Available online at www.sciencedirect.com

SciVerse ScienceDirect

Advances in Mathematics 244 (2013) 1–15

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

Higher-order Erdős–Szekeres theorems☆

Marek Eliáš[a,*], Jiří Matoušek[a,b]

[a] *Department of Applied Mathematics, Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic*
[b] *Institute of Theoretical Computer Science, ETH Zurich, 8092 Zurich, Switzerland*

Received 13 December 2011; accepted 29 April 2013
Available online 29 May 2013

Communicated by Erwin Lutwak

Abstract

Let $P = (p_1, p_2, \ldots, p_N)$ be a sequence of points in the plane, where $p_i = (x_i, y_i)$ and $x_1 < x_2 < \cdots < x_N$. A famous 1935 Erdős–Szekeres theorem asserts that every such P contains a monotone subsequence S of $\lceil \sqrt{N} \rceil$ points. Another, equally famous theorem from the same paper implies that every such P contains a convex or concave subsequence of $\Omega(\log N)$ points.

Monotonicity is a property determined by pairs of points, and convexity concerns triples of points. We propose a generalization making both of these theorems members of an infinite family of Ramsey-type results. First we define a $(k+1)$-tuple $K \subseteq P$ to be *positive* if it lies on the graph of a function whose kth derivative is everywhere nonnegative, and similarly for a *negative* $(k+1)$-tuple. Then we say that $S \subseteq P$ is *kth-order monotone* if its $(k+1)$-tuples are all positive or all negative.

We investigate a quantitative bound for the corresponding Ramsey-type result (i.e., how large kth-order monotone subsequence can be guaranteed in every N-point P). We obtain an $\Omega(\log^{(k-1)} N)$ lower bound ($(k-1)$-times iterated logarithm). This is based on a quantitative Ramsey-type theorem for *transitive colorings* of the complete $(k+1)$-uniform hypergraph (these were recently considered by Pach, Fox, Sudakov, and Suk).

For $k = 3$, we construct a geometric example providing an $O(\log \log N)$ upper bound, tight up to a multiplicative constant. As a consequence, we obtain similar upper bounds for a Ramsey-type theorem for *order-type homogeneous* subsets in $\mathbb{R}^3$, as well as for a Ramsey-type theorem for hyperplanes in $\mathbb{R}^4$ recently used by Dujmović and Langerman.

Keywords: Erdős–Szekeres theorem; Ramsey theory; Hypergraph; Order type; Divided difference; kth-order monotone subset

☆ Extended abstract appeared in [5].

* Corresponding author.
E-mail addresses: elias@kam.mff.cuni.cz (M. Eliáš), matousek@kam.mff.cuni.cz (J. Matoušek).

0001-8708/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.aim.2013.04.020

1. Introduction

In this paper we mainly consider sets $P = \{p_1, p_2, \dots, p_N\}$ of points in the plane, where $p_i = (x_i, y_i)$. We always assume that no two of the x-coordinates coincide, and unless stated otherwise, we also assume that the p_i are numbered so that $x_1 < x_2 < \dots < x_N$ (the same also applies to subsets of P, which we will enumerate in the order of increasing x-coordinates).

Two theorems of Erdős and Szekeres. Among simple results in combinatorics, only few can compete with the following one in beauty and usefulness:

Theorem 1.1 (*Erdős–Szekeres on Monotone Subsequences [7]*). *For every positive integer n, among every $N = (n-1)^2 + 1$ points $p_1, \dots, p_N \in \mathbb{R}^2$ as above, one can always choose a monotone subset of at least n points, i.e., indices $i_1 < i_2 < \dots < i_n$ such that either $y_{i_1} \le y_{i_2} \le \dots \le y_{i_n}$ or $y_{i_1} \ge y_{i_2} \ge \dots \ge y_{i_n}$.*

See, for example, Steele [17] for a collection of six nice proofs and some applications. For many purposes, it is more natural to view the above theorem as a purely combinatorial result about permutations, but here we prefer the geometric formulation (which is also similar to the one in the original Erdős–Szekeres paper).

Another result of the same paper of Erdős and Szekeres is the following well-known gem in discrete geometry:[1]

Theorem 1.2 (*Erdős–Szekeres on Convex/Concave Configurations [7]*). *For every positive integer n, among every $N = \binom{2n-4}{n-2} + 1 \approx 4^n/\sqrt{n}$ points $p_1, \dots, p_N \in \mathbb{R}^2$ as above, one can always choose a convex configuration or a concave configuration of n points, i.e., indices $i_1 < i_2 < \dots < i_n$ such that the slopes of the segments $p_{i_j} p_{i_{j+1}}$, $j = 1, 2, \dots, n-1$, are either monotone nondecreasing or monotone nonincreasing.*

See, e.g., [13,11] for proofs and surveys of developments around this result.

k-general position. To simplify our forthcoming discussion, at some places it will be convenient to assume that the considered point sets are in a "sufficiently general" position. Namely, we define a set P to be in *k-general position* if no $k+1$ points of P lie on the graph of a polynomial of degree at most $k-1$. In particular, 1-general position requires that no two y-coordinates coincide, and 2-general position means the usual general position, i.e., no three points collinear.

kth-order monotone subsets. Here we propose a view of Theorems 1.1 and 1.2 as the first two members in an infinite sequence of Ramsey-type results about planar point sets.[2]

In Theorem 1.1, monotonicity of a subset is a property of *pairs* of points of the subset, and actually, it suffices to look at pairs of consecutive points. Similarly, convexity or concavity of a configuration in Theorem 1.2 is a property of triples, and again it is enough to look at consecutive triples.

In the former case, we are considering the slope of the segment determined by a pair of points, which can be thought of as the first derivative. In the latter case, a triple is convex iff its

[1] Somewhat unfortunately, the name Erdős–Szekeres theorem refers to Theorem 1.1 in some sources and to Theorem 1.2 or similar statements in other sources.

[2] There is also a (trivial) 0th member, namely, the statement that in every P, at least half of the points either have all y-coordinates nonnegative or have or all y-coordinates nonpositive.

points lie on the graph of a smooth convex function, i.e., one with nonnegative second derivative everywhere.

With this point of view, it is natural to define a $(k+1)$-tuple $K \subseteq P$ to be *positive* if it lies on the graph of a function whose k-th derivative (exists and) is everywhere nonnegative, and similarly for a *negative* $(k+1)$-tuple (in Section 2, we will provide several other, equivalent characterizations of these properties). Then we say that an arbitrary subset $S \subseteq P$ is *kth-order monotone* if its $(k+1)$-tuples are all positive or all negative.

First-order monotonicity is obviously equivalent to monotonicity as in Theorem 1.1, and second-order monotonicity is equivalent to convexity/concavity as in Theorem 1.2. We will also see (Lemma 2.5) that, to certify kth-order monotonicity, it is enough to consider all $(k+1)$-tuples of *consecutive* points.

Let us remark that every $(k+1)$-tuple K is positive or negative, and moreover, if K is in k-general position, it cannot be both positive and negative (Corollary 2.3). We will write $\mathrm{sgn}(K) = +1$ if K is positive and $\mathrm{sgn}(K) = -1$ if K is negative.

