University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Katarzyna Kowalska
Student no. 371053

Approximation and Parametrized Algorithms for Segment Set Cover

Master's thesis
in COMPUTER SCIENCE

Supervisor:
dr Michał Pilipczuk
Instytut Informatyki

June 2020

Supervisor's statement

Hereby I confirm that the presented thesis was prepared under my supervision and that it fulfils the requirements for the degree of Master of Computer Science.

Date   Supervisor's signature

Author's statement

Hereby I declare that the presented thesis was prepared by me and none of its contents was obtained by means that are against the law.

The thesis has never before been a subject of any procedure of obtaining an academic degree.

Moreover, I declare that the present version of the thesis is identical to the attached electronic version.

Date   Author's signature

Abstract

The work presents a study of different geometric set cover problems. It mostly focuses on segment set cover and its connection to the polygon set cover.

Keywords

set cover, geometric set cover, FPT, W[1]-completeness, APX-completeness, PCP theorem, NP-completeness

Thesis domain (Socrates-Erasmus subject area codes)

11.3 Informatyka

Subject classification

D. Software
D.127. Blabalgorithms
D.127.6. Numerical blabalysis

Tytuł pracy w języku polskim

Algorytmy parametryzowania i trudność aproksymacji problemu pokrywania zbiorów odcinkami na płaszczyźnie

Contents

Chapter 1

Introduction

The Set Cover problem is one of the most common NP-complete problems. [tutaj referencja] We are given a family of sets and have to choose the smallest subfamily of these sets that cover all their elements. This problem naturally extends to settings were we put different weights on the sets and look for the subfamily of the minimal weight. This problem is NP-complete even without weights and if we put restrictions on what the sets can be. One of such variants is Vertex Cover problem, where sets have size 2 (they are edges in a graph).

In this work we focus on another such variant where the sets correspond to some geometric shapes and only some points of the plane have to be covered. When these shapes are rectangles with edges parallel to the axis, the problem can be proven to be W[1]-complete (solution of size k cannot be found in $n^{o}(k)$ time), APX-complete (for suffciently small $\epsilon > 0$, the problem does not admit $1+\epsilon$-approximation scheme) [refrencje].

Some of these settings are very easy. Set cover with lines parallel to one of the axis can be solved in polynomial time.

There is a notion of δ-expansions, which loosen the restrictions on geometric set cover. We allow the objects to cover the points after δ-expansion and compare the result to the original setting. This way we can produce both FPT and EPTAS for the rectangle set cover with δ-extensions [referencje].

Our contribution. In this work, we prove that unweighted geometric set cover with segments is fixed parameter tractable (FPT).

Moreover, we show that geometric set cover with segments is APX-complete for unweighted axis-parallel segments, even with 1/2-extensions. So the problem for very thin rectangles also cannot admit PTAS. Therefore, in the efficient polynomial-time approximation scheme (EPTAS) for *fat polygons* by [Har-Peled and Lee, 2009], the assumption about polygons being fat is necessary.

Finally, we show that geometric set cover with weighted segments in 3 directions is W[1]-complete. However, geometric set cover with weighted segments is FPT if we allow δ-extension.

This result is especially interesting, since it's counter-intuitive that the unweighed setting is FPT and the weighted setting is W[1]-complete. Most of such problems (like vertex cover or [wiecej przykladow]) are equally hard in both weighted and unweighted settings.

Chapter 2

Definitions

2.1. Geometric Set Cover

In the geometric set cover problem we are are given $\mathcal{P}$ – a set of objects, which are connected subsets of the plane, $\mathcal{C}$ – a set of points in the plane. The task is to choose $\mathcal{R} \subseteq \mathcal{P}$ such that every point in $\mathcal{C}$ is inside some element from $\mathcal{R}$ and $|\mathcal{R}|$ is minimized.

In the parametrized setting for a given k, we only look for a solution $\mathcal{R}$ such that $|\mathcal{R}| \leq k$.

In the weighted setting, there is some given weight function $f : \mathcal{P} \to \mathbb{R}^+$, and we would like to find a solution $\mathcal{R}$ that minimizes $\sum_{R \in \mathcal{R}} f(R)$.

2.2. Approximation

Let us recall some definitions related to optimization problems that are used in the following sections.

Definition 2.1. A **polynomial-time approximation scheme (PTAS)** for a minimization problem Π is a family of algorithms $\mathcal{A}_\epsilon$ for every $\epsilon > 0$ such that $\mathcal{A}_\epsilon$ takes an instance I of Π and in polynomial time finds a solution that is within a factor $(1 + \epsilon)$ of being optimal. That means the reported solution has weight at most $(1 + \epsilon)opt(I)$, where $opt(I)$ is the weight of an optimal solution for I.

Definition 2.2. A problem Π is **APX-hard** if assuming P $\neq$ NP, there exists $\epsilon > 0$ such that there is no polynomial-time $(1 + \epsilon)$-approximation algorithm for Π.

2.3. δ-extensions

TODO PLACEHOLDER for introductory text

δ-extensions is one of the modifications to a problem, that makes geometric set cover problem easier, it has been already used in literature (place some refrence here).

Definition 2.3 (δ-extensions for center-symmetric objects)**.** For any $\delta > 0$ and a center-symmetric object L with centre of symmetry $S = (x_s, y_s)$, the **δ-extension** of L is the object $L^{+\delta} = \{(1 + \delta) \cdot (x - x_s, y - y_s) + (x_s, y_s) : (x, y) \in L\}$, that is, $L^{+\delta}$ is the image of L under homothety centered at S with scale $(1 + \delta)$

The geometric set cover problem with δ-extensions is a modified version of geometric set cover where:

- We need to cover all the points in $\mathcal{C}$ with objects from $\{P^{+\delta} : P \in \mathcal{P}\}$ (which always include no fewer points than the objects before δ-extensions);

- We look for a solution that is no larger than the optimum solution for the original problem. Note that it does not need to be an optimal solution in the modified problem.

Formally, we have the following.

Definition 2.4 (Geometric set cover problem with δ-extensions)**.** The geometric set cover problem with δ-extensions is the problem where for an input instance $I = (\mathcal{P}, \mathcal{C})$, the task is to output a solution $\mathcal{R} \subseteq \mathcal{P}$ such that the δ-extended set $\{R^{+\delta} : R \in \mathcal{R}\}$ covers $\mathcal{C}$ and is no larger than the optimal solution for the problem without extensions, i.e. $|\mathcal{R}| \leq |opt(I)|$.

TODO: Some text

Definition 2.5 (Geometric set cover PTAS with δ-extensions)**.** We define a PTAS for geometric set cover with δ-extensions as a family of algorithms $\{\mathcal{A}_{\delta,\epsilon}\}_{\delta,\epsilon>0}$ that each takes as an input instance $I = (\mathcal{P}, \mathcal{C})$, and in polynomial-time outputs a solution $\mathcal{R} \subseteq \mathcal{P}$ such that the δ-extended set $\{R^{+\delta} : R \in \mathcal{R}\}$ covers $\mathcal{C}$ and is within a $(1 + \epsilon)$ factor of the optimal solution for this problem without extensions, i.e. $(1 + \epsilon)|\mathcal{R}| \leq |opt(I)|$.

Chapter 3

APX-hardness geometric set cover problem

In this section we analyze whether there exists a PTAS for geometric set cover for rectangles. We show that we can restrict this problem to a very simple setting: segments parallel to axes and allow (1/2)-extension, and the problem is still APX-hard. Note that segments are just degenerated rectangles with one side being very narrow.

Our results can be summarized in the following theorem and this section aims to prove it.

Theorem 3.1. ***(axis-parallel segment set cover with 1/2-extension is APX-hard).*** *Unweighted geometric set cover with axis-parallel segments in 2D (even with 1/2-extension) is APX-hard. That is, assuming $P \neq NP$, there does not exist a PTAS for this problem.*

Theorem 3.1 implies the following.

Corollary 3.1. ***(rectangle set cover is APX-hard).*** *Unweighted geometric set cover with axis-parallel rectangles (even with 1/2-extension) is APX-hard.*

We prove Theorem 3.1 by taking a problem that is APX-hard and showing a reduction. For this problem we choose MAX-(3,3)-SAT which we define below.

3.1. MAX-(3,3)-SAT and statement of reduction

Definition 3.1. MAX-3SAT is the following maximization problem. We are given a 3-CNF formula, and need to find an assignment of variables that satisfies the most clauses.

Definition 3.2. MAX-(3,3)-SAT is a variant of MAX-3SAT with an additional restriction that every variable appears in exactly 3 clauses and every clause contains exactly 3 literals of 3 different variables. Note that thus, the number of clauses is equal to the number of variables.

In our proof of Theorem 3.1 we use hardness of approximation of MAX-(3,3)-SAT proved in [Håstad, 2001] and described in Theorem 3.2 below.

Definition 3.3 (α-satisfiable MAX-3SAT formula)**.** MAX-3SAT formula with m clauses is at most α-satisfiable, if every assignment of variables satisfies no more than αm clauses.

Theorem 3.2. ***[Håstad, 2001]*** *For any $\epsilon > 0$, it is NP-hard to distinguish satisfiable (3,3)-SAT formulas from at most $(7/8 + \epsilon)$-satisfiable (3,3)-SAT formulas.*

Given an instance I of MAX-(3,3)-SAT, we construct an instance J of axis-parallel segment set cover problem, such that for a sufficiently small $\epsilon > 0$, a polynomial time $(1+\epsilon)$-approximation algorithm for J would be able to distinguish whether an instance I of MAX-(3,3)-SAT is fully satisfiable or is at most $(7/8+\epsilon)$-satisfiable. However, according to Theorem 3.2 the latter problem is NP-hard. This would imply P = NP, contradicting the assumption.

