Charles University in Prague

Faculty of Mathematics and Physics

MASTER THESIS

Marek Eliáš

Erdős-Szekeres type theorems

Department of Applied Mathematics (KAM)

Supervisor of the master thesis:	prof. RNDr. Jiří Matoušek DrSc.
Study programme:	Computer Science
Specialization:	Discrete Models and Algorithms

Prague 2012

I would like to thank to my supervisor for many advices and care throughout preparation of this thesis. I would also like to thank János Pach for kindly discussing some of the results of Fox et al. [FPSS11] with us. I also thank Günter Rote for informing us about about his refutation of our conjecture and for permission to present it.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In .... date .... signature

Název práce: Věty Erdős-Szekeresova typu

Autor: Marek Eliáš

Katedra: Katedra aplikované matematiky (KAM)

Vedoucí diplomové práce: prof. RNDr. Jiří Matoušek DrSc.

Abstrakt: Nech $P = (p_1, p_2, \ldots, p_N)$ je postupnosť bodov v rovine, kde $p_i = (x_i, y_i)$ a $x_1 < x_2 < \cdots < x_N$. Slávna Erdős–Szekeresova veta z roku 1935 hovorí, že každá taká postupnosť P obsahuje monotónnu podpostupnosť S dĺžky $\lceil \sqrt{N} \rceil$. Iná, podobne slávna veta z toho istého článku hovorí, že každá taká postupnosť P obsahuje konvexnú alebo konkávnu podpostupnosť dĺžky $\Omega(\log N)$. Najprv definujeme $(k+1)$-ticu $K \subseteq P$ ako pozitívnu, keď leží na grafe funkcie s nezápornou k-tou deriváciou a podobne tiež negatívnu $(k+1)$-ticu. Ďalej hovoríme, že $S \subseteq P$ je monotónna k-teho rádu, keď jej $(k+1)$-tice sú buďto všetky pozitívne alebo všetky negatívne. V tejto práci skúmame kvantitatívne odhady pre zodpovedajúce Ramseyovské funkcie. Dostávame $\Omega(\log^{(k-1)} N)$ ako dolný odhad. Taktiež uvádzame vylepšené odhady pre súvisiace problémy ako Order types a One-sided sets of hyperplanes.

Klíčová slova: Erdős–Szekeresova veta, Ramseyova teória, order type, podielová diferencia, monotónna množina k-teho rádu

Title: Erdős-Szekeres type theorems

Author: Marek Eliáš

Department: Department of Applied Mathematics (KAM)

Supervisor: prof. RNDr. Jiří Matoušek DrSc.

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. 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. In this thesis we investigate quantitative bound for the corresponding Ramsey-type result. We obtain an $\Omega(\log^{(k-1)} N)$ lower bound ($(k-1)$-times iterated logarithm). We also improve bounds for related problems: Order types and One-sided sets of hyperplanes.

Keywords: Erdős–Szekeres theorem, Ramsey theory, order type, divided difference, k-th-order monotone subset

Contents

Preface

In this thesis we provide a generalisation of the two well-known theorems of Erdős and Szekeres and apply our results to related problems of order type and one-sided sets of hyperplanes. We introduce a concept of kth-order monotone sets to generalise monotonicity and convexity/concavity which are considered in original Erdős–Szekeres theorems.

Chapter 1 provides theoretical background and tools which are used throughout the thesis. In Chapter 2 we define kth-order monotonicity itself and outline several aspect of this definition. Transitive colorings of hypergraphs introduced by Fox et al. [FPSS11] turned out to be very useful when proving bounds for kth-order monotone sets; their properties are discussed in Chapter 3. Bounds for kth-order monotone sets are proved in Chapter 4. Probably the most interesting result is the construction providing a lower bound for 3rd-order monotone subsets in Theorem 4.2. Improved bounds for order types and one-sided sets of hyperplanes are provided in Chapter 5 and a list of unsolved related problems can be found in Chapter 6.

Most of the results in this thesis was written jointly with professor Matoušek in a paper [EM11] which will be published soon.

Chapter 1

Introduction

In this chapter we provide necessary theoretical background needed in proofs of our results. The first two sections provide insight into most important results of Ramsey theory which served as the main motivation of this work and are needed for basic understanding of its subject. Section 1.3 gives a summary of a recently published paper by Fox et al. [FPSS11] which introduces very important definition of transitive hypergraph coloring. Our results imply better bounds for two problems described in Sections 1.4 and 1.5. In the last section we recall technical propositions mostly from mathematical analysis and polynomial interpolation which are used mainly in the next chapter when describing properties of kth-order monotonicity.

1.1 Ramsey

The history of Ramsey theory begin in 1930 by article of Frank P. Ramsey [Ram30] although many of the results nowadays belonging to Ramsey theory were known before. More about history of Ramsey theory and many similar results can be found in book of Graham, Rothschild, and Spencer [GRS80]. We refer to this book also for proofs of the two following theorems.

Theorem 1.1 (Ramsey [Ram30])**.** *For every $k, r, n \in \mathbb{N}$ there exist a number $N \in \mathbb{N}$ such that for every coloring of the k-tuples of the $[N]$ by r colors there is a subset $T \subseteq [N]$ of size n such that all k-tuples of elements of T have the same color. We denote $R_k^r(n)$ the smallest such N and if $r = 2$ we write only $R_k(n)$.*

Infinite version is provided in the original words of Ramsey for r-tuples and μ colors as cited in [GRS80].

Theorem 1.2 (Infinite version of Ramsey's theorem [Ram30])**.** *Let Γ be an infinite class, and μ and r positive integers; and let all those sub-classes of Γ which have exactly r members, or, as we may say, let all r-combinations of the members of Γ be divided in any manner into μ mutually exclusive classes C_i ($i = 1, 2, \ldots, \mu$), so that every r-combination is a member of one and only one C_i; then, assuming the Axiom of Selections, Γ must contain an infinite sub-class Δ such that all the r-combinations of the members of Δ belong to the same C_i.*

The best known lower and upper bounds of $R_k(n)$ are stated in the following theorem. For recent improvement and more detailed overview of the known bounds see paper of Conlon, Fox and Sudakov [CFS11].

Theorem 1.3 (Bounds for Ramsey function). *For any $n \in \mathbb{N}$ $R_2(n) = 2^{\Theta(n)}$ and for $k \geq 3$*

$$\mathrm{twr}_{k-1}(\Omega(n^2)) \leq R_k(n) \leq \mathrm{twr}_k(O(n)),$$

where $\mathrm{twr}_1(x) = x$ *and* $\mathrm{twr}_{i+1}(x) = 2^{\mathrm{twr}_i(x)}$.

1.2 Erdős and Szekeres

Two famous Ramsey-type results of Erdős and Szekeres consider large sets of points in plane.

Theorem 1.4 (Erdős-Szekeres on monotone subsequences [ES35]). *For every positive integer n among every $N = (n-1)^2 + 1$ points $p_1, \ldots, p_N \in \mathbb{R}^2$, where $p_i = (x_i, y_i)$ and $x_1 < \cdots < x_N$, there is a monotone subset of at least n points. This means that there are indices $i_1 < \cdots < i_n$ such that $y_{i_1} \leq \cdots \leq y_{i_n}$ or $y_{i_1} \geq \cdots \geq y_{i_n}$.*

Theorem 1.5 (Erdős–Szekeres on convex/concave configurations [ES35]). *For every positive integer n among every $N = \binom{2n-4}{n-2} + 1 \sim 4^n/\sqrt{n}$ points $p_1, \ldots, p_N \in \mathbb{R}^2$, where $p_i = (x_i, y_i)$ and $x_1 < \cdots < x_N$, there is a convex configuration or a concave configuration of at least n points. This means that there are indices $i_1 < \cdots < i_n$ such that the slopes of the segments $p_{i_j}p_{i_{j+1}}, j = 1, \ldots, n-1$ are either all nondecreasing or all nonincreasing.*

1.3 Fox et al.

A work on related problems were published recently by Fox et al. [FPSS11]. They also considered a problem similar to that of Erdős and Szekeres:

Definition 1 (monotone path [FPSS11]). Let $\mathcal{H}_N^k = ([N], \binom{[N]}{k})$ be a hypergraph and n a positive integer. For $j_1 < \ldots < j_n$ we call sequence of hyperedges

$$\big\{\{j_i, j_{i+1}, \ldots, j_{i+k-1}\}\big\}_{i=1}^{n-k+1}$$

a monotone path of length n.