Ramsey's theorem, quantitative bounds, and transitive colorings. Using the just mentioned facts, one can immediately derive a Ramsey-type theorem for kth-order monotone subsets from Ramsey's theorem.

Proposition 1.3. *For every k and n there exists N such that every N-point planar set in k-general position contains an n-point kth-order monotone subset.*

Proof. We recall Ramsey's theorem (for two colors; see, e.g., Graham, Rothschild, and Spencer [10]): for every ℓ and n there exists N such that for every coloring of the set $\binom{X}{\ell}$ of all ℓ-element subsets of an N-element set X there exists an n-element *homogeneous* set $Y \subseteq X$, i.e., a subset in which all ℓ-tuples have the same color. The smallest N for which the claim holds is usually denoted by $R_\ell(n)$.

In our case, we set $X = P$ and color each $(k+1)$-tuple $K \subseteq P$ with the color $\mathrm{sgn}(K) \in \{\pm 1\}$. Then homogeneous subsets are exactly kth-order monotone subsets. $\square$

Let us denote by $\mathrm{ES}_k(n)$ the smallest value of N for which the claim in this proposition holds. We have $\mathrm{ES}_1(n) \le (n-1)^2 + 1$ and $\mathrm{ES}_2(n) \le \binom{2n-4}{n-2} + 1$ according to Theorems 1.1 and 1.2, respectively; moreover, these inequalities actually hold with equality [7]. Our main goal is to estimate the order of magnitude of $\mathrm{ES}_k(n)$ for $k \ge 3$.

The above proof gives $\mathrm{ES}_k(n) \le R_{k+1}(n)$. However, for $k = 1$, and most likely for all k, the order of magnitude of $R_{k+1}(n)$ is much larger than that of $\mathrm{ES}_k(n)$. Indeed, considering k fixed and n large, the best known lower and upper bounds on $R_{k+1}(n)$ are of the form[3] $R_2(n) = 2^{\Theta(n)}$ and, for $k \ge 2$,

$$\mathrm{twr}_k(\Omega(n^2)) \le R_{k+1}(n) \le \mathrm{twr}_{k+1}(O(n)),$$

where the tower function $\mathrm{twr}_k(x)$ is defined by $\mathrm{twr}_1(x) = x$ and $\mathrm{twr}_{i+1}(x) = 2^{\mathrm{twr}_i(x)}$. It is widely believed that the upper bound is essentially the truth. This belief is supported by known bounds for more than two colors, where the lower bound for $(k+1)$-tuples is also a tower of height $k+1$;

[3] We employ the usual asymptotic notation for comparing functions: $f(n) = O(g(n))$ means that $|f(n)| \le C|g(n)|$ for some C and all n, where C may depend on parameters declared as constants (in our case on k); $f(n) = \Omega(g(n))$ is equivalent to $g(n) = O(f(n))$; and $f(n) = \Theta(g(n))$ means that both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

see Conlon, Fox, and Sudakov [3] for a recent improvement and more detailed overview of the known bounds.

The coloring of the $(k+1)$-tuples in the above proof of Proposition 1.3 is not arbitrary. In particular, it has a property we call *transitivity* (see Lemma 2.5). Transitive colorings were introduced earlier in the recent preprint Fox et al. [8, Section 6], under the same name.

To define a transitive coloring in general, we need to consider a hypergraph whose vertex set is linearly ordered; without loss of generality, we can identify it with the set $[N] := \{1, 2, \ldots, N\}$. A coloring $c\colon \binom{[N]}{\ell} \to [m]$ is *transitive* if, for every $i_1, \ldots, i_{\ell+1} \in [N], i_1 < \cdots < i_{\ell+1}$, whenever the ℓ-tuples $\{i_1, \ldots, i_\ell\}$ and $\{i_2, \ldots, i_{\ell+1}\}$ have the same color, then *all* ℓ-element subsets of $\{i_1, \ldots, i_{\ell+1}\}$ have the same color. Let $R_\ell^{\text{trans}}(n)$ denote the Ramsey number for transitive colorings, i.e., the smallest N such that any transitive coloring of the complete ℓ-uniform hypergraph on $[N]$ contains an n-element homogeneous subset. We have the following bound.[4]

Theorem 1.4. *For $k = 1, 2$, we have $R_{k+1}^{\text{trans}}(n) = \mathrm{ES}_k(n)$, and for every fixed $k \geq 3$,*

$$\mathrm{ES}_k(n) \leq R_{k+1}^{\text{trans}}(n) \leq \mathrm{twr}_k(O(n)).$$

We note that Fox et al. [8] proved the slightly weaker upper bound $R_{k+1}^{\text{trans}}(n) \leq \mathrm{twr}_k(O(n \log n))$.

The proof of Theorem 1.4 is given in Section 3. The inequality $\mathrm{ES}_k(n) \leq R_{k+1}^{\text{trans}}(n)$ is clear (since every N-point set in k-general position provides a transitive coloring of $\binom{[N]}{k+1}$). The upper bounds for $R_2^{\text{trans}}(n)$ and $R_3^{\text{trans}}(n)$ follow by translating the proofs of Theorems 1.1 and 1.2 to the setting of transitive colorings almost word by word, and they are contained in [8]. The upper bound on $R_{k+1}^{\text{trans}}(n)$ is then obtained by induction on k, with $k = 3$ as the base case, following one of the usual proofs of Ramsey's theorem.

A set with no large third-order monotone subsets. For $k \leq 2$, the numbers $\mathrm{ES}_k(n)$ (and thus $R_{k+1}^{\text{trans}}(n)$) are known exactly. Our perhaps most interesting result is an asymptotically matching lower bound for $\mathrm{ES}_3(n)$.

Theorem 1.5. *For all $n \geq 2$ we have $R_4^{\text{trans}}(2n+1) \geq \mathrm{ES}_3(2n+1) \geq 2^{2^{n-1}} + 1$. Consequently, $\mathrm{ES}_3(n) = 2^{2^{\Theta(n)}}$.*

The proof is given in Section 4. A Ramsey function with known doubly exponential growth seems to be rare in geometric Ramsey-type problems (a notable example is a result of Valtr [18]).

Order types. Here we change the setting from the plane to $\mathbb{R}^d$ and we consider an ordered sequence $P = (p_1, p_2, \ldots, p_N)$ in $\mathbb{R}^d$. This time we do *not* assume the first coordinates to be increasing. For simplicity, we assume P to be in general position, which now means that no $d+1$ points of P lie on a common hyperplane.

We recall that *order type* of P specifies the orientation of every $(d+1)$-tuple of points of P, and it this way, it describes purely combinatorially many of the geometric properties of P. More formally, the order type of P is the mapping $\chi\colon \binom{[N]}{d+1} \to \{-1, +1\}$, where for a $(d+1)$-tuple $I = \{i_1, \ldots, i_{d+1}\}, i_1 < i_2 < \cdots < i_{d+1}, \chi(I) := \operatorname{sgn}\det M(p_{i_1}, p_{i_2}, \ldots, p_{i_{d+1}})$, where

[4] By inspecting the proof of the next theorem, it is easy to verify that the transitivity condition is not used in full strength—it suffices to assume only that the subsets obtained by omitting one of i_2, i_3 have the same color.

