Coloring Geometric Hypergraphs

Dömötör Pálvölgyi

Doctoral dissertation submitted to the
Hungarian Academy of Sciences

Budapest, 2023

Contents

1 Introduction

Recent developments in technology led to many new fascinating challenges in computational and combinatorial geometry. Nowadays we are all constantly surrounded by electronic gadgets fulfilling various roles. The complex interconnection network of these objects can be described by geometric graphs and hypergraphs. The edges of such graphs are defined by the metric properties of the space they are embedded in. For example, in case of a surveillance problem, an area might be monitored by sensors with a given location and range; here the the points of the area form the vertices of the hypergraph and the range of each sensor forms a hyperedge.

Any mathematical problem that involves partitioning into groups can be modeled and visualized by colorings. These include many practical problems, such as job scheduling and bin packing, that have important real world applications. For example, in case of the surveillance problem, suppose that each sensor also has an associated lifetime for which it can remain active, and our goal is to create a time schedule which determines when each sensor is active, so that the whole area is constantly monitored, for as long as possible. Or, in case of solar powered devices with a fixed amount of active hours per day, our goal is to determine the feasibility of maintaining the surveillance all day. Suppose that the lifetime of each sensor is the same, and for each we need to pick a time slot during which it stays active. The different time slots can be represented by the colors, and the question becomes a coloring problem in a hypergraph.

There are many different questions one can ask about graph colorings, as these notions are general enough to capture a wide variety of problems. The most famous example, Ramsey's theorem about coloring graphs without large monochromatic cliques, comes up in several seemingly unrelated fields and have motivated a large part of the research in combinatorics. This result is, in fact, about hypergraph colorings, where the vertices of the hypergraph are the edges of the underlying graph, and the hyperedges are the cliques. In case of hypergraph colorings, there are several versions of colorings one can consider, all generalizing the notion of proper coloring for graphs. Depending on the condition required to be satisfied for each single hyperedge, we can differentiate several colorings; these include proper coloring, polychromatic coloring, rainbow coloring, strong coloring, conflict-free coloring, unique-maximum coloring, and odd coloring.

The topic of this dissertation will primarily be proper and polychromatic coloring of geometric hypergraphs. The systematic study of such questions was initiated by Pach in 1980, but some examples go back earlier. Perhaps the most famous example is the Hales-Jewett theorem, which can be geometrically phrased as follows. For every naturals k and m there is a d such that in any k-coloring of the points of the d-dimensional grid of width m, having m^d points, there is a line through m of the points that are all of the same color.

We will also primarily focus on hypergraphs where each hyperedge contains many points. A typical question that we will try to answer is the following. For which naturals k and m can we k-color any planar point set such that every disk containing at least m points has at least two points of different colors?

1.1 Abstract hypergraphs

There are several different terminologies for hypergraphs. We mainly follow the one used by extremal combinatorists; see [36].

A hypergraph $\mathcal{H} = (V, E)$ is a collection of sets E over a base set V. The elements of V are called the vertices of the hypergraph and the elements of E the hyperedges, or just simply the edges of the hypergraph. For a hypergraph $\mathcal{H}$, we denote its vertices by $V(\mathcal{H})$ and its edges by $E(\mathcal{H})$; the $\mathcal{H}$ is omitted when it leads to no confusion. If all sets of E are different, the hypergraph is simple; otherwise, it is called multihypergraph. For technical reasons, we will assume that $\emptyset \notin E$ and $E \neq \emptyset$. A hypergraph is finite if V and E are finite sets. A hypergraph is m-uniform if all of its edges have size m and it is m-heavy if all of its edges have size at least m. Edges with at least m vertices are called m-heavy; the m is omitted when it leads to no confusion.

The incidence matrix $M(\mathcal{H})$ of $\mathcal{H} = (V, E)$ is a matrix whose rows and columns are indexed by V and E, respectively, such that $M(v, e) = 1$ if $v \in e$ and $M(v, e) = 0$ if $v \notin e$. Note that the order of the rows and columns is arbitrary. The dual of the hypergraph $\mathcal{H} = (V, E)$ is the hypergraph $\mathcal{H}^* = (E, V)$ where the containment relation $\in^*$ of $\mathcal{H}^*$ is defined as $e \in^* v$ if $v \in e$. Note that $M(\mathcal{H}^*) = M^T(\mathcal{H})$, the transpose of the matrix $M(\mathcal{H})$.

$\mathcal{H}' = (V', E')$ is a subhypergraph of $\mathcal{H} = (V, E)$ if $V' \subset V$ and $E' \subset E$. For a subset of the vertices $X \subset V$ we define the trace of $\mathcal{H}$ on X as $\mathcal{H}[X] = (X, E \cap X)$, that is, $V(\mathcal{H}[X]) = X$ and $E(\mathcal{H}[X]) = \{e \cap X \mid e \in E(\mathcal{H})\}$. The subhypergraph of a trace is called a (sub)configuration or (sub)pattern. Note that the incidence matrix of a subpattern of $\mathcal{H}$ is a submatrix of $M(\mathcal{H})$. A family of hypergraphs $\mathcal{F}$ is hereditary if it is closed for taking subhypergraphs and traces, i.e., for taking subpatterns. This is equivalent to the family of incidence matrices being closed for taking submatrices.

A k-coloring of a hypergraph $\mathcal{H} = (V, E)$ is a map from V to $\{1, \ldots, k\}$. There are several different version of hypergraph colorings of interest.

- A coloring is proper if every edge contains two differently colored vertices. The least k for which a proper k-coloring of $\mathcal{H}$ exists is denoted by $\chi(\mathcal{H})$.
- A coloring is polychromatic if every edge contains a vertex of each k colors; note that such a coloring is only possible if $\mathcal{H}$ is k-heavy. The largest k for which a polychromatic k-coloring of $\mathcal{H}$ exists is denoted by $\chi_{pc}(\mathcal{H})$.
- A coloring is rainbow if all vertices of each edge have different colors; note that such a coloring is possible only if every edge has size at most k. The largest k for which a rainbow k-coloring of $\mathcal{H}$ exists is denoted by $\chi_{rb}(\mathcal{H})$.
- A coloring is strong if it is polychromatic for the k-heavy part of $\mathcal{H}$, while rainbow on the rest. The largest k for which a strong k-coloring of $\mathcal{H}$ exists is denoted by $\chi_s(\mathcal{H})$.
- A coloring is conflict-free if every edge has a vertex whose color differs from the color of all other vertices of that edge. The least k for which a conflict-free k-coloring of $\mathcal{H}$ exists is denoted by $\chi_{cf}(\mathcal{H})$.

- A coloring is unique-maximum if in each edge e the largest[1] color in e occurs exactly once in e. The least k for which a unique-maximum k-coloring of $\mathcal{H}$ exists is denoted by $\chi_{um}(\mathcal{H})$.
- A coloring is odd if for each edge there is a color that occurs in it an odd number of times. The least k for which an odd k-coloring of $\mathcal{H}$ exists is denoted by $\chi_{odd}(\mathcal{H})$.

Note that all of these parameters might be infinite if $\mathcal{H}$ is not finite; in this case we could study the rate as they tend to infinity as a function of the number of the vertices for finite $\mathcal{H}$ from the given family, but this is not the topic of this dissertation.[2] We also have the following relationship among the above chromatic numbers by definition: $\chi(\mathcal{H}) \le \chi_{odd}(\mathcal{H}) \le \chi_{cf}(\mathcal{H}) \le \chi_{um}(\mathcal{H})$ and $\chi_s(\mathcal{H}) \le \chi_{pc}(\mathcal{H}), \chi_{rb}(\mathcal{H})$. Moreover, $\chi(\mathcal{H}) = 2$ if and only if $\chi_{pc}(\mathcal{H}) \ge 2$ if and only if $\chi_s(\mathcal{H}) \ge 2$.

In the rest of the dissertation we will only consider proper and polychromatic colorings; for a survey on conflict-free colorings of geometric hypergraphs, see [73].

We can naturally extend the notions of chromatic numbers to families of hypergraphs by taking the maximum/minimum values over members of the family. This way, $\chi(\mathcal{F}) = \max \chi(\mathcal{H})$ and $\chi_{pc}(\mathcal{F}) = \min \chi_{pc}(\mathcal{H})$ where $\mathcal{H}$ is any finite hypergraph from the family $\mathcal{F}$. Note that finiteness is a technical condition that is needed in most proofs, while it is more convenient to state certain statements if we allow $\mathcal{F}$ to contain infinite hypergraphs as well in the definition. In this thesis we will only study $\mathcal{F}$ that are hereditary, for example, all (finite) subconfigurations of some typically infinite hypergraph; see Section 1.2.

Sometimes a suitable coloring might not exist because of the non-heavy edges of $\mathcal{H}$, as typically for these it is harder to satisfy the requirements of the coloring. Because of this, we denote by $\chi_m(\mathcal{F})$ the least k for which for some m a proper k-coloring exists for every m-heavy $\mathcal{H} \in \mathcal{F}$, i.e., $\chi(\mathcal{H}) \ge k$; if there is no such k, then let $\chi_m(\mathcal{F}) = \infty$. Note that here the m in the subscript of χ_m is *not* a variable, unlike at other places.

We could define a similar parameter for polychromatic colorings, but we instead define $m_k(\mathcal{F})$ to be the least m (if exists) for which every m-heavy $\mathcal{H} \in \mathcal{F}$ is polychromatic k-colorable, i.e., $\chi_{pc}(\mathcal{H}) \ge k$; if there is no such m, then let $m_k(\mathcal{F}) = \infty$. Note that by definition $m_2(\mathcal{F}) < \infty$ if and only if $\chi_m(\mathcal{F}) = 2$, and we always have $m_k(\mathcal{F}) \le m_{k+1}(\mathcal{F})$. Moreover, the following might also be true.

Conjecture 1.1. *If $m_2(\mathcal{F}) < \infty$, then $m_k(\mathcal{F}) < \infty$ for every k for any hereditary $\mathcal{F}$.*

This beautiful conjecture is only known to hold if $m_2(\mathcal{F}) = 2$; this is a classic result of Berge [8], who proved that in this case $m_k(\mathcal{F}) = k$, and also characterized these families. The case $m_2(\mathcal{F}) = 3$ is already wide open.

Conjecture 1.1 first "came up" in 2009 during the writing of a survey [63], until then it was simply believed to be true by the (few) people working on related questions for geometric families. Later it was popularized by the author at several venues, such as Oberwolfach meetings and MathOverflow, and other variants also emerged. The strongest form is the following, conjectured by Keszegh and the author.

[1] Here we use that the coloring maps to numbers.

[2] For some nice questions of geometric flavor about the growth rate, see [18, 20].

Conjecture 1.2. *$m_k(\mathcal{F}) \le (k-1)(m_2(\mathcal{F})-1)+1$ for every k for any hereditary $\mathcal{F}$.*

Compare the striking similarity of the formula to the Calder-Eckhoff conjecture about Radon numbers of abstract convexity spaces; as that conjecture has already been refuted [14], just to be on the safe side, let us also mention three alternative, weaker conjectures:[3]
$m_k(\mathcal{F}) \le C \cdot k \cdot m_2(\mathcal{F})$,
$m_k(\mathcal{F}) \le \mathrm{poly}(k, m_2(\mathcal{F}))$,
$m_k(\mathcal{F}) \le C(m_2(\mathcal{F})) \cdot k$.

There are several (geometric) families for which the inequality of Conjecture 1.2 is sharp; several examples of such families are given later. We also know a polynomial dependence in several special cases, see Theorem 5.2.

1.1.1 Coloring unions of hypergraphs

Claim 1.3. *$\mathcal{H}_1 = (V, E_1)$ and $\mathcal{H}_2 = (V, E_2)$ be two hypergraphs on a common vertex set V. Then $\chi(\mathcal{H}_1 \cup \mathcal{H}_2) \le \chi(\mathcal{H}_1) \cdot \chi(\mathcal{H}_2)$.*

Proof. Take the direct product of the colorings for $\mathcal{H}_1$ and $\mathcal{H}_2$. □

Corollary 1.4. *$\chi_m(\mathcal{H}_1 \cup \mathcal{H}_2) \le \chi_m(\mathcal{H}_1) \cdot \chi_m(\mathcal{H}_2)$.*

In case of proper colorings, these bounds can be sharp, already for graphs (2-uniform hypergraphs). But if we start with a stronger condition, then we can get a better bound.

Lemma 1.5 (Damásdi-Pálvölgyi [24])**.** *Let $\mathcal{H}_1, \dots, \mathcal{H}_{k-1}$ be hypergraphs on a common vertex set V. If $\mathcal{H}_1, \dots, \mathcal{H}_{k-1}$ are each polychromatic k-colorable, then $\chi(\bigcup_{i=1}^{k-1} \mathcal{H}_i) \le k$.*

Proof. Let $c_i : V \to \{1, \dots, k\}$ be a polychromatic k-coloring of $\mathcal{H}_i$. Choose $c(v) \in \{1, \dots, k\}$ such that it differs from each $c_i(v)$. We claim that c is a proper k-coloring of $\bigcup_{i=1}^{k-1} \mathcal{H}_i$. To prove this, it is enough to show that for every edge $H \in \mathcal{H}_i$ and for every color $j \in \{1, \dots, k\}$, there is a $v \in H$ such that $c(v) \ne j$. We can pick $v \in H$ for which $c_i(v) = j$. This finishes the proof. □

Corollary 1.6. *For any families $\mathcal{F}_1, \dots, \mathcal{F}_{k-1}$ if $m_k(\mathcal{F}_1), \dots, m_k(\mathcal{F}_{k-1}) < \infty$, then $\chi_m(\mathcal{F}_1 \cup \dots \cup \mathcal{F}_{k-1}) \le k$.*

Lemma 1.5 is sharp in the sense that for every k there are $k-1$ hypergraphs such that each is polychromatic k-colorable but their union is not properly $(k-1)$-colorable. For example, take a $(k-1)$-dimensional grid of width k, $V = \{(i_1, \dots, i_{k-1}) \mid 1 \le i_j \le k\}$, and let a k-tuple $v^1, \dots, v^k \in V$ be an edge of $\mathcal{H}_j$ if the j-th coordinates are all different, i.e., $\{v_j^1, \dots, v_j^k\} = \{1, \dots, k\}$. A simple induction argument shows that $\chi(\bigcup_{j=1}^{k-1} \mathcal{H}_j) = k$.

[3]The equivalent of the below last weakening of the Calder-Eckhoff conjecture has been recently proved by the author [68], while the equivalent of the first weakening is still open. The equivalent of Conjecture 1.1 would be a very easy statement for Radon numbers.

1.1.2 Connection to ε-nets

Let us also mention that polychromatic colorings are also related to ε-nets. If $m_k(\mathcal{F}) = m$, then the elements of the base set $V(\mathcal{F})$ can be partitioned into k parts such that each part intersects every m-heavy $F \in \mathcal{F}$. Letting $n = |V(\mathcal{F})|$ and setting $\varepsilon = m/n$, each of the k parts forms a strong ε-net for $\mathcal{F}$. By the pigeonhole principle, the size of one of these parts will be at most $n/k = m/(\varepsilon k)$. This implies that if $m_k(\mathcal{F}) = O(k)$, then we have an $O(1/\varepsilon)$ size strong ε-net for $\mathcal{F}$.

1.2 Geometric hypergraphs

A hypergraph $\mathcal{H} = (V, E)$ is called geometric if its structure is derived from some geometric configuration in some space. If $\mathcal{R}$ is a range space, a family of sets in some geometric space, then the hypergraph whose vertices are the points of the underlying space, and whose edges are the sets of $\mathcal{R}$ with the natural containment relation, is called the primal hypergraph induced by $\mathcal{R}$. An example of a range space is the collection of disks in the plane, and its primal hypergraph is $\mathcal{H}(disks) = (\mathbb{R}^2, disks)$ where $v \in e$ if the point corresponding to v is contained in the disk corresponding to e. Similarly, the hypergraph whose vertices are the sets of $\mathcal{R}$, and whose edges are the points of the underlying space with the reverse of the natural containment relation, is called the dual hypergraph induced by $\mathcal{R}$. We will be interested in the hereditary family $\mathcal{F}(disks)$ that consists of the (finite) subconfigurations of $\mathcal{H}(disks)$, or its dual, $\mathcal{F}^*(disks)$. In short, we will call $\mathcal{F}(disks)$ the family of disks, and we will also refer to parameters, such as $\chi_m(\mathcal{F}(disks))$, simply as "$\chi_m$ of disks," when it leads to no confusion. Also, for brevity, instead of parameters of the dual of the family, such as $\chi_m(\mathcal{F}^*(disks))$, we write simply "$\chi_m^*$ of disks." If a hypergraph $\mathcal{H}$ belongs to the family of disks, we will also say that $\mathcal{H}$ is realizable by disks and, similarly, if $\mathcal{H}^*$ belongs to the family of disks, then we say that $\mathcal{H}$ has a dual realization with disks.

For a geometric family $\mathcal{F}$ a statement of the form $\chi_m \leq k$ is equivalent to the following: There is an $m = m(\mathcal{F})$ such that any finite set of points P can be k-colored such that if for some $F \in \mathcal{F}$ we have $|P \cap F| \geq m$, then not all the points in $P \cap F$ have the same color. Similarly, $m_k \leq m$ is equivalent to the following: Any finite set of points P can be k-colored such that if for some $F \in \mathcal{F}$ we have $|P \cap F| \geq m$, then all k colors occur among the points in $P \cap F$. The statement $m_k^* \leq m$ about the dual range space is equivalent to the following: Any finite collection of sets $F_1, \ldots, F_n \in \mathcal{F}$ can be k-colored such that if for some point we have $|\{F_i \mid p \in F_i\}| \geq m$, then all k colors occur among the sets $\{F_i \mid p \in F_i\}$. This latter can also be rephrased as any finite m-fold covering of any subset of the underlying space is decomposable into k disjoint coverings.

The simplest example to consider is the family of intervals in $\mathbb{R}$. A result, sometimes attributed to Tibor Gallai, is that for intervals $m_k = k$ and $m_k^* = k$; for finite hypergraphs both of these follow from a simple induction. It is not hard to generalize this result to the family whose members are the sets formed by the union of t intervals. In this case $m_k = t(k-1) + 1$, which means that these families satisfy the inequality of Conjecture 1.2 with equality for all $m_k \geq 2$. The dual case is entirely different as $m_2^* = \infty$ for any $t \geq 2$.

Another classic example is the family of lines in $\mathbb{R}^2$. It was observed in [65] that a generic projection from a sufficiently high dimensional grid, $\{1, \ldots, m\}^d$, to the plane shows that $\chi_m = \infty$ using the Hales-Jewett theorem and, by point-line duality, also $\chi_m^* = \infty$.

As we mentioned earlier, in this dissertation we want to avoid issues with infinite hypergraphs, that is why χ_m and m_k have been defined for finite hypergraphs from the family; for some of the issues with infinite hypergraphs, see [67], while for some positive results, see [54]. Because of this, we typically do not need to specify whether the underlying sets are open or closed, as it usually follows by a perturbation argument that the finite hypergraphs of the respective families are the same.

1.2.1 Relationships among geometric hypergraphs

As just mentioned, the hypergraphs realizable by open disks, for example, are the same as the hypergraphs realizable by closed disks, because in any finite realization after an appropriate perturbation the incidences remain the same, but no points will fall on the boundary of the disks. In this section, we will describe some further equivalences and containments among geometric families. For a large poset describing many of these containments, see the webpage `https://coge.elte.hu/cogezoo.html`, designed by Keszegh, and maintained by him and the author.

Another important equivalence [61, 63] is that if a family is the collection of some (or all) translates of one set in $\mathbb{R}^d$, then the primal and dual hypergraphs induced by the range spaces are the same. Indeed, consider a family $\mathcal{C} = \{C_i \mid i \in I\}$ of translates of a set $C \subset \mathbb{R}^d$ and a set of points $P \subseteq \mathbb{R}^d$. Suppose, without loss of generality, that C contains the origin 0. For every $i \in I$, let c_i denote the point of C_i that corresponds to $0 \in C$. In other words, we have $\mathcal{C} = \{C + c_i \mid i \in I\}$. Assign to each $p \in P$ a translate of $-C$, the reflection of C about the origin, by setting $C_p^* = -C + p$. Observe that

$$p \in C_i \iff c_i \in C_p^*$$

which proves that the same hypergraphs are realizable by primal and dual range spaces. In the early papers more focus was put on studying dual range spaces, but mainly range spaces of translates were studied, when the two problems are anyhow equivalent. A set was defined as cover-decomposable [61] if any sufficiently thick covering of the *whole* plane by the translates of the given set can be decomposed into two disjoint coverings. For a bounded open set S it follows from a compactness argument that cover-decomposability follows from the statement that for (the family defined by the range space of) the translates of S we have $\chi_m = 2$. For more about the connection and results, see [63].

Finally, some so-called *dynamic* versions of hypergraph families can be shown to be equivalent to other (normal) hypergraph families. In a dynamic hypergraph the vertices are ordered as $v_1, \ldots, v_n$, and each prefix $v_1, \ldots, v_i$ induces a certain hypergraph from a given family; for the precise statement, see Definition 4.15. A geometric example is the hypergraph induced by an ordered set of points, $p_1, \ldots, p_n \in \mathbb{R}$, where the range space is formed by all intervals. The edges of this hypergraph are the sets of the form $I \cap \{p_1, \ldots, p_i\}$ where I is an interval and $1 \le i \le n$. A good way to visualize this problem is that the points "appear" in the given order on the number line. This dynamic interval hypergraph family is in fact the same as the family of hypergraphs realizable by so-called bottomless rectangles. A subset of $\mathbb{R}^2$ is called a bottomless rectangle if it is of the form $\{(x, y) : \ell \le x \le r, y \le t\}$. This equivalence was used in [7]] to prove Theorem 2.20, and later a similar relationship was used between dynamic quadrants and octants [45] to prove Theorem 2.24; for the details, see Chapter 3. Several further results about dynamic versions can be found in Section 4.5.

An easy containment relation among hypergraph families is that hypergraphs realizable by halfspaces (in any dimension) are a subfamily of the hypergraphs realizable by unit balls, which are a subfamily of the hypergraphs realizable by balls. This latter containment is a special case of the fact that hypergraphs realizable by the translates of a set are always a subfamily of the hypergraphs realizable by the homothetic copies[4] of the same set.

Let us mention one more containment that proved quite useful for obtaining interesting results. Note that if a hypergraph is realizable by axis-parallel bottomless rectangles, then it is also realizable by the homothetic copies of a fixed triangle, as the sides of the bottomless rectangles can be slightly bent to meet. And any hypergraph $\mathcal{H}$ realizable by the homothets of a fixed triangle is also realizable by octants in $\mathbb{R}^3$, where by octant we mean a subset of $\mathbb{R}^3$ of the form $(-\infty, x_0) \times (-\infty, y_0) \times (-\infty, z_0)$. This follows from embedding the plane realizing $\mathcal{H}$ with triangles into $\mathbb{R}^3$ as the $x + y + z = 0$ plane (see Figure 1).

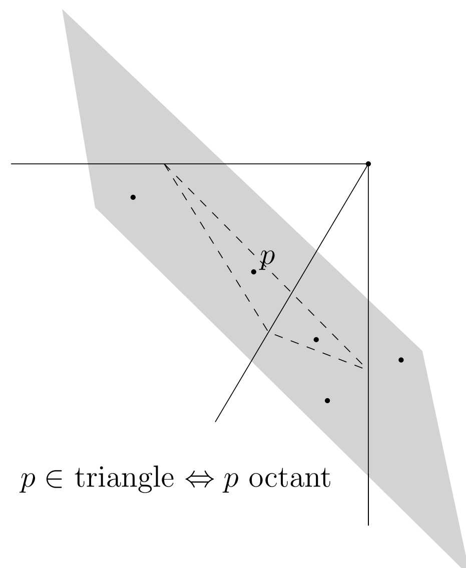


Figure 1: Octants give a richer family than homothetic copies of a triangle, because every homothet of the triangle depicted on the shaded plane can be obtained as the intersection of an octant with the plane.

1.2.2 Generalized Delaunay triangulation

To obtain the generalized Delaunay triangulation with respect to some bounded convex body C of a point set P, define a graph $\mathcal{DT} = \mathcal{DT}(P, C)$ whose vertex set is P, and two points of P are connected by an edge if they are covered by a homothet of C not containing any other point of P. It is well-known (see e.g., [51]) that this graph is connected and planar for any P and C, such that the vertices are mapped to the respective points of P and the edges are segments. Each inner face of this embedding is covered by a homothet of C not containing any points in its interior, and all inner faces are triangles if the points are in general position with respect to C,

[4] A homothetic copy, or homothet, of a set is a scaled and translated copy of it (rotations are *not* allowed). We also require the scaling factor to be positive—some other papers call this a *positive* homothet.

meaning that no four points fall on the boundary of a homothet of C. If this is not the case, then we triangulate the remaining inner faces arbitrarily to obtain the graph $\mathcal{DT}$. Since because of this last step $\mathcal{DT}$ might not be unique, we should say one of the Delaunay triangulations but to avoid complications we always fix one of the Delaunay triangulations and denote that by $\mathcal{DT}$.

We recall a few simple statements about Delaunay triangulations, which also appeared in [2].

Proposition 1.7. *If C' is a homothet of C, the points $C' \cap P$ induce a connected subgraph of $\mathcal{DT}(P, C)$.*

Corollary 1.8 ([2])**.** *If C' is a homothet of C and e is an edge of $\mathcal{DT}$ that crosses C' and splits it into two parts, then one of these parts does not contain any point from P.*

We will also use the following related claim.

Claim 1.9. *The intersection of a convex polygon with the boundary of its homothetic copy is always connected.*

2 Survey of results

This chapter is a survey of the most interesting results about the parameters χ_m and m_k of geometric families. It can be considered as an updated version of the survey [63], joint work with Pach and Tóth. In fact, we rely very little on [63], as it became outdated practically already by the time it appeared, as many breakthrough results came soon after. Some of the proofs of the results mentioned in this chapter can be found in later chapters. For a summary of some important results, see the Summary Table on the last page, and for a more complete, always up-to-date version of the known results, see the website `https://coge.elte.hu/cogezoo.html`.

Let $\mathcal{C}$ be a family of sets in $\mathbb{R}^d$, and let $P \subseteq \mathbb{R}^d$. We say that $\mathcal{C}$ is an *m-fold covering of* P if every point of P belongs to at least m members of $\mathcal{C}$. A 1-fold covering is simply called a *covering*. Clearly, the union of m coverings is an m-fold covering. We will be mostly interested in the case when P is a finite set of points or the whole space $\mathbb{R}^d$.

Sphere packings and coverings have been studied for centuries, partially because of their applications in crystallography, Diophantine approximation, number theory, and elsewhere. The research in this field has been dominated by density questions of the following type: What is the most "economical" (i.e., least dense) m-fold covering of space by unit balls or by translates of a fixed convex body? It is suggested by many classical results and physical observations that, at least in low-dimensional spaces, the optimal arrangements are typically periodic, and they can be split into several lattice-like coverings [31, 32]. Does a similar phenomenon hold for all sufficiently "thick" multiple coverings, without any assumption on their densities?

About 15 years ago, a similar problem was raised for *large scale ad hoc sensor networks*; see Feige et al. [30] and Buchsbaum et al. [13]. In the—by now rather extensive—literature, it is usually referred to as the *sensor cover problem*. In its simplest version it can be phrased as follows. Suppose that a large region P is monitored by a set of sensors, each having a circular range of unit radius and each powered by a battery of unit lifetime. Suppose that every point of P is within the range of at least m sensors, that is, the family of ranges of the sensors, $\mathcal{C}$, forms an m-fold covering of P. If $\mathcal{C}$ can be split into k coverings $\mathcal{C}_1, \dots, \mathcal{C}_k$, then the region can be monitored by the sensors for at least k units of time. Indeed, at time i, we can switch on all sensors whose ranges belong to $\mathcal{C}_i$ $(1 \le i \le k)$. We want to maximize k, in order to guarantee the longest possible service. Of course, the first question is the following, raised by Pach in 1980.

Problem 2.1 (Pach [60]). *Is it true that every m-fold covering of the plane with unit disks splits into two coverings, provided that m is sufficiently large?*

In a long unpublished manuscript, Mani and Pach [56] claimed that the answer to this question was in the affirmative with $m \le 33$. Pach [66] warned that this "has never been independently verified." Winkler [77] even conjectured that the statement is true with $m = 4$. For more than 30 years, the prevailing conjecture has been that for any open plane *convex body* (i.e., bounded convex set) C, there exists a positive integer $m = m(C)$ such that every m-fold covering of the plane with translates of C splits into two coverings. This conjecture was proved in [61] for centrally symmetric convex polygons C. It took almost 25 years to generalize this statement to all convex polygons [75, 69, 37]. Thus, this question has propelled research in the area for decades. Eventually, it was surprisingly answered in the negative by the author.[5] Later, the counterexample has been extended to other smooth bodies as well in a paper with Pach [62], see Section 7.

In the remaining part of this chapter, we will look at coloring results about hypergraphs of range spaces of natural geometric families, mostly focusing on primal range spaces and finite hypergraphs.

[5] D. Pálvölgyi, Indecomposable coverings with unit discs, preprint https://arxiv.org/abs/1310.6900v1, 2013.

2.1 Translates of polygons

As mentioned, in a series of papers, it was proved that all open convex polygons are cover-decomposable. These proofs relied on the following results about finite hypergraphs.

Theorem 2.2 (Pach [61]). *For the translates of any centrally symmetric convex polygon* $\chi_m = 2$.

Theorem 2.3 (Tardos-Tóth [75]). *For the translates of any triangle* $\chi_m = 2$.

Theorem 2.4 (Pálvölgyi-Tóth [69]). *For the translates of any convex polygon* $\chi_m = 2$.

Note that the above statement cannot be extended to all polygons.

Theorem 2.5 (Pach-Tardos-Tóth [65]). *For the translates of any concave quadrilateral* $\chi_m > 2$.

While an upper bound for χ_m has not been studied in detail, it can shown that in fact $\chi_m = 3$.[6] Triangulate Q to obtain two triangles, T_1 and T_2. Using Theorem 2.3, for the translates of each triangle we have $\chi_m = 2$, that is, there is an m such that any finite m-heavy hypergraph whose edges are translates of T_i can be 2-colored for $i = 1, 2$. Finally, by applying Lemma 1.5 to these two hypergraphs, we get a proper 3-coloring for translates of Q that contain at least $2m-1$ points, thus $\chi_m = 3$. A similar argument shows that $\chi_m \leq n-1$ for the translates of any n-gon.

A larger class of concave polygons for whose translates $\chi_m > 2$ was given by the author [67], where it turned out that the pairs of angles that can be found in the polygon play an important role. In fact, based on these results, the translates of polygons have been completely classified with respect to $\chi_m = 2$.

Theorem 2.6 (Pálvölgyi-Tóth [69]). *For the translates of a polygon* $\chi_m = 2$ *if and only if any pair of its convex angles*[7] *are such that either one angle contains the other one, or the same pair could occur in a convex polygon.*

For families with $\chi_m = 2$, the growth rate of the function m_k has been also extensively studied. The only case when m_2 was also studied explicitly is the translates of triangles, for which $5 \leq m_2 \leq 9$ as a corollary of Theorems 3.2 and 2.24. First, in [61] it was shown that for any centrally symmetric convex polygon P the parameter m_k exists and is bounded by an exponentially fast growing function of k. In [75] a similar result was established for triangles, and in [69] for convex polygons. However, all these results were improved to the optimal linear bound in a series of papers.

Theorem 2.7 (Pach-Tóth [66]). *For the translates of any centrally symmetric convex polygon* $m_k = O(k^2)$.

Theorem 2.8 (Aloupis et al. [5]). *For the translates of any centrally symmetric convex polygon* $m_k = O(k)$.

[6] Earlier a similar argument of István Kovács (personal communication) gave the weaker bound $\chi_m \leq 4$.

[7] An angle is called a convex angle if it is smaller than π.

Theorem 2.9 (Gibson-Varadarajan [37]). *For the translates of any convex polygon* $m_k = O(k)$.

Although in [37] only convex polygons have been studied, their proof also works for any polygon such that any pair of its convex angles are such that either one angle contains the other one, or the same pair could occur in a convex polygon.

The situation, however, is completely different in higher dimensions.

Theorem 2.10 (Pálvölgyi [67]). *For the translates of any polyhedron* $\chi_m > 2$.

The proof is based on the observation that for any polyhedron P, either there is a plane that intersects P in a concave polygon for which $\chi_m > 2$ because of Theorem 2.6, or there are two parallel planes that intersect P in two polygons such that taking one convex angle from each they could not form a convex polygon. In both cases, we can take a plane in space and a family of translates of P that realize the planar construction from [67] in this plane with the translates of P.

2.2 Homothets of polygons

Recall that a homothetic copy, or homothet, of a set is a scaled and translated copy of it (rotations are *not* allowed). Colorings of geometric range spaces induced by homothetic copies of polygons have been studied much less than translates of polygons. Note that it follows from the planarity of the generalized Delaunay graph (Section 1.2.2) that for homothets $\chi_m \le 4$.

The first paper in which specifically homothets were studied proved the following.

Theorem 2.11 (Cardinal et al. [16])**.** *For the homothets of a triangle* $m_k \le 144k^8$*.*

Their method was developed further by Keszegh and the author.

Theorem 2.12 (Keszegh-Pálvölgyi [47])**.** *For the homothets of a triangle the inequality* $m_k \le m_2 \cdot k^{\log_2(2m_2-1)}$ *holds.*

Since as a corollary of Theorem 3.1 the best bound for m_2 of the homothets of a triangle is currently 9, this gives $m_k \le 9 \cdot k^{4.09}$. Moreover, in [47] a general method was developed which implies that if $m_2 < \infty$ for the homothets of any convex polygon, then m_k grows polynomially; see Corollary 5.4. This method can be found in full detail in Chapter 5. However, apart from triangles, the square (and its affine transformations, i.e., parallelograms) is the only convex polygon for whose homothets $m_2 < \infty$ has been proved.

Theorem 2.13 (Ackerman-Keszegh-Vizer [2])**.** *For the homothets of a square* $m_2 \le 215$*.*

Combining this with Theorems 5.2 and 5.6, we get the following bound.

Corollary 2.14 (Ackerman-Keszegh-Vizer [2])**.** *For the homothets of a square* $m_k = O(k^{8.75})$*.*

The situation, surprisingly, is completely different for the dual range spaces of homothets. For the homothets of a triangle we have $m_k^* = O(k^{5.09})$ from Corollary 2.26 because if a hypergraph is (dual) realizable by the homothets of a triangle, then it is also realizable by octants. However, for the homothets of other polygons Kovács showed the following, building on the construction from [67].

Theorem 2.15 (Kovács [53])**.** *For the homothets of any non-triangle polygon* $\chi_m^* > 2$*.*

This means that for the homothets of squares $\chi_m = 2$ but $\chi_m^* > 2$. Just like for translates, the finiteness of χ_m for the homothets of any polygon follows from $\chi_m = 2$ for the homothets of triangles. A similar bound can be also proved for χ_m^*, which we sketch here for the homothets of a quadrangle Q. Suppose that we have a finite collection $\mathcal{Q}$ of homothets of Q. Divide Q to two triangles, T_1 and T_2. Partition the points of the plane covered by at least $2m-1$ members of $\mathcal{Q}$ into P_1 and P_2 such that every point $p_i \in P_i$ is covered by at least m homothets of T_i, formed by the respective triangles of the quadrangles from $\mathcal{Q}$. From Corollary 2.26, if m is large enough, we get a polychromatic 3-coloring of the members of $\mathcal{Q}$ for the m-heavy edges of the dual hypergraph induced by the points of P_1, and another polychromatic 3-coloring of the members of $\mathcal{Q}$ for the m-heavy edges of the dual hypergraph induced by the points of P_2. Finally, Lemma 1.5 implies $\chi_m^* \le 3$.

The bound on χ_m was improved for *convex* polygons as follows.

Theorem 2.16 (Keszegh-Pálvölgyi [50])**.** *For the homothets of a convex polygon* $\chi_m \leq 3$.

The proof of this result can be found in Chapter 6.

We would like to remark that known constructions do not exclude the possibility that for convex polygons $\chi_m = 2$ might also hold; this statement has been already mentioned to hold for triangles and squares, and it is also known in the special case when all homothets are stabbed by a common point, for example, if they all contain the origin [22].

Keszegh and the author conjectured that Theorem 2.16 can be extended to every plane convex set C, i.e., that there is an $m = m(C)$ such that any finite set of points admits a 3-coloring such that any homothetic copy of C that contains at least m points contains two points with different colors. (The special case when C is a disk has been posed earlier by Keszegh [42].) However, these conjectures have been disproved in [22]; see Section 7.10.

2.3 Axis-parallel boundaries

Besides polygons, other natural families are ones defined by axis-parallel boundaries. After intervals, the simplest range space is the family of (positive) quadrants, i.e., sets of the form $[x_0,\infty)\times[y_0,\infty)$, for which several easy arguments give the following claim.

Claim 2.17. *For positive quadrants $m_k = m_k^* = k$.*

If instead of positive quadrants, we allow all four quadrants, then $m_k = O(k)$ is a simple corollary of Theorem 2.9 for squares, while $m_k^* \le 4k-3$ follows from the previous claim by applying it to the four families separately.

Another simple shape family is axis-parallel strips, i.e., the sets of the form $[x_1,x_2]\times\mathbb{R}$ and $\mathbb{R}\times[y_1,y_2]$. If only horizontal or only vertical strips are considered, then the obtained family is isomorphic to the family of intervals in $\mathbb{R}$. If both are allowed, we have the following results.

Theorem 2.18 (Aloupis et al. [6])**.** *For axis-parallel strips $1.5k-1 \le m_k \le 2k-1$ and $m_k^* \le 2k-1$.*

In the same paper some generalizations for higher dimensions are also proved. Define a slab as the section between two parallel hyperplanes in $\mathbb{R}^d$.

Theorem 2.19 (Aloupis et al. [6])**.** *For d-dimensional axis-parallel slabs $2\lceil(2d-1)k/2d\rceil \le m_k \le k(4\ln k+\ln d)$ and $\lfloor k/2\rfloor d+1 \le m_k^* \le d(k-1)+1$.*

A lot of work was put into determining these parameters for *bottomless rectangles*. Recall that a subset of $\mathbb{R}^2$ is called a (closed) bottomless rectangle if it is of the form $\{(x,y) : \ell \le x \le r, y \le t\}$. The range space of bottomless rectangles were first studied by Keszegh [42], who showed $m_2 = 4$ and $m_2^* = 3$. Later the following general upper bound was given for m_k.

Theorem 2.20 (Asinowski et al. [7])**.** *For bottomless rectangles $1.67k-2.5 \le m_k \le 3k-2$.*

Therefore, bottomless rectangles are known to satisfy Conjecture 1.2 but we do not know whether the inequality is tight or not. Much less is known about their dual range space. The best bound, $m_k^* = O(k^{5.09})$, follows from the respective result for octants, Corollary 2.26; for better bounds in some special cases and for the reason why known methods fail, see [15]. Also, it was recently proved that for the union of bottomless rectangles and horizontal strips $\chi_m > 2$ [19], while from Theorems 2.18, 2.20 and Lemma 1.5 we get that $\chi_m \le 3$, thus $\chi_m = 3$.