The following lemma encapsulates the properties of the reduction described in this section, and it allows us to prove Theorem 3.1.

Lemma 3.1. *Given an instance S of MAX-(3,3)-SAT with n variables and optimum value $opt(S)$, we can construct an instance I of geometric set cover with axis-parallel segments in 2D, such that:*

(1) For every solution X of instance I, there exists a solution of S that satisfies at least $15n - |X|$ clauses.

(2) For every solution of instance S that satisfies w clauses, there exists a solution of I of size $15n - w$.

(3) Every solution with $1/2$-extensions of I is also a solution to the original instance I.

Therefore, the optimum size of a solution of I is $opt(I) = 15n - opt(S)$.

We prove Lemma 3.1 in subsequent sections, but meanwhile let us prove Theorem 3.1 using Lemma 3.1 and Theorem 3.2.

Proof of Theorem 3.1. Consider any $0 < \epsilon < 1/(15 \cdot 8)$.

Let us assume that there exists a polynomial-time $(1+\epsilon)$-approximation algorithm for unweighted geometric set cover with axis-parallel segments in 2D with $(1/2)$-extensions. We construct an algorithm that solves the problem stated in Theorem 3.2, thereby proving that P = NP.

Take an instance S of MAX-(3,3)-SAT to be distinguished and construct an instance of geometric set cover I using Lemma 3.1. We now use the $(1+\epsilon)$-approximation algorithm for geometric set cover on I. Denote the size of the solution returned by this algorithm as $approx(I)$. We prove that if in S one can satisfy at most $(\frac{7}{8}+\epsilon)n$ clauses, then $approx(I) \geq 15n - (\frac{7}{8}+\epsilon)n$ and if S is satisfiable, then $approx(I) < 15n - (\frac{7}{8}+\epsilon)n$.

Assume S satisfiable. From the definition of S being satisfiable, we have:

$$opt(S) = n.$$

From Lemma 3.1 we have:

$$opt(I) = 14n.$$

Therefore,

$$approx(I) \leq (1+\epsilon)opt(I) = 14n(1+\epsilon) = 14n + 14\epsilon \cdot n =$$
$$= 14n + (15\epsilon - \epsilon)n < 14n + \left(\frac{1}{8} - \epsilon\right)n = 15n - \left(\frac{7}{8} + \epsilon\right)n.$$

Assume S is at most $\left(\frac{7}{8}+\epsilon\right)$ satisfiable. From the defintion of S being at most $\left(\frac{7}{8}+\epsilon\right)n$ satisfiable, we have:

$$opt(S) \leq \left(\frac{7}{8} + \epsilon\right)n$$

From Lemma 3.1 we have:

$$opt(I) \geq 15n - \left(\frac{7}{8} + \epsilon\right) n$$

Since a solution to I with $\frac{1}{2}$-extension is also a solution without any extention, by Lemma 3.1 (3), we have:

$$approx(I) \geq opt(I) = 15n - \left(\frac{7}{8} + \epsilon\right) n$$

Therefore, by using the assumed $(1+\epsilon)$-approximation algorithm, it is possible to distinguish the case when S is satisfiable from the case when it is at most $(\frac{7}{8}+\epsilon)n$ satisfiable, it suffices to compare $approx(I)$ with $15n - (\frac{7}{8}+\epsilon)n$. Hence, the assumed approximation algorithm cannot exist, unless P = NP. □

3.2. Reduction

We proceed to the proof of Lemma 3.1. That is, we show a reduction from the MAX-(3,3)-SAT problem to geometric set cover with segments parallel to axis. Moreover, the obtained instance of geometric set cover will be robust to 1/2-extensions (have the same optimal solution after 1/2-extension).

The construction will be composed of 2 types of gadgets: **VARIABLE-gadgets** and **CLAUSE-gadgets**. CLAUSE-gadgets will be constructed using two **OR-gadgets** connected together.

3.2.1. VARIABLE-gadget

VARIABLE-gadget is responsible for choosing the value of a variable in a CNF formula. It allows two minimum solutions of size 3 each. These two choices correspond to the two Boolean values of the variable corresponding to this gadget.

Points. Define points a, b, c, d, e, f, g, h as follows, where $L = 12n$:

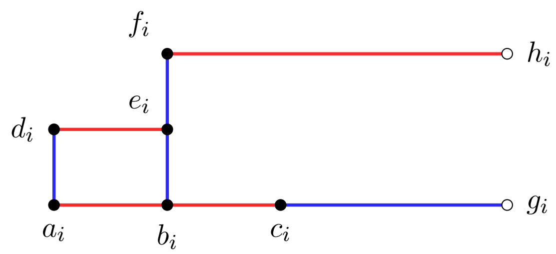

Figure 3.1: **VARIABLE-gadget.** We denote the set of points marked with black circles as $\mathsf{pointsVariable}_i$, and they need to be covered (are part of the set $\mathcal{C}$). Note that some of the points are not marked as black dots and exists only to name segments for further reference. We denote the set of red segments as $\mathsf{chooseVariable}_i^{false}$ and the set of blue segments as $\mathsf{chooseVariable}_i^{true}$.

$$\begin{array}{llll} a = (-L, 0) & b = (-\frac{2}{3}L, 0) & c = (-\frac{1}{3}L, 0) & d = (-L, 1) \\ e = (-\frac{2}{3}L, 1) & f = (-\frac{2}{3}L, 2) & g = (L, 0) & h = (L, 2) \end{array}$$

Let us define:

$$\mathsf{pointsVariable} = \{a, b, c, d, e, f\}$$

and, for any $1 \le i \le n$,

$$\mathsf{pointsVariable}_i = \mathsf{pointsVariable} + (0, 4i).$$

We denote $a_i = a + (0, 4i)$ etc.

Segments. Let us define:

$$\mathsf{chooseVariable}_i^{true} = \{(a_i, d_i), (b_i, f_i), (c_i, g_i)\}$$

$$\mathsf{chooseVariable}_i^{false} = \{(a_i, c_i), (d_i, e_i), (f_i, h_i)\}$$

$$\mathsf{segmentsVariable}_i = \mathsf{chooseVariable}_i^{true} \cup \mathsf{chooseVariable}_i^{false}$$

Lemma 3.2. *For any $1 \le i \le n$, points in $\mathsf{pointsVariable}_i$ can be covered using 3 segments from $\mathsf{segmentsVariable}_i$.*

Proof. We can use either set $\mathsf{chooseVariable}_i^{true}$ or $\mathsf{chooseVariable}_i^{false}$. $\square$

Lemma 3.3. *For any $1 \le i \le n$, points in $\mathsf{pointsVariable}_i$ can not be covered with fewer than 3 segments from $\mathsf{segmentsVariable}_i$.*

Proof. No segment of $\mathsf{segmentsVariable}_i$ covers more than one point from $\{d_i, f_i, c_i\}$, therefore $\mathsf{pointsVariable}_i$ can not be covered with fewer than 3 segments. $\square$

Lemma 3.4. *For every set $A \subseteq \mathsf{segmentsVariable}_i$ such that A covers $\mathsf{pointsVariable}_i$ and $(c_i, g_i), (f_i, h_i) \in A$, it holds that $|A| \ge 4$.*

Proof. No segment from $\mathsf{segmentsVariable}_i$ covers more than one point from $\{a_i, e_i\}$, therefore $\mathsf{pointsVariable}_i - \{c_i, f_i, g_i, h_i\}$ can not be covered with fewer than 2 segments. $\square$

3.2.2. OR-gadget

OR-segment connects input and output segments that are connected to other parts of constructions.

Output segment is part of OR-segment, but iunput is not.

For every solution $\mathcal{R}$ of the whole construction. Define $\mathcal{R}'$ as intersection of $\mathcal{R}$ and the gadget segments. Minimum solution of OR-gadget has size w, i.e. $\mathcal{R}' \le w$. *output* segments can be part of $\mathcal{R}'$ only if $input_x$ or $input_y$ are part of the chosen solution $\mathcal{R}$. If none of them are chosen, then solution containing *output* segment has weight at least $w + 1$. Therefore the following formula holds:

$$output \in \mathcal{R}' \wedge |\mathcal{R}'| = w \Rightarrow (x \in R) \vee (y \in R)$$

Only 3 points that belong to this segment: $l_{i,j}, p_{i,j}, v_{i,j}$ can be covered by segment not from the OR-gadget.

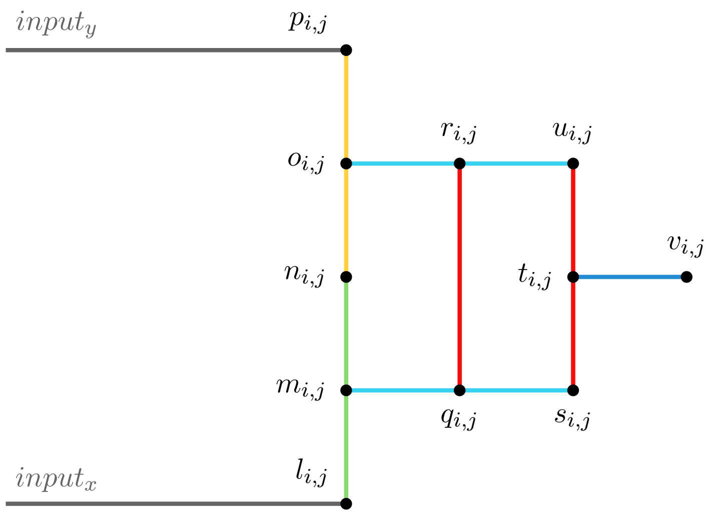

Figure 3.2: **OR-gadget.** Segments from $\mathsf{chooseOr}_{i,j}^{false}$ are red, segments from $\mathsf{chooseOr}_{i,j}^{true}$ are blue (both light blue and dark blue), segments from $\mathsf{orMoveVariable}_{i,j}$ are green and yellow. Dark blue segment is the *output* segment. Grey segments $input_x$ and $input_y$ are input segments that are not part of $\mathsf{segmentsOr}_{i,j}$.