$M(q_1, \ldots, q_{d+1})$ is the $(d+1) \times (d+1)$ matrix whose jth column is $(1, q_j)$, i.e., 1 followed by the vector of the d coordinates of q_j. See, e.g., Goodman and Pollack [9] or [11] for more background about order types.

From Ramsey's theorem for $(d+1)$-tuples, we can immediately derive a Ramsey-type result for order types: for every d and n there exists N such that every N-point sequence contains an n-point subsequence in which all the $(d+1)$-tuples have the same orientation (we call such a subsequence *order-type homogeneous*). Let us write $\mathrm{OT}_d(n)$ for the smallest such N.

In Section 5 we first observe that, by simple and probably well known considerations, $\mathrm{OT}_1(n) = (n-1)^2 + 1$ and $\mathrm{OT}_2(n) = 2^{\Theta(n)}$. For $d \geq 3$, the upper bound for $\mathrm{OT}_d(n)$ from the Ramsey argument above is $\mathrm{OT}_d(n) \leq R_{d+1}(n) \leq \mathrm{twr}_{d+1}(O(n))$. In particular, for $\mathrm{OT}_3(n)$ this upper bound is triply exponential; in Section 5 we prove a doubly exponential lower bound. A recent paper by Conlon et al. [2] provides a doubly exponential upper bound for $\mathrm{OT}_3(n)$; see below.

Proposition 1.6. *For all d and n, $\mathrm{OT}_d(n) \geq \mathrm{ES}_d(n)$. In particular,* $\mathrm{OT}_3(n) = 2^{2^{\Omega(n)}}$.

A Ramsey-type result for hyperplanes. Let us consider a finite set H of hyperplanes in $\mathbb{R}^d$ in general position (every d intersecting at a single point). Let us say that H is *one-sided* if $V(H)$, the vertex set of the arrangement of H, lies completely on one side of the coordinate hyperplane $x_d = 0$.

Let $\mathrm{OSH}_d(n)$ be the smallest N such that every set H of N hyperplanes in $\mathbb{R}^d$ in general position contains a one-sided subset of n hyperplanes. Ramsey's theorem for d-tuples immediately gives $\mathrm{OSH}_d(n) \leq R_d(n)$ (a d-tuple gets color $+1$ if its intersection has a positive last coordinate, and color -1 otherwise).

Matoušek and Welzl [12] observed that, actually, $\mathrm{OSH}_2(n) = \mathrm{ES}_1(n) = (n-1)^2 + 1$, and applied this in a range-searching algorithm. Recently Dujmović and Langerman [4] used the existence of $\mathrm{OSH}_d(n)$ (essentially Lemma 9 in the arXiv version of their paper) to prove several interesting results, such as a ham-sandwich and centerpoint theorems for hyperplanes.

In Section 5 we show that lower bounds for kth-order monotone subsets in the plane can be translated into lower bounds for OSH_d.

Proposition 1.7. *We have* $\mathrm{OSH}_d(n) \geq \mathrm{ES}_{d-1}(n)$*, and in particular,* $\mathrm{OSH}_3(n) = 2^{\Omega(n)}$ *and*[5] $\mathrm{OSH}_4(n) = 2^{2^{\Omega(n)}}$.

The lower bounds for $\mathrm{OSH}_d(n)$ can also be translated into lower bounds in the theorems of Dujmović and Langerman. For example, in their ham-sandwich theorem, we have d collections $H_1, \ldots, H_d$ of hyperplanes in $\mathbb{R}^d$, each of size N, and we want a hyperplane g such that in each H_i, we can find disjoint subsets A_i, B_i of n hyperplanes each such $V(A_i)$ lies on one side of g and $V(B_i)$ on the other side.

To derive a lower bound for the smallest necessary N, we fix d affinely independent points $p_1, \ldots, p_d$ in the $x_d = 0$ hyperplane, and a set H of N hyperplanes in general position with no one-sided subset of size n. We let H_i be an affinely transformed copy of H such that all of $V(H_i)$ lies very close to p_i. Then every potential ham-sandwich hyperplane g for these H_i has to be almost parallel to the $x_d = 0$ hyperplane, and thus there cannot be A_i, B_i of size n for all i.

[5] An exponential lower bound for OSH_3 was known to the authors of [12], and perhaps to others as well, but as far as we know, it has not appeared in print.

The work of Fox et al. While preparing a draft of the present paper, we learned about a recent preprint of Fox, Pach, Sudakov, and Suk [8]. They investigated various combinatorial and geometric problems inspired by Theorems 1.1 and 1.2, and as was mentioned above, among others, they introduced transitive colorings,[6] but mainly they studied a related but different Ramsey-type quantity: let $N_\ell(q, n)$ be the smallest integer N such that, for every coloring of $\binom{[N]}{\ell}$ with q colors, there exists an n-element $I = \{i_1, \dots, i_n\} \subseteq [N]$, $i_1 < \cdots < i_n$, inducing a *monochromatic monotone path*, i.e., such that all the ℓ-tuples of the form $\{i_j, i_{j+1}, \dots, i_{j+\ell-1}\}$, $j = 1, 2, \dots, n - \ell + 1$, have the same color.

They note that $R_\ell^{\mathrm{trans}}(n) \le N_\ell(2, n)$, and they obtained the following bounds for $N_\ell(2, n)$: $N_2(2, n) = \mathrm{ES}_1(n)$, $N_3(2, n) = \mathrm{ES}_2(n)$, and for every fixed $k \ge 3$,

$$\mathrm{twr}_k(\Omega(n)) \le N_{k+1}(2, n) \le \mathrm{twr}_k(O(n \log n)).$$

As we mentioned after Theorem 1.4, this also yields an upper bound for $R_{k+1}^{\mathrm{trans}}(n)$ only slightly weaker than the one in that theorem.

After publication of a conference version of this paper [5], closely related works on semialgebraic predicates of Bukh and Matoušek [1] and Conlon et al. [2] appeared. Conlon et al. prove that the bounds from the Ramsey theorem can be improved by one exponential for every semialgebraic predicate, including nontransitive ones, and that in a sufficiently large dimension this is the best possible.

Open problems.