For the family of axis-parallel rectangles in the plane, we only have negative results.

Theorem 2.21 (Chen-Pach-Szegedy-Tardos [21])**.** *For axis-parallel rectangles $\chi_m = \infty$.*

Theorem 2.22 (Pach-Tardos [64])**.** *For axis-parallel rectangles $\chi_m^* = \infty$.*

Cardinal noticed that orthants of $\mathbb{R}^d$ for $d \ge 4$ can simulate the axis-parallel rectangles of an appropriate subplane of $\mathbb{R}^4$. To state this precisely for $d = 4$, define a (positive) *sedecimant* in $\mathbb{R}^4$ as the set of points $\{(x,y,z,w) \mid x \ge x_0, y \ge y_0, z \ge z_0, w \ge w_0\}$.

Corollary 2.23 (Cardinal[8]). *For positive sedecimants $\chi_m = \infty$.*

Proof. From Theorem 2.21, for any k and m there is a finite planar point set S such that for every k-coloring of S there is an axis-parallel rectangle that contains exactly m points of S, all of the same color. Place this construction on the $\Pi = \{(x,y,z,w) \mid x+y = 0,\ z+w = 0\}$ subplane of $\mathbb{R}^4$. A sedecimant $\{(x,y,z,w) \mid x \geq x_0, y \geq y_0, z \geq z_0, w \geq w_0\}$ intersects Π in $\{(x,y,z,w) \mid x_0 \leq x = -y \leq -y_0, z_0 \leq z = -w \leq -w_0\}$, which is a rectangle whose sides are parallel to the lines $\{x+y = 0, z = w = 0\}$ and $\{x = y = 0, z+w = 0\}$, respectively. Taking these perpendicular lines as axes, thus any "axis-parallel" rectangle of Π is realizable by an appropriate sedecimant, and the theorem follows. □

Kolja Knauer (personal communication) observed that all axis-parallel rectangles of a subplane of $\mathbb{R}^3$ can be cut out in a similar way by the homothets of a (regular) tetrahedron. Indeed, let Δ be the tetrahedron whose vertices are $(1,0,1), (-1,0,1), (0,-1,-1)$, and $(0,1,-1)$. Let Δ_h be a translate of Δ by $-1 < h < 1$ parallel to the z-axis. The intersection of Δ_h with the plane $\Pi = \{(x,y,z) \mid z = 0\}$ yields an axis-parallel rectangle $R_h = \Delta_h \cap \Pi$. The ratio of the sides of R_h depends on h, and can take any value, as it tends to $\pm\infty$ as $h \to \pm 1$. It follows that by taking a homothetic scaling of D_h, we can obtain any axis-parallel rectangle. Just like in the proof of Corollary 2.23, we obtain by [21] that for any c and m there is a finite point set $S \subset \mathbb{R}^3$ such that for every c-coloring of S there is a homothet of Δ that contains exactly m points of S, all of the same color. Therefore, $\chi_m = \infty$ for homothets of simplices.

Between quadrants and sedecimants, only the case of octants had been open, until the following result.

Theorem 2.24 (Keszegh-Pálvölgyi [45, 49]). *For positive octants $5 \leq m_2 \leq 9$.*

A detailed proof of this result is given in Chapter 3. Let us also mention that it was proved in [9] that 6-heavy hypergraphs realizable by octants are always proper 3-colorable.

For m_k, no upper bound followed from the method of [45]. Later only the very weak bound of $m_k \leq 12^{2^k}$ was proved [46]. Then a general method was developed in a series of papers by Cardinal et al. [16, 17] and by Keszegh and the author [47] which culminated in the following polynomial bound.

Theorem 2.25 (Cardinal et al. [17]). *For positive octants $m_k \leq m_2 \cdot k^{\log_2(2m_2-1)+1}$.*

The combination of Theorems 2.24 and 2.25 implies the following bound.

Corollary 2.26. *For positive octants $m_k = O(k^{5.09})$.*

Finally, we mention that if instead of positive octants, we allow all 8 octants, then it follows from the methods of [67] that $\chi_m > 2$, while the boundedness of χ_m follows from Lemma 1.5, or from simply taking the direct products of the colorings for different octants.

[8] Cardinal (personal communication) stated $\chi_m > 2$ using the same reduction based on [65] about axis-parallel rectangles; Corollary 2.23 is only more general because we use a stronger result, Theorem 2.21, about axis-parallel rectangles.

2.4 Disks

The first important result about shapes whose boundary does not consist of fixed direction segments was the following.

Theorem 2.27 (Smorodinsky-Yuditsky [74]). *For halfplanes* $m_k = 2k-1$ *and* $2k-1 \le m_k^* \le 3k-2$.

Later it was shown by Fulek [34] that for the dual range space of halfplanes the lower bound is sharp for $k = 2$, i.e., $m_2^* = 3$. Keszegh and the author managed to generalize Theorem 2.27 to pseudohalfplane arrangements as well, which can be defined as follows. A *pseudoline arrangement* is a collection of simple curves, each of which splits $\mathbb{R}^2$ into two unbounded parts, such that any two curves intersect at most once. A *pseudohalfplane* is the region on one side of a pseudoline in such an arrangement.

Theorem 2.28 (Keszegh-Pálvölgyi [48]). *For pseudohalfplanes* $m_k = 2k-1$ *and* $2k-1 \le m_k^* \le 3k-2$.

Note that hypergraphs realizable by halfplanes are also realizable by unit discs, but this later family turned out to be much richer: The author proved $\chi_m > 2$ for them, answering Problem 2.1 in the negative; later the counterexample has been extended to other smooth bodies as well in a paper with Pach.

Theorem 2.29 (Pach-Pálvölgyi [62]). *Let C be any convex body in the plane that has two parallel supporting lines such that C is strictly convex in some neighborhood of the two points of tangencies.*[9] *Then for any positive integer m, there exists a 3-chromatic m-uniform hypergraph that is realizable with translates of C, therefore, $\chi_m > 2$.*

Recall that if $\mathcal{F}$ consists of the translates or homothets of some planar convex body, then it follows from the planarity of generalized Delaunay-triangulations that $\chi_m \le 4$. This left only the following question open for translates: Is there for any planar convex body C a positive integer m such that no 4-chromatic m-uniform hypergraph is realizable with the translates of C? This was answered recently by Damásdi and the author.

Theorem 2.30 (Damásdi-Pálvölgyi [24]). *For any planar convex body C $\chi_m \le 3$, i.e., there is a positive integer m such that any finite point set P in the plane can be three-colored in a way that there is no translate of C containing at least m points of P, all of the same color.*

This has been hitherto known to hold only when C is a polygon (in which case 2 colors suffice according to Theorem 2.6, and 3 colors are known to be enough even for homothets according to Theorem 2.16) and for pseudodisk families that intersect in a common point [1] (which generalizes the case when C is unbounded, in which case 2 colors suffice [62]).

Note that an earlier version of the proof [23] only worked if C is a disk, and while the generalization to other convex bodies with a smooth boundary seemed feasible,

[9] This condition can be relaxed to require only one smooth neighborhood on the boundary of C. This was discovered later by Damásdi and the author [24], see Figure 30.

there was no direct way to extend it to arbitrary convex bodies. The proof of Theorem 2.30 relies on a surprising connection to two other famous results, the solution of the two dimensional case of the Illumination conjecture (Levi [55]), and a recent solution of the Erdős-Sands-Sauer-Woodrow conjecture (Bousquet, Lochet, and Thomassé [12]). In fact, a generalization of the latter result is used, proved also in [24].

The main part of the proof of Theorem 2.30 can be found in Section 7.11.

About homothetic copies, $\chi_m \leq 4$ turned out to be sharp for most bodies, including the disk, disproving earlier conjectures [42, 50], and improving [65], which established $\chi_m > 2$ for disks.

Theorem 2.31 (Damásdi-Pálvölgyi [22])**.** *Let C be any convex body in the plane that has two parallel supporting lines such that C is strictly convex in some neighborhood of the two points of tangencies. For any positive integer m, there exists a 4-chromatic m-uniform hypergraph that is realizable with homothets of C, therefore, $\chi_m = 4$.*

The proof of Theorem 2.31 can be found in Section 7.10.

Balls and halfspaces in higher dimensions have not been studied yet in detail. We are aware of only two unpublished observations.

Damásdi noted that for halfspaces in $\mathbb{R}^3$ we have $4 \leq \chi_m \leq \chi \leq 5$. The lower bound follows from Theorem 2.31 by projecting up the point set P to a paraboloid, while the upper bound follows from coloring the points inside the convex hull with one color, and the points of the convex hull with four more, using the planarity of the Delaunay triangulation on the surface.

The author noted that halfspaces in $\mathbb{R}^5$ can cut out any conic section (given by $Ax^2 + Bxy + Cy^2 + Dx + Ey + F \leq 0$) from the surface $x_3 = x_1^2, x_4 = x_1x_2, x_5 = x_2^2$, and thus the Hales-Jewett hypergraph can be realized, implying $\chi_m = \infty$. There are other constructions as well that are realizable in a high enough dimension, but we do not know anything else about χ_m in $\mathbb{R}^3$ or in $\mathbb{R}^4$ for halfspaces and (unit) balls.

2.5 Further results and open problems

The following notion came up while studying geometric hypergraphs (for the connection, see Section 7.7 and [48]).

Definition 2.32. *For $A \subset [n] = \{1, 2, \dots, n\}$, denote by a_i the i^{th} smallest element of A. For two equal sized sets, $A, B \subset [n]$, we write $A \leq B$ if $a_i \leq b_i$ for every i.*

An m-uniform hypergraph on the vertex set $[n]$ is called a shift-chain *if its hyperedges are totally ordered by the relation $\leq$.*

Problem 2.33. *Does there exist an integer m such that every m-uniform shift-chain is 2-colorable?*

Fulek [33] managed to find by an exhaustive computer search a 3-uniform shift-chain that is not 2-colorable (see Figure 2), but for larger m the question is wide open, despite having received significant attention [11].

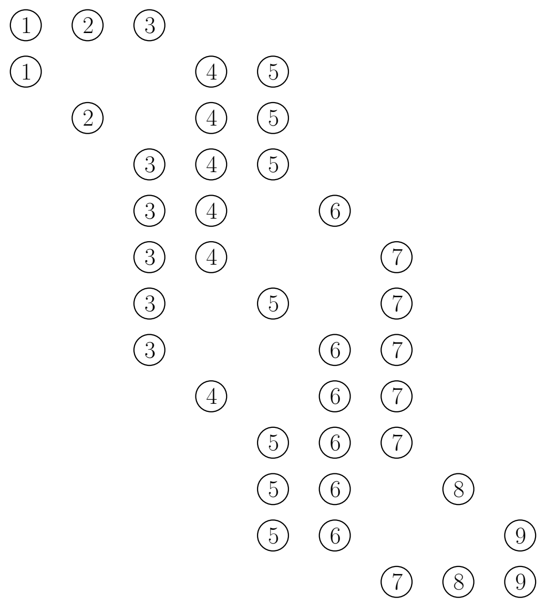


Figure 2: A shift-chain of 13 triples, each of which corresponds to a row. For any 2-coloring of the 9 vertices, one of the triples is monochromatic.

The problem to study the respective colorings of arithmetic progressions was raised recently by the author.[10] For a set $D \subset \mathbb{N}$, denote by $\mathcal{A}_D$ the family of arithmetic progressions of $\mathbb{N}$ with difference $d \in D$ (of any length). For any $S \subset \mathbb{N}$, these progressions can be interpreted as the hyperedges of a hypergraph over S, where the edges are formed by the intersections of the arithmetic progressions from $\mathcal{A}_D$ with S.

Problem 2.34. *What can we say about the χ_m and m_k parameters of the above hypergraph families for various D?*

[10]See `https://coge.elte.hu/kutverseny21.pdf` (in Hungarian).

Note that if $D = \mathbb{N}$, then van der Waerden's theorem implies that $\chi_m = \infty$. On the other hand it is quite easy to establish that $\chi_m = 2$ and $m_k = O(k)$ if D is finite, and also that the value of χ_m depends only on the divisibility lattice of D. The author proposed Problem 2.34 as it is a natural question in combinatorial number theory, while it also connects to geomeric families: If $D = \{2^i \mid i \in \mathbb{N}\}$, then this hypergraph contains all finite hypergraphs realizable by bottomless rectangles. A further interesting subfamily, $\mathcal{A}_D^\infty \subset \mathcal{A}_D$, consists of the arithmetic progressions with an infinite number of elements. Some partial results have been obtained about the χ_m parameter of these families for some D by several students working on these questions.

Finally, without defining the exact problem and results, let us also mention that polychromatic colorings of the edges of graphs have also been studied in this context, see Bollobás et al. [10].

3 Octants

This chapter is based on the article [49], using ideas from the articles [45, 46], all joint work with Keszegh.

By an *octant*, in this chapter we mean an open subset of $\mathbb{R}^3$ of the form $(-\infty, x_0) \times (-\infty, y_0) \times (-\infty, z_0)$ and the point (x_0, y_0, z_0) is called the *apex* of the octant. Note that this is a negative octant, just the opposite of a positive octant.

To establish the upper bound of Theorem 2.24, we need to prove the following statement.

Theorem 3.1. *Any finite set of points in* $\mathbb{R}^3$ *can be colored with two colors such that any octant containing at least* 9 *points contains both colors, thus* $m_2 \leq 9$.

We also give the following construction, which will prove the lower bound of Theorem 2.24.

Theorem 3.2. *For every triangle* T *there is a finite point set* P *such that for every two-coloring of* P *there is a translate of* T *that contains exactly* 4 *points and all of these have the same color.*

This also implies $m_2 \geq 5$ for octants, as they give a richer family (recall Figure 1).

Now we present the proofs of the above two theorems.

3.1 Proof of Theorem 3.1

We will prove the following dynamic planar version of the problem. In this chapter, by a (negative) *quadrant* or *wedge* we mean a subset of the plane of the form $(-\infty, x) \times (-\infty, y)$. We have to two-color a finite ordered planar point set $\{p_1, p_2, \dots, p_n\}$ such that for every t every quadrant that contains at least 9 points from $P_t = \{p_1, \dots, p_t\}$ contains both colors. The equivalence of this dynamic planar version and the statement of Theorem 3.1 is implied by the following containment-reversing bijections: an octant with apex (x, y, z) is mapped to the point (x, y) that "appears" at time z, while a point with coordinates (a, b) is mapped to a mirrored quadrant with apex (a, b), i.e., to the subset $(a, \infty) \times (b, \infty)$; for a more detailed discussion, see Section 1.2.1.

A way to imagine this problem is that the points "appear" in order and at step t we have to color the new point, p_t. This is impossible to do in an online setting [44], as we will see in Chapter 4, i.e., without knowing in advance which points will come in which order. Moreover, it was shown by Cardinal et al. [17] that such a coloring is even impossible in a so-called *semi-online* model, where points can be colored at any time after their arrival as long as every octant with 9 (or any other constant number of) points contains both colors. Our strategy builds a forest on the points such that any time any quadrant containing at least 9 points contains two adjacent points from the same tree-component. Therefore, after all the points arrived, any proper two-coloring of the forest will be such that any octant containing at least 9 points contains both colors. We start by introducing some notation (see also Figure 3).

Definition 3.3. *We say that point* $p = (p_x, p_y)$ *is* northwest *(in short NW) from point* $q = (q_x, q_y)$ *if and only if* $p_x < q_x$ *and* $p_y > q_y$. *In this case we also say that* q *is* southeast *(in short SE) from* p *and that* p *and* q *are* incomparable.

(a) p is NW from q, q is SE from p, p and q are incomparable

(b) p is SW from q, q is NE from p, p and q are comparable

Figure 3: Definition 3.3: the possible relations of two points.

Similarly, we say that point $p = (p_x, p_y)$ is southwest *(in short SW) from point $q = (q_x, q_y)$ if and only if $p_x < q_x$ and $p_y < q_y$. In this case we also say that q is* northeast *(in short NE) from p and that p and q are* comparable*.*

We can suppose that all points have different coordinates, as by a slight perturbation we can only get more subsets of the points contained in a quadrant (without losing others).

At any step t, we define a graph G_t (which is actually a forest) on the points of P_t and a vertex set S_t of pairwise incomparable points called the *staircase*, recursively. At the beginning G_0 is the empty graph and S_0 is the empty set. A point on the staircase is called a *stair-point*. Thus, before the t^{th} step we have a graph G_{t-1} on the points of P_{t-1} and a set S_{t-1} of pairwise incomparable points. In the t^{th} step we add p_t to our point set obtaining P_t and we will define the new staircase, S_t, and also the new graph, G_t, containing G_{t-1} as a subgraph. Before the exact definition of S_t and G_t, we make some more definitions and fix some properties that will be maintained during the process (see Figure 4).

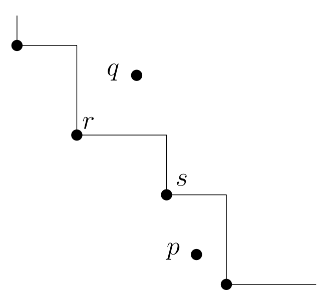


Figure 4: Definition 3.4: p is below the staircase, q is above the staircase, r and s are neighboring staircase points and r is the left neighbor of s.

Definition 3.4. *We say that a point p of P_t is* above *the staircase if there exists a stair-point $s \in S_t$ such that p is NE from s. If p is not above or on the staircase, then we say that p is* below *the staircase. A point below (resp. above) the staircase is called a* below-point *(resp.* above-point*). At any time t, we say that two points of S_t are* neighbors *if their x-coordinates are consecutive among the x-coordinates of the stair-points. (Note that this does not mean that they are connected in the graph.) We* *also say that p is the* left *(resp.* right*) neighbor of q if p and q are neighbors and the x-coordinate of p is less (resp. more) than the x-coordinate of q.*

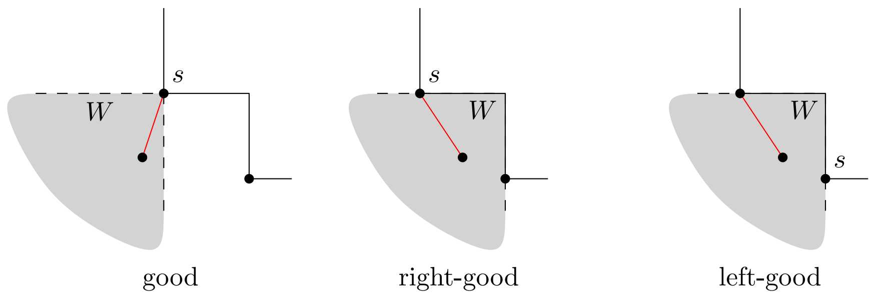


Figure 5: The stair-point s is good (resp. right-good, left-good) if W contains two points that are connected by an edge.

Definition 3.5. *In any step t, we say that a point p is* good *if any wedge containing p already contains two points connected by an edge, which are thus forced to get different colors (see Figure 5). I.e., at any time after t, a wedge containing p will contain points of both colors in the final coloring. A stair-point p is* almost-good *if for at least one of its neighbors, q, it is true that any wedge containing p and q contains two points connected by an edge of G_t. Additionally, if q is the left neighbor of p, then we say that p is* left-good*, and if q is the right neighbor of p, then we say that p is* right-good.

Notice that the good points and the neighbors of the good points are always almost-good. In fact, good points are also left- and right-good, and a left (resp. right) neighbor of a good point is right (resp. left) good.

Now we can state the properties we maintain at any time t.

Property 1. All above-points are good.

Property 2. All stair-points are almost-good.

Property 3. All below-points are in different components of G_t.

Property 4. G_t is a forest.

For $t = 0$, all these properties are trivially true. Whenever a new point arrives, we execute the following operations (see also Figure 6) repeatedly as long as it is possible, in any order. This will ensure that the properties remain true.

1Above: If an above-point p is not good, then we connect p by an edge with a stair-point that is SW from p.

2Comparable: If for some below-points p, q we have that q is NE from p, then connect them by an edge and put q on the staircase.

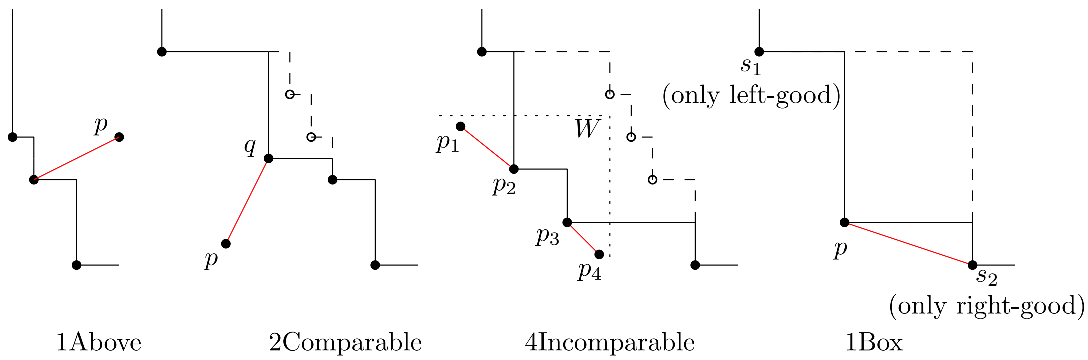


Figure 6: The operations maintaining the properties.

4Incomparable: Suppose there are no comparable below-points and there is a wedge W that lies entirely below the staircase and contains four incomparable points, q_1, q_2, q_3, and q_4, in order of their x-coordinates. Then connect q_1 with q_2 and also q_3 with q_4, and put q_2 and q_3 on the staircase.

1Box: Suppose there are no comparable below-points, and suppose s_1 and s_2 are two neighboring stair-points, s_1 is NW from s_2, s_1 is left-good but not right-good while s_2 is right-good but not left-good and p is a point in the rectangle defined by the two opposite vertices s_1 and s_2. We connect p and s_2, and put p on the staircase.

Now we have to verify that the properties remain true after executing an operation. First, we make the following observation, which implies that we will have to verify Property 2 only for the new stair-points.

Observation 3.6. *If a stair-point is left-good (resp. right-good, resp. good), then if after an operation this point is still a stair-point, then it remains left good (resp. right-good, resp. good).*

Proof. Notice that if a stair-point gets a new left-neighbor, then the new neighbor is either good, or right-good. Similarly, if a stair-point gets a new right-neighbor, then the new neighbor is either good, or left-good. □

Proposition 3.7. *After executing any operation, Properties* 1-4 *remain true.*

Proof. We check each operation and each property.

1Above: For Property 1, notice that p necessarily has to be the newly arrived point, p_t, and it becomes good after the operation. Property 2 obviously remains true. For Properties 3 and 4 we use that as p is the newly arrived point, before the operation it is not connected to any other point.

2Comparable: Property 1 remains true as only points NE from q became above-points and thus they all have the edge pq SW from them. Property 2 remains true as the only new stair-point is q, which became good. Properties 3 and 4 remain true as before the operation p and q were in different tree-components, which are then connected.

4Incomparable: Property 1 remains true as, using that no below-points were comparable, any point that became an above-point has either both q_1 and q_2, or both q_3 and q_4 SW from it. Property 2 remains true as there are only two new stair-points: q_2 becomes left-good and q_3 becomes right-good. Properties 3 and 4 remain true as before the operation q_1, q_2, q_3, q_4 were all in different tree-components, and after the operation two-two of these are connected in a suitable way.

1Box: Property 1 remains true as, using that no below-points were comparable, there are no new above-points. Property 2 remains true as p is the only new stair-point and it is right-good. Properties 3 and 4 remain true if p and s_2 are in different tree-components. This will be proved in Lemma 3.8 and Lemma 3.9.

Lemma 3.8. *If there is no below-point in the tree-component T_s of a stair-point s, then this remains true, i.e., later during the process the component containing s will never contain a below-point.*

Proof. Suppose that there is no below-point in the tree-component T_s of a stair-point s. This trivially remains true when a new point arrives (before doing operations). Then a simple case analysis shows that none of the operations can introduce a below-point to the tree-component T_s of a stair-point s:

1Above: Either T_s does not change or only p (an above-point) is added to it.

2Comparable: Only the components of the below-points p and q are joined, as T_s must be a different tree from these two (as it contained no below-point), T_s does not change.

4Incomparable: Only the components of the below-points $q_1, \dots, q_4$ change, as T_s must be a different tree from these (as it contained no below-point), T_s does not change.

1Box: Either T_s does not change or $s_2 \in T_s$ in which case T_s is joined with the tree T_p containing p. In the latter case in T_p the only below-point was p (by Property 3), which after the operation becomes a stair-point, so the new tree containing s, $T'_s = T_s \cup T_p$ still does not contain a below-point after this operation. □

Lemma 3.9. *Suppose s is a stair-point and b is a below-point in the tree-component T_s containing s. If s is right-good but not left-good, then b is lower than s, that is, b has a smaller y-coordinate than s.*

Proof. By Observation 3.6, we know that if s is right-good but not left-good, then this was also true at the time when s became a stair-point. A simple case analysis of the operations shows that at the time when s becomes a stair-point, the statement holds:

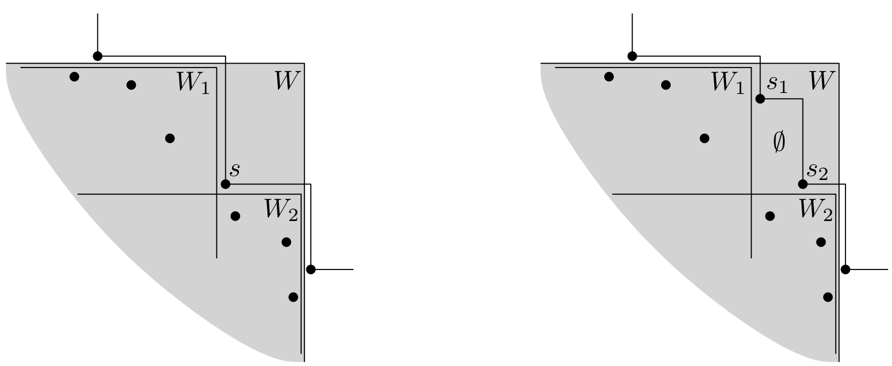
1Above: This is not a possible case as in this case no point becomes a stair-point.

2Comparable: This is not a possible case as necessarily s plays the role of q in the operation, in which case s is good and thus also left-good, contradicting our assumption on s.

4Incomparable: Necessarily s plays the role of q_3 in the operation, thus b is q_4 and so it is lower than s, as required.

1Box: Necessarily s plays the role of p in the operation, thus after the operation the below-point in T_s is necessarily the point which was the below point of T_{s_2} before the operation. As s_2 was right-good but not left-good, by induction this below point was lower than s_2, thus it is also lower than $s = p$, as required.

If after some step T_s stops to have a below-point, then by Lemma 3.8 this remains true and so there can be no below-point b in T_s as required by the lemma and we are done. Otherwise, if T_s still has a below-point, then by Property 3 there is exactly one below-point b in T_s, it is lower than s, and we have to check that after any operation the below-point in T_s remains below s. The only operation in which the below-point b in T_s could go higher is *4Incomparable* such that b plays the role of q_2. If $b = q_2$ is SW from s, then s goes above the staircase, thus stops being a stair-point as required by the lemma and we are done. If $b = q_2$ is SE from s, then the whole wedge W must be lower than s, and then the new below-point in T_s becomes q_1, also lower than s. This finishes the proof of the lemma and also of Proposition 3.7. ☐



(a) (b)

Figure 7: A monochromatic wedge can contain at most 8 points.

Now we can finish the proof of the dynamic dual version, and thus also of Theorem 3.1, by showing that taking any (partial) two-coloring of the forest G_t constructed using the above operations, at all times (i.e., for every prefix set $\{p_1, \dots, p_t\}$ of the point set), any quadrant W containing at least 9 points contains both colors. Fix the time after the arrival of the point p_t (and after we repeatedly applied the operations as long as possible). Thus no more operations can be applied, in particular there are no two comparable below-points otherwise we could apply operation *2Comparable*. If W contains an above-point, it contains both colors as all above-points are good. If W contains at most one stair-point, s, then by "splitting" W at s (see Figure 7(a)), we get two quadrants that do not contain any stair-point, but contain all other points that W contains. One of these two quadrants must contain at least 4 below-points, thus we could apply operation *4Incomparable*, a contradiction. If W contains at least 3 stair-points, then it contains a stair-point s such that both neighbors of s are also in W. As every stair-point is almost-good, W must contain both colors. Finally, if W contains exactly two (neighboring) stair-points, s_1 NW from s_2, then the only way for W to be monochromatic is if s_1 is left-good but not right-good and s_2 is right-good but not left-good. Therefore, there can be no points in the rectangle formed by s_1 and s_2, as otherwise we could apply operation *1Box*, a contradiction. At least one of the two quadrants obtained by "splitting" W at s_1 and s_2 (see Figure 7(b)), must contain at least 4 below-points, thus we could apply operation *4Incomparable*, a contradiction.

3.2 Proof of Theorem 3.2

Here we construct for any triangle T a finite point set P such that for every two-coloring of P there is a translate of T that contains exactly 4 points and all of these have the same color. As the construction is quite hard to describe precisely, we refer to Figures 8 and 9 for the details, and we only give an informal description below. With a simple case analysis, we will show that in any two-coloring, there is a monochromatic triangle with exactly 4 points.

The "big picture" is on Figure 8 that shows what the construction looks like from far. The thicker triangles denote families of triangles that are very close to each other. The center part has only three points, p_1, p_2 and p_3. Two of these, without loss of generality, p_1 and p_3, must receive the same color, say blue.

After this, we look more closely at the family $\mathcal{T}_2$ that consists of the subfamilies $\mathcal{T}_{2,1}$, $\mathcal{T}_{2,2}$, $\mathcal{T}_{2,3}$ and $\mathcal{T}_{2,4}$, see the left part of Figure 9. Unless the triangle $T_{2,0}$ is monochromatic, at least one of $p_{2,0,1}$, $p_{2,0,2}$, $p_{2,0,3}$ and $p_{2,0,4}$ must be blue. Without loss of generality, we suppose $p_{2,0,3}$ is blue.

After this we look more closely at the family $\mathcal{T}_{2,3}$ that consists of the triangles $T_{2,3,1}$, $T_{2,3,2}$, $T_{2,3,3}$ and $T_{2,3,4}$, see the right part of Figure 9. Unless the triangle $T_{2,3,0}$ is monochromatic, at least one of $p_{2,3,1}$, $p_{2,3,2}$, $p_{2,3,3}$ and $p_{2,3,4}$ must be blue. But if $p_{2,i,3}$ is blue, then $T_{2,i,3}$ is monochromatic. This finishes the proof.

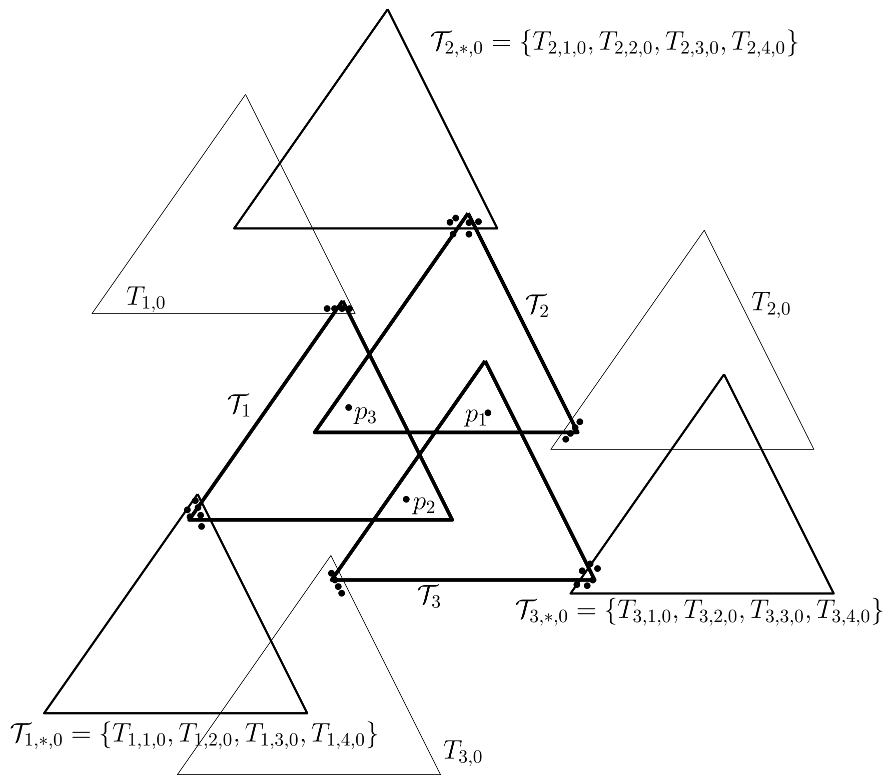
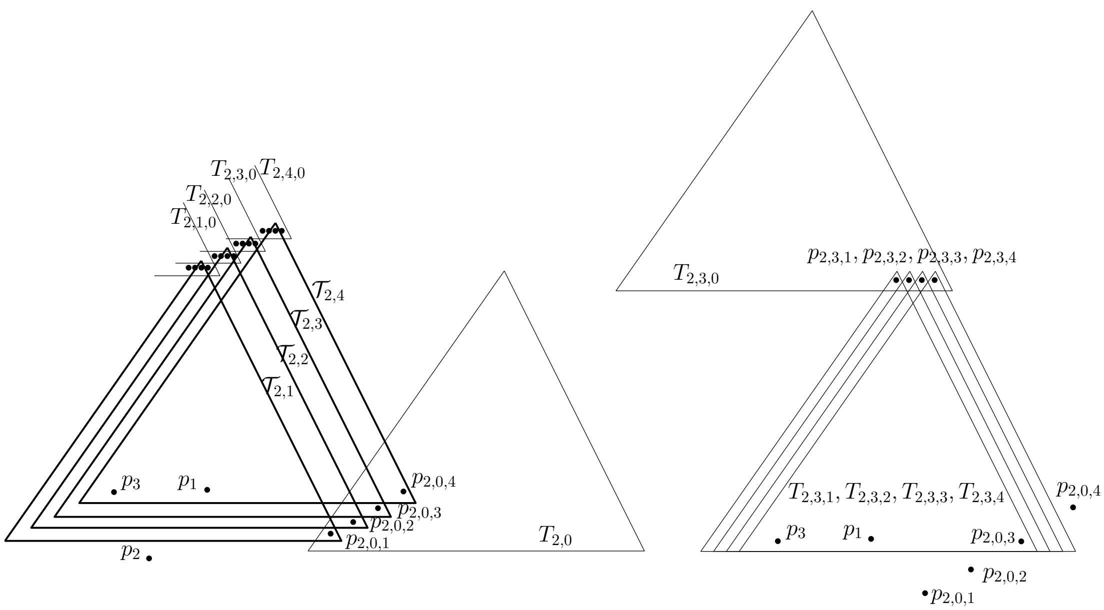


Figure 8: The construction in which there is always a monochromatic triangle with 4 points.



(a) zooming into $\mathcal{T}_2$

(b) zooming into $\mathcal{T}_{2,3}$

Figure 9: Zooming into the construction in two stages.

4 Online and dynamic colorings

This chapter is based on the article [44], joint work with Keszegh and Lemons.

In this chapter we consider proper online colorings of hypergraphs defined by geometric regions. We prove that there is an online coloring method that colors N intervals of the real line using $\Theta(\log N/k)$ colors such that for every point p, contained in at least k intervals, not all the intervals containing p have the same color. We also show that this is best possible. We also prove the corresponding result about online coloring quadrants in the plane that are translates of a given fixed quadrant. These results contrast to the results of the previous chapter, which showed that in the dynamic setting 2 colors are enough to color quadrants if $k \leq 9$ (independent of N). We also consider coloring intervals in the dynamic setting. In all cases we present efficient coloring algorithms as well.

Despite the well-known applications of the study of colorings of geometric graphs and hypergraphs, relatively little attention has been paid to the online and dynamic versions of these problems. In online coloring problems, the set of objects to be colored is not known beforehand; objects come to be colored one-by-one and a proper coloring must be maintained at all times. We give asymptotically tight bounds on the number of colors necessary to properly online color wedges in the plane and then relate these to bounds on the number of colors necessary to properly online color intervals.

In dynamic coloring, objects come online and must be colored one by one, such that a valid coloring is maintained at each step, yet the objects and their order are known in advance. Such problems can be used to solve corresponding offline higher dimensional problems, as we have seen earlier. This was the motivation to revisit the problem about dynamic coloring intervals.

4.1 Definitions and description of main results

In this chapter, a *wedge* in the plane is a set of points $\{(x,y) \in \mathbb{R}\times\mathbb{R} | x \leq x_0 \ y \leq y_0\}$ for some x_0, y_0. We are interested in coloring with c colors a set of wedges (or intervals) such that for any point contained in at least k wedges (respectively intervals) not all the wedges (intervals) containing that point have the same color. We will often refer to this simply as a coloring of the set of points; the parameters k and c will be obvious from the context. For a collection of c-colored points in the plane, we define the associated *color-vector* to be a vector of length c where the i^{th} coordinate is the size of the largest wedge consisting only of points with color i. The size of the color-vector refers to the sum of its coordinates.

Note that given a set of wedges (or intervals) we can define a hypergraph whose vertices are the wedges (intervals) and whose edges consist exactly of those subsets of the vertices such that there exists a point contained exactly in the corresponding wedges (intervals). Then the coloring problem above is exactly the problem of finding a proper coloring of this hypergraph if we disregard those edges which contain less than k points.

For convenience, we will work with the equivalent dual-form of the wedge-coloring: a finite set of points S in the plane is called *k-properly c-colored* if S is colored with c colors such that every wedge intersecting S in at least k points contains at least two points of different colors. A wedge containing points of only one color is said to be *monochromatic*.

A reformulation of Theorem 2.24 is the following statement: Any finite family of wedges in the plane can be colored dynamic with 2 colors such that any point contained by at least $k = 9$ of these wedges is not monochromatic.

Gábor Tardos asked (personal communication) whether such a coloring can be achieved in a completely online setting, possibly with a larger k and more colors (again such that all large wedges are non-monochromatic). It is easy to see that 2 colors are not enough to guarantee colorful wedges (i.e. there may be arbitrarily large monochromatic wedges) even when the points are restricted to a diagonal line. However, 3 colors (and $k = 2$) are enough if the points are restricted to a diagonal, see [45]. We prove that in general, for any c and k, there exists a method of placing points in the plane such that any online-coloring of these points with c colors will result in the creation of monochromatic wedges of size at least k.

Next, we consider the cases when either c or k is fixed. When c is fixed, we derive upper bounds on k (in terms of n) for which there is always a k-proper online c-coloring of a set of n points. For fixed k, we derive bounds on c for which there is always a k-proper online c-coloring of n points.

The rest of this chapter is organized as follows.
In Section 4.2 we study online coloring points with respect to wedges.
In Section 4.3 we show how our results on online colorings of wedges directly relates to the online coloring of intervals.
In Section 4.4 we describe our results on dynamic colorings of intervals.
Finally, in Section 4.5, based on parts from [49], we investigate various dynamic hypergraphs defined by intervals on a line.

4.2 Online coloring wedges

Our first result is a negative answer to the question of Tardos: for every c and k, there exists a method of placing points in the plane such that any online-coloring of these points with c colors will result in the creation of monochromatic wedges of size at least k. Actually we prove a stronger statement.

