ac}$ either contains all the vertices of the component, or it contains none of them. The only component which contains a vertex lying on $\overleftrightarrow{ab}$ and a vertex not on $\overleftrightarrow{ab}$ is C. Therefore, each component defines a line, and those lines are different for two different components. If G contains more than k components, they define more than k different lines.

Finally, let $n \geq k_T + k \cdot k_1$. We have shown that G contains at most k_T isolated vertices, therefore there are at least $k \cdot k_1$ vertices in the 2-edge-connected components. Also, there are at most k components, therefore, by the pidgeonhole principle, at least one of them contains k_1 vertices. □

Lemma 5.8. *For every two integers k and k_2, there exists an integer k_1, such* *that every graph G with at least k_1 vertices and at most k different lines contains two vertices with a distance more than k_2.*

Proof. Let $k_1 = R(k_K, k_2)$, and let G be a graph with at least k_1 vertices. For a contradiction, assume that the distance of every two vertices in G is at most k_2. Constuct a complete graph L, such that $V_L = V_G$. For every two vertices u and v, let $c(u, v) = d(u, v)$. By Ramsey's theorem 5.4, there is a complete graph in L with all edges of the same color. This means that there are k_K vertices in G, such that the distance between every two of them is equal. Now, similarly as for complete graphs, note that no line defined by two of those vertices can contain a third one. Therefore, every pair defines a different line. Since there are at least k_K vertices, they define more than k different lines, which is a contradiction. $\square$

Lemma 5.9. *For every two integers k and k_3, there exists an integer k_2, such that every graph G with two vertices with a distance of at least k_2 and at most k different lines contains two vertices u and v, such that the shortest path between u and v has at least k_3 vertices, and the degree of every inner vertex of the path is 2.*

Proof. Let u' and v' be vertices of G, such that $d(u', v') = k_2$. Consider a shortest path P between u' and v'. Let us denote all the vertices on this path with degree greater than two *branched.* We will bound the number of branched vertices.

Consider the graph induced by P and all neighbors of the branched vertices. Since the path is shortest, no vertex not on the path can be adjanced to two vertices with a distance of three or more. Observe that every branched vertex must be in one of the following configurations:

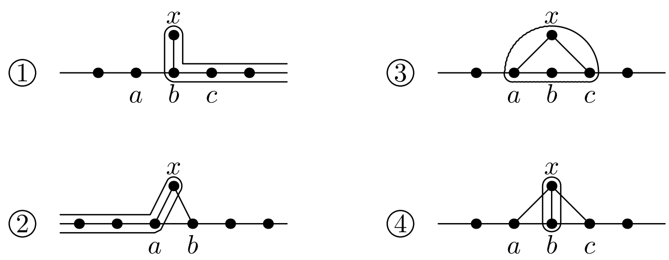

Figure 5.4: Possible configurations of branched vertices. We will explain the highlighted lines in the proof.

We will show an algorithm which gives us several different "important" lines. Throughout the algorithm we will mark all of the branched vertices. We start with all branched vertices unmarked, and proceed in four steps.

(1) While there is an unmarked vertex in the configuration ④, select the line $\overleftrightarrow{xb}$ as important, and mark the vertices a, b, and c. Observe, for example by considering the shortest paths, that out of the vertices of P, the line $\overleftrightarrow{xb}$ contains only the vertex b. We are always selecting a different vertex, therefore every such line is different. If there is no unmarked vertex in the configuration ④, proceed to step two.
(2) While there is an unmarked vertex in the configuration ③, select the line $\overleftrightarrow{xb}$ as important, and mark the vertices a, b, and c. Again, observe that the line $\overleftrightarrow{xb}$ contains from P only the vertices a, b and c. Therefore, every such line is different, and they are also different from lines obtained in step one. If there is no unmarked vertex in the configuration ③, proceed to step three.
(3) While there is an unmarked vertex in the configuration ②, select the line $\overleftrightarrow{xa}$ as important, and mark the vertices a and b. Observe that the line contains all the vertices of the path up to a. Again, two such lines are different from each other and from all the lines obtained in previous steps. If there is no unmarked vertex in the configuration ②, proceed to step three.
(4) Last, while there is an unmarked vertex in the configuration ①, select the line $\overleftrightarrow{xc}$ as important, and mark the vertex b. This line contains all the vertices of P, starting from b. All the selected lines are different from each other and from all previously selected lines.

