On a Generalization of the de Bruijn-Erdős Theorem

Cathryn Supko

A thesis
in
The Department
of
Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Computer Science at
Concordia University
Montréal, Québec, Canada

July 2014

Concordia University

School of Graduate Studies

This is to certify that the thesis prepared

By: Cathryn Supko

Entitled: On a Generalization of the de Bruijn-Erdős Theorem

and submitted in partial fulfillment of the requirements for the degree of

Master of Computer Science

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair

Dr. Volker Haarslev

______________________________ Examiner

Dr. Thomas Fevens

______________________________ Examiner

Dr. Brigitte Jaumard

______________________________ Supervisor

Dr. Vašek Chvátal

Approved by ______________________________

Chair of Department or Graduate Program Director

__________ 20_____ ______________________________

Dean of Faculty

Abstract

On a Generalization of the de Bruijn-Erdős Theorem

Cathryn Supko

The de Bruijn-Erdős Theorem from combinatorial geometry states that every set of n noncollinear points in the plane determine at least n distinct lines. Chen and Chvátal conjecture that this theorem can be generalized from the Euclidean metric to all finite metric spaces with appropriately defined lines. The purpose of this document is to survey the evidence given thus far in support of the Chen-Chvátal Conjecture. In particular, it will include recent work which provides an $\Omega(\sqrt{n})$ lower bound on the number of distinct lines in all metric spaces without a universal line.

Acknowledgements

First and foremost, I would like to thank my gracious supervisor Dr. Vašek Chvátal for his guidance throughout the last few years. I would also like to thank Pierre Aboulker and Rohan Kapadia for putting up with me on a regular basis, Xiaomin Chen for his helpful insights, and Samuel Laferriere and Ben Inch for their feedback on this document. Last but certainly not least, I would like to thank my family for their constant encouragement and support.

Contribution of Authors

The results which appear in the final chapter of this document are based on joint work with Pierre Aboulker, Xiaomin Chen, Guangda Huzhang, and Rohan Kapadia. They are the result of many sessions of collaboration during which we each contributed equally to their development.

Contents

Chapter 1

Introduction

Our starting point is a special case of a classic theorem due to de Bruijn and Erdős ([14]) from an area of combinatorics known as incidence geometry. This result is one of two theorems known as "the de Bruijn-Erdős Theorem" and asserts the following:

Theorem 1.1. *Every noncollinear set of n points in the plane determines at least n distinct lines.*

As the name suggests, theorems from incidence geometry are less concerned with angles and lengths of line segments than with the incidences of points and lines. It was noted by Erdős that Theorem 1.1 can in fact be viewed as a consequence of another well-known result from incidence geometry referred to as the Sylvester-Gallai Theorem [16]:

Theorem 1.2. *For every set S of at least 3 noncollinear points in the plane, there exists a line which passes through exactly two points of S.*

One notion which is significant in incidence geometry is that of *betweenness*: in the plane, it is said that a point y is *between* points x and z, denoted by $[xyz]$, if $d(x,z) = d(x,y) + d(y,z)$ where d is the Euclidean distance metric. The set of points contained in a line determined by two points x and y, which we will denote by $\overline{xy}$, can be defined in these terms as

$$\overline{xy} = \{x, y\} \cup \{z : [zxy]\} \cup \{z : [xzy]\} \cup \{z : [xyz]\}.$$

Several axiomatizations of betweenness in the Euclidean plane have been established by mathematicians including but not limited to Pasch [21], Peano [22], Hilbert [17], and Coxeter [13]. Coxeter used seven of his ten axioms of planar betweenness in his proof of Theorem 1.2 which appears in [13], making his axiomatic system of particular interest to us. The following is a list of the axioms of planar betweenness on which his proof relies:

(i) There are at least two points.

(ii) If a and b are two distinct points, then there is at least one point c such that $[abc]$.

(iii) If $[abc]$ then $a \neq c$.

(iv) If $[abc]$, then $[cba]$ but not $[bca]$.

(v) If a and b are distinct points on the line $\overline{cd}$, then c is on the line $\overline{ab}$.

(vi) If $\overline{ab}$ is a line, then there is a point c not on this line.

(vii) If a, b, and c are three points not lying on the same line such that $[bcd]$ and $[cea]$ then there is a point f on the line $\overline{de}$ for which $[afb]$.

The lack of a need for explicit reference to the Euclidean metric in the proof of Theorem 1.2 suggests that perhaps there are other structures for which incidence theorems make sense. Moreover, the concept of betweenness need not be restricted to Euclidean distances: *metric betweenness,* a notion introduced by Menger in [20], is a ternary relation in a metric space (X, d) such that

$$[xyz] \iff d(x,z) = d(x,y) + d(y,z).$$

In [9], Chen and Chvátal use the idea of metric betweenness to suggest a possible generalization of the de Bruijn-Erdős Theorem. Similar to the definition of a line in the plane in terms of betweenness, they define a line determined by two points x and y in a metric space (X, d) as

$$\overline{xy} = \{x, y\} \cup \{z : [zxy]\} \cup \{z : [xzy]\} \cup \{z : [xyz]\}.$$

If a line $\overline{xy}$ contains the entirety of X, then $\overline{xy}$ is called *universal.*

Many, but not all, of Coxeter's axioms for planar betweenness also hold for metric betweenness. As a result, lines in general metric spaces exhibit very different behaviour from lines in the plane. To make this more concrete, we will consider the following example: let

$$X = \{v_1, v_2, v_3, v_4, v_5\}, \tag{1.1}$$

$$d(v_i, v_j) = 1 \text{ if } j = i+1 \mod 5 \text{ or } j = i-1 \mod 5,$$

and

$$d(v_i v_j) = 2 \text{ otherwise},$$

where i and j are integers between 1 and 5. In this metric space, we can see that

$$\overline{v_1 v_2} = \{v_1, v_2, v_3, v_5\}$$

$$\overline{v_1 v_3} = \{v_1, v_2, v_3\},$$

and

$$\overline{v_2v_4} = \{v_2, v_3, v_4\}.$$

This should be considered peculiar not only because the intersection of the two lines $\overline{v_1v_3}$ and $\overline{v_2v_4}$ exceeds one, but also because the line $\overline{v_1v_3}$ is a proper subset of the line $\overline{v_1v_2}$, both scenarios being forbidden in the Euclidean plane.

In this framework, Chen and Chvátal posed the following conjecture:

Conjecture 1.3. *Every finite metric space (X, d) such that $|X| \geq 2$ either has a universal line or the points of X determine at least $|X|$ distinct lines.*

Although there have been several partial results in the direction of a positive response, Conjecture 1.3 remains wide open today. As previously mentioned, in the Euclidean metric Theorem 1.1 is a corollary of Theorem 1.2. It is therefore natural to ask if the same is true in all metric spaces. Unfortunately, as is exhibited by the example provided in (1.1), a generalized version of Theorem 1.2 with this definition for lines in metric spaces is false: every line determined by two points at distance 1 contains precisely four points and every line determined by points at distance 2 contains precisely three points. The metric space in this example does, however, have at least as many lines as it has points: no two lines in this metric space are equal and hence there are 10 distinct lines. Chvátal conjectured and Chen proved (in [12] and [8] respectively) that with a slightly different definition of lines, Theorem 2 can be generalized to arbitrary metric spaces; however, as shown by Chen and Chvátal in [9], the de Bruijn-Erdős Theorem does not generalize in that setting.

Although it therefore appears that a proof of Conjecture 1.3 relying on a generalized version of Theorem 1.2 will not be possible, the definition of lines in metric spaces as given above does seem to be suitable for a generalization of Theorem 1.1. Evidence for this has thus far been concentrated on special cases of metric spaces. We will now state a fairly comprehensive list of known results, noting that all relevant definitions will be provided in the chapters which follow. In particular, Conjecture 1.3 has been verified in the following cases:

- *Every metric space (X, d) where X is a set of points in the plane no two of which share an $x-$ or $y-$ coordinate and d is the Manhattan metric either has a universal line or contains at least $|X|$ distinct lines* (Kantor and Patkós, [19]).

- *Every 1-2 metric space (X, d) either has a universal line or contains at least $|X|$ distinct lines* (Chvátal, [11]).

- *Every metric space (X, d) induced by a chordal graph either has a universal line or contains at least $|X|$ distinct lines* (Beaudou, Bondy, Chen, Chiniforooshan, Chudnovsky, Chvátal, Fraiman, and Zwols, [5]).

- *Every metric space (X, d) induced by a distance-hereditary graph either has a universal line or contains at least $|X|$ distinct lines* (Aboulker and Kapadia, [2]).

Further partial results have previously been achieved in more general settings:

- *Every metric space (X, d) where X is a set of points in the plane and d is the Manhattan metric either has a universal line or contains at least $|X|/37$ distinct lines* (Kantor and Patkós, [19]).
- *The vertices of every 3-uniform hypergraph (X, H) with $|X| = n \geq 2$ determine at least $(2-o(1)) \log n$ distinct lines* (Aboulker, Bondy, Chen, Chiniforooshan, Chvátal, and Miao,[1]).
- *Every metric space induced by a connected graph G either has a universal line or has $n^{2/7}/(2^{8/7})$ distinct lines* (Chiniforooshan and Chvátal, [10]).

The following stronger partial results were recently obtained by the author and her coauthors :

- *Every metric space induced by a connected graph on at least 2 vertices either has a universal line or has at least $(2n/3)^{1/2}$ distinct lines.*
- *Every metric space (X, d) contains either a universal line or at least $n^{1/2}/2$ distinct lines.*

The purpose of this thesis is to survey the results which have verified Conjecture 1.3 for particular metric spaces and to give the proofs of these new results. Many of the details of the proofs of the previous results will be given throughout the following sections so as to illuminate some of the key ideas which led to these improvements.