We denote $N_k(q, n)$ a smallest number such that for every coloring of $\mathcal{H}_{N_k(q,n)}^k$ there exists a monochromatic monotone path of size n.

Theorem 1.6 (on monotone paths [FPSS11]). *For every n and q following holds:*

$$\mathrm{twr}_{k-1}(cn^{q-1}) \leq N_k(q, n) \leq t_{k-1}(c'n^{q-1} \log n)$$

Another important definition is stated in this paper, which we later use to prove our main results.

Definition 2 (Transitive coloring of a hypergraph). Let $\mathcal{H}([N], \binom{[N]}{k}))$ be a k-uniform hypergraph with linear ordering of vertices. We say that a coloring $\chi \colon \binom{[N]}{k} \to \{\pm 1\}$ is transitive if for any vertices $i_1 < \cdots < i_{k+1}$, $i_j \in [n]$ the following holds: whenever $\chi(i_1, \ldots, i_k)$ equals $\chi(i_2, \ldots, i_{k+1})$, then all k-element subsets of $\{i_1, \ldots, i_{k+1}\}$ have the same color.

Two $(k+1)$-tuples from the previous definition which differ only in the first point of the first $(k+1)$-tuple and in the last point of the second $(k+1)$-tuple are called subsequent $(k+1)$-tuples.

1.4 Order types

Order types were considered in the paper by Goodman and Pollack [GP93] and also in [Mat02].

Definition 3 (Order type)**.** Let $P = (p_1, \ldots, p_N)$ be a sequence of points in $\mathbb{R}^d$ and we do not assume the first coordinate to be increasing. Order type of P is a mapping $\chi\colon \binom{[N]}{d+1} \to \{-1, +1\}$. Mapping $\chi(I)$ specifies an orientation of a $(d+1)$-tuple $I = \{i_1, \ldots, i_{d+1}\}$, $i_1 < i_2 < \cdots < i_{d+1}$, where $\chi(I) := \operatorname{sgn}\det M(p_{i_1}, p_{i_2}, \ldots, p_{i_{d+1}})$, where

$$M(q_1, \ldots, q_{d+1}) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_{d+1} \\ | & | & \cdots & | \end{pmatrix},$$

i.e. the j-th column consists of 1 followed by the vector of the d coordinates of q_j.

It is a direct consequence of Ramsey's theorem (Theorem 1.1) that for any d and n there is such N that every sequence of N points in $\mathbb{R}^d$ contains an n-point subsequence with all $(d+1)$-tuples having the same orientation. The smallest such N we denote $\mathrm{OT}_d(n)$. By Theorem 1.3 we get the following upper bound:

Theorem 1.7 (Upper bound from Ramsey's theorem)**.**

$$\mathrm{OT}_d(n) = \mathrm{twr}_{d+1}(O(n))$$

1.5 One-sided sets of hyperplanes

Now we consider a finite sets of hyperplanes in $\mathbb{R}^d$. This problem was previously studied by Matoušek and Welzl [MW92] and later by Dujmović and Langerman [DL11].

Definition 4 (One-sided set of hyperplanes)**.** Let H be a finite set of hyperplanes in $\mathbb{R}^d$ in general position (every d hyperplanes intersect at a single point). We say that H is one-sided if the intersection of every d-tuple from H lies on the same side of the coordinate hyperplane $x_d = 0$.

We denote $\mathrm{OSH}_d(n)$ the Ramsey function for one-sided sets of hyperplanes, i. e. the smallest number N such that any set of N hyperplanes contains a one-sided set of n hyperplanes. The existence of $\mathrm{OSH}_d(n)$ was used by Dujmović and Langerman [DL11] to prove several interesting results. In Section 5.2 we provide a lower bound for $\mathrm{OSH}_d(n)$ which can also be translated into lower bounds for these problems. Here is a direct consequence of Ramsey's theorem:

Theorem 1.8 (Upper bound for one-sided set of hyperplanes)**.**

$$\mathrm{OSH}_d(n) = \mathrm{twr}_k(O(n))$$

1.6 Polynomial interpolation

In this section we mention definitions and theorems which we use mainly in Chapter 2 while developing our definition of kth-order monotonicity and later in Sections 5.1 and 5.2.

Definition 5 (Divided Difference [Phi03, Eq. 1.22])**.** Let $p_1, \ldots, p_{k+1}$ be points in the plane, where $p_i = (x_i, y_i)$ and all x_i are distinct but not necessarily increasing. We define divided difference $\Delta_k(p_1, \ldots, p_{k+1})$ by the following recursive formula:

$$\Delta_0(p_i) := y_i$$
$$\Delta_j(p_{i_1}, p_{i_2}, \ldots, p_{i_{j+1}}) := \frac{\Delta_{j-1}(p_{i_2}, \ldots, p_{i_{j+1}}) - \Delta_{j-1}(p_{i_1}, \ldots, p_{i_j})}{x_{i_{j+1}} - x_{i_1}}$$

For example $\Delta_1(p_1, p_2)$ equals to the slope of the line p_1p_2. It should be noted that Phillips [Phi03] uses different notation for divided differences: $\Delta(p_1, \ldots, p_{k+1})$ is there written as $f[x_1, \ldots, x_{k+1}]$, where f is a function such that $f(x_i) = y_i$ for all points $p_i = (x_i, y_i)$.

Following lemma is mentioned in [Phi03] as a corollary of Theorem 1.1.1:

Lemma 1.9. *The divided difference $\Delta(p_1, \ldots, p_{k+1})$ is a symmetric function of its arguments, meaning that it is unchanged if we rearrange the p_j in any order.*

Another important property of divided difference is the following lemma of Cauchy. It is a generalisation of the Mean Value Theorem.

Theorem 1.10 (Cauchy [Phi03, Eq. 1.33])**.** *Let f be a function such that the kth derivative $f^{(k)}$ exists everywhere on the interval (x_1, x_{k+1}). Let $p_1, \ldots, p_{k+1}$ be the points such that $p_i = (x_i, f(x_i))$. Then there exists $\xi \in (x_1, x_{k+1})$ such that*

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

Now we proceed to interpolation theorems by Newton and Vandermonde.

Theorem 1.11 (Newton's interpolation [Phi03, Eq. 1.19])**.** *Let $p_1, \ldots, p_{k+1} \in \mathbb{R}^2$ where $p_i = (x_i, y_i)$ are points with distinct x-coordinates. Then the unique polynomial f of degree at most k whose graph contains $p_1, \ldots, p_{k+1}$ is given by*

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

Theorem 1.12 (Vandermonde's interpolation [Phi03, Eq. 1.6])**.** *Let $p_1, \ldots, p_{k+1}$, $p_i = (x_i, y_i)$ be points in $\mathbb{R}^2$ with distinct x-coordinates. Then the set of equations*

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{k+1} & x_{k+1}^2 & \cdots & x_{k+1}^k \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{k+1} \end{pmatrix} \tag{1.1}$$

has a unique solution $(a_0, a_1, \ldots, a_k)$ and the polynomial

$$f(x) = a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0$$

interpolates the points $p_1, \ldots, p_{k+1}$.

The leftmost matrix in the equation (1.1) is called Vandermonde matrix. It possesses the following useful property:

Lemma 1.13 (Determinant of Vandermonde matrix [Phi03, Eq. 1.8])**.** *Let V be a Vandermonde matrix as in the equation* (1.1)*. Then*

$$\det V = \prod_{i<j} (x_j - x_i).$$

1.7 Convexity

In this section we provide several definitions of convexity/concavity. Our attempt to generalise these definitions can be found in the next chapter.

Definition 6 (Convex function [Hö94, Def. 1.1.1])**.** Function f is called convex if the graph of f lies below the chord between any two points lying on the graph.

This definition is similar to an ancient definition of concavity by Archimedes from a work called On the Sphere and the Cylinder (Definition 2).

Definition 7 (Concave line by Archimedes[1])**.** I call concave in the given direction a line such that whenever two points are taken, which are on that line, the straight lines between these points fall either in that direction from the line, or some in that direction though some along the line itself, although none in the other direction.

Now we provide a definition of a discrete set of points in the plane similar to Definition 6.