Observe that at the end of the algorithm, all branched vertices are marked. Now, we have assumed that the graph contains at most k different lines. Therefore, we have executed a step of this algorithm at most k times. Since in each step we mark at most three vertices, P contains at most $3k$ branched vertices.

Also note that the vertices marked in each step form a continuous subpath of P, therefore they divide P into at most $k+1$ continuous subpaths consisting only of unbranched vertices. Now, let $d(u', v') > 3k + (k+1) \cdot k_3$. At most $3k$ vertices are branched, so there are at least $(k+1) \cdot k_3$ unbranched vertices in at most $k+1$ continuous subpaths. Consequently, by the pidgeonhole principle, there is a subpath with at least k_3 unbranching vertices. Recall that all unbranched vertices have a degree of 2, so this subpath is the path the lemma asks for. $\square$

Proof of Theorem 5.6. For a contradiction, assume that for any n there is a normalized graph G with n vertices and at most k lines.

The first lemma states that we can choose n large enough for G to definitely contain a 2-edge-connected subgraph G' with at least k_1 vertices. The second lemma states that we can choose k_1 large enough for G' to contain two vertices with a distance at least k_2. The third lemma states that we can choose k_2 large enough for G' to contain a path P between vertices u and v with every inner vertex of degree two, such that $d(u, v) > \frac{k_C}{2}$. Now, note that G' is 2-edge-connected, therefore there is at least one other path uv, which is edge-disjoint with P. Let Q be the shortest such path. Since the degree of all the inner vertices of P is 2, the paths must be vertex-disjoint as well.

Finally, observe that the cycle formed by the union of P and Q is a distance subgraph of G': if there was a shorter path between two vertices of the cycle, the chosen paths would not be shortest. This circle contains at least k_C vertices, therefore it defines more than k different lines. Consequently, by the distance subgraph lemma, the graph G', and also G, must contain more than k different lines as well. $\square$

Chapter 6

Weighted graphs

In this chapter we will briefly consider weighted graph. We will assign a real positive length to each edge of the graph and consider the length of a path be the sum of lengths of its edges.

Observation 6.1. *Every finite discrete metric space (X, ρ) is equivalent to a weighted graph.*

Proof. Consider a complete weighted graph G, let $V_G = X$. For every two vertices x and y, let $d(x, y) = \rho(x, y)$ in the metric space. □

Theorem 6.2 (Weighted cycles)**.** *Let a graph G be a weighted cycle C_n. Then exactly one of the following holds:*

(1) The graph G contains a universal line.
(2) All the $\binom{n}{2}$ lines in G are different.

Proof. We will first show that at least one of the statements is true.

We will prove the theorem by a geometric analogy. Let us map the vertices of G on a circle, so that the angle between two consecutive vertices is proportional to the length of the edge connecting them in G. See Figure 6.1.

The distance of two vertices a and b in G is equal to their distance on the circle in this mapping. Observe that vertices c, such that $[acb]$, are mapped to points on the shorter arc ac in the circle. Let a' be a point opposite to a on the circle, similarly define b'. Observe that vertices c, such that $[cab]$, are mapped to points on the shorter arc $b'a$. Similarly, vertices c with $[abc]$ are mapped to the arc ba'. Therefore, the line $\overleftrightarrow{ab}$ contains vertices mapped to points on the longer arc $a'b'$. We will call this arc *line arc* of $\overleftrightarrow{ab}$, and we will call the shorter arc $a'b'$ *outer arc* of the line. Again, see Figure 6.1.

If the point a lies opposite to the point b, for all points c on the circle holds $[acb]$, therefore the line $\overleftrightarrow{ab}$ is universal. Similarly, if there is a vertex mapped to either of the points a' or b', it defines a universal line.