Points.

$$\begin{array}{llll} l_0 = (0,0) & m_0 = (0,1) & n_0 = (0,2) & o_0 = (0,3) \\ p_0 = (0,4) & q_0 = (1,1) & r_0 = (1,3) & s_0 = (2,1) \\ t_0 = (2,2) & u_0 = (2,3) & v_0 = (3,2) & \end{array}$$

$$x_{i,j} = (10i + 3 + 3j, 4n + 2j)$$

For integers i, j, define $\{l_{i,j}, m_{i,j} \dots v_{i,j}\}$ as $\{l_0, m_0 \dots v_0\}$ shifted by $x_{i,j}$, ie. $l_{i,j} = l_0 + x_{i,j}$ etc.

Note that $v_{i,0} = l_{i,1}$ (see Figure 3.3)

$$\mathsf{pointsOr}_{i,j} = \{l_{i,j}, m_{i,j}, n_{i,j}, o_{i,j}, p_{i,j}, q_{i,j}, r_{i,j}, s_{i,j}, t_{i,j}, u_{i,j}\}$$

Note that $\mathsf{pointsOr}_{i,j}$ does not include the point $v_{i,j}$

Segments. We define names subsets of segments, to refer to them in lemmas.

$$\mathsf{chooseOr}_{i,j}^{false} = \{(q_{i,j}, r_{i,j}), (s_{i,j}, u_{i,j})\}$$

$$\mathsf{chooseOr}_{i,j}^{true} = \{(m_{i,j}, s_{i,j}), (o_{i,j}, u_{i,j}), (t_{i,j}, v_{i,j})\}$$

$$\mathsf{orMoveVariable}_{i,j} = \{(l_{i,j}, n_{i,j}), (n_{i,j}, p_{i,j})\}$$

Segments in OR-gadget:

$$\mathsf{segmentsOr}_{i,j} = \mathsf{chooseOr}_{i,j}^{false} \cup \mathsf{chooseOr}_{i,j}^{true} \cup \mathsf{orMoveVariable}_{i,j}$$

Lemma 3.5. *For any $1 \le i \le n, j \in \{0,1\}$ and $x \in \{l_{i,j}, p_{i,j}\}$, points in $\mathsf{pointsOr}_{,i,j} - \{x\} \cup \{v_{i,j}\}$ can be covered with 4 segments from $\mathsf{segmentsOr}_{i,j}$.*

Proof. We can do that using one segment from $\mathsf{orMoveVariable}_{i,j}$, the one that does not cover x, and all segments from $\mathsf{chooseOr}_{i,j}^{true}$. □

Lemma 3.6. *For any $1 \leq i \leq n, j \in \{0,1\}$, points in $\mathsf{pointsOr}_{i,j}$ can be covered with 4 segments from $\mathsf{segmentsOr}_{i,j}$.*

Proof. We can do that using segments from $\mathsf{orMoveVariable}_{i,j} \cup \mathsf{chooseOr}_{i,j}^{false}$. □

3.2.3. CLAUSE-gadget

A CLAUSE-gadget is responsible for determining whether variable values assigned in variable gadgets satisfy the corresponding clause in the input formula ρ. It has a minimum solution of weight w if and only if the clause is satisfied, i.e. at least one of the respective variables is assigned a correct value. Otherwise, a minimum solution of weight $w+1$. This way, by analyzing the minimum solution for the whole problem, we can tell how many clauses it was possible to satisfy in the optimum solution of ϕ.

The CLAUSE-gadgets consist of two OR-gadgets. It would be inconvenient to position the CLAUSE-gadents in between the very long variable segments. Instead, we use a simple auxiliary gadget to *transfer* whether the segment is in a solution, i.e. segments $(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,0}, z_{i,1})$. Each gadget consists of two segments $(x_{i,0}, x_{i,1}), (x_{i,1}, a)$. These are the only segments that can cover $x_{i,1}$. If $x_{i,0}$ is already covered by some other gadget, we can cover $x_{i,1}$ by the other segment covering another point from the gadget, say a. If $x_{i,0}$ is not covered, then the only way to cover $x_{i,0}$ is to use segment $(x_{i,0}, x_{i,1})$. Intuitively, the two segments *transfer* the state of $x_{i,0}$ onto a, but there are less restrictions on where a can be placed, simplifying the construction.

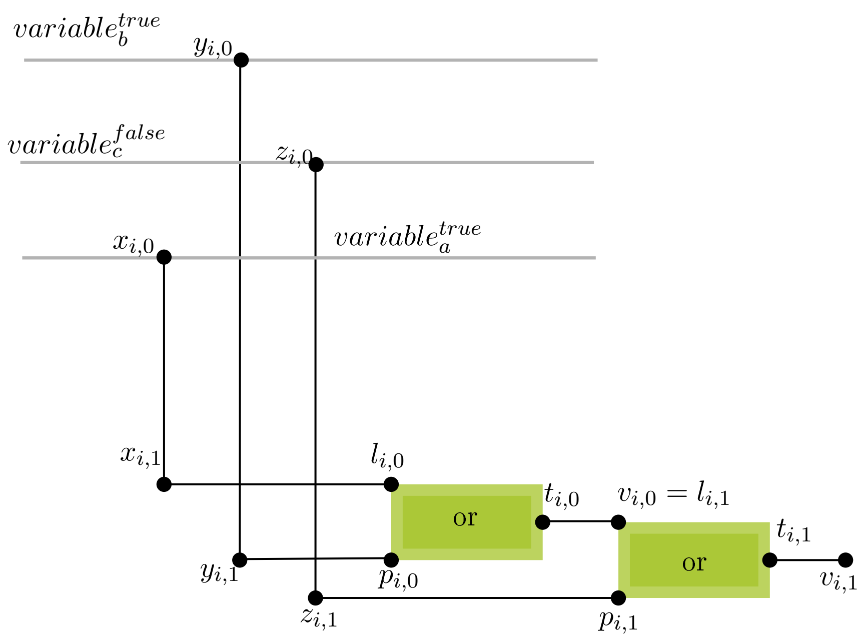

Figure 3.3: **CLAUSE-gadget for a clause** $a \vee b \vee \neg c$**.** Every green rectangle is an OR-gadget. y-coordinates of $x_{i,0}$, $y_{i,0}$ and $z_{i,0}$ depend on the variables in the i-th clause. Grey segments corresponds to the values of variables satistying the i-th clause.

Points. TODO: Rephrase it

Assuming clause $C_i = a \vee b \vee c$. For a literal w, let $idx(w)$ be the index of the variable in w, and $neg(w)$ be the Boolean value whether the variable is negated in w or not.

$$\begin{aligned} x_{i,0} &= (10i+1, 4\cdot idx(a) + 2\cdot neg(c)) & x_{i,1} &= (10i+1, 4(n+1)) \\ y_{i,0} &= (10i+2, 4\cdot idx(b) + 2\cdot neg(b)) & y_{i,1} &= (10i+2, 4(n+1)+4) \\ z_{i,0} &= (10i+3, 4\cdot idx(c) + 2\cdot neg(c)) & z_{i,1} &= (10i+3, 4(n+1)+6) \end{aligned}$$

$$\mathsf{moveVariable}_i = \{x_{i,j} : j \in \{0,1\}\} \cup \{y_{i,j} : j \in \{0,1\}\} \cup \{z_{i,j} : j \in \{0,1\}\}$$

$$\mathsf{pointsClause}_i = \mathsf{moveVariable}_i \cup \mathsf{pointsOr}_{i,0} \cup \mathsf{pointsOr}_{i,1} \cup \{v_{i,1}\}$$

Note that $v_{i,0} = l_{i,1}$.

Segments.

$$\begin{aligned} \mathsf{segmentsClause}_i \quad = \quad & \{(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,0}, z_{i,1}), (x_{i,1}, l_{i,0}), (y_{i,1}, p_{i,0}), (z_{i,1}, p_{i,1}), \} \cup \\ & \cup\ \mathsf{segmentsOr}_{i,0} \cup \mathsf{segmentsOr}_{i,1} \end{aligned}$$

Lemma 3.7. *For any $1 \le i \le n$ and $a \in \{x_{i,0}, y_{i,0}, z_{i,0}\}$, there is a set $\mathsf{solClause}_i^{true,a} \subseteq \mathsf{segmentsClause}_i$ with $|\mathsf{solClause}_i^{true,a}| = 11$ that covers all points in $\mathsf{pointsClause}_i - \{a\}$.*

Proof. For $a = x_{i,0}$ (analogous proof for $y_{i,0}$): First we use Lemma 3.5 twice with excluded $x = l_{i,0}$ and $x = l_{i,1} = v_{i,0}$, resulting with 8 segments in $\mathsf{chooseOr}_{i,0}^{true} \cup \mathsf{chooseOr}_{i,1}^{true}$ which cover all required points apart from $x_{i,1}, y_{i,0}, y_{i,1}, z_{i,0}, z_{i,1}, l_{i,0}$. We cover those using additional 3 segments: $\{(x_{i,1}, l_{i,0}), (y_{i,0}, y_{i,1}), (z_{i,0}, z_{i,1})\}$

For $a = z_{0,i}$: Using Lemma 3.6 and Lemma 3.5 with $x = p_{i,1}$, we obtain 8 segments in $\mathsf{chooseOr}_{i,0}^{false} \cup \mathsf{chooseOr}_{i,1}^{true}$ which cover all required points apart from $x_{i,0}, x_{i,1}, y_{i,0}, y_{i,1}, z_{i,1}, p_{i,1}$. We cover those using additional 3 segments: $\{(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,1}, p_{i,1})\}$. ☐

Lemma 3.8. *For any $1 \le i \le n$ there is a set $\mathsf{solClause}_i^{false} \subseteq \mathsf{segmentsClause}_i$ with $|\mathsf{solClause}_i^{false}| = 12$ that covers all points in $\mathsf{pointsClause}_i$.*

Proof. Using Lemma 3.6 twice we can cover $\mathsf{pointsOr}_{i,0}$ and $\mathsf{pointsOr}_{i,1}$ with 8 segments. To cover the remaining points we additionally use: $\{(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,0}, z_{i,1}), (t_{i,1}, v_{i,1})\}$
☐

Lemma 3.9. *For any $1 \le i \le n$:*

(1) points in $\mathsf{pointsClause}_i$ can not be covered using any subset of segments from $\mathsf{segmentsClause}_i$ of size smaller than 12;

(2) points in $\mathsf{pointsClause}_i - \{x_{i,0}, y_{i,0}, z_{i,0}\}$ can not be covered using any subset of segments from $\mathsf{segmentsClause}_i$ of size smaller than 11.