Theorem 4.1. *There exists a method to give $N = 2^n - 1$ points in a sequence such that for any online-coloring method using c colors there will be c monochromatic wedges, $W_1, W_2, \dots, W_c$, and nonnegative integers $x_1, \dots, x_c$ such that for each i, the wedge W_i contains exactly x_i points colored with color i and $\sum x_i \geq n + 1$ if $n \geq 2$.*

Corollary 4.2. *No online-coloring method using c colors can avoid to make a monochromatic wedge of size $k + 1$ for some sequence of $N = 2^{ck} - 1$ points.*

Proof of Theorem 4.1. By induction on the size of the color-vector. Clearly, one point gives a color-vector of size 1. Two points guarantee a color-vector of size 2 if they are placed diagonally from each other. Now we can place the third point diagonally between the first two if they had a different color or place it diagonally below them to get a color-vector of size 3 for three points.

By the inductive hypothesis, using at most $2^{n-1} - 1$ points, we can force a color-vector with size n. Moving southeast (i.e. so that all the new points are southeast from the previous ones) we repeat the procedure, again using at most $2^{n-1} - 1$ points we can force a second color-vector with sum n. If the two color-vectors are different then the whole point set has a color-vector of size at least $n + 1$. If they are the same then we put an additional point southwest from all the points of the first set of points but not interfering with the second set of points. Then as this point is colored with some color, i, the i^{th} coordinate of the first color-vector increases. (The rest of its coordinates will becomes 0.) Together with the wedges found in the second set of points, we can see that the size of the color-vector of the whole point set increased by one. Altogether we used at most $2(2^{n-1} - 1) + 1 = 2^n - 1$ points, as desired. □

What happens if c or k is fixed? The case when $c = 2$ was considered, e.g., in [45]. It is not hard to see that using $2k-1$ points, the size of the largest monochromatic wedge can be forced to be at least k and this is the best possible. For $c = 3$ a quadratic (in k) number of points is needed to force a monochromatic wedge of size k:

Proposition 4.3. *The following statements hold.*

1. *There exists a method of placing k^2 points such that any online 3-coloring of these points produces a monochromatic wedge of size at least k.*
2. *There exists a method of online 3-coloring $k^2 - 1$ points such that all monochromatic wedges have size less than k.*

We use the following terminology. Given a collection of 3-colored points in the plane, we say a new, uncolored point x is a *potential member* of a monochromatic wedge W, if by giving x the color of W, the size of W increases. Furthermore if x is a potential member of W, and giving x a color different from the color of W destroys the wedge W, then x *threatens* the wedge W.

Proof. To prove the first statement, consider the largest monochromatic wedge of each color after some points have already been placed and colored. Moving in the northwest / southeast directions label the wedges W_1, W_2 and W_3. It is clear that W_2 lies between the other two wedges. Note that it is possible to place a new point directly southwest of the points in W_2 such that the point is a potential member of all three wedges but only threatens W_2. Thus if the point is assigned the color of one of the other wedges (say W_1), the size of W_1 increments while W_3 remains the same and W_2 is destroyed (it is no longer monochromatic). Now suppose W_2 is not larger than either of the other two wedges. In this case, a point is placed, as described above, such that it is a potential member of all three wedges. If the point is assigned the color of W_1 or W_3 then W_2 is destroyed and while W_1 (or W_3) moves from size i to $i + 1$, the $j \le i$ points of W_2 are rendered ineffective for forming monochromatic wedges. On the other hand suppose W_2 has size larger than (at least) one of the other wedges (say W_3). Then we forget the wedge W_3 and proceed as above where there is a new wedge W_0 (of size 0) between W_1 and W_2. (We can think that in the previous step W_2 increased from i to $i + 1$ while the $j \le i$ points of W_3 are destroyed.) As we proceed in this way, the sizes of the two "side" wedges only increase at each step while the "middle" wedge may be reduced to size 0 at some steps. However, the j points of the middle wedge are only destroyed when a side wedge increases from i to $i+1$ when $i \geq j$. Thus by destroying at most $2\binom{k}{2}$ points we can guarantee that the two side wedges have size $k-1$. Adding at most k more points to the middle, a monochromatic wedge of size k is guaranteed.

To prove the second statement we must assign colors to the points to avoid a monochromatic k-wedge. When a new point, x is given, consider those wedges of which x is a potential member. Note that at most two of these are not threatened by x. Let s be the size of the smallest wedge, W which is not threatened by x but of which x is a potential member. If by giving x the color of W at least s points are destroyed among the wedge(s) threatened by x, then give x this color. Otherwise give x the color different from the two non-threatened wedges. In this way we guarantee that a wedge only increases in size from i to $i+1$ if at the same time i other points are destroyed (i.e. rendered ineffective) or if two other wedges of size $i+1$ already exist. Therefore if only $2\sum_{i=1}^{k-2} i + 3(k-1) = k^2 - 1$ vertices are online-colored, we can avoid a monochromatic k-wedge. □

For $c \geq 4$ we can give an exponential (in ck) lower bound for the worst case:

Theorem 4.4. *For $c \geq 4$ we can online color with c colors any set of $N = O(1.22074^{ck})$ points such that throughout the process there is no monochromatic wedge of size k. Moreover, if c is large enough, then we can even online color $N = O(1.46557^{ck})$ points.*

Proof. Denote the colors by the numbers $\{1, \ldots, c\}$. A wedge refers to both an area in the plane as well as the collection of placed points which fall within that area. For brevity, we will often refer to maximal monochromatic wedges as simply wedges. If a wedge is not monochromatic, we will specifically note it. At each step, we define a partition of all the points which have come online in such a way that each set in the partition contains exactly one maximal monochromatic wedge. Two maximal monochromatic wedges are called *neighbors* if they are contained within a larger (non-monochromatic) wedge which contains no other monochromatic wedges. If the placement and coloring of a point cause a wedge to no longer be monochromatic, that wedge has been *killed*.

We now describe how to color a new point given that we have already colored some (or possibly no) points. If the new point is Northeast of an earlier point, it is given a different color from the earlier point. In this case no new wedges are created and no wedges increase in size. Otherwise the new point will eventually be part of a wedge. We want to make sure that the color of the point is distinct from its neighbors colors. In particular, consider the (at most) two wedges which are neighbors of wedges containing the point but which do not actually contain the point. From the c colors we disregard these two colors. From the remaining, we choose the color which first minimizes the size of the wedge containing the point and secondly minimizes the color (as a number from 1 to c.) This means that our order of preference is first to have size 1 wedge of color 1, then a size 1 wedge of color 2, …, size 1 wedge of color c, size 2 wedge of color 1, … etc. These rules determine our algorithm. For an illustration see Figure 10 for a 5-coloring, where the new vertex v cannot get the neighboring colors 2 and 4 and by our order of preference it gets color 3, thus introducing a (monochromatic) wedge (of color 3) of size 2. Now we have to see how effective this coloring algorithm is.

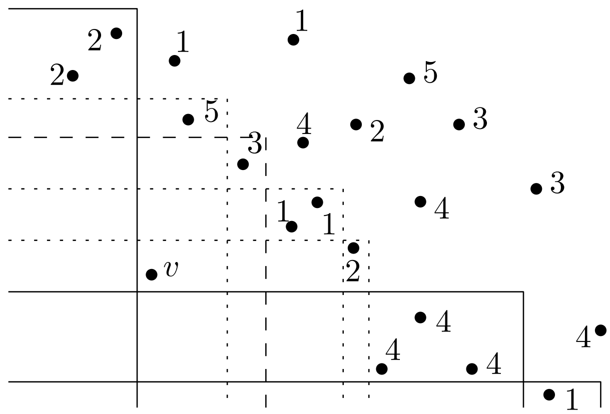


Figure 10: A general step of the coloring in the proof of Theorem 4.4.

To prove the theorem we show that the partition set associated with the newly created (or incremented) wedge is relatively large. Suppose this wedge is of color i and size j. Let $A_{i,j}$ denote the smallest possible size of the associated partition set. One can regard $A_{i,j}$ as the least number of points that are required to "build" a wedge W of size i of color j. For simplicity, we also use the notation $B_{c(i-1)+j} = A_{i,j}$. Note that this notation is well defined as j is always less than c. Thus we have $B_1 = B_2 = B_3 = 1$ and $B_4 = 2$.

It follows from our preferences that $B_i \leq B_j$ if $i \leq j$. Our goal is to give a good lower bound on $A_{k,1} = B_{c(k-1)+1}$.

Notice that when we create a new wedge, it will kill many points that were contained in previous wedges. More precisely, from our preferences we have $B_i \geq 1 + B_{i-3} + B_{i-4} + \ldots + B_{i-c}$ (where $B_i = 0$ for $i \leq 0$). Note that B_{i-1} and B_{i-2} is missing from this sum because the coloring method must choose a color different from the new points two to-be-neighbors'.

From the solution of this recursion we know that the magnitude of B_i is at least q^i where q is the (unique, real, > 1) solution of $q^c = (q^{c-2} - 1)/(q - 1)$, which is equivalent to $q^{c+1} = q^c + q^{c-2} - 1$. Moreover, since trivially $B_i \geq 1 \geq q^{i-c}$ if $i \leq c$, from the recursion we also have $B_i \geq q^{i-c}$ for all i. If we suppose $c \geq 4$, then $q \geq 1.22074$ and from this $B_{c(k-1)+1} \geq 2.22074^{k-2}$. As c tends to infinity, q tends (from below) to the real root of $q^3 = q^2 + 1$, which is ≥ 1.46557. From this we obtain that $B_{c(k-1)+1} \geq 1.46557^{c(k-2)}$ if c is large. Also, in the special case $k = 2$, we get the well known sequence A000930 (see OEIS), which is at least 1.46557^c, if c is big enough. $\square$

Summarizing, if we have $c \geq 4$ colors, the smallest N_0 number of points that forces a monochromatic wedge of size k is exponential in ck. Thus, if the number of colors, c, is given, these bounds give an estimate of $\Theta(\log N/c)$ on the size of the biggest monochromatic wedge in the worst case.

If we consider k fixed (and we want to use as few colors as possible), by the above bound the number of colors needed to avoid a monochromatic wedge of size k is $O(\log N/k)$ for $k \geq 1$.

Corollary 4.5. *There is a method to online color N points in the plane using at most $O(\log N/k)$ colors such that all monochromatic wedges have size strictly less than k.*

Recall that Theorem 4.1 stated that $N = 2^n - 1$ points can always force a size $n+1$ color-vector (for the definition see the proof of Theorem 4.1). We remark that Theorem 4.4 implies a lower bound close to this bound too. Indeed, fix, e.g., $c = 4$ and $k = \lceil n/4 \rceil$. If the number of points is at most $N = O(1.22074^n) = O(1.22074^{ck})$ then by Theorem 4.4 there is an online coloring such that at any time there is no monochromatic wedge of size k, thus the color-vector is always at most $4(k-1) < n$.

Suppose now that k is fixed and we want to use as few colors as possible without knowing in advance how many points will come, i.e. for k fix we want to minimize c without knowing N. To solve this, we alter our previous algorithm. (Note that we could also easily adjust the algorithm if for an unknown N we want to minimize $\min(c,k)$, or ck, the answer would be still logarithmic in N.) All this comes with the price of loosing a bit on the base of the exponent. The following theorem implies that for $k = 2$ (and thus also for any $k \geq 2$) we can color any set of $N = O(1.0905^{ck})$ points and if k is big enough then we can color any set of $N = O(1.1892^{ck})$ points.

Theorem 4.6. *For fixed $k \geq 1$ we can color a countable set of points such that for any c, and any $n < 2^{\lfloor (c+1)/4 \rfloor (k-1)}$, the first n points of the set are k-properly c-colored.*

Proof. We need to define a coloring algorithm and prove that it uses many colors only if there were many points. The coloring and the proof is similar to the proof of Theorem 4.4, we only need to change our preferences when coloring and because of this the analysis of the performance of the algorithm differs slightly too. We fix a c and an $N < 2^{\lfloor (c+1)/4 \rfloor (k-1)}$ for which we will prove the claim of the theorem (the coloring we define can obviously not depend on c or N, but it depends on k). Denote the colors by the numbers $\{1, 2, \dots, c-1, c, \dots\}$.

We can suppose again that every new point will be on the actual diagonal. When we add a point its color must still be different from its to-be-neighbors' and together with this point we still cannot have a monochromatic wedge of size k. Our primary preference now is that we want to keep $\lfloor c/4 \rfloor$ small where c is the color of the new point (as a number). Our secondary preference is that the size of the biggest wedge containing the new point should be small.

This means that our order of preference is first to have size 1 wedge of color 1, then a size 1 wedge of color 2, size 1 wedge of color 3, a size 1 wedge of color 4, a size 2 wedge of color 1, ..., a size $k-1$ wedge of color 1, size $k-1$ wedge of color 2, a size $k-1$ wedge of color 3, a size $k-1$ wedge of color 4, size 1 wedge of color 5, size 2 wedge of color 5, ... etc. These rules determine our algorithm, now we have to see how effective it is.

$A_{i,j}$ is defined as in the proof of Theorem 4.4. We only need to prove that $A_{i,j} \geq 2^{(k-1)(\lfloor j/4 \rfloor + i - 1)}$ as this means that if the algorithm uses the color $c+1$, then we had at least $A_{1,c+1} \geq 2^{(k-1)\lfloor (c+1)/4 \rfloor} > N$ points, a contradiction. Recall that $A_{i,j}$ denotes the least number of points that are required to "build" a wedge W of size i of color j.

We prove by induction. First, $A_{1,1} = A_{1,2} = A_{1,3} = A_{1,4} = 1$ indeed. By our preferences, whenever we introduce a size one wedge with color j, we had to kill at least two (four minus the two forbidden colors of the neighbors of the new point) points that have colors from the previous 4-tuple of colors and are contained in monochromatic wedges of size $k-1$. Thus $A_{1,j} \geq A_{k-1,4(\lfloor j/(4-1)\rfloor+1)} + A_{k-1,4(\lfloor j/(4-1)\rfloor+2} \geq 2 \cdot 2^{(k-1)(\lfloor j/(4-1)\rfloor+k-2)} = 2 \cdot 2^{(k-1)(\lfloor j/4\rfloor-1} = 2^{(k-1)(\lfloor j/4\rfloor+1-1}$. If we introduce a wedge of size $i > 1$ with color j, we had to kill at least two points that have colors from the same 4-tuple of colors as j and are contained in monochromatic wedges of size $i-1$. Thus in this case $A_{i,j} \geq A_{i-1,4(\lfloor j/4\rfloor+1} + A_{i-1,4(\lfloor j/4\rfloor+2} \geq 2 \cdot 2^{(k-1)(\lfloor j/4\rfloor+i-2)} = 2^{(k-1)(\lfloor j/4\rfloor+i-1}$. □

Proposition 4.7. *The online coloring methods, guaranteed by the second part of Proposition 4.3 and Theorems 4.4 and 4.6, run in $O(n \log n)$ time to color the first n points (even if we have a countable number of points and n is not known in advance).*

The proof of this proposition is omitted as it follows easily from the analysis of the algorithms.

4.3 Online coloring intervals

This section deals with the following *interval coloring problem.* Given a finite family of intervals on the real line, we want to online color them with c colors such that throughout the process if a point is covered by at least k intervals, then not all of these intervals have the same color.

Proposition 4.8. *The interval coloring problem is equivalent to a restricted case of the point with respect to wedges coloring problem, where we care only about the wedges with apex on the line L defined by $y = -x$.*

Proof. Consider the natural bijection of the real line and L. Associate to every point p of L the wedge with apex p and associate with every interval $I = (x_1, -x_1), (x_2, -x_2)$ of L the point $(x_1, -x_2)$. It is easy to see that $p \in I$ if and only if the point associated to I is contained in the wedge associated to p. □

Corollary 4.9. *Any upper bound on the number of colors necessary to (online) color wedges in the plane is also an upper bound for the number of colors necessary to (online) color intervals in $\mathbb{R}$.*

Also the lower bounds of Theorem 4.1 and of Proposition 4.3 follow for intervals easily by either repeating the proofs for intervals or by Observation 4.10:

Observation 4.10. *The proofs of Theorem 4.1 and of the first part of Proposition 4.3 can be easily modified such that all the relevant wedges have their apex on the line $y = -x$.*

In particular, we have the following.

Corollary 4.11. *There is a method to online color N intervals in $\mathbb{R}$ using $\Theta(\log N/k)$ colors such that for every point x, contained in at least k intervals, there exist two intervals containing x of different colors.*

As we have seen the results about intervals follow in a straightforward way from the results about wedges. Thus all the statements we proved hold for online coloring wedges, also hold for intervals, however, it seems unlikely that the exact bounds are the same. Thus, we would be happy to see (small) examples where there is a distinction. As the next section shows, there is a difference between the exact bounds for dynamic coloring wedges and intervals.

4.4 Dynamic coloring of intervals

A dynamic coloring of an ordered collection of intervals $\{I_t\}_{t=1}^n$, is a coloring ϕ such that for every k, the sub-collection $\{I_t\}_{t=1}^k$ is properly colored under ϕ.

Theorem 4.12. *Any finite family of intervals on the line can be dynamic colored with* 2 *colors such that at any time any point contained by at least* 3 *of these intervals is not monochromatic.*

We exploit an idea used in Chapter 3; instead of online coloring the intervals we online build a labelled acyclic graph (i.e., a forest) with the following properties. Each interval will correspond to a vertex in this graph (there might be other vertices in the graph as well). The final coloring of the intervals will then be generated from this graph. In particular, to define a two-coloring, we will assign each edge in the forest one of two labels, "different" or "same". For an arbitrary coloring of exactly one vertex in each component (tree) of the graph, there is a unique extension to a coloring of the whole graph compatible with the labelling, i.e., such that each edge labelled "same" is adjacent to vertices of the same color and each edge labelled "different" is adjacent to vertices of different colors. In Chapter 3 all the edges were labelled "different" so it was actually a simpler variant of our current scheme. As we will see, this idea can also be generalized to more than two colors.

We denote the color of an interval I by $\phi(I)$, the left (resp. right) endvertex of I by $\ell(I)$ (resp. by $r(I)$). These vertices are real numbers, and so they can be compared.

Proof of Theorem 4.12. Let $\{I_t\}_{t=1}^n$ be the given enumeration of the intervals to be dynamic colored. We first build the forest and then show that the coloring defined by this forest works. As we build the forest we will maintain also a set of intervals, called the active intervals (not necessarily a subset of the given set of intervals). At any time t the vertices of the actual forest correspond to the intervals of $\{I_t\}_{t=1}^n$ and the set of current or past active intervals. The set of active intervals will change during the process, but we maintain that the following properties hold any time.

1. Every point of the line is covered by at most two active intervals.
2. No active interval contains another active interval.
3. A point is either forced by the labelling to be contained in original intervals of different colors or it is contained in the same number of active intervals as original intervals, and additionally the labelling forces these original intervals to have the same colors as these active intervals.
4. Each tree in the forest contains exactly one vertex that corresponds to an active interval.

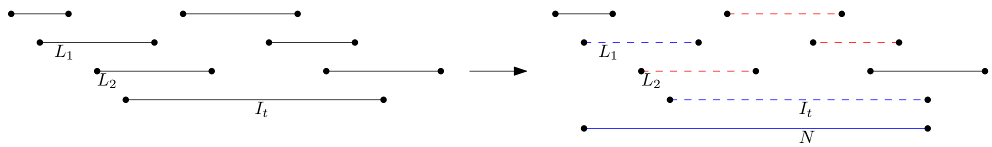


Figure 11: A general step in the proof of Theorem 4.12.

The last property ensures that a coloring of the active intervals determines a unique coloring of all the intervals which is compatible with the labelling of the forest.

Note that in the third property one or two of the original intervals can actually coincide with one or two of the active intervals.

For the first step we simply make the first interval active; our forest will consist of a single vertex corresponding to this interval. In general, at the beginning of step t, we have a list of active intervals, $\mathcal{J}_{t-1}$. Consider the t^{th} interval, I_t. If I_t is covered by an active interval, $J \in \mathcal{J}_{t-1}$, then we add I_t to the forest and connect it to J with an edge labelled "different". Note that there is at most one such active interval. If there is no active interval containing I_t, we add I_t to the set of active intervals and also add a corresponding vertex to the forest. Now if there are active intervals contained in I_t, these are all deactivated (removed from the set of active intervals) and each is connected to I_t in the graph with an edge labelled "different". Note that this way any point covered by these inactivated intervals will be covered by intervals of both colors.

It remains to ensure that no point is contained within three active intervals. If there still do exist such points, by induction they must be contained within I_t. Let L_1 and L_2 be the (at most) two active intervals covering $\ell(I_t)$ such that $\ell(L_1) < \ell(L_2)$ (if both of them exist.) Similarly, let R_1 and R_2 be the (at most) two active intervals covering $r(I_t)$ such that $\ell(R_1) < \ell(R_2)$ (if both of them exist.) We note that the L_i and R_j cannot be the same, as such an interval would cover I_t. Also, no other active intervals can intersect I_t, as they would necessarily be contained in I_t. Without loss of generality, we can assume that both L_1 and L_2 exist. If R_1 and R_2 also both exist, deactivate L_1, L_2, I_t, R_1 and R_2 and activate a new interval $N = L_1 \cup I_t \cup R_2$ (and add a corresponding vertex to the graph). In the graph, connect L_1, I_t and R_2 to N with edges labelled "same". Connect L_2 and R_1 to N with edges labelled "different". Otherwise, if at most one active edge contains $r(I_t)$ we deactive L_1 and L_2 and connect these to the new interval $N = L_1 \cup I_t$ (again with edges labelled "same" and "different", respectively), also we deactive I_t and connect it to N with an edge labelled "same". Figure 11 is an illustration of this case when the active interval N is assigned color blue and deactivated intervals are shown with dashed lines.

This way within a given step, any point which is contained in (at least) two intervals deactivated during the step, is forced by the labelling to be contained in intervals of different colors. For any other point v the number of original intervals containing v remains the same as the number of active intervals covering v (both remains the same or both increases by 1). The first three properties were maintained and also it is easy to check that the graph remains a forest such that in each component there is a unique active interval.

At the end of the process any coloring of the final set of active intervals extends to a coloring of all the intervals (compatible with the labelling of the graph). We have to prove that for this coloring at any time any point contained by at least 3 of these intervals is not monochromatic. By induction at any time $t < n$ the coloring is compatible with the graph at that time, thus by induction any point contained by at least 3 of these intervals is not monochromatic. Now at time n, if the active intervals are colored, the extension (by induction) is such that every point not in I_n is either covered by at most two original intervals or it is covered by intervals of both colors. On the other hand, from the way we defined the graph, we can see that points covered by I_n and contained in at least 3 intervals are covered by intervals of both colors as well. Indeed, by the properties maintained, if a point v is not covered by intervals of both colors than it is covered by as many active intervals as original intervals. Yet, no point is covered by more than 2 active intervals at any time, thus v is covered by no more than 2 active and thus no more than 2 original intervals. □

Theorem 4.13. *Any finite family of intervals on the line can be dynamic colored with* 3 *colors such that at any time any point contained by at least* 2 *of these intervals is not monochromatic.*

Proof. We proceed similarly to Theorem 4.12. In particular, we require the same four properties from active intervals, although the second two need some modifications. Now instead of a labelled graph we define rules of the following form: some interval I (original or auxiliary) gets a different color from at most two other intervals J_1, J_2. We say that I depends from J_1, J_2, otherwise I is independent. If there is an order on the intervals such that an interval depends only on intervals later in this order then starting with any coloring of the independent intervals and then coloring the dependent ones from the last going backwards we can naturally extend this coloring to all the intervals such that the coloring is compatible with the rules (i.e. I gets a color different from the color of J_1, J_2 for all dependent triples). For a representation with directed acyclic graphs - showing more clearly the similarities with the previous proof - see the proof of Theorem 4.14.

1. Every point of the line is covered by at most two active intervals.
2. No active interval contains another active interval.
3. A point of the line is either forced by the rules to be contained in original intervals of different colors or it is contained in the same number of active intervals as original intervals, and additionally the rules force these original intervals to have the same colors as these active intervals.
4. An interval is independent if and only if it is an active interval.

The first two properties ensure the following structure on the set of active intervals. Define a chain as a sequence of active intervals such that everyone intersects the one before and after it in the chain. The set of active intervals can be partitioned into disjoint chains. The last property guarantees that any coloring of the active intervals extends naturally and uniquely to a coloring of all the intervals which is compatible with the rules.

We will define the rules such that if we start by a proper coloring of the active intervals then the extension is a dynamic coloring (as required by the theorem) of the original set of intervals. Note that in the previous proof we started with an arbitrary coloring of the active intervals, which was not necessarily proper, thus now we additionally have to take care that a proper coloring of the active intervals extends to a coloring which is a proper coloring of the active intervals at any previous time as well.

In the first step we add I_1 and activate it. In the induction step we add I_t to the set of active intervals. If I_t is covered by an active interval or by two intervals of a chain, then we deactivate I_t and the rule is that we give a color to it differing from the color(s) of the interval(s). If I_t does not create a triple intersection, it remains activated. Otherwise, denote by L (resp. R) the interval with the leftmost left end (resp. rightmost right end) that covers a triple covered point. We distinguish two cases.

Case i) If I_t is not covered by one chain, then either L or R is I_t, or L and R are not in the same chain. In either case we deactivate all intervals covered by $N = L \cup I_t \cup R$, except for L, I_t and R. The rule to color the now deactivated intervals is that they get a color different from I_t, in an alternating way along their chains starting from L and R.

It is easy to check that the four properties are maintained.

Given a proper coloring of the active intervals at step n by our rules it extends to a proper coloring of the active intervals in the previous step. Thus by induction at any time $t < n$ for any point v it is either covered by differently colored intervals or it is covered by at most one interval. For time n it is either covered by differently colored intervals or it is covered by as many original intervals as active intervals, and they have the same set of colors (by the third property). As the coloring was proper on the active intervals, v is either covered by two original intervals and then two active intervals which have different colors, thus the original intervals have different colors as well, or v is covered by at most one active and thus by at most one original interval.

Case ii) If I_t is covered by one chain, then L and R both differ from I_t. We deactivate all intervals covered by $L \cup I_t \cup R$ (including I_t), except for L and R. Notice that apart from I_t these intervals are all between L and R in this chain.

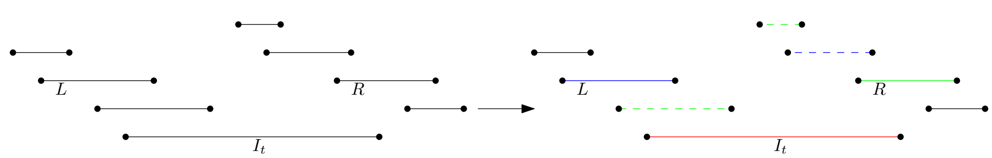


Figure 12: Case i) of Theorem 4.13

If we deactivated an odd number of intervals this way (so an even number from the chain), then we insert the new active interval L' that we get from L by prolonging the right end of L such that L' and R intersect in an epsilon short interval. We deactivate L and the rule is to color it the same as we color L'. The rule to color the deactivated I_t is to color it differently from the color of L' (or, equivalently, L) and R. The rule to color the deactivated intervals of the chain is to color them in an alternating way using the colors of L and R (in a final proper coloring of the active intervals they get different colors as L' and R intersect). If we deactivated an even number of intervals

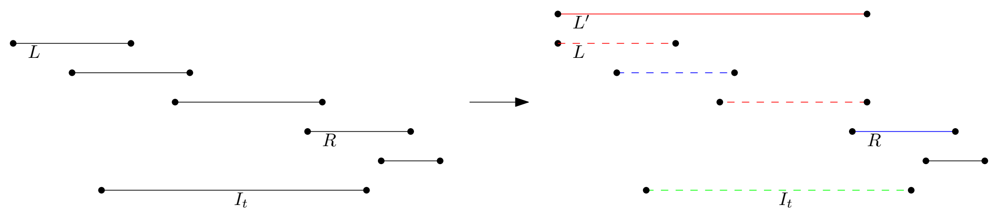


Figure 13: Case ii), odd subcase of Theorem 4.13

this way (so an odd number from the chain), then we deactivate L and R as well and add a new active interval $N = L \cup I_t \cup R$. The rule to color the deactivated I_t is to color it differently from the color of N. The rule to color the deactivated intervals of the chain is to color them in an alternating way using the color of N (L and R get this color) and the color that is different from the color of N and I_t.

It is easy to check that the four properties are maintained. Also, similarly to the previous case, it can be easily checked that if we extend a proper coloring of the active intervals then for its extension it is true at any time (for time $t < n$ by induction, otherwise by the way we defined the rules) that every point is either covered by at most one original interval or it is covered by intervals of different colors. □

Theorem 4.14. *Colorings guaranteed by Theorem 4.12 and Theorem 4.13 can be found in $O(n \log n)$ time.*

Proof. Instead of a rigorous proof we provide only a sketch, the easy details are left to the reader. In both algorithms we have n intervals, thus n steps. In each step we define a bounded number of new active intervals, thus altogether we have cn regular and active intervals. We always maintain the (well-defined) left-to-right order of the active intervals. Also we maintain an order of the (active and regular) intervals such that an interval's color depends only on the color of one or two intervals' that are later in this order. This order can be easily maintained as in each step the new interval and the new active intervals come at the end of the order. We also save for each interval the one or two intervals which it depends on. This can be imagined as the intervals represented by vertices on the horizontal line arranged according to this order and an acyclic directed graph on them representing the dependency relations, thus each edge goes backwards and each vertex has indegree at most two (at most one in the first algorithm, i.e. the graph is a directed forest in that case). In each step we have to update the order of active intervals and the acyclic graph of all the intervals, this can be done in $c \log n$ time plus the time needed for the deletion of intervals from the order. Although the latter can be linear in a step, yet altogether during the whole process it remains cn, which is still ok. At the end we just color the vertices one by one from right to left following the rules, which again takes only cn time. Altogether this is $cn \log n$ time. □

These problems are equivalent to (offline) colorings of bottomless rectangles in the plane. Using this notation, Theorem 4.13 and Theorem 4.12 were proved already in [43] and [42], yet those proofs are quite involved and they only give quadratic time algorithms, so these results are improvements regarding simplicity of proofs and efficiency of the algorithms. The algorithms in [43] and [42] proceed with the intervals in backwards order and the intervals are colored immediately (in each step many of them are also recolored), this might be a reason why a lot of recolorings are needed there (which we don't need in the above proofs), adding up to quadratic time algorithms (contrasting the near-linear time algorithms above).

4.5 Dynamic interval hypergraphs

In this section, based on parts from [49], we investigate two-coloring geometric *dynamic hypergraphs* defined by intervals on a line. The vertices of a dynamic hypergraph are ordered and they "appear" in this order. Knowing in advance the whole ordered hypergraph, our goal is to color the whole vertex set such that at all times any edge restricted to the vertices that have "arrived so far" is non-monochromatic if it contains at least m vertices that have arrived so far. The exact definitions are as follows.

Definition 4.15. *For a hypergraph $\mathcal{H}(V,\mathcal{E})$ with an ordered set of vertices, $V = \{v_1, v_2, \dots, v_n\}$, we define the* dynamic closure *of $\mathcal{H}$ as the hypergraph on the same vertex set and with edge set $\{E \cap \{v_1, v_2, \dots, v_i\} : E \in \mathcal{E}, 1 \le i \le |V|\}$. A hypergraph with an order on its vertices is* dynamic *if it is its own dynamic closure. A hypergraph is m-proper two-colorable if V can be two-colored such that for every i and $E \in \mathcal{E}$ if $|E| \ge m$, then E contains both colors. For a family of (ordered) hypergraphs, $\{\mathcal{H}_i \mid i \in I\}$, define $m(\{\mathcal{H}_i \mid i \in I\})$ as the smallest number m such that every (ordered) hypergraph in the family is m-proper two-colorable.*

Let $m_{oct} = m(\text{Point2Octant})$, where Point2Octant is the family of hypergraphs realizable by octants. From Theorem 2.24 we know that $5 \le m_{oct} \le 9$. Let DPoint2Quadrant be the family of the dynamic closures of ordered hypergraphs on ordered finite planar point sets where a subset is an edge if and only if there is a quadrant containing exactly this subset of the point set. As we noted already in the previous section, in [45, 63] it was also shown (not using this terminology) that the hypergraph family Point2Octant is the same as the (ordered) hypergraph family DPoint2Quadrant (regarding the hypergraphs in it without the vertex orders). Summarizing:

Observation 4.16 ([45, 63])**.** *DPoint2Quadrant equals Point2Octant and therefore* $m(\text{Point2Octant}) = m(\text{DPoint2Quadrant}) = m_{oct}$.

The set of all intervals on the real line is denoted by $\mathcal{I}_{\mathbb{R}}$. Note that we are dealing with finitely many objects, so it does not matter if the intervals are closed or open. We study the following five hypergraph families and their dynamic closures defined by points and intervals on the real line. (For their relations that we will establish, see Figure 16.)

Point2Int: Vertices: a finite point set;
Edges: subsets of the vertex points contained in an interval $I \in \mathcal{I}_{\mathbb{R}}$.

Int2Point: Vertices: a finite set of intervals;
Edges: subsets of the vertex intervals containing a point $p \in \mathbb{R}$.

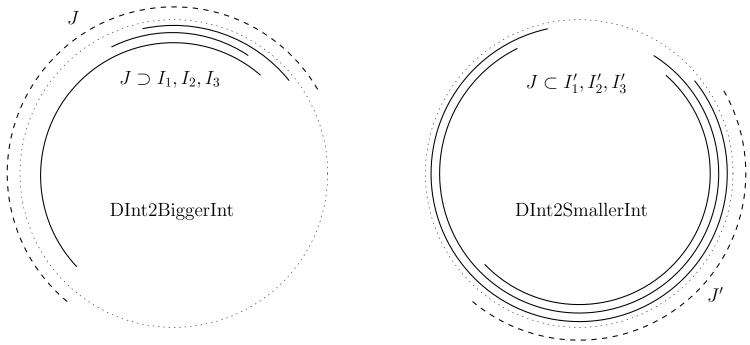


Figure 14: Int2BiggerInt equals Int2SmallerInt.

Int2BiggerInt: Vertices: a finite set of intervals;
Edges: subsets of the vertex intervals contained in an interval $I \in \mathcal{I}_{\mathbb{R}}$.

Int2SmallerInt: Vertices: a finite set of intervals;
Edges: subsets of the vertex intervals containing an interval $I \in \mathcal{I}_{\mathbb{R}}$.

Int2CrossInt: Vertices: a finite set of intervals;
Edges: subsets of the vertex intervals intersecting an interval $I \in \mathcal{I}_{\mathbb{R}}$.

DH: When H is a hypergraph family, DH is the hypergraph family that contains all the dynamic closures of the family H (with all orderings of their vertex sets).

Observation 4.17. *If for two non-ordered hypergraph families, A is a subfamily of B, then for their dynamic closures, DA is a subfamily of DB. Thus, if A and B are equal, then DA and DB are also equal.*

Now we study the relations among the above five hypergraph families. By exchanging points with small enough intervals we get that the family Point2Int is a subfamily of Int2BiggerInt, and the family Int2Point is a subfamily of Int2SmallerInt, while both Point2Int and Int2Point are subfamilies of Int2CrossInt. By definition and using Observation 4.17, this implies, e.g., $m(\text{DPoint2Int}) \le m(\text{DInt2BiggerInt})$.

We are only aware of earlier papers studying the first two variants. It follows from a greedy algorithm that $m(\text{Point2Int}) = 2$ and $m(\text{Int2Point}) = 2$. It was shown in [43] that $m(\text{DPoint2Int}) = 4$, and later this was generalized for k-colors in [7]. It was also shown in [43] that $m(\text{DInt2Point}) = 3$, and later this proof was simplified in [44]. It is interesting to note that for the DPoint2Int m-proper coloring problem there is a so-called *semi-online* algorithm, that can maintain an appropriate partial m-proper coloring of the points arrived so far, while it was shown in [17] that no semi-online algorithm can exist for m-proper coloring DInt2Point. Here we mainly study the other three hypergraph families.

Proposition 4.18. *Int2BiggerInt equals Int2SmallerInt and DInt2BiggerInt equals DInt2SmallerInt.*

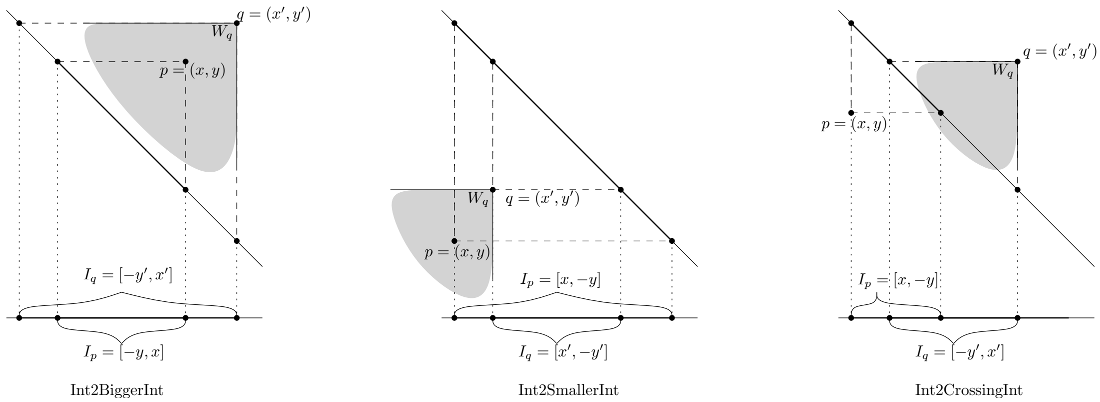


Figure 15: Left: Int2BiggerInt equals Point2Quadrant, Center: Int2SmallerInt equals Point2Quadrant, Right: Int2CrossingInt is a subfamily of Point2Quadrant.

Proof. By Observation 4.17 it is enough to prove the first statement. Notice that in both Int2BiggerInt and Int2SmallerInt, we can suppose that the left endpoint of any vertex interval is to the left of the right endpoint of any vertex interval, as swapping a right endpoint with a left endpoint which is next to (and to the right to) it does not change the hypergraphs. Thus, without loss of generality, there is a point that is in all the vertex intervals. Instead of a line, imagine that the vertex intervals of a hypergraph of Int2BiggerInt are the arcs of a circle such that none of them contains the bottommost point of the circle and all of them contains the topmost point.[11] This is clearly equivalent to the version when the vertex intervals are on the line. Similarly, we can imagine that the vertex intervals of a hypergraph of Int2SmallerInt are the arcs of a circle such that none of them contains the topmost point of the circle and all of them contains the bottommost point. Taking the complement of each arc transforms the families into each other, see Figure 14. □

Lemma 4.19. *Int2BiggerInt and Int2SmallerInt are both equal to Point2Quadrant, while Int2CrossingInt is a family of subhypergraphs of hypergraphs from the above, and the same holds for the dynamic variants.*

Proof. By Observation 4.17, it is enough to prove the statements about the non-dynamic families. For an illustration for the proof, see Figure 15. Recall that a quadrant is a set of the form $(-\infty, x) \times (-\infty, y)$ for some apex (x, y). We can suppose that all points of the point set are in the North-Eastern halfplane above the line ℓ defined by the function $x + y = 0$, i.e., $x + y > 0$ for every $p = (x, y)$. For each point $p = (x, y)$ we define an interval, $I_p = [-y, x]$. Quadrants that lie entirely below ℓ do not contain points from P. For the quadrants with apex above ℓ, a quadrant whose apex is at q contains the point p if and only if I_q contains I_p. This shows that the hypergraphs in Point2Quadrant and in Int2BiggerInt are the same.