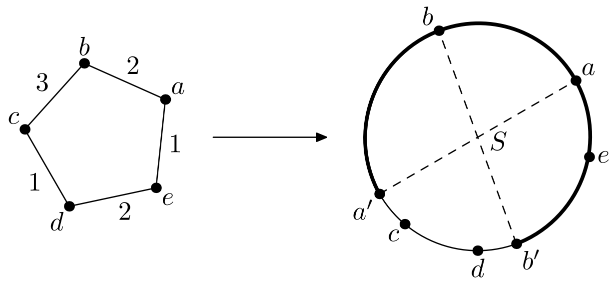

Figure 6.1: Mapping of the points of the cycle to a circle.

Now let $\overleftrightarrow{ab}$ be an edge line. If the outer arc of a line $\overleftrightarrow{ab}$ contains no mapped vertices, the line is universal. If this arc contains at least two points c and d, observe that the outer arc of $\overleftrightarrow{cd}$ is a subset of the shorter arc ab, therefore it is empty, and $\overleftrightarrow{cd}$ is universal. Consequently, if there is no universal line, the outer arc of each edge line contains exactly one point.

Observe that the outer arcs are disjoint, and they cover the entire circle. Therefore, each point is contained in exactly one outer arc. We can define a bijection between edge lines and points on the circle, as shown in Figure 6.2.

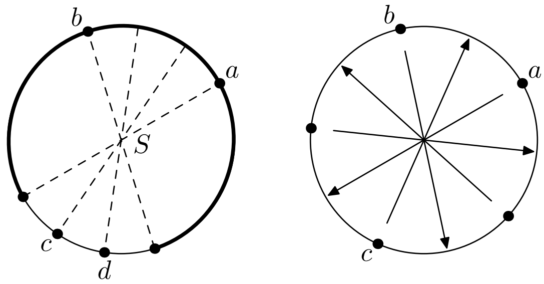

Figure 6.2: A bijection between edge lines and points on the circle.

Now consider a long line $\overleftrightarrow{ab}$. Note that the outer arc of $\overleftrightarrow{ab}$ is an union of outer arcs of edge lines defined by edges on the shortest path ab. Similarly, the line arc of $\overleftrightarrow{ab}$ is an intersection of line arcs of these edge lines. Now observe the outer arcs of two lines ab and cd. We can see that since there is a point in each of the line arcs defined by edge lines, two outer arcs of the lines must contain different points. Therefore, the lines must be different. It does not matter which two lines we choose, therefore every two lines in G are different.

Now we show that both statements cannot be true at the same time. Let $\overleftrightarrow{ab}$ be an universal line. We are going to show that G contains at least two equal lines. As above, we can see that either a and b are opposite points, or the outer arc of $\overleftrightarrow{ab}$ is empty.

If a and b are opposite, consider a point c closest to a and a point d closest to b. Without loss of generality, let $d(a, c) \leq d(b, d)$. Consider the line $\overleftrightarrow{ac}$. Its outer arc is the shorter arc bc', and since d was chosen as the closest point to b, this arc is empty. Consequently, the line $\overleftrightarrow{ac}$ is universal and equal to $\overleftrightarrow{ab}$.

Now, let there be no pair of opposite points, and let there be no points in the outer arc of an edge line $\overleftrightarrow{ab}$. Let c be a neighbor of b different from a. Consider the outer arcs of $\overleftrightarrow{ac}$ and of $\overleftrightarrow{bc}$. Since the outer arc of $\overleftrightarrow{ab}$ is empty, these two arcs contain precisely points in the outer arc of $\overleftrightarrow{bc}$. Therefore, the lines $\overleftrightarrow{ac}$ and $\overleftrightarrow{bc}$ are equal. □

Corollary 6.3. *In a simple cycle C_k, there are $\binom{k}{2}$ different lines if k is odd, and $\binom{k}{2} - k - \frac{k}{2} + 1$ lines if k is even.*

Proof. Perform the mapping of vertices to a circle again. In the first case, the outer arc of each edge line contains exactly one point, therefore all the lines are different. In the second case, all the lines are different, except for k edge lines and $\frac{k}{2}$ lines defined by opposite points, which are all universal. □

Observe that some of the results introduced in previous chapters hold for weighted graphs as well. In several cases we need to construct an *normalized weighted graph* by removing all the edges which do not influence the metric. In particular, we remove edges ab, for which the length of the edge ab is greater than or equal to the length of the shortest path between a and b not containing this edge. Note that this does not change any distances in the graph.