Definition 8 (Convex set of points)**.** Let $P \subseteq \mathbb{R}^2$ be a finite set of points. The set P is called convex if for every two points $p, q \in P$ all points of P with x-coordinate between p and q lie below the segment pq.

[1] Ἐπὶ τὰ αὐτὰ δὴ κοίλην καλῶ τὴν τοιαύτην γραμμήν, ἐν ᾗ ἐὰν δύο σημείων λαμβανομένων ὁποιωνοῦν αἱ μεταξὺ τῶν σημείων εὐθεῖαι ἤτοι πᾶσαι ἐπὶ τὰ αὐτὰ πίπτουσιν τῆς γραμμῆς, ἢ τινὲς μὲν ἐπὶ τὰ αὐτά, τινὲς δὲ κατ᾽ αὐτῆς, ἐπὶ τὰ ἕτερα δὲ μηδεμία. [Šír11], [Arc70]
Maybe "on that side" would be a better translation of "ἐπὶ τὰ αὐτά" although we find "in that direction" more convenient for this definition.

Chapter 2

On kth-order monotonicity

2.1 Introduction

There are two original Erdős-Szekeres theorems: the first one for monotone subsequences (1.4) and the second one for convex/concave configurations (1.5) and we try to generalize these notions of monotonicity and convexity/concavity to higher orders. We know from mathematical analysis that monotonicity is related to the first derivative, and convexity and concavity are related to the second derivative. And actually our initial definition is analytic and we define monotonicity of k-th order using derivatives. The following definition describes the property of points which we look for in our Ramsey-type results.

Definition 9 (kth-order monotonicity)**.** We say that a $(k+1)$-tuple is positive if it lies on the graph of a function whose k-th derivative exists and is everywhere non-negative. On the other hand we say that a $(k+1)$-tuple is negative if it lies on the graph of a function whose k-th derivative exists and is everywhere non-positive. An arbitrary set of points is said to be kth-order monotone if all of its $(k+1)$-tuples are positive or all are negative.

The 1st-order monotonicity is equivalent to monotonicity as in Theorem 1.4 and the 2nd-order monotonicity is equivalent to convexity/concavity as in Theorem 1.5. Although there are some interesting questions. The first one is whether there exists a single function with k-th derivative non-negative/non-positive everywhere which would intersect the whole kth-order monotone set. This question was answered negatively for $k = 3$ by Günter Rote [Rot12], and details and a generalisation for all k are provided in Section 2.4. Another question is what do such kth-order monotone sets look like. The convexity/concavity itself is a geometric concept and in Section 2.3 we provide geometric definitions equivalent to Definition 9.

Sometimes we need the points to be in a "sufficiently" general position so that we have no $(k+1)$-tuples which are both positive and negative:

Definition 10 (k-general position)**.** A set P is in k-general position if no $k+1$ points of P lie on the graph of a single polynomial of degree at most $k-1$.

We denote $ES_k(n)$ the smallest N such that every set of N points in the plain in k-general position contains a kth-order monotone set of size n. The existence of such N is a direct consequence of Ramsey's theorem (1.1).

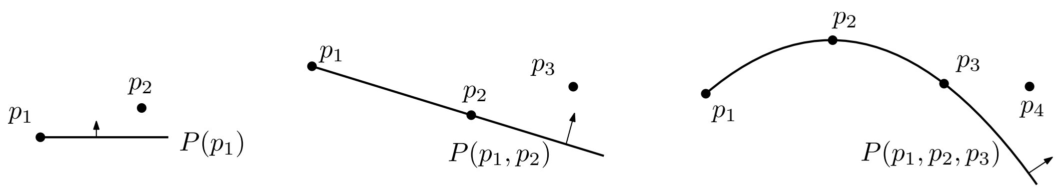

Figure 2.1: monotone, convex and 3rd-order monotone sets

2.2 Definition using divided difference

Theorem 2.1 (Definition of kth-order monotonicity using divided difference)**.** *Let $P = p_1, \ldots, p_n$ be a set of points in $\mathbb{R}^2$. Then P is kth-order monotone if and only if divided differences of all $(k+1)$-tuples of points in P have the same sign.*

Proof of this theorem follows directly from the following lemma.

Lemma 2.2. *A $(k+1)$-tuple K is positive if and only if $\Delta(K) \geq 0$.*
A $(k+1)$-tuple K is negative if and only if $\Delta(K) \leq 0$.

Proof. Let $p_1, \ldots, p_{k+1}$ be points of K where $p_i = (x_i, y_i)$. We use Newton's interpolation (Theorem 1.11) to define $f(x)$ — a polynomial of degree k passing through all points of K:

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

The leading coefficient of $f(x)$ is $\Delta(p_1, \ldots, p_{k+1})$, so that the k-th derivative of the polynomial $f(x)$ is exactly the divided difference of the whole K:

$$f^{(k)}(x) = \Delta(p_1, \ldots, p_{k+1}) = \Delta(K)$$

Now we know that whenever the difference $\Delta(K)$ is non-negative, the $f(x)$ is the function passing through all points of K with k-th derivative existing and everywhere non-negative, and therefore the set K is positive. And similarly whenever $\Delta(K)$ is non-positive, the set K is negative.

To prove the opposite implications we use the Cauchy's Lemma (Theorem 1.10). We consider an arbitrary function f passing through all points of K which has derivatives to the order of k everywhere. Then by Cauchy's Lemma there exists a point $\xi \in (x_1, x_{k+1})$ such that $\Delta(K) \cdot k! = f^{(k)}(\xi)$ and $\operatorname{sgn} \Delta(K) = \operatorname{sgn} f^{(k)}(\xi)$. Therefore if $\Delta(K) < 0$ the K can't be positive and if $\Delta(K) > 0$ then K can't be negative. And the lemma is proved. □

2.3 Geometric interpretation

It is clear that increasing duples and convex triples posses an interesting feature: the last point always lies above the polynomial interpolating rest of the points as is illustrated in Figure 2.1. Generally, this property holds for all orders of monotonicity. Moreover following lemma is true:

Lemma 2.3. *Let $K = \{p_1, \ldots, p_{k+1}\}$, $p_i = (x_i, y_i)$ 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 $\operatorname{sgn} K = (-1)^{k-i}$ if p_i lies below the graph of f_i, and $\operatorname{sgn} K = (-1)^{k+1-i}$ if p_i lies above the graph.*

Proof. Let f be the polynomial of degree at most k passing through all points of K. We use Newton's interpolation (Theorem 1.11), but with the points reordered so that p_i comes last, and we get that

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

Using this with $x = x_i$, we get

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

Divided differences are invariant under permutations (Lemma 1.9), and so $\operatorname{sgn} \Delta_k(p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_{k+1}, p_i) = \operatorname{sgn} K$. 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. □

The following theorem provides a definition of kth-order monotone set similar to that of convex set of points (Definition 8).

Theorem 2.4. *Let $S = \{p_1, \ldots, p_n\}$ be a set of points in $\mathbb{R}^2$ where $p_i = (x_i, y_i)$ and $x_1 < \cdots < x_n$. The set S is kth-order positive if and only if the following holds:*
We choose arbitrary k-tuple $K = \{p_{i_1}, \ldots, p_{i_k}\} \in S$ where $i_1 < \cdots < i_k$ and denote f a polynomial of degree at most $(k-1)$ interpolating all points of K. Then for every $l \geq 1$ each point p_j such that $i_{k-2l} < j < i_{k-2l+1}$ lies above the graph of f, and each point $p_{j'}$ such that $i_{k-2l+1} < j' < i_{k-2l+2}$ lies below the graph of f. Similarly for a kth-order negative set.

Proof. Let S be a kth-order positive set. We assume for a contradiction that there are k points $p_{i_1}, \ldots, p_{i_k}$ and a point p_j that lies on the wrong side of the graph of the polynomial f interpolating points $p_{i_1}, \ldots, p_{i_k}$. W. l. o. g. $i_{k-2l} < j < i_{k-2l+1}$ for some l and p_j lies below the graph of f. Then by Lemma 2.3 the sign of $(k+1)$-tuple $p_{i_1}, \ldots, p_{i_k}, p_j$ is -1 a contradiction.