[11]Without the extra condition regarding the bottommost point, we could define a circular variant of the problem whose parameter m can be at most one larger than m(DInt2BiggerInt) but we omit discussing this here.

The equivalence of Int2SmallerInt and Point2Quadrant already follows from Proposition 4.18, but we could give another proof in the above spirit, by supposing that for all points $p = (x, y)$ we have $x + y < 0$, moreover, that for every quadrant intersecting some of the points there is a quadrant containing the same set of points whose apex $q = (x', y')$ has $x' + y' < 0$. Now for each point $p = (x, y)$ we can define the interval $I_p = [x, -y]$ and proceed as before. Note that this gives another proof for Proposition 4.18.

Finally, taking a H in Int2CrossingInt, it is isomorphic to the subhypergraph of some H' in Point2Quadrant where in H for all points $p = (x, y)$ we have $x + y < 0$ and we take only the edges corresponding to quadrants whose apex $q = (x', y')$ has $x' + y' > 0$. Now for each point $p = (x, y)$ below ℓ we define $I_p = [x, -y]$, and for each point $q = (x', y')$ above ℓ we define $I_q = [-y', x']$, and proceed as before. This finishes the proof of the theorem. □

As it was shown in [69] that $m(\text{Point2Quadrant}) = 2$, and so it follows that we also have $m(\text{Int2BiggerInt}) = m(\text{Int2SmallerInt}) = m(\text{Int2CrossingInt}) = 2$.

Surprisingly, we could not find a direct proof for the fact that DInt2CrossingInt is a family of subhypergraphs of hypergraphs from the families DInt2BiggerInt and DInt2SmallerInt.

From Theorems 3.1 and 3.2, and Lemma 4.19 we obtain the following.

Corollary 4.20. $5 \le m(\text{DInt2BiggerInt}) = m(\text{DInt2SmallerInt}) = m_{oct} \le 9$.

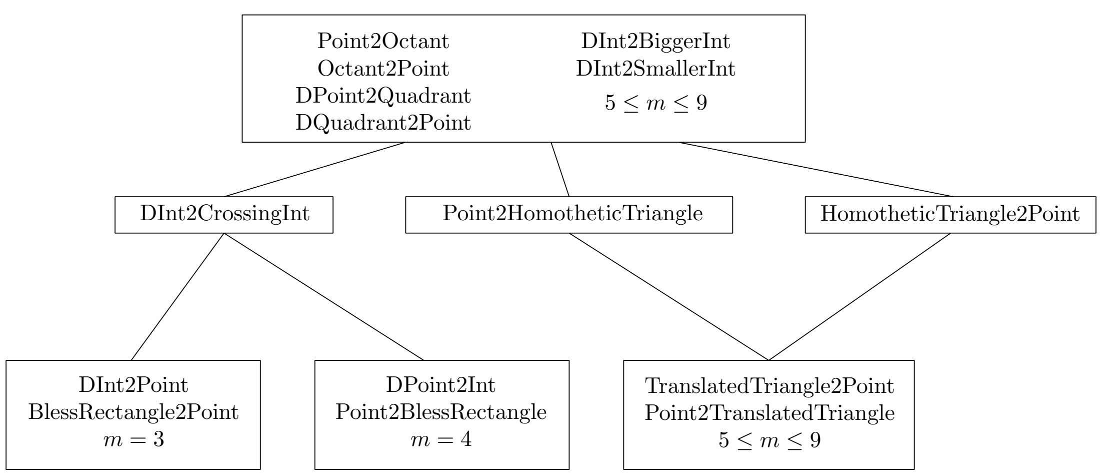


Figure 16: Diagram of proved and related hypergraph results.

5 Self-coverability

This chapter is based on the article [47], joint work with Keszegh.

Definition 5.1. *A collection of closed geometric sets $\mathcal{S}$ is* self-coverable *if there exists a* self-coverability function *f such that for any $S \in \mathcal{S}$ and for any finite point set $P \subset S, |P| = k$ there exists a subcollection $\mathcal{S}' \subset \mathcal{S}$, $|\mathcal{S}'| \le f(k)$ such that $\cup \mathcal{S}' = S$ but no point of P is in the interior of an $S' \in \mathcal{S}'$.*

Note that by definition all points of P are outside or on the boundary of regions from $\mathcal{S}'$. Also, points outside or on the boundary of S are irrelevant, thus we can and will assume that all points of P are in the interior of S.

E.g., it is easy to see that (closed)[12] axis-parallel rectangles are self-coverable with $f(k) = k+1$ and that all discs in the plane (or, in fact, the homothets of any set that is a concave polygon or a set with a smooth boundary) are not self-coverable as already $f(1)$ does not exist.

The motivation to study this notion is the following theorem, which is a generalization of a result contained implicitly in Cardinal et al. [16].

Theorem 5.2. *Suppose that $\mathcal{S}$ is self-coverable with a monotone self-coverability function f for which $f(k) > k$, and that $m_2 \le m$ for $\mathcal{S}$, i.e., any finite set of points can be colored with two colors such that if $S \in \mathcal{S}$ contains at least m points, then S contains both colors. Then $m_k \le m(f(m-1))^{\lceil \log k \rceil - 1}$, i.e., any finite set of points can be colored with k colors such that if $S \in \mathcal{S}$ contains at least $m(f(m-1))^{\lceil \log k \rceil - 1} \le k^d$ points, then S contains all k colors. (Here d is a constant that depends only on $\mathcal{S}$.)*

Our main results about self-coverability are about homothets of convex polygons where we prove the following.

Theorem 5.3. *The family of all homothets of a given convex polygon C is self-coverable with $f(k) \le ck$ where the constant c depends only on C.*

In other words, given a closed convex polygon C and a collection of k points in its interior, we can take ck homothets of C whose union is C such that none of the homothets contains any of the given points in its interior.

Corollary 5.4. *If for the homothets of any convex polygon $m_2 < \infty$, then $m_k \le k^d$.*

For triangles and squares we could even determine the exact value of f.

Theorem 5.5. *The family of all homothets of a given triangle is self-coverable with $f(k) = 2k+1$ and this is sharp.*

Theorem 5.6. *The family of all homothets of a square is self-coverable with $f(k) = 2k+2$ and this is sharp.*

We also show that the constant in Theorem 5.3 cannot depend only on the number of vertices of P as even for a quadrangle it can be arbitrarily big.

Theorem 5.7. *For every c there exists a quadrangle Q such that the family of all homothets of Q is self-coverable with f, then $f(k) \ge ck$.*

In the rest of this chapter, we prove Theorem 5.2 in Section 5.1, then Theorem 2.3 and 5.6 in Section 5.2 and finally Theorem 5.3 and 5.7 in Section 5.3.

[12] Every polygon is considered to be closed in this chapter, unless stated otherwise.

5.1 Proof of Theorem 5.2

Proof of Theorem 5.2. Suppose that $\mathcal{S}$ is self-coverable with self-coverability function f and $m_2 \le m$, i.e., any finite set of points P can be colored with two colors such that any member of $\mathcal{S}$ with at least m points contains both colors. Now we show by induction on k that any finite set of points can be colored with k colors such that any member of $\mathcal{S}$ with at least $m_k = m(f(m-1))^{\lceil \log k \rceil - 1}$ points contains all k colors.

Suppose we already know the statement for a and b, from this we establish it for $k = ab$. Color P with a colors using induction and denote the color classes by $P_1, \ldots, P_a$. Now color each of these color classes with b colors using induction. We claim that this coloring is good for $k = ab$. By contradiction, say $S \in \mathcal{S}$ does not contain all colors. This means that for some $1 \le i \le a$ we have $|S \cap P_i| \le m_b - 1$. Using the self-coverability of $\mathcal{S}$, cover $S \setminus P_i$ with $f(m_b-1)$ sets of $\mathcal{S}$. Using the monotonicity of f, one of these covering sets contains at least $\lceil \frac{m_k - m_b + 1}{f(m_b-1)} \rceil$ points of P but no points of P_i. This contradicts that our coloring with a colors was good if $\lceil \frac{m_k - m_b + 1}{f(m_b-1)} \rceil \ge m_a$.

Using the above argument for $b = 2$, we can see that $m_k = m(f(m-1))^{\lceil \log k \rceil - 1}$ satisfies $\lceil \frac{m_k - m + 1}{f(m-1)} \rceil = m_{\lceil \frac{k}{2} \rceil}$, thus we are done (using the monotonicity of m_k if k is odd). □

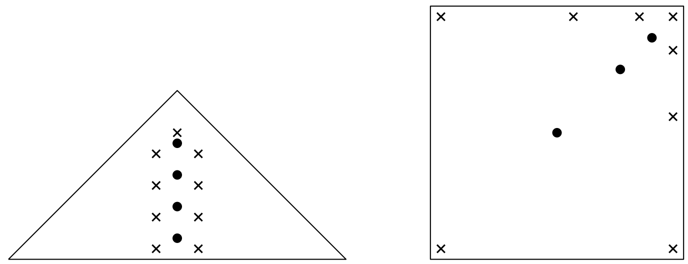
5.2 Self-coverability of triangles and squares

Proof of Theorem 5.5. We now prove that for the family $\mathcal{T}$ of homothets of a given triangle T we have $f(k) = 2k + 1$. First by affine transformations we can transform the triangle to any other triangle, thus it is enough to prove the statement for one triangle. Further, by homothetic symmetry it is enough to prove self-coverability of one fixed size triangle T_0. Thus we can assume that T has the three vertices $(0,0), (2,0), (1,1)$.

First we begin by giving a set P of k points for which $2k + 1$ triangles are indeed needed to cover T_0. Let the set of points be on a vertical line passing through the vertex $(1,1)$, i.e. all points of P have coordinates $(1,y); 0 < y < 1$. Let ϵ be a small positive constant and for each point $(1,y)$ of P assign two dummy points $(1-2\epsilon, y-\epsilon)$ and $(1+2\epsilon, y-\epsilon)$ inside T_0 (i.e., we choose ϵ such that $\epsilon < 1-y$ for all points of P). Put an additional dummy point at coordinate $(1, y_{max} + \epsilon)$ above the highest point $(1, y_{max})$ of P. It is easy to see that if ϵ is small enough then any triangle from $\mathcal{T}$ contained in T_0 and not containing a point of P in its interior can cover at most one dummy point in its interior. As to cover T_0 in particular we need to cover all dummy points, so we need at least $2k+1$ triangles. See Figure 17(a) for an illustration.

Now it is enough to prove that at most $2k+1$ triangles are always enough to cover T_0. We prove this by induction on k. For $k = 0$ we can cover T_0 by itself. If $k \ge 1$, then take the[13] point $p \in P$ with the smallest y-coordinate and denote it by $y(p)$. Denote by H_y the halfplane with the horizontal line $y = y(p)$ as its boundary containing an infinite positive ray on the y-axis. Apply induction on $P \setminus p$ and the triangle $T_1 = T_0 \cap H_{y(p)}$.

[13]For simplicity we suppose that there is only one such point p yet the proof can be easily modified to the case when there are multiple points with the same y-coordinate, in which case we have to handle all these points in one step of the induction.



(a) (b)

Figure 17: Lower bound constructions for self-covering (a) a triangle and (b) a square.

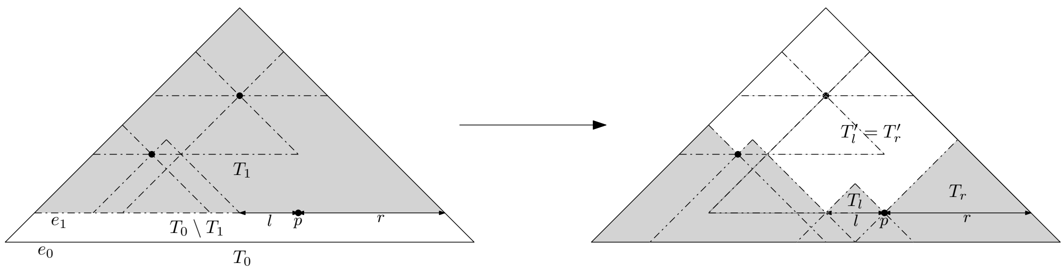


Figure 18: Extending a covering of T_1 and adding the two triangles T_ℓ and T_r.

See Figure 18 for an illustration. We get a collection $\mathcal{S}_1$ of at most $2k-1$ triangles (homothetic to T_0) covering T_1. Denote by $\mathcal{S}_a$ those triangles from $\mathcal{S}_1$ whose bottom edge e is on the bottom edge e_1 of T_1 but e does not contain p in its interior (thus p can be a vertex of such a triangle).

Now we can scale all triangles in $\mathcal{S}_a$, their top vertex as the center of scaling, so that their bottom edge goes onto the bottom edge e_0 of T_0. Our new collection will consist of these scaled triangles and $\mathcal{S}_1 \setminus \mathcal{S}_a$. The points of e_1 not covered by the scaled triangles form an interval and p cuts this interval into a left interval ℓ and a right interval r. Now the triangle T_ℓ is (well) defined to be the triangle which intersects e_1 in exactly ℓ and has its bottom edge on e_0. Similarly, T_r is (well) defined to be the triangle which intersects e_1 in exactly r and has its bottom edge on e_0. We claim that these two triangles do not contain a point from P in their interior. Indeed, first of all there are no points of P under $H_{y(p)}$. Second, in the inductive construction there must be a triangle T'_ℓ whose bottom edge contains the whole ℓ, thus T'_ℓ contains no point of P in its interior and $T_\ell \cap H_{y(p)} \subseteq T'_1$. The interior of T_r is similarly disjoint from P.

We have seen that none of the triangles in this new collection of at most $2k-1+2 = 2k+1$ triangles contains a point of P in its interior. Now we finish the proof by showing that this collection of triangles covers T_0. T_1 is trivially covered by induction. For an arbitrary point $q \in T_0 \setminus T_1$ at least one of the two diagonal lines (these are the lines parallel to the non-horizontal edges of T) across q intersects e_1 in a point q'. If q' is on ℓ (or respectively r), then q is covered by T_ℓ (or respectively by T_r). If none of these happens then q is covered by one of the scaled triangles. $\square$

Proof of Theorem 5.6. First we begin by giving a set P of k points for which $2k+2$ squares are indeed needed to cover a square $R = [0,1] \times [0,1]$ if $k \geq 1$. The points are

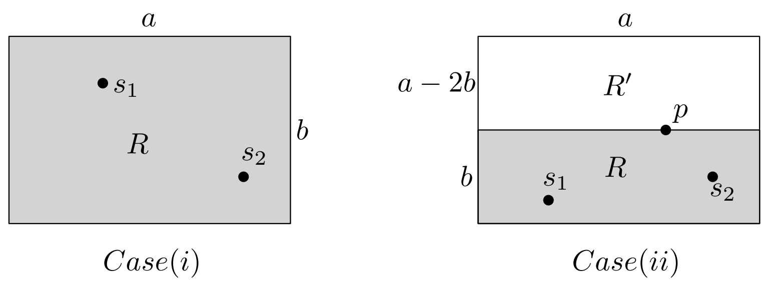


Figure 19: the rectangle R to be covered by Lemma 5.8, in Case (ii) the squares can cover parts of $R' \setminus R$ as well, but are not allowed to contain p in their interior.

on one of the diagonals of R, the ith point has coordinates $(1-1/2^i, 1-1/2^i)$. Let ϵ be a small positive constant and for each point $(1-1/2^i, 1-1/2^i)$ of P assign two dummy points $(1-1/2^{(i-1)}+i\epsilon, 1-\epsilon)$ and $(1-\epsilon, 1-1/2^{(i-1)}+i\epsilon)$ inside R. Put an additional dummy point at coordinate (ϵ,ϵ) and $(1-\epsilon, 1-\epsilon)$. It is easy to see that if ϵ is small enough ($\epsilon < 1/2^k(k+1)$ is sufficient) then any square contained in R and not containing a point of P in its interior can cover at most one dummy point. As to cover R in particular we need to cover all dummy points, so we need at least $2k+2$ squares. See Figure 17(b) for an illustration.

Now it is enough to prove that at most $2k+2$ squares are always enough to cover a square. We again proceed by induction but we need a more general statement, Lemma 5.8. The theorem follows from this lemma by taking R to be a square.

The following lemma states that if the ratio of the two sides of an axis-parallel rectangle R is at most 2 then it can be covered by $2k+2$ axis-parallel squares (while not covering the point set of size k), whereas if the ratio of the sides is bigger, then we can cover R such that the squares may hang out over the top edge of R but only until a limited height, and not covering an additional fixed point p on the top edge.

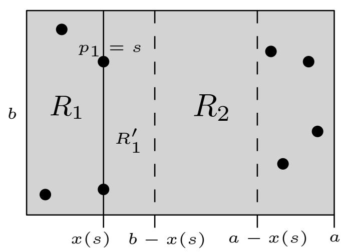
Lemma 5.8. *Given an axis-parallel rectangle R with width a, height $b \le a$ and a point set $P \subset R$, $|P| = k$ and a point p on the top edge of R, there is a collection $\mathcal{R}$ of at most $2k+2$ axis-parallel squares covering R, none of them containing a point from $P \cup \{p\}$ in their interior, such that*

(i) if $a/2 \le b$ then $\cup\mathcal{R} = R$,

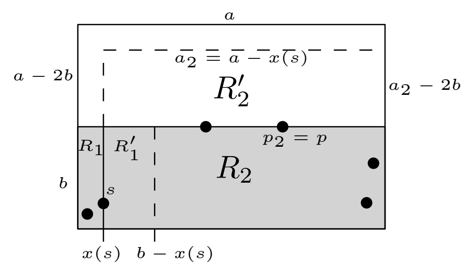
(ii) if $b < a/2$ then $R \subseteq \cup\mathcal{R} \subseteq R \cup R'$, where R' is a rectangle whose bottom side is the top side of R and its height is $b' = a - 2b > 0$.

Note that p is not in P. Also note that in the first case points of P on the boundary of R do not matter while in the second case points of P on the top edge of R are not irrelevant and can modify the choice of squares. See Figure 19 for an illustration.

Proof. We can suppose that the bottom left corner of R is the origin $(0,0)$. We prove the two cases simultaneously by induction on k. Both cases will be quite similar, we always cut the rectangle through some point of P into two as equal parts as possible and apply induction on both parts. We denote the x- and y-coordinate of a point s by $x(s)$ and by $y(s)$, respectively. For a rectangle Q we denote by $intQ$ its interior.



(a) Case (i), applying induction on the two parts of R, Case (ii) on R_1 and Case (i) on R_2.



(b) Case (ii), applying induction on the two parts of R, Case (ii) on R_1 and Case (ii) on R_2.

Figure 20: Proof of Lemma 5.8

If $k = 0$ then in Case (i) it is trivial to cover R using two squares. In Case (ii) we can suppose, without loss of generality, that $x(p) \geq a/2$. Now put a square of height $\min(a-b, x(p))$ in the bottom left corner of R and a square of height $max(b, a - x(p))$ in the bottom right corner of R, so by definition these do not contain p in their interior. It is easy to check that they cover R and are contained in R'.

Next suppose $k > 0$ and we are in Case (i). If there exists $s \in P$ such that $b/2 \leq x(s) \leq a - b/2$ then cut the rectangle R into two parts R_1, R_2 by a vertical line through s and then by induction (Case (i)) we can find squares covering R_1 and R_2, together they cover R and the number of the squares is at most $2k_1 + 2 + 2k_2 + 2 = 2(k_1 + k_2 + 1) + 2 \leq 2k + 2$, where $k_1 = |P \cap intR_1|, k_2 = |P \cap intR_2|$.

If there is no such s then choose s to be the point of P which is closest to the vertical halving line of R, i.e. for which $|x(s) - a/2|$ is minimal. Without loss of generality, we can suppose that $x(s) < a/2$ and thus we also know that $x(s) < b/2$. We again cut by the vertical line through s. To get a covering of the right rectangle R_2 we can apply the induction hypothesis with Case (i). For the left rectangle R_1 we apply the induction hypothesis with Case (ii) by setting $p_1 := s$, $R_1 := R_1$ and R_1' being the part of R between the vertical lines at x-coordinate $x(s)$ and $b - x(s)$. The two set of squares together cover R and as $b - x(s) < b < a$ implies $R_1' \subset R$, they do not hang out from R. We need to check if the covering of R_1 does not interfere with the points in $P \cap intR_1$. This is true if $intR_1'$ does not contain points from P, which follows from the fact that there is no point of P with x-coordinate between $x(s)$ and $a - x(s)$ and the right edge of R_1' has x-coordinate $b - x(s) < a - x(s)$. Finally, the number of squares we used is again at most $2k_1 + 2 + 2k_2 + 2 = 2(k_1 + k_2 + 1) + 2 \leq 2k + 2$, where $k_1 = |P \cap intR_1|, k_2 = |(P \cap R_2) \setminus \{s\}|$. See Figure 20(a) for an illustration.

Suppose now that $k > 0$ and we are in Case (ii). Similarly to the previous case, if there exists $s \in P$ such that $b/2 \leq x(s) \leq a - b/2$ then cut the rectangle R into two parts R_1, R_2 by a vertical line through s and then by induction (using Case (i) or Case(ii), see details below) we can find squares covering R_1 and R_2, together they cover R and the number of the squares is at most $2k_1 + 2 + 2k_2 + 2 = 2(k_1 + k_2 + 1) + 2 \leq 2k + 2$, where $k_1 = |P \cap intR_1|, k_2 = |P \cap intR_2|$. The induction is done in the following way. We just consider R_2, R_1 can be handled in the same way. If the ratio of the two sides of the rectangle is at most 2, then we can simply apply induction Case (i). If the ratio of the sides is bigger than 2, i.e., $b < (a - x(s))/2$, then we apply induction Case (ii) with $p_2 = p$ if p is on the top edge of R_2 and choosing an arbitrary p_2 on the top edge of R_2 if p is not on the top edge of R_2.

If there is no such s then choose s to be the point of P which is closest to the vertical halving line of R, i.e. for which $|x(s) - a/2|$ is minimal. Without loss of generality, we can suppose that $x(s) < a/2$ and thus we also know that $x(s) < b/2$. We again cut by the vertical line through s. Now in the same way as in Case (i) we can apply induction Case (ii) on the left part R_1. It is easy to see that by the choice of s, R_1' corresponding to R_1 is contained by R and does not contain points from P in its interior. On the right part R_2, again, if the ratio of the two sides b and $a - x(s)$ is at most 2, then we can simply apply induction Case (i). If the ratio of the sides of R_2 is bigger than 2, i.e., $b < (a - x(s))/2$, then we apply induction Case (ii) on R_2 with $p_2 = p$ if p is on the top edge of R_2 and choosing an arbitrary p_2 on the top edge of R_2 if p is not on the top edge of R_2. It is easy to see that the rectangle R_2' corresponding to R_2 is contained in R'. Thus, the two set of squares we get by induction again cover the whole R, are contained in $R \cup R'$, none of the squares contains p in its interior and the number of squares is at most $2k_1 + 2 + 2k_2 + 2 = 2(k_1 + k_2 + 1) + 2 \leq 2k + 2$, where $k_1 = |P \cap R_1|, k_2 = |P \cap (R_2 \setminus R_1)|$. See Figure 20(b) for an illustration of this case. $\square$

5.3 Self-coverability of convex polygons using Delaunay triangulation

In this section we prove Theorem 5.3. Since the proof is a bit complicated, to illustrate it, first we will reprove Theorem 5.5, then Theorem 5.6 (with a worse self-coverability function) and only then prove Theorem 5.3 in its full generality.

The proof uses the notion of generalized Delaunay triangulation defined in Section 1.2.2.

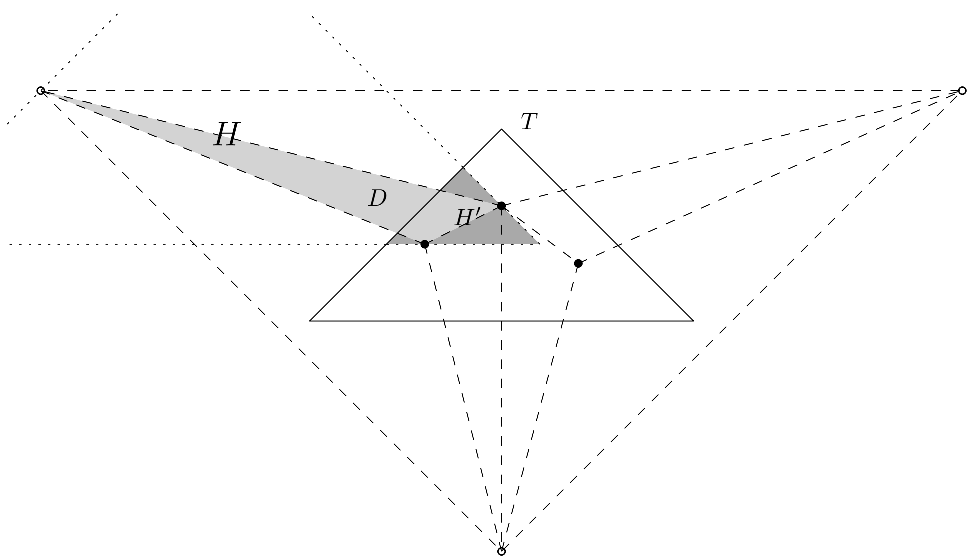


Figure 21: Proof of Theorem 5.5

Second proof of Theorem 5.5. (See Figure 21.) We add three new points to P which are far, outside of T, and form a reflected copy of T. Denote the new point set by P'. In the Delaunay triangulation determined by T, these three points will be all connected, making all the faces triangles. Using Euler's formula, there are $k+3$ vertices and thus $2(k+3)-4$ faces, so we have $2k+1$ inner faces, all of which can be covered by a homothet of T not containing any point of P' in its interior.

The only problem is that these homothets might extend beyond the boundary of T. But it is easy to see that for any homothet H of T the triangle $H' = H \cap T$ is also a homothet of T, so these give at most $2k+1$ covering triangles. $\square$

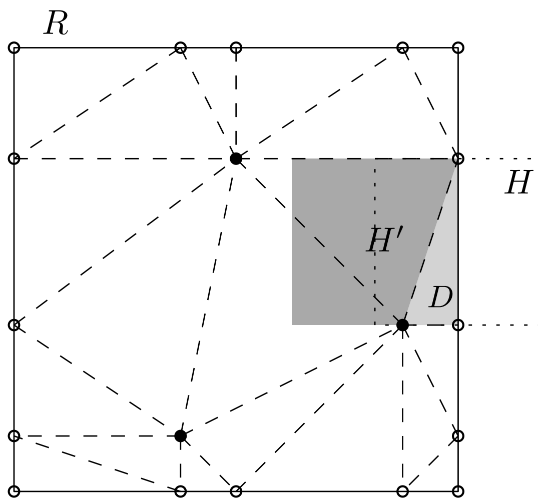


Figure 22: Proof of Theorem 5.6

Second proof of Theorem 5.6. with worse self-coverability function. (See Figure 22.) Similarly to the proof for triangles, we add a few points to P and we denote the new point set by P'. Now all new points will be on the boundary of the square R. The new points are obtained as follows. For each $p \in P$ project it orthogonally to all four sides of R and add it to P'. Also we add the four corners of R to P'. Thus $|P'| = 5k+4$ if all vertices of P have different coordinates, which we suppose from now on for simplicity. In the Delaunay triangulation determined by R, all boundary points will be connected to their neighbors (on the boundary), making all inner faces triangles. The infinite outer face has $4k+4$ vertices. Using Euler's formula, there are $5k+4$ vertices and thus $2 \cdot (5k+4) - 4 - (4k+1) = 6k+3$ faces. The $6k+2$ triangular inner faces partition R and all of them can be covered by a homothet of R not containing any point of P' in its interior.

Again, the problem is that these homothets might extend beyond the boundary of R. To take care of this, observe first that each square H that extends beyond R intersects only one side s of R. For each such H, push H perpendicularly to s until its outer side overlaps with s, call the pushed square H'. This way no point $p \in P$ can get into the interior of H' since then the projection of p to s would have been also inside H. So we get $f(k) \leq 6k+2$. $\square$

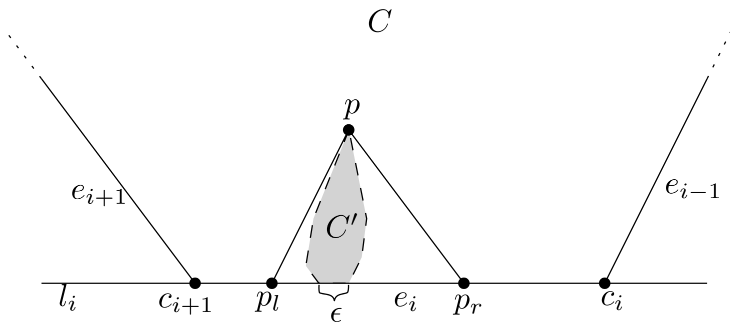


Figure 23: Proof of Theorem 5.3

Proof of Theorem 5.3. Let C be an arbitrary convex polygon. Denote its vertices in clockwise order by $c_0c_1\dots c_{n-1}$ and its sides by $e_i = c_ic_{i+1}$. We will again add some points to P to define P' and take the Delaunay triangulation of P' with respect to C. All the added points lie on the boundary or outside of C and their positions depend on the point set P as follows.

For each $p \in P$ and side e_i (indices are always modulo n) of C we do the following. First draw two lines through p such that the first is parallel to e_{i-1} and the second is parallel to e_{i+1}. These intersect the supporting line ℓ_i of e_i in two points, p_ℓ and p_r. (See Figure 23 for an illustration.) Any homothet C' of C that intersects ℓ_i and has p on its boundary contains a point of the $\overline{p_\ell p_r}$ segment. (Here we allow C' to contain more points in its interior.) Take a C' for which the length of the intersection of (the closure of) C' and ℓ_i is minimal and denote this minimum by ϵ. It is easy to see that ϵ is well-defined and positive.

Also observe that $|p_\ell p_r|/\epsilon$ depends only on C and i and is independent of the position of p, since translating p parallel to e_i only shifts the $pp_\ell p_r$ triangle with the same quantity, while moving p perpendicularly to e_i only scales $pp_\ell p_r$, thus scaling both ϵ and $|p_\ell p_r|$ by the same value.

Now put $N_i = \lfloor |p_\ell p_r|/\epsilon \rfloor$ evenly spaced points on the segment $\overline{p_\ell p_r}$, so that the distance between any two of them is less than ϵ. Moreover, add one point very close to p_ℓ and another very close to p_r onto ℓ_i but not on $\overline{p_\ell p_r}$, such that the distances from them to the next point is still less than ϵ. We add these $N_i + 2$ points to P'. Since the distance between any two of them is less than ϵ, any homothet of C with p on its boundary and intersecting ℓ_i contains one of the just added $N_i + 2$ points in its interior by the definition of p_ℓ, p_r and ϵ. The number of points we added depends only on C and i, not on p. Repeating this for all points and edges of C, we add at most $k\sum_{i=1}^{n}(N_i + 2) = O(k)$ points. (Here the constant in the $O(.)$ notation depends on C.)

For each vertex c_i of C we add 3 points to P'. First put a small homothetic copy C'_i touching C from inside whose respective vertex $c'_i = c_i$ and add c_i and the two neighboring vertices c'_{i-1} and c'_{i+1} of C'_i to P'.

To ensure that the outerface would not intersect the interior of C, we add n more points to P' which are far, outside of C, and form a reflected copy of C. This way these n points will be the vertices of the outerface.

In the last two stages we added $4n = O(1)$ vertices.

In the Delaunay triangulation of P' the inner faces are triangles, with a homothet of C covering each triangle not containing any point of P' in its interior. We claim that there is no Delaunay triangle whose side crosses the boundary of C and thus the triangles inside C partition C (this way Delaunay edges form a cycle along the boundary of C). Furthermore, we claim that the homothetic copies covering the triangles inside C are contained in C.

Consider any Delaunay triangle T. By Claim 1.9, if T intersects two different sides of C, the homothet D covering T must also contain one of C's vertices, c_i on its boundary and thus we necessarily have $C'_i \subset D$ and as the vertices neighboring c_i in C'_i were added to P', D must be completely inside C. If T intersects only one side of C and the interior of C, then it must contain a point p from P. The homothet D covering T must also contain p on its boundary and thus it would also contain a point of P' in its interior if it crossed the boundary of C.

Using Euler's formula, there are $O(k)$ vertices and thus at most $O(k)$ faces. Each face inside C is a triangle which can be covered by a homothet of C which does not intersect the boundary of C and does not contain any point of P' in its interior. $\square$

Finally, we prove Theorem 5.7, which states that even for a quadrangle we may need many points. For that we basically prove that while by the above upper bound states that at most $k\sum_{i=1}^{n}(N_i+5)$ copies are enough to cover C, also at least $k \min_i N_i$ copies are necessary to cover C.

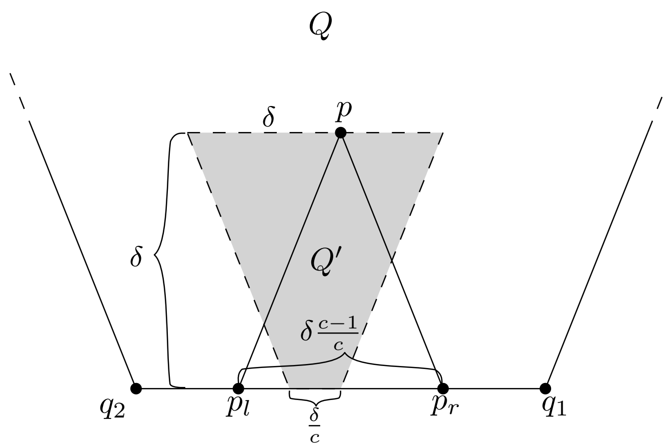


Figure 24: Proof of Theorem 5.7

Proof of Theorem 5.7. Given $c > 1$, let Q be a symmetric trapezoid with vertices q_1, q_2, q_3, q_4 in clockwise order, with two horizontal edges, the bottom edge, $\overline{q_1q_2}$ has length $1/c$, while the length of the top edge, $\overline{q_3q_4}$, and the height of Q are both equal to 1. We show that for Q we have $f(k) \geq (c-1)k$.

Put a point p very close to the bottom edge of Q, say the distance of p from the bottom edge is δ. We define p_ℓ and p_r as in the previous proof (see Figure 24). Evidently, in the self-cover of Q, the points of $p_\ell p_r$ can only be covered by homothetic copies of Q whose upper edge touches or is below p, thus have height at most δ. The length of the top edge of such a Q' is thus also at most δ and thus the length of its bottom edge is at most δ/c. We also know that $|p_\ell p_r| = \delta(1-1/c)$. Thus to cover the points of $p_\ell p_r$ we need at least $\frac{\delta(1-1/c)}{\delta/c} = c-1$ such homothets.

Now if instead of one, we put k points very close to the bottom edge, but far from each other, then for each point we need $c-1$ homothets to cover the respective segment on the boundary of Q, thus altogether we need at least $(c-1)k$ homothets, as claimed. $\square$

6 3-coloring for homothets of polygons

This chapter is based on the article [50], joint work with Keszegh.

In Section 6.1 we summarize the main idea of our proof.
In Section 6.2 we prove Theorem 2.16, using the proof method of [2].
In Section 6.3 we examine stabbed families, following [22].
In Section 6.4 we briefly discuss further related topics.

6.1 Framework

In this section we outline the main idea behind the proof of Theorem 2.16. Consider the generalized Delaunay triangulation $\mathcal{DT} = \mathcal{DT}(S, D)$. We will take an initial coloring of S that has some nice properties. More specifically, we need a 3-coloring for which the assumptions of the following lemma hold for $c = 3$ and for some constant t that only depends on D.

Lemma 6.1. *For every convex polygon D for every c and t there is an m such that if for a c-coloring of a point set S and a set of points $R \subset S$ and for every homothet D'*

(i) if $D' \cap S$ is monochromatic with at least t vertices, D' contains a point of R,

(ii) if D' contains t points from R colored with the same color, D' also contains a point from $S \setminus R$ that has the same color,

then there is a c-coloring of S such that no homothet that contains at least m points of S is monochromatic.

To prove Lemma 6.1, we use Theorem 5.3 about self-coverability of convex polygons.

Proof of Lemma 6.1. The proper c-coloring will be simply taking the c-coloring given in the hypothesis, and recoloring each vertex in R arbitrarily to a different color. Now we prove the correctness of this new coloring. Let D' be a homothet of D containing at least m points (where m is to be determined later).

Suppose first that D' contains $m \geq ct$ points from R. Using the pigeonhole principle, D' contains at least t points from R that originally had the same color. Using (ii), D' will have a point both in R and in $S \setminus R$ that had the same color. These points will have different colors after the recoloring, thus D' will not be monochromatic.

Otherwise, suppose that D' contains m points of which less than ct are from R. Apply Theorem 5.3 with D' and $R' = D' \cap R$. This gives $c_D ct$ homothets (where c_D comes from Theorem 5.3), each of which might contain at most three points on their boundaries (which include the points from R'), thus by the pigeonhole principle at least one homothet, D'', contains no points from R and at least $\frac{m-3c_D ct}{c_D ct}$ points from $S \setminus R$. If we set $m = c_D ct(t+3)$, this is at least t. Thus, by (i), D'' was not monochromatic before the recoloring. As the recoloring does not affect points in $S \setminus R$, after the recoloring D'' (and so also D') still contains two points that have different colors. Thus $m = c_D ct(t+3)$ is a good choice for m in both cases. $\square$

Therefore, to prove Theorem 2.16, we only need to show that we can find a coloring with three colors that satisfy the conditions of Lemma 6.1 for some t.

6.2 Proof of Theorem 2.16

In this section we prove Theorem 2.16, that is, we show that for every convex polygon D there is an m such that any finite set of points S admits a 3-coloring such that there is no monochromatic homothet of D that contains at least m points. If one could find a 3-coloring where every monochromatic component of $\mathcal{DT}$ is bounded, then that would immediately prove Theorem 2.16. This, however, is not true in general [52], only for bounded degree graphs [28], but the $\mathcal{DT}$ can have arbitrarily high degree vertices for any convex polygon, thus we cannot apply this result. Instead, we use the following result (whose proof is just a couple of pages).

Theorem 6.2 (Poh [70]; Goddard [38])**.** *The vertices of any planar graph can be* 3*-colored such that every monochromatic component is a path.*

To prove Theorem 2.16, apply Theorem 6.2 to $\mathcal{DT}$ to obtain a 3-coloring where every monochromatic component is a path. It follows from Lemma 6.1 that it is sufficient to show that for $t = 4n + 12$ (where n denotes the number of sides of D) there is a set of points $R \subset S$ for which

(i) for every homothet D' if $D' \cap S$ is monochromatic with at least t vertices, D' contains a point of R,

(ii) for every homothet D' if D' contains t points from R colored with the same color, D' also contains a point from $S \setminus R$ that has the same color.