Proof of (1). No segment in $\mathsf{segmentsClause}_i$ covers more than 1 point from

$$\{x_{i,0}, y_{i,0}, z_{i,0}, l_{i,0}, p_{i,0}, q_{i,0}, u_{i,0}, v_{i,0} = l_{i,1}, p_{i,1}, q_{i,1}, u_{i,1}, v_{i,1}\}.$$

Therefore we need to use at least 12 segments. ☐

Proof of (2). We can define disjoint sets X, Y, Z such that $X \cup Y \cup Z \subseteq \mathsf{pointsClause}_i - \{x_{i,0}, y_{i,0}, z_{i,0}\}$ such that there are no segments in $\mathsf{segmentsClause}_i$ covering points from different sets. And we prove a lower bound for each of these sets. First, let:

$$X := \{x_{i,1}, y_{i,1}, z_{i,1}\}.$$

No two points in X can be covered with one segment of $\mathsf{segmentsClause}_i$, so it must be covered with 3 different segments.

$$Y = \mathsf{pointsOr}_{i,0} - \{l_{i,0}, p_{i,0}\}$$

$$Z = \mathsf{pointsOr}_{i,1} - \{l_{i,1}, p_{i,1}\}$$

For both Y and Z we can check all of the subsets of 3 segments of $\mathsf{segmentsClause}_i$ to conclude that none of them cover the considered, so both Y and Z have to be covered with disjoint sets of 4 segments each.

TODO: Funny fact, neither Y nor Z does not have independent set of size 4.

Therefore $\mathsf{pointsClause}_i - \{x_{i,0}, y_{i,0}, z_{i,0}\}$ must be covered with at least $3 + 4 + 4 = 11$ segments from $\mathsf{segmentsClause}_i$. $\square$

3.2.4. Summary

Add some smart lemmas that sets will be exclusive to each other.

Lemma 3.10. ***Robustness to 1/2-extensions***. *For every segment $s \in \mathcal{P}$, s and $s^{+1/2}$ cover the same points from $\mathcal{C}$.*

Proof. We can just check every segment. Most of the segments s are collinear only with points that lay on s, so trivially $s^{+\frac{1}{2}}$ cannot cover more points than s does.

TODO: list problematic segments here

In the same gadget: $(n_{i,j}, p_{i,j})$ does not cover $m_{i,j}$ and symmetrically. $(t_{i,j}, v_{i,j})$ does not cover $n_{i,j}$. $(o_{i,0}, u_{i,0})$ does not cover $m_{i,1}$ and symmetrically. $(y_{i,1}, p_{i,0})$ does not cover $n_{i,j}$.

From different gadgets: (b_i, f_i) after $\frac{1}{2}$-extensions does not cover b_{i+1} point.

VARIABLE-gadget's (a_i, c_i) after $\frac{1}{2}$-extensions does not cover any points $x_{i,0}, y_{i,0}$ or $z_{i,0}$ from CLAUSE-gadget.

$\square$

3.2.5. Summary of construction

We define:

$$\mathcal{C} := \bigcup_{1 \le i \le n} \mathsf{pointsVariable}_i \cup \mathsf{pointsClause}_i$$

$$\mathcal{P} := \bigcup_{1 \le i \le n} \mathsf{segmentsVariable}_i \cup \mathsf{segmentsClause}_i$$

The subsequent sections define these sets.

We prove some properties of different gadgets. Every segment for a gadget will only cover points in this gadget (will not interact with any diferent gadget), so we can prove lemmas *locally*.

TODO: y axis is increasing values downward on figures (not upwards like in normal).

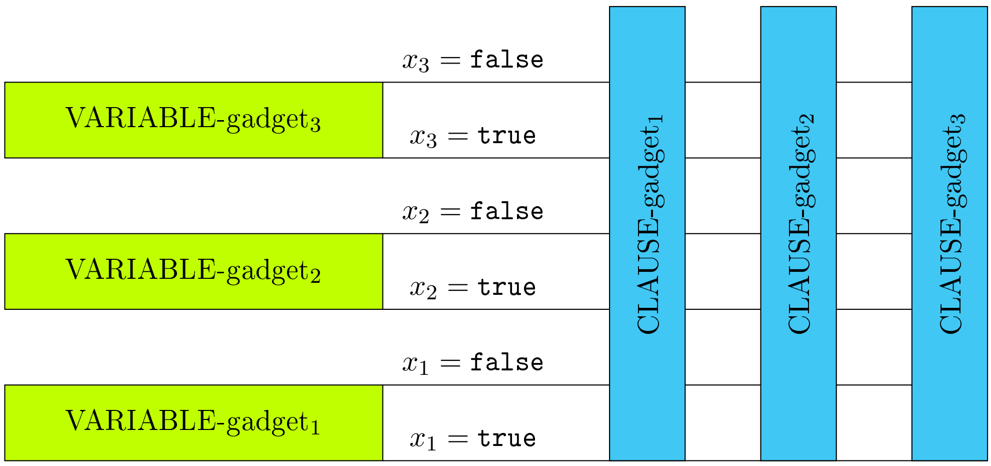

Figure 3.4: **Schema of the whole construction.**
General layout of VARIABLE-gadgets and CLAUSE-gadgets and how they interact with each other.

3.3. Construction lemmas and proof of Lemma 3.1

In order to prove Lemma 3.1 we introduce several auxiliary lemmas proving properties of the construction described in the previous section.

Consider an instance S of MAX-(3,3)-SAT of size n with optimum solution satisfying k clauses. Let us construct an instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover as described in Section 3.2 for the instance S of MAX-(3,3)-SAT.

Lemma 3.11. *Instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover admits a solution of size $15n - k$.*

Proof. Let the clauses in S be $c_1, c_2 \dots c_n$ and the variables be $x_1, x_2 \dots x_n$. Let the variable assignment in the optimum solution to S be $\phi : \{x_1, x_2 \dots x_n\} \to \{\texttt{true}, \texttt{false}\}$.

We cover every VARIABLE-gadget with solution described in Lemma 3.2, where in the i-th gadget we choose the set of segments corresponding to the value of $\phi(x_i)$.

For every clause that is satisfied, say c_i, let us name the variable that is $\texttt{true}$ in it as x_i and point corresponding to x_i in $\mathsf{pointsClause}_i$ as a. Points in $\mathsf{pointsClause}_i$ are covered with set $\mathsf{solClause}_i^{true,a}$ described in Lemma 3.7. For every clause that is not satisfied, say c_j, points in $\mathsf{pointsClause}_j$ are covered with set $\mathsf{solClause}_i^{false}$ described in Lemma 3.8.

Formally we define sets responsible for choosing variable assignment and satisfing clauses, R_i and C_i respectively, as following:

$$R_i = \begin{cases} \mathsf{chooseVariable}_i^{true} & \text{if } \phi(x_i) = \texttt{true} \\ \mathsf{chooseVariable}_i^{false} & \text{if } \phi(x_i) = \texttt{false} \end{cases}$$

$$C_i = \begin{cases} \mathsf{solClause}_i^{true,a} & \text{if } c_i \text{ satisfied by literal corresponding to point } a \\ \mathsf{solClause}_i^{false} & \text{if } c_i \text{ not satisfied} \end{cases}$$

$$\mathcal{R} = \bigcup_{i=1}^{n} \{R_i \cup C_i : 1 \le i \le n\}.$$

This set covers all the points from $\mathcal{C}$, because the sets R_i, C_i individually cover their corresponding gadgets, as proved in the respective lemmas.

All of these sets are disjoint, so the size of the obtained solution is:

$$|\mathcal{R}| = \sum_{i=1}^{n} R_i + \sum_{i=1}^{n} C_i = 3n + 11k + 12(n-k) = 15n - k. \quad \square$$

Lemma 3.12. *Suppose we have a solution $\mathcal{R}$ of the instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover. Then there exists a solution $\mathcal{R}'$, such that $|\mathcal{R}'| \leq |\mathcal{R}|$, and for each VARIABLE-gadget $\mathcal{R}'$ contains at most one of the segments (c_i, g_i) and (f_i, h_i).*

Proof. Assume that we have $\{(c_i, g_i), (f_i, h_i)\} \subseteq \mathcal{R}$ for some i. We will show how to modify $\mathcal{R}$ into $\mathcal{R}'$, such that the number of such i decreases, while $\mathcal{R}'$ is still a valid solution of $(\mathcal{C}, \mathcal{P})$, and $|\mathcal{R}'| \leq |\mathcal{R}|$. Then, by repeating this procedure, we can eventually construct a solution satisfying the property from the Lemma.