Let S be a set satisfying the condition of the theorem. Assume for a contradiction that there is a negative $(k+1)$-tuple $\{p_1, \ldots, p_{k+1}\} \subseteq S$. Then by Lemma 2.3 the point p_k lies above the graph of the polynomial interpolating points $p_1, \ldots, p_{k-1}, p_{k+1}$ a contradiction. □

2.4 Nonexistence of a global function

Lemma 2.5 (3rd-order positive set with no global function [Rot12])**.** *There exists a 3rd-order positive set such that there is no function f passing through all points of the set whose 3rd derivative exists and is everywhere positive.*

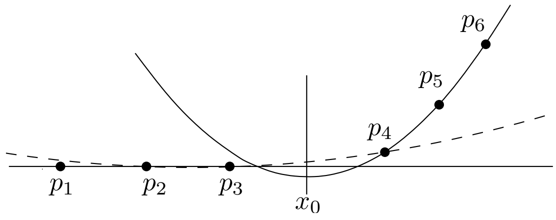

Figure 2.2: Rote's example: a 6-point 3rd-order positive set in 3-general position that does not lie on the graph of any function with nonnegative 3rd derivative.

Proof. Fig. 2.2 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 2.3: 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. And by Cauchy's Lemma (Theorem 1.10) a strictly negative quadruple of points lying on the graph of f implies existence of a point where $f^{(3)}$ is strictly negative — a contradiction. □

Lemma 2.6 (kth-order positive set with no global function for k odd)**.** *For every k odd there exists a kth-order positive set such that there is no function f passing through all points of the set whose k-th derivative exists and is everywhere positive.*

Proof. We use a very similar example as in previous lemma. We have k points $p_1, \ldots, p_k$ on x-axis and another k points $p_{k+1}, \ldots, p_{2k}$ lying above x-axis on the graph of function g which is a function x^{k-1} shifted and scaled such that there are exactly two intersections with x-axis and they occur in the interval (x_k, x_{k+1}) where $(x_i, y_i) = p_i$. Using Lemma 2.3 it can be easily seen that the set of points $p_1, \ldots, p_{2k}$ is kth-order positive. We fix x_0 an x-coordinate between the two intersections of g with the x-axis.

Assuming for contradiction that there is an f passing through all points $p_1, \ldots, p_{2k}$ with k-th derivative everywhere non-negative, we consider the point $q := (x_0, f(x_0))$. For the $(k+1)$-tuple $\{p_1, \ldots, p_k, q\}$ to be positive, q has to lie above the x-axis, but the $(k+1)$-tuple $\{q, p_{k+1}, \ldots, p_{2k}\}$ is positive only if q lies below the graph of g — a contradiction. □

Lemma 2.7 (kth-order positive set with no global function for k even)**.** *For every k even there exists a kth-order positive set such that there is no function f passing through all points of the set whose k-th derivative everywhere exists and is positive.*

Proof. We have again k points on the x-axis and k points on a polynomial of degree $k-1$ as in Figure 2.3. Points $p_{k+1}, \ldots, p_{2k}$ define a unique polynomial g

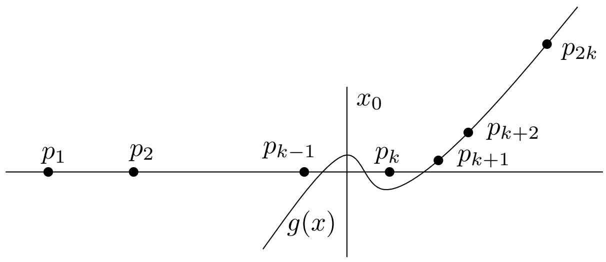

Figure 2.3: kth-order monotone point set with no global function for k even

of degree $k-1$ (odd number) so that the g can be easily enforced to have exactly three intersections with x-axis: one between p_k and p_{k+1} and two between p_{k-1} and p_k. The set $p_1, \ldots, p_{2k}$ is kth-order positive as can be shown using Lemma 2.3. If we want the $(k+1)$-tuple $q, p_{k+1}, \ldots, p_{2k}$ to be positive, we need q to lie above the polynomial g because k is now even. We fix an x-coordinate x_0 between two intersection of g with the x-axis between p_{k-1} and p_k.

Assuming for contradiction that there is an f passing through all points $p_1, \ldots, p_{2k}$ with k-th derivative everywhere non-negative, we consider the point $q := (x_0, f(x_0))$. For the $(k+1)$-tuple $\{p_1, \ldots, p_{k-1}, q, p_k\}$ to be positive, q has to lie below the x-axis, but the $(k+1)$-tuple $\{q, p_{k+1}, \ldots, p_{2k}\}$ is positive only if q lies above the graph of g — a contradiction. □

2.5 Transitivity vs. kth-order monotonicity

In this section we prove that every coloring of k-tuples of points by their sign is transitive. On the other hand, all transitive colorings have a set of points with corresponding signs only in duples.

Lemma 2.8. *Let $P = \{p_1, \ldots, p_N\}$ 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 $L = \{p_1, \ldots, p_{k+2}\}$ with $\operatorname{sgn}\{p_1, \ldots, p_{k+1}\} = \operatorname{sgn}\{p_2, \ldots, p_{k+2}\} = +1$, and we fix $i \in \{2, \ldots, k+1\}$. 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 $\operatorname{sgn}(L \setminus \{p_i\}) = +1$ by Lemma 2.3.

Since $\operatorname{sgn}(L \setminus \{p_1\}) = +1$, we have $f_{1,k+2}(x_{k+2}) < y_{k+2}$ (Lemma 2.3 again), and so it suffices to prove $f_{i,k+2}(x_{k+2}) < f_{1,k+2}(x_{k+2})$.

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 2.3). 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. □

Lemma 2.9. *Let $\chi : \binom{[n]}{2} \to \{-1, +1\}$ be a transitive coloring. We can construct a set of points $p_1, \ldots, p_n \in \mathbb{R}^2$ such that $\operatorname{sgn} \Delta(p_i, p_j) = \chi(i, j)$ for all $i, j \in [n]$.*

Proof. We set $p_1 := (1, 0)$ and continue by induction. We need to add the i-th point, its x-coordinate will be i. Let P be a set of indices $j < i$ such that $\chi(j, i) = +1$. Similarly $N = \{j < i \mid \chi(j, i) = -1\}$. This means that points with indices belonging to P should have smaller y-coordinate then p_i and points with indices belonging to N should have larger. Now we prove that for any $l \in P$ and $m \in N$ the y-coordinate of p_l is smaller then p_m. If so, we can set $p_i = (i, y_i)$ such that $y_l < y_i < y_m$ for all $l \in P$ and $m \in N$.

For contradiction suppose that there is $l \in P$ and $m \in N$ such that y-coordinate of p_l is larger then p_m.

- if $l < m$ then $\chi(l, m) = -1$ but $\chi(m, i) = -1$ and from transitivity also $\chi(l, i) = -1$ — a contradiction with choice of l
- if $l > m$ then $\chi(m, l) = +1$ but $\chi(l, i) = +1$ and from transitivity also $\chi(m, i) = +1$ — a contradiction with choice of m

□

Lemma 2.10. *For every $k \geq 2$ there is a transitive coloring $\chi\colon \binom{[k+2]}{k+1} \to \{-1, +1\}$ such that there is no set of points $p_1, \dots, p_{k+2}$ such that for every $(k+1)$-tuple $\{i_1, \dots, i_{k+1}\} \in \binom{[k+2]}{k+1}$ is $\chi(i_1, \dots, i_{k+1}) = \operatorname{sgn} \Delta(p_{i_1}, \dots, p_{i_{k+1}})$.*

Proof. We set $\chi(1, \dots, k+1) := +1$ and $\chi(2, \dots, k+2) := -1$. Now we can define colors of the rest of the $(k + 1)$-tuples arbitrarily and χ will be still transitive. We set just $\chi(1, \dots, k, k + 2) := -1$ and $\chi(1, 3, \dots, k + 2) := +1$; the coloring of other $(k + 1)$-tuples is not important.

