Discrete Comput Geom 19:427–435 (1998)

Discrete & Computational
Geometry

Canonical Theorems for Convex Sets*

J. Pach[1,2] and J. Solymosi[2]

[1]Courant Institute, NYU,
251 Mercer Street, New York, NY 10012, USA

[2]Mathematical Institute of the Hungarian Academy of Sciences,
Pf 127, H-1364 Budapest, Hungary
{pach,solymosi}@math-inst.hu

Communicated by Jacob E. Goodman

Abstract. Let $\mathcal{F}$ be a family of pairwise disjoint compact convex sets in the plane such that none of them is contained in the convex hull of two others, and let r be a positive integer. We show that $\mathcal{F}$ has r disjoint $\lfloor c_r n \rfloor$-membered subfamilies $\mathcal{F}_i$ $(1 \le i \le r)$ such that no matter how we pick one element F_i from each $\mathcal{F}_i$, they are in convex position, i.e., every F_i appears on the boundary of the convex hull of $\bigcup_{i=1}^{r} F_i$. (Here c_r is a positive constant depending only on r.) This generalizes and sharpens some results of Erdős and Szekeres, Bisztriczky and Fejes Tóth, Bárány and Valtr, and others.

1. Introduction

In their classical paper written in 1935, Erdős and Szekeres [ES1], [E] proved that for every $r \ge 3$, there exists an integer $f(r)$ such that any set of at least $f(r)$ points in the plane has r elements in convex position. This result has inspired a lot of research in combinatorial geometry and in Ramsey theory (see, e.g., [BDV], [GRS], [H], [PA], [TV], and [V]).

It follows that if n is much larger than $f(r)$, then every n-element point set P contains many r-tuples in convex position. For instance, Solymosi [S] showed that for a suitable constant $c_r > 0$, we can select a sequence of $c_r n$ distinct elements from P, whose any r consecutive members are in convex position. In the case $r = 4$, Nielsen [N] and, in general, Bárány and Valtr [BV] proved the following stronger result:

* The first author was supported by NSF Grant CCR-94-24398, PSC-CUNY Research Award 667339, OTKA-T-020914.

Theorem A. *For any fixed $r \geq 4$, there is a constant $c_r > 2^{-2^{6r}}$ satisfying the following condition*:

Every n-element point set P in general position in the plane has r pairwise disjoint subsets P_i $(1 \leq i \leq r)$ such that $|P_i| \geq \lfloor c_r n \rfloor$ and no matter how we pick one point from each P_i, they are in convex position.

This result provides a *canonical* way to find many convex r-gons in a sufficiently large point set in the plane.

Bisztriczky and Fejes Tóth [BF] found the following generalization of the Erdős–Szekeres theorem to families of pairwise disjoint compact convex sets in the plane. We say that such a family $\mathcal{F}$ is in *general position* if none of its members is contained in the convex hull of the union of two others. $\mathcal{F}$ is said to be in *convex position* if none of its members is contained in the convex hull of the union of the others.

Theorem B. *For every $r \geq 3$, there exists an integer $g(r)$ such that any family of at least $g(r)$ pairwise disjoint compact convex sets in general position in the plane has r members in convex position.*

In Section 3 of this paper, we apply a straightforward counting argument suggested in [S] to establish the following common generalization of Theorems A and B:

Theorem 1. *For every $r \geq 4$, there is a positive constant $c_r = 2^{-O(r^2)}$ with the following property*:

Every family $\mathcal{F}$ of n pairwise disjoint compact convex sets in general position in the plane has r disjoint $\lfloor c_r n \rfloor$-membered subfamilies $\mathcal{F}_i$ $(1 \leq i \leq r)$ such that no matter how we pick one set from each $\mathcal{F}_i$, they are always in convex position.

It is worth mentioning that in the special case when all members of $\mathcal{F}$ are single points, our proof shows that the statement of Theorem A is true with a much better constant c_r than the one given in [BV].

The proofs of the next three theorems follow the same scheme.