To construct $\mathcal{R}'$, we remove either (c_i, g_i) or (f_i, h_i) from $\mathcal{R}$, and then add one extra segment to make $\mathcal{R}'$ valid. Recall that the i-th VARIABLE-gadget corresponds to variable x_i in S. As every variable in S is used in exactly 3 clauses, one of the ways of setting x_i (to either `true` or `false`) must satisfy at least 2 clauses. If that setting is $x_i = \texttt{true}$, then we remove (f_i, h_i), otherwise we remove (c_i, g_i). Now, there exists at most one CLAUSE-gadget which needs adjustment to make $\mathcal{R}'$ valid; we do that by adding $(t_{j,1}, v_{j,1})$ to $\mathcal{R}'$.

TODO: Can we really just remove one segment and add another one? I'd think we need to "restructure" $\mathcal{R}$ around $\textsf{pointsVariable}_i$ (saving one segment due to Lemma 3.3 and Lemma 3.4) and then again restructure $\mathcal{R}$ around the clause that we need to fix? $\square$

Lemma 3.13. *Suppose we have a solution $\mathcal{R}$ of the instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover that is of size w. Then there exists a solution of S that satisfies at least $15n - w$ clauses.*

Proof. Let the clauses in S be $c_1, c_2 \ldots c_n$ and the variables be $x_1, x_2 \ldots x_n$. Given a solution $\mathcal{R}$ of the instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover, we use Lemma 3.12 to modify $\mathcal{R}$ such that for any i it contains at most one of (c_i, g_i) and (f_i, h_i); this may decrease the cost of $\mathcal{R}$, but that does not matter in the subsequent construction. To simplify notation, in the remainder of this proof we use $\mathcal{R}$ to refer to the modified solution.

Given $\mathcal{R}$, we construct a solution of S by constructing an assignment of variables $\phi : \{x_1, x_2 \ldots x_n\} \to \{\texttt{true}, \texttt{false}\}$ that satisfies at least $15n - w$ clauses in S.

Variables. Recall that due to Lemma 3.12, $\mathcal{R}$ contains at most one of (c_i, g_i) and (f_i, h_i). We define the value $\phi(x_i)$ for the variable x_i as follows:

$$\begin{cases} \phi(x_i) = \texttt{true} & \text{if } (c_i, g_i) \in \mathcal{R} \\ \phi(x_i) = \texttt{false} & \text{if } (f_i, h_i) \in \mathcal{R} \\ \phi(x_i) = \texttt{false} & \text{otherwise} \end{cases} \tag{3.1}$$

Moreover, from Lemma 3.3 we get $|\textsf{pointsVariable}_i \cap \mathcal{R}| \geq 3$ for every i.

Clauses For a clause $c_i = x \vee y \vee z$, $\mathcal{R}$ needs to use at least 11 segments to cover $\textsf{pointsClause}_i - \{x_{i,0}, y_{i,0}, z_{i,0}\}$ in CLAUSE-gadget (Lemma 3.9).

TODO: maybe put something with cases and names of sets as above

Moreover, if all of the points $\{x_{i,0}, y_{i,0}, z_{i,0}\}$ are not covered by the segments from $\mathcal{R} \cap \textsf{pointsVariable}_i$, then $\mathcal{R}$ needs to cover $\textsf{pointsClause}_i$ with at least 12 segments by Lemma 3.9.

TODO: Maybe remove section below, because we do this calculation at the end anyway

We covered CLAUSE-gadget with at least 11 or at least 12 segments:

$$\left| \bigcup_{i=1}^{n} \mathsf{segmentsClause}_i \cap \mathcal{R} \right| \geq 11n + a$$

where a is the number of clauses where none of the points $x_{i,0}, y_{i,0}, z_{i,0}$ were covered by $\mathcal{R} \cap \mathsf{segmentsVariable}_j$ for their respective variable x_j.

Satisfied clauses with chosen variable assignment. Consider a clause, say c_i. If none of the points $x_{i,0}, y_{i,0}, z_{i,0}$ in $\mathsf{pointsClause}_i$ were covered by segments from $\mathcal{R} \cap \mathsf{segmentsVariable}_j$, this clause is not satisfied by assignment ϕ.

If one of these points is covered by segments from VARIABLE-gadget (TODO better this or $\mathcal{R} \cap \mathsf{segmentsVariable}_j$), then denote this point as t and say it corresponds to variable x_j. Consider the cases of choosing value of $\phi(x_j)$ in equation (3.1).

If $\mathcal{R}$ contains exactly one of the segments (c_j, g_j) and (f_j, h_j), then the value $\phi(x_j)$ satisfies c_i.

If $\mathcal{R}$ contains neither (c_j, g_j) nor (f_j, h_j), then it is impossible that t is covered by segments in $\mathcal{R} \cap \mathsf{segmentsVariable}_j$.

This means that ϕ satisfies all but at most a clauses in S.

To conclude, we proved that given a solution of $(\mathcal{C}, \mathcal{P})$ of size w, we have constructed a variables assignment ϕ that satisfies at least $n - a$ clauses of S. Finally, note that

$$w \geq 3n + 11(n - a) + 12a = 3n + 11n + a = 14n + a,$$

hence

$$15n - w \leq 15n - 14n - a = n - a.$$

So ϕ satisfies at least $15n - w$ clauses of S. □

We are ready to conclude the proof of Lemma 3.1.

Proof of Lemma 3.1. By Lemma 3.11, we know that there exists a solution to $(\mathcal{C}, \mathcal{P})$ of size $15n - k$, so:

$$opt((\mathcal{C}, \mathcal{P})) \leq 15n - k.$$

Since the optimum solution of S satisfies k clauses, then according to Lemma 3.13:

$$opt((\mathcal{C}, \mathcal{P})) \geq 15n - k.$$

Therefore, the solution given by Lemma 3.11 of size $15n - k$ is an optimum solution to the instance $(\mathcal{C}, \mathcal{P})$. □

Chapter 4

Fixed-parameter tractable algorithm for geometric set cover problem

In this chapter we show two fixed-parameter tractable algorithms for geometric set cover problem in two different settings. Section 4.1 shows a fixed-parameter tractable algorithm for geometric set cover with unweighted segments. The reminder of the chapter presents a fixed-parameter tractable algorithm for geometric set cover with weighted segments with δ-extensions. We show an algorithm for the setting with δ-extensions, because the original problem with weights is W[1]-hard, as we show in Chapter 5.

We start with a shared definition for this problem. We define *extreme points* for a set of collinear points.

Definition 4.1. For a set of collinear points C, **extreme points** are the ends of the smallest segment that covers all points from set C.

If C consists of one point or is empty, then there exists 1 or 0 extreme points respectively.

4.1. Fixed-parameter tractable algorithm for unweighted segments

In this section we consider fixed-parameter tractable algorithms for unweighted geometric set cover with segments. The setting where segments are limited to be axis-parallel (or limited to a constant number of directions) has an FPT algorithm already present in literature. We present an FPT algorithm for geometric set cover with unweighted segments, where segments are in arbitrary directions.

4.1.1. Axis-parallel segments

You can find this simple algorithm in Parametrized Algorithms book [Cygan et al., 2015].

We show an $\mathcal{O}(2^k)$-time branching algorithm. In each step, the algorithm selects a point a which is not yet covered, branches to choose one of the two directions, and greedily chooses a segment a in that direction to cover. This proceeds until either all points are covered or k segments are chosen.

Let us take the point $a = (x_a, y_a)$ which is the smallest among points that are not yet covered in the lexicographic ordering of points in $\mathbb{R}^2$. We need to cover a with some of the remaining segments.

Branch over the choice of one of the coordinates (x or y); without loss of generality, let us assume we chose x. Among the segments lying on line $x = x_a$, we greedily add to the solution the one that covers the most points. As a was the smallest in the lexicographical order, all points on the line $x = x_a$ have the y-coordinate larger than y_a. Therefore, if we denote the greedily chosen segment as s, then any other segment on the line $x = x_a$ that covers a can only cover a (possibly improper) subset of points covered by s. Thus, greedily choosing s is optimal.

In each step of the algorithm we add one segment to the solution, thus each branch can stop at depth k. If no branch finds a solution, then that means a solution of size at most k does not exist.

Remark 4.1. *The same algorithm can be used for segments in d directions, where we branch over d choices of directions, and it runs in complexity $\mathcal{O}(d^k)$.*

4.1.2. Segments in arbitrary directions

In this section we consider the setting where segments are not constrained to a constant number of directions. We present a fixed-parameter tractable algorithm, parametrized by the size of the solution.

Theorem 4.1. ***(FPT for segment cover).*** *There exists an algorithm that given a family $\mathcal{P}$ of segments (in any direction), a set of points $\mathcal{C}$ and a parameter k, runs in time $k^{O(k)} \cdot (|\mathcal{C}| \cdot |\mathcal{P}|)^2$, and outputs a solution $\mathcal{R} \subseteq \mathcal{P}$ such that $|\mathcal{R}| \leq k$ and $\mathcal{R}$ covers all points in $\mathcal{C}$, or determines that such a set $\mathcal{R}$ does not exist.*

We will need the following lemmas proving properties of any instance of the problem.

