Discrete Applied Mathematics 171 (2014) 137–140

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Note

Number of lines in hypergraphs

Pierre Aboulker [a], Adrian Bondy [b], Xiaomin Chen [c], Ehsan Chiniforooshan [d], Vašek Chvátal [a,*,1], Peihan Miao [e]

[a] Concordia University, 1455 De Maisonneuve Blvd. West, Montréal, Québec H3G 1M8, Canada
[b] 221 rue St. Jacques, 75005 Paris, France
[c] Shanghai Jianshi LTD, 2301, 20 North Chaling Road, Shanghai, 200032, China
[d] Google Kitchener-Waterloo, 151 Charles Street West, Suite 200, Kitchener, Ontario, N2G 1H6, Canada
[e] Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang District, Shanghai, 200240, China

ARTICLE INFO

Article history:
Received 28 August 2013
Received in revised form 10 February 2014
Accepted 11 February 2014
Available online 1 March 2014

Keywords:
Finite metric spaces
Lines in hypergraphs
De Bruijn–Erdős theorem
Extremal combinatorics

ABSTRACT

Chen and Chvátal introduced the notion of lines in hypergraphs; they proved that every 3-uniform hypergraph with n vertices either has a line that consists of all n vertices or else has at least $\log_2 n$ distinct lines. We improve this lower bound by a factor of $2 - o(1)$.

A classic theorem in plane geometry asserts that every noncollinear set of n points in the plane determines at least n distinct lines. (It is easy to see that the validity of this theorem implies the validity of its extensions to Euclidean spaces of all dimensions.) As noted by Erdős [9] in 1943, this theorem is a corollary of the Sylvester–Gallai theorem (which asserts that for every noncollinear set V of finitely many points in the plane, some line goes through precisely two points of V); it is also a special case of a combinatorial theorem proved by De Bruijn and Erdős [8] in 1948. In 2006, Chen and Chvátal [5] suggested that it might be generalized to all metric spaces. More precisely, for any two distinct points u, v in a Euclidean space, line $\overline{uv}$ in this space can be characterized as

$$\overline{uv} = \{p : \mathrm{dist}(p, u) + \mathrm{dist}(u, v) = \mathrm{dist}(p, v) \text{ or } \mathrm{dist}(u, p) + \mathrm{dist}(p, v) = \mathrm{dist}(u, v) \text{ or } \mathrm{dist}(u, v) + \mathrm{dist}(v, p) = \mathrm{dist}(u, p)\},$$

where dist is the Euclidean metric; in an arbitrary metric space (V, dist), the same relation may be taken for the definition of line $\overline{uv}$. The resulting family of lines may have strange properties: for instance, a line can be a proper subset of another [5, p. 2102]. Nevertheless, fragments of Euclidean geometry might translate to the framework of metric spaces. In particular, Chen and Chvátal asked:

($\star$) *True or false? Every metric space on n points, where $n \geq 2$, either has at least n distinct lines or else has a line that is universal in the sense of consisting of all n points.*

* Corresponding author.
E-mail addresses: pierreaboulker@gmail.com (P. Aboulker), adrian.bondy@sfr.fr (A. Bondy), gougle@gmail.com (X. Chen), chiniforooshan@alumni.uwaterloo.ca (E. Chiniforooshan), chvatal@cse.concordia.ca (V. Chvátal), sandy656692@gmail.com (P. Miao).
1 Canada Research Chair in Discrete Mathematics.

http://dx.doi.org/10.1016/j.dam.2014.02.008
0166-218X/

There is some evidence that the answer to (⋆) may be 'true'. For instance, Kantor and Patkós [11] proved that

- if no two of n points ($n \geq 2$) in the plane share their x- or y-coordinate, then these n points with the L_1 metric either induce at least n distinct lines or else they induce a universal line.

(For sets of n points in the plane that are allowed to share their coordinates, [11] provides a weaker conclusion: these n points with the L_1 metric either induce at least $n/37$ distinct lines or else they induce a universal line.) Chvátal [7] proved that

- every metric space on n points where $n \geq 2$ and each nonzero distance equals 1 or 2 either has at least n distinct lines or else has a universal line.