Chapter 2

Survey of results

2.1 Preliminaries and notation

Throughout what follows, we will always use n to denote the number of points in a metric space and m to denote the number of lines in a metric space, provided it is clear to which metric space we are referring. We say that a metric space (X, d) has the *de Bruijn-Erdős property* if the metric space either has a universal line or the points of the metric space determine as many distinct lines as there are points in the metric space. Keeping this in mind, we can see that the following statements are equivalent:

- Conjecture 1.3 is true.
- Every metric space (X, d) either contains a universal line or $m \geq n$.
- Every metric space (X, d) has the de Bruijn-Erdős property.

These three characterizations will be used interchangeably. We will now proceed with our chronicle.

2.2 Lines in hypergraphs

A hypergraph (V, H) is a set of *vertices* V together with a set H of *hyperedges*, which are subsets of the vertex set. It is said that a hypergraph is *k-uniform* if each of its hyperedges has cardinality precisely k. We can translate the definition of lines in metric spaces into the language of hypergraphs; indeed, from a metric space (X, d) we can construct the hypergraph $(X, H(d))$ which contains a vertex corresponding to each point in the ground set of (X, d) and has hyperedges

$$H(d) = \{\{x, y, z\} : [xyz]\}.$$

Chen and Chvátal define a line determined by vertices x and y in a 3-uniform hypergraph $(X, H(d))$ as

$$\overline{xy} = \{x, y\} \cup \{z : \{x, y, z\} \in H(d)\}.$$

Note that the lines determined by the vertices of the hypergraph $(X, H(d))$ will be precisely the same as the lines determined by the points of (X, d). Using this framework and the fact that in the Euclidean plane each pair of points can belong to precisely one maximal line, the original de Bruijn-Erdős Theorem can be stated as:

If every two vertices of (X, H) belong to precisely one maximal line and there is no line that contains the entirety of X then the number of distinct lines determined by the vertices of X is at least $|X|$.

In fact, since for every metric space (X, d) a 3-uniform hypergraph $(X, H(d))$ can be constructed such that the lines determined by the points in the ground set of the metric space are the same as the lines in this hypergraph, a proof of a lower bound on the number of lines in a 3-uniform hypergraph also serves as a proof of a lower bound on the number of lines in metric spaces. It is for this reason that lines in hypergraphs are of interest to us. Some partial results in this direction appear in [4], [1], and [10]. Unfortunately, as shown in [14], there exist examples of 3-uniform hypergraphs which have no universal line and fewer than n lines:

Theorem 2.1. *There exist positive constants n_0 and c such that if $n \geq n_0$ then $m \leq c^{\sqrt{\ln(n)}}$.*

Theorem 2.1 does not, however, disprove Conjecture 1.3. Although a 3-uniform hypergraph $(X, H(d))$ can be constructed from every metric space (X, d) such that the two share the same lines, it is not the case that every 3-uniform hypergraph corresponds to a metric space. For example, it has been shown that the hypergraph which is constructed from the Fano plane does not arise from any metric space (see [12]).

As a result, we can infer that our study of lines in arbitrary 3-uniform hypergraphs may not immediately allow us to reach the desired lower bound on the number of distinct lines determined by the points of a metric space, but can certainly still provide insight. The best known lower bound on the number of lines in hypergraphs, due to Aboulker, Bondy, Chen, Chiniforooshan, Chvátal, and Miao ([1]) was also the best known lower bound on the number of lines in arbitrary metric spaces before the results which will appear in Chapter 4 were proved:

Theorem 2.2. *The vertices of every 3-uniform hypergraph (X, H) with $|X| = n \geq 2$ determine at least $(2 - o(1)) \log n$ distinct lines.*

We will present the proof of Theorem 2.2 as it appears in [1]. Their proof relies on a result of which Theorem 2.2 is an improvement and is due to Beaudou, Bondy, Chen, Chiniforooshan, Chudnovsky, Chvátal, Fraiman and Zwols (see [4]):

Theorem 2.3. *The vertices of every 3-uniform hypergraph (X, H) with $|X| = n \geq 2$ determine at least $\log n$ distinct lines.*

It also requires the use of the following lemmas.

Lemma 2.4. *For every $\epsilon > 0$, there exists $\delta > 0$ such that for every positive integer N,*

$$\sum_{i<\delta N} \binom{N}{i} \le 2^{\epsilon N}.$$

Proof of Lemma 2.4. A special case of Bernstein's inequality from [6] states that for every positive integer N and every integer k between 0 and $\lfloor N/2 \rfloor$,

$$\sum_{i=0}^{k} \binom{N}{i} \le \left(\frac{N}{k}\right)^k \left(\frac{N}{N-k}\right)^{N-k}.$$

Letting δ be such that $\lim_{\delta \to 0+} \left(\frac{e}{\delta}\right)^{\delta} = 1$,

we then have that

$$\left(\frac{N}{k}\right)^k \left(\frac{N}{N-k}\right)^{N-k} \le \left(\frac{eN}{k}\right)^k \le 2^{\epsilon N}.$$

□

Let L be the set of lines determined by the vertices of a hypergraph (X, H). We will denote $|L|$ by m. Define the mapping $\alpha : X \to 2^m$ to be such that

$$\alpha(x) = \{\ell \in L : x \in \ell\}$$

for every $x \in X$ and the mapping $\beta : X \to 2^m$ to be such that

$$\beta(x) = \{\overline{xw} : w \ne x, w \in X\}$$

for every $x \in X$.

Lemma 2.5. *If $f : X \to 2^m$ is a mapping such that $\beta(x) \subset f(x) \subset \alpha(x)$ for all $x \in X$, then f is one-to-one and $\{f(x) : x \in X\}$ is an antichain.*

Proof. We first observe that from our definitions of α and β, $\beta(x) \subset \alpha(x)$ for all $x \in X$. To show that $\{f(x) : x \in X\}$ is an antichain, it is therefore enough to show that $\beta(x) - \alpha(y) \ne \emptyset$ for every $x, y \in X$ as this will guarantee that $f(x) \not\subset f(y)$ for all $x, y \in X$. We will do so by considering the line $\overline{xy}$ for arbitrary distinct vertices x and y in X. We may assume that $\overline{xy}$ is not universal and so there exists a point z such that $z \notin \overline{xy}$. This means that $\{x, y, z\}$ is not a hyperedge of (X, H) and hence $\overline{xz} \in \beta(x) - \alpha(y)$. □

Lemma 2.6. *If x, y and z are vertices such that $\overline{xy} = \overline{xz}$, then $\alpha(y) \cap \beta(x) = \alpha(z) \cap \beta(x)$.*

Proof. To prove Lemma 2.6, it will suffice to show that if $\overline{xy} = \overline{xz}$ and $y \in \overline{xw}$, then $z \in \overline{xw}$. This is indeed the case: if $y \in \overline{xw}$ then $\{x, y, w\}$ must be a hyperedge of (X, H). This implies that $w \in \overline{xy}$, and so from our assumption that $\overline{xy} = \overline{xz}$ it must also be the case that $w \in \overline{xz}$. Hence, $\{x, z, w\} \in H$ and so indeed $z \in \overline{xw}$. □

Lemma 2.7. *If $n \geq 2$ and S is a nonempty set of s vertices which spans a set T of t lines, then*

$$m - t \geq \log(n - s) - s\log(t).$$

Proof. Denote the vertices of S by $v_1, ..., v_s$. We note that since we require that S is nonempty, it must also be the case that $t > 0$. We define a mapping $\gamma(x) : (V - S) \to T^s$ to be such that

$$\gamma(x) = (\overline{xv_1}, \overline{xv_2}, ..., \overline{xv_s}).$$

Since $T = \cup_{i=1}^{s}\beta(v_i)$, if there are vertices y and z in $V - S$ such that $\gamma(y) = \gamma(z)$, then by Lemma 2.6, $\alpha(y) \cap \beta(v_i) = \alpha(z) \cap \beta(v_i)$ for all v_i in S. It then follows that $\alpha(y) \cap T = \alpha(z) \cap T$. Combining this with Lemma 2.5 gives us that if y and z are distinct and $\gamma(y) = \gamma(z)$ then $\alpha(y) - T \neq \alpha(z) - T$.

We denote by C a subset of $V - S$ where γ is constant, and note that there exists such a set of size at least $(n - s)/t^s$. It now follows that

$$\frac{(n-s)}{t^s} \leq |C| \leq 2^{m-t};$$

hence, $\log(n - s) - s\log(t) \leq m - t$ as desired. □

Proof of Theorem 2.2. Letting ϵ be any fixed positive real number, we will show that for sufficiently large n, $m \geq (2 - 4\epsilon)\log n$. To do so, we consider a maximal set S of vertices which spans at least $((\delta \log n)/2)|S|$ lines, where δ is as in Lemma 2.4. We denote $|S|$ by s and the number of lines spanned by S as t. We may assume that $t < \log n$ since $m \geq t$ and so otherwise we would have the result of Theorem 2.2 immediately. This implies that we may also assume that $s < 4/\delta$.

If $t > 0$, then by Lemma 2.7 we now have that $m - t \geq (1 - o(1))\log n$; if $t = 0$ then the same bound on $m - t$ follows from Theorem 2.3. We may therefore also assume that $t \leq m/2$.