We can see that results we depending only on the relation of betweenness hold here as well. All the results about line structure and distinguishing of lines hold. Results about bridges, articulations and cuts are independent on the weights. On the other hand, properties of edge lines are lost.

Chapter 7

Equilinearity and Graph Reconstruction

We can consider graphs and their lines in the opposite way. Is it possible to find a graph with a given set of lines? Are there some non-isomorphic graphs that have exactly the same lines? We would like to answer some of these questions in this section.

We will use the following notation. We write $\mathrm{ls}(G)$ for the set of different lines and $\mathrm{LS}(G)$ for the multiset of all $\binom{n}{2}$ lines. Also, sometimes is useful to know which points define each line. The *tagged line* $\overleftrightarrow{ab}$ is a pair $(\{a,b\}, \overleftrightarrow{ab})$ for $a \neq b$. We will write $\mathrm{LS}_t(G)$ for the multiset of all the tagged lines.

We can note one general thing about reconstruction algorithms working only for certain graph classes. We can easily defect if such an algorithm obtains a set of lines which does not come from the specified graph class. In this case, the algorithm either fails or outputs a graph. If it does not fail, we can take the constructed graph, calculate its lines and compare them with the input. If this does not yield the set of lines we started with, the algorithm has made an error, which means some of the assumptions about the graph did not hold.

7.1 Equilinearity

How good description of a graph G is its $\mathrm{ls}_t(G)$ or $\mathrm{LS}_t(G)$? Some graphs are uniquely determined, for example, the complete graph is the only graph where every line contains only two vertices. However, for some ls_t or LS_t there are multiple non-isomorfic graphs with the given lines. We will call such graphs *e*quilinear.

Definition 7.1. Let G and H be two graphs, such that $|V(G)| = |V(H)|$ and $G \ncong H$. Graphs G and H are *weakly equilinear*, if $\mathrm{ls}(G) = \mathrm{ls}(H)$, and *strongly equilinear*, if $\mathrm{LS}(G) = \mathrm{LS}(H)$.

Obviously, strong equilinearity is a stronger property. When two graphs are strongly equilinear, they are also weakly equilinear. The smallest example of a pair of strongly equilinear graphs is C_4 and P_3, in both cases all the lines are universal.

Moreover, there exists an infinite number of such pairs. Consider the following construction of pairs G_n and H_n, graphs with $n+2$ vertices $a, b, c_1, \ldots, c_n$. Both graph contain a complete subgraph induced by the vertices $c_1, \ldots, c_n$. In G_n, ab is an edge and a is adjacent to all the vertices c_i. In H_n, ab is not an edge, and both a and b are adjacent to all the vertices c_i. Observe that the two graphs are strongly equilinear.

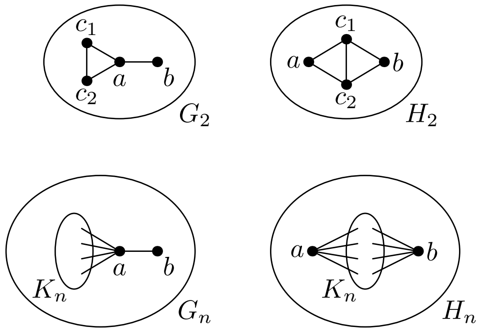

Figure 7.1: The construction for $n = 2$ and for a general n.

In the following sections, we consider the following problem: given one of the representations of lines of a graph, are we able to construct a graph with the same lines? Note that we are always able to examine lines in all possible graphs with the required number of points and compare them to the given description. Therefore, we are only interested in polynomial-time algorithms.

7.2 Reconstruction of trees

As shown in Section 4.3, the line structure of trees is highly regular.