Lemma 4.1. *Given an instance $(\mathcal{P}, \mathcal{C})$ of the segment cover problem, without a loss of generality we can assume that no segment covers a superset of what another segment covers. That is, for any distinct $A, B \in \mathcal{P}$, we have $A \cap C \not\subseteq B \cap C$ and $A \cap C \not\supseteq B \cap C$.*

Proof. Trivial. □

Lemma 4.2. *Given an instance $(\mathcal{P}, \mathcal{C})$ of the segment cover problem transformed by Lemma 4.1, if there exists a line L with at least $k + 1$ points on it, then there exists a subset $A \subseteq \mathcal{P}$, of size at most k, such that every solution $\mathcal{R}$ with $|\mathcal{R}| \leq k$ satisfies $|A \cap \mathcal{R}| \geq 1$. Moreover, such a subset can be found in polynomial time.*

Proof. Let us enumerate the points from $\mathcal{C}$ that lie on L as $x_1, x_2, \ldots x_t$ in the order in which they appear on L. Our proposed set is defined as:

$$A := \{\text{segment that have the leftmost point in } x : x \in x_1, x_2, \ldots x_k\}.$$

If such segment does not exist for any of these points, then set A has smaller size. We prove the lemma by contradition. Let us assume that there exists a solution $\mathcal{R}$ of size at most k, such that $\mathcal{R} \cap A = \emptyset$.

Every segment that is not collinear with L can cover at most one of the points that lie on this line. Hence if all segments from $\mathcal{R}$ were not collinear with L, then $\mathcal{R}$ would cover at most k points on line L, but L had at least $k + 1$ different points from $\mathcal{C}$ on it.

Therefore we know that one of the segments from $\mathcal{R}$ must be collinear with L and at most $k - 1$ segments can be not collinear with L. Segments from $\mathcal{R}$, that are not collinear with L can cover at most $k - 1$ points among $\{x_1, x_2, \ldots x_k\}$, therefore at least one of these points must be covered by segments from $\mathcal{R}$. We take leftmost point from $\{x_1, x_2, \ldots x_k\}$ that is covered in $\mathcal{R}$ by segment collinear with L and name it a. After transformation from Lemma

4.1 there is only one segment that starts in a, therefore this segment must be in both $\mathcal{R}$ and A.

This contradiction concludes the proof that $|A \cap \mathcal{R}| \geq 1$ for any solution $\mathcal{R}$ of size at most k. $\square$

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1.

We will prove this theorem by presenting a branching algorithm that works in desired complexity. It branches over the choice of segments to cover the lines with *a lot* of points, then solves a small instance (where every line has at most k points) by checking all possible solutions.

Algorithm. First we use Lemma 4.1.

Next, we present a recursive algorithm. Given an instance of the problem:

(1) If there exist a line with at least $k+1$ points from $\mathcal{C}$, we branch over choice of adding to the solution one of the at most k possible segments provided by Lemma 4.2; name this segment s and name set of points from $\mathcal{C}$ that lie on s as S. Then we find a solution $\mathcal{R}$ for the instance $(\mathcal{C} - S, \mathcal{P} - \{s\})$, and parameter $k-1$. We return $\mathcal{R} \cup \{s\}$.

(2) If every line has at most k points on it and $|\mathcal{C}| > k^2$, then answer `NO`.

(3) If $|\mathcal{C}| \leq k^2$, solve the problem by brute force: check all subsets of $\mathcal{P}$ of size at most k.

Correctness. Lemma 4.2 proves that at least one segment that we branch over in (1) must be present in every solution $\mathcal{R}$ with $|\mathcal{R}| \leq k$. Therefore, the recursive call can find a solution, provided there exists one.

In (2) the answer is no, because every line covers no more than k points from $\mathcal{C}$, which implies the same about every segment from $\mathcal{P}$. Under this assumption we can cover only k^2 points with a solution of size k, which is less than $|\mathcal{C}|$.

Checking all possible solutions in (3) is trivially correct.

Complexity. In the leaves of the recursion we have $|\mathcal{C}| \leq k^2$, so $|\mathcal{P}| \leq k^4$, because every segment can be uniquely identified by the two extreme points it covers (by Lemma 4.1). Therefore, there are $\binom{k^4}{k}$ possible solutions to check, each can be checked in time $O(k|\mathcal{C}|)$. Thus, (3) takes time $k^{O(k)}$.

In this branching algorithm our parameter k is decreased with every recursive call, so we have at most k levels of recursion with branching over k possibilites. Candidates to branch over can be found on each level in time $O((|\mathcal{C}| \cdot |\mathcal{P}|)^2)$.

Reduction from Lemma 4.1 can be implemented in time $O(|\mathcal{C}|^2|\mathcal{P}|)$.

It follows that the overall complexity is $O((|\mathcal{C}| \cdot |\mathcal{P}|)^2 \cdot k^{O(k)})$ $\square$

4.2. Fixed-parameter tractable algorithm for weighted segments with δ-extensions

In this section we consider a geometric set cover problem for weighted segments relaxed with δ-extensions. We show that this problem admits an FPT algorithm when parametrized with the size of the solution and δ. In the next chapter we show that the assumption about the problem being relaxed with δ-extensions is necessary: we prove that geometric set cover problem for weighted segments is W[1]-hard, which means there does not exist an FPT algorithm parametrized by solution size for it.

Theorem 4.2 (FPT for weighted segment cover with δ-extensions)**.** *There exists an algorithm that given a family $\mathcal{P}$ of n weighted segments (in any direction), a set of m points $\mathcal{C}$, and parameters k and $\delta > 0$, runs in time $f(k, \delta) \cdot (nm)^c$ for some computable function f and a constant c, and outputs a set $\mathcal{R} \subseteq \mathcal{P}$ such that $|\mathcal{R}| \leq k$ and $\mathcal{R}^{+\delta}$ covers all points in $\mathcal{C}$ and weight of $\mathcal{R}$ is not greater than weight of optimum solution of size at most k for this problem without δ-extensions, or determines that such a set $\mathcal{R}$ does not exist.*

To solve this problem we will introduce a lemma about choosing a *dense* subset of points. Dense subset of points for a set of collinear points C and parameters k and δ is a subset of C, such that if we cover it with at most k segments, these segments after δ-extensions will cover all of the points from C.

We will prove that such set of size bounded by some function $f(k, \delta)$ always exists. In later part of the section Lemma 4.3 will allow us to find a kernel for our original problem.

Definition 4.2. For a set of collinear points C, a subset $A \subseteq C$ is (k, δ)**-dense** if for any set of segments R that covers A and such that $|R| \leq k$, it holds that $R^{+\delta}$ covers C.

Lemma 4.3. *For any set of collinear points C, $\delta > 0$ and $k \geq 1$, there exists a (k, δ)-dense set $A \subseteq C$ of size at most $(2 + \frac{2}{\delta})^k$. Moreover, there exists an algorithm that computes the (k, δ)-dense set in time $O(|C| \cdot (2 + \frac{2}{\delta}))$.*

Proof of Lemma 4.3. We prove this for a fixed δ by induction over k.

Inductive hypothesis. For any set of collinear points C, there exists a set A such that:

- A is subset of C,
- A is (ℓ, δ)-dense for every $1 \leq \ell \leq k$,
- $|A| \leq (2 + \frac{2}{\delta})^k$,
- extreme points from C are in A.

Base case for $k = 1$**.** It is sufficient that A consists of extreme points of C.

If they are covered with one segment, it must be a segment that includes the extreme points from C, so it covers the whole set C.

There are at most 2 extreme points in C and $2 < 2 + \frac{2}{\delta}$.

Inductive step. Assuming inductive hypothesis for any set of collinear points C and for parameter k, we will prove it for $k + 1$.

Let s be the minimal segment that includes all points from C. That is, the extreme points of C are endpoints of s.

We define $M = \lceil 1 + \frac{2}{\delta} \rceil$ subsegments of s by splitting s into M closed segments of equal length. We name these segments v_i and $|v_i| = \frac{|s|}{M}$ for each $1 \leq i \leq M$.

Let C_i be the subset of C consisting of points laying on v_i.

Let t_i be the segment with endpoints being the extreme points of C_i. It might be a degenerate segment if C_i consists of one point or empty if C_i is empty.

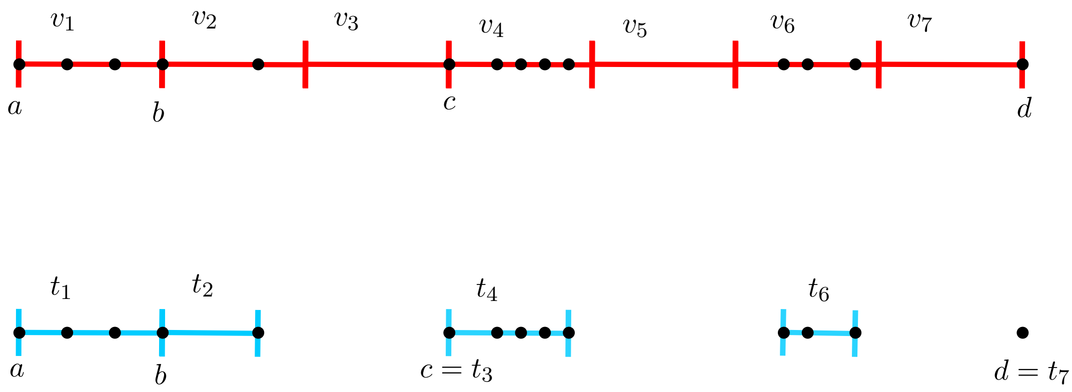

Figure 4.1: **Example of segments v_i and t_i.**
Example for $M = 7$ and some set of points (marked with black circles). Upper picture shows segments v_i and lower picture shows segments t_i on the same set of points. a and b are extreme points and therefore segment s ends in a and b. Red segments denote split into M segments of equal length v_i. Blue segments denote segments t_i. t_5 is an empty segment, because there are no points that lie on segment v_5. Segments t_3 and t_7 are degenerated to one point – c and d respectively. Segments t_1 and t_2 share one point b.

You can see an example of such segments v_i and t_i in Figure 4.1.