We will now consider the maximal set R of vertices of X such that $\beta(y) \cap T = \beta(z) \cap T$ for every pair of vertices $y, z \in R$. Note that $|R| \geq 2^t$. By Lemma 2.5, we know that β is one-to-one and so $\beta(y) - T$ must be distinct from $\beta(z) - T$ for every $y, z \in R$. We can also see that since S is maximal, the set $\beta(y) - T$ includes fewer than $(\delta \log n)/2$ lines as otherwise it could have been added to S. Combining these two observations with Lemma 2.4 gives us that

$$|R| \leq \sum_{i<\delta \log n/2} \binom{m-t}{i} \leq \sum_{i<\delta(m-t)} \binom{m-t}{i} \leq 2^{\epsilon(m-t)} \leq 2^{\epsilon m}$$

for all n such that $\log n/2 < m - t$. It follows that

$$n \leq 2^t|R| \leq 2^{t+\epsilon m} \leq 2^{(\epsilon+1/2)m} \leq 2^{m/(2-4\epsilon)}$$

and hence $(2 - 4\epsilon)\log n \leq m$. □

2.3 Integral metrics

A metric space (X, d) is called *integral* if for every choice of distinct points x and y in X, $d(x, y)$ is an integer. In particular, (X, d) is called a k-metric if $d(x, y)$ is always an integer between 1 and k. For general k, the best lower bound on the number of distinct lines in a $k-$metric is no different from the bound on general metric spaces; however, much more is known about 1-2 metric spaces, which, as the name suggests, are metric spaces where every non-zero distance is either 1 or 2. In [10], Chiniforooshan and Chvátal showed the following:

Theorem 2.8. *The number of distinct lines determined by the vertices of a 1-2 metric space on n points is at least* $(1 + o(1))2^{-7/3}n^{4/3}$.

Theorem 2.8 guarantees that every sufficiently large 1-2 metric space has the de Bruijn-Erdős property. This was later improved by Chvátal in [11], in which he showed that in fact all 1-2 metric spaces have the de Bruijn-Erdős property.

Theorem 2.9. *All 1-2 metric spaces on at least 2 points have the de Bruijn-Erdős property.*

We will now present Chvátal's proof of Theorem 2.9.

Proof. A 1-2 metric space (X, d) is *critical* if it is a minimal counter example to Theorem 2.9. We will prove Theorem 2.9 by showing through a series of claims that a critical 1-2 metric space does not exist. This will then guarantee that in fact every 1-2 metric space must have the de Bruijn-Erdős property.

We begin with a claim whose proof is very straight-forward. Its statement will require the following definition: two points x and y in a metric space are called *twins* if $d(x, y) = 2$ and for every vertex $z \in X - \{x, y\}$, $d(x, z) = d(y, z)$.

Claim 2.9.1. If v_1, v_2, v_3 and v_4 are four distinct points in a 1-2 metric space then:

(i) if $d(v_1, v_2) \neq d(v_3, v_4)$ then $\overline{v_1v_2} \neq \overline{v_3v_4}$,

(ii) if $d(v_1, v_2) = d(v_2, v_3) = 2$ then $\overline{v_1v_2} \neq \overline{v_2v_3}$, and

(iii) if $d(v_1, v_2) = d(v_2, v_3) = 1$ and v_1 and v_3 are not twins then $\overline{v_1v_2} \neq \overline{v_2v_3}$.

Proof of Claim 2.9.1. To verify (i), we first observe that if either $v_1 \notin \overline{v_3v_4}$ or $v_4 \notin \overline{v_1v_2}$ then the two lines are not equal. We may therefore assume that $v_1 \in \overline{v_3v_4}$ and $v_4 \in \overline{v_1v_2}$. Since (X, d) is a 1-2 metric space, it must therefore be the case that $d(v_3, v_1) = d(v_4, v_1) = 1$, from which we can deduce that $d(v_2, v_4) = 2$. This now implies that $v_2 \notin \overline{v_3v_4}$, and hence the two lines are not equal.

To verify (ii), we note that since $d(v_1, v_2) = d(v_2, v_3) = 2$ and $d(v_1, v_3) \in \{1, 2\}$, there is no formulation of the triangle inequality for which these three values will hold with strict equality. Therefore, v_3 will not be contained in $\overline{v_1v_2}$ and the two lines cannot be equal.

To verify (iii), we first note that since v_1 and v_3 are not twins, either $d(v_1, v_3) = 1$ or $d(v_1, v_3) = 2$ and there must exist a vertex x such that $d(v_1, x) \neq d(v_3, x)$. If it is the former then the two lines are not equal as v_1 would not be contained in $\overline{v_2 v_3}$, and so we may assume it is the latter. Without loss of generality, we may assume that $d(v_1, x) = 1$ and $d(v_3, x) = 2$. If $d(v_2, x) = 1$, then $x \in \overline{v_2 v_3}$ and $x \notin \overline{v_2 v_1}$ and hence the two lines are not equal. Otherwise, $d(v_2, x) = 2$ and so $x \in \overline{v_1 v_2}$ and $x \notin \overline{v_2 v_3}$ and again the lines are not equal.

Claim 2.9.2. For every pair u, v of twins in a critical 1-2 metric space (X, d), there exists a third point w in X such that $d(u, w) = d(v, w) = 2$ and $d(x, y) = 1$ whenever $x \in \{u, v, w\}$ and $y \notin \{u, v, w\}$.

Proof of Claim 2.9.2. Since S is a critical 1-2 metric space, we know that S does not have the de Bruijn-Erdős property whereas $S - \{v\}$ does. From this fact, we know that the line $\overline{uv}$ is not universal in S and hence there exists a point w such that $d(v, w) = d(u, w) = 2$ in S. We will show that w has the other desired properties stated in the conclusion of this claim, i.e. that $d(w, y) = d(v, y) = 1$ for all $y \in S - \{u, v, w\}$. We will show that this is the case by proving that $\overline{wu}$ must be a universal line in $S - \{v\}$: to see that this is the case, we will consider each of the lines $\overline{xy} \cap \{u, v\}$ where x and y are distinct vertices in $S - \{v\}$. We can see that since u and v are twins, if x and y are both distinct from u then in S the line $\overline{xy}$ will either contain both of u and v or neither u nor v. If one of x or y is equal to u, then the line $\overline{xy}$ will contain v if and only if $d(x, y) = 1$. The existence of w therefore ensures that there are fewer lines in $S - \{v\}$ then there are in S. From our assumption that S is critical, it must be the case that $S - \{v\}$ has a universal line $\overline{xy}$ such that in S $\overline{xy}$ does not contain v. Our previously stated observations about the intersection of lines in $S - \{v\}$ with $\{u, v\}$ reveal that the only candidate for such a universal line is determined by u and a vertex at distance 2 from u. We may therefore assume that w is this vertex. Since $\overline{wu}$ is universal line $S - \{v\}$ and u and v are twins, it follows that in S, $d(w, y) = d(v, y) = d(u, y) = 1$ for all vertices y in $S - \{u, v, w\}$ as promised.

Claim 2.9.3. No critical 1-2 metric space contains a pair of twins.

Proof of Claim 2.9.3. Let $S = (X, d)$ be a critical metric space. We will prove Claim 2.9.3 by assuming to the contrary that S does contain a pair of twins (u, v). To do so, we will consider the maximal set $Y = \{Y_1, ..., Y_k\}$ of pairwise disjoint 3-point subsets such that $d(x, y) = 1$ if $x \in Y_i$ and $y \notin Y_i$. and $d(x, y) = 2$ if $x, y \in Y_i$ for all integers i such that $1 \leq i \leq k$. By Claim 2.9.2 and our assumption that S contains a pair of twins, we are guaranteed that $k \geq 1$. We will be able to use this fact to show that the vertices of X determine at least $|X|$ distinct lines, which is a contradiction to our assumption that S does not have the de Bruijn -Erdős Property.

To see that this is the case, we will consider the sets of lines $L_1 = \{\overline{xy} : x, y \in Y_i, 1 \leq i \leq k\}$ and $L_2 = \{\overline{xy} : x \in Y_1, y \in X - Y\}$.

If $\overline{xy}$ is in L_1 and $\{x, y, z\} \in Y$ then by the definition of Y, $\overline{xy} = X - \{z\}$ and so

$$|L_1| = 3k. \tag{2.1}$$

If $\overline{xy}$ is in L_2, then $\overline{xy} = Y_1 \cup Z_y$ where $Z_y = \{z : d(y, z) = 2, z \in X - Y\}$. By the maximality of Y and k and Claim 2.9.2, we are guaranteed that if $y \neq w$ and y and w are in $X - Y$, then $Z_y \neq Z_w$. We now have that

$$|L_2| = |X| - 3k. \tag{2.2}$$

Further, we from these descriptions of the lines in L_1 and L_2 we can see that $L_1 \cap L_2 = \emptyset$, and so combining (2.1) and (2.2),

$$m \geq |L_1| + |L_2| = 3k + |X| - 3k = |X|$$

as desired.

To complete our proof of Theorem 2.9, we let T be a maximal set of lines that are all equal. The combination of Claim 2.9.1 and Claim 2.9.3 guarantees that T is either a set of lines such that no single point determines more than one line in T, or a set of lines of size at most 4. This implies that $|T|$ is at most $\max\{(n-1)/2, 4\}$ and hence

$$m \geq \frac{\binom{n}{2}}{|T|} \geq n$$

for all $n \geq 7$. If $n \leq 7$, then the it can be shown by a routine case analysis that (X, d) has the de Bruijn-Erdős property and this portion of the proof will be omitted.