Now we describe how to select R. First, partition every monochromatic path that has at least t vertices into subpaths, called *sections*, such that the number of vertices of each section is at least $\frac{t}{4}$ but at most $\frac{t}{2}$. We call such a section *cuttable* if there is a monochromatic homothet of D that contains all of its points. R will consist of exactly one point from each cuttable section. These points are selected arbitrarily from the non-extremal points of each section, except that they are required to be non-adjacent on their monochromatic path. Since each section has at most two end points and n extremal points, we can select such a point from each section if $\frac{t}{4} \geq n + 3$. For an $r \in R$ we denote its section by σ_r and a (fixed) monochromatic homothet containing σ_r by D_r.

Now we prove that R satisfies the requirements (i) and (ii).

To prove (i), suppose that a homothet D' is monochromatic with at least t vertices. Using Proposition 1.7, the subgraph induced on these vertices is connected. As any monochromatic connected component is a path, D' contains at least t consecutive vertices of a monochromatic path, and thus also a section. Because of D' this section is cuttable, and thus contains a point of R.

To prove (ii), suppose that a homothet D' contains t points from R colored with the same color, red. Denote these points by R'. For each $r \in R'$, the neighbors of r in σ_r are red but not in R, thus they must be outside D', or otherwise (ii) holds and we are done. Denote the geometric embedding of the two edges adjacent to r in σ_r by Λ_r. Therefore, Λ_r will intersect the boundary of D' in two points for each $r \in R'$. We claim that these two intersection points usually fall on the same side of D', i.e., they are not separated along the boundary by a vertex.

Proposition 6.3. *Both intersection points of Λ_r and the boundary of D' are on the same side of D' for all but at most n points of $r \in R'$.*

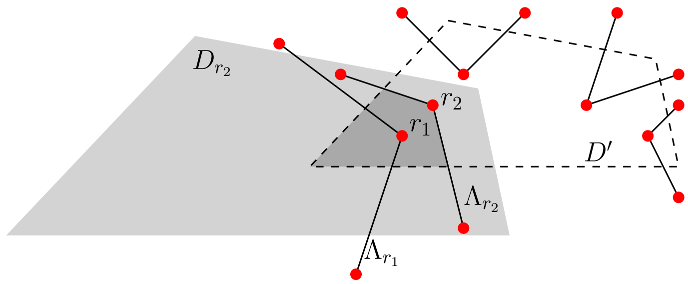


Figure 25: Proof of Proposition 6.3.

Proof. Suppose that there are more than n points $r \in R'$ for which Λ_r intersects D' in two sides. For each such point $r \in R'$, for (at least) one of the two (one convex and one non-convex) cones whose sides are the halflines starting in Λ_r, denoted by C_r, we have $C_r \cap D' \subset D_r$. Since the intersection $C_r \cap D'$ is a connected curve, it contains a vertex of D'. Using the pigeonhole principle, there are two points, $r_1, r_2 \in R'$, such that C_{r_1} and C_{r_2} contain the same vertex of D'. (See Figure 25.) As $\Lambda_{r_1} \cap \Lambda_{r_2} = \emptyset$, we have (without loss of generality) $r_1 \in C_{r_2}$, which also implies $r_1 \in D_{r_2}$. But using Proposition 1.7, r_1 must have a neighbor in D'. Since this neighbor is also in D_{r_2}, it has to be red. As the red neighbors of any red point of R are not in R, we have found a red point from $S \setminus R$ in D', proving (ii). □

Divide the points $r \in R'$ for which Λ_r intersects only one side of D' into n groups, $R'_1, \dots, R'_n$, depending on which side is intersected. By the pigeonhole principle there is a group, R'_i, that contains at least $\frac{t-n}{n} \geq 3$ points. Suppose, without loss of generality, that the side ab intersected by Λ_r for $r \in R'_i$ is horizontal, bounding D' from below. For each $r \in R'_i$, fix and denote by x_r a point from σ_r whose y-coordinate is larger than the y-coordinate of r. (Such a point exists because no $r \in R$ is extremal in σ_r.) Denote the path from r to x_r in σ_r by P_r, and the neighbor of r in P_r by q_r.

The geometric embedding of P_r starts above ab with r, then goes below ab as $q_r \notin D'$, and finally x_r is again above the line ab. Denote the first intersection (starting from r) of the embedding of the path P_r with the line ab by $\alpha_r = r\bar{q}_r \cap \bar{ab}$, and the next intersection by β_r. Since $|R'_i| \geq 3$, without loss of generality, there are $r_1, r_2 \in R'_i$ such that β_{r_1} is to the left of α_{r_1} and β_{r_2} is to the left of α_{r_2}. For readability and simplicity, let $x_i = x_{r_i}$, $P_i = P_{r_i}$, $q_i = q_{r_i}$, $\alpha_i = \alpha_{r_i}$, $\beta_i = \beta_{r_i}$.

Without loss of generality, suppose that α_1 is to the left of α_2. Recall that P_2 contains only red points, of which only r_2 is in R. Therefore, no other vertex of P_2 can be in D'. If β_2 is to the right of α_1, then one of the edges of P_2 would separate r_1 and r_2 in the sense described in Corollary 1.8. (See Figure 26.) As this cannot happen, β_2 is to the left of α_1.

This implies that $q_1 \notin P_2$ is in the convex hull of P_2 below the ab line. Take the point $q \in S \setminus P_2$ with the smallest y-coordinate such that q is in the convex hull of P_2 below the ab line. As q is not an extremal point of S, it is connected in $\mathcal{DT}$ to some

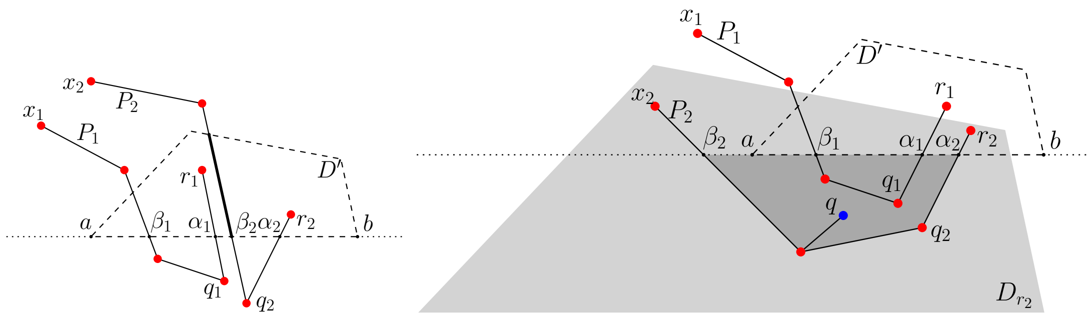


Figure 26: The two cases at the end of the proof of Theorem 2.16. To the left, β_2 is to the right of α_1, the part of the edge splitting D' is bold. To the right, β_2 is to the left of α_1, the shaded regions must contain q_1.

point in S whose y-coordinate is smaller (because the faces of $\mathcal{DT}$ are triangles). By the definition of q, this neighbor must be in P_2. As the end vertices of P_2, r_2 and x_2, are above the ab line, q is connected to an inner vertex of a monochromatic red path. Since every monochromatic component is a path, q cannot be red. The homothet D_{r_2} contains the red vertices of P_2 and thus all the points in the convex hull of P_2. But D_{r_2} is monochromatic, so it cannot contain the non-red point q, a contradiction.

This finishes the proof of Theorem 2.16.

6.3 Stabbed homothets of polygons

it is also interesting to consider the case when our family contains stabbed homothetic copies of a given convex set. Based on parts of [22], we show the following for polygons.

Theorem 6.4 (Damásdi-Pálvölgyi [22]). *For stabbed convex polygons $m_k = O(k)$, i.e., for any convex polygon P there is a integer t_P such that for any positive integer k any finite point set S can be k-colored such that any positive homothetic copy of P, that contains some fixed point $o \notin S$ and $t_P k$ points from S, will contain all k colors.*

Proof. Let P be a fixed polygon and let $v_1, v_2, \dots, v_n$ be the vertices of P. Let $\alpha_{ij} = \inf\{\angle v_i q v_j \mid i \in \{1, \dots, n\}, q \in P \setminus \overline{v_i v_j}\}$. It is not hard to see that this angle α_{ij} is minimized for a vertex v. This also implies $\alpha_{ij} > 0$. Let $\alpha = \min_{i,j} \alpha_{ij}$.

Divide the plane into regions around o such that each region is a cone whose angle is less than α. A homothet of P that contains o can have at most one vertex in each region, otherwise they would be visible from an angle α from o. Let P' and P'' be homothetic copies of P that contain o. We will show that P' and P'' intersects at most once in each region.

Suppose there is a region R where they intersect twice. The boundary of R intersects at most two sides of each polygon, since they are convex and contain o. There are essentially two ways P' and P'' can intersect (see Figure 27). Either the two intersection points of P' and P'' fall on the same side of one of the polygons, or not.

We will show that these are not possible for two homothets. In the first case let s be the side that contains the two intersections. In the second case let s be one of the

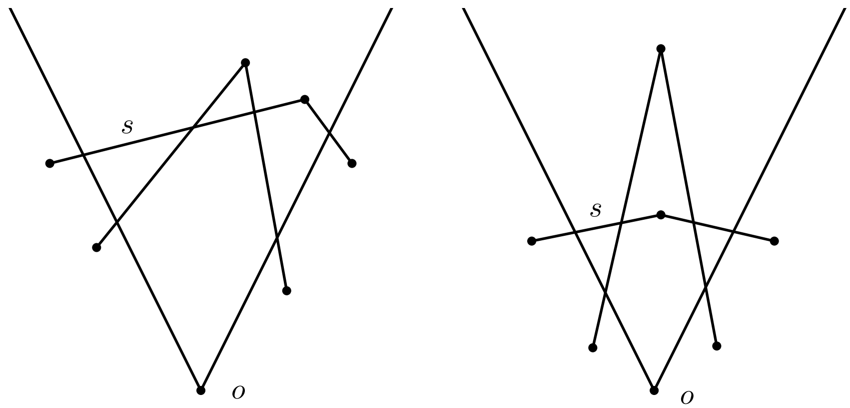


Figure 27: The two ways P' and P'' can intersect each other twice in a region.

sides that ends at the vertex in R that is closer to o. Let n be the normal vector of the side s and let us look at the extremal points of P' and P'' in the direction of n. Since they are homothets of each other, the extremal points should form a side of the polygon in both P' and P''. But for one of them we get only a vertex as an extremal point. Therefore two homothets meet at most once in each region. This implies that in each region they form pseudohalfplanes, thus we can apply Theorem 2.28 separately to each region to color the points. □

We remark that with similar methods it can be proved that for the stabbed translates of any plane convex set $m_k = O(k)$ [22].

6.4 Further remarks

Combining Theorems 2.16 and 5.3, for any convex polygon, D, and for any finite point set, S, we can first find a 3-coloring of S using Theorem 2.16 such that every large (in the sense that it contains many points of S) homothet of D contains two differently colored points, then using Theorem 5.3 we can conclude that every very large homothet of D contains many points from at least two color classes, and finally we can recolor every color class separately using Theorem 2.16. This proves that for every k there is a 3^k-coloring such that every large homothet of D contains at least 2^k colors. Of course, the colors that we use when recoloring need not be different for each color class, so we can also prove for example that there is a 6-coloring such that every large homothet of D contains at least 3 colors. What are the best bounds of this type that can be obtained?

7 Disks and other smooth bodies

The main result of this chapter is that $\chi_m > 2$ for unit disks and for most other smooth bodies. As most of this chapter is based on a slightly older paper [62] (joint work with Pach), many statements use the terminology of cover decomposition, i.e., colorings of the dual range space. In our case, this is equivalent to colorings of the primal range space, as we only deal with translates. It also has statements about infinite coverings, like the following main result.

Theorem 7.1. *For every positive integer m, there exists an m-fold covering of the plane with open unit disks that cannot be split into 2 coverings.*

Our construction can be generalized as follows.

Theorem 7.2. *Let C be any open plane convex set, which has two parallel supporting lines with positive curvature at their points of tangencies. Then, for every positive integer m, there exists an m-fold covering of the plane with translates of C that cannot be split into 2 coverings.*

It follows from Theorem 2.4 (using Theorem 2.6 and compactness) that for every open convex polygon Q, there exists a smallest positive integer $m(Q)$ such that every $m(Q)$-fold covering of the plane with translates of Q splits into 2 coverings. We have that $\sup m(Q) = \infty$, where the sup is taken over all convex polygons Q. Otherwise, we could approximate the unit disk with convex n-gons with n tending to infinity. By compactness, we would conclude that the unit disk C satisfies $m(C) < +\infty$, which contradicts Theorem 7.1. This leaves the following question open.

Problem 7.3. *Does there exist, for any $n > 3$, an integer $m(n)$ such that every convex n-gon Q satisfies $m(Q) \leq m(n)$?*

For any triangle T, there is an affine transformation of the plane that takes it into an equilateral triangle T_0. Therefore, we have $m(T) = m(T_0)$ and $m(3)$ is finite. For every $n \geq 4$, Problem 7.3 is open.

In spite of our sobering negative answer to Problem 2.1 and its analogues in higher dimensions (cp. [56]), there are important classes of multiple coverings such that all of their members are splittable. According to our next, somewhat counter-intuitive result, for example, any m-fold covering of $\mathbb{R}^d$ with unit balls can be split into 2 coverings, provided that no point of the space is covered by too many balls. (We could innocently believe that heavily covered points make it only easier to split an arrangement.)

Theorem 7.4. *For every $d \geq 2$, there exists a positive constant c_d with the following property. For every positive integer m, any m-fold covering of $\mathbb{R}^d$ with unit balls can be split into two coverings, provided that no point of the space belongs to more than $c_d 2^{m/d}$ balls.*

Theorem 7.4, first stated in [56], was one of the first geometric applications of the Lovász local lemma [26], and it was included in [4]. Here, we establish a more general statement (see Theorem 7.29).

One may also believe that *unbounded* convex sets behave even worse than the bounded ones. It turns out, however, that this is not the case.

Theorem 7.5. *Let C be an unbounded open convex set and let P be a finite set of points in the plane. Then every 3-fold covering of $P \subset \mathbb{R}^2$ with translates of C can be split into two coverings of P.*

Here we remark that later in [48] Theorem 7.5 was generalized to arbitrary pseudohalfplane arrangements, see Theorem 2.28.

Using a standard compactness argument, Theorem 7.5 also holds if P is any *compact* set in the plane. However, Theorem 7.5 does not generalize to higher dimensions. Indeed, it follows from the proof of Theorem 7.1 that, for every positive integer m, there exists a finite family $\mathcal{C}$ of open unit disks in the plane and a finite set $P \subset \mathbb{R}^2$ such that $\mathcal{C}$ is an m-fold covering of P that cannot be split into two coverings. Consider now an unbounded convex cone C' in $\mathbb{R}^3$, whose intersection with the plane $\mathbb{R}^2$ is an open disk. Take a system of translates of C' such that their intersections with the plane coincide with the members of $\mathcal{C}$. These cones form an m-fold covering of P that cannot be split into two coverings.

For interesting technical reasons, the proof of Theorem 7.5 becomes much easier if we restrict our attention to multiple coverings of the *whole plane*. In fact, in this case, we do not even have to consider *multiple* coverings! Moreover, the statement remains true in higher dimensions.

Proposition 7.6. *Let C be an unbounded line-free open convex set in $\mathbb{R}^d$. Then every covering of $\mathbb{R}^d$ with translates of C can be split into two, and hence into infinitely many, coverings.*

The reason why we assume here that C is *line-free* (i.e., does not contain a full line) is the following. If C contains a straight line, then it can be obtained as the direct product of a line ℓ and a $(d-1)$-dimensional open convex set C'. Any arrangement $\mathcal{C}$ of translates of C in $\mathbb{R}^d$ is combinatorially equivalent to the $(d-1)$-dimensional arrangement of translates of C', obtained by cutting $\mathcal{C}$ with a hyperplane orthogonal to ℓ. In particular, the problem whether an m-fold covering of $\mathbb{R}^d$ with translates of C can be split into two coverings reduces to the respective question about m-fold coverings of $\mathbb{R}^{d-1}$ with translates of C'.

Proposition 7.6 is false already in the plane without the assumption that C is open. However, every 2-fold covering of the plane with translates of an unbounded C can be split into two coverings. We omit the proof as it reduces to a simple claim about intervals.

However, in higher dimensions, the similar claim is false.

Theorem 7.7. *There is a bounded convex set $C' \subset \mathbb{R}^3$ with the following property. One can construct a family of translates of $C = C' \times [0, \infty) \subset \mathbb{R}^4$ which covers every point of $\mathbb{R}^4$ infinitely many times, but which cannot be split into two coverings.*

The construction given in Section 7.8 is based on an example of Naszódi and Taschuk [59], and explores the fact that the boundary of C' can be rather "erratic." We do not know whether sufficiently thick coverings of $\mathbb{R}^3$ by translates of an unbounded line-free convex set can be split into two coverings or not.

In the sequel, we will study the equivalent of the above questions for primal range spaces. As we talk about translates, Theorem 7.1 can be rephrased in the following form.

Theorem 7.1'. *For every $m \geq 2$, there is a set of points $P^* = P^*(m)$ in the plane with the property that every open unit disk contains at least m elements of P^*, and no matter how we color the elements of P^* with two colors, there exists a unit disk such that all points in it are of the same color.*

A set system *not* satisfying this condition is said to have *property B* (in honor of Bernstein) or is 2-*colorable* (see [58, 25, 71]). Generalizations of this notion are related to *conflict-free colorings* [29] and have strong connections, e.g., to the theory of ε-*nets*, *geometric set covers* and to *combinatorial game theory* [41, 65, 3, 76, 35].

The rest of this chapter is organized as follows. In the next three sections, we prove Theorem 7.1' in 3 steps. In Section 7.1, we exhibit a family of non-2-colorable m-uniform hypergraphs $\mathcal{H}(k,\ell)$ which first appeared in [67]. In Section 7.2, we construct planar "realizations" of these hypergraphs, where the vertices correspond to points and the (hyper)edges to unit disks, preserving the incidence relations. In Section 7.3, we show how the same hypergraph can be easily realized with halfdisks, a fact that was missed in [62] and was noticed later, in [24]. In Section 7.4, we extend this construction, without violating the colorability condition, so that every disk contains at least m points. In Section 7.5, we modify these steps in order to establish Theorem 7.2, a generalization of Theorem 7.1 to bounded plane convex bodies with a smooth boundary. Sections 7.7 and 7.8 contain the proofs of our results related to multiple coverings with *unbounded* convex sets: Theorem 7.5, Proposition 7.6, and Theorem 7.7. The proof of a more general version of Theorem 7.4, using the Lovász local lemma, can be found in Section 7.9. Finally, in Sections 7.10 and 7.11 we give the proofs of Theorems 2.31 and 2.30, the newest results from [22] and [24].

7.1 A family of non-2-colorable hypergraphs $\mathcal{H}(k,\ell)$

In this section we define, for any positive integers k and ℓ, the abstract hypergraph $\mathcal{H}(k,\ell)$ with vertex set $V(k,\ell)$ and edge set $E(k,\ell)$ that was first used in [67]. The hypergraphs $\mathcal{H}(k,\ell)$ are defined recursively. The edge set $E(k,\ell)$ will be the disjoint union of two sets, $E(k,\ell) = E_R(k,\ell) \uplus E_B(k,\ell)$, where the subscripts R and B stand for red and blue. All edges belonging to $E_R(k,\ell)$ will be of size k, all edges belonging to $E_B(k,\ell)$ will be of size ℓ. In other words, $\mathcal{H}(k,\ell)$ is the union of a k-*uniform* and an ℓ-*uniform* hypergraph. If $k = \ell = m$, we get an m-uniform hypergraph.

Definition 7.8. *Let k and ℓ be positive integers.*

1. *For $k = 1$, let $V(1,\ell)$ be an ℓ-element set.*
   Set $E_R(1,\ell) := V(1,\ell)$ and $E_B(1,\ell) := \{V(1,\ell)\}$.
2. *For $\ell = 1$, let $V(k,1)$ be a k-element set.*
   Set $E_R(k,1) := \{V(k,1)\}$ and $E_B(k,1) := V(k,1)$.
3. *For any $k,\ell > 1$, we pick a new vertex p, called the* root*, and let*

$$V(k,\ell) := V(k-1,\ell) \uplus V(k,\ell-1) \uplus \{p\},$$

$$E_R(k,\ell) := \{e \cup \{p\} : e \in E_R(k-1,\ell)\} \cup E_R(k,\ell-1),$$

$$E_B(k,\ell) := E_B(k-1,\ell) \uplus \{e \cup \{p\} : e \in E_B(k,\ell-1)\}.$$

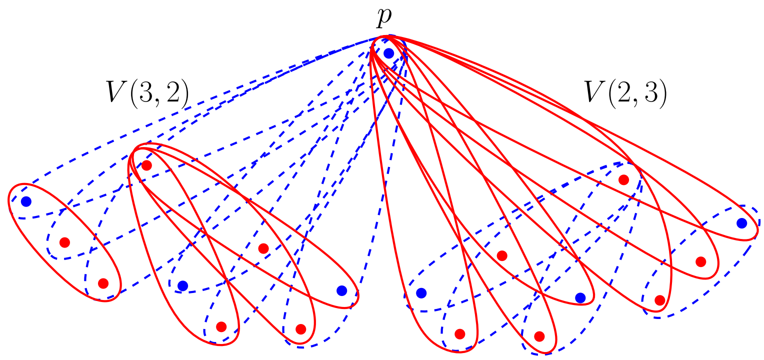


Figure 28: The hypergraph $\mathcal{H}(3,3)$ with (arbitrarily) 2-colored vertices. There is a blue (dashed) set with 3 blue vertices or a red (solid) set with 3 red vertices.

By recursion, we obtain that

$$|V(k,\ell)| = \binom{k+\ell}{k} - 1,$$

$$|E_R(k,\ell)| = \binom{k+\ell-1}{k}, \ |E_B(k,\ell)| = \binom{k+\ell-1}{\ell},$$

$$|E(k,\ell)| = |E_R(k,\ell)| + |E_B(k,\ell)| = \binom{k+\ell}{k}.$$

Lemma 7.9 ([67])**.** *For any positive integers k, ℓ, the hypergraph $\mathcal{H}(k,\ell)$ is not 2-colorable. Moreover, for every coloring of $V(k,\ell)$ with red and blue, there is an edge in $E_R(k,\ell)$ such that all of its k vertices are red or an edge in $E_B(k,\ell)$ such that all of its ℓ vertices are blue.*

For completeness, here we include the proof of Lemma 7.9 from [67]. The induction on two parameters, k and ℓ, is similar to the proof of Ramsey's theorem by Erdős and Szekeres [27].

Proof. We will prove that for every coloring of $V(k,\ell)$ with red and blue, there is an edge in $E_R(k,\ell)$ such that all of its k vertices are red or an edge in $E_B(k,\ell)$ such that all of its ℓ vertices are blue.

Suppose first that $k = 1$. If any vertex in $V(1,\ell)$ is red, then it is a red singleton edge in $\mathcal{H}(1,\ell)$. If all vertices in $V(1,\ell)$ are blue, then the (only) edge $V(1,\ell) \in E_B(1,\ell)$ contains only blue points. Analogously, the assertion is true if $\ell = 1$.

Suppose next that $k,\ell > 1$. Assume, without loss of generality, that the root p is red. Consider the subhypergraph $\mathcal{H}(k-1,\ell) \subset \mathcal{H}(k,\ell)$ induced by the vertices in $V(k-1,\ell)$. If it has a monochromatic red edge $e \in E_R(k-1,\ell)$, then $e \cup \{p\} \in E_R(k,\ell)$ is red. If there is a monochromatic blue edge in $E_B(k-1,\ell)$, then we are again done, because it is also an edge in $E_B(k,\ell)$. □

For other interesting properties of the hypergraphs $\mathcal{H}(k,\ell)$ related to hereditary discrepancy, see Matoušek [57].

7.2 Geometric realization of the hypergraphs $\mathcal{H}(k,\ell)$

The aim of this section is to establish the following version of Theorem 7.1' that states that $\chi_m > 2$ for unit disks.

Theorem 7.1". *For every $m \geq 2$, there exists a* finite *point set $P = P(m) \subset \mathbb{R}^2$ and a* finite *family of unit disks $\mathcal{C} = \mathcal{C}(m)$ with the property that every member of $\mathcal{C}$ contains at least m elements of P, and no matter how we color the elements of P with two colors, there exists a disk in $\mathcal{C}$ such that all points in it are of the same color.*

We *realize* the hypergraph $\mathcal{H}(k,\ell)$ defined in Section 7.1 with points and disks. The vertex set $V(k,\ell)$ is mapped to a point set $P(k,\ell) \subset \mathbb{R}^2$, and the edge sets, $E_R(k,\ell)$ and $E_B(k,\ell)$, to families of open unit disks, $\mathcal{C}_R(k,\ell)$ and $\mathcal{C}_B(k,\ell)$, so that a vertex belongs to an edge if and only if the corresponding point is contained in the corresponding disk. The geometric properties of this realization are summarized in the following lemma.

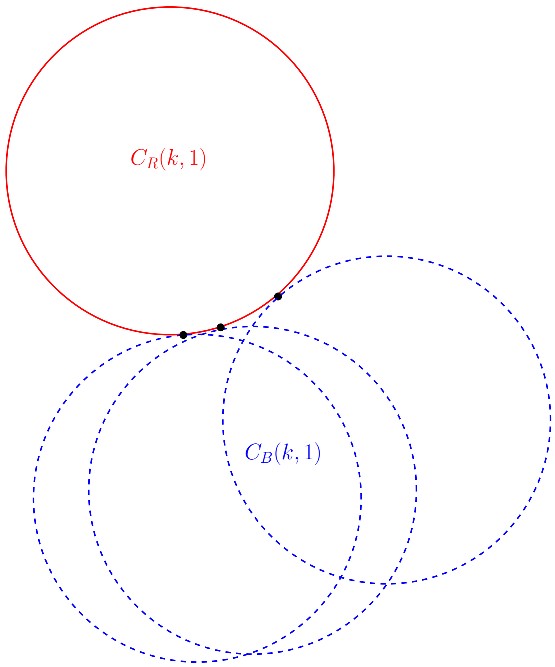
Given two unit disks C, C', let $d(C,C')$ denote the distance between their centers. We fix an orthogonal coordinate system in the plane so that we can talk about the *topmost* and the *bottommost* points of a disk.[14]

Lemma 7.10. *For any positive integers k,ℓ and for any $\varepsilon > 0$, there is a finite point set $P = P(k,\ell)$ and a finite family of open unit disks $\mathcal{C}(k,\ell) = \mathcal{C}_R(k,\ell) \cup \mathcal{C}_B(k,\ell)$ with the following properties.*

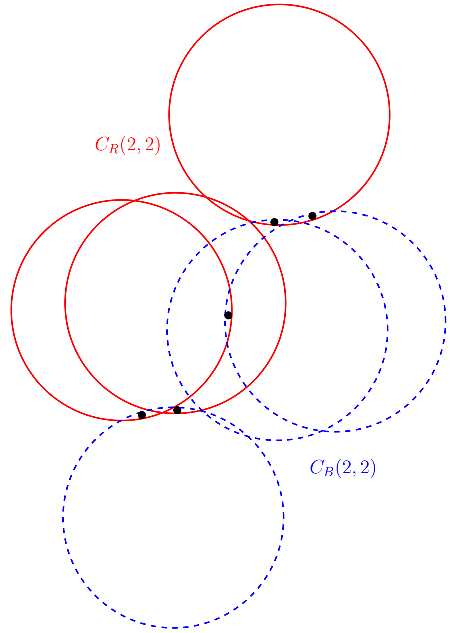
1. *Any disk $C \in \mathcal{C}_R(k,\ell)$ (resp. $\mathcal{C}_B(k,\ell)$) contains precisely k (resp. ℓ) points of P.*
2. *For any coloring of P with red and blue, there is a disk in $\mathcal{C}_R(k,\ell)$ such that all of its points are red or a disk in $\mathcal{C}_B(k,\ell)$ such that all of its point are blue. In fact, P and $\mathcal{C}(k,\ell)$* realize *the abstract hypergraph $\mathcal{H}(k,\ell)$ in the above sense.*
3. *For the coordinates (x,y) of any point from P, we have $-\varepsilon < x < \varepsilon$ and $-\varepsilon^2 < y < \varepsilon^2$.*
4. *For the coordinates (x,y) of the center of any disk from $\mathcal{C}_R(k,\ell)$, we have $-\varepsilon < x < \varepsilon$ and $-\varepsilon^2 < y - 1 < \varepsilon^2$.*
5. *For the coordinates (x,y) of the center of any disk from $\mathcal{C}_B(k,\ell)$, we have $-\varepsilon < x < \varepsilon$ and $-\varepsilon^2 < y + 1 < \varepsilon^2$.*
6. *The topmost and the bottommost points of a disk $C \in \mathcal{C}(k,\ell)$ are not covered by the closure of any other member of $\mathcal{C}(k,\ell)$.*

Looking at our construction from "far away" the two families $\mathcal{C}_R$ and $\mathcal{C}_B$ look like two touching disks, with all points of P very close to the touching point. The segments connecting the centers of disks from different families are almost vertical with all members of $\mathcal{C}_R$ lying "above" all members of $\mathcal{C}_B$. We prove the lemma by induction. Most conditions are needed for the induction to go through. Condition 6 is an exception: it will be used in Section 7.4.

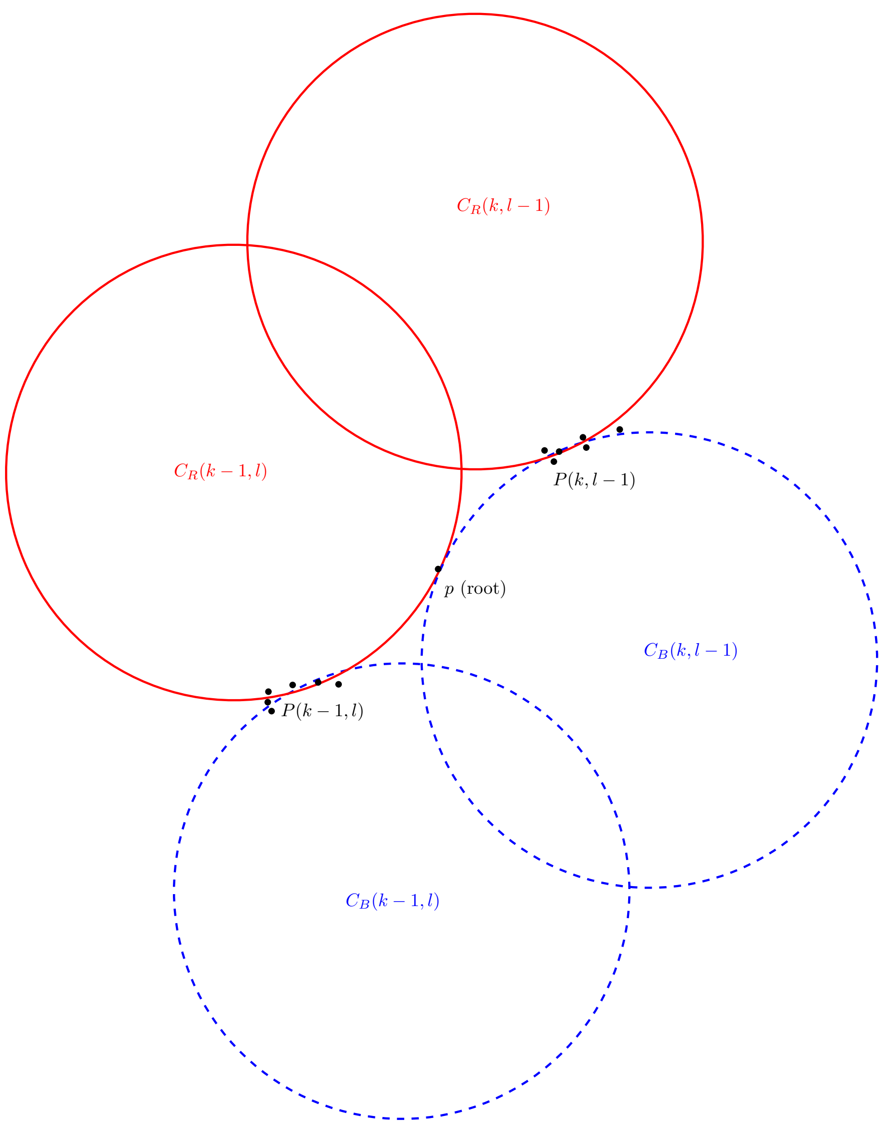
[14] Beware that these extremal points are not inside the open disk.



(a) Starting step: $\mathcal{C}(k,1)$.



(b) $\mathcal{C}(2,2)$ magnified (and a bit distorted for visibility).



(c) Induction step.

Figure 29: The construction.

Proof. We give a recursive construction. We can assume that $\varepsilon < 1/10$. It is easy to see that, for $k = 1$ or $\ell = 1$, there exists such a family of unit disks for any $\varepsilon > 0$, see Figure 29(a). The family $\mathcal{C}(2,2)$ is depicted in Figure 29(b), where the main idea of the induction may already be visible.

Suppose that $k, \ell \geq 2$ and we have already constructed $P(k-1,\ell)$ and $\mathcal{C}(k-1,\ell)$, and $P(k,\ell-1)$ and $\mathcal{C}(k,\ell-1)$, for some $\varepsilon(k-1,\ell) < \varepsilon/100$ and $\varepsilon(k,\ell-1) < \varepsilon/100$, respectively. To obtain $P(k,\ell)$, we place the root p of $\mathcal{H}(k,\ell)$ into the origin $(0,0)$, and we shift (translate) $P(k-1,\ell)$ and $P(k,\ell-1)$ into new positions such that their roots are at $(-\varepsilon/3, -\varepsilon^2/10)$ and $(\varepsilon/3, \varepsilon^2/10)$, respectively. With a slight abuse of notation, the shifted copies will also be denoted $P(k-1,\ell)$ and $P(k,\ell-1)$. See Figure 29(c). In this way, it is guaranteed that for the coordinates (x,y) of any point of P, we have

$$-\varepsilon < -(\varepsilon/3 + \varepsilon(k-1,\ell) + \varepsilon(k,\ell-1)) < x < \varepsilon/3 + \varepsilon(k-1,\ell) + \varepsilon(k,\ell-1) < \varepsilon$$

and

$$-\varepsilon^2 < -(\varepsilon^2/10 + \varepsilon^2(k-1,\ell) + \varepsilon^2(k,\ell-1)) < y < \varepsilon^2/3 + \varepsilon^2(k-1,\ell) + \varepsilon^2(k,\ell-1) < \varepsilon^2.$$

Thus, **property 3** of the lemma holds.

The family $\mathcal{C}(k,\ell)$ is defined as the union of two previously defined families, $\mathcal{C}(k-1,\ell)$ and $\mathcal{C}(k,\ell-1)$, translated by the same vectors as $P(k-1,\ell)$ and, resp. $P(k,\ell-1)$ were. Again, we use the same symbols to denote the translated copies. To verify **properties 4 and 5**, we only have to repeat the above calculations, with the y-coordinates being shifted 1 higher (resp. 1 lower).

Now we show that our set of points $P(k,\ell)$ and set of disks $\mathcal{C}(k,\ell)$ realize the hypergraph $\mathcal{H}(k,\ell)$ (**properties 1 and 2**). It is easy to see that if $C \in \mathcal{C}_R(k-1,\ell)$ and $s \in P(k,\ell-1)$, then $s \notin C$ but $p = (0,0) \in C$. The coordinates of the center of C are $\big(-\varepsilon/3 \pm \varepsilon(k-1,\ell), 1 - \varepsilon^2/10 \pm \varepsilon^2(k-1,\ell)\big)$ (where here and in the following, $\pm z$ denotes a number that is between $-z$ and z), so the distance of p from C is at most $(\varepsilon/3 + \varepsilon(k-1,\ell))^2 + (1 - \varepsilon^2/10 + \varepsilon^2(k-1,\ell))^2 < 1$. On the other hand, the coordinates of s are $\big(\varepsilon/3 \pm \varepsilon(k,\ell-1), \varepsilon^2/10 \pm \varepsilon^2(k,\ell-1)\big)$, thus the square of its distance from the center of C is at least

$$\big(2\varepsilon/3 - \varepsilon(k-1,\ell) - \varepsilon(k,\ell-1)\big)^2 + \big(1 - 2\varepsilon^2/10 - \varepsilon^2(k-1,\ell) - \varepsilon^2(k,\ell-1)\big)^2 > 1.$$

Analogously, if $C \in \mathcal{C}_B(k,\ell-1)$ and $s \in P(k-1,\ell)$, then $s \notin C$ but $p = (0,0) \in C$.

Let $C \in \mathcal{C}_R(k,\ell-1)$ and $s \in P(k-1,\ell)$. We prove that $p, s \notin C$. The coordinates of the center of C are $\big(\varepsilon/3 \pm \varepsilon(k,\ell-1), 1 + \varepsilon^2/10 \pm \varepsilon(k,\ell-1)\big)$. Therefore, the distance of p from the center of C is at least $(\varepsilon/3 - \varepsilon(k,\ell-1))^2 + (1 + \varepsilon^2/10 - \varepsilon(k,\ell-1))^2 > 1$. The calculation for s is similar in the case $C \in \mathcal{C}_R(k-1,\ell)$. Analogously, we have that if $C \in \mathcal{C}_B(k-1,\ell)$ and $s \in P(k,\ell-1)$, then $p, s \notin C$. As the disks in $\mathcal{C}(k,\ell-1)$ (resp. $\mathcal{C}(k-1,\ell)$) contain precisely the same points of $P(k,\ell-1)$ (resp. $P(k-1,\ell)$, as before the shift, we have obtained a geometric realization of $\mathcal{H}(k,\ell)$, and properties 1 and 2 hold.

It remains to prove that the topmost and the bottommost points of a disk $C \in \mathcal{C}(k,\ell)$ are not covered by any other member of $\mathcal{C}(k,\ell)$ (**property 6**). Using that our construction and disks are centrally symmetric, it is enough to prove the statement for the topmost points. If $C \in \mathcal{C}_R(k,\ell-1)$, the coordinates of its topmost point are

$\big(\varepsilon/3 \pm \varepsilon(k, \ell-1), 2+\varepsilon^2/10 \pm \varepsilon^2(k, \ell-1)\big)$. If $C \in \mathcal{C}_R(k-1, \ell)$, the coordinates of its topmost point are $\big(-\varepsilon/3 \pm \varepsilon(k-1, \ell), 2-\varepsilon^2/10 \pm \varepsilon^2(k-1, \ell)\big)$. If $C \in \mathcal{C}_B(k, \ell-1)$, the coordinates of its topmost point are $\big(\varepsilon/3 \pm \varepsilon(k, \ell-1), -2+\varepsilon^2/10 \pm \varepsilon^2(k, \ell-1)\big)$. If $C \in \mathcal{C}_B(k-1, \ell)$, the coordinates of its topmost point are $\big(-\varepsilon/3 \pm \varepsilon(k-1, \ell), -2-\varepsilon^2/10 \pm \varepsilon^2(k-1, \ell)\big)$.