1. We have obtained reasonably tight bounds for $\mathrm{ES}_3(n)$, but the gaps are much more significant for $\mathrm{ES}_k(n)$ with $k \ge 4$. According to the cases $k = 1, 2, 3$, one may guess that $\mathrm{ES}_k(n)$ is of order $\mathrm{twr}_k(\Theta(n))$, and thus that stronger lower bounds are needed. On the other hand, a surprising result of Bukh and Matoušek [1] says that a Ramsey function for any k-ary semialgebraic predicate in $\mathbb{R}^1$ is at most $2^{2^{Cn}}$. A similar result for $\mathbb{R}^2$ would imply that $\mathrm{ES}_k(n)$ can be bounded from above by a tower function of a constant height independent of k. This question looks both interesting and challenging.
2. A perhaps more manageable task might be a better lower bound for $R_k^{\mathrm{trans}}(n)$, $k \ge 4$. A natural approach would be to imitate the Stepping-Up Lemma used for lower bounds for the Ramsey numbers $R_k(n)$ (see, e.g., [3]). But so far we have not succeeded in this, since even if we start with a transitive coloring of k-tuples, we could not guarantee transitivity for the coloring of $(k + 1)$-tuples.
3. As for order-type homogeneous sequences and one-sided subsets of hyperplanes, recent results of Conlon et al. [2] imply nearly tight upper bounds for $\mathrm{OT}_3(n)$ and $\mathrm{OSH}_4(n)$. They also give upper bounds for larger dimensions, namely $\mathrm{OT}_d(n) \le \mathrm{twr}_d(n^c)$ and $\mathrm{OSH}_d(n) \le \mathrm{twr}_{d-1}(cn^2 \log n)$, but a question of exact magnitude of these functions is still open.
4. In an earlier version of this paper we asked whether $n \log n$ can be replaced by n in the upper bound for the quantity $N_\ell(2, n)$ considered by Fox et al. [8]. This question was answered positively in a recent paper of Moshkovitz and Shapira [14].
5. In our definition of kth-order positivity, every $(k + 1)$-tuple of points should lie on the graph of a function with a nonnegative kth derivative, and different functions can be used for different $(k + 1)$-tuples. In an earlier version of this paper, we conjectured that, assuming

[6] With still another geometric source of such colorings besides the Erdős–Szekeres theorems, namely, noncrossing convex bodies in the plane.

k-general position, a single function should suffice for all $(k+1)$-tuples; in other words, that every kth-order monotone finite set in k-general position lies on a graph of a k-times differentiable function $f: \mathbb{R} \to \mathbb{R}$ whose kth derivative is everywhere nonnegative or everywhere nonpositive.

However, Rote [16] disproved this for $k = 3$ (while the cases $k = 1, 2$ do hold, as is not hard to check). With his kind permission, we reproduce his example at the end of Section 2.

Naturally, this opens up interesting new questions: How can one characterize point sets lying on the graph of a function whose kth derivative is positive everywhere? Is there a Ramsey-type theorem for such sets, and if yes, how large is the corresponding Ramsey function?

2. On the definition of kth-order monotonicity

Here we provide several equivalent characterizations of kth-order monotonicity of planar point sets and some of their properties. First we recall several known results.

Divided differences and Newton's interpolation. Let $p_1, p_2, \ldots, p_{k+1}$ be points in the plane, $p_i = (x_i, y_i)$, where the x_i are all distinct (but not necessarily increasing). We recall that the *kth divided difference* $\Delta_k(p_1, p_2, \ldots, p_{k+1})$ is defined recursively as follows:

$$\Delta_0(p_1) := y_1$$

$$\Delta_k(p_1, p_2, \ldots, p_{k+1}) := \frac{\Delta_{k-1}(p_2, p_3, \ldots, p_{k+1}) - \Delta_{k-1}(p_1, p_2, \ldots, p_k)}{x_{k+1} - x_1}.$$

For example, $\Delta_1(p_1, p_2)$ equals the slope of the line $p_1 p_2$. In general, the kth divided difference is related to the kth derivative as follows (see, e.g., [15, Eq. 1.33]; note that the case $k = 1$ is the Mean Value Theorem):

Lemma 2.1 (*Cauchy*). *Let the points $p_1, \ldots, p_{k+1}$, $a := x_1 < x_2 < \cdots < b := x_{k+1}$, lie on the graph of a function f such that the kth derivative $f^{(k)}$ exists everywhere on the interval (a, b). Then there exists $\xi \in (a, b)$ such that*

$$\Delta_k(p_1, \ldots, p_{k+1}) = \frac{f^{(k)}(\xi)}{k!}.$$

We will also need the following result (see, e.g., [15, Eqs. 1.11–1.19]).

Lemma 2.2 (*Newton's Interpolation*). *Let $p_1, \ldots, p_{k+1} \in \mathbb{R}^2$ be points with distinct x-coordinates (here we need not assume that the x-coordinates are increasing). Then the unique polynomial f of degree at most k whose graph contains $p_1, \ldots, p_{k+1}$ is given by*

$$f(x) = \sum_{i=1}^{k+1} \left(\Delta(p_1, \ldots, p_i) \prod_{j=1}^{i-1} (x - x_j) \right)$$

In particular, the coefficient of x^k is $\Delta(p_1, \ldots, p_{k+1})$, and it equals $f^{(k)}(x)/k!$ (which is a constant function).

We recall that a $(k+1)$-tuple $K = \{p_1, \ldots, p_{k+1}\}$ was defined to be positive if it is contained in the graph of a function having a nonnegative kth derivative everywhere. We obtain the following equivalent characterization:

Corollary 2.3. *A $(k+1)$-tuple $K = \{p_1, \dots, p_{k+1}\}$ is positive iff $\Delta_k(p_1, \dots, p_{k+1}) \ge 0$ (and similarly for a negative $(k+1)$-tuple). If K is in k-general position, we have* $\mathrm{sgn} K = \mathrm{sgn}\Delta_k(p_1, \dots, p_{k+1})$.

Proof. If K is contained in the graph of f with $f^{(k)} \ge 0$ everywhere, then $\Delta_k(p_1, \dots, p_{k+1}) \ge 0$ by Lemma 2.1.

Conversely, if $\Delta_k(p_1, \dots, p_{k+1}) \ge 0$, then by Lemma 2.2, the unique polynomial of degree at most k whose graph contains K is the required function with nonnegative kth derivative.

If, moreover, K is in k-general position, then $\Delta_k(p_1, \dots, p_{k+1}) \ne 0$, and so K cannot be both kth-order positive and kth-order negative by Lemma 2.1. $\square$

We will also need the following criterion for the sign of a $(k+1)$-tuple.

Lemma 2.4. *Let $K = \{p_1, p_2, \dots, p_{k+1}\}$ be a $(k+1)$-tuple of points in k-general position, $x_1 < \cdots < x_{k+1}$, let $i \in [k+1]$, and let f_i be the (unique) polynomial of degree at most $k-1$ whose graph passes through the points of $K \setminus \{p_i\}$. Then $\mathrm{sgn} K = (-1)^{k-i}$ if p_i lies below the graph of f_i, and $\mathrm{sgn} K = (-1)^{k+1-i}$ if p_i lies above the graph.*

Let f be the polynomial of degree at most k passing through all of K. We use Newton's interpolation (Lemma 2.2), but with the points reordered so that p_i comes last, and we get that

$$f(x) = f_i(x) + \Delta_k(p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_{k+1}, p_i) \prod_{j \in [k+1] \setminus \{i\}} (x - x_j).$$

Using this with $x = x_i$, we get

$$\begin{aligned} \mathrm{sgn}(y_i - f_i(x_i)) &= \mathrm{sgn}(f(x_i) - f_i(x_i)) \\ &= \mathrm{sgn}\Delta_k(p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_{k+1}, p_i) \cdot \mathrm{sgn} \prod_{j \in [k+1] \setminus \{i\}} (x_i - x_j). \end{aligned}$$

Divided differences are invariant under permutations of the points (as can be seen, e.g., from Lemma 2.2, since the interpolating polynomial does not depend on the order of the points), and so $\mathrm{sgn}\Delta_k(p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_{k+1}, p_i) = \mathrm{sgn} K$. Finally, the product

$$\prod_{j \in [k+1] \setminus \{i\}} (x_i - x_j)$$

has $k+1-i$ negative factors, thus its sign is $(-1)^{k+1-i}$, and the lemma follows. $\square$

It remains to prove transitivity.