We use the inductive hypothesis to choose (k, δ)-dense sets A_i for sets C_i. Note that if $|C_i| \leq 1$, then $A_i = C_i$ and it is still a (k, δ)-dense set for C_i.

Then we define $A = \bigcup_{i=1}^{M} A_i$. Thus A includes the extreme points of C, because they are included in the sets A_1 and A_M.

Size of each A_i is at most $(2 + \frac{2}{\delta})^k$ from the inductive hypothesis, therefore size of A is at most:

$$M\left(2 + \frac{2}{\delta}\right)^k = \left\lceil 1 + \frac{2}{\delta} \right\rceil \cdot \left(2 + \frac{2}{\delta}\right)^k \leq \left(2 + \frac{2}{\delta}\right)^{k+1}.$$

Proof that A is (k, δ)-dense for C. Let us take any cover of A with $k + 1$ segments and call it $\mathcal{R}$.

For every segment t_i, if there exists a segment x in $\mathcal{R}$ that is disjoint with t_i, then we have a cover of A_i with at most k segments using $\mathcal{R} - \{x\}$. Since A_i is (k, δ)-dense for t_i and C_i, then $(\mathcal{R} - \{x\})^{+\delta}$ covers C_i. So $\mathcal{R}^{+\delta}$ covers C_i as well.

If there exists a segment t_i for which a segment x as defined above does not exist, then all $k + 1$ segments that cover A_i intersect with t_i. You can see an example of such segments in Figure 4.2. Note that there may exist only one such segment t_i. From the inductive hypothesis endpoints of s are in A_1 and A_M respectively, so $\mathcal{R}$ must cover them. For each endpoint of s, there exists a segment that starts in this endpoint and ends somewhere in t_i. Let us call these two segments y and z. It follows that: $|y| + |z| + |t_i| \geq |s|$. Since $|t_i| \leq |v_i| = \frac{|s|}{M} \leq \frac{|s|}{1+\frac{2}{\delta}} = \frac{|s|\delta}{\delta+2}$, we have $\max(|y|, |z|) \geq |s|(1 - \frac{\delta}{\delta+2})/2 = \frac{|s|}{\delta+2}$.

After δ-extension, the longer of these segments will expand at both ends by at least:

$$\max(|y|, |z|)\delta \geq \frac{|s|\delta}{\delta + 2} = \frac{|s|}{1 + \frac{2}{\delta}} \geq \frac{|s|}{M} = v_i \geq t_i.$$

Therefore, the longer of segments y and z will cover the whole segment t_i after δ-extension.

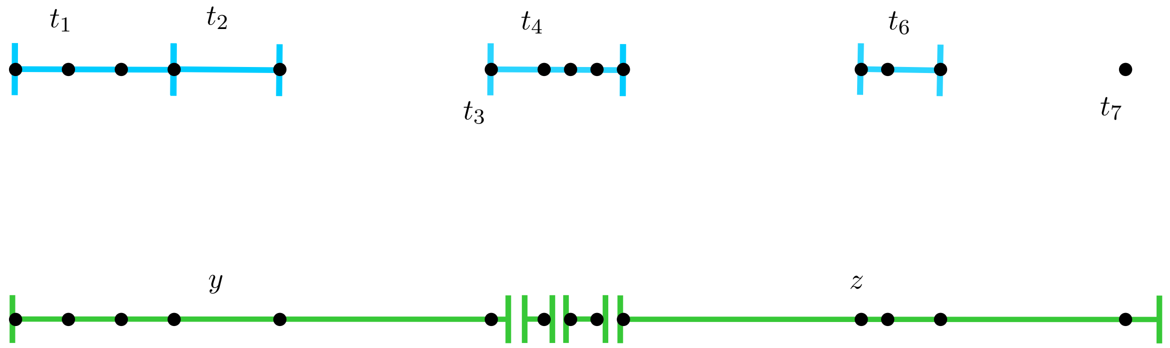

Figure 4.2: **Example of all $k+1$ segments intersecting with one segment t_i.** Both pictures show the same set $\mathcal{C}$ (black circles), the same as in Figure 4.1. The upper picture shows blue segments t_i for $M = 7$. The lower picture shows green segments – solution $\mathcal{R}$ of size 4. All segments from $\mathcal{R}$ intersect with t_4. Segments z and y are named on the picture.

We conclude that $\mathcal{R}^{+\delta}$ covers C_i.

Since $C = \bigcup_{i=1}^{M} C_i$, it follows that $\mathcal{R}^{+\delta}$ covers C.

Algorithm. We can simulate the inductive proof by a recursive algorithm with the following complexity:

$$O\left(|C| + \frac{1}{\delta}\right) + O\left(k(2 + \frac{1}{\delta})^k\right).$$

□

Let us now formulate some claims about the properties for the problem parametrized by the solution size. These properties provide bounds for different objects in the problem instance, that help us to find small kernel of the problem or claim that the minimal solution of this instance must be above some treshold.

Definition 4.3. A line in $\mathbb{R}$ is **long** if there are at least $k+1$ points from $\mathcal{C}$ on it.

Claim 4.1. *If there are more than k different long lines, then $\mathcal{C}$ can not be covered with k segments.*

Proof. We prove the claim by contradiction. Let us assume that we have at least $k+1$ different long lines in our instance of the problem and solution $\mathcal{R}$ of size at most k covering points $\mathcal{C}$.

Choose any long line L. Every segment from $\mathcal{R}$, which is not collinear with L, covers at most one point that lies on L.

L is long, so there are at least $k+1$ points from $\mathcal{C}$ that lie on L. That implies that there must be a segment in $\mathcal{R}$ that is collinear with L.

Since we have at least $k+1$ different long lines, then there are at least $k+1$ segments in $\mathcal{R}$ collinear with different lines. It contradicts with the assumption that $|\mathcal{R}| \leq k$. □

Claim 4.2. *If there are more than k^2 points from $\mathcal{C}$ that do not lie on any long line, then $\mathcal{C}$ can not be covered with k segments.*

Proof. We prove the claim by contradiction. Let us assume that we have at least k^2+1 points from $\mathcal{C}$ that do not lie on any long line, call this set A, and a solution $\mathcal{R}$ of size at most k covering points $\mathcal{C}$.

For every segment s from $\mathcal{R}$ it covers at most k points from A. It holds because if s covered at least $k+1$ points from A, then the line in the direction of s would be a long line and that contradicts of definiton of A.

If every segment from $\mathcal{R}$ covers at most k points from A and $|\mathcal{R}| \le k$, then at most k^2 points from A are covered by $\mathcal{R}$ and that contradicts the fact that $\mathcal{R}$ is a solution of the given geometric set cover instance. □

We are now ready to give a proof of Theorem 4.2.

Proof of Theorem 4.2. Applying the claims 4.1 and 4.2, if we have more than k different long lines or more than k^2 points from $\mathcal{C}$ that do not lie on any long line, then we answer that there is no solution of size at most k.

Otherwise, we can split $\mathcal{C}$ into at most $k+1$ sets:

- D, at most k^2 points that do not lie on any long line;
- C_i, points that lie on the i-th long line.

Sets C_i do not need to be disjoint.

Then for every set C_i we can use Lemma 4.3 to obtain a (k,δ)-dense set A_i for C_i with $|A_i| \le (2+\frac{2}{\delta})^k$.

Then we have a set $\mathcal{C}' = D \cup (\bigcup A_i)$ of size at most $k^2 + k(2+\frac{2}{\delta})^k$. Observe that if we have a solution $\mathcal{R}$ of size at most k that covers $\mathcal{C}'$, then $\mathcal{R}^{+\delta}$ covers $\mathcal{C}$.

$\mathcal{C}$ is separated into several parts – sets D and C_i. Points from D are covered by $\mathcal{R}$, because D is part of $\mathcal{C}'$. Each A_i is covered, because A_i is part of $\mathcal{C}'$; A_i is a (k,δ)-dense set for C_i, therefore $\mathcal{R}^{+\delta}$ covers C_i.

After that we shrunk down $\mathcal{C}$ to $\mathcal{C}'$ of size $f(k,\delta)$ for some computable function f. Then we would like to shrink down $\mathcal{P}$ to some set of relevant segments of bounded size as well.

For every pair of points $\mathcal{C}'$, we can choose one segment from $\mathcal{P}$ that have the lowest weight among segments that cover these points or decide there is no segment that cover them. Call this set $\mathcal{P}'$ and name these segments **interesting**. There are at most $|\mathcal{C}'|^2$ different segments in $\mathcal{P}'$.

We need to show that when we cover $\mathcal{C}'$ with segments from $\mathcal{P}'$ we achieve the same minimal solution as when we cover them with segments from $\mathcal{P}$. In order to prove this, consider a minimal solution $\mathcal{R}$ that covers $\mathcal{C}'$ with segments from $\mathcal{P}'$ and take any segment s from $\mathcal{R}$. Let us look at the points from $\mathcal{C}'$ that lie on s and call this set of points F. F is a set of collinear points for course. We can cover F with any segment that covers extreme points of F, because all other points lay on the segment between these points. Therefore we can change s to an interesting segment s' and interesting segments are defined in such a way, that s' has weight no larger than weight of s.

This has complexity $O(|\mathcal{C}'|^2|\mathcal{P}|)$ and produces shrunk down set of segments $\mathcal{P}'$ of size $f(k,\delta)$ for some computable function f.

Then we can iterate over all subsets of $\mathcal{P}'$ and choose the set with the lowest sum of weights that cover $\mathcal{C}'$. This solution would have weight not larger than optimal solution for the problem without extension, because we iterate over all posibilities of covering the subset of $\mathcal{C}'$. □

Chapter 5

W[1]-hardness for weighted segments in 3 directions

In this chapter we consider geometric set cover problem with weighted segments. Theorem 5.1 proves that this problem is W[1]-hard when parametrized by the size of the solution. We additionally restrict the problem to only use segments in three directions to achieve a stronger result. W[1]-hardness is proved by reduction to a grid tiling problem, which was introduced in [Marx, 2007].