□

Chapter 3

Graph metrics

Another special case of an integral metric space which has been relatively well studied in this context are metric spaces induced by graphs. Before proceeding with our discussion, we will need several basic graph theoretic definitions, all of which can be found in [7]. A *graph* $G = (V, E)$ consists of a set of *vertices* $V(G)$ and a set of *edges* $E(G)$ which are unordered pairs of the vertices of G. If u and v are vertices of G and $\{u, v\}$ is an edge of G then u and v are said to be *adjacent* and are called *neighbours*; otherwise, u and v are *nonadjacent.* The set of all neighbours of a fixed vertex $v \in V(G)$ is called the *neighbourhood* of v and is denoted $N(v)$. A *path* P of length t in a graph G is a sequence of distinct vertices $p_1, p_2, ..., p_t$ in $V(G)$ such that for every value of i from 1 to $t-1$, $\{p_i, p_{i+1}\}$ is an edge. A graph is called *connected* if for every pair of vertices u, v in the graph there exists a path P with u and v as endpoints. If there exists a pair of vertices u and v for which no such path exists, then the graph is called *disconnected.*

The distance between vertices u and v of G, denoted $d(u, v)$ is defined to be the number of edges in a shortest path from u to v. The *diameter* of a graph G, denoted by $diam(G)$, is defined to be $max_{x,y \in G} d(x, y)$. Throughout what follows, we will assume that the graphs are connected; without this assumption there may be pairs of vertices for which their distance is not well-defined. The ground set $V(G)$ of a graph together with this distance function induce a metric space for every connected graph G. To see that $(V(G), d)$ is indeed a metric space, we note that it satisfies all of the necessary conditions associated with being a metric space:

- $d(u, v) \geq 0$ for all $u, v \in V(G)$ (there cannot be a path of negative length),
- $d(u, v) = 0$ if and only if $u = v$,
- $d(u, v) = d(v, u)$ for all $u, v \in V(G)$ (since the edges are not directed, every path from u to v is also a path from v to u), and
- $d(u, v) \leq d(u, w) + d(w, v)$ for all $u, v, w \in V(G)$ (the length of the shortest path between two vertices is at most the length of a walk between those two vertices which goes through a

particular vertex).

The remainder of this chapter will be split in three parts: the first two will pertain to several classes of graphs for which Conjecture 1.3 has been verified, and the third will give the proof of the best previously known lower bound on the number of distinct lines in metric spaces induced by arbitrary connected graphs.

3.1 Preliminary examples

In this section we will first define three classes of graphs and then show that if a graph G is contained in one of these three classes then G has the de Bruijn-Erdős Property.

A *clique* K_n is a graph where every pair of vertices is adjacent. A set of vertices that are pairwise nonadjacent are a *stable set*. A graph is *bipartite* if the vertex set of G can be partitioned into two sets X and Y such that X and Y are each stable sets. In a similar vein, a graph G is *complete k-partite* if the vertex set of G can be partitioned into sets $X_1, ..., X_k$ such that two vertices $u \in X_i$ and $v \in X_j$ are adjacent if and only if $i \neq j$ for all i and j between 1 and k. Notice that this definition requires that X_i is a stable set for every i such that $1 \leq i \leq k$. We will now show that if G is in one of these three classes, then the metric space induced by G has the De Bruijn-Erdős property.

We will first show that if G is a clique then the metric space induced by G either contains a universal line or at least n distinct lines. Note that if $n = 1$ then G is a single vertex and so no lines can be determined in this metric space. If $n = 2$ then G is an edge $\{u, v\}$ and the line determined by these two vertices will contain the entire vertex set since they in fact are the entire vertex set. It is therefore enough to consider G such that $n \geq 3$.

Fact 3.1. *Every metric space induced by a clique K_n on $n \geq 3$ vertices contains at least $\binom{n}{2}$ distinct lines.*

Proof. We will show that K_n has $\binom{n}{2}$ lines by proving that for every choice of vertices u and v in K_n, $\overline{uv} = \{u, v\}$. To do so, we let u, w and v be arbitrary distinct vertices in K_n and we will show that w is not contained in the line $\overline{uv}$. By the definition of a clique, $d(u, v) = d(w, v) = d(u, w) = 1$. From this we can see that

- $d(u, v) = 1 < 1 + 1 = d(u, w) + d(w, v)$,
- $d(u, w) = 1 < 1 + 1 = d(u, v) + d(v, w)$, and
- $d(v, w) = 1 < 1 + 1 = d(v, u) + d(u, w)$.

We can therefore conclude that w is not contained in $\overline{uv}$. Since w is an arbitrary vertex in the graph, it follows that the only vertices contained in the line $\overline{uv}$ are u and v. We know from the definition of a line that if u, v, w and z are points in a metric space, then $\overline{uv} = \overline{wz}$ only if w and z are contained in $\overline{uv}$; hence, for every pair of vertices u and v in K_n there cannot exist a line determined by vertices w and z such that $\overline{uv} = \overline{wz}$. Fact 3.1 now follows. □

We will now show that all metric spaces induced by bipartite graphs have the de Bruijn-Erdős Property.

Fact 3.2. *Every metric space induced by a bipartite graph G on n vertices with bipartition (X, Y) contains a universal line.*

Proof. Let u and v be adjacent vertices in G. By the definition of a bipartite graph, u and v must be in different parts of the partition. Without loss of generality, we may assume that u is contained in X and v is contained in Y. We will prove Fact 3.2 by showing that if w is an arbitrary vertex in G distinct from u and v then w must be contained in the line $\overline{uv}$. To do so, we first observe that since $d(u, v) = 1$, a vertex w will be excluded from the line $\overline{uv}$ if and only if $d(u, w) = d(v, w)$. It is therefore enough to show that no such vertex exists. We will proceed by considering the parity of $d(u, w)$ and $d(v, w)$.

First we will suppose that w is an arbitrary vertex in X, and we will let $P = \{u, p_1, ...p_s, w\}$ and $Q = \{v, q_1, ...q_t, w\}$ be shortest paths from u to w and from v to w respectively. Since there are no edges between two vertices which are in the same part, we can see that p_i is in Y if and only if i is an odd number between 1 and s and p_i is in X if and only if i is an even number between 1 and s. We are assuming that w is in X, and so p_s must be in Y and hence s must be odd. There are $s + 1$ edges in P and so we can conclude that $d(u, w) = s + 1$ is an even number. Similarly, q_i is contained in X if and only if i is an odd number between 1 and t, and q_i is in Y if and only if i is an even number between 1 and t. From our assumption that w is in X, we can deduce that p_t is in Y and so t must be even. Since $d(v, w) = t + 1$ we can conclude that $d(v, w)$ must be odd. From the difference in the parity of these two values, we can conclude that there does not exist a vertex in X such that $d(u, w) = d(v, w)$. The symmetric argument will hold by swapping X for Y, and so in fact if w is an arbitrary vertex in G then the parity of $d(u, w)$ and $d(v, w)$ must be different; hence, every vertex of G will be contained in $\overline{uv}$ and therefore it is universal. □

We will complete this section by considering complete k-partite graphs. Note that if $k = 1$ and the graph is connected then the graph must consist of a single vertex, and so no lines will be defined in this metric space. If $k = 2$ then the graph is a complete bipartite graph and so the fact that the metric space induced by this graph has the de Bruijn-Erdős Property follows from Fact 3.2. It is therefore enough to consider the case of when $k \geq 3$.

Fact 3.3. *Every metric space induced by a complete $k-$partite graph G on n vertices with partition $\{X_1, ..., X_k\}$ and $k \geq 3$ contains at least n distinct lines.*

Proof. We will prove Fact 3.3 by considering the cardinality of each X_i for $1 \leq i \leq k$. If there exists a value of i such that $|X_i| = 2$, then the line determined by the two vertices u and v contained in X_i will be universal; indeed, by the definition of a complete $k-$partite graph we are guaranteed that $d(u, v) = 2$ and $d(u, w) = d(v, w) = 1$ for every vertex w distinct from u and v. This implies that for every vertex w in G,

$$d(u, v) = 2 = 1 + 1 = d(u, w) + d(w, v)$$

and hence $\overline{uv}$ is universal.

We may now assume that $|X_i| = 1$ or $|X_i| \geq 3$ for all i. We will count the lines determined by the vertices of G by grouping them into two sets as follows:

- $\overline{uv} \in L_1$ if $u \in X_i, v \in X_j$, and $i \neq j$, and
- $\overline{uv} \in L_2$ if $u \in X_i, v \in X_j$, and $i = j$.

Let $\overline{uv}$ be a line in L_1 where i and j are fixed distinct integers between 1 and k and u and v are arbitrary vertices in X_i and X_j respectively. From the definition of a complete k-partite graph, we know that $d(u, v) = 1$, $d(u, w) = d(v, w) = 1$ if $w \notin X_i \cup X_j$, $d(u, w) = 1$ and $d(v, w) = 2$ if $w \in X_j$, and $d(u, w) = 2$ and $d(v, w) = 1$ if $w \in X_i$. Together these facts imply that $\overline{uv} = \{X_i \cup X_j\}$ for every $u \in X_i$ and $v \in X_j$ and so

$$|L_1| = \binom{k}{2}. \tag{3.1}$$

We will now consider the line $\overline{uv} \in L_2$ where i is a fixed integer between 1 and k and u and v are arbitrary vertices in X_i. We know that $d(u, v) = 2$, $d(u, w) = d(v, w) = 2$ if $w \in X_i$, and $d(u, w) = d(v, w) = 1$ if $w \notin X_i$. This implies that $\overline{uv} = \{\cup_{\ell \neq i} X_\ell \cup \{u, v\}\}$ and hence

$$|L_2| = \sum_{i=1}^{k} \binom{|X_i|}{2}. \tag{3.2}$$

We note that

$$|L_1| = \binom{k}{2} \geq k \geq \sum_{i:|X_i|=1} |X_i|.$$

Combining (3.1) and (3.2) now gives us that

$$\begin{aligned}
|L_1| + |L_2| &= \binom{k}{2} + \sum_{i=1}^{k} \binom{|X_i|}{2} \\
&= \binom{k}{2} + \sum_{i:|X_i|\geq 3} \binom{|X_i|}{2} \\
&\geq \sum_{i:|X_i|=1} |X_i| + \sum_{i:|X_i|\geq 3} |X_i| \\
&= n.
\end{aligned}$$

Therefore in this case G will have at least n lines as required to complete the proof.

□

3.2 Subgraphs

A *subgraph* H of a graph G is a graph obtained from G by deleting a subset of the vertices and the edges of G. An *induced subgraph* H of a graph G is a graph whose vertex set is obtained from deleting a subset of vertices from G and whose edge set contains the edge $\{u, v\}$ if and only if $u, v \in V(H)$ and $\{u, v\} \in E(G)$. We denote the distance between vertices u and v in a subgraph H by $d_H(u, v)$. A subgraph H of G is called *isometric* if $d_G(u, v) = d_H(u, v)$ for every u and v in $V(H)$. Note that an isometric subgraph is always an induced subgraph; however, an induced subgraph is not necessarily isometric.