Lemma 2.5. *Let $P = \{p_1, \dots, p_N\}$ be a point set in k-general position. Then the 2-coloring of $(k+1)$-tuples $K \in \binom{P}{k+1}$ by their sign is transitive.*

Proof. We consider a $(k+2)$-tuple $L = \{p_1, \dots, p_{k+2}\}$ with

$$\mathrm{sgn}\{p_1, \dots, p_{k+1}\} = \mathrm{sgn}\{p_2, \dots, p_{k+2}\} = +1$$

and we fix $i \in \{2, \dots, k+1\}$. Let $f_{i,k+2}$ be the polynomial of degree at most $k-1$ passing through $L \setminus \{p_i, p_{k+2}\}$, and similarly for $f_{1,k+2}$. Our goal is to show that $f_{i,k+2}(x_{k+2}) < y_{k+2}$, since this gives $\mathrm{sgn}(L \setminus \{p_i\}) = +1$ by Lemma 2.4.

Since $\mathrm{sgn}(L \setminus \{p_1\}) = +1$, we have $f_{1,k+2}(x_{k+2}) < y_{k+2}$ (Lemma 2.4 again), and so it suffices to prove $f_{i,k+2}(x_{k+2}) < f_{1,k+2}(x_{k+2})$.

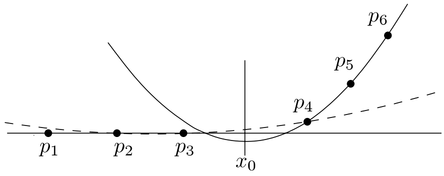

Fig. 1. Rote's example: a 6-point 3rd-order positive set in 3-general position that does not lie on the graph of any function with nonnegative 3rd derivative.

Let us consider the polynomial $g := f_{1,k+2} - f_{i,k+2}$; as explained above, our goal is proving $\operatorname{sgn} g(x_{k+2}) = +1$. To this end, we first determine $\operatorname{sgn} g(x_1)$: We have $f_{i,k+2}(x_1) = y_1$ and $\operatorname{sgn}(y_1 - f_{1,k+2}(x_1)) = (-1)^k$ (using $\operatorname{sgn}(L \setminus \{p_1\}) = +1$ and Lemma 2.4). Hence $\operatorname{sgn} g(x_1) = (-1)^{k-1}$.

Next, we observe that g is a polynomial of degree at most $k - 1$, and it vanishes at $x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{k+1}$. These are $k - 1$ distinct values; thus, they include all roots of g, and each of them is a simple root. Consequently, g changes sign $(k - 1)$-times between x_1 and x_{k+2}. Hence, finally, $\operatorname{sgn} g(x_{k+2}) = (-1)^{k-1} \operatorname{sgn} g(x_1) = +1$ as claimed. □

Rote's example. Fig. 1 shows a 6-point set $P = \{p_1, \ldots, p_6\}$ in 3-general position (no four points on a parabola). It is easy to check 3rd-order positivity using Lemma 2.4: By transitivity, it suffices to look at 4-tuples of consecutive points. For $p_1, \ldots, p_4$ we use the parabola through p_1, p_2, p_3 (which actually degenerates to the x-axis); for $p_2, \ldots, p_5$ we use the dashed parabola through p_2, p_3, p_4 (which is very close to the x-axis in the relevant region); and for $p_3, \ldots, p_6$, the parabola through p_4, p_5, p_6 (drawn full).

It remains to check that P does not lie on the graph of a function f with $f^{(3)} \geq 0$ everywhere. Assuming for contradiction that there is such an f, we consider the point $q := (x_0, f(x_0))$, where x_0 is such that the full parabola is below the x-axis at x_0. For the 4-tuple $\{p_1, p_2, p_3, q\}$ to be positive, q has to lie above the x-axis, but the 4-tuple $\{q, p_4, p_5, p_6\}$ is positive only if q lies below the parabola through p_4, p_5, p_6—a contradiction.

3. Upper bounds on the Ramsey numbers for transitive colorings

In this section we prove Theorem 1.4. As we mentioned in the remark following that theorem, it suffices to establish the case $k \geq 3$.

Thus, we want to prove that $R_{k+1}^{\text{trans}}(n) \leq \operatorname{twr}_k(C_k n)$ for all n and for every $k \geq 3$, with suitable constants C_k depending on k. As the base of the induction we use $R_3^{\text{trans}}(n) \leq 4^n$, which, as was remarked earlier, follows by imitating the proof of Theorem 1.2.

Thus, let $k \geq 3$ be fixed, let n be given, and let us set $M := R_k^{\text{trans}}(n)$. We will prove that

$$R_{k+1}^{\text{trans}}(n) \leq N := 2^{M^k}. \tag{1}$$

Theorem 1.4 then follows from this recurrence and from the fact that $2^{\operatorname{twr}_{k-1}(n)^k} \leq \operatorname{twr}_k(kn)$ for $k \geq 3$, which is easy to check.

To prove (1), we follow an inductive proofs of Ramsey's theorem going back to Erdős and Rado [6]. Let $\chi : \binom{[N]}{k+1} \to \{1, 2\}$ be an arbitrary transitive 2-coloring. We set $A_{k-1} := \{1, 2, \ldots, k-1\}$ and $X_{k-1} := [N] \setminus A_{k-1}$. For $i = k, k+1, \ldots, M$ we will inductively construct sets $A_i, X_i \subseteq [N]$ such that

(i) $A_i < X_i$ (i.e., all elements of A_i precede all elements of X_i);
(ii) $|A_i| = i$ and $|X_i| \geq |X_{i-1}|/2^{M^{k-1}}$; and
(iii) the color of a $(k+1)$-tuple whose first k elements all belong to A_i does not depend on its last element; in other words, for $K \in \binom{A_i}{k}$ and $x, y \in A_i \cup X_i$ with $K < \{x, y\}$, we have $\chi(K \cup \{x\}) = \chi(K \cup \{y\})$.

For the inductive step, suppose that A_i and X_i have already been constructed. We let x_i be the smallest element of X_i, we set $A_{i+1} := A_i \cup \{x_i\}$, and we write $X'_i := X_i \setminus \{x_i\}$.