A polygonal path $p_1 p_2 \ldots p_r$ in the plane or in space, is called *ε-straight* if $\angle p_{i-1} p_i p_{i+1} > \pi - \varepsilon$, $1 < i < r$ (see [ES2] and [P]). The *length* of a polygonal path is the number of its *vertices*.

Theorem 2. *For every $d \geq 2, r \geq 3$, and $\varepsilon > 0$, there exists a positive constant $c = c^d_{r,\varepsilon}$ with the following property*:

Every n-element point set P in general position in Euclidean d-space has r pairwise disjoint subsets P_i $(1 \leq i \leq r)$ with at least $\lfloor cn \rfloor$ elements such that no matter how we pick a point from each P_i, they always form an ε-straight polygonal path.

Theorem 3. *For every $r, s \geq 2$, there exists a positive constant $c = c_{r,s} = (rs)^{-O(r)}$ with the following property*:

Let $\mathcal{F}$ be a family of n compact convex sets in the plane, no s of which are pairwise intersecting. Then $\mathcal{F}$ has r disjoint $\lfloor cn \rfloor$-membered subfamilies $\mathcal{F}_i$ $(1 \leq i \leq r)$ such that no two sets belonging to distinct subfamilies have a point in common.

Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. For any positive integer r, let $G(r)$ denote the graph obtained from G by replacing each vertex $v \in V(G)$ by r vertices, v_i $(1 \le i \le r)$, and connecting v_i and u_j by an edge if and only if $vu \in E(G)$ $(1 \le i, j \le r)$.

Theorem 4. *For every $c > 0, r \ge 1$, there exists a constant $c_r > 0$ with the following property*:

Let T be any tree of at most $c_r n$ vertices. Then every graph G with n vertices and at least cn^2 edges has a subgraph isomorphic to $T(r)$.

The letters $c, c_r, c_{r,\varepsilon}$, etc., appearing in different theorems denote unrelated positive constants depending on r, ε, etc.

2. Proofs of Theorems 2 and 3

To establish Theorem 2, we need the following straightforward generalization of a result from [ES2]:

Lemma 2.1. *There exists a constant $c > 0$ such that any set of at least $k^{(c/\varepsilon)^{d-1}}$ points in Euclidean d-space has k elements that form an ε-straight polygonal path of length k.*

Proof of Theorem 2. Let ε, d, and r be fixed, and set $k = 2r - 1$. By Lemma 2.1, there exists an integer $K = K(\varepsilon, d, r)$ such that any set of K points in d-space contains k elements that form the vertex set of an $\varepsilon/3$-straight polygonal path Π. Notice that if we skip every other vertex of Π, then we obtain a polygonal path Π' with r vertices, which is ε-straight. The sequence formed by the $r - 1$ vertices we skipped is called the *support* of Π.

Consider now any set P of n points in the plane. Clearly, P contains at least

$$\binom{n}{K}\Big/\binom{n-k}{K-k} = \binom{n}{k}\Big/\binom{K}{k}$$

different $\varepsilon/3$-straight polygonal paths of length k, and at least

$$\frac{\binom{n}{k}\Big/\binom{K}{k}}{n!/(n-r+1)!} > \frac{n^r}{K^{2r-1}}$$

of them must share the same support S.

Let P_i denote the set of all elements of P that occur as the $(2i - 1)$st vertex in some $\varepsilon/3$-straight polygonal path of length k, whose support is S $(1 \le i \le r)$. These sets meet the requirements of the theorem. In particular, for every i, we have

$$|P_i| > \frac{n^r}{K^{2r-1}}\Big/\prod_{j\ne i}|P_j| > \frac{n}{K^{2r-1}}. \qquad \square$$

The proof of Theorem 3 uses the same idea. We need a little preparation.