Fact 3.4. *If H is an isometric subgraph of G, then the number of lines in the metric space induced by G is at least the number of lines in the metric space induced by H.*

Proof. This follows directly from the definition of an isometric subgraph. □

Several classes of graphs can be defined in terms of their subgraphs. A familiar example is a *tree*, which is a connected graph which does not contain a cycle as a subgraph. Classifying graphs in terms of their forbidden subgraphs is useful because frequently we are able to exploit known structural properties of these classes to prove that they have other desirable properties (for example, the de Bruijn-Erdős property). This will be the case for each of the classes of graphs discussed in this section. We will begin by showing that trees have the de Bruijn-Erdős property:

Fact 3.5. *Every metric space induced by a tree on n vertices contains a universal line.*

Proof. Let T be a tree on n vertices, u be a leaf of T, and v be the neighbour of u in T. Since T is a tree, the only path from u to an arbitrary vertex w in T must go through v; therefore, $d(u, w) = d(u, v) + d(v, w)$ for every vertex w in T and hence $\overline{uv}$ is universal. □

The remainder of this section will discuss two other classes of graphs defined in terms of their subgraphs for which Conjecture 1.3 has been verified.

3.2.1 Chordal graphs

A graph is *chordal* if it does not contain any cycles with at least 4 vertices as an induced subgraph. Beaudou, Bondy, Chen, Chiniforoooshan, Chudnovsky, Chvátal, Fraiman, and Zwols proved in [5] that every metric space induced by a chordal graph has the de Bruijn- Erdős property. The remainder of this subsection will give their proof of this result.

Theorem 3.6. *Every metric space induced by a chordal graph G either has a universal line or has at least n distinct lines.*

Their argument relies on the following theorem due to Dirac about the structure of chordal graphs whose proof will be omitted (see [15]).

Theorem 3.7. *If G is a connected chordal graph on at least two vertices then there exist two distinct vertices u and v such that $N(u)$ and $N(v)$ each induce a clique.*

Proof of Theorem 3.6. A vertex v is called simplicial if the subgraph induced by $N(v)$ is a clique. By Theorem 3.7, if G is a chordal graph then G contains two simplicial vertices. We will denote one of them by s and will denote two other distinct vertices in G by x and y. We write $d(s, x)$ as j and $d(s, y)$ as k. Without loss of generality, we may assume that $j \leq k$. To prove this theorem, we will begin by proving the following claim:

Claim 3.6.1. If $\overline{sx} = \overline{sy}$ then $\overline{xy}$ is a universal line in G.

Proof of Claim 3.6.1. We define the set $A_\ell(s)$ to be the set of vertices z in G such that $d(s, z) = \ell$. Our proof of Claim 3.6.1 will require validation of the following statements:

(i) If $\overline{sx} = \overline{sy}$ then $d(s, y) = d(s, x) + d(x, y)$.

(ii) If $\overline{sx} = \overline{sy}$ then x is a cutvertex in G separating y from s.

The combination of these two statements will be almost enough prove our claim: together they will allow us to show that every vertex w in G is contained in the line $\overline{xy}$.

We will first verify (i). To do so, we must show that $d(s, x) < d(s, y) + d(x, y)$ and $d(x, y) < d(x, s) + d(s, y)$. The former is guaranteed by the assumption that k is at least j and the fact that $d(x, y)$ is at least 1. The latter is guaranteed by the fact that G is simplicial: to see this, we will select one vertex in $A_1(s)$ which is on a shortest path from x to s and one vertex in $A_1(s)$ which is on a shortest path from y to s and denote these vertices by w and z respectively. If w is equal to z then

$$d(x, y) \leq (j - 1) + (k - 1) < j + k.$$

Otherwise, w and z are adjacent since $A_1(s)$ is a clique and so

$$d(x, y) \leq (j - 1) + 1 + (k - 1) < j + k.$$

This proves (i).

We will now prove (ii). To do so, we note that $A_j(s)$ is a cutset of G which separates s from y. We define B to be a minimal subset of $A_j(s)$ which is a cutset separating s from y. By our assumption that $\overline{sx} = \overline{sy}$ and (i), we can see that x will be contained in B. We denote by z an arbitrary vertex in B, and will prove our claim by showing that in fact $z = x$. We will do so by considering $d(x,z)$. We first note that $d(x,z) \leq 1$: to see this, note that since x and z are contained in a minimal cutset separating y from s, there exists a path from x to z both in the component S containing s and in the component Y containing y in $G - B$. We will denote by P and Q the shortest paths from x to z in S and Y respectively. If $d(x,z)$ is at least two then the lengths of both P and Q must be at least 2, from which we can infer that $P \cup Q$ would be a chordless cycle with at least 4 vertices. This cannot happen since G is a chordal graph; hence $d(x,z) \leq 1$. If $d(x,z) = 1$, then P must be a path of length 2; otherwise G is not chordal. This implies that there exists a vertex w in $Y \cap A_{j+1}(s)$; however, the existence of w guarantees that z is contained in $\overline{sy}$ and not $\overline{sx}$, contradicting our assumption that $\overline{sx} = \overline{sy}$. To see this, we first note that $d(w,s) = d(w,x) + d(x,s)$ and so w is contained in $\overline{sx}$. It must therefore also be contained in $\overline{sy}$ and be such that $d(s,y) = d(s,w) + d(w,y)$. Since w is adjacent to z, we can then infer that $d(s,y) = d(s,z) + d(z,w) + d(w,y) = d(s,z) + d(z,y)$ and hence that z is contained in $\overline{sy}$. On the other hand, since $d(s,x) = d(s,z) = j$ and $d(x,z) = 1$ we can see that z will not be included in $\overline{sx}$. We have now proved that $d(x,z) < 1$, and so indeed $d(x,z) = 0$ and x must be the only vertex in B.

To complete the proof of Claim 3.6.1, we may assume that there exists a pair of vertices x and y such that $\overline{sx} = \overline{sy}$; otherwise, G contains at least $n-1$ lines which all contain s and so we have the result. We proceed by showing that an arbitrary vertex w in the graph will be contained in $\overline{xy}$. We define P to be a shortest path from w to s and Q to be a shortest path from w to y. By (ii), x is a cutvertex separating y from s and so at least one of P and Q must contain x. If Q contains x then by our definition of Q, $d(w,y) = d(w,x) + d(x,y)$. This implies that w is indeed contained in $\overline{xy}$.

We may now assume that P contains x. By our definition of P, we have that $d(s,w) = d(s,x) + d(x,w)$ which implies that w is contained in $\overline{sx}$. Recall that $\overline{sx} = \overline{sy}$, from which we can infer that w must also be contained in $\overline{sy}$. We know that $d(w,y) < d(w,s) + d(s,y)$ since s is simplicial. If $d(s,w) = d(s,y) + d(y,w)$, then by subtracting $d(s,x)$ from both sides of this equation we obtain that

$$d(x,w) = d(s,w) - d(s,x) = d(s,y) + d(y,w) - d(s,x) = d(x,y) + d(y,w).$$

This implies that in this case $\overline{xy}$ contains w. Else it must be the case that $d(s,y) = d(s,w) + d(w,y)$. Subtracting $d(s,x)$ from both sides we have that

$$d(x,y) = d(s,y) - d(s,x) = d(s,w) - d(s,x) + d(w,y) = d(x,w) + d(w,y).$$

Here we can again deduce that w is contained in $\overline{xy}$. This completes the proof of Claim 3.6.1.

Claim 3.6.1 is almost enough to prove the theorem. Indeed, Claim 3.6.1 guarantees that G will either have a universal line or at least $n-1$ distinct lines. To complete the proof we need to only provide one additional line or show that G contains a universal line. This can be done by considering the neighbourhood of s. If $|A_1(s)| = 1$ then the line determined by s and the single vertex in its neighbourhood is a universal line. If $|A_1(s)|$ is at least 2 then we label two distinct vertices in $A_1(s)$ by w and z and claim that the line $\overline{wz}$ is distinct from every line $\overline{sx}$. To see this, we observe that every line $\overline{sx}$ will contain s; however, $d(w, z) = d(z, s) = d(s, w) = 1$ and so the line $\overline{wz}$ does not contain s. We are now guaranteed that the vertices of metric space induced by G determine at least n distinct lines, as desired.

□

3.2.2 Distance-hereditary graphs

A graph is called *distance-hereditary* if every connected induced subgraph is also an isometric subgraph. In [2], Aboulker and Kapadia proved the following:

Theorem 3.8. *Every metric space (X, d) induced by a distance-hereditary graph G has the de Bruijn-Erdős property.*

We will not reproduce the entirety of their proof here, but will give the key ideas which are used. Similar to the proof of Theorem 3.6, their proof relies on other results which characterize the structure of graphs in this class. The three key such results that they make use of are:

(a) Any cycle of length at least 5 in a distance-heredity graph has two crossing chords.

(b) If x is a vertex in a distance-hereditary graph, u and v are adjacent vertices such that $d(u, x) = d(v, x) = i$, and z is a vertex such that $d(z, x) = i - 1$, then z is adjacent to u if and only if z is adjacent to v.

(c) Any 2-connected distance-hereditary graph with at least four vertices has two disjoint pairs of twins.