Let us call two elements $x, y \in X'_i$ *equivalent* if we have, for every $K \in \binom{A_{i-1}}{k-1}$, $\chi(K \cup \{x_i, x\}) = \chi(K \cup \{x_i, y\})$. There are $\binom{i}{k-1}$ possible choices of K, and hence there are at most $2^{\binom{i}{k-1}} < 2^{M^{k-1}}$ equivalence classes. We choose $X_{i+1} \subseteq X'_i$ as the largest equivalence class. Then (i), (iii) obviously hold for A_{i+1} and X_{i+1}, and we have

$$|X_{i+1}| \geq (|X_i| - 1)/(2^{M^{k-1}} - 1) \geq |X_i|/2^{M^{k-1}}$$

(since $i \leq M$ and thus we have $|X_i| \geq N/(2^{M^{k-1}})^{i-1} = 2^{M^k-(i-1)M^{k-1}} \geq 2^{M^{k-1}}$). This finishes the inductive construction of A_i and X_i.

In this way, we construct the sets $A := A_M$ and X_M (note that $|X_M| \geq 1$ by (ii)). Let x be the first element of X_M, and let us define a 2-coloring $\chi^*\colon \binom{A}{k} \to \{1, 2\}$ of the k-tuples of A by $\chi^*(K) := \chi(K \cup \{x\})$.

We claim that, crucially, χ^* is transitive (which is not entirely obvious). So we consider elements $a_1 < a_2 < \cdots < a_{k+1}$ of A, and we suppose that

$$\chi^*(\{a_1, \ldots, a_k\}) = \chi^*(\{a_2, \ldots, a_{k+1}\}) =: c$$

We want to show that $\chi^*(\{a_1, \ldots, a_{k+1}\} \setminus \{a_i\}) = c$ for every $i = 2, 3, \ldots, k$. We have $c = \chi^*(\{a_1, \ldots, a_k\}) = \chi(\{a_1, \ldots, a_k, x\}) = \chi(\{a_1, \ldots, a_{k+1}\})$ (by definition and by the independence of χ of the last element), and $c = \chi^*(\{a_2, \ldots, a_{k+1}\}) = \chi(\{a_2, \ldots, a_{k+1}, x\})$. Next we use the transitivity of χ on the $(k+2)$-tuple $(a_1, \ldots, a_{k+1}, x)$, obtaining

$$\chi(\{a_1, \ldots, a_{k+1}, x\} \setminus \{a_i\}) = c = \chi^*(\{a_1, \ldots, a_{k+1}\} \setminus \{a_i\})$$

as needed.

Now we can apply the inductive hypothesis to A, which yields an n-element subset of A homogeneous with respect to χ^*, and this subset is homogeneous with respect to χ as well, finishing the proof of Theorem 1.4. □

4. A lower bound for ES_3

Here we prove Theorem 1.5, a lower bound for $\mathrm{ES}_3(2n+1)$. We proceed by induction on n; the goal is to construct a set P_n of $N := 2^{2^{n-1}}$ points with no $(2n+1)$-point third-order monotone subset. The induction starts for $n = 2$ with an arbitrary P_2 of size $2^{2^1} = 4$.

In the inductive step, given P_n, we will construct P_{n+1} so that $|P_{n+1}| = |P_n|^2$; then the bound on the size of P_n clearly holds.

We may assume that $P = P_n$ is in 3-general position (this can always be achieved by a small perturbation). By an affine transformation we also make sure that $P \subset [1, 2] \times [0, 1]$; or actually, $P \subset [1, 1.9] \times [0, 1]$ so that there is some room for perturbation. Moreover, there is a small

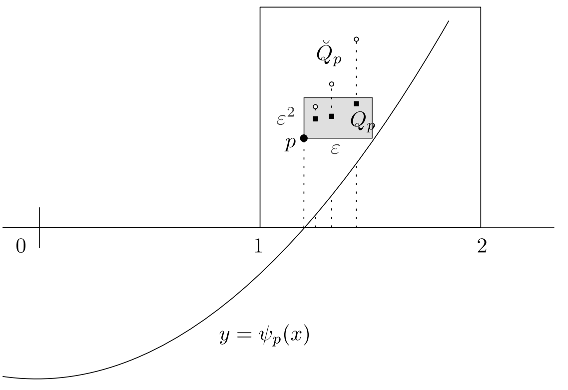

Fig. 2. A schematic illustration of the construction of P_{n+1}.

$\delta > 0$ such that if P' is obtained from P by moving each point arbitrarily by at most δ, then P' is still in 3-general position, the order of the points of P' along the x-axis is the same as that for P, and the sign of every 4-tuple in P' is the same as the sign of the corresponding 4-tuple in P.

The construction. The construction of P_{n+1} from $P = P_n$ as above proceeds in the following steps.

1. We choose a sufficiently large number $A = A(P)$ (the requirements on it will be specified later), and we set $\varepsilon := 1/A^2$.
2. For every point $p \in P$, let Q_p be the image of P under the affine map that sends the square $[1, 2] \times [0, 1]$ to the axis-parallel rectangle of width ε, height ε^2, and with the lower left corner at p; see Fig. 2.
3. Let $\psi_p(x) = Ax^2 + C_p$ be a quadratic function, where A is as above and C_p is chosen so that $\psi_p(x(p)) = 0$ (where $x(p)$ is the x-coordinate of p). Let $\check{Q}_p$ be the set obtained by "adding ψ_p to Q_p", i.e., by shifting each point $(x, y) \in Q_p$ vertically upwards by $\psi_p(x)$. We set $P_{n+1} := \bigcup_{p \in P} \check{Q}_p$. We call the $\check{Q}_p$ the *clusters* of P_{n+1}.

First we check that each cluster $\check{Q}_p$ lies close to p.

Lemma 4.1. *Each $\check{Q}_p$ is contained in an $O\left(\sqrt{\varepsilon}\right)$-neighborhood of p.*

Proof. Writing $p = (x_0, y_0)$, the set Q_p obviously lies in the 2ε-neighborhood of p, and the maximum amount by which a point of Q_p was translated upwards is at most

$$\psi_p(x_0 + \varepsilon) = A\left((x_0 + \varepsilon)^2 - x_0^2\right) = A(2x_0\varepsilon + \varepsilon^2) = O\left(\sqrt{\varepsilon}\right). \quad \square$$

Here is a key property of the construction.

Lemma 4.2 (*Slope Lemma*)**.** *Let λ be a parabola passing through three points of P_{n+1} that belong to three different clusters, or a line passing through two points of different clusters. Let μ be a parabola passing through three points of a single cluster $\check{Q}_p$, or a line passing through two such points. Then the maximum slope (first derivative) of λ on the interval* $[1, 2]$ *is smaller than the minimum slope of μ on* $[1, 2]$*, provided that A was chosen sufficiently large.*

Proof. Clearly, the maximum slope of any such λ can be bounded from above by some finite number depending only on P but not on A. Thus, it suffices to show that, with A large, for every μ as in the lemma, the minimum slope is bounded from below by A.

First let us assume that μ is a parabola passing through three points of $\check{Q}_p$, where $p = (x_0, y_0)$, let $\tilde{\mu}$ be the parabola passing through the corresponding three points of P, and let the equation of $\tilde{\mu}$ be $y = ax^2 + bx + c$.