Let $\mathcal{F}$ be a family of n compact convex sets in the plane. Assume without loss of generality that no two members of $\mathcal{F}$ have a common vertical tangent line. For $C \in \mathcal{F}$, let $\pi(C)$ denote the projection of C onto the x-axis. Following [LMPT], we define four partial orders, $\prec_1$, $\prec_2$, $\prec_3$, and $\prec_4$, on $\mathcal{F}$. For any two disjoint sets $A, B \in \mathcal{F}$:

1. $A \prec_1 B$ if $\pi(A) \subseteq \pi(B)$ and A lies below B ("below" means in the y-axis direction).
2. $A \prec_2 B$ if $\pi(A) \subseteq \pi(B)$ and A lies above B.
3. $A \prec_3 B$ if the left endpoint of $\pi(B)$ is to the right of the left endpoint of $\pi(A)$, the right endpoint of $\pi(B)$ is to the right of the right endpoint of $\pi(A)$, and in the part where $\pi(A)$ and $\pi(B)$ overlap (if any), A lies above B.
4. $A \prec_4 B$ if the left endpoint of $\pi(B)$ is to the right of the left endpoint of $\pi(A)$, the right endpoint of $\pi(B)$ is to the right of the right endpoint of $\pi(A)$, and in the part where $\pi(A)$ and $\pi(B)$ overlap (if any), A lies below B.

Lemma 2.2 [LMPT]. *Any family of more than $(k-1)^4(s-1)$ compact convex sets in the plane, no s of which have pairwise nonempty intersections, contains k members that form a chain with respect to one of the relations $\prec_1, \prec_2, \prec_3, \prec_4$.*

Proof of Theorem 3. Setting $K = (k-1)^4(s-1)+1$ and $k = 2r-1$, we obtain, just as in the previous proof, that there exists $1 \le j \le 4$ such that $\mathcal{F}$ has at least $\frac{1}{4}\binom{n}{k}\big/\binom{K}{k}$ chains $\mathcal{C}$ of length k with respect to $\prec_j$. If we skip every other element of $\mathcal{C}$, we obtain a chain $\mathcal{C}'$ of length r. The chain $\mathcal{C}\backslash\mathcal{C}'$ is called the *support* of $\mathcal{C}$. It follows that at least

$$\frac{\frac{1}{4}\binom{n}{k}\Big/\binom{K}{k}}{\binom{n}{r-1}} > \frac{n^r}{K^{2r-1}}$$

chains C share the same support S.

For every i $(1 \le i \le r)$, let $\mathcal{F}_i$ denote the set of all members of $\mathcal{F}$ that occur as the $(2i-1)$st smallest element of a chain in $(\mathcal{F}, \prec_j)$ with length k and support S. It is clear that no two sets belonging to distinct $\mathcal{F}_i$'s have a point in common. The same estimation as at the end of the proof of Theorem 2 gives that $|\mathcal{F}_i| \ge n/K^{2r-1}$ for every i. □

It is possible that the following far-reaching generalization of Theorem 3 is also true. For every $s \ge 2$, there exists a constant $c = c_s > 0$ with the property that any family of n compact connected sets in the plane, no s of which have pairwise nonempty intersections, has at least cn pairwise disjoint members. We have been unable to decide whether this statement holds for families of straight-line segments.

3. Proof of Theorem 1

We follow the same approach as in the previous section. The proof is based on a stronger version of Theorem B.

Lemma 3.1 [PT]. *For every $k \geq 3$, any family of 2^{4k} pairwise disjoint compact convex sets in general position in the plane has k members in convex position.*

Let $\mathcal{F}$ be a family of n pairwise disjoint compact convex sets in general position in the plane. Assume without loss of generality that no three members of $\mathcal{F}$ have a common tangent line and no two have a common vertical tangent.

Applying first Lemma 2.2 and then Lemma 3.1, we obtain that for every k, any 2^{16k}-membered subfamily of $\mathcal{F}$ contains k sets in convex position that form a chain with respect to one of the relations $\prec_j$ $(1 \leq j \leq 4)$.

Set $k = 4r-2$, $K = 2^{16k}$. Just as in the proof of Theorem 3, it follows that there exists $1 \leq j \leq 4$ such that $\mathcal{F}$ has at least $\frac{1}{4}\binom{n}{k}/\binom{K}{k}$ chains $\mathcal{B} = (B_1 \prec_j B_2 \prec_j \cdots \prec_j B_k)$, whose members are in convex position. We distinguish two substantially different cases according to the value of j.