If $C \in \mathcal{C}_R(k, \ell-1)$, by the induction hypothesis, its topmost point cannot be covered by any other disk from $\mathcal{C}(k, \ell-1)$. Nor can it be covered by any other disk, as the topmost points of all other disks are below it (i.e., have smaller y-coordinates). If $C \in \mathcal{C}_R(k-1, \ell)$, then the square of the distance of its topmost point from the center of some $C' \in \mathcal{C}_R(k, \ell-1)$ is at least

$$\big(2\varepsilon/3 - \varepsilon(k, \ell-1) - \varepsilon(k-1, \ell)\big)^2 + \big(1 - 2\varepsilon^2/10 - \varepsilon^2(k, \ell-1) - \varepsilon^2(k-1, \ell)\big)^2 > 1.$$

If $C \in \mathcal{C}_B(k, \ell-1)$, then the distance of its topmost point from the center of some $C' \in \mathcal{C}_R(k-1, \ell)$ is also at least

$$\big(2\varepsilon/3 - \varepsilon(k, \ell-1) - \varepsilon(k-1, \ell)\big)^2 + \big(1 - 2\varepsilon^2/10 - \varepsilon^2(k, \ell-1) - \varepsilon^2(k-1, \ell)\big)^2 > 1.$$

In all other cases, trivially, the corresponding distances are also larger than 1. This completes the proof of property 6 and hence the lemma. □

7.3 Halfdisks

Theorem 7.2 states that, if C is a plane convex body with two antipodal points at which the curvature is positive, then for every m, there exists an m-fold covering of $\mathbb{R}^2$ with translates of C that does not split into two coverings. We also know that this statement is false for any convex polygon. But what happens if C "almost satisfies" the condition concerning the antipodal point pair?

Problem 7.11. *Does there exist an integer m such that every m-fold covering of $\mathbb{R}^2$ with translates of an open semidisk splits into two coverings?*

This question was answered later in [24], where it was later observed that the construction presented in this chapter is in fact easily realizable by halfdisks and similar shapes, see Figure 30.

7.4 Adding points to P—Proof of Theorem 7.1'

In this section, we extend the proof of Theorem 7.1" to establish Theorem 7.1' (which is equivalent to Theorem 7.1). Note that the only difference between Theorems 7.1" and 7.1' is that in the latter it is also required that every unit disk of the plane contains at least m elements of the point set $P^* = P^*(m)$. The set $P = P(m, m)$ constructed in Lemma 7.10, does not satisfy this condition. In order to fix this, we will add all points *not* in $\cup\mathcal{C}(m, m)$ to the set P (or rather a sufficiently dense discrete subset of $\mathbb{R}^2 \setminus \cup\mathcal{C}(m, m)$). In order to show that the resulting set P^* meets the requirements of Theorem 7.1', all we have to show is the following.

Lemma 7.12. *No (open) unit disk $C \notin \mathcal{C}(k, \ell)$ is entirely contained in $\cup\mathcal{C}(k, \ell)$.*

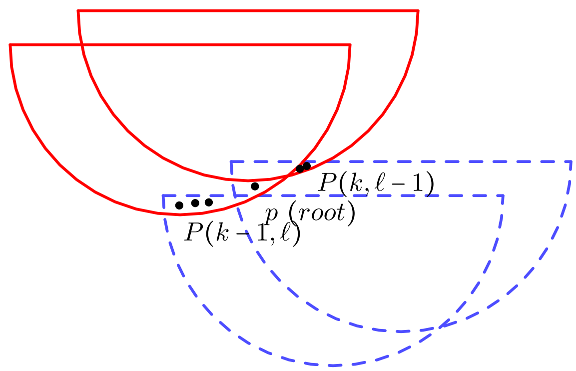


Figure 30: The recursive step of the construction with unit halfdisks.

For future purposes, we prove this statement in a slightly more general form. In what follows, we only assume that C is an open convex body with a unique topmost point t and a unique bottommost point b,[15] which divide the boundary of C into two *closed arcs*. They will be referred to as the *left boundary arc* and a *right boundary arc*.

Definition 7.13. *A collection $\mathcal{C}$ of translates of C is said to be* exposed *if the topmost and bottommost points of its members do not belong to the closure of any other member of $\mathcal{C}$.*

By the last condition in Lemma 7.10, the collections of disks $\mathcal{C}(k, \ell)$ constructed in the previous section are exposed. We prove the following generalization of Lemma 7.12.

Lemma 7.14. *Let $\mathcal{C}$ be a finite exposed collection of translates of an open convex body C with unique topmost and bottommost points. If $C \notin \mathcal{C}$, then $C \nsubseteq \cup\mathcal{C}$.*

For the proof, we need a simple observation.

Claim 7.15. *If the right boundary arcs of two translates of C intersect, then the closure of one of the translates must contain the topmost or bottommost point of the other.*

Proof. Let C_1 and C_2 be the two translates, and let γ_i denote the closed convex curve formed by the right boundary arc of C_i and the straight-line segment connecting its two endpoints (the topmost and the bottommost points of C_i). The curves γ_1 and γ_2 are translates of each other, and since they intersect, they must *cross* twice. (At a *crossing*, one curve comes from the exterior of the other, then it shares an arc with it, which may be a single point, and enters the interior.) It cannot happen that both crossings occur between the right boundary arcs, because they are convex and translates of each other. Therefore, one of the two crossings involves the straight-line segment of one the curves, say, γ_1. But since the condition is that the right boundary arcs intersect, one of the two endpoints of this straight-line segment, either the topmost or the bottommost point of C_1, lies in the closure of C_2 $\square$

[15] Recall that these extremal points are not inside the open C.

Proof of Lemma 7.14. Suppose, for contradiction, that $C \subseteq \cup\mathcal{C}$. By removing some members of $\mathcal{C}$ if necessary, we can assume that $\mathcal{C}$ is a *minimal* collection of translates that covers C. As $C \notin \mathcal{C}$, there cannot be only one translate in $\mathcal{C}$. If $\mathcal{C}$ consists of only two translates, $\{C_1, C_2\} = \mathcal{C}$, then the topmost point of C, t, must lie on the segment whose endpoints are the topmost points of C_1 and C_2, t_1 and t_2. But in this case, if C is covered, either $t_1 \in C_2$ or $t_2 \in C_1$, contradicting the assumption that $\mathcal{C}$ is exposed. From this we can conclude that if $C_1, C_2 \in \mathcal{C}$ intersect inside C, their boundaries must also intersect inside C, otherwise either $C \subset C_1 \cup C_2$, or $C \cap C_1 \subset C_2$ or $C \cap C_2 \subset C_1$, contradicting the minimality of $\mathcal{C}$. Fix a $p \in C$ that lies on the boundary of C_1 and C_2, inside C. Take a $C_3 \in C$ that covers p. Since C_3 is open, it covers a small neighborhood of p, thus we have $C_1 \cap C_2 \cap C_3 \cap C \neq \emptyset$.

None of the topmost and bottommost points of these three translates can be covered by C, otherwise, it would also be covered by another member of $\mathcal{C}$, contradicting the assumption that $\mathcal{C}$ is exposed. Thus, C intersects either the left or the right boundary arc of every C_i. Without loss of generality, suppose that C intersects the right boundary arcs of C_1 and C_2. These right boundary arcs must intersect inside C, otherwise $C_1 \cap C \subseteq C_2 \cap C$ or $C_2 \cap C \subseteq C_1 \cap C$, and $\mathcal{C}$ would not be minimal. Therefore, we can apply Claim 7.15 to conclude that one of them must contain the topmost or bottommost point of the other. $\square$

Remark 7.16. *In the construction described in Lemma 7.10, every disk in $\mathcal{C}(m,m)$ contains at most $|P(m,m)| < 2^{2m}$ points. At the last stage, we added many new points to P. We can keep the maximum number of points of P lying in a unit disk bounded from above by a function $f(m)$. What is the best upper bound? The bound given by our construction depends on $\varepsilon(m,m) \leq 100^{-2m}\varepsilon(1,1)$.*

7.5 Other convex bodies—Proof of Theorem 7.2

Throughout this section, C denotes an open plane convex body which has two parallel supporting lines with positive curvature at the two points of tangencies. To prove Theorem 7.2, by duality, it is sufficient to establish the analogue of Theorem 7.1', where the role of unit disks is played by translates of C.

Theorem 7.2'. *For every $m \geq 2$, there is a set of points $P^* = P^*(m)$ in the plane with the property that every translate of C contains at least m elements of P^*, and no matter how we color the elements of P^* with two colors, there exists a translate of C such that all points in it are of the same color.*

As in the case of disks, after defining the hypergraphs $\mathcal{H}(k,\ell)$, the proof consists of two steps:

Step 1: We find a geometric realization of $\mathcal{H} = \mathcal{H}(k,\ell)$ with translates of C, i.e., a finite point set P representing the vertices and a collection $\mathcal{C}$ of translates of C representing the hyperedges of $\mathcal{H}$ such that a point of P lies in a member of $\mathcal{C}$ if and only if the corresponding vertex belongs to the corresponding hyperedge. We show that $\mathcal{C}$ is an exposed family.

Step 2: We show that no translate of C is entirely contained in $\cup\mathcal{C}$, unless $C \in \mathcal{C}$. Thus, we can add all the points not in $\cup\mathcal{C}(k,\ell)$ to the points of P to ensure that every translate of C contains many points.

In Section 7.4, we have shown that Step 2 can be completed, provided that $\mathcal{C}$ is exposed (see Lemma 7.14). Therefore, here we concentrate on Step 1.

Without loss of generality, we can assume that C has unique bottommost and topmost points, b and t, resp., at which the curvature is positive. After applying an affine transformation, we can also attain that the line bt is vertical. Let r_b and r_t denote the reciprocals of the curvatures at b and t, respectively. If we place b at the origin, then, for every $\delta > 0$, in a small neighborhood of b, the boundary of C will lie between the parabolas $y = (1-\delta)r_b x^2$ and $y = (1+\delta)r_b x^2$. Analogously, if we place t at the origin, then in a small neighborhood of it, the boundary of C will lie between the parabolas $y = -(1-\delta)r_t x^2$ and $y = -(1+\delta)r_t x^2$. We find a geometric realization using the following lemma.

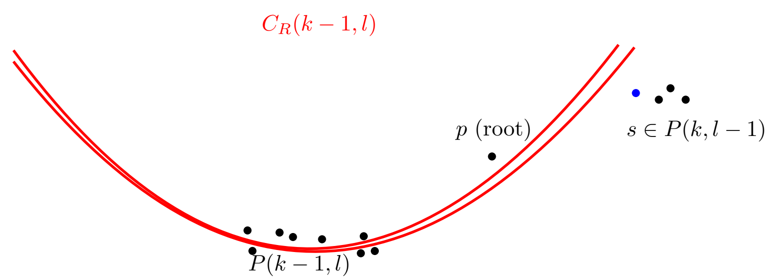


Figure 31: Parabolas enclosing the boundary of C.

Lemma 7.10'. *For any positive integers k,ℓ and for any $\varepsilon > 0$, there is a finite point set $P = P(k,\ell)$ and a finite family of translates of C, $\mathcal{C}(k,\ell) = \mathcal{C}_R(k,\ell) \cup \mathcal{C}_B(k,\ell)$ with the following properties.*

1. *Any translate from $\mathcal{C}_R(k,\ell)$ (resp. $\mathcal{C}_B(k,\ell)$) contains precisely k (resp. ℓ) points of P.*
2. *For any coloring of P with red and blue, there is a translate from $\mathcal{C}_R(k,\ell)$ such that all of its points are red or a translate from $\mathcal{C}_B(k,\ell)$ such that all of its point are blue. In fact, P and $\mathcal{C}(k,\ell)$* realize *the abstract hypergraph $\mathcal{H}(k,\ell)$ in the above sense.*
3. *For the coordinates (x,y) of any point from P, we have $-\varepsilon < x < \varepsilon$ and $-\varepsilon^2 < y < \varepsilon^2$.*
4. *For the coordinates (x,y) of the bottommost point of any translate from $\mathcal{C}_R(k,\ell)$, we have $-\varepsilon < x < \varepsilon$ and $-\varepsilon^2 < y < \varepsilon^2$.*
5. *For the coordinates (x,y) of the topmost point of any translate from $\mathcal{C}_B(k,\ell)$, we have $-\varepsilon < x < \varepsilon$ and $-\varepsilon^2 < y < \varepsilon^2$.*
6. *The topmost and the bottommost points of translate from $\mathcal{C}(k,\ell)$ are not covered by the closure of any other member of $\mathcal{C}(k,\ell)$.*

Proof. Using an affine transformation, we can suppose that $r_t, r_b < 1$. We fix a δ that is small enough compared to r_t and r_b, and an $\varepsilon = \varepsilon(k,\ell)$ that is small enough compared to δ, r_t and r_b, but big enough compared to $\varepsilon(k, \ell-1)$ and $\varepsilon(k-1, \ell)$. (To keep the presentation simple, we omit the exact required dependencies here.) We will use that the boundary of C in a $2\varepsilon(r_t + r_b)$ neighborhood around t and b is between the (above mentioned) pairs of parabolas, $y = (1-\delta)r_b x^2$ and $y = (1+\delta)r_b x^2$, and $y = -(1-\delta)r_t x^2$ and $y = -(1+\delta)r_t x^2$.

If $k = 1$ or $\ell = 1$, the construction is trivial. For $k, \ell \geq 2$, assume that the point sets $P(k-1, \ell)$ and $P(k, \ell-1)$, and the families of translates of C, $\mathcal{C}(k-1, \ell)$ and $\mathcal{C}(k, \ell-1)$, have already been defined, and that they satisfy all conditions in the lemma. To obtain $P(k, \ell)$, we place the root p of $\mathcal{H}(k, \ell)$ into the origin $(0,0)$, and we shift $P(k-1, \ell)$ and $P(k, \ell-1)$ such that their roots are at $(-r_t\varepsilon, -(1+2\delta)r_b r_t^2 \varepsilon^2)$ and $(r_b\varepsilon, (1+2\delta)r_t r_b^2 \varepsilon^2)$, respectively. The family of translates $\mathcal{C}(k, \ell)$ is defined as the union of the families $\mathcal{C}(k-1, \ell)$ and $\mathcal{C}(k, \ell-1)$ translated by the same vectors, as $P(k-1, \ell)$ and $P(k, \ell-1)$, respectively.

To verify **properties 3, 4, and 5**, we need that $-\varepsilon < -r_t\varepsilon, r_b\varepsilon < \varepsilon$ and $-\varepsilon^2 < -(1+2\delta)r_b r_t^2 \varepsilon^2, (1+2\delta)r_t r_b^2 \varepsilon^2 < \varepsilon^2$, which hold since $r_t, r_b < 1$ and δ is small. Notice that where we have omitted $\varepsilon(k, \ell-1)$ and $\varepsilon(k-1, \ell)$ from these equations to keep the calculations simple. This we can do as the difference of the two sides depends on ε, which we can select to be sufficiently large compared to $\varepsilon(k, \ell-1)$ and $\varepsilon(k-1, \ell)$. We will also omit dependencies of $\varepsilon(k, \ell-1)$ and $\varepsilon(k-1, \ell)$ later.

To verify **properties 1 and 2**, we have to show that for any $C \in \mathcal{C}_R(k-1, \ell)$, the origin $p = (0,0)$ belongs to C, but no point $s \in P(k, \ell-1)$ does, provided that $\varepsilon > 0$ is sufficiently small. To see this, fix $C \in \mathcal{C}_R(k-1, \ell)$. The equation of the parabola that touches C from the inside at its bottommost point is approximately $y = (1+\delta)r_b(x + r_t\varepsilon)^2 - (1+2\delta)r_b r_t^2 \varepsilon^2$. If $x = 0$, the value of y is $(1+\delta)r_b(r_t\varepsilon)^2 - (1+2\delta)r_b r_t^2\varepsilon^2 = -\delta r_b r_t^2 \varepsilon^2$. This is negative, which means that $p = (0,0)$ lies above the parabola. Thus, we have $p \in C$. Analogously, if $C \in \mathcal{C}_B(k, \ell-1)$ and $s \in P(k-1, \ell)$, then $s \notin C$ but $p = (0,0) \in C$.

On the other hand, the equation of the parabola that touches C at its bottommost point from the outside is approximately $y = (1-\delta)r_b(x + r_t\varepsilon)^2 - (1+2\delta)r_b r_t^2 \varepsilon^2$. If $x = r_b\varepsilon \pm \varepsilon(k, \ell-1)$ the value of y at x is approximately

$$(1-\delta)r_b(r_b\varepsilon + r_t\varepsilon)^2 - (1+2\delta)r_b r_t^2 \varepsilon^2 = \left((1-\delta)(r_b^3 + 2r_b^2 r_t) - 3\delta r_b r_t^2\right)\varepsilon^2 \geq \left(r_b^3 + O(\delta)\right)\varepsilon^2.$$

Therefore, $s = (r_b\varepsilon \pm \varepsilon(k, \ell-1), (1+2\delta)r_t r_b^2 \varepsilon^2 \pm \varepsilon^2(k, \ell-1))$ is below the parabola, if δ is small enough, thus $s \notin C$.

Let $C \in \mathcal{C}_R(k, \ell-1)$ and $s \in P(k-1, \ell)$. We prove that $p, s \notin C$. The equation of the parabola that touches C from the outside at its bottommost point is approximately $y = (1-\delta)r_t(x - r_b\varepsilon)^2 - (1+2\delta)r_t r_b^2 \varepsilon^2$. If $x = 0$, the value of y is $(1+\delta)r_t(-r_b\varepsilon)^2 - (1+2\delta)r_t r_b^2 \varepsilon^2 = -\delta r_b r_t^2 \varepsilon^2 < 0$, thus $p \in C$. The calculation for s is similar in the case $C \in \mathcal{C}_R(k-1, \ell)$. Analogously, we have that if $C \in \mathcal{C}_B(k-1, \ell)$ and $s \in P(k, \ell-1)$, then $p, s \notin C$. As the translates in $\mathcal{C}(k, \ell-1)$ (resp. $\mathcal{C}(k-1, \ell)$) contain precisely the same points of $P(k, \ell-1)$ (resp. $P(k-1, \ell)$, as before the shift, we have obtained a geometric realization of $\mathcal{H}(k, \ell)$, and properties 1 and 2 hold.

It remains to prove that the topmost and the bottommost points of a translate $\mathcal{C}(k, \ell)$ are not covered by any other member of $\mathcal{C}(k, \ell)$ (**property 6**). Using that our construction is symmetric, it is enough to prove the statement for the topmost points. Recall that the line connecting b and t is vertical and denote their distance, the height of C, by h.

The coordinates of the topmost points of translates from $\mathcal{C}_R(k,\ell-1)$ are approximately $(r_b\varepsilon+h,(1+2\delta)r_tr_b^2\varepsilon^2+h)$. The coordinates of the topmost points of translates from $\mathcal{C}_R(k-1,\ell)$ are approximately $(-r_t\varepsilon+h,-(1+2\delta)r_br_t^2\varepsilon^2+h)$. The coordinates of the topmost points of translates from $\mathcal{C}_B(k,\ell-1)$ are approximately $(r_b\varepsilon,(1+2\delta)r_tr_b^2\varepsilon^2)$. The coordinates of the topmost points of translates from $\mathcal{C}_B(k-1,\ell)$ are approximately $(-r_t\varepsilon,-(1+2\delta)r_br_t^2\varepsilon^2)$.

If $C_1\in\mathcal{C}_R(k,\ell-1)$, by the induction hypothesis, its topmost point cannot be covered by any other $C_2\in\mathcal{C}(k,\ell-1)$. Nor can it be covered by any other translate, as the topmost points of all other translates are below it (i.e., have smaller y-coordinates). If $C_1\in\mathcal{C}_R(k-1,\ell)$, then the vector connecting it to the topmost point of some $C_2\in\mathcal{C}_R(k,\ell-1)$ is approximately the same as the vector connecting a point $s\in P(k-1,\ell)$ to the topmost point of some $C'\in\mathcal{C}_B(k,\ell-1)$. As we have seen earlier that $s\notin C'$, the same calculation shows that the topmost point of C_1 is not in C_2. If $C_1\in\mathcal{C}_R(k-1,\ell)$ and $C_2\in\mathcal{C}_B(k,\ell-1)$ or $C_2\in\mathcal{C}_B(k-1,\ell)$, then the topmost point of C_2 is lies below the topmost point of C_1. If $C_1\in\mathcal{C}_B(k,\ell-1)$ and $C_2\in\mathcal{C}(k,\ell-1)$, by induction the topmost point of C_1 is not in C_2. If $C_1\in\mathcal{C}_B(k,\ell-1)$ and $C_2\in\mathcal{C}(k-1,\ell)$, then the topmost point of C_1 is approximately at the same place as the points of $P(k,\ell-1)$ which are avoided by C_2, and the same calculation works here. Similarly, if $C_1\in\mathcal{C}_B(k-1,\ell)$ and $C_2\in\mathcal{C}(k-1,\ell)$, we can use induction, and if $C_1\in\mathcal{C}_B(k-1,\ell)$ and $C_2\in\mathcal{C}(k,\ell-1)$, we can use that the topmost point of C_1 is approximately at the same place as the points of $P(k-1,\ell)$ which are avoided by C_2, the same calculation works here. This completes the proof of property 6 and hence the lemma. □

7.6 Higher dimensions

In this section, based on parts from [50], we study the following natural extension of the problem to higher dimensions. Given a finite set of points $S\in\mathbb{R}^d$ and a family $\mathcal{F}$, can we c-color S such that every $F\in\mathcal{F}$ contains at least two colors?

For balls, however, we do not know of any counterexamples, even though a 3-dimensional *Delaunay triangulation* of any number of points might induce a complete graph (for a recent proof, see [39]). We find it quite surprising that while in the plane convex polygons admit polychromatic colorings and disks do not, in the space it might be vice versa. We could only prove the following weaker statement.

Theorem 7.17. *For every m there is a finite set of points $S\in\mathbb{R}^3$ such that for any 3-coloring of S there is a unit ball that contains exactly m points of S, all of the same color.*

Earlier such a construction with unit balls was only known for 2-colorings [65]. Although by now Theorem 7.17 became a simple corollary of Theorem 2.31, we will sketch its original proof from [50], as it shows how the planar construction from the previous sections can be improved in higher dimensions.

Abstract hypergraph.

First we define the abstract hypergraph that will be realized with unit balls. It is a straight-forward generalization of the hypergraph from Section 7.1. This time the induction will be on three parameters, k,ℓ and m. For any k,ℓ,m we define the (multi)hypergraph $\mathcal{H}(k,\ell,m) = (V(k,\ell,m), \mathcal{E}(k,\ell,m))$ recursively. The edge set $\mathcal{E}(k,\ell,m)$ will be the disjoint union of three sets, $\mathcal{E}(k,\ell,m) = \mathcal{E}_1(k,\ell,m) \uplus \mathcal{E}_2(k,\ell,m) \uplus \mathcal{E}_3(k,\ell,m)$. All edges belonging to $\mathcal{E}_1(k,\ell,m)$ will be of size k, all edges belonging to $\mathcal{E}_2(k,\ell,m)$ will be of size ℓ, and all edges belonging to $\mathcal{E}_3(k,\ell,m)$ will be of size m. We will prove that in every 3-coloring of $\mathcal{H}(k,\ell,m)$ with colors c_1, c_2 and c_3 there will be an edge in $\mathcal{E}_i(k,\ell,m)$ such that all of its vertices are colored c_i for some $i \in \{1,2,3\}$. If $k = \ell = m$, we get an m-uniform hypergraph that cannot be properly 3-colored.

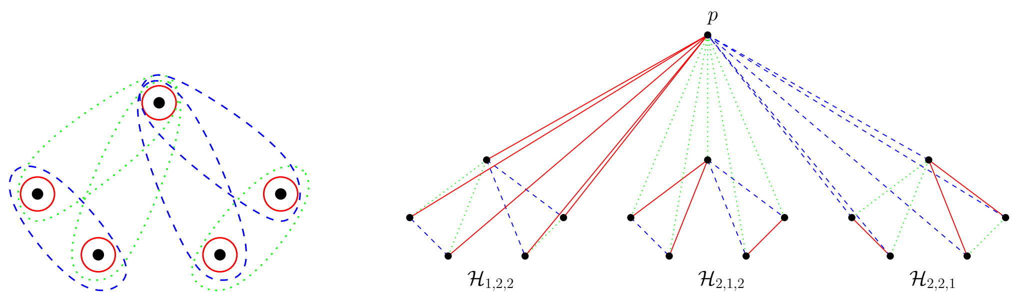


Figure 32: $\mathcal{H}(1,2,2)$ drawn with sets (left) and $\mathcal{H}(2,2,2)$ drawn as graph (right). Different colors represent the edges from the different families $\mathcal{E}_i$.

Now we give the recursive definition. Define $\mathcal{H}(1,1,1)$ as a hypergraph on one vertex with three edges containing it, with one edge in each of $\mathcal{E}_1(1,1,1)$, $\mathcal{E}_2(1,1,1)$ and $\mathcal{E}_3(1,1,1)$. If at least one of k,ℓ,m is bigger than 1, define $\mathcal{H}(k,\ell,m)$ recursively from $\mathcal{H}(k-1,\ell,m)$, $\mathcal{H}(k,\ell-1,m)$, $\mathcal{H}(k,\ell,m-1)$ by adding a "new" vertex p as follows.

$$V(k,\ell,m) = V(k-1,\ell,m) \uplus V(k,\ell-1,m) \uplus V(k,\ell,m-1) \uplus \{p\}.$$

If $k = 1$, then $\mathcal{E}_1(1,\ell,m) = \{\{v\} : v \in V(1,\ell,m)\}$, otherwise

$$\mathcal{E}_1(k,\ell,m) = \{e \cup \{p\} : e \in \mathcal{E}_1(k-1,\ell,m)\} \uplus \mathcal{E}_1(k,\ell-1,m) \uplus \mathcal{E}_1(k,\ell,m-1).$$

Similary, if $\ell = 1$, then $\mathcal{E}_2(k,1,m) = \{\{v\} : v \in V(k,1,m)\}$, otherwise

$$\mathcal{E}_2(k,\ell,m) = \{e \cup \{p\} : e \in \mathcal{E}_2(k,\ell-1,m)\} \uplus \mathcal{E}_2(k-1,\ell,m) \uplus \mathcal{E}_2(k,\ell,m-1),$$

and if $m = 1$, then $\mathcal{E}_3(k,\ell,1) = \{\{v\} : v \in V(k,\ell,1)\}$, otherwise

$$\mathcal{E}_3(k,\ell,m) = \{e \cup \{p\} : e \in \mathcal{E}_3(k,\ell,m-1)\} \uplus \mathcal{E}_3(k-1,\ell,m) \uplus \mathcal{E}_3(k,\ell-1,m).$$

Lemma 7.18. *In every 3-coloring of $\mathcal{H}(k,\ell,m)$ with colors c_1, c_2 and c_3 there is an edge in $\mathcal{E}_i(k,\ell,m)$ such that all of its vertices are colored c_i for some $i \in \{1,2,3\}$. Therefore, $\mathcal{H}(k,\ell,m)$ has no proper 3-coloring.*

The proof is a simple modification of the respective statement from [67].

Proof. If $k = \ell = m = 1$, the statement holds. Otherwise, suppose, without loss of generality, that the color of p is c_1. If $k = 1$, we are done as $\{p\} \in \mathcal{E}_1(1,\ell,m)$. Otherwise, consider the copy of $\mathcal{H}(k-1,\ell,m)$ contained in $\mathcal{H}(k,\ell,m)$. If it contains an edge in $\mathcal{E}_2(k-1,\ell,m)$ or $\mathcal{E}_3(k-1,\ell,m)$ such that its vertices are all colored c_2 or all colored c_3, respectively, we are done. Otherwise, it contains an $e \in \mathcal{E}_1(k-1,\ell,m)$ such that its vertices are all colored c_1. But then all the vertices of $(e \cup \{p\}) \in \mathcal{E}_1(k,\ell,m)$ are also all colored c_1, we are done. $\square$

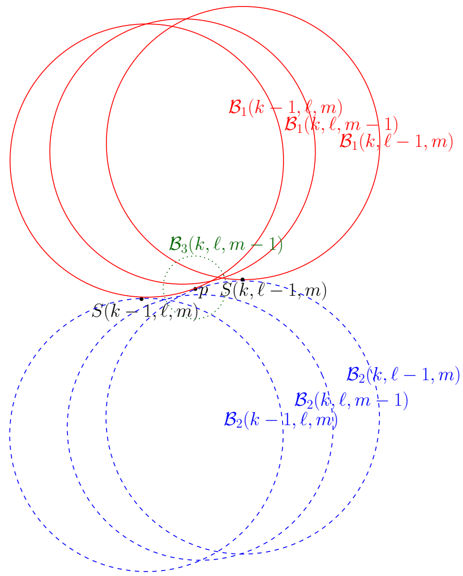


Figure 33: The intersection of $\mathcal{H}(k,\ell,m)$ with the $z=0$ plane. Point sets/collections of balls that are at distance $O(\varepsilon^5)$ are represented by a single point/ball. As the balls $\mathcal{B}_3(k-1,\ell,m)$ intersect in a $O(\varepsilon^5)$ vicinity of $S(k-1,\ell,m)$ and the balls $\mathcal{B}_3(k,\ell-1,m)$ intersect in a $O(\varepsilon^5)$ vicinity of $S(k,\ell-1,m)$, they are not drawn to avoid overcrowding the picture.

Geometric realization.

Now we sketch how to realize $\mathcal{H}(k,\ell,m)$ by unit balls in $\mathbb{R}^3$. The construction will build on the construction of [62], where the edges belonging to $\mathcal{E}_1(k,\ell,1)\cup\mathcal{E}_2(k,\ell,1)$ of $\mathcal{H}(k,\ell,1)$ were realized by unit disks.

The vertices $V(k,\ell,m)$ will be embedded as a point set, $S(k,\ell,m)$, and the edge set $\mathcal{E}_i(k,\ell,m)$ as a collection of unit balls, $\mathcal{B}_i(k,\ell,m)$, where a point is contained in a ball if and only if the respective vertex is in the respective edge. All the points of $S(k,\ell,m)$ will be placed in a small neighborhood of the origin. The centers of the balls from $\mathcal{B}_1(k,\ell,m)$, $\mathcal{B}_2(k,\ell,m)$ and $\mathcal{B}_3(k,\ell,m)$ will be close to $(0,-1,0)$, $(0,1,0)$ and $(0,0,-1)$, respectively. The realization of $\mathcal{H}(1,1,1)$ contains only one point, the origin, and one ball in each family, centered appropriately close to the required center.

Suppose that not all of k,ℓ,m are 1, and we have already realized the hypergraphs $\mathcal{H}(k-1,\ell,m)$, $\mathcal{H}(k,\ell-1,m)$ and $\mathcal{H}(k,\ell,m-1)$. Place the new point p in the origin, and shift the corresponding realizations (i.e., the point sets, $S(k-1,\ell,m)$, $S(k,\ell-1,m)$ and $S(k,\ell,m-1)$, and the collection of balls, $\mathcal{B}(k-1,\ell,m)$, $\mathcal{B}(k,\ell-1,m)$ and $\mathcal{B}(k,\ell,m-1)$) by the following vectors, where $\varepsilon=\varepsilon(k,\ell,m)$ is a small enough number, but such that $\varepsilon(k-1,\ell,m)$, $\varepsilon(k,\ell-1,m)$ and $\varepsilon(k,\ell,m-1)$ are all $O(\varepsilon^5(k,\ell,m))$.

1. Shift $\mathcal{H}(k-1,\ell,m)$ by $(2\varepsilon-1.5\varepsilon^3, 2\varepsilon^2, 0)$.
2. Shift $\mathcal{H}(k,\ell-1,m)$ by $(-2\varepsilon+1.5\varepsilon^3, -2\varepsilon^2, 0)$.
3. Shift $\mathcal{H}(k,\ell,m-1)$ by $(0,0,2\varepsilon^2)$.

For an illustration, see Figure 33.

Proposition 7.19. *The above construction realizes $\mathcal{H}(k,\ell,m)$.*

The proof of this proposition is a routine calculation, we only show some parts.

Proof. Denote by o_B the center of the ball B and denote by $dist(p,q)$ the Euclidean distance of two points p,q.

1. $p \in B \in \mathcal{B}_1(k-1,\ell,m)$:

$$dist^2(p,o_B) = (2\varepsilon - 1.5\varepsilon^3)^2 + (1-2\varepsilon^2)^2 + O(\varepsilon^5) = 1 - 2\varepsilon^4 + O(\varepsilon^5) < 1.$$

2. $p \notin B \in \mathcal{B}_1(k,\ell-1,m)$:

$$dist^2(p,o_B) = (2\varepsilon - 1.5\varepsilon^3)^2 + (1+2\varepsilon^2)^2 + O(\varepsilon^5) = 1 + 4\varepsilon^2 + O(\varepsilon^3) > 1.$$

3. $p \notin B \in \mathcal{B}_1(k,\ell,m-1)$:

$$dist^2(p,o_B) = 1^2 + (2\varepsilon^2)^2 + O(\varepsilon^5) = 1 + 4\varepsilon^4 + O(\varepsilon^5) > 1.$$

4. If $s \in S(k,\ell-1,m)$, then $s \notin B \in \mathcal{B}_1(k-1,\ell,m)$:

$$dist^2(s,o_B) = (4\varepsilon - 3\varepsilon^3)^2 + (1-4\varepsilon^2)^2 + O(\varepsilon^5) = 1 + 8\varepsilon^2 + O(\varepsilon^3) > 1.$$

5. If $s \in S(k,\ell-1,m)$, then $s \notin B \in \mathcal{B}_1(k,\ell,m-1)$:

$$dist^2(s,o_B) = (2\varepsilon - 1.5\varepsilon^3)^2 + (1-2\varepsilon^2)^2 + (2\varepsilon^2)^2 + O(\varepsilon^5) = 1 + 2\varepsilon^4 + O(\varepsilon^5) > 1.$$

6. If $s \in S(k,\ell-1,m)$, then $s \notin B \in \mathcal{B}_3(k,\ell,m-1)$:

$$dist^2(s,o_B) = (2\varepsilon - 1.5\varepsilon^3)^2 + (2\varepsilon^2)^2 + (1-2\varepsilon^2)^2 + O(\varepsilon^5) = 1 + 2\varepsilon^4 + O(\varepsilon^5) > 1.$$

7. If $s \in S(k,\ell,m-1)$, then $s \notin B \in \mathcal{B}_1(k-1,\ell,m)$:

$$dist^2(s,o_B) = (2\varepsilon - 1.5\varepsilon^3)^2 + (1-2\varepsilon^2)^2 + (2\varepsilon^2)^2 + O(\varepsilon^5) = 1 + 2\varepsilon^4 + O(\varepsilon^3) > 1.$$

The other incidences can be checked similarly and thus Proposition 7.19 follows. □

Lemma 7.18 and Proposition 7.19 imply Theorem 7.17 by selecting $k = \ell = m$, therefore this also finishes the proof of Theorem 7.17.

7.7 Special shift-chains—Proof of Theorem 7.5

Throughout this section, P denotes a fixed set of n points in the plane, no two of which have the same x-coordinate, and C is a fixed open convex set that contains a vertical upward half-line. First, we recall the Definition 2.32: of shift-chains, and introduce a special variant of them.[16]

[16] In later papers, these were called *ABA-free hypergraphs*; see [48].

Definition 7.20. *For $A \subset [n] = \{1, 2, \dots, n\}$, denote by a_i the i^{th} smallest element of A. For two equal sized sets, $A, B \subset [n]$, we write $A \leq B$ if $a_i \leq b_i$ for every i.*

An m-uniform hypergraph on the vertex set $[n]$ is called a shift-chain *if its hyperedges are totally ordered by the relation $\leq$. A shift-chain $\mathcal{H}$ is* special *if for any two hyperedges, $A, B \in \mathcal{H}$ with $A \leq B$, we have* $\max(A \setminus B) < \min(B \setminus A)$.

For any integer m and real number x, let $C(m; x)$ denote the translate of C which

a. contains exactly m points of P,

b. can be obtained from C by translating it through a vector with x-coordinate x,

c. and has minimum y-coordinate, among all translates satisfying **a** and **b**.

The union of all translates of C through every vector that has x-coordinate x is a vertical strip (or an open half-plane or the whole plane), denoted by $S(x)$. If $S(x)$ contains precisely m points for some x, then in condition **c**, the minimum y-coordinate is $y = -\infty$, and we set $C(m; x) = S(x)$. If $S(x)$ contains fewer than m points, then $C(m; x)$ is undefined.

Proposition 7.21. *Let $p_1, p_2, \dots, p_n$ denote the elements of P, listed in the increasing order of their x-coordinates. Then the m-uniform hypergraph consisting of the sets $P(x) = \{i \in [n]\,;\, p_i \in C(m; x)\}$, over all $x \in \mathbb{R}$, is a special shift-chain.*

Proof. Notice that if $x < x'$, then the boundary of $C(m; x)$ intersects the boundary of $C(m; x')$ precisely once. Therefore, every element of $(C(m; x) \setminus C(m; x')) \cap P$ is to the left of all elements of $(C(m; x') \setminus C(m; x)) \cap P$. This means that $P(x) \leq P(x')$. □

In view of the duality described at the end of the introduction, Theorem 7.5 is an immediate corollary of the following statement.

Theorem 7.22. *For any $m \geq 3$, every m-uniform special shift-chain is 2-colorable. Moreover, such a coloring can be constructed in linear time.*

An example found by Fulek [33] (depicted on Figure 2) shows that Theorem 7.22 is false without assuming that the shift-chain is special.

Proof of Theorem 7.22. The proof breaks into several simple claims. In the rest of this section, $\mathcal{H}$ denotes a fixed 3-uniform special shift-chain on $[n] = \{1, 2, \dots, n\}$. For simplicity, a hyperedge (triple) $\{a, b, c\} \in \mathcal{H}$ with $a < b < c$ will be denoted by $\{a < b < c\}$.

Claim 7.23. *If $\{a < b < c\} \in \mathcal{H}$ and $\{a' < b < c'\} \in \mathcal{H}$, then $a' = a$ or $c' = c$.*

Proof. Otherwise, $\{a < b < c\} \setminus \{a' < b < c'\} = \{a < c\}$ and $\{a' < b < c'\} \setminus \{a < b < c\} = \{a' < c'\}$ would not be separated, contradicting our assumption that $\mathcal{H}$ is special. □

Define a digraph, $D = D(\mathcal{H})$ with vertex set $[n]$ and edge set E, as follows. For any $b < c$, the directed edge $bc \in E$ if and only if there exist $a, a' \in [n]$, $a \neq a'$, such that $\{a < b < c\} \in \mathcal{H}$ and $\{a' < b < c\} \in \mathcal{H}$. Analogously, for any $a < b$, the directed edge $ba \in E$ if and only if there exist $c, c' \in [n]$, $c \neq c'$, such that $\{a < b < c\} \in \mathcal{H}$ and $\{a < b < c'\} \in \mathcal{H}$. According to Claim 7.23, the out-degree of every vertex of D is at most one. Note that an edge may appear in E with both orientations ab and ba.

Claim 7.24. *The directed graph D can be constructed by a linear time algorithm.*

Proof. $\mathcal{H}$, as any 3-uniform shift-chain on n vertices has at most $3n-8$ hyperedges. Suppose that they are listed in an arbitrary order, and process them one-by-one. Suppose the next triple is $\{a < b < c\}$.