By the construction of $\check{Q}_p$, the affine map transforming P to Q_p sends a point with coordinates (x, y) to the point $(\varepsilon(x-1) + x_0, \varepsilon^2 y + y_0)$. Calculation shows that the image of $\tilde{\mu}$ under this affine map has the equation $y = ax^2 + (2a\varepsilon + b\varepsilon - 2ax_0)x + c'$, where the value of the absolute term c' need not be calculated since it does not matter. Hence the minimum slope of this curve on $[1, 2]$ is bounded from below by $-(8|a| + 4|a|\varepsilon + 2|b|\varepsilon + 8|a|)$. Finally, μ is obtained by adding $\psi_p(x) = Ax^2 + C_p$ to this curve, and the minimum slope of ψ_p on $[1, 2]$ is at least $2A$.

Next, let μ be a line passing through two points $q, r \in \check{Q}_p$. Let us choose another point $s \in \check{Q}_p$ and consider the parabola μ' through q, r, s. By the Mean Value Theorem, the slope of μ equals the slope of μ' at some point between q and r, and the latter is at least A by the above. The lemma is proved. □

Let $K = \{p_1, p_2, p_3, p_4\} \subseteq P_{n+1}$ be a 4-tuple, $p_i = (x_i, y_i)$, $x_1 < \cdots < x_4$. We assign a *type* to K, which is an ordered partition of 4 given by the distribution of K among the clusters; for example, K has type $1+1+2$ if the first point p_1 lies in some $\check{Q}_p$, p_2 lies in $\check{Q}_{p'}$ for $p' \neq p$, and $p_3, p_4 \in \check{Q}_{p''}$, $p'' \neq p, p'$.

The next lemma shows that the sign K is determined by its type. We provide a complete classification, although we will not use all of the types in the subsequent proof.

Lemma 4.3. *Let $K = \{p_1, p_2, p_3, p_4\} \subseteq P_{n+1}$ be a 4-tuple. If K is of type $1+1+1+1$ or 4, then the sign of K is the same as that of the corresponding 4-tuple in P. Otherwise, the sign of K is determined by its type as follows:*

- *for types $3+1$ and $1+3$ it is -1;*
- *for types $1+1+2$ and $2+1+1$ it is $+1$;*
- *for type $1+2+1$ it is -1; and*
- *for type $2+2$ it is $+1$.*

Proof. Since the transformation that converts P into $\check{Q}_p$ preserves the types of 4-tuples, the statement for type 4 is clear. The statement for type $1+1+1+1$ follows since, by Lemma 4.1, K is obtained by a sufficiently small perturbation of the corresponding 4-tuple in P (this gives one of the lower bounds on A, since we need the bound in Lemma 4.1 to be smaller than the δ considered at the beginning of our description of the construction).

The statements for the remaining types are obtained by simple application of the slope lemma (Lemma 4.2) together with Lemma 2.4. Namely, for type $3+1$, we get that the parabola through p_1, p_2, p_3 lies above p_4 (by comparing its slope to the slope of the line p_3p_4); see Fig. 3. For type $1+3$ we similarly get that p_1 lies above the parabola through p_2, p_3, p_4, and so the sign is -1 in both of these cases.

For type $1+1+2$, the segment p_3p_4 is steeper than the parabola through $p_1p_2p_3$, and so the sign is $+1$. Similarly for type $2+1+1$ we get that p_1 lies below the parabola through p_2, p_3, p_4, which again gives sign $+1$. For type $1+2+1$, p_3 lies above the parabola through p_1, p_2, p_4, giving sign -1. Finally, for type $2+2$, the segment p_1p_2 is steeper than p_2p_3, thus the parabola

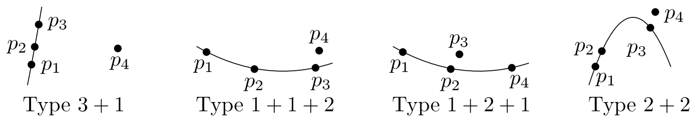

Fig. 3. Determining the signs of 4-tuples by type.

through p_1, p_2, p_3 is concave, and hence its slope at p_3 and after it is no larger than the slope of the segment $p_2 p_3$. Thus, p_4 lies above this parabola and the sign is $+1$ as claimed. □

Finishing the proof of Theorem 1.5. It remains to show that P_{n+1} contains no $(2n + 3)$-point third-order monotone subset.

For contradiction, suppose that $M \subseteq P_{n+1}$ is such a $(2n + 3)$-point subset. Let $2n + 3 = n_1 + n_2 + \cdots + n_s$ be the type of M (i.e., M has $n_i \geq 1$ points in the ith leftmost cluster it intersects). By the inductive assumption we have $s \leq 2n$ and $n_i \leq 2n$ for all i.

Let $n_a = \max_i n_i$ and $n_b = \max_{i \neq a} n_i$ be the two largest among the n_i. For convenience, let us assume $a < b$; the case $a > b$ is handled symmetrically. We distinguish three cases.

First, if $n_a \geq 3$ and $n_b \geq 2$, then we can select 4-tuples of types $3 + 1$ and $2 + 2$ from the corresponding two clusters, which have different signs, and so M is not homogeneous.

Second, if $n_a \geq 3$ and $n_b = 1$, then we have at least three n_i equal to 1 (since $n_a \leq 2n$), and at least two of them lie on the same side of the cluster corresponding to n_a, say to the right of it. Then we can select 4-tuples of types $3 + 1$ and $2 + 1 + 1$, again of opposite signs.

Third, if $n_a = 2$, then there are at least two other clusters of size 2. From these three 2-element clusters, we can select 4-tuples of types $2 + 2$ and $1 + 2 + 1$, again of opposite signs.

This exhausts all possibilities ($n_a = 1$ cannot happen, because $s \leq 2n$), and Theorem 1.5 is proved. □

5. Order types and one-sided sets of hyperplanes

First we substantiate the two claims made above Proposition 1.6, concerning OT_1 and OT_2. For $d = 1$, an order-type homogeneous sequence in $\mathbb{R}^1$ is just a monotone sequence of real numbers, so $\mathrm{OT}_1(n) = (n-1)^2 + 1$ by Theorem 1.1.

In a similar spirit, it is easy to check that a planar order-type homogeneous sequence corresponds to the vertices of a convex n-gon, enumerated in a clockwise or counterclockwise order. Thus, $\mathrm{OT}_2(n) \geq \mathrm{ES}_2(\lceil n/2 \rceil) = 2^{\Omega(n)}$. On the other hand, given any N-point sequence, we can first select a subsequence of $\left\lceil \sqrt{N} \right\rceil$ points with increasing or decreasing x-coordinates, and then we select a convex or concave configuration from it. Thus, by Theorem 1.2, we have $\mathrm{OT}_2(n) = 2^{O(n)}$.

Proof of Proposition 1.6. For a point $p = (x, y) \in \mathbb{R}^2$, we define the point $\tilde{p} := (x, x^2, \ldots, x^{d-1}, y) \in \mathbb{R}^d$.