Case 1: $j = 1$.
Let $\mathcal{B}$ be any chain of length $k = 4r - 2$ with respect to $\prec_1$, whose members are in convex position. Then $\mathcal{B}$ has a subchain $\mathcal{C} = (C_1, C_2, \ldots, C_{2r-1})$ with the following property. Each C_i contributes to the boundary of the convex hull conv $\bigcup_{i=1}^{2r-1} C_i$ at least one point to the left of C_1, or each C_i contributes to bd conv $\bigcup_{i=1}^{2r-1} C_i$ at least one point to the right of C_1. In the former case we call $\mathcal{C}$ a *left-convex* chain and in the latter case a *right-convex* chain. Thus, there are at least

$$\frac{\frac{1}{4}\binom{n}{k}\Big/\binom{K}{k}}{2\binom{n-2r+1}{2r-1}}$$

different chains $\mathcal{C} = (C_1 \prec_1 C_2 \prec_1 \cdots \prec_1 C_{2r-1})$ of the same type, say, left-convex. Define the *support* of $\mathcal{C}$ as the subchain $\mathcal{C}^* \subseteq \mathcal{C}$ formed by the even-numbered elements, i.e., let

$$\mathcal{C}^* = (C_2, C_4, C_6, \ldots, C_{2r-2}).$$

Clearly, there are at least

$$\frac{\frac{1}{4}\binom{n}{k}\Big/\binom{K}{k}}{2\binom{n-2r+1}{2r-1}\binom{n}{r-1}} > \frac{n^r}{K^{4r-2}}$$

different left-convex chains $\mathcal{C} = (C_1, C_2, \ldots, C_{2r-1})$ which have the same support. These chains are called *standard*. Let

$$(\bar{C}_2, \bar{C}_4, \bar{C}_6, \ldots, \bar{C}_{2r-2})$$

denote the common support of the standard chains. We will refer to this sequence as the *standard support*.

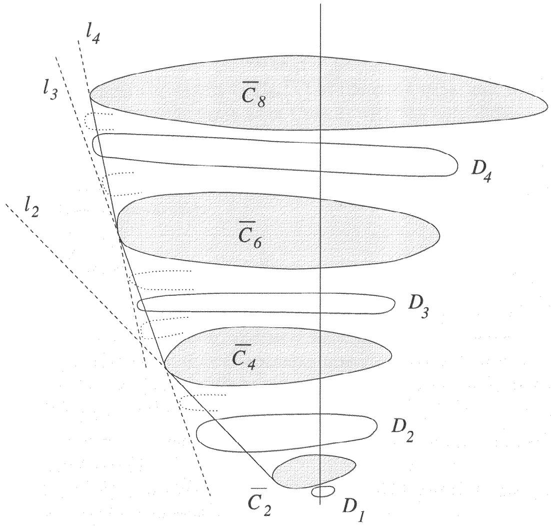

Fig. 1

For any t $(1 \le t \le r)$, let $\mathcal{F}_t$ denote the family of all members of $\mathcal{F}$ that occur as the $(2t-1)$st element $C_{2t-1} \in \mathcal{C}$ for some standard chain $\mathcal{C}$. We have

$$|\mathcal{F}_t| > \frac{n^r}{K^{4r-2}} \Big/ \prod_{s \neq t} |\mathcal{F}_s| > \frac{n}{K^{4r-2}}.$$

It remains to show that for every choice $D_t \in \mathcal{F}_t$ $(1 \le t \le r)$, the sets $D_1, D_2, \ldots, D_r$ are in convex position (see Fig. 1). To see this, consider the left-hand side ∂ of the boundary of the union of all members of the standard support. ∂ consists of nonempty portions of the boundaries of the sets $\bar{C}_{2t}$ $(1 \le t < r)$, separated by straight-line segments. For any $1 < t < r$, let l_t denote the common tangent line of the sets $\bar{C}_{2t-2}$ and $\bar{C}_{2t}$ with the property that every other member of the standard support is on its right-hand side. To finish the proof in Case 1, it is sufficient to notice that every $D_t \in \mathcal{F}_t$ has at least one point to the left of l_t, while all members of $\bigcup_{s \neq t} \mathcal{F}_s$ lie on the right-hand side of l_t. Therefore, $D_1, D_2, \ldots, D_r$ are in convex position.