Assume for a contradiction that there exists a set of points $P = \{p_1, \dots, p_{k+2}\}$, $p_i = (x_i, y_i)$, $x_1 < \cdots < x_{k+2}$ such that for all $i = 1, 2, k+1, k+2$ following holds: $\operatorname{sgn} \Delta(P \setminus \{p_i\}) = \chi([k + 2] \setminus \{i\})$. We denote $f_{i,j}$ a polynomial of degree at most $k-1$ interpolating points of $P \setminus \{p_i, p_j\}$. The point p_{k+1} must lie above the graph of $f_{k+1,k+2}$ thanks to Lemma 2.3. Since $\operatorname{sgn} \Delta(P \setminus \{p_{k+1}\}) = -1$ the point p_{k+2} must lie bellow the graph of $f_{k+1,k+2}$ and since $\operatorname{sgn} \Delta(P \setminus \{p_2\}) = +1$, p_{k+2} must lie above the graph of $f_{2,k+2}$.

We know that $p_1, p_3, \dots, p_k$ are the only intersections of $f_{k+1,k+2}$ and $f_{2,k+2}$ as they are both of degree at most $k - 1$. But $f_{2,k+2}$ passes through the point p_{k+1} which lies above $f_{k+1,k+2}$ and therefore $f_{2,k+2}(x) > f_{k+1,k+2}(x)$ for all $x > x_k$. But we need $f_{k+1,k+2}(x_{k+2}) > y_{k+2} > f_{2,k+2}(x_{k+2})$ — a contradiction. □

Chapter 3

Ramsey numbers for transitive colorings

In this chapter we provide bounds for Ramsey numbers for transitively colored hypergraphs. We denote $\mathrm{R}_k^{trans}(n)$ the corresponding Ramsey function. We have an asymptotically matching bounds for k up to 4. For greater k there is an upper bound $\mathrm{R}_k^{trans}(n) = \mathrm{twr}_{k-1}(O(n))$ and the only known lower bound for $k \geq 4$ is $\mathrm{R}_k^{trans}(n) \geq \mathrm{R}_4^{trans}(n-k+4)$.

3.1 Upper bounds

We know that coloring of $k+1$-tuples of points in plane by their signs is transitive (Lemma 2.8). Now we will explore properties of transitive colorings. Firstly we will show, that both Erdős-Szekeres theorems hold for transitively colored hypergraphs.

Theorem 3.1 (Erdős-Szekeres theorem for transitively colored duples)**.**

$$\mathrm{R}_2^{trans}(n) \leq (n-1)^2 + 1$$

Proof. Let χ be a transitive coloring of $\mathcal{H}([N], \binom{[N]}{2})$. Thanks to Lemma 2.9 we can construct a sequence of points $p_1, \ldots, p_N \in \mathbb{R}^2$ such that for $i < j$ a point p_i has lower y-coordinate then p_j whenever $\chi(i,j) = -1$ and greater whenever $\chi(i,j) = +1$. Using Theorem 1.4 we find a monotone subsequence $p_{i_1}, \ldots, p_{i_n}$ and the set of indices $\{i_1, \ldots, i_n\}$ is a requested monochromatic subset of $\mathcal{H}$. □

Transitivity and ordering of vertices by x-coordinate are the only properties of convex sets in the plane that are used by original proof of Erdős and Szekeres so the proof of the next theorem is simply repeating it word by word. We provide it for completeness.

Theorem 3.2 (Erdős-Szekeres theorem for transitively colored triples)**.** *For any $n \in \mathbb{N}$ there is N of size $O(2^n)$, that every hypergraph $\mathcal{H}\big([N], \binom{[N]}{3}\big)$ with linearly ordered vertices and a transitive coloring of its edges have a monochromatic subset of size at least n.*

Proof. Let $N := f_2(k,l)$ be a number that any transitively colored hypergraph $\mathcal{H}\big([N], \binom{[N]}{3}\big)$ has a positive subset of size k or a negative subset of size l for $k, l \geq 3$. It is clear that $f_2(3,l) = f_2(k,3) = 3$.

We will prove the following recursion:

$$f_2(k,l) = f_2(k-1,l) + f_2(k,l-1) - 1$$

We suppose by induction that the theorem is proved for $f_2(k-1,l)$ and for $f_2(k,l-1)$. Let $N = f_2(k-1,l)+f_2(k,l-1)-1$. We consider arbitrary hypergraph $\mathcal{H}\big([N], \binom{[N]}{3}\big)$ and his transitive coloring χ. Let us suppose that $\mathcal{H}$ does not contain a negative set of size l. Then $\mathcal{H}$ must contain a positive set of size $k-1$. We denote $E = \{p \mid p \text{ is the last vertex of some positive set with } k-1 \text{ vertices}\}$. Then $[N] \setminus E$ does not contain a positive subset of size $k-1$ which implies that $\big|[N] \setminus E\big| < f_2(k-1,l)$ and therefore $|E| \geq f_2(k,l-1)$. If E contains a positive set of size k, then we are done. Else it must contain a negative set $N = \{p, y, y_3, \dots, y_{l-1}\}$. There should be a subset $P = \{x_1, \dots, x_{k-3}, x, p\} \subseteq [N]$. Now we show that if $\chi(x,p,y) = 1$ then set $P \cup \{y\}$ is positive and if $\chi(x,p,y) = -1$ then set $\{x\} \cup N$ is negative.

W. l. o. g. let $\chi(x,p,y) = 1$. For every $x' \in N \setminus \{x,p\}$ there is $\chi(x',x,p) = 1$ and $\chi(x,p,y) = 1$ so from transitivity also $\chi(x',p,y) = \chi(x',x,y) = 1$. For every $x', x'' \in N \setminus \{x,p\}, x'' < x$; there is $\chi(x'',x',p)$, $\chi(x',x,p) = 1$ and $\chi(x',x,p) = 1$. And therefore from transitivity $\chi(x'',x',y) = 1$. We have considered all triples containing y so the theorem is proved. □

Now we know the value of $R_3^{\text{trans}}(n)$. We use proof of Ramsey's theorem for hypergraphs by Erdős and Rado [ER52]. Nevertheless our induction has a better start thanks to Theorem 3.2. We will use the following lemma:

Lemma 3.3. *Let χ be a transitive coloring of k-tuples of set $[n]$ so that for all $i_1 < \cdots < i_{k-1} < n$ and every $i_k \geq i_{k-1}+1$ there is $\chi(i_1, \dots, i_{k-1}, i_k) = \chi(i_1, \dots, i_{k-1}, i_{k-1}+1)$. We define χ^* coloring of $(k-1)$-tuples by following:*

$$\chi^*(i_1, \dots, i_{k-1}) := \chi(i_1, \dots, i_{k-1}, i_{k-1}+1)$$

Then χ^* is transitive on $[n-1]$.

Proof. Let we chose $i_1 < i_2 < \cdots < i_k < n$. From the definition of χ^* we now that following holds:

$$\chi^*(i_1, \dots, i_{k-1}) = \chi(i_1, \dots, i_k)$$
$$\chi^*(i_2, \dots, i_k) = \chi(i_2, \dots, i_k, i_k+1)$$

In the case that $\chi^*(i_1, \dots, i_{k-1})$ equals $\chi(i_1, \dots, i_k)$, the colloring $\chi(i_1, \dots, i_k)$ must be equal to $\chi(i_2, \dots, i_k, i_k+1)$. If we chose arbitrary $(k-1)$-tuple A from $i_1, \dots, i_k$, we know, that

$$\chi^*(A) = \chi(A \cup \{i_k+1\}) = \chi(x_2, \dots, x_k, x_k+1) = \chi^*(i_2, \dots, i_k)$$

because χ is transitive. □

Theorem 3.4 (Ramsey theorem for transitively colored hypergraphs)**.** *For any $k \in \mathbb{N}$ there exist $N \in \mathbb{N}$ of size $\mathrm{twr}_k(O(n))$ such that in every hypergraph $\mathcal{H} = \big([N], \binom{[N]}{k}\big)$ with a transitive coloring χ there is a monochromatic set of size n.*

Proof. We proceed by induction. For $k = 3$ this is implied by Theorem 3.2. We set $M := R_k^{\text{trans}}(n)$, $N := 2^{M^k}$ and show that $R_{k+1}^{\text{trans}}(n) \le N$ for $k > 3$.