Every connected undirected graph induces a metric space on its vertex set, where $\text{dist}(u, v)$ is the usual graph-theoretic distance between vertices u and v (defined as the smallest number of edges in a path from u to v). It is easy to see that

- every metric space induced by a connected bipartite graph on n vertices, where $n \geq 2$, has a universal line.

A *chordal graph* is a graph that contains no induced cycle of length four or more. Beaudou, Bondy, Chen, Chiniforooshan, Chudnovsky, Chvátal, Fraiman, and Zwols [2] proved that

- every metric space induced by a connected chordal graph on n vertices, where $n \geq 2$ either has at least n distinct lines or else has a universal line.

A *distance-hereditary graph* is a graph G such that, for every connected induced subgraph H of G and for every pair u, v of vertices of H, the distance between u and v in H equals the distance between u and v in G. Aboulker and Kapadia [1] proved that

- every metric space induced by a connected distance-hereditary graph on n vertices, where $n \geq 2$, either has at least n distinct lines or else has a universal line.

Chiniforooshan and Chvátal [6] proved that

- every metric space induced by a connected graph on n vertices either has $\Omega(n^{2/7})$ distinct lines or else has a universal line;

recently, Pierre Aboulker, Rohan Kapadia, and Cathryn Supko improved the exponent 2/7 to 1/2. (Their manuscript is in preparation.)

A *hypergraph* is an ordered pair (V, H) such that V is a set and H is a family of subsets of V; elements of V are the *vertices* of the hypergraph and members of H are its *hyperedges*; a hypergraph is called *k-uniform* if each of its hyperedges consists of k vertices. The definition of lines in a metric space (V, dist) that was our starting point depends only on the 3-uniform hypergraph (V, H) where $H = \{\{a, b, c\} : \text{dist}(a, b) + \text{dist}(b, c) = \text{dist}(a, c)\}$: we have

$$\overline{uv} = \{u, v\} \cup \{p : \{u, v, p\} \in H\}.$$

Chen and Chvátal [5] proposed to take this relation for the definition of line $\overline{uv}$ in an arbitrary 3-uniform hypergraph (V, H). With this definition, the combinatorial theorem of De Bruijn and Erdős [8] can be stated as follows:

- if no four vertices in a 3-uniform hypergraph carry two or three hyperedges, then, except when one of the lines in this hypergraph is universal, the number of lines is at least the number of vertices and the two numbers are equal if and only if the hypergraph belongs to one of two simply described families.

Here, four vertices u, v, w, x are said to *carry* a hyperedge if this hyperedge is one of $\{u, v, w\}, \{u, v, x\}, \{u, w, x\}, \{v, w, x\}$; the assumption of the theorem can be rephrased by saying that every four vertices carrying at least one hyperedge carry precisely one hyperedge or four of them.

Beaudou, Bondy, Chen, Chiniforooshan, Chudnovsky, Chvátal, Fraiman, and Zwols [3] generalized the De Bruijn–Erdős theorem by allowing any four vertices to carry three hyperedges:

- if no four vertices in a 3-uniform hypergraph carry two hyperedges, then, except when one of the lines in this hypergraph is universal, the number of lines is at least the number of vertices and the two numbers are equal if and only if the hypergraph belongs to one of three simply described families.

In particular, if the 'metric space' in (⋆) is replaced by '3-uniform hypergraph where no four vertices carry two hyperedges', then the answer is 'true'. Without the assumption that no four vertices carry two hyperedges, the answer is 'false' [5, Theorem 3]: there are arbitrarily large 3-uniform hypergraphs where no line is universal and yet the number of lines is only $\exp(O(\sqrt{\ln n}))$. Nevertheless, even this variation of (⋆) can be answered 'true' if the desired lower bound on the number of lines is weakened enough [5, Theorem 4]:

- Every 3-uniform hypergraph with n vertices either has at least $\lg n + \frac{1}{2} \lg \lg n + \frac{1}{2} \lg \frac{\pi}{2} - o(1)$ distinct lines or else has a universal line.

(We follow the convention of letting lg stand for the logarithm to base 2.) The purpose of our note is to improve this lower bound by a factor of $2 - o(1)$.

All our hypergraphs are 3-uniform. We let V denote the vertex set, we let $\mathcal{L}$ denote the line set, and we write $n = |V|$, $m = |\mathcal{L}|$. The number of hyperedges, which we call *hedges*, is irrelevant to us. We assume throughout that no line is universal.