1. If b is a middle vertex for the first time, store it, together with both of its *neighbors*, a and c.
2. If b is a middle vertex for the second time, decide if it was a or c that has been previously stored as one of its neighbors. (By Claim 7.23, we know that one of them was.) If it is a, add ba to E, if it is c, add bc.
3. Otherwise, do not add any new edge, and pass to the next triple. □

Claim 7.25. *For $a < b < c$ (or $c < b < a$) it is not possible that $ac \in E$ and $ba \in E$.*

Proof. Suppose $ac, ba \in E$. By definition, this means that there exist $\{x < a < c\} \in \mathcal{H}$ and $\{a < b < y\} \in \mathcal{H}$, for some x and some $y \neq c$. Obviously, with respect to the ordering of the triples, we have $\{x < a < c\} \prec \{a < b < y\}$. The sets $\{x < a < c\} \setminus \{a < b < y\} = \{x < c\}$ and $\{a < b < y\} \setminus \{x < a < c\} = \{b < y\}$ are not separated, because the maximal element of the first set, c, is larger than the minimal element of the second set, b. This contradicts our assumption that $\mathcal{H}$ was a *special* shift-chain. □

Claim 7.26. *For $a < b < c < d$ it is not possible that $bd \in E$ and $ca \in E$.*

Proof. This would mean that there exist $\{x < b < d\} \in \mathcal{H}$ with $x \neq a$ and $\{a < c < y\} \in \mathcal{H}$ with $y \neq d$. These two triples are disjoint, but not separated, contradicting the assumption that $\mathcal{H}$ is special. □

If a directed graph T can be obtained from a directed tree oriented toward its root r, by possibly adding one of the edges pr entering the root also with the reverse orientation rp, then it is called a *quasi-tree*. Note that in this case, we can also think of T as a quasi-tree rooted in p.

Claim 7.27. *The graph D is the vertex-disjoint union of quasi-trees.*

Proof. As no vertex of D has out-degree larger than 1, it is enough to show that D has no directed cycle of length larger than 2. Suppose there is such a directed cycle, and denote its smallest and largest elements by a and b, respectively. By Claim 7.25, we have that $ab \notin E$ and $ba \notin E$. Let $ya \in E$ and $ax \in E$ be the incoming edge and the outgoing edge of the cycle at a. Again, by Claim 7.25, we have $a < x < y < b$. There is a directed path from x to b, and along this path there is a first edge uv with $u < y$ and $v > y$. But then the edges $ya, uv \in E$ would contradict Claim 7.26, as $a < u < y < v$. □

Now we are in a position to find a 2-coloring of $\mathcal{H}$ in linear time. For every $\{a < b < c\} \in \mathcal{H}$, we will guarantee that the color of its middle vertex, b, will differ from the color of a or the color of c.

First, using breadth-first search, we properly 2-color each connected component. Hence, it will be guaranteed that if the out-degree of b is non-zero, then all triples of the form $\{a < b < c\} \in \mathcal{H}$ contain both colors. Then assign to each vertex $x \in [n]$ an edge $ab \in E$ such that $a < x < b$ or $b < x < a$, provided that such an edge exists. This can be done in linear time, but here we omit the details.

For every $\{x < y < z\} \in \mathcal{H}$ that does not yet contain both colors, its middle vertex y has out-degree zero. If there is an edge $bc \in E$ assigned to y such that $b < y < c$, then there are two different hyperedges $\{a < b < c\}, \{a' < b < c\} \in \mathcal{H}$. Either $a \neq x$ or $a' \neq x$, and thus, necessarily, we have $c = z$. We color y with the same color as b (i.e., differently from $c = z$), so that $\{x < y < z\}$ contains both colors. Note that if the in-degree of y is non-zero, then a simple case analysis shows that the only possibility is $zy \in E$. Thus, this color agrees with the color given earlier to y from its connected component.

In a similar manner, if there is an edge $ab \in E$ assigned to y such that $a < y < b$, then necessarily $a = x$, and we can color y with the same color as b (i.e., differently from $a = x$).

Finally, if there are uncolored vertices, color them in increasing order so that when y is colored, if there are $\{x < y < z\} \in \mathcal{H}$, then y gets the opposite color as x. (This step is well defined, because it follows from the fact that the out-degree of y is zero, that there is only one triple $\{x < y < z\}$ with the above property.)

This completes the proof of Theorem 7.22, as it follows that the middle vertex of any triple will have a different color from another vertex of the triple. □

7.8 Covering space with unbounded convex sets

Every open, unbounded, line-free convex set C is contained in a half-space with inner normal vector $\vec{v}$ such that for any $c \in C$, the half-line emanating from c and pointing in the direction of $\vec{v}$ lies entirely in C. We can assume, without loss of generality, that $\vec{v}$ is the unit vector $e_d = (0, 0, \dots, 0, 1)$, pointing *vertically upwards*, and that C lies in the *upper half-space*.

First, we prove Proposition 7.6, according to which every covering of $\mathbb{R}^d$ with translates of a set C satisfying the above conditions can be split into two, and hence into infinitely many, coverings. We prove a slight generalization of this statement, in which C is not required to be convex.

Proposition 7.6'. *Let C be an open set in the upper half-space of $\mathbb{R}^d$, which has the property that, for every $c \in C$, the half-line starting at c and pointing vertically upwards belongs to C. Then every covering of $\mathbb{R}^d$ with translates of C splits into two coverings.*

Proof. For any positive integer i, let B_i denote the *closed* $(d-1)$-dimensional ball of radius i around the origin in the coordinate hyperplane $x_d = 0$. Let $\mathcal{C}$ be a covering of $\mathbb{R}^d$ with translates of C. As the members of $\mathcal{C}$ cover the whole d-dimensional space, they also cover the $(d-1)$-dimensional ball $B_1 \times \{0\}$, orthogonal to the x_d-axis. This set is *compact* and the members of $\mathcal{C}$ are *open*. Therefore, there is a *finite* subfamily $\mathcal{C}_1 \subset \mathcal{C}$ which covers $B_1 \times \{0\}$. Choose a number $z_1 < 0$ such that all members of $\mathcal{C}_1$ lie strictly above the hyperplane $x_d = z_1$, and consider the $(d-1)$-dimensional ball $B_2 \times \{z_1\}$. Select a finite family $\mathcal{C}_2 \subset \mathcal{C}$ that covers this ball and a number $z_2 < z_1$ such that all members of $\mathcal{C}_2$ lie strictly above the hyperplane $x_d = z_2$. Proceeding like this, we can construct an infinite sequence of disjoint finite subfamilies $\mathcal{C}_1, \mathcal{C}_2, \dots \subset \mathcal{C}$ and a sequence of reals $z_0 := 0 > z_1 > z_2 > \dots$ tending to $-\infty$ such that $\mathcal{C}_i$ covers the $(d-1)$-dimensional ball $B_i \times \{z_{i-1}\}$.

Let p be any point of $\mathbb{R}^d$ which is at distance r from the d^{th} coordinate axis and whose d^{th} coordinate is p_d. Notice that p lies above some point of every $(d-1)$-dimensional ball $B_i \times \{z_{i-1}\}$ such that $i \geq r$ and $z_{i-1} \leq p_d$. Consequently, p is covered by the corresponding families $\mathcal{C}_i$. Hence, $\mathcal{C}_1 \cup \mathcal{C}_3 \cup \mathcal{C}_5 \cup \dots$ and $\mathcal{C}_2 \cup \mathcal{C}_4 \cup \mathcal{C}_6 \cup \dots$ are two disjoint subfamilies of $\mathcal{C}$, each of which covers the whole space. □

Next, we establish Theorem 7.7, which shows that starting from 4-dimensions, Proposition 7.6 is false if we drop the assumption that C is an *open* set.

Proof of Theorem 7.7. We have to prove that there is a convex, bounded (not open) set $C' \subset \mathbb{R}^3$ such that $\mathbb{R}^4$ can be covered by translates of $C = C' \times [0,\infty)$ so that every point of $\mathbb{R}^4$ is covered infinitely many times, but this covering cannot be decomposed into two.

The set C' will be the convex hull of $\cup_{i=1}^{\infty}(C_i \times \{\frac{1}{i^2}\})$, where each C_i is in $\mathbb{R}^2$, and thus $C_i \times \{\frac{1}{i^2}\}$ lies in the plane determined by the equation $z = \frac{1}{i^2}$ of $\mathbb{R}^3$. Each C_i is the union of an open disk, defined by the inequality $x^2 + (y - \frac{1}{i})^2 < 1$, and a part of its boundary defined as follows. A point belongs to the boundary of C_i if and only if it can be represented as $(x, \sqrt{1-x^2} + \frac{1}{i})$, where $x \in [0,1]$ and the i^{th} digit of x after the "decimal" point in binary form is 1. For each i, denote the set of such x's by C_i^*. Therefore, C_i^* is the disjoint union of 2^{i-1} closed intervals.

Note that C' is neither closed, nor open. Clearly, C' is a bounded set, as it is contained in the box $[-1,1]\times[-1,2]\times[0,1]$. Observe that for every i, the point $(0, \frac{1}{i}, \frac{1}{i^2})$, the center of the disk $C_i \times \{\frac{1}{i^2}\}$, lies the plane $x = 0$, on the parabola $z = y^2$. Hence, for each i and for every point $p \in \mathbb{R}^3$ whose third coordinate is $\frac{1}{i^2}$ and first coordinate is non-negative, p belongs to the boundary of C' if and only if it is of the form $(x, \sqrt{1-x^2} + z, z^2)$ with $x \in C_i^*$ and $z = \frac{1}{i}$. To see this, it is enough to notice that no point of this form can be obtained as a convex combination of other points in C'.

Now we describe an infinite-fold covering $\mathcal{C}$ of $\mathbb{R}^4$ with translates of C that cannot be decomposed into two coverings. Let $X = \{(x, \sqrt{1-x^2}, 0, -w) \mid x \in [0,1], w \in [0,\infty)\}$. For every point $x \notin X$, select an arbitrary translate of C that covers x and does not intersect X. (It is easy to see that such a translate always exists.) Let $\mathcal{C}$ consist of all these translates, and for every i $(i = 1, 2, \dots)$, the translate of C through the vector $(0, -\frac{1}{i}, -\frac{1}{i^2}, -i)$, denoted by $\hat{C}_i$.

Notice that the $\hat{C}_i$ covers $(x, \sqrt{1-x^2}, 0, -w) \in X$ if and only if $x \in C_i^*$ and $w \leq i$. This implies that every point of X is covered by infinitely many members of $\mathcal{C}$, because every number has a representation with infinitely many digits that are 1.

It remains to show that $\mathcal{C}$ cannot be split into two coverings. This is a direct consequence of the following statement: For any $I \subset \mathbb{N}$ for which $\mathbb{N} \setminus I$ is infinite, there is a point $(x, \sqrt{1-x^2}, 0, 0) \in X$ that is not covered by $\cup_{i \in I} \hat{C}_i$.

To prove this statement when I is *infinite*, define the i^{th} digit of x as 1 if and only if $i \notin I$. Since this is only one binary representation of x, we have $x \notin \cup_{i\in I}\hat{C}_i^*$ and $(x, \sqrt{1-x^2}, 0, 0) \notin \cup_{i \in I}\hat{C}_i$. If I is *finite*, it can be extended to an infinite set such that $\mathbb{N} \setminus I$ remains infinite. Thus, this case can be reduced to the case when I is infinite. □

7.9 Bounded coverings

We prove Theorem 7.4 in a somewhat more general form. For the proof we need the following consequence of the Lovász local lemma.

Lemma 7.28 (Erdős-Lovász [26]). *Let $k, m \geq 2$ be integers. If every edge of a hypergraph has at least m vertices and every edge intersects at most $k^{m-1}/4(k-1)^m$ other edges, then its vertices can be colored with k colors so that every edge contains at least one vertex of each color.*

Let $\mathcal{C}$ be a class of subsets of $\mathbb{R}^d$. Given n members $C_1, \ldots, C_n$ of $\mathcal{C}$, assign to each point $x \in \mathbb{R}^d$ a *characteristic vector* $c(x) = (c_1(x), \ldots, c_n(x))$, where $c_i(x) = 1$ if $x \in C_i$ and $c_i(x) = 0$ otherwise. The number of distinct characteristic vectors shows how many "pieces" $C_1, \ldots, C_n$ cut the space into. The *dual shatter function* of $\mathcal{C}$, denoted by $\pi^*_{\mathcal{C}}(n)$, is the maximum of this quantity over all n-tuples $C_1, \ldots, C_n \in \mathcal{C}$. For example, when $\mathcal{C}$ is the family of open *balls* in $\mathbb{R}^d$, it is well known that

$$\pi^*_{\mathcal{C}}(n) \leq \binom{n-1}{d} + \sum_{i=0}^{d} \binom{n}{i} \leq n^d,$$

provided that $2 \leq d \leq n$.

Theorem 7.29. *Let $\mathcal{C}$ be a class of open sets in $\mathbb{R}^d$ with diameter at most D and volume at least v. Let $\pi(n) = \pi^*_{\mathcal{C}}(n)$ denote the dual shatter function of $\mathcal{C}$, and let B^d denote the unit ball in $\mathbb{R}^d$. Then, for every positive integer m, any m-fold covering of $\mathbb{R}^d$ with members of $\mathcal{C}$ splits into two coverings, provided that no point of the space is covered more than $\frac{v}{(2D)^d Vol B^d}\pi^{-1}(2^{m-3})$ times, where $Vol B^d$ is the volume of B^d.*

Proof. Given an m-fold covering of $\mathbb{R}^d$ in which no point is covered more than M times, define a hypergraph $\mathcal{H} = (V, E)$, as follows. Let V consist of all members of $\mathcal{C}$ that participate in the covering. To each point $x \in \mathbb{R}^d$, assign a (hyper)edge $e(x)$: the set of all members of the covering that contain x. (Every edge is counted only once.) Since every point x is covered by at least m members of $\mathcal{C}$, every edge $e(x) \in E$ consists of at least m points.

Consider two edges $e(x), e(y) \in E$ with $e(x) \cap e(y) \neq \emptyset$. Then there is a member of $\mathcal{C}$ that contains both x and y, so that y must lie in the ball $B(x, D)$ of radius D around x. Hence, all members of the covering that contain y lie in the ball $B(x, 2D)$ of radius $2D$ around x. Since the volume of each of these members is at least v, and no point of $B(x, 2D)$ is covered more than M times, we obtain that $B(x, D)$ can be intersected by at most $M Vol B(x, 2D)/v = M(2D)^d Vol B^d/v$ members of the covering. By the definition of the dual shatter functions, those members of the covering that intersect $B(x, D)$ cut $B(x, D)$ into at most $\pi(M(2D)^d Vol B^d/v)$ pieces, each of which corresponds to an edge of $\mathcal{H}$. Therefore, for the maximum number N of edges of $\mathcal{H}$ that can intersect the same edge $e(x) \in E$, we have

$$N \leq \pi(M(2D)^d Vol B^d/v).$$

According to Lemma 7.28 (for $k = 2$), in order to show that the covering can be split into two, i.e., the hypergraph $\mathcal{H}$ is 2-colorable, it is sufficient to assume that $N \leq 2^{m-3}$. Comparing this with the previous inequality, the result follows. $\square$

In the special case where $\mathcal{C}$ is the class of unit balls in $\mathbb{R}^d$, we have $v = Vo\ell B^d$, $D = 2$, and so the upper bound on $\pi^*_{\mathcal{C}}(n)$ implies, $\pi^{-1}(z) \geq z^{1/d}$. Thus, we obtain Theorem 7.4 with $c_d = 2^{-2d-3/d}$.

If we want to decompose an m-fold covering into $k > 2$ coverings, then the above argument shows that it is sufficient to assume that

$$\pi(M(2D)^d Vo\ell B^d/v) \leq k^{m-1}/4(k-1)^m.$$

In case of unit balls, this holds for $M \leq c_{k,d}(1 + \frac{1}{k-1})^{m/d}$ with $c_{k,d} = k^{-1/d}4^{-d-1/d}$.

Two sets are *homothets* of each other if one can be obtained from the other by a dilation with positive coefficient followed by a translation. It is easy to see [40] that for $d = 2$, the dual shatter function of the class $\mathcal{C}$ consisting of all homothets of a fixed convex set C is at most $n^2 - n + 2 \leq n^2$, for every $n \geq 2$. In this case, Theorem 7.29 immediately implies

Corollary 7.30. *Every m-fold covering $\mathcal{C}$ of the plane with homothets of a fixed convex set can be decomposed into two coverings, provided that no point of the plane belongs to more than $2^{(m-11)/2}$ members of $\mathcal{C}$.*

Naszódi and Taschuk [59] constructed a convex set C in $\mathbb{R}^3$ such that the dual shatter function of the class of all translates of C cannot be bounded from above by any polynomial of n. Therefore, for translates of C, the above approach breaks down. We do not know how to generalize Theorem 7.4 from balls to arbitrary convex bodies in $\mathbb{R}^d$, for $d \geq 3$.

7.10 $\chi_m = 4$ for general disks

Here we prove Theorem 2.31, based on parts of [22]. In this section all disks are assumed to be open. Let $\mathcal{P}$ be a point set and let $\mathcal{D}$ be a family of disks. An important folklore observation that we will use many times is that small perturbations of the points and the disks will not change the hypergraph $\mathcal{H}(\mathcal{P}, \mathcal{D})$. To put this into more precise terms, we will say that two points are *ε-close* if their distance is less than ε and two disks/circles are *ε-close* if their centers are ε-close and the difference of their radii is also smaller than ε.

Observation 7.31. *Suppose we have a finite point set $\mathcal{P}$ and a finite set of open disks $\mathcal{D}$ such that none of the points lie on the boundary of any of the disks. Then there exists an $\varepsilon > 0$ such that replacing each disk with any ε-close disk and each point with any ε-close point will not change the hypergraph $\mathcal{H}(\mathcal{P}, \mathcal{D})$.*

First we prove the following result about stabbed disks.

Theorem 7.32 (Damásdi-Pálvölgyi [22])**.** *For any positive integer m, there exists an m-uniform hypergraph that is not two-colorable and that permits a planar realization with disks that all contain some fixed point.*
Moreover, there is a realization $(\mathcal{P}, \mathcal{D})$, where the boundary of each disk from $\mathcal{D}$ is arbitrarily close to a given circle, on this circle we are given $|\mathcal{P}|$ points, and an arbitrarily small neighborhood of each of these points contains a point from $\mathcal{P}$.

Our main lemma is the following. Combined with Observation 7.31, this gives us a way to perturb the points of $\mathcal{P}$ with changing only a small part of the hypergraph $\mathcal{H}(\mathcal{P},\mathcal{D})$.

Lemma 7.33. *If $\varepsilon > 0$, C is a circle, and $a, b_1, \ldots, b_t, c$ are points on C in this order, then there is a circle C' and points $b'_1, \ldots, b'_t$ on C' with the following properties.*

1. *C' is ε-close to C.*
2. *b'_i is ε-close to b_i for each $i \in [t]$.*
3. *C' intersects C between a and b_1, and between b_t and c.*
4. *Each b'_i is outside of C.*

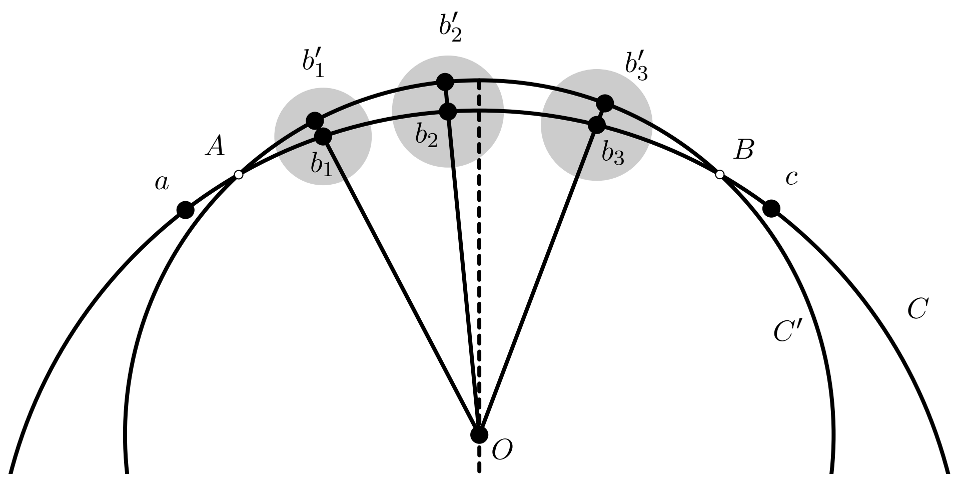


Figure 34: Illustration for Lemma 7.33.

Proof. Choose points A and B between between a and b_1, and between b_k and c, respectively (see Figure 34). Choose O on the perpendicular bisector of AB, close to the center of C. C' will be the circle centered at O passing through A and B. Project $b_1, \ldots, b_t$ onto C' using O as center. If O is close enough to the center of C, this will clearly satisfy the requirements. $\square$

Hypergraphs based on rooted trees

The following hypergraph construction was used in [65] to create several counterexamples for coloring problems.

Definition 7.34. *For every rooted tree T, let $\mathcal{H}(T)$ denote the hypergraph on vertex set $V(T)$, whose hyperedges are all sets of the following two types.*

1. *Sibling hyperedges: for each vertex $v \in V(T)$ that is not a leaf, take the set $S(v)$ of all children of v.*

2. *Descendent hyperedges: for each leaf $v \in V(T)$, take the set of all vertices along the unique path $Q(v)$ from the root to v.*

It is easy to see that $\mathcal{H}(T)$ is not two-colorable for any T. Either there is a monochromatic sibling edge, or we can follow the color of the root down to a leaf, finding a monochromatic descendent edge. We can create an m-uniform hypergraph by choosing T to be the complete m-ary tree of depth m. The non-two-colorable construction of Pach, Tardos and Tóth is also based on these hypergraphs.

Theorem 7.35 (Pach-Tardos-Tóth [65])**.** *For every rooted tree T, the hypergraph $\mathcal{H}(T)$ permits a planar realization $(\mathcal{P}, \mathcal{D})$ with disks in general position such that every disk $D \in \mathcal{D}$ has a point on its boundary that does not belong to the closure of any other disk $D' \in \mathcal{D}$.*

In order to be able to later build a point set that is not three-colorable, we will first extend Theorem 7.35 by showing that we can require the points to be close to a prescribed set of concyclic points, and require the disks to be close to the circle that contains the prescribed points. (We lose the property that every disk $D \in \mathcal{D}$ has a point on its boundary that does not belong to the closure of any other disk $D' \in \mathcal{D}$, so strictly speaking Theorem 7.36 is not a generalization of Theorem 7.35.)

Theorem 7.36. *If $\gamma > 0$, C is a circle and $q_1, q_2, \dots, q_n$ are distinct points on C, then for any rooted tree T on n vertices, the hypergraph $\mathcal{H}(T)$ permits a planar realization $(\mathcal{P}, \mathcal{D})$ with disks such that*

(I) $\mathcal{P} = \{p_1, \dots, p_n\}$, and each p_i is γ-close to q_i for all $i \in [n]$.

(II) Every $D \in \mathcal{D}$ is γ-close to C.

Theorem 7.36 clearly implies Theorem 7.32, so it is sufficient to establish Theorem 7.36.

An important property of rooted trees is that we can order their vertices in a special way. For a vertex v, let $\mathrm{Des}(v)$ denote all descendants of v. Keszegh and the author [48] have used that the vertices of T can be ordered such that

1. For each vertex $v \in V(T)$ the vertices in $S(v)$ are consecutive and they appear in the order later than v.

2. Furthermore, suppose $S(v) = \{r_1, \dots, r_k\}$ and they are in this order. Then the vertices $r_1, \dots, r_{k-1}, r_k, \mathrm{Des}(r_k), \mathrm{Des}(r_{k-1}), \dots, \mathrm{Des}(r_1)$ are ordered like this, and the rest of the vertices of T are not in this interval. (The internal order of each $\mathrm{Des}(r_i)$ is not specified by this statement.)

Call an order satisfying these properties a *siblings first order* [1] of T. Such an ordering can be constructed in a straight-forward way.

Proof of Theorem 7.36. In the planar realization of $\mathcal{H}(T)$, the vertices will correspond to the points q_i according to an (arbitrary) fix siblings first order. We start by showing that the sibling hyperedges can be easily realized and we only need to consider the descendent hyperedges.

Sibling hyperedges

From the properties of the siblings first order we know that for each v the vertices of $S(v)$ are consecutive, i.e., the points q_i corresponding to $S(v)$ are consecutive along the circle C. We apply Lemma 7.33 to find a disk that is γ-close to C, and contains exactly the points of $S(v)$. We can also ensure that no point q_i lies on the boundary of this disk. We repeat this for each $v \in V(\mathcal{H})$, until each sibling hyperedge is realized. Let $\mathcal{D}_{\text{Sib}}$ denote the set of these disks. We apply Observation 7.31 to $(\{q_1, \dots, q_n\}, \mathcal{D}_{\text{Sib}})$ to get ε_{Sib}. That is, if each point p_i is ε_{Sib}-close to q_i, then $(\mathcal{P}, \mathcal{D}_{\text{Sib}})$ will still represent the sibling hyperedges. Therefore, it is enough to show that we can realize the descendent hyperedges for every $\gamma > 0$.

Descendent hyperedges

It is useful to realize a slightly extended hypergraph. In $\mathcal{H}(T)$, we have a descendent hyperedge for each leaf. Now we will create a descendent hyperedge for non-leaf vertices too. So let $Q(v)$ contain the vertices of the path from the root to v, and for each vertex v, we will realize the hyperedge $Q(v)$. The disk realizing $Q(v)$ will be denoted by $B(v)$. Let this extended hypergraph be denoted by $\mathcal{H}'(T)$. We will realize not only $\mathcal{H}(T)$, but $\mathcal{H}'(T)$. See Figure 35 for an example where we omitted the sibling edges.

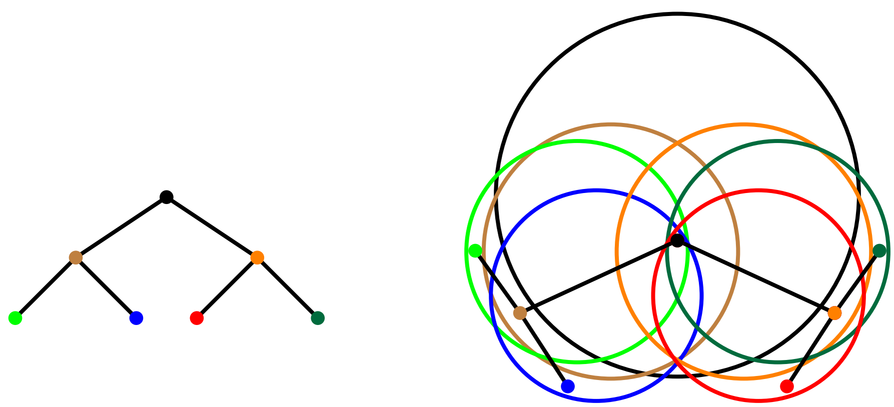


Figure 35: Disks realizing all paths from the root.

We construct a required point set by the following algorithm. Let $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ denote the vertices of T in a siblings first order. We will create the planar realization gradually, step by step. The algorithm starts by setting $p_i = q_i$ and in each step we will modify the position of some of the p_i-s. That is, we identify the vertices of the hypergraph with points in the plane, and we will update the position of the vertices until they realize the hypergraph $\mathcal{H}'(T)$.

During the algorithm, we need to change the position of the points many times without altering the hyperedges that we have already realized. For this reason, we introduce a set of fixed points, $\mathcal{P}_{\text{fix}}$, and a set of disks, $\mathcal{D}_{\text{Des}}$, that corresponds to the descendent hyperedges. Once a point is in $\mathcal{P}_{\text{fix}}$ we will not change its position anymore. Every disk that we create will be immediately added to $\mathcal{D}_{\text{Des}}$, and we will never change its position. The unfixed points, i.e., the points in $\mathcal{P} \setminus \mathcal{P}_{\text{fix}}$, will always be kept on the boundary of $\bigcup_{D \in \mathcal{D}_{\text{Des}}} D$. Furthermore, if we add $B(v)$ to $\mathcal{D}_{\text{Des}}$ in the k-th step of the algorithm, $B(v) \cap \mathcal{P}$ will remain the same after we finished the k-th step, i.e., it will contain exactly the points of the path $Q(v)$ in its interior. Moreover, all descendants of v will be on the boundary of $B(v)$ after the k-th step, but later they will be moved off.

The structure of the algorithm is the following. We go through the vertices in the siblings first order and for each we do the following. By the time we arrive at vertex p_k, the disk $B(p_k)$ representing the path $Q(p_k)$ from the root to p_k, will be already realized. For each child of p_k we will add a new disk, close to $B(p_k)$, that also realizes the same path from the root to p_k. Then we will move the points such that each new disk contains exactly one child of p_k. We have summarised the structure of the algorithm in the following pseudo code, while the phases of a step are depicted in Figure 36.

Algorithm 1: Structure of the algorithm

Set $p_i = q_i$, $\mathcal{P}_{\text{fix}} = \varnothing, \mathcal{D}_{\text{Des}} = \varnothing$.
Add the disk of C, $B(p_1)$, to $\mathcal{D}_{\text{Des}}$.
Move p_1 inside $B(p_1)$ and add p_1 to $\mathcal{P}_{\text{fix}}$.
for $k = 1$ **to** n **do**
  for *each child* r_i *of* p_k **do**
    Add a disk $B(r_i)$ representing $Q(p_k)$ to $\mathcal{D}_{\text{Des}}$. `/* Using Lemma 7.33 */`
    Move $\{r_i, \dots, r_\ell\} \cup \text{Des}(r_\ell) \cup \dots \cup \text{Des}(r_i)$ to the boundary of $B(r_i)$.
  end
  for *each child* r_i *of* p_k **do**
    Move r_i inside $B(r_i)$. `/* now` $B(r_i)$ `represents` $Q(r_i)$ `*/`
    Add r_i to $\mathcal{P}_{\text{fix}}$.
  end
end

To be able to do these steps we also need a parameter δ that will ensure that the points do not move too much and the disks are close to each other. There will be only three kinds of operations that we do during the algorithm.

a) We update the position of a vertex by moving it at a distance less than the current value of δ. If it reached its final position, we add it to $\mathcal{P}_{\text{fix}}$.

b) We add a new disk to $\mathcal{D}_{\text{Des}}$ that is δ-close to one of the disks already in $\mathcal{D}_{\text{Des}}$.

c) We decrease the value of δ.

We start with $\delta = \min(\gamma/n^2, \varepsilon_{\text{Sib}})$. (The reason for this is explained later.) Every time a disk is added or the position of a vertex is changed, we use Observation 7.31 for $(\mathcal{P}_{\text{fix}}, \mathcal{D}_{\text{Des}})$ to update δ to a smaller value, if needed. After any given update of δ, we only move points at distance less than δ. This ensures two things. Firstly, the points do not move far from their original position, and secondly, if we take a new disk that is δ-close to one of the disks, then it contains the same points of $\mathcal{P}_{\text{fix}}$.

The algorithm makes an initial adjustment on p_1, and then there is one step (the k-th step) for each point p_k. After each step the following properties will hold.

(i) Each point p_i has either reached its final position or it lies on the boundary of the disk $B(w)$ where w is the lowest ancestor of p_i for which $B(w)$ is already defined. Also, in the second case p_i does not belong to the closure of any other disk $D' \in \mathcal{D}_{\text{Des}}$.

(ii) Suppose p_j is the parent of p_i in T. Then in the j-th step the point p_i is added to $\mathcal{P}_{\text{fix}}$ and the disk $B(p_i)$ is added to $\mathcal{D}_{\text{Des}}$.

(iii) Each disk in $\mathcal{D}_{\text{Des}}$ contains those points of P that correspond to the appropriate hyperedge of $\mathcal{H}'(T)$.

During the initial adjustment we update p_1 to lie inside C but δ-close to q_1. We add the disk corresponding to the circle C to $\mathcal{D}_{\text{Des}}$. We add p_1 to $\mathcal{P}_{\text{fix}}$ and then we update δ by applying Observation 7.31 for $(\mathcal{P}_{\text{fix}}, \mathcal{D}_{\text{Des}})$. Clearly properties (i), (ii) and (iii) are satisfied.

In the k-th step we update the points in $\text{Des}(p_k)$. (If p_k is a leaf, we continue with the next step.) The process is depicted in Figure 36. From properties (i) and (ii) we know that at the start of the k-th step every point in $\text{Des}(p_k)$ lies on the boundary of $B(p_k)$, and they do not belong to any disk in $\mathcal{D}_{\text{Des}}$. Suppose $S(p_k) = \{r_1, \dots, r_\ell\}$. To maintain the three properties, we want to add the disks $B(r_1), \dots, B(r_\ell)$, and by the end of the k-th step we want to place the points of $\text{Des}(r_i)$ on the boundary of $B(r_i)$.

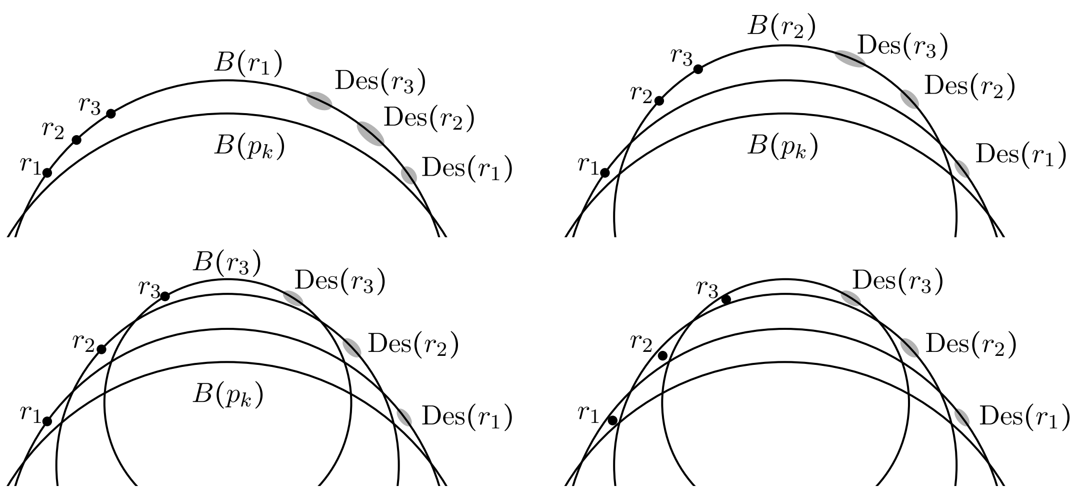


Figure 36: Phases of the k-th step of the algorithm.

To achieve this, we apply Observation 7.31 and Lemma 7.33 for each child in the following way. First apply Observation 7.31 for the points in $\text{Des}(p_k)$ and disks in $\mathcal{D}_{\text{Des}} \setminus \{B(p_k)\}$. Since the points in $\text{Des}(p_k)$ do not belong to the boundary of any other disk this is possible. We update δ to the value of ε obtained from Observation 7.31 if $\varepsilon < \delta$.

Then we apply Lemma 7.33 for the boundary of the disk $B(p_k)$, such that the b_i-s are the points $\{r_1, \dots, r_\ell\} \cup \text{Des}(r_\ell) \cup \dots \cup \text{Des}(r_1)$ and ε is chosen to be the current value of δ. The points a and c have to be chosen carefully. We know that the points in $\{r_1,\dots,r_\ell\}\cup \mathrm{Des}(r_\ell)\cup\dots\cup \mathrm{Des}(r_1)$ lie on an arc of $B(p_k)$ that is not covered by any disk in $\mathcal{D}_{\mathrm{Des}}$. We choose a and c on the two ends of this arc such that they are also not covered by any disk. Lemma 7.33 gives us a circle C', this defines $B(r_1)$, which is added to $\mathcal{D}_{\mathrm{Des}}$. The position of the points in $\{r_1,\dots,r_\ell\}\cup \mathrm{Des}(r_\ell)\cup\dots\cup \mathrm{Des}(r_1)$ is updated according to the result of Lemma 7.33. As usual, δ is also updated.

When we apply Observation 7.31 for the i-th time ($i>1$) we apply it for the points $\{r_i,\dots,r_\ell\}\cup \mathrm{Des}(r_\ell)\cup\dots\cup \mathrm{Des}(r_i)$ and disks in $\mathcal{D}_{\mathrm{Des}}\setminus\{B(r_{i-1})\}$.

When we apply Lemma 7.33 for the i-th time ($i>1$), we apply it for the boundary of $B(r_{i-1})$, such that the b_i-s are the points in $\{r_i,\dots,r_\ell\}\cup \mathrm{Des}(r_\ell)\cup\dots\cup \mathrm{Des}(r_i)$ and ε is the current value of δ. The point a is chosen between r_{i-1} and r_i. The point c is chosen after the points of $\{r_i,\dots,r_\ell\}\cup \mathrm{Des}(r_\ell)\cup\dots\cup \mathrm{Des}(r_i)$, but before the points of $\mathrm{Des}(r_{i-1})\cup\dots\cup D(r_1)$. If $\mathrm{Des}(r_{i-1})\cup\dots\cup D(r_1)$ is empty, c is chosen before the arc of $B(r_{i-1})$ reaches any other disk. In the i-th case we get $B(r_i)$, which is added to $\mathcal{D}_{\mathrm{Des}}$, and we get new positions for the points $\{r_i,\dots,r_\ell\}\cup \mathrm{Des}(r_\ell)\cup\dots\cup \mathrm{Des}(r_i)$ on $B(r_i)$. We update δ after each application of Lemma 7.33.

Finally, we finish the k-th step by moving $r_1,r_2,\dots,r_\ell$ inside $B(r_1),B(r_2),\dots,B(r_\ell)$ respectively, but at most δ far. Since r_i was lying on the boundary of $B(r_1)$, we can also ensure that they are not moved into any other disk. Then we add $r_1,\dots,r_\ell$ to $\mathcal{P}_{\mathrm{fix}}$. We update δ again by applying Observation 7.31 to $(\mathcal{P}_{\mathrm{fix}},\mathcal{D}_{\mathrm{Des}})$.

Let us see why properties $(i),(ii),(iii)$ are maintained in the course of the algorithm. Suppose they are true after the $(k-1)$-th step.

The first part of property (i) and property (ii) are maintained since we have created the disks $B(r_1),\dots,B(r_\ell)$, and by the end of the k-th step the points of $\mathrm{Des}(r_i)$ are on the boundary of $B(r_i)$. Points $r_1,\dots,r_\ell$ were added to $\mathcal{P}_{\mathrm{fix}}$.

The second part of property (i) could be violated in two ways. It could be that one of the new disks covers a point it should not. We have always chosen a and c such that this is avoided. The other possible violation is that we move a point into a disk. This is avoided, since if a point lies on the boundary of disk D before moving it, then we have updated δ for the disks in $\mathcal{D}_{\mathrm{Des}}\setminus\{D\}$ right before moving the point. Also the movement is done by Lemma 7.33 so v cannot move into D.