Theorem 7.2 (Reconstruction of a tree)**.** *Let T be a tree. We are able to construct a tree isomorphic to T, given only* $\mathrm{ls}(T)$.

Proof. We will show an algorithm to do this.

Let us denote vertices with a degree of at least three as *large*, and all the other vertices as *small*. By a *non-branching path* we mean a path with a degree of every inner vertex equal to two. Recall Observation 4.6 about trees.

The algorithm works in three steps. First, we identify all the large vertices of the tree. Second, we find out all the edges between large vertices. Last, we connect all the small vertices.

(*1*) Consider the following asymmetric relation R defined on the vertices of the tree. Let a and b be vertices, then $(a, b) \in R$ if and only if every line containing a contains also b.

Let us observe basic properties of this relation. Let a be a large vertex, we show that for no vertex x is $(a, x) \in R$. Consider three neighbors of a, let us denote them b, c and d. By the tree observation, we obtain that every neighbor of a does not lie on at least one of the lines $\overleftrightarrow{bc}$, $\overleftrightarrow{cd}$ and $\overleftrightarrow{bd}$.

On the other hand, let a be a small vertex. Consider a maximal non-branching path containing a, let us denote it P. Observe that for every vertex b, $(a, b) \in R$ if and only if b lies on this path: by Observation 4.7, if a lies on a line $\overleftrightarrow{xy}$, this line has to contain both ends of P.

Consider the endpoints of P, denote them b and c. Each of these points must be either a leaf, or a large vertex. If it is a large vertex, by the previous observation $(b, a) \in R$ cannot hold.

These two observations tell us all we need to know about the relation R for the purpose of the reconstruction. We construct the relation from the set of lines, and we identify all the large vertices. Moreover, we find all the maximal non-branching paths in the graph and which large vertices are its endpoints.

Consider the oriented graph of the relation R, a graph where ab is an edge if $(a, b) \in R$. Non-branching paths of the tree form a complete oriented subgraph, and additionally, there is an edge going from every vertex of the path to the endpoint(s) of the path. Examine Figure 7.2.

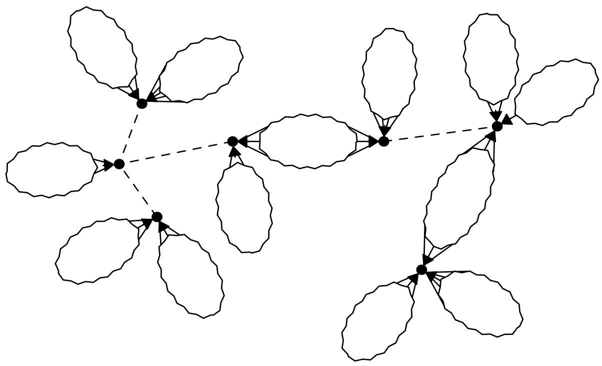

Figure 7.2: The graph of the relation R of a tree. The "bubbles" are complete directed graphs, the dashed lines are edges of the tree we do not know yet.

(*2*) Now, we need to find out edges between large vertices. Let a and b be two large vertices. Observe which points lie in the intersection of all lines containing both a and b. We will show that these are exactly the points of the path ab. Clearly, points of this path lie on every line containing a and b, because every line induces a connected subgraph. Now, let a_1 and a_2 be the neighbors of a not lying on the path ab. Similarly, let b_1 and b_2 be such neighbors of b. Observe that $\overleftrightarrow{a_1b_1} \cap \overleftrightarrow{a_2b_2} =$ the path ab. Therefore, the intersection of all such lines must be the path ab.

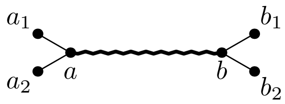

Figure 7.3: The intersection of $\overleftrightarrow{a_1b_1}$ and $\overleftrightarrow{a_2b_2}$ is the path ab.

Now, we see that if a and b are adjanced, this path contains only these two vertices.