Let us define mappings $\alpha, \beta : V \to 2^{\mathcal{L}}$ by

$$\alpha(x) = \{L \in \mathcal{L} : x \in L\} \quad \text{and} \quad \beta(x) = \{\overline{xw} : w \neq x\}.$$

Note that $\beta(x) \subseteq \alpha(x)$ for all x. The proof of the lower bound

$$m \geq \lg n \tag{1}$$

in [5, Theorem 4] relies on the observation that α is one-to-one. This observation can be generalized as follows.

Lemma 1. *If $f : V \to 2^{\mathcal{L}}$ is a mapping such that $\beta(x) \subseteq f(x) \subseteq \alpha(x)$ for all x, then $x \neq y \to f(x) \not\subseteq f(y)$.*

Proof. We only need to prove that $\beta(x) - \alpha(y) \neq \emptyset$ whenever $x \neq y$. To do this, we use the assumption that $\overline{xy}$ is not universal: there is a point z such that $z \notin \overline{xy}$. This means that $\{x, y, z\}$ is not a hedge, and so $\overline{xz} \in \beta(x) - \alpha(y)$. □

Lemma 2. *If x, y, z are vertices such that $\overline{xy} = \overline{xz}$, then $\alpha(y) \cap \beta(x) = \alpha(z) \cap \beta(x)$.*

Proof. If $y \in \overline{xw}$, then $\{x, w, y\}$ is a hedge, and so $w \in \overline{xy} = \overline{xz}$, and so $\{x, z, w\}$ is a hedge, and so $z \in \overline{xw}$. □

We define the *span* of a subset S of V to be $\cup_{x \in S} \beta(x)$.

Lemma 3. *If $n \geq 2$ and a nonempty set of s vertices has a span of t lines, then*

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

Proof. Given a nonempty set of s vertices and its span T of t lines, enumerate the vertices in S as $x_1, x_2, \ldots, x_s$. Note that $t > 0$ (since $n \geq 2$ and $s > 0$) and define a mapping $\psi : (V - S) \to T^s$ by

$$\psi(v) = (\overline{x_1 v}, \overline{x_2 v}, \ldots, \overline{x_s v}).$$

If y, z are vertices in $V - S$ such that $\psi(y) = \psi(z)$, then Lemma 2 guarantees that $\alpha(y) \cap \beta(x_i) = \alpha(z) \cap \beta(x_i)$ for every x_i in S and so (since $T = \cup_{i=1}^{s} \beta(x_i)$) $\alpha(y) \cap T = \alpha(z) \cap T$. This and Lemma 1 (with $f = \alpha$) together imply that $\alpha(y) - T \neq \alpha(z) - T$ whenever $\psi(y) = \psi(z)$ and $y \neq z$. It follows that $|C| \leq 2^{m-t}$ for every subset C of $V - S$ on which ψ is constant. Since at least one of these sets C has at least $(n - s)/t^s$ points, we conclude that $(n - s)/t^s \leq 2^{m-t}$. □

Lemma 4. *For every positive ε, there is a positive δ such that*

$$\sum_{i < \delta N} \binom{N}{i} \leq 2^{\varepsilon N} \quad \textit{for all positive integers } N.$$

Proof. A special case of an inequality proved first by Bernstein [4,10] asserts that

$$\sum_{i=0}^{k} \binom{N}{i} \leq \left(\frac{N}{k}\right)^k \left(\frac{N}{N-k}\right)^{N-k} \quad \text{for all } k = 0, 1, \ldots, \lfloor N/2 \rfloor;$$

setting $x = k/(N - k)$ in the inequality $1 + x \leq e^x$, we find that

$$\left(\frac{N}{k}\right)^k \left(\frac{N}{N-k}\right)^{N-k} \leq \left(\frac{eN}{k}\right)^k.$$

Since $\lim_{t \to \infty}(1 + \ln t)/t = 0$, there is t_0 such that

$$t \geq t_0 \Rightarrow \frac{1 + \ln t}{t} \leq \varepsilon \ln 2,$$

and so

$$k \leq t_0^{-1} N \Rightarrow \left(\frac{eN}{k}\right)^k = \exp\left(N \cdot \frac{1 + \ln(N/k)}{N/k}\right) \leq 2^{\varepsilon N}. \quad \square$$

Theorem 1. $m \geq (2 - o(1)) \lg n$.