As for property (iii), note that the only new disks in $\mathcal{D}_{\mathrm{Des}}$ are $B(r_1),\dots,B(r_\ell)$. Since (iii) was true before the step, $B(p_k)$ contains the vertices of $Q(p_k)$ which are in $\mathcal{P}_{\mathrm{fix}}$. When we add $B(r_1)$, it is δ-close to $B(p_k)$, so it contains exactly the points in $Q(p_k)$. Similarly, when $B(r_i)$ is added, it is δ-close to $B(r_{i-1})$, so each of $B(r_1),\dots,B(r_\ell)$ contains the vertices of $Q(p_k)$. When finally we move the points $r_1,r_2,\dots,r_\ell$ inside $B(r_1),B(r_2),\dots,B(r_\ell)$, respectively, we achieve that $B(r_i)$ contains the vertices of $Q(r_i)$.

We also need to check property (iii) for the disks that were already in $\mathcal{D}_{\mathrm{Des}}$. Consider a disk $D\in\mathcal{D}_{\mathrm{Des}}$. No point inside D was moved, since they are in $\mathcal{P}_{\mathrm{fix}}$. The points in $\mathcal{P}\setminus\mathcal{P}_{\mathrm{fix}}$ remain outside of D, since r_i only moves into one of the new disks and the rest of the points remain on the boundary of $\bigcup\limits_{D\in\mathcal{D}_{\mathrm{Des}}} D$. Hence property (iii) remains true.

Finally, we need to show that the point set we constructed satisfies the requirements in Theorem 7.36. Let $\mathcal{D}$ contain the disks in $\mathcal{D}_{\mathrm{Sib}}$ and those disks in $\mathcal{D}_{\mathrm{Des}}$ that correspond to descendent edges that end at a leaf. Property (iii) and the argument for the sibling edges gives that $(\mathcal{P},\mathcal{D})$ is a planar representation of $\mathcal{H}(T)$.

For property (I), note that each point moves less than n^2 times since in the k-th step they move less than n times. We started with $\delta=\gamma/n^2$, so each point is at most $n^2\frac{\gamma}{n^2}=\gamma$ far from their original position. The first disk was given by C and each disk was taken δ-close to a previous disk, in at most n steps reaching back to the first disk. Hence, all disks are γ/n close to C, implying (II). □

A point set that is not three-colorable

Now we turn to the proof of Theorem 2.31 by realizing for every positive integer m a non-three-colorable m-uniform hypergraph with disks.

We introduce a general operation for constructing hypergraphs that are not c-colorable. Essentially the same construction was used in [1].

Definition 7.37. *Suppose we have a hypergraph $\mathcal{A}$, a hypergraph $\mathcal{B}$ and let F be an edge of $\mathcal{A}$. Then we define the hypergraph $\mathcal{A}$* extended by $\mathcal{B}$ through F*, denoted by $\mathcal{A}+_F\mathcal{B}$, as follows. We construct the new hypergraph starting from $\mathcal{A}$, by replacing F with $|V(\mathcal{B})|$ edges of the form $F\cup\{x_i\}$, where x_i is a new vertex. Then we add a set of edges such that they form the hypergraph $\mathcal{B}$ on these new vertices.*

Suppose $\mathcal{A}$ is m-uniform and $\mathcal{B}$ is $(m+1)$-uniform. Then, if we extend $\mathcal{A}$ by $\mathcal{B}$ through each edge of $\mathcal{A}$, we get a hypergraph that is $(m+1)$-uniform.

Claim 7.38. *Suppose $\mathcal{A}$ is not c-colorable and $\mathcal{B}$ is not $(c-1)$-colorable. Then any extension of $\mathcal{A}$ by $\mathcal{B}$ is not c-colorable.*

Proof. Fix an edge F of $\mathcal{A}$, and assume that we are given a c-coloring of $\mathcal{A}+_F\mathcal{B}$. Since any coloring of $\mathcal{A}$ using c-colors has a monochromatic edge, and we have only "lost" the edge F, our only chance is to color the extended hypergraph such that the vertices of F are monochromatic. These are contained in each new edge, implying that none of the new vertices have the same color as the vertices of F. But then the copy of $\mathcal{B}$ is $(c-1)$-colored, and thus one of its edges is monochromatic, a contradiction. □

For any positive integer i, let $\mathcal{G}_i$ denote the hypergraph that has i vertices and only a single edge, that contains all the i vertices. Clearly, $\mathcal{G}_1$ is not c-colorable for any c (because it has an edge with just one vertex) and $\mathcal{G}_i$ is not 1-colorable. Hence, we can build non-c-colorable hypergraphs starting from these trivial ones and using them in the extensions.

Observation 7.39. *For any rooted tree T the hypergraph $\mathcal{H}(T)$ can be built with a sequence of extensions starting from $\mathcal{G}_1$, where each extending hypergraph is one of the $\mathcal{G}_i$-s.*

Now we are ready to define the non-three-colorable hypergraphs that we will realize with points and disks. For every positive integer m let $\mathcal{H}^2(m)=\mathcal{H}(T)$ where T is the m-ary tree of depth m. Clearly, $\mathcal{H}^2(m)$ is m-uniform and not two-colorable. We define non-three-colorable m-uniform hypergraphs $\mathcal{H}^3(m)$ based on them inductively.

First we create a sequence of hypergraphs. Let $\mathcal{F}_1(m) = \mathcal{G}_1$ and let v denote the single vertex of it. For for $i > 1$, let $\mathcal{F}_i(m)$ be the hypergraph that we get if we extend each edge of $\mathcal{F}_{i-1}(m)$ that contains v by $\mathcal{H}^2(m)$. (Note that the order in which we extend the edges does not matter.) Note that $\mathcal{F}_i(m)$ has only two types of edges. There are the edges that contain v; each of these contains exactly i vertices. (These are like the descendent edges.) And there are the edges that were added in a copy of $\mathcal{H}^2(m)$, each of these contain exactly m vertices. (These are a bit like the sibling edges.) Therefore, $\mathcal{F}_m(m)$ is m-uniform. Also, by Claim 7.38 and induction, no $\mathcal{F}_i(m)$ is three-colorable.

Let $\mathcal{H}^3(m) = \mathcal{F}_m(m)$. For example, $\mathcal{F}_1(2)$ is $\mathcal{G}_1$ and $\mathcal{H}^2(2)$ is a K_3. Extending $\mathcal{G}_1$ through its single edge by a K_3 gives us a K_4, so $\mathcal{H}^3(2)$ is just K_4.

Realizing $\mathcal{H}^3(m)$

Since $\mathcal{H}^3(m)$ was built by a sequence of extensions from $\mathcal{G}_1$, it is enough to show that we can realize each extension geometrically. This step is essentially the same as in [65].

Lemma 7.40. *Suppose a hypergraph $\mathcal{A}$ is realized with $(\mathcal{P},\mathcal{D})$ and there is a set $\mathcal{E} \subset \mathcal{D}$ such that every disk $D \in \mathcal{E}$ has a point on its boundary that does not belong to the closure of any other disk $D' \in \mathcal{D}$. Then we can also realize $\mathcal{A} +_F \mathcal{H}(T)$ for any rooted tree T and any edge $F \in \mathcal{E}$, with a pair $(\mathcal{P}',\mathcal{D}')$, such that any disk $D \in \mathcal{E}$ and every new copy of F has a point on its boundary that does not belong to the closure of any other disk $D' \in \mathcal{D}'$.*

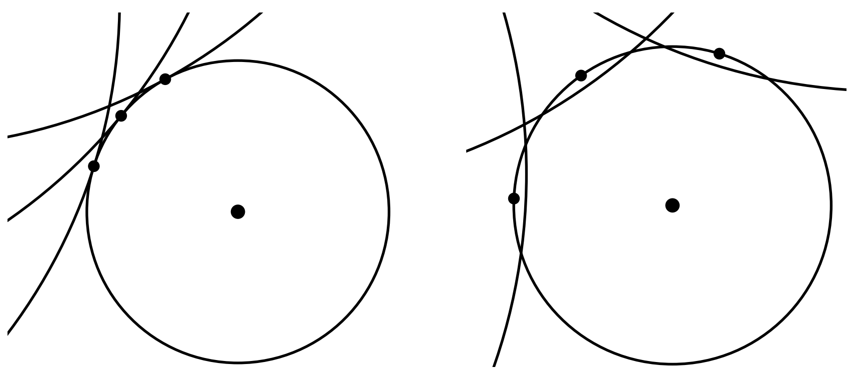


Figure 37: Realizing an extension.

Proof. Suppose $D_F \in \mathcal{D}$ is the disk realizing the edge F. Then D_F has a point p on its boundary that does not belong to the closure of any other disk $D' \in \mathcal{D}$. Take a small circle C that is tangent to D at p. If we chose the radius of C to be small enough, then C will not intersect any of the disks. Now take $n = |V(T)|$ copies of D_F and rotate them slightly around the center of C. If all the rotations are small, the resulting disks will be ε-close to D_F and by Observation 7.31 they will contain the same points as D_F. Also, if the angles of the rotations are different, then each new disk has a point on its boundary that does not belong to the closure of any other disk.

Then we enlarge each copy of D_F slightly, so that they each intersect C but some part of their boundary remains uncovered by other disks. Place the points $q_1, \dots, q_n$ in the intersections of C with these enlarged copies (one point in each), and use Theorem 7.36 to realize $\mathcal{H}(T)$. Since each disk in the realization of $\mathcal{H}(T)$ is close to C, they do not contain any point of $\mathcal{P}$. □

Using Lemma 7.40 we can construct $\mathcal{H}^3(m)$. We start from a single disk and a vertex v inside it. Then, at each extension we apply Lemma 7.40 with $\mathcal{E}$ chosen to be the set of disks that contain v. This completes the proof of Theorem 2.31.

7.11 $\chi_m = 3$ for unit disks

Here we prove Theorem 2.30, based on parts of [24]. The proof is based on three results, which we state now. Combining Theorem 2.28 with Lemma 1.5 for $k = 3$, we obtain the first result that we will use.

Corollary 7.41. *Any 5-heavy hypergraph realizable by two pseudohalfplane families is proper 3-colorable, i.e., given a finite set of points and two different pseudohalfplane arrangements in the plane, the points can be 3-colored such that every pseudohalfplane that contains at least 5 points contains two differently colored points.*

The second needed result is a theorem of Bousquet, Lochet and Thomassé, which solved a conjecture of Erdős, and of Sands, Sauer and Woodrow [72].

Theorem 7.42 (Bousquet-Lochet-Thomassé [12]). *For every k, there exists an integer $f(k)$ such that if D is a complete multidigraph whose arcs are the union of k quasi orders*[17]*, then D has a dominating set of size at most $f(k)$.*

We will need the following generalization of Theorem 7.42.

Theorem 7.43 (Damásdi-Pálvolgyi [24]). *For every pair of positive integers k and ℓ, there exist an integer $f(k,\ell)$ such that if $D = (V, A)$ is a complete multidigraph whose arcs are the union of k quasi orders $\prec_1, \dots, \prec_k$, then V contains a family of pairwise disjoint subsets S_i^j for $i \in [k]$, $j \in [\ell]$ with the following properties:*

- $|\bigcup_{i,j} S_i^j| \leq f(k,\ell)$
- *For each vertex $v \in V \setminus \bigcup_{i,j} S_i^j$ there is an $i \in [k]$ such that for each $j \in [\ell]$ there is an edge of $\prec_i$ from a vertex of S_i^j to v.*

The third result that we use is a consequence of the planar case of Hadwiger's Illumination conjecture settled by Levi [55]. Let $\mathbb{S}^{d-1}$ denote the unit sphere in $\mathbb{R}^d$. For a convex body C, let ∂C denote the boundary of C and let $int(C)$ denote its interior. A direction $u \in \mathbb{S}^{d-1}$ *illuminates* $b \in \partial C$ if $\{b + \lambda u : \lambda > 0\} \cap int(C) \neq \emptyset$.

[17] A quasi order $\prec$ is a reflexive and transitive relation, but it is not required to be antisymmetric, so $p \prec q \prec p$ is allowed, unlike for partial orders.

Conjecture 7.44. *The boundary of any convex body in $\mathbb{R}^d$ can be illuminated by 2^d or fewer directions. Furthermore, the 2^d lights are necessary if and only if the body is a parallelepiped.*

In [24] an interesting connection was made between the Illumination conjecture and pseudolines. Roughly speaking, the Illumination conjecture implies that for any convex body in the plane the boundary can be broken into three parts such that the translates of each part behave similarly to pseudolines, i.e., we get three pseudoline arrangements from the translates of the three parts.

To put this into precise terms, we need some technical definitions and statements. Fix a body C and an injective parametrization of ∂C, $\gamma : [0,1] \to \partial C$, that follows ∂C counterclockwise. For each point p of ∂C there is a set of possible tangents touching at p. Let $g(p) \subset \mathbb{S}^1$ denote the Gauss image of p, i.e., $g(p)$ is the set of unit outernormals of the tangent lines touching at p. Note that $g(p)$ is an arc of $\mathbb{S}^1$ and $g(p)$ is a proper subset of $\mathbb{S}^1$.

Let $g_+ : \partial C \to \mathbb{S}^1$ be the function that assigns to p the counterclockwise last element of $g(p)$. (See Figure 38 left.) Similarly let g_- be the function that assigns to p the clockwise last element of $g(p)$. Thus, $g(p)$ is the arc from $g_-(p)$ to $g_+(p)$. Let $|g(p)|$ denote the length of $g(p)$.

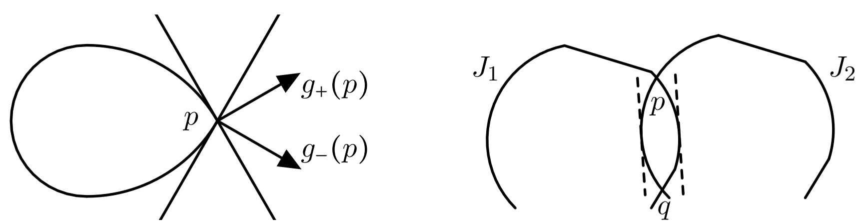


Figure 38: Extremal tangents at a boundary point (on the left) and parallel tangents on two intersecting translates (on the right).

This can be shown to imply the following, which is the last result that we need.

Lemma 7.45 (Damásdi-Pálvölgyi [24])**.** *For a convex body C, which is not a parallelogram, and an injective parametrization γ of ∂C, we can pick $0 \le t_1 < t_2 < t_3 \le 1$ and $\varepsilon > 0$ such that $|g(\gamma_{[t_1-\varepsilon,t_2+\varepsilon]})|$, $|g(\gamma_{[t_2-\varepsilon,t_3+\varepsilon]})|$ and $|g(\gamma_{[t_3-\varepsilon,t_1+\varepsilon]})|$ are each strictly smaller than π.*

Now we can start the proof of Theorem 2.30. We need to show that for any planar convex body C there is a positive integer m such that any finite point set P in the plane can be three-colored in a way that there is no translate of C containing at least m points of P, all of the same color.

If C is a parallelogram, then our proof method fails. Luckily, for translates (even for homothets) of parallelograms we know that even $\chi_m = 2$ holds [61, 2]. So from now on we assume that C is not a parallelogram.

We partition P into several parts, and for each part P_i, we divide the translates of C into three families such that two of the families each form a pseudohalfplane arrangement over P_i, while the third family will only contain translates that are automatically non-monochromatic. Then Corollary 7.41 would provide us a proper three-coloring but this is not done directly. First, we divide the plane using a grid, and then in each small square we will use Theorem 7.43 to discard some of the translates of C at the cost of a bounded number of points.

The first step of the proof is a classic divide and conquer idea [61]. We chose a constant $r = r(C)$ depending only on C and divide the plane into a grid of squares of side length r. Since each translate of C intersects some bounded number of squares, by the pidgeonhole principle we can find for any positive integer m another integer m' such that the following holds: each translate $\hat{C}$ of C that contains at least m' points intersects a square Q such that $\hat{C} \cap Q$ contains at least m points. For example, choosing $m' = m(diam(C)/r + 2)^2$ is sufficient, where $diam(C)$ denotes the diameter of C. Therefore, it is enough to show the following localized version of Theorem 2.30, since applying it separately for the points in each square of the grid provides a proper three-coloring of the whole point set.

Theorem 7.46. *There is a positive integer m such that for any convex body C there is a positive real r such that any finite point set P in the plane that lies in a square of side length r can be three-colored in a way that there is no translate of C containing at least m points of P, all of the same color.*

We will show that m can be chosen to be $f(3,2) + 13$ according to Theorem 7.43, independently of C.

Proof. We pick r the following way. First we fix an injective parametrization γ of ∂C and then fix t_1, t_2, t_3 and ε according to Lemma 7.45. Let ℓ_1, ℓ_2, ℓ_3 be the tangents of C touching at $\gamma(t_1), \gamma(t_2)$ and $\gamma(t_3)$. Let $K_{1,2}$, $K_{2,3}$, $K_{3,1}$ be the set of tri-partition cones bordered by ℓ_1, ℓ_2, ℓ_3, such that $K_{i,i+1}$ is bordered by ℓ_i on its counterclockwise side, and by ℓ_{i+1} on its clockwise side (see Figure 39 left, and note that we always treat $3+1$ as 1 in the subscript).

For a translate $\hat{C}$ of C we will denote by $\hat{\gamma}$ the translated parametrization of $\partial\hat{C}$, i.e., $\hat{\gamma}(t) = \gamma(t) + v$ if $\hat{C}$ was translated by vector v. Our aim is to choose r small enough to satisfy the following two properties for each $i \in [3]$.

(A) Let $\hat{C}$ be a translate of C, and Q be a square of side length r such that $\partial\hat{C} \cap Q \subset \hat{\gamma}_{[t_i+\varepsilon/2, t_{i+1}-\varepsilon/2]}$ (see Figure 39 right). Then for any translate K of $K_{i,i+1}$ whose apex is in $Q \cap \hat{C}$, we have $K \cap Q \subset \hat{C}$. (I.e., r is small with respect to C.)

(B) Let $\hat{C}$ be a translate of C, and Q be a square of side length r such that the curve $\hat{\gamma}_{[t_i-\varepsilon/2, t_{i+1}+\varepsilon/2]}$ intersects Q. Then $\partial\hat{C} \cap Q \subset \hat{\gamma}_{[t_i-\varepsilon, t_{i+1}+\varepsilon]}$. (I.e., r is small compared to ε.)

We show that an r satisfying properties (A) and (B) can be found for $i = 1$. The argument is the same for $i = 2$ and $i = 3$, and we can take the smallest among the three resulting r-s.

First, consider property (A). Since the sides of K are parallel to ℓ_1 and ℓ_2, the portion of K that lies "above" the segment $\overline{\hat{\gamma}(t_1)\hat{\gamma}(t_2)}$ is in $\hat{C}$. Hence, if we choose r small enough so that Q cannot intersect $\overline{\hat{\gamma}(t_1)\hat{\gamma}(t_2)}$, then property (A) is satisfied.

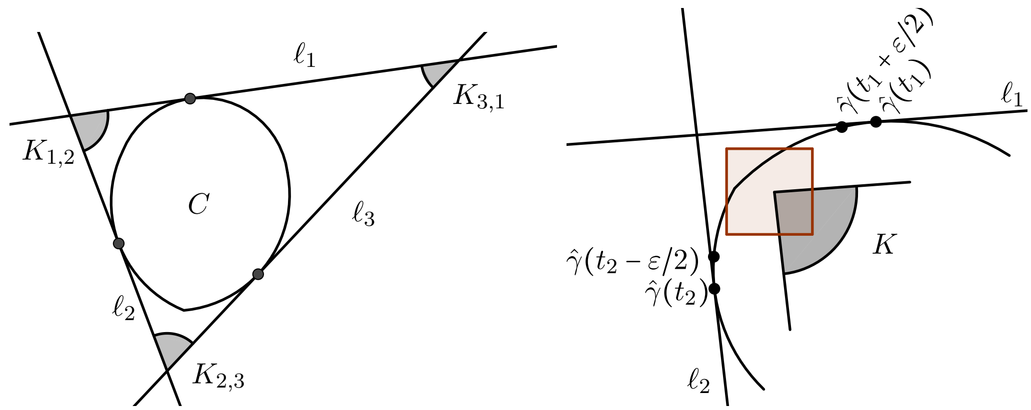


Figure 39: Selecting the cones (on the left) and Property (A) (on the right).

For example, choosing r to be smaller than $\frac{1}{\sqrt{2}}$ times the distance of the segments $\overline{\hat\gamma(t_1)\hat\gamma(t_2)}$ and $\overline{\hat\gamma(t_1+\varepsilon/2)\hat\gamma(t_2-\varepsilon/2)}$ works.

Using that γ is a continuous function on a compact set, we can pick r such that property (B) is satisfied. Therefore, there is an r satisfying properties (A) and (B).

The next step is a subdivision of the point set P using Theorem 7.43. Apply Theorem 7.43 for the graph given by the union of $\prec_{K_{1,2}}$, $\prec_{K_{2,3}}$ and $\prec_{K_{3,1}}$. It is easy to see that this is indeed a complete multidigraph on P.

We apply Theorem 7.43 with $k=3$ and $\ell=2$, resulting in subsets S_i^j for $i\in[3], j\in[2]$. Let $S=\bigcup_{i\in[3],j\in[2]} S_i^j$. For each point $p\in P\setminus S$ there is an i such that $\prec_{K_{i,i+1}}$ has an edge from a vertex of $S_{i,1}$ and $S_{i,2}$ to p. Let P_1,P_2,P_3 be the partition of $P\setminus S$ according to this i value.

We start by coloring the points of S. Color the points of $S_{1,1}\cup S_{2,1}\cup S_{3,1}$ with the first color and color the points of $S_{1,2}\cup S_{2,2}\cup S_{3,2}$ with the second color.

Note that m is at least $f(3,2)+13$. Any translate of C that contains $f(3,2)+13$ points of P must contain 5 points from either P_1,P_2 or P_3. (Note that the cone might contain all points of S). Therefore it is enough to show that for each $i\in[3]$ the points of P_i can be colored with three color such that no translate of C that contains at least 5 points of P_i is monochromatic.

Consider P_1, the proof is the same for P_2 and P_3. We divide the translates of C that intersect Q into four groups. Let $\mathcal{C}_0$ denote the translates where $\hat C\cap Q=Q$. Let $\mathcal{C}_1$ denote the translates for which $\partial\hat C\cap Q\subset\hat\gamma_{[t_1+\varepsilon/2,t_2-\varepsilon/2]}$. Let $\mathcal{C}_2$ denote the translates for which $\partial\hat C\cap Q\cap\hat\gamma_{[t_2-\varepsilon/2,t_3]}\neq\emptyset$. Let $\mathcal{C}_3$ denote the remaining translates for which $\partial\hat C\cap Q\cap\hat\gamma_{[t_3,t_1+\varepsilon/2]}\neq\emptyset$.

We do not need to worry about the translates in $\mathcal{C}_0$, as Q itself will not be monochromatic.

Take a translate $\hat C$ from $\mathcal{C}_1$ and suppose that it contains a point $p\in P_1$. By Theorem 7.43, there is an edge of $\prec_{K_{1,2}}$ from a vertex of $S_{1,1}$ to p and another edge from a vertex of $S_{1,2}$ to p. I.e., the cone $p+K_{1,2}$ contains a point from $S_{1,1}$ and another point from

$S_{1,2}$, and hence it is not monochromatic. From property (A) we know that every point in $(p+K_{1,2}) \cap P$ is also in $\hat{C}$. Therefore, $\hat{C}$ is not monochromatic.

Now consider the translates in $\mathcal{C}_2$. From property (B) we know that for these translates we have $\partial\hat{C} \cap Q \subset \hat{\gamma}_{[t_2-\varepsilon, t_3+\varepsilon]}$. By the definition of t_1, t_2 and t_3, we know that this implies that any two translates from $\mathcal{C}_2$ intersect at most once on their boundary within Q, i.e., they behave as pseudohalfplanes. To turn the translates in $\mathcal{C}_2$ into a pseudohalfplane arrangement as defined earlier, we can do as follows. For a translate $\hat{C}$, replace it with the convex set whose boundary is $\hat{\gamma}_{[t_2-\varepsilon, t_3+\varepsilon]}$ extended from its endpoints with two rays orthogonal to the segment $\overline{\hat{\gamma}(t_2-\varepsilon)\hat{\gamma}(t_3+\varepsilon)}$. This new family provides the same intersection pattern in Q and forms a pseudohalfplane arrangement. We can do the same with the translates in $\mathcal{C}_3$. Therefore, by Corollary 7.41 there is a proper three-coloring for the translates in $\mathcal{C}_2 \cup \mathcal{C}_3$. $\square$

References

[1] E. Ackerman, B. Keszegh, D. Pálvölgyi, Coloring hypergraphs defined by stabbed pseudo-disks and ABAB-free hypergraphs, SIAM J. on Discrete Math, 34, 2250–2269, 2020.

[2] E. Ackerman, B. Keszegh, M. Vizer, Coloring points with respect to squares, Discrete and Computational Geometry 58, 757–784, 2017.

[3] N. Alon, A non-linear lower bound for planar epsilon-nets, Discrete and Computational Geometry 47, 235–244, 2012.

[4] N. Alon, J. H. Spencer, The probabilistic method, third edition, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, Hoboken, NJ, 2008.

[5] G. Aloupis, J. Cardinal, S. Collette, S. Langerman, D. Orden, P. Ramos, Decomposition of multiple coverings into more parts, Discrete and Computational Geometry 44, 706–723 2010.

[6] G. Aloupis, J. Cardinal, S. Collette, S. Imahori, M. Korman, S.Langerman, O. Schwartz, S. Smorodinsky, P. Taslakian, Colorful Strips, Graphs and Combinatorics 27, 327–339, 2011.

[7] A. Asinowski, J. Cardinal, N. Cohen, S. Collette, T. Hackl, M. Hoffmann, K. Knauer, S. Langerman, M. Lason, P. Micek, G. Rote, T. Ueckerdt, Coloring Hypergraphs Induced by Dynamic Point Sets and Bottomless Rectangles, Algorithms and Data Structures, Lecture Notes in Computer Science 8037, 73–84, 2013.

[8] C. Berge, Balanced matrices, Mathematical Programming 2, 19–31, 1972.

[9] A. Biniaz, P. Bose, J. Cardinal, M. Payne, Three-Coloring Three-Dimensional Uniform Hypergraphs, in Proceedings of the 31st Canadian Conference on Computational Geometry (CCCG 2019).

[10] B. Bollobás, D. Pritchard, T. Rothvoß, A. Scott, Cover-decomposition and polychromatic numbers, SIAM J. Discrete Math. 27, 240–256, 2013.

[11] B. Bosek, S. Czerwinski, M. Debski, J. Grytczuk, Z. Lonc, P. Rzazewski, Coloring Chain Hypergraphs, in 20 Years of the Faculty of Mathematics and Information Science, A colletion of research papers in mathematics, Politechnika Warszawska, Warszawa, 2020.

[12] N. Bousquet, W. Lochet, S. Thomassé, A proof of the Erdos-Sands-Sauer-Woodrow conjecture, J. Combin. Theory Ser. B 137, 316–319, 2019.

[13] A. L. Buchsbaum, A. Efrat, S. Jain, S. Venkatasubramanian, and K. Yi, Restricted strip covering and the sensor cover problem, in Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2007), 1056–1063, ACM, New York, 2007.

[14] B. Bukh, Radon partitions in convexity spaces, arXiv:1009.2384, unpublished.

[15] J. Cardinal, K. Knauer, P. Micek, D. Pálvölgyi T. Ueckerdt, N. Varadarajan, Colouring bottomless rectangles and arborescences, to appear.

[16] J. Cardinal, K. Knauer, P. Micek, T. Ueckerdt, Making Triangles Colorful, Journal of Computational Geometry 4, 240–246, 2013.

[17] J. Cardinal, K. Knauer, P. Micek, T. Ueckerdt, Making Octants Colorful and Related Covering Decomposition Problems, SIAM J. Discrete Math. 28, 1948–1959, 2014.

[18] T. Chan, Conflict-free coloring of points with respect to rectangles and approximation algorithms for discrete independent set, in Proceedings of the twenty-eighth annual Symposium on Computational Geometry (SoCG 2012), ACM, New York, 2012.

[19] V. Chekan, T. Ueckerdt, Polychromatic Colorings of Unions of Geometric Hypergraphs, in Proceedings of Graph-Theoretic Concepts in Computer Science (WG 2022), LNCS 13453, Springer, Cham.

[20] K. Chen, A. Fiat, H. Kaplan, M. Levy, J. Matoušek, E. Mossel, J. Pach, M. Sharir, S. Smorodinsky, U. Wagner, E. Welzl, Online Conflict-Free Coloring for Intervals, SIAM J. Comput. 36, 1342–1359, 2006.

[21] X. Chen, J. Pach, M. Szegedy, G. Tardos, Delaunay graphs of point sets in the plane with respect to axis-parallel rectangles, Random Struct. Algorithms 34, 11-23, 2009.

[22] G. Damásdi, D. Pálvölgyi, Realizing an m-uniform four-chromatic hypergraph with disks, Combinatorica 42, 1027–1048 (2022).

[23] G. Damásdi, D. Pálvölgyi, Unit disks hypergraphs are three-colorable, extended abstract appeared in Extended Abstracts EuroComb 2021, Trends in Mathematics 14, 483–489, 2021.

[24] G. Damásdi, D. Pálvölgyi, Three-chromatic geometric hypergraphs, to appear in Journal of the European Mathematical Society.

[25] P. Erdős, On a combinatorial problem, Nordisk Mat. Tidskr. 11, 5–10, 1963.

[26] P. Erdős, L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, in Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. II, 609–627. Colloq. Math. Soc. János Bolyai, 10, North-Holland, Amsterdam, 1973.

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

[28] L. Esperet, G. Joret, Colouring planar graphs with three colours and no large monochromatic, Combinatorics, Probability and Computing, 23, 551–570, 2014.

[29] G. Even, Z. Lotker, D. Ron, and S. Smorodinsky, Conflict-free colorings of simple geometric regions with applications to frequency assignment in cellular networks, SIAM J. Comput. 33, 94–136, 2003.

[30] U. Feige, M. M. Halldórsson, and G. Kortsarz, Approximating the domatic number, SIAM J. Comput. 32, 172–195, 2002/03.

[31] G. Fejes Tóth, New results in the theory of packing and covering, in Convexity and its applications, 318–359, Birkhäuser, Basel, 1983.

[32] G. Fejes Tóth and W. Kuperberg, A survey of recent results in the theory of packing and covering, in New trends in discrete and computational geometry, 251–279, Algorithms Combin. 10, Springer, Berlin, 1993.

[33] R. Fulek, personal communication, 2010.

[34] R. Fulek, Coloring geometric hypergraphs defined by an arrangement of half-planes, in Proceedings of 22nd Canadian Conference on Computational Geometry (CCCG 2010), 71–74.

[35] H. Gebauer, H. Gebauer, Disproof of the neighborhood conjecture with implications to SAT, Combinatorica 32, 573–587, 2012.

[36] D. Gerbner, B. Patkós, Extremal Finite Set Theory, Chapman and Hall/CRC, 2018.

[37] M. Gibson, K. Varadarajan, Decomposing coverings and the planar sensor cover problem, Discrete and Computational Geometry 46, 313–333, 2011.

[38] W. Goddard, Acyclic colorings of planar graphs, Discrete Math. 91, 91–94, 1991.

[39] B. Gonska, A. Padrol, Neighborly inscribed polytopes and Delaunay triangulations, Advances in Geometry 16, 349–360, 2016.

[40] B. Grünbaum, Venn diagrams and independent families of sets, Math. Mag. 48, 12–23, 1975.

[41] D. Haussler, E. Welzl, ϵ-nets and simplex range queries, Discrete Comput. Geom. 2, 127–151, 1987.

[42] B. Keszegh, Coloring half-planes and bottomless rectangles, Computational Geometry: Theory and Applications 45, 495–507, 2012.

[43] B. Keszegh, Weak conflict free colorings of point sets and simple regions, in Proceedings of the 19th Canadian Conference on Computational Geometry (CCCG 2007), 97–100, 2007.

[44] B. Keszegh, N. Lemons, D. Pálvölgyi, Online and quasi-online colorings of wedges and intervals, Order 33, 389–409, 2016.

[45] B. Keszegh, D. Pálvölgyi, Octants are Cover Decomposable, Discrete and Computational Geometry 47, 598–609, 2012.

[46] B. Keszegh, D. Pálvölgyi, Octants are cover-decomposable into many coverings, Computational Geometry 47, 585–588, 2014.

[47] B. Keszegh, D. Pálvölgyi, Convex Polygons are Self-Coverable, Discrete and Computational Geometry 51, 885–895, 2014.

[48] B. Keszegh and D. Pálvölgyi, An abstract approach to polychromatic coloring: shallow hitting sets in ABA-free hypergraphs and pseudohalfplanes, Journal of Computational Geometry 10, 1–26, 2019.

[49] B. Keszegh, D. Pálvölgyi, More on Decomposing Coverings by Octants, Journal of Computational Geometry 6, 300–315, 2015.

[50] B. Keszegh, D. Pálvölgyi, Proper Coloring of Geometric Hypergraphs, Discrete and Computational Geometry, 62, 674–689, 2019.

[51] R. Klein, Concrete and Abstract Voronoi Diagrams. Lecture Notes in Computer Science 400, Springer, 1989.

[52] J. M. Kleinberg, R. Motwani, P. Raghavan, S. Venkatasubramanian, Storage management for evolving databases, 38th Annual Symposium on Foundations of Computer Science (FOCS '97), 353–362, 1997.

[53] I. Kovács, Indecomposable coverings with homothetic polygons, Discrete and Computational Geometry 53, 817–824, 2015.

[54] I. Kovács, G. Tóth, Multiple coverings with closed polygons, Electronic Journal of Combinatorics 22, 18 pages, 2015.

[55] F. W. Levi, Überdeckung eines Eibereiches durch Parallelverschiebung seines offenen Kerns, Arch. Math. (Basel), 6, 369–370, 1955.

[56] P. Mani-Levitska, J. Pach, Decomposition problems for multiple coverings with unit balls, manuscript, 1986.

[57] J. Matoušek, The determinant bound for discrepancy is almost tight, Proc. Amer. Math. Soc. 141, 451–460, 2013.

[58] E. W. Miller, On a property of families of sets, C. R. Soc. Sci. Varsovie 30, 31–38, 1937.

[59] M. Naszódi, S. Taschuk, On the transversal number and VC-dimension of families of positive homothets of a convex body, Discrete Math. 310, 77–82, 2010.

[60] J. Pach, Decomposition of multiple packing and covering, Diskrete Geometrie 2. Kolloq. Math. Inst. Univ. Salzburg, 169–178, 1980.

[61] J. Pach, Covering the plane with convex polygons, Discrete and Computational Geometry 1, 73–81, 1986.

[62] J. Pach, D. Pálvölgyi, Unsplittable coverings in the plane, Advances in Mathematics 302, 433–457, 2016.

[63] J. Pach, D. Pálvölgyi, G. Tóth, Survey on Decomposition of Multiple Coverings, in Geometry, Intuitive, Discrete, and Convex (I. Bárány, K. J. Böröczky, G. Fejes Tóth, J. Pach eds.), Bolyai Society Mathematical Studies 24, 219–257, Springer-Verlag, 2014.

[64] J. Pach, G. Tardos, Coloring axis-parallel rectangles, Journal of Combinatorial Theory. Series A 117, 776–782. 2010.

[65] J. Pach, G. Tardos, G. Tóth, Indecomposable coverings, Discrete Geometry, Canadian Mathematical Bulletin 52, 451–463, 2009.

[66] J. Pach, G. Tóth, Decomposition of multiple coverings into many parts, Computational Geometry: Theory and Applications 42, 127–133, 2009.

[67] D. Pálvölgyi, Indecomposable coverings with concave polygons, Discrete and Computational Geometry 44, 577–588, 2010.

[68] D. Pálvölgyi. Radon numbers grow linearly, Discrete and Computational Geometry 68, 165–171, 2022.

[69] D. Pálvölgyi, G. Tóth, Convex polygons are cover-decomposable, Discrete and Computational Geometry 43, 483–496, 2010.

[70] K. S. Poh, On the linear vertex-arboricity of a planar graph, J. of Graph Theory, 14, 73–75, 1990.

[71] J. Radhakrishnan, A. Srinivasan, Improved bounds and algorithms for hypergraph 2-coloring, Random Structures Algorithms 16, 4–32, 2000.

[72] B. Sands, N. Sauer, R. Woodrow, On monochromatic paths in edge-coloured digraphs, J. Combin. Theory Ser. B 33, 271–275, 1982.

[73] S. Smorodinsky, Conflict-free coloring and its applications, in Geometry, Intuitive, Discrete, and Convex (I. Bárány, K. J. Böröczky, G. Fejes Tóth, J. Pach eds.), Bolyai Society Mathematical Studies 24, Springer-Verlag, 331–389, 2014.

[74] S. Smorodinsky, Y. Yuditsky, Polychromatic coloring for half-planes, J. Combin. Theory Ser. A, 119, 146–154, 2012.

[75] G. Tardos, G. Tóth, Multiple coverings of the plane with triangles, Discrete and Computational Geometry 38, 443–450, 2007.

[76] K. Varadarajan, Weighted geometric set cover via quasi-uniform sampling, in Proceedings of the 2010 ACM International Symposium on Theory of Computing (STOC 2010), 641–647, ACM, New York, 2010.

[77] P. Winkler, on page 137 of Mathematical mind-benders, A K Peters, Wellesley, MA, 2007. See also: P. Winkler, Puzzled: covering the plane, Commun. ACM 52, 112, 2009.

A Summary Table

basic shape	translates	homothets
bottomless rectangle	$\chi_m = 2$ [42] $m_k \le 2k - 1$ [15, 48]	$\chi_m = 2$ [42] $1.67k - 2.5 \le m_k \le 3k - 2$ [7]
triangle	$\chi_m = 2$ [75] $m_k = O(k)$ [37]	$\chi_m = 2$ [45] $m_k = O(k^{4.09})$ [17] + [49]
convex polygon	$\chi_m = 2$ [69] $m_k = O(k)$ [37]	$2 \le \chi_m \le 3$ [50] $\chi_m = 2$ for squares [2]
non-convex polygon*	$3 \le \chi_m$ [67]	$3 \le \chi_m$ [67]
stabbed convex polygon	$\chi_m = 2$ [69] $m_k = O(k)$ [37]	$\chi_m = 2$ [22] $m_k = O(k)$ [22]
stabbed disk	$\chi_m = 2$ [22] $m_k = O(k)$ [22]	$\chi_m \le 3$ [1] $\chi_m = 3$ [22]
disk	$\chi_m = 3$ [23, 62]	$\chi_m = 4$ [22]
smooth body**	$\chi_m = 3$ [24, 62]	$\chi_m = 4$ [22]

Summary of some important results related about proper and polychromatic colorings of planar geometric hypergraphs. For a family, χ_m denotes the smallest positive integer k for which there is an integer m such that every finite set of points can be k-colored such that any member of the family with at least m points will contain at least two colors. For k colors, m_k denotes the smallest integer m such that any finite set of points can be k-colored such that any member of the family with at least m points will contain all k colors.

* There are some very special non-convex polygons for which $\chi_m = 2$—for the complete classification, see [69].

** The lower bound only holds for most smooth bodies—for details, see [24].