Finally, from the relation R we can easily connect all the small vertices on non-branching paths to the endpoints of the path. However, note that we will never be able to distinguish the order of vertices on such a path. But this does not matter, because regardless of this order the resulting tree will be isomorphic to the desired result. □

Corollary 7.3. *No two non-isomorphic trees are equilinear.*

Proof. For a given set of lines, the algorithm always builds exactly one tree. Therefore, no two non-isomorphic trees can have the same sets of lines. □

Chapter 8

Conclusions

In the article we have considered a generalization of the de Brujin-Erdős theorem for discrete metric spaces induced by graphs. We have not been able to prove the Klee-Wagon conjecture, but we have obtained vast knowledge about the structure of lines in graphs.

We have proved several observations useful in easily determining whether a vertex lies on a line, and whether two lines are different. We have seen that the line structure is closely connected to shortest paths in the graph.

Then we have considered bridges, cuts, and separators and we have proved that lines in separated parts are closely connected. We have seen an interesting connection between cuts and universal lines: every bridge defines a universal line, and the union of lines defined by a cut contains all the vertices of the graph. We conjecture that if two graphs satisfy the Klee-Wagon conjecture, their articulation join satisfies it as well.

Next, we have analyzed several interesting classes of graphs. We have been able to completely analyze the structure of several graph classes, such as trees, cycles, or complete k-partite graphs. We conjecture that the conditions described in the Klee-Wagon can be strengthened. We have completely described graphs where every line is universal, and we have given numerous examples of graphs where every line is different, such as the complete graph K_n. It is still unclear which graphs exactly belong in the latter class. We have shown a surprising property between lines in a graph with a diameter of 2 and its complement. We have not yet found any similar relationships for graphs with a larger diameter.

We have also considered the class of graphs with a given number of different lines. We have completely described these graphs for several small numbers, and proved that such classes are often infinite. However, the infinite property holds only because of one specific operation which adds vertices, but preserves the lines.

We have briefly touched the problem for general finite discrete metric spaces, interpreted as weighted graphs. Many of the previous results hold here as well, and we have been able to count the number of lines in a cycle easily using weighted graphs. However, several notions simple in simple graphs are difficult to interpret for weighted graphs.

Last, we have evaluated the reverse problem: given only a representation of the lines in a graph, what can we tell about the graph itself? We have shown that in some cases, the graph can not be reconstructed uniquely, as there are sets of non-isomorphic graphs with exactly the same structure of lines. However, for specific classes of graphs, trees and graphs with a diameter of 2, we can completely reconstruct the graph using only minimal information about the lines.

Several interesting questions are still open:

- Prove the Klee-Wagon conjecture 1.4 for specific graph classes, especially if the class itself does not guarantee fulfillment of one of the conditions trivially.
- Is the articulation join neutral in regard to the Klee-Wagon conjecture?
- What is the minimal number of lines in a graph without a universal line?
- Which graphs have all the lines different?
- Describe the class of k-linear graphs for some $k > 3$, or at least show for which integers k this classes are empty or finite.
- Improve one of the bounds in Theorems 5.5 or 5.6 not to use Ramsey numbers.
- Which graphs can we reconstruct from their line description? It is possible to create at least one graph satisfying a given line description in polynomial time?

Acknowledgments

Most of all, we would like to thank Jan Kratochvíl and Pavel Valtr for introducing us to the problem and guiding our research. We would also like to thank Ondra Bílka, Martin Doucha, Tomaš Gavenčiak, Martin Kupec and Jan Volec for their insights and help in proving several of the results in this work.

Finally, we would like to thank Galina Jirásková for her generous help with the technical side of putting the work together.

Bibliography

[1] J. J. Sylvester, Mathematical Question 11851, *Educational Times* **59** (1893), p. 98.

[2] V. Chvátal, Sylvester-Gallai theorem and metric betweenness, *Discrete & Computational Geometry* **31** (2004), p. 175 - 195.

[3] X. Chen, The Sylvester-Chvatal Theorem, *Discrete & Computational Geometry* **35** (2006).

[4] N. G. de Bruijn and P. Erdős, On a combinatorial problem, *Indagationes Mathematicae* **10** (1948), p. 421 - 423.

[5] V. Klee and S. Wagon, Old and New Unsolved Problems in Plane Geometry and Number Theory, Mathematical Association of America, 1991.