Let $\chi\colon \binom{[N]}{k+1} \to \{\pm 1\}$ be an arbitrary transitive 2-coloring of $[N]$. We iteratively construct a sequence $a_1 < \cdots < a_M$ so that χ has all properties demanded by lemma 3.3 on $A = \{a_1, \dots, a_M\}$. At the beginning we set $A_{k-1} := \{1, 2, \dots, k-1\}$ and $X_{k-1} := [N] \setminus A_{k-1}$. For $i = k, k+1, \dots, M$ we construct sets $A_i, X_i \subseteq [N]$ in the following way:

1. We say that x and y are equivalent if for all $T \in \binom{A_{i-1}}{k-1}$ is $\chi(T \cup \{x\})$ equal to $\chi(T \cup \{y\})$. Let C be the largest of the equivalence classes.
2. We chose a_i the smallest element of C and $X_i := C \setminus \{a_i\}$.

In i-th step there are $2^{\binom{i}{k-1}} < 2^{M^{k-1}}$ equivalence classes and therefore $|X_i| \ge |X_{i-1}| \cdot 2^{-\binom{i}{k-1}} - 1 \ge |X_{i-1}| \cdot 2^{M^{k-1}}$. Clearly we can make M steps because $|X_M| \ge N \cdot (2^{-M^{k-1}})^M = 2^{M^k} \cdot 2^{-M^k} = 1$. We set $A := A_M$.

Coloring χ restricted to A is clearly transitive. Let x be the smallest element of X_M (x does not belong to A). Coloring χ satisfies requirements of Lemma 3.3 on $A \cup \{x\}$. Now we define a coloring $\chi^*\colon \binom{A}{k} \to \{\pm 1\}$ by $\chi^*(K) := \chi(K \cup \{x\})$. Then by Lemma 3.3 we know that χ^* is transitive. From inductive hypothesis there is a monochromatic n-element subset of A with respect to χ^*. Clearly A is monochromatic also with respect to χ. □

3.2 Lower bounds

Lower bounds for R_2^{trans} and R_3^{trans} are direct consequences of Theorems 1.4, 1.5 and Lemma 2.8:

Theorem 3.5 (Lower bounds for R_2^{trans} and R_3^{trans})**.** *For all $n \in \mathbb{N}$ following holds:* $\mathrm{R}_2^{trans}(n) \ge (n-1)^2 + 1$ *and* $\mathrm{R}_3^{trans}(n) = \Omega(2^n)$.

Proof. We know that the coloring of duples resp. triples of vertices by their sign is transitive (Lemma 2.8). Therefore the examples giving the lower bounds of 1.4 and 1.5 can be directly transformed to examples which imply lower bounds for R_2^{trans} and R_3^{trans}. □

Now we provide a construction of a large 4-uniform hypergraph with only a small monochromatic subset.

Theorem 3.6 (Lower bound for R_4^{trans})**.** *For all $n \ge 2$ we have*

$$\mathrm{R}_4^{trans}(2n+1) \ge 2^{2^{n-1}} + 1$$

This means that $\mathrm{R}_4^{trans}(n) = \mathrm{twr}_3(\Omega(n))$.

Proof. Inductively we construct a 4-uniform hypergraph $\mathcal{H}_n$ with no monochromatic subgraph of size $2n+1$. We begin with $\mathcal{H}_2$ on $2^{2^n} = 4$ vertices with one hyperedge which has clearly no monochromatic set of size 5.

In the inductive step we replace every vertex of $\mathcal{H}_n$ with a new copy of $\mathcal{H}_n$. Formally we can write that $V(\mathcal{H}_{n+1}) = \{(u, v) \mid u, v \in V(\mathcal{H}_n)\}$. Vertices will be ordered in an alphabetic ordering, which is illustrated in Figure 3.1. The copies of

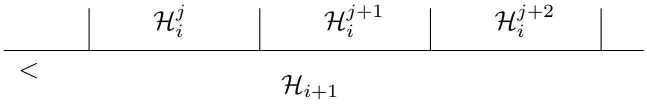

Figure 3.1: vertex ordering in $\mathcal{H}_{n+1}$

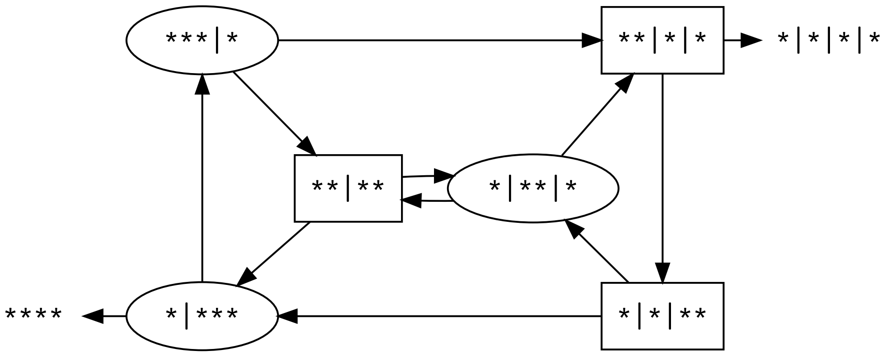

Figure 3.2: Coloring of quadruple types: symbol '|' means border of a cluster, oval frame means negative and square frame means positive color. All potentially subsequent types are connected with an arrow.

$\mathcal{H}_n$ are called clusters. Since $|V(\mathcal{H}_{n+1})| = |V(\mathcal{H}_n)|^2$ we have $|V(\mathcal{H}_{n+1})| = 2^{2^{n+1}}$.

Coloring of a quadruple depends on its type. The type is an ordered partition of 4 given by the distribution of the quadruple among clusters. All quadruples whose vertices belong to the same cluster (type 4) have the same color as in $\mathcal{H}_n$. All quadruples which have all vertices from a different cluster (type 1+1+1+1) also have the same color as in $\mathcal{H}_n$. For the rest of the types, there is a constant color for each type. The colors are following:

- 3+1 has color -1
- 2+2 has color $+1$
- 1+3 has color -1
- 2+1+1 has color $+1$
- 1+2+1 has color -1
- 1+1+2 has color $+1$

The coloring of the types is illustrated in Figure 3.2

Now we need to show that every monochromatic set in $\mathcal{H}_{n+1}$ can be of size at most $2n+2$ and that suggested coloring is transitive.

Let X be a monochromatic subset of $V(\mathcal{H}_{n+1})$ and Y be the largest subset of X which is contained inside a single cluster. From induction we know that $|Y| \leq 2n$ and that X can cross at most $2n$ clusters.

If $|Y| = 1$, then every points of X belongs to a different cluster and therefore $|X| \leq 2n$.

If $|Y| \geq 3$ then we will show, that X contains only one point smaller and only one point greater than all points of Y. For a contradiction, assume that there are points $x_1 < x_2 \in X \setminus Y$ such that for all $y \in Y$ x_1, x_2 are smaller than y.

Let $y_1 < y_2 < y_3$ be points of Y. Then $(x_2; y_1, y_2, y_4)$ is a quadruple of type 1+3, but $(x_1, x_2; y_1, y_2)$ is a quadruple of type 2+2 which has the opposite color — a contradiction. Similarly for two points of $X \setminus Y$ greater than all points of Y.

If $|Y| = 2$ and there is at most one different cluster which contains two points of X, then $|X| \leq 2n + 2$. Otherwise let $a_1 < a_2 < b_1 < b_2 < c_1 < c_2$ be points of X such that a_1 and a_2 belong to the same cluster, similarly for b_1, b_2 and c_1, c_2. Then $(a_2; b_1, b_2; c_1)$ is a quadruple of type 1+2+1 and $(a_1, a_2; b_1, b_2)$ is a quadruple of type 2+2 which has the opposite color — a contradiction. □

Theorem 3.7 (Lower bound for $\mathrm{R}_k^{trans}, k > 4$)**.** *For any $k > 4$ the following holds:*

$$\mathrm{R}_k^{trans}(n) \geq \mathrm{R}_4^{trans}(n - k + 4)$$

Proof. We proceed by induction. The case $k = 4$ is proved due to Theorem 3.5. We have a transitive coloring $\chi \colon \binom{[\mathrm{R}_k^{trans}(n)]}{k} \to \{-1, +1\}$ with no monochromatic n-point set and we construct a transitive coloring of $(k+1)$-tuples of the same set of points with no monochromatic $(n+1)$-point set.