The term *twins* in (c) is used in a similar sense here as it is used in our discussion of 1-2 metrics: a pair of vertices u and v are twins if $N(u)-v = N(v)-u$. See a paper of Howorka [18] as a reference for (a) and a paper of Bandelt and Mulder [3] for (b) and (c). Together (a), (b), and (c) are used to prove the following lemmas about lines in distance hereditary graphs:

Lemma 3.9. *Let G be a connected distance hereditary graph and let $\{x, y\}$ be an edge of G. Then either the edge $\{x, y\}$ is contained in a triangle in G or $\overline{xy}$ is universal.*

Lemma 3.10. *Let G be a connected distance hereditary graph and x, a, and b be three distinct vertices of G. If $\overline{xa} = \overline{xb}$, then either $\overline{ab}$ is a universal line or $[axb]$.*

They then use a strategy similar to that used in the proof of Theorem 2.9: they assume that a metric space (X, d) induced by a distance-hereditary graph G is a minimal counter example to Conjecture 1.3, and then through a series of claims related to the existence of a pair of twins in the graph deduce that such a minimal counter example does not exist.

3.3 All graph metrics

In this subsection we will present the proof of Chiniforooshan and Chvátal from [10] which gives the best previously known bound for all metric spaces which are induced by graphs.

Theorem 3.11. *Every metric space induced by a connected graph G either has a universal line or has $n^{2/7}/(2^{8/7})$ distinct lines.*

The proof of Theorem 3.11 will rely on the following two Lemmas.

Lemma 3.12. *Every metric space induced by a connected graph G with diameter $diam(G)$ either has a universal line or has at least $(n/diam(G))^{2/3}/4$ distinct lines.*

Lemma 3.13. *Every metric space induced by a connected graph G with diameter $diam(G)$ either has a universal line or has at least $\sqrt{diam(G)/2}$ distinct lines.*

In order to prove Lemma 3.12, we will need to use the following additional Lemma about lines in hypergraphs. As discussed previously, for every metric space (X, d) we can construct a 3-uniform hypergraph $(V, H(d))$ whose vertex set determines the same lines as (X, d). Therefore, by showing that the hypergraph associated with the metric space induced by a graph G satisfies the assumptions of this lemma, we will be able to apply its conclusion. Here we will denote by K_4^3 the 3-uniform hypergraph on 4 vertices with all 4 possible hyperedges.

Lemma 3.14. *Let H be a 3-uniform hypergraph, x be a vertex of H, and $T = \{v_1, ..., v_t\}$ be a set of the vertices of H excluding x such that x, v_i, v_j and v_k do not induce a K_4^3 in H for every $1 < i < j < k \leq t$. Then H contains at least $(2|T|)^{2/3}/4$ distinct lines.*

Proof of Lemma 3.14. Throughout the proof, we may assume that $|T| > 4$ as otherwise $(2|T|)^{2/3}/4 \leq 1$ and so the conclusion follows immediately.

We define $S = \{v_1, ..., v_s\}$ to be a maximal set of vertices in T such that $\overline{xv_i} = \overline{xv_j}$ for every i and j between 1 and s. If $|S| \leq (2|T|)^{1/3}$, then H contains at least $|T|/|S| \geq (2|T|)^{2/3}/4$ distinct lines determined by x and the vertices in T. We may therefore assume that $|S| > (2|T|)^{1/3}$. We will complete the proof by showing that every line induced by a pair of vertices in S is distinct. To see this, we will consider the line $\overline{v_iv_j}$ for a fixed choice of i and j. We note that if v_k is an arbitrary vertex in S distinct from v_i and v_j, then the sets $\{x, v_i, v_j\}, \{x, v_j, v_k\}$ and $\{x, v_i, v_k\}$ must be hyperedges in H. Further, since v_i, v_j, and v_k are in T we know that these four vertices do not induce a K_4^3 and hence $\{v_i, v_j, v_k\}$ is not a hyperedge in H. From this we can infer that v_k is not contained $\overline{v_iv_j}$. Since v_k is an arbitrary vertex in S, we can moreover conclude that $\overline{v_iv_j} \cap S$ is precisely v_i and v_j. It is therefore indeed the case that each line determined by a pair of vertices in S is distinct. Hence, in H we have that

$$m \geq \binom{|S|}{2} \geq \frac{(2|T|)^{1/3}((2|T|)^{1/3} - 1)}{2} \geq \frac{(2|T|)^{2/3}}{4}$$

as desired. $\square$

Proof of Lemma 3.12. Let x be an arbitrary vertex in G. We can partition the vertex set of G into sets $A_1, ..., A_{diam(G)}$ such that for every k between 1 and $diam(G)$, A_k is the set containing all vertices y in G such that $d(x, y) = k$. We can see that there exists an integer ℓ between 1 and $diam(G)$ such that

$$|A_\ell| \geq \frac{n-1}{diam(G)} \geq \frac{n}{2diam(G)}.$$

We denote one fixed set A_ℓ where this inequality is satisfied by T. We will prove Lemma 3.12 by showing that from G we can construct a hypergraph $(X, H(d))$ such that the set of vertices in T satisfy the assumption of Lemma 3.14. Since $|T| \geq n/(2diam(G))$, applying Lemma 3.14 will immediately imply the result.

To proceed, we will now consider the hypergraph $(X, H(d))$ where $X = V(G)$ and

$$\{x, y, z\} \in H(d) \implies d(x, z) = d(x, y) + d(y, z).$$

To see that T satisfies the assumption of Lemma 3.14, we will consider three arbitrary vertices v_i, v_j, and v_k in T such that $\{x, v_i, v_j\}, \{x, v_j, v_k\}$ and $\{x, v_i, v_k\}$ are in $H(d)$. By our definition of T, $d(x, v_i) = d(x, v_j) = d(x, v_k) = \ell$ and hence if these three sets are hyperedges in $H(d)$ then

$$\begin{aligned} 2\ell &= d(v_i, v_j) = & d(x, v_i) + d(x, v_j) \\ &= d(v_j, v_k) = & d(x, v_j) + d(x, v_k) \\ &= d(v_i, v_k) = & d(x, v_i) + d(x, v_k). \end{aligned}$$

Since all of the quantities $d(v_i, v_j), d(v_j, v_k)$ and $d(v_i, v_k)$ are equal, we can deduce that $\{v_i, v_j, v_k\}$ is not in $H(d)$. The vertices v_i, v_j, and v_k were chosen arbitrarily from T and hence there are no three vertices of T which induce a K_4^3 with x. This shows that the assumptions of Lemma 3.14 hold and so our proof of Lemma 3.12 is now complete.

$\square$

Proof of Lemma 3.13. We define $P = p_0...p_{diam(G)}$ to be an isometric path of length $diam(G)$ in G and define S to be a maximal set of of edges in P such that $\overline{p_i p_{i+1}} = \overline{p_j p_{j+1}}$. If $|S| \leq \sqrt{2diam(G)}$ then the pairs of endpoints of the edges of P determine at least $diam(G)/|S| \geq \sqrt{diam(G)/2}$ distinct lines and so we have the desired result. We may therefore assume that $|S| \geq \sqrt{2diam(G)}$. To complete the proof, we select a vertex q which is not contained in the line $\overline{p_i p_{i+1}}$ generated by the endpoints of the edges in S. We may assume that q exists; otherwise, $\overline{p_i p_{i+1}}$ is universal and again we have the desired conclusion. We will show that for every three vertices p_i, p_j, and p_k which are endpoints of edges that are contained in S, at least two of $\overline{qp_i}, \overline{qp_j}$ and $\overline{qp_k}$ are distinct. This will guarantee that $m \geq |S|/2 \geq \sqrt{diam(G)/2}$ and hence will be enough to complete the proof of Lemma 3.13.

To prove this claim, we will begin by showing that if $i < \ell$, p_ℓ is an endpoint of an edge in S, and p_ℓ is contained in $\overline{qp_i}$ then $d(p_i, p_\ell) = d(p_i, q) + d(q, p_\ell)$. We will show that this is the case by eliminating the two other formulations of the triangle inequality as possibilities: we can see that

$$d(q, p_i) = d(q, p_{i+1}) \leq d(q, p_\ell) + d(p_\ell, p_{i+1}) < d(q, p_\ell) + d(p_\ell, p_i)$$

and

$$d(q, p_\ell) \leq d(q, p_{i+1}) + d(p_{i+1}, p_\ell) = d(q, p_i) + d(p_{i+1}, p_\ell) < d(q, p_i) + d(p_i, p_\ell).$$

The only remaining possible formulation of the triangle inequality which would allow p_ℓ to be contained in $\overline{qp_i}$ requires that $d(p_i, p_\ell) = d(p_i, q) + d(q, p_\ell)$. Since we may assume that $i < j < k$, it must therefore follow that if p_j and p_k are both contained in $\overline{qp_i}$ then $d(p_i, p_j) = d(p_i, q) + d(q, p_j)$ and $d(p_i, p_k) = d(p_i, q) + d(q, p_k)$.

We will now assume that $\overline{qp_i} = \overline{qp_k}$ since otherwise our claim follows immediately. From our assumptions that q is not contained in the line $\overline{p_k p_{k+1}}$ and that $d(p_k, p_{k+1}) = 1$, it must be the case that $d(p_k, q) = d(p_{k+1}, q)$. This, together with our observation that $d(p_i, p_k) = d(p_i, q) + d(q, p_k)$, implies that

$$d(p_i, p_k) = d(p_i, q) + d(q, p_k) = d(p_i, q) + d(q, p_{k+1}) \geq d(p_i, p_{k+1}) = k + 1 - i.$$

It must therefore be the case that either $d(p_i, q) > j - i$ or $d(p_k, q) > k - j$. If $d(p_i, q) > j - i$, then $d(p_i, p_j) = j - i < d(p_i, q) + d(q, p_j)$ and so in this case $p_j \notin \overline{qp_j}$. Otherwise, $d(p_k, q) > k - j$ from which we can see that $d(p_j, p_k) = k - j < d(p_j, q) + d(q, p_k)$ and so again $p_j \notin \overline{qp_k}$. Since p_j will not be contained in $\overline{qp_k}$ in both of these cases, we are guaranteed that $\overline{qp_k} \neq \overline{qp_j}$ and so our claim follows. $\square$

The proof of Theorem 3.11 now follows easily by combining these lemmas.