To prove that $\mathrm{ES}_d(n) \leq \mathrm{OT}_d(n)$, we consider a set $P = \{p_1, \ldots, p_N\} \subset \mathbb{R}^2$ in d-general position, $p_i = (x_i, y_i)$, where $N = \mathrm{ES}_d(n) - 1$ and $x_1 < \cdots < x_N$, with no dth-order monotone subset of n points. It suffices to prove that the sequence $\tilde{P} := (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N)$ has no n-point order-type homogeneous subsequence. This follows from the next lemma.

Lemma 5.1. *For every $(d + 1)$-tuple $(p_1, \ldots, p_{d+1})$ of points in $\mathbb{R}^2$, $x_1 < \cdots < x_{d+1}$, we have* $\mathrm{sgn}(\{p_1, \ldots, p_{d+1}\}) = \mathrm{sgn} \det M(\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{d+1})$, *where $M(q_1, \ldots, q_{d+1})$ is the matrix from the definition of order type above Proposition 1.6.*

Proof. By Lemma 2.2 and Corollary 2.3, the sign of $\{p_1, \ldots, p_{d+1}\}$ equals the sign of the coefficient a_d of the unique polynomial $f(x) = \sum_{j=0}^{d} a_j x^j$ of degree at most d whose graph passes through the points $p_1, \ldots, p_{d+1}$.

The vector $a = (a_0, \ldots, a_d)$ can be expressed as the solution of the linear system $Va = y$, where $y = (y_1, \ldots, y_{d+1})$ and V is the *Vandermonde matrix* with $v_{ij} = x_i^{j-1}$, $i, j = 1, 2, \ldots, d+1$. By Cramer's rule, we obtain

$$a_d = \frac{\det W}{\det V},$$

where W stands for the matrix V with the last column replaced with the vector y. As is well known, $\det V = \prod_{1 \le i < j \le d+1} (x_j - x_i)$, and since $x_1 < \cdots < x_{d+1}$, we have $\det V > 0$. Thus, $\operatorname{sgn} a_d = \operatorname{sgn} \det W$. Finally, we have

$$W = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{d-1} & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{d+1} & x_{d+1}^2 & \ldots & x_{d+1}^{d-1} & y_{d+1} \end{pmatrix} = M(\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{d+1})^T.$$

The lemma follows, and Proposition 1.6 is proved. □

Proof of Proposition 1.7. The proof is very similar to the previous one. This time we start with a set $P = \{p_1, \ldots, p_N\} \subset \mathbb{R}^2$ in $(d-1)$-general position, $p_i = (x_i, y_i)$, where $N = \mathrm{ES}_{d-1}(n) - 1$ and $x_1 < \cdots < x_N$, with no $(d-1)$th-order monotone subset of n points. We define a collection $H = \{h_1, \ldots, h_N\}$ of N hyperplanes in $\mathbb{R}^d$, where h_i is given by

$$h_i = \left\{ (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d : \sum_{j=1}^{d} x_i^{j-1} \xi_j = y_i \right\}.$$

The intersection point $\xi = (\xi_1, \ldots, \xi_d)$ of, say, $h_1, \ldots, h_d$ is the solution of the linear system $V\xi = y$, where V is the $d \times d$ Vandermonde matrix this time, $v_{ij} = x_i^{j-1}$. Cramer's rule then gives that the dth coordinate ξ_d, whose sign we are interested in, equals $(\det W)/(\det V)$, where W is obtained from V by replacing the last column with y.

As we saw in the proof of Proposition 1.6, $(\det W)/(\det V)$ also expresses the leading coefficient in the polynomial of degree $d-1$ passing through $p_1, \ldots, p_d$, and thus its sign equals $\operatorname{sgn} \Delta_{d-1}(p_1, \ldots, p_d)$. It follows that one-sided subsets of H precisely correspond to $(d-1)$st-order monotone subsets in P, and the proposition is proved. □

Acknowledgments

The authors would like to thank János Pach for kindly discussing some of the results of Fox et al. [8] with them. They also thank Günter Rote for informing them about his refutation of their conjecture and for permission to present it in this paper. They were supported by the ERC Advanced Grant No. 267165. The first author was partially supported by the Charles University Grant GAUK 1262213.

References

[1] B. Bukh, J. Matoušek, Erdős–Szekeres-type statements: Ramsey function and decidability in dimension 1, 2012. Arxiv preprint arXiv:1207.0705.

[2] D. Conlon, J. Fox, J. Pach, B. Sudakov, A. Suk, Ramsey-type results for semi-algebraic relations, 2013. Arxiv preprint arXiv:1301.0074.

[3] D. Conlon, J. Fox, B. Sudakov, An improved bound for the stepping-up lemma, Discrete Appl. Math. 161 (2013) 1191–1196.

[4] V. Dujmović, S. Langerman, A center transversal theorem for hyperplanes and applications to graph drawing, in: Proc. 27th ACM Symposium on Computational Geometry, 2011, pp. 117–124. Full version arXiv:1012.0548.

[5] M. Eliáš, J. Matoušek, Higher-order Erdős–Szekeres theorems, in: Proceedings of the 2012 Symposium on Computational Geometry, SoCG '12, ACM, New York, NY, USA, 2012, pp. 81–90.

[6] P. Erdős, R. Rado, Combinatorial theorems on classifications of subsets of a given set, Proc. Lond. Math. Soc. 3 (1952) 417–439.

[7] P. Erdős, G. Szekeres, A combinatorial problem in geometry, Compos. Math. 2 (1935) 463–470.

[8] J. Fox, J. Pach, B. Sudakov, A. Suk, Erdős–Szekeres-type theorem for monotone paths and convex bodies, Proceedings of the London Mathematical Society 105 (5) (2012) 953–982.

[9] J.E. Goodman, R. Pollack, Allowable sequences and order types in discrete and computational geometry, in: J. Pach (Ed.), New Trends in Discrete and Computational Geometry, in: Algorithms and Combinatorics, vol. 10, Springer, Berlin etc., 1993, pp. 103–134.

[10] R.L. Graham, B.L. Rothschild, J. Spencer, Ramsey Theory, J. Wiley & Sons, New York, 1990.

[11] J. Matoušek, Lectures on Discrete Geometry, Springer, New York, 2002.

[12] J. Matoušek, E. Welzl, Good splitters for counting points in triangles, J. Algorithms 13 (1992) 307–319.

[13] W. Morris, V. Soltan, The Erdős–Szekeres problem on points in convex position—a survey, Bull. Amer. Math. Soc. (N.S.) 37 (4) (2000) 437–458.

[14] G. Moshkovitz, A. Shapira, Ramsey theory, integer partitions and a new proof of the Erdős–Szekeres theorem, 2012. Arxiv preprint arXiv:1206.4001.

[15] G.M. Phillips, Interpolation and Approximation by Polynomials, Springer, Berlin etc., 2003.

[16] G. Rote, Private communication, February 2012.

[17] M.J. Steele, Variations on the monotone subsequence theme of Erdős and Szekeres, in: D. Aldous, et al. (Eds.), Discrete Probability and Algorithms, in: IMA Volumes in Mathematics and its Applications, vol. 72, Springer, Berlin etc., 1995, pp. 111–131.

[18] P. Valtr, Open caps and cups in planar point sets, Discrete Comput. Geom. 37 (2004) 365–567.