Case 2: $j = 3$.
Let $\mathcal{B}$ be any chain of length $k = 4r - 2$ with respect to $\prec_3$, whose members are in convex position. Then $\mathcal{B}$ has a subchain $\mathcal{C} = (C_1, C_2, \ldots, C_{2r-1})$ with the following property. Each C_i contributes at least one point to the upper portion of bd conv $\bigcup_{i=1}^{2r-1} C_i$, or each C_i contributes at least one point to the lower portion of bd conv $\bigcup_{i=1}^{2r-1} C_i$. In the former case, $\mathcal{C}$ is called a *upper-convex* chain, and in the latter one, a *lower-convex* chain. Thus,

there are at least

$$\frac{\frac{1}{4}\binom{n}{k}\Big/\binom{K}{k}}{2\binom{n-2r+1}{2r-1}}$$

different chains $\mathcal{C} = (C_1 \prec_3 C_2 \prec_3 \cdots \prec_3 C_{2r-1})$ of the same type, say, upper-convex. The rest of the argument is exactly the same as in Case 1, with the only difference that in place of the left-hand side ∂ of the boundary of the union of all members of the standard support, we have to consider its *upper side*.

The cases $j = 2$ and $j = 4$ are symmetric counterparts of the above two cases, so they do not have to be treated separately.

4. Proof of Theorem 4

Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. Assume that $|V(G)| = n$ and $|E(G)| \geq cn^2$ for some constant $c > 0$, and let r be a fixed positive integer.

First, we would like to show that G contains many complete bipartite subgraphs $K_{r,r}$ with r vertices in its classes. The proof is based on the following simple statement, discovered by Erdős, which is a weak version of a result by [KST].

Lemma 4.1. *For every $r \geq 1$ and every $\gamma > 0$, there exists a positive integer $n_0 = n_0(r, \gamma)$ with the following property*:

Every graph G_0 with n_0 vertices and at least γn_0^2 edges contains a complete bipartite subgraph $K_{r,r}$ with r vertices in each of its classes.

Let x denote the number of n_0-element subsets of $V(G)$ which induce a subgraph of G with at least γn_0^2 edges. Then we have

$$x\binom{n_0}{2} + \left(\binom{n}{n_0} - x\right)\gamma n_0^2 > |E(G)|\binom{n-2}{n_0-2} \geq cn^2\binom{n-2}{n_0-2},$$

which yields

$$x > \binom{n}{n_0}\frac{c(n_0-1) - \gamma n_0}{(n_0-1)/2 - \gamma n_0}.$$

Thus, for $\gamma = c/2$, $n_0 > 2$, we obtain $x > \frac{c}{2}\binom{n}{n_0}$.

Set $n_0 = n_0(r, \gamma) = n_0(r, c/2)$. By Lemma 4.1, every subgraph of G with n_0 vertices and at least $(c/2)n_0^2$ edges contains at least one copy of $K_{r,r}$. Thus, the number y of complete bipartite subgraphs $K_{r,r}$ of G satisfies

$$y \geq \frac{x}{\binom{n-2r}{n_0-2r}} > \frac{\frac{c}{2}\binom{n}{n_0}}{\binom{n-2r}{n_0-2r}} > \frac{cn^{2r}}{2n_0^{2r}}.$$

Suppose for simplicity that n is divisible by r, and consider all possible partitions of $V(G)$ into classes of size r. The number of these partitions is

$$p(n,r) = \frac{\binom{n}{r}\binom{n-r}{r}\binom{n-2r}{r}\cdots\binom{r}{r}}{(n/r)!}.$$