We define χ^*, the coloring of $(k+1)$-tuples, as follows:

$$\chi^*(i_1, \ldots, i_k, i_{k+1}) := \chi(i_1, \ldots, i_k)$$

Now we prove that χ^* is transitive. Let $\chi^*(i_1, \ldots, i_{k+1}) = \chi^*(i_2, \ldots, i_{k+2}) = c$. Then $\chi(i_1, \ldots, i_k) = \chi(i_2, \ldots, i_{k+1}) = c$ and by transitivity also every k-tuple $K \in \binom{\{i_1, \ldots, i_{k+1}\}}{k}$ have the same color. Let I be a $(k+1)$-tuple of elements from $\{i_1, \ldots, i_{k+2}\}$ and K be a set of first k points of I. It is clear that $K \subseteq \{i_1, \ldots, i_{k+1}\}$ and therefore $\chi(K) = c$ and by the definition of χ^* also $\chi^*(I) = c$.

Let $S = \{i_1, \ldots, i_{n+1}\}$ be a monochromatic set with respect to χ^*. Then by the definition of χ^* the set $\{i_1, \ldots, i_n\}$ is monochromatic with respect to χ — a contradiction. □

Chapter 4

Ramsey numbers for kth-order monotone subsets

The upper bound is a simple corollary of Theorem 3.4.

Theorem 4.1 (Upper bound for ES_k)**.** *For every $k \geq 2$ the following bound holds:*

$$\mathrm{ES}_k(n) = \mathrm{twr}_k(O(n))$$

Proof. We know that every coloring of $(k+1)$-tuples by their signs is transitive (Lemma 2.8) so that the bound of the Theorem 3.4 applies also on kth-order monotone subsequences of points. □

Now we use the Theorem 3.6 to construct a point set with only a small 3rd-order monotone subsets.

Theorem 4.2 (Lower bound for ES_3)**.** *For all $n \geq 2$ we have*

$$\mathrm{ES}_3(2n+1) \geq 2^{2^{n-1}} + 1$$

The proof of this theorem can be found in the end of this chapter after details of the construction are provided.

We start with P_2 as an arbitrary set of $2^{2^1} = 4$ points in $\mathbb{R}^2$ in 3-general position. Now we proceed similarly as in the proof of Theorem 3.6: when constructing the set P_{n+1} from P_n we replace each point of P_n by a tiny and deformed copy of P_n. We use the deformation to enforce the same coloring of the types of quadruples as in the Theorem 3.6.

1. By an affine transformation we make sure that P_n is inside $[1,2] \times [0,1]$ or better inside $[1,1.9] \times [0,0.9]$ so that we have enough room for perturbation.

2. There is a small δ such that in a set P' obtained from P_n by moving each point arbitrarily by at most δ, the P' stays in general position and moreover the ordering of vertices by x-coordinate and signs of all quadruples of points stays the same.

3. We choose a sufficiently large number $A = A(P_n)$ as in Lemma 4.4 and we set $\varepsilon := \frac{1}{A^2}$.

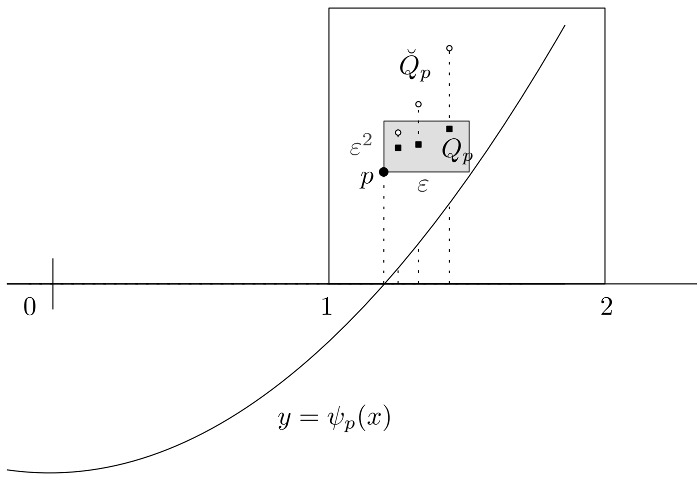

Figure 4.1: A schematic illustration of the construction of P_{n+1}.

4. For every point $p \in P_n$ let Q_p be the image of P_n 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 the lower left corner at p, see Figure 4.1.

5. Let $\psi_p(x) = Ax^2 + C_p$ be a quadratic function where C_p is a constant chosen so that $\psi_p(x_p) = 0$ (x_p is an x-coordinate of the point p). Let $\breve{Q}_p$ 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 $P_{n+1} := \bigcup_{p \in P_n} \breve{Q}_p$. We call the $\breve{Q}_p$ the *clusters* of P_{n+1}. As these transformations does not affect a 3-general position of $\breve{Q}_p$, the whole P_{n+1} is also in 3-general position.

Lemma 4.3. *Each $\breve{Q}_p$ is contained in an $O(\sqrt{\varepsilon}\,)$-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

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

□

Lemma 4.4 (Slope lemma)**.** *There is a constant A depending only on P_n such that following holds: Whenever λ is a parabola passing through three points of P_{n+1} each from a different cluster or a line passing through two points from different clusters and μ is a parabola passing through points inside a single cluster or a line passing through two such points then the maximum slope of λ on the interval* $[1,2]$ *is smaller then the minimum slope of μ on* $[1,2]$.

Proof. Clearly, the maximum slope of any such λ can be bounded above by some finite number depending only on P_n itself. It suffices to show that in every $\breve{Q}_p$ the parabola or line μ defined by points of $\breve{Q}_p$ has the minimum slope on $[1,2]$ bounded from below by A.

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

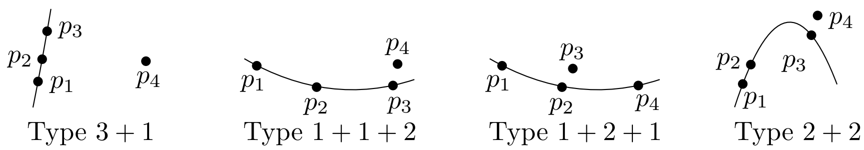

Figure 4.2: Determining the signs of quadruples by type

By the construction of $\breve{Q}_p$, the affine map transforming P_n 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 doesn't 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 $q, r \in \breve{Q}_p$. Let us choose another point $s \in \breve{Q}_p$ and consider the parabola μ' through q, r, s. By 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. □

Proof of theorem 4.2. We know that $|P_{n+1}| = |P_n|^2$ and therefore $|P_n| = 2^{2^{n-1}}$. Now we prove that in every induction step the coloring of all types of quadruples is the same as in the proof of Theorem 3.6 and thereby P_n does not contain any 3rd-order monotone subset of $2n + 1$ points.

Several types are illustrated in Figure 4.2. We denote the points in a quadruple p_1, p_2, p_3, p_4.

- type 1+1+1+1: thanks to the definition of δ and Lemma 4.3 we know that the sign of this type as was in P_n is preserved
- type 4: affine transformation and adding a polynomial to the point set does not change the signs of quadruples
- type 3+1: from Lemma 4.4 we know that the parabola through p_1, p_2, p_3 lies above p_4 and from Lemma 2.3 the sign of the quadruple is -1.
- type 1+3: the parabola through p_2, p_3, p_4 lies belove p_1 and therefore the sign is -1.
- type 1+1+2: the segment p_3p_4 is steeper then the parabola through p_1, p_2, p_3 as in Figure 4.2 and the sign is $+1$.
- type 2+1+1: the segment p_1p_2 is steeper then the parabola through p_1, p_3, p_4 which means that p_2 is above this parabola and by Lemma 2.3 the sign is $+1$.
- type 1+2+1: the segment p_2p_3 is steeper then the parabola through p_1, p_2, p_4 as in Figure 4.2 and the sign is -1
- type 2+2: This is the most complex case: We know that the segment p_1p_2 is steeper than p_2p_3 thus the parabola through p_1, p_2, p_3 must be concave.

Therefore its slope by p_3 is no larger than the slope of segment p_2p_3 which is smaller then the slope of segment p_3p_4. Thus, the point p_4 lies above the parabola through p_1, p_2, p_3 and the sign is $+1$.

We have analyzed all possible types and the proof is finished. □

Chapter 5

Applications

5.1 Order type problem

There does not seem to be any kind of transitivity in order types and therefore the only upper bound we know is that of Theorem 1.7. Although all examples for kth-order monotone subsets can be converted to examples for order type. This way we get the same lower bounds as for $\mathrm{ES}_k(n)$.