Proof. Given any positive ε, we will prove that $m \geq (2 - 4\varepsilon) \lg n$ for all sufficiently large n. To do this, let δ be as in Lemma 4 and consider a largest set S of vertices whose span T has at least $(0.5\delta \lg n) \cdot |S|$ lines (this S may be empty). Writing $s = |S|$ and $t = |T|$, we may assume that

$$t < 2 \lg n$$

(else we are done since $m \geq t$), and so $s < 4/\delta$. Now

$$m - t \geq (1 - o(1)) \lg n :$$

this follows from Lemma 3 when $t > 0$ and from (1) when $t = 0$. In turn, we may assume that

$$t \le 0.5m$$

(else $0.5m > m-t \ge (1-o(1)) \lg n$ and we are done). Finally, consider a largest set R of vertices such that $\beta(y)\cap T = \beta(z)\cap T$ whenever $y, z \in R$ and note for future reference that $|R| \ge n/2^t$. Since β is one-to-one (Lemma 1), all the sets $\beta(y) - T$ with $y \in R$ are distinct; by maximality of S, each of them includes less than $0.5\delta \lg n$ lines (else y could be added to S); it follows that (when n is large enough to make $0.5 \lg n$ less than $m - t$)

$$|R| \le \sum_{i<0.5\delta \lg n} \binom{m-t}{i} \le \sum_{i<\delta(m-t)} \binom{m-t}{i} \le 2^{\varepsilon(m-t)} \le 2^{\varepsilon m},$$

and so

$$n \le 2^t|R| \le 2^{t+\varepsilon m} \le 2^{(0.5+\varepsilon)m} \le 2^{m/(2-4\varepsilon)}. \quad \square$$

Acknowledgments

The work whose results are reported here began at a workshop held at Concordia University in April 2013. We are grateful to the Canada Research Chairs program for its generous support of this workshop. We also thank Laurent Beaudou, Nicolas Fraiman, and Cathryn Supko for their stimulating conversations during the workshop and the two anonymous referees for their thoughtful comments that helped to improve the presentation of our results.

References

[1] P. Aboulker, R. Kapadia, The Chen–Chvátal conjecture for metric spaces induced by distance-hereditary graphs, 2013, arXiv:1312.3214 [math.MG].
[2] L. Beaudou, A. Bondy, X. Chen, E. Chiniforooshan, M. Chudnovsky, V. Chvátal, N. Fraiman, Y. Zwols, A De Bruijn–Erdős theorem for chordal graphs, 2012, arXiv:1201.6376v1 [math.CO].
[3] L. Beaudou, A. Bondy, X. Chen, E. Chiniforooshan, M. Chudnovsky, V. Chvátal, N. Fraiman, Y. Zwols, Lines in hypergraphs, Combinatorica 33 (2013) 633–654.
[4] S. Bernstein, On a modification of Chebyshev's inequality and of the error formula of Laplace, in: Section Mathématique des Annales Scientifiques des Institutions Savantes de l'Ukraine, vol. 1, 1924, pp. 38–49 (in Russian).
[5] X. Chen, V. Chvátal, Problems related to a de Bruijn—Erdős theorem, Discrete Appl. Math. 156 (2008) 2101–2108.
[6] E. Chiniforooshan, V. Chvátal, A de Bruijn–Erdős theorem and metric spaces, Discrete Math. Theor. Comput. Sci. 13 (2011) 67–74.
[7] V. Chvátal, A de Bruijn–Erdős theorem for 1–2 metric spaces, Czechoslovak Math. J. (2012) arXiv:1205.1170 [math.CO] (in press).
[8] N.G. De Bruijn, P. Erdős, On a combinatorial problem, Indag. Math. 10 (1948) 421–423.
[9] P. Erdős, Three point collinearity, Amer. Math. Monthly 50 (1943). Problem 4065, p. 65. Solutions in Vol. 51 (1944), 169–171.
[10] W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Amer. Statist. Assoc. 58 (1963) 13–30.
[11] I. Kantor, B. Patkós, Towards a de Bruijn–Erdős theorem in the L_1-metric, Discrete Comput. Geom. 49 (2013) 659–670.