For every partition P, construct a graph $G(P)$ whose vertices are the classes V_i $(1 \le i \le n/r)$ of the partition, and two vertices V_i and V_j are connected by an edge of $G(P)$ if and only if G contains all edges running between them. By averaging over all partitions, we find that there exists a P such that the number of edges of $G(P)$ is at least

$$\frac{yp(n-2r,r)}{p(n,r)} > \frac{cn^{2r}}{2n_0^{2r}} \frac{\binom{\frac{n}{r}}{r}\binom{\frac{n-r}{r}}{}}{\binom{n}{r}\binom{n-r}{r}} > c\left(\frac{r}{e}\right)^{2r} \cdot \left(\frac{n}{r}\right)^2.$$

We can now finish the proof of the theorem by applying to $G(P)$ the following simple assertion, whose proof is left to the reader.

Lemma 4.2. *For any $C > 0$, every graph with N vertices and at least CN^2 edges contains every tree of at most $CN/2$ vertices as a subgraph.*

Hence, Theorem 4 is true with $c_r = (c/2)(r/e)^{2r}$.

References

[BV] I. Bárány and P. Valtr, A positive fraction Erdős–Szekeres theorem, *Discrete and Computational Geometry*, this issue, pp. 335–342.

[BDV] A. Bialostocki, P. Dierker, and B. Voxman, Some notes on the Erdős–Szekeres theorem, *Discrete Mathematics*, **91** (1991), 231–238.

[BF] T. Bisztriczky and G. Fejes Tóth, A generalization of the Erdős–Szekeres convex *n*-gon theorem, *Journal reine und angewandte Mathematik*, **395** (1989), 167–170.

[E] P. Erdős, *The Art of Counting*, *Selected Writings* (J. Spencer, ed.), MIT Press, Cambridge, MA, 1973.

[ES1] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Mathematica*, **2** (1935), 463–470.

[ES2] P. Erdős and G. Szekeres, On some extremum problems in elementary geometry, *Annales Universitatis Scientiarum Budapestinensis*, *Eötvös*, *Sectio Mathematica*, **III–IV** (1960–61), 53–62.

[GRS] R. Graham, B. Rothschild, and J. Spencer, *Ramsey Theory* (2nd edn), Wiley-Interscience, New York, 1990.

[H] J. D. Horton, Sets with no empty 7-gons, *Canadian Mathematical Bulletin*, **26** (1983), 482–484.

[KST] T. Kővári, V. T. Sós, and P. Turán, On a problem of K. Zarankiewicz, *Colloquium Mathematicum*, **3** (1954), 50–57.

[LMPT] D. Larman, J. Matoušek, J. Pach, and J. Törőcsik, A Ramsey-type result for planar convex sets, *Bulletin of the London Mathematical Society*, **26** (1994), 132–136.

[N] M. J. Nielsen, Transverse matchings on a finite planar set (Manuscript), University of Idaho, Moscow, 1995.

[P] J. Pach, A remark on transversal numbers, In: *The Mathematics of Paul Erdős*, Vol. II (R. Graham and J. Nešetřil, eds.), Springer-Verlag, Heidelberg, 1997, pp. 310–317.

[PA] J. Pach and P. K. Agarwal, *Combinatorial Geometry*, Wiley-Interscience, New York, 1995.

[PT] J. Pach and G. Tóth, A generalization of the Erdős–Szekeres theorem to disjoint convex sets, *Discrete and Computational Geometry*, this issue, pp. 437–445.

[SA] M. Sharir and P. K. Agarwal, *Davenport–Schinzel Sequences and Their Geometric Applications*, Cambridge University Press, Cambridge, 1995.

[S] J. Solymosi, *Combinatorial Problems in Finite Ramsey Theory*, Master's Thesis, Eötvös University, Budapest, 1988.

[TV] G. Tóth and P. Valtr, Note on the Erdős–Szekeres theorem, *Discrete and Computational Geometry*, this issue, pp. 457–459.

[V] P. Valtr, Several Results Related to the Erdős–Szekeres Theorem (Doctoral Dissertation), Department of Applied Mathematics, Charles University, Prague, 1996.

Received April 30, 1997, *and in revised form August* 5, 1997.