Theorem 5.1 (Lower bound for Order types)**.** *For all $d \geq 1$ following holds:*

$$\mathrm{OT}_d(n) \geq \mathrm{ES}_d(n)$$

Specially $\mathrm{OT}_3(n) = 2^{2^{\Omega(n)}}$.

Proof. Let $p_1, \ldots, p_{\mathrm{ES}_d(n)} \in \mathbb{R}^2$ be a sequence of points ordered by x-coordinate with no dth-order monotone subset of length $n+1$ as in Theorem 4.2. We denote $p_i = (x_i, y_i)$. Now we construct a set of points $q_1, \ldots, q_{\mathrm{ES}_d(n)} \in \mathbb{R}^d$ where $q_i = (x_i, x_i^2, \ldots, x_i^{d-1}, y)$.

The proof of the theorem follows from the fact that the sign of the orientation of every $(d+1)$-tuple $q_{i_1}, \ldots, q_{i_{d+1}}$ is the same as the sign of the $(d+1)$-tuple $p_{i_1}, \ldots, p_{i_{d+1}}$ where $i_1 < \cdots < i_{d+1}$.

To prove this we use Vandermonde interpolation to get the sign of the coefficient near x^d of the polynomial interpolating points $p_{i_1}, \ldots, p_{i_{d+1}}$. This is the Vandermonde matrix of these points:

$$V := \begin{pmatrix} 1 & x_{i_1} & x_{i_1}^2 & \cdots & x_{i_1}^d \\ 1 & x_{i_2} & x_{i_2}^2 & \cdots & x_{i_2}^d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{i_{d+1}} & x_{i_{d+1}}^2 & \cdots & x_{i_{d+1}}^d \end{pmatrix}$$

Let $a = (a_0, \ldots, a_d)$ be the coefficients of the interpolating polynomial, this means a is a solution to the system $Va = y$ where $y = (y_{i_1}, \ldots, y_{i_{d+1}})$ are y-coordinates of points $p_{i_1}, \ldots, p_{i_{d+1}}$. For the coefficient a_d we get the following equation by Cramer's rule:

$$a_d = \frac{\det V_{(d+1)\to y}}{\det V}$$

Thanks to Lemma 1.13 the determinant $\det V$ is always positive because $x_{i_1} < \cdots < x_{i_{d+1}}$. So that $\operatorname{sgn} \Delta(p_{i_1}, \ldots, p_{i_{d+1}}) = \operatorname{sgn} a_d = \operatorname{sgn} \det V_{(d+1)\to y}$. Moreover the matrix $V_{(d+1)\to y}$ is in fact transposed matrix $M(q_{i_1}, \ldots, q_{i_{d+1}})$ and therefore

$$\operatorname{sgn} \Delta(p_{i_1}, \ldots, p_{i_{d+1}}) = \operatorname{sgn} \det M(q_{i_1}, \ldots, q_{i_{d+1}})^T = \operatorname{sgn} \det M(q_{i_1}, \ldots, q_{i_{d+1}})$$

and this is the definition of the orientation of $q_{i_1}, \ldots, q_{i_{d+1}}$. □

5.2 One-sided sets of hyperplanes

The situation is similar to the order type problem. The only known upper bound is that of Theorem 1.8 although the lower bounds for $\mathrm{ES}_{d-1}(n)$ apply also on $\mathrm{OSH}_d(n)$.

Theorem 5.2 (Lower bound for one-sided sets of hyperplanes)**.** *For any $d \geq 2$ following holds:*

$$\mathrm{OSH}_d(n) \geq ES_{d-1}(n)$$

Proof. Let $N := \mathrm{ES}_{d-1}(n)$ and $P = \{p_1, \ldots, p_N\}$ where $p_i = (x_i, y_i)$ be a set of points in $(d-1)$-general position with no $(d-1)$th-order monotone set of size $n+1$. We define a collection $H = \{h_1, \ldots, h_N\}$ of hyperplanes in $\mathbb{R}^d$ such that

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

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 $(\det V_{(d+1)\to y})/(\det V)$.

As we saw in the proof of Theorem 5.1, $(\det V_{(d+1)\to y})/(\det V)$ 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 $(d-1)$th-order monotone subsets in P, and the theorem is proved. □

Chapter 6

Open problems

Lower bounds for $\mathrm{ES}_k(n)$. We have obtained reasonably tight bounds for $\mathrm{ES}_3(n)$, but the gaps are much more significant for $\mathrm{ES}_k(n)$ with $k \geq 4$. 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, but a possibility of a better upper bound shouldn't also be overlooked. This question looks both interesting and challenging.

Lower bounds for $\mathrm{R}_k^{trans}(n)$. A perhaps more manageable task might be a better lower bound for $\mathrm{R}_k^{trans}(n)$, $k \geq 4$. 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., [CFS11]). 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.

Bounds for Order types. For $\mathrm{OT}_3(n)$ we have the lower bound of $2^{2^{\Omega(n)}}$, but upper bound only $\mathrm{twr}_4(O(n))$ directly from Ramsey's theorem. It seems that the colorings given by the order type are not transitive in any reasonable sense, and we have no good guess of which of the upper and lower bounds should be closer to the truth.

Bounds for One-sided sets of hyperplanes. Similar comments apply to the problem with one-sided subsets of planes in $\mathbb{R}^3$ (concerning $\mathrm{OSH}_3(n)$), and the higher-dimensional cases are even more widely open.

Monotone paths. Another interesting question is 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. [FPSS11].

Characterization of sets having global function. 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 k-general position, a single function should suffice for all $(k+1)$-tuples; in other words, that every kth-order monotone finite set finite set in k-general position lies on a graph of a k-times differentiable function $f\colon \mathbb{R} \to \mathbb{R}$ whose kth derivative is everywhere nonnegative or everywhere nonpositive. However this is disproved now in Section 2.4. 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?

Bibliography

[Arc70] Archimède. *I: De la sphère et du cylindre.* Ch. Mugler, Paris, 1970.

[CFS11] D. Conlon, J. Fox, and B. Sudakov. An improved bound for the stepping-up lemma. *Discrete Applied Mathematics*, 2011. In press.

[DL11] V. Dujmović and S. Langerman. A center transversal theorem for hyperplanes and applications to graph drawing. In *Proc. 27th ACM Symposium on Computational Geometry*, pages 117–124, 2011. Full version arXiv:1012.0548.

[EM11] Marek Eliáš and Jiří Matoušek. Higher-Order Erdős–Szekeres Theorems. Arxiv preprint 1111.3824, 2011. *Proc. ACM Symposium on Computational Geometry*, in press.

[ER52] P. Erdős and R. Rado. Combinatorial theorems on classifications of subsets of a given set. *Proc. London Math. Soc.*, 3:417–439, 1952.

[ES35] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[FPSS11] J. Fox, J. Pach, B. Sudakov, and A. Suk. Erdős–Szekeres-type theorem for monotone paths and convex bodies. Arxiv preprint 1105.2097v1, 2011. *Proc. London Math. Soc.*, in press.

[GP93] J. E. Goodman and R. Pollack. Allowable sequences and order types in discrete and computational geometry. In J. Pach, editor, *New Trends in Discrete and Computational Geometry*, volume 10 of *Algorithms and Combinatorics*, pages 103–134. Springer, Berlin etc., 1993.

[GRS80] R. Graham, B. Rothschild, and J. Spencer. *Ramsey Theory.* Wiley-Interscience, New York, 1980.

[Hö94] Lars Hörmander. *Notions of Convexity.* Birkhäuser, Boston, 1994.

[Mat02] J. Matoušek. *Lectures on Discrete Geometry.* Springer, New York, 2002.

[MW92] J. Matoušek and Emo Welzl. Good splitters for counting points in triangles. *J. Algorithms*, 13:307–319, 1992.

[Phi03] George M. Phillips. *Interpolation and approximation by polynomials.* Springer, Berlin etc., 2003.

[Ram30] F. P. Ramsey. On a problem of formal logic. *Proc. London Math. Soc.*, 30:264–286, 1930.

[Rot12] G. Rote. Private communication, February 2012.

[Šír11] Zbyněk Šír. *Řecké matematické texty*. OIKOYMENH, Prague, 2011.