Proof of Theorem 3.11. If G is a graph with diameter at least $2^{-9/7} n^{4/7}$ then the result follows from Lemma 3.13; otherwise, it follows from Lemma 3.12. $\square$

Chapter 4

New results

Here we will present recent improvements on the lower bounds for the number of distinct lines in graph metrics as well as general metric spaces. This is joint work with Pierre Aboulker, Xiaomin Chen, Guangda Huzhang, and Rohan Kapadia.

4.1 A new bound for graph metrics

This subsection will be dedicated to the proof of the following improvement of Theorem 3.11:

Theorem 4.1. *Every metric space induced by a connected graph on at least* 2 *vertices either has a universal line or has at least* $\sqrt{2n/3}$ *distinct lines.*

The proof of Theorem 4.1 will rely on the following two lemmas, the first of which applies specifically to metric spaces induced by graphs and the second of which can be applied to all metric spaces.

Lemma 4.2. *If* G *is a graph without a universal line, then*

$$m \geq \frac{2n}{3diam(G)}.$$

For the next lemma we will need the following definition: an *isometric path* in a metric space is a sequence of points $p_1, ..., p_t$ such that $d(p_i, p_j) = d(p_i, p_k) + d(p_k, p_j)$ for all $1 \leq i \leq k \leq j \leq t$.

Lemma 4.3. *If* (X, d) *is a metric space which contains an isometric path* $P = \{p_1, ..., p_t\}$ *with* t *points, then* (X, d) *either has a universal line or contains at least* t *distinct lines.*

Note that in metric spaces induced by graphs, the maximal number of vertices in an isometric path will be equal to the diameter of the graph, and so in that case Lemma 4.3 implies that the points of such a metric space will determine at least $diam(G)$ distinct lines or a universal line.

Using Lemma 4.3, we were in fact able to generalize this lower bound on the number of lines to a $\Omega(\sqrt{n})$ bound for all metric spaces. The proof of that result will be presented in the next section.

Proof of Lemma 4.2. Let x and y be two vertices in G such that $d(x, y) = diam(G)$. We define the set $A_j(x)$ for every j between 1 and $diam(G)$ to be the vertices z such that $d(x, z) = j$.

We now fix $j = diam(G)/2$ and partition the set $A_j(x)$ into sets $B_1, ..., B_\ell$ such that vertices u and v are both contained in a set B_i for some i between 1 and ℓ if and only if $\overline{xu} = \overline{xv}$. We note that if $diam(G)$ is odd then in fact A_j is the empty set. If A_j is nonempty, then without loss of generality we may assume that sets B_i with i between 1 and an integer k which is at most ℓ are each singletons, and that the sets B_i with i between $k + 1$ and ℓ contain at least 2 vertices.

We partition the vertex set of $G - \{x\}$ into three sets in the following way:

- v is in R if $d(x, v) > j$.
- v is in S if v is in $\bigcup_{i=1}^{k} B_i$
- v is in T if $d(x, v) < j$ or if v is in $\bigcup_{i=k+1}^{\ell} B_i$.

We will denote $|R|$ by r, $|S|$ by s, and $|T|$ by t. Our proof of Lemma 4.2 will require the following claim:

Claim 4.2.1. Let i be an integer such that $k < i \leq \ell$ and w and z be vertices in G such that w is in B_i and z is in $A_{j+1}(x)$. Then z is not adjacent to w.

Proof of Claim 4.2.1. By our definition of B_i, in addition to w there exists a vertex u in B_i such that $\overline{xw} = \overline{xu}$. Since these two lines are equal, we must have that $d(u, w) = d(x, u) + d(x, w) = 2j = diam(G)$. We know that $diam(G)$ is greater than 2 since otherwise the conclusion follows from Theorem 2.9, and so if the distance from u to w is equal to $diam(G)$ then u and w cannot both be adjacent to z; hence, at least one of u and w is not adjacent to z. Without loss of generality, we may assume that u is not adjacent to z. We will denote $d(z, u)$ by a. We now observe that $1 < a \leq diam(G)$, which leads us to conclude that z is not contained in $\overline{xu}$:

$$d(x, u) = j < j + 1 + a = d(x, z) + d(z, u),$$

$$d(x, z) = j + 1 < j + a = d(x, u) + d(u, z),$$

and

$$d(u, z) = a < j + j + 1 = d(u, x) + d(x, z).$$

It must therefore also be the case that z is not contained in $\overline{xw}$. Writing $d(z, w)$ as b, we will complete the proof by showing that $1 < b \leq diam(G)$ for the following inequalities to be satisfied:

$$d(x, w) = j < j + 1 + b = d(x, z) + d(z, w)$$

$$d(z, w) = b < j + j + 1 = d(z, x) + d(x, w),$$

and

$$d(x,z) = j + 1 < j + b = d(x,w) + d(w,z).$$

Indeed, these inequalities will hold only if $1 < b \leq diam(G)$. We can now conclude that z and w are not adjacent, completing our proof of Claim 4.2.1.

Claim 4.2.1 will be useful to us as it allows us to assume that all of the vertices in T are at distance greater than $diam(G)/2$ from y. To be able to proceed, we will also need to rely on the following claim:

Claim 4.2.2: Let a be an integer strictly larger than $diam(G)/2$ and let x, w, and z be distinct vertices in G such that w and z are both contained in $A_a(x)$. Then $\overline{xw}$ is distinct from $\overline{xz}$.

Proof of Claim 4.2.1. To see that these lines must be distinct, we will show that under these assumptions z is not contained in $\overline{xw}$. Since w and z are distinct vertices contained in $A_a(x)$, we have that $d(x,w) = d(x,z) = a$ and $d(w,z) = b$ where b is an integer which is positive and at most $diam(G)$. From these facts we can determine that there is no formulation of the triangle inequality which will hold with strict equality for these 3 vertices:

$$d(x,w) = a < a + b = d(x,z) + d(w,z),$$

$$d(x,z) = a < a + b = d(x,w) + d(w,z),$$

and

$$d(w,z) = b \leq diam(G) < a + a = d(x,w) + d(x,z).$$

Hence, z is not contained in $\overline{xw}$ and so the two lines are distinct. This concludes the proof of our claim.

We now let $C = diam(G)/(diam(G) - 1)$. We can see that showing that $m \geq n/(Cdiam(G))$ will imply our desired result: indeed, if $diam(G) \leq 2$, then a stronger result follows from Theorem 2.9, and so we may assume that $diam(G) \geq 3$. This implies that $C = diam(G)/(diam(G)-1) \leq 3/2$ and hence if $m \geq n/Cdiam(G)$ then $m \geq 2n/(3diam(G))$. We will now be able to conclude the proof.

If $s \geq n/Cdiam(G)$, then from our definition of the sets B_i and S, each of the lines determined by x and a vertex in S will be distinct and so Lemma 4.2 follows.

If $r \geq 2C$, then there exists a value of $i > j$ such that

$$|A_i| \geq \frac{r}{j} \geq \frac{n}{2Cj} = \frac{n}{Cdiam(G)}.$$

By Claim 4.2.2, each of the lines determined by x and a vertex in $A_i(x)$ for this value of i will be distinct, and so again in this case Lemma 4.2 follows.

We may now assume that $s < n/(Cdiam(G))$ and $r < n/(2Cdiam(G))$ and hence

$$t = n - 1 - (s + t) > n - 1 - \frac{n}{Cdiam(G)} - \frac{n}{2Cdiam(G)} \geq n - \frac{n}{Cdiam(G)} - \frac{n}{2C}.$$

As noted previously, all of the vertices in T are at distance greater than j from y. We can therefore find a value of $i > j$ such that

$$|A_i(y)| \geq \frac{t}{j} \geq \frac{2(n - \frac{n}{Cdiam(G)} - \frac{n}{2C})}{diam(G)} = \frac{n}{Cdiam(G)}$$

with $C = d/(d-1)$ as prescribed. Lemma 4.2 now follows from Claim 4.2.2.

□

Proof of Lemma 4.3. Throughout our proof, we may assume that (X, d) does not contain a universal line as otherwise the conclusion of our lemma follows immediately. We can therefore select a vertex q_i for each line $\overline{p_i p_{i+1}}$ such that q_i is not contained in $\overline{p_i p_{i+1}}$. We do not require that q_i and q_j be distinct for distinct values of i and j. We claim that for every choice of distinct i and j, $\overline{p_i q_i} \neq \overline{p_j q_j}$. Without loss of generality, we will assume that $i < j$. To see that the lines $\overline{p_i q_i}$ and $\overline{p_j q_j}$ are distinct, we first observe that if $j = i + 1$ then q_i is not contained in $\overline{p_i p_j}$ and hence p_j is not contained in $\overline{p_i q_i}$.

We may therefore assume that $j > i + 1$. If p_j is not contained in $\overline{p_i q_i}$ then it is clear that $\overline{p_j q_j} \neq \overline{p_i q_i}$; otherwise, one of the following must hold:

(i) $d(p_i, q_i) = d(p_i, p_j) + d(p_j, q_i)$,

(ii) $d(q_i, p_j) = d(q_i, p_i) + d(p_i, p_j)$, or

(iii) $d(p_i, p_j) = d(p_i, q_i) + d(q_i, p_j)$.

If (i) or (ii) holds, then since P is an isometric path q_i must also be contained in $\overline{p_i p_{i+1}}$, which contradicts our choice of q_i. If (iii) holds, then it must also be true that $d(p_i, p_{j+1}) = d(p_i, q_i) + d(q_i, p_{j+1})$ since P is an isometric path. This shows that p_{j+1} is contained in $\overline{p_i q_i}$. On the other hand, by our choice of q_j we know that p_{j+1} is not contained in $\overline{p_j q_j}$ and hence the two lines cannot be equal.

We are therefore guaranteed that the $t - 1$ lines $\overline{p_i q_i}$ will all be distinct as promised. To find an additional line which is distinct from $\overline{q_i p_i}$ for every i between 1 and $t - 1$, we observe that the line $\overline{p_1 p_2}$ will fit this purpose: $\overline{p_1 p_2}$ will contain all of the vertices in P, whereas each $\overline{p_i q_i}$ excludes p_{i+1}. □

Proof of Theorem 4.1. We may assume that G is a graph without a universal line; indeed, if G contains a universal line then the result holds trivially. This assumption allows us to apply Lemma 4.2 and Lemma 4.3. The combination of these two lemmas is enough to prove the theorem: if

$diam(G)$ is at most $\sqrt{2n/3}$ then the result follows from Lemma 4.3; otherwise, it follows from Lemma 4.2. □

4.2 A new bound for metric spaces

This section will provide a generalization of Theorem 4.1.

Theorem 4.4. *Every metric space (X, d) contains either a universal line or at least $\sqrt{n}/2$ distinct lines.*

Proof of Theorem 4.4. Let x and y be points in X such that the distance between x and y is the greatest distance possible between two points in the metric space. We denote by t the number of vertices in the largest isometric path in (X, d) and we will proceed by considering the value of t.
Case 1. $t \geq \sqrt{n}/2$. In this case the result follows immediately from Lemma 4.3.
Case 2. $t < \sqrt{n}/2$. We define three subsets of the points of X in the following way:

- $z \in X_1$ if $d(z, x) > d(x, y)/2$
- $z \in X_2$ if $d(z, y) > d(x, y)/2$
- $z \in X_3$ if $z \in X - (X_1 \cup X_2)$.

Note that each point in X will be contained in at least one of these subsets and that if z is in X_3 then $d(z, x) = d(z, y) = d(x, y)/2$. We define a partial order on X_1 where $w \prec z$ if $d(x, z) = d(x, w) + d(w, z)$. Similarly, we define a partial order on X_2 where $w \prec z$ if $d(y, z) = d(y, w) + d(w, z)$. To see that this relation does indeed define a partial order, we note that these relations satisfy the following necessary conditions:

(i) $w \prec w$,

(ii) if $w \prec z$ and $z \prec w$, then $w = z$, and

(iii) if $u \prec w$ and $w \prec z$ then $u \prec z$.

Condition (i) is satisfied due to the fact that $d(w, w) = 0$ and hence $d(x, w) = d(x, w) + d(w, w)$. Since d is a metric, we know that $d(a, b) = 0$ if and only if $a = b$ and so (ii) follows. The validity of (iii) can be derived from the fact that d is a metric space and from the assumptions of condition (iii); indeed, from this we know that

$$0 = d(x, u) + d(u, w) - d(x, w) \tag{4.1}$$

and

$$0 = d(x, w) + d(w, z) - d(x, z). \tag{4.2}$$

Adding 4.1 to 4.2, we have that

$$0 = d(x,u) + d(u,w) + d(w,z) - d(x,z).$$

We would now like to deduce that $d(x,u) + d(u,z) = d(x,z)$, which will grant us (iii). Since d is a metric, we know that $d(x,u) + d(u,z) \geq d(x,z)$. They in fact must be equal as otherwise we would have that

$$0 = d(x,u) + d(u,w) + d(w,z) - d(x,z) \geq d(x,u) + d(u,z) - d(x,z) > 0,$$

a contradiction. Therefore, this relation does define a partial order as promised. Further, we observe that due to the transitivity of this relation which is guaranteed by (iii), if a set of vertices P is a chain in one of these partial orders then P is an isometric path with x or y as an endpoint.

We will complete our proof by considering the cardinality of each of the X_is. If $|X_1| \geq n/4$, then there is a set C of size at least $n/(4t)$ which is an antichain in X_1. We claim that if w and z are arbitrary vertices in C then $\overline{xw} \neq \overline{xz}$. To see this, we note that since C is an antichain, it must be true that $d(x,z) < d(x,w) + d(w,z)$ and $d(x,w) < d(x,z) + d(w,z)$. Further, since w and z are in X_1,

$$d(w,z) \leq d(x,y) < d(w,x) + d(x,z).$$

This shows that w is not contained in $\overline{xz}$ and hence $\overline{xz} \neq \overline{xw}$. Since w and z are arbitrary vertices in C, there are indeed $|C|$ distinct lines determined by x and each of the vertices of C.

Swapping x for y where necessary, the same argument shows that if $|X_2| \geq n/4$, then there are at least $n/(4t)$ distinct lines determined by y and the vertices in a maximal antichain in X_2. By our assumption that $t < \sqrt{n}/2$, in both of these cases we get that $m \geq n/(4\sqrt{n}/2) > \sqrt{n}/2$ as desired.

We may therefore assume that $|X_3| > n/2$. To complete our proof, we partition the vertices of $|X_3|$ into sets Y_i such that w and z are both in Y_i if and only if $\overline{xw} = \overline{xz}$. If $|Y_i| \leq \sqrt{n}$ for every choice of i then there are at least $n/(2\sqrt{n})$ distinct lines determined by x and the vertices contained in X_3. Otherwise, there exists a value of i such that $|Y_i| > \sqrt{n}$. We observe that for every choice of w and z in Y_i, since $d(x,w) = d(x,z) = d(x,y)/2$ and $\overline{xw} = \overline{xz}$ it must follow that $d(w,z) = d(x,y)$. This implies that for every choice of u, v, w, and z in Y_i, $\overline{uv} \neq \overline{wz}$. In this case we therefore have that

$$m \geq \binom{|Y_i|}{2} \geq \binom{\sqrt{n}}{2} > \sqrt{n}/2.$$

This concludes our proof. □

4.3 Conclusion and Future Directions

Although the results given here are a large step forward, there is still a long way to go towards determining the validity of Conjecture 1.3. One potential way to move forward would be to find a better lower bound in terms of the number of points in the largest isometric path in a metric space as this would immediately improve Theorem 4.4. It may also be fruitful to find a completely different strategy to improve the lower bound on all graphs and then try to generalize this approach to all metric spaces as Theorem 4.1 generalizes to Theorem 4.4. We note, however, that relying on graph metrics may be unnecessarily limiting as in most cases techniques which rely on induction cannot be applied. It may therefore be advantageous to consider $k-$metrics for values of k larger than 2 yet still sufficiently small to get a handle on.

Bibliography

[1] P. Aboulker, A. Bondy, X. Chen, E. Chiniforooshan, V. Chvátal, and P. Miao, Number of lines in hypergraphs, To appear in *Discrete Applied Mathematics*. `arXiv:1308.5393v1 [math.CO]`

[2] P. Aboulker, R. Kapadia, The Chen-Chvátal conjecture for metric spaces induced by distance-hereditary graphs, `arXiv:1312.3214 [math.CO]`

[3] H.J. Bandelt, H.M. Mulder, Distance-hereditary graphs, *J.Combin. Theory Ser. B* **41** (1986), 182-208.

[4] L. Beaudou, A. Bondy, X. Chen, E. Chiniforooshan, M. Chudnovsky, V. Chvátal, N. Fraiman, and Y. Zwols, Lines in hypergraphs, `arXiv:1112.0376v1 [math.CO]`. To appear in *Combinatorica.*

[5] L. Beaudou, A. Bondy, X. Chen, E. Chiniforooshan, M. Chudnovsky, V. Chvatal, N. Fraiman, and Y. Zwols, A De Bruijn-Erdős theorem for chordal graphs, `arXiv:1201.6376 [math.CO]`

[6] S. Bernstein, On a modification of Chebyshev's inequality and of the error formula of Laplace, *Section Mathématiques des Annales Scientifiques des Institutions Savantes de l'Ukraine* **1** (1924), 38-49.

[7] A. Bondy, U.S.R. Murty, Graph Theory, Springer, New York, 2008.

[8] X. Chen, The Sylvester-Chvátal theorem, *Discrete Computational Geometry* **35** (2006) 193-199.

[9] X. Chen and V. Chvátal, Problems related to a de Bruijn - Erdős theorem, *Discrete Applied Mathematics* **156** (2008), 2101 – 2108.

[10] E. Chiniforooshan and V. Chvátal, A de Bruijn - Erdős theorem and metric spaces, *Discrete Mathematics & Theoretical Computer Science* **13** (2011), 67 – 74.

[11] V. Chvátal, A de Bruijn-Erdős theorem for 1-2 metric spaces, to appear in *Czechoslovak Mathematical Journal* **31**. `arXiv:1205.1170 [math.CO]`

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

[13] H. S. M. Coxeter, Intro to Geometry, Wiley, New York, 1961.

[14] N. G. De Bruijn, P. Erdős, On a combinatorial problem, *Indagationes Mathematicae* **10** (1948), 421 – 423.

[15] G.A. Dirac, On rigid circuit graphs, *Abh. Math. Sem. Univ. Hamburg* **25** (1961), 71-76.

[16] P. Erdős, Three Point Collinearity, *American Mathematical Monthly* **50** (1943), Problem 4605, p. 65. Solutions in Vol **51** (1944), 169-171.

[17] D. Hilbert, Grundlagen der Geometrie, Teubner, Stuttgart, 1899.

[18] E. Howorka, A characterization of distance-hereditary graphs, *Quart. J. Math. Oxford Ser. 2* **28** (1977) 417-420.

[19] I. Kantor, B. Patkós, Towards a de Bruijn-Erdős theorem in the L1-metric, *Discrete & Computational Geometry* **49** (2013), 659 – 670.

[20] K. Menger, Untersuchungen über allgemeine Metrik. *Mathematische Annalen* **100** (1928), 75-163.

[21] M. Pasch, Vorlesungen über neuere Geometrie, Teubner, Leipzig, 1882.

[22] G. Peano, I principii di Geometria logicamente esposti, Fratelli Bocca, Torino, 1889.