Definition 5.1. Line is **right-diagonal** if it is described by linear function $y = -x + d$ for any $d \in \mathbb{R}$. Segment is **right-diagonal** if its direction is a right-diagonal line.

Theorem 5.1. *Consider the problem of covering a set $\mathcal{C}$ of points by selecting at most k segments from a set of segments $\mathcal{P}$ with non-negative weights $w : \mathcal{P} \to \mathbb{R}^+$ so that the weight of the cover is minimal. Then this problem is W[1]-hard when parametrized by $\sqrt{k}$ and assuming ETH, there is no algorithm for this problem with running time $f(k) \cdot (|\mathcal{C}| + |\mathcal{P}|)^{o(\sqrt{k})}$ for any computable function f. Moreover, this holds even if all segments in $\mathcal{P}$ are axis-parallel or right-diagonal.*

Theorem 5.1 is also true for less restricted problem where segments have any direction. In order to prove Theorem 5.1 we will show reduction from a W[1]-hard problem. We introduce the grid tiling problem, which is proven to be W[1]-complete in literature.

Definition 5.2. We define **powerset** of a set A, denoted as $\mathsf{Pow}(A)$, as the set of all subsets of A, ie. $\mathsf{Pow}(A) = \{B : B \subseteq A\}$.

Definition 5.3. In the **grid tiling** problem we are given integers n and k, and a function $f : \{1 \dots k\} \times \{1 \dots k\} \to \mathsf{Pow}(\{1 \dots n\} \times \{1 \dots n\})$ specifying the set of allowed tiles for each cell of a $k \times k$ grid. The task is to find functions $x, y : \{1 \dots k\} \to \{1 \dots n\}$ that assign colors from $\{1 \dots n\}$ to respectively columns and rows of the grid, so that $(x(i), y(j)) \in f(i, j)$ for all valid i and j, or conclude that such an assignment does not exist.

In short, in grid tiling problem you need to assign numbers to rows and columns in such a way, that for every pair of a row and a column, the pair of colors assigned to the row and column belongs to the allowed set tiles for this pair. The next theorem describes the complexity of this problem, which is W[1]-hard when parametrized by the size of the grid.

Theorem 5.2. ***[Marx, 2007]*** *Grid tiling is W[1]-hard when parametrized by k and assuming ETH, there is no $f(k) \cdot n^{o(\sqrt{k})}$-time algorithm solving the grid tiling problem for any computable function f.*

	$x(1) = 3$	$x(2) = 1$	$x(3) = 3$	$x(4) = 7$
$y(4) = 1$	$(2,1)$; $(2,2)$; $\mathbf{(3,1)}$; $(3,9)$	$\mathbf{(1,1)}$; $(3,1)$	$\mathbf{(3,1)}$; $(7,2)$	$(2,1)$; $\mathbf{(7,1)}$
$y(3) = 1$	$(2,1)$; $\mathbf{(3,1)}$; $(4,2)$; $(8,2)$	$\mathbf{(1,1)}$; $(1,3)$	$\mathbf{(3,1)}$; $(4,3)$	$\mathbf{(2,2)}$; $\mathbf{(7,1)}$
$y(2) = 6$	$(2,6)$; $\mathbf{(3,6)}$	$(1,2)$; $\mathbf{(1,6)}$; $(2,6)$	$(2,6)$; $\mathbf{(3,6)}$	$(2,6)$; $\mathbf{(7,6)}$
$y(1) = 4$	$(2,4)$; $(2,6)$; $\mathbf{(3,4)}$; $(3,9)$	$\mathbf{(1,4)}$; $(1,9)$	$\mathbf{(3,4)}$; $(3,9)$	$(2,9)$; $\mathbf{(7,4)}$

Figure 5.1: **Example of a grid tiling instance with its solution.**
In the first row and column of the table you can see the solution: functions x and y. The tiles used in this solution are marked in **bold**. If we instead chose the tiles marked in blue (whenever there is one, taking the tile marked in **bold** otherwise), then that corresponds to setting $x(1) = 2$, and would also form a correct solution. On the other hand, if we instead chose the tiles marked in red (as before), then that corresponds to setting $y(1) = 9$ and $x(4) = 2$ and that would **not** form a correct solution. Even though the first row is correct, tile with coordinates (3,4) requires tile (2,1), not (2,2).

The reminder of this section is proving Theorem 5.1 by reduction of a grid tiling problem instance to a geometric set cover instance. That proves the W[1]-hardness of the geometric set cover problem, because if we could solve it with an FPT algorithm, then we could also solve the grid tiling problem (which we reduced to the geometric set cover). Therefore geometric set cover with setting described in Theorem 5.1 is at least as hard as the grid tiling problem.

Construction. We start with an instance of the grid tiling problem (n, k, f). The instance consists of:

- size of the grid k,
- number of colors n,
- function of allowed tiles $f : \{1, \ldots, k\} \times \{1, \ldots, k\} \to \mathsf{Pow}(\{1, \ldots, n\} \times \{1, \ldots, n\})$.

We construct an instance $(\mathcal{C}, \mathcal{P}, w, 3k^2 + 2k)$ of geometric set cover as follows.

First, let us choose any bijection $\mathsf{order} : \{1, \ldots, n^2\} \to \{1, \ldots, n\} \times \{1, \ldots, n\}$.

Define $\mathsf{match}_v(i, j)$ and $\mathsf{match}_h(i, j)$ as boolean functions denoting whether two points share x or y coordinate:

$$\mathsf{match}_v(i,j) \text{ is } \texttt{true} \iff \mathsf{order}(i) \text{ and } \mathsf{order}(j) \text{ have the same x coordinate},$$

$$\mathsf{match}_h(i,j) \text{ is } \texttt{true} \iff \mathsf{order}(i) \text{ and } \mathsf{order}(j) \text{ have the same y coordinate}.$$

Points. For $1 \le i, j \le k$ and $1 \le t \le n^2$ define points:

$$h_{i,j,t} := (i \cdot (n^2 + 1) + t, j \cdot (n^2 + 1)),$$

$$v_{i,j,t} := (i \cdot (n^2 + 1), j \cdot (n^2 + 1) + t).$$

Let us define sets H and V as:

$$H := \{h_{i,j,t} : 1 \le i, j, \le k, 1 \le t \le n^2\},$$

$$V := \{v_{i,j,t} : 1 \leq i, j, \leq k, 1 \leq t \leq n^2\}.$$

Let $\epsilon = \frac{1}{2k^2}$. For a point $p = (x, y)$ we define points:

$$p^L := (x - \epsilon, y),$$
$$p^R := (x + \epsilon, y),$$
$$p^U := (x, y + \epsilon),$$
$$p^D := (x, y - \epsilon).$$

Then we define the point set as follows:

$$\mathcal{C} := H \cup \{p^L : p \in H\} \cup \{p^R : p \in H\} \cup V \cup \{p^U : p \in V\} \cup \{p^D : p \in V\}.$$

Definition 5.4. For every point $p \in H$, we name point p^L its **left guard** and point p^R its **right guard**.

Similarily for every points $p \in V$, we name point p^D its **lower guard** and point p^U its **upper guard**.

Segments. For $1 \leq i, j \leq k$ and $1 \leq t_1, t_2 \leq n^2$ define segments:

$$\begin{aligned}
\mathsf{hor}_{i,j,t_1,t_2} &:= (h^R_{i,j,t_1}, h^L_{i+1,j,t_2}),\\
\mathsf{ver}_{i,j,t_1,t_2} &:= (v^U_{i,j,t_1}, v^D_{i,j+1,t_2}),\\
\mathsf{horBeg}_{i,t} &:= (h^L_{1,i,1}, h^L_{1,i,t}),\\
\mathsf{horEnd}_{i,t} &:= (h^R_{k,i,t}, h^R_{k,i,n^2}),\\
\mathsf{verBeg}_{i,t} &:= (v^D_{i,1,1}, v^D_{i,1,t}),\\
\mathsf{verEnd}_{i,t} &:= (v^U_{i,k,t}, v^U_{i,k,n^2}).
\end{aligned}$$

Next, we define sets of vertical and horizontal segments:

$$\begin{aligned}
\mathsf{HOR} &:= \{\mathsf{hor}_{i,j,t_1,t_2} : 1 \leq i < k, 1 \leq j \leq k, 1 \leq t_1, t_2 \leq n^2, \mathsf{match}_h(t_1, t_2) \text{ holds}\}\\
&\cup \ \{\mathsf{horBeg}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2\}\\
&\cup \ \{\mathsf{horEnd}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2\},
\end{aligned}$$

$$\begin{aligned}
\mathsf{VER} &:= \{\mathsf{ver}_{i,j,t_1,t_2} : 1 \leq i \leq k, 1 \leq j < k, 1 \leq t_1, t_2 \leq n^2, \mathsf{match}_v(t_1, t_2) \text{ holds}\}\\
&\cup \ \{\mathsf{verBeg}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2\}\\
&\cup \ \{\mathsf{verEnd}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2\}.
\end{aligned}$$

You can see an example of these segments in Figure 5.3.

Finally, we also define a set of right-diagonal segments:

$$\mathsf{DIAG} := \{(h_{i,j,t}, v_{i,j,t}) : 1 \leq i, j \leq k, 1 \leq t \leq n^2, \mathsf{order}(t) \in f(i,j)\}.$$

You can see an example of such segments in Figure 5.2.

Every segment in DIAG connects points $(i(n^2+1)+t, j\cdot(n^2+1))$ and $(i \cdot (n^2+1), j(n^2+1)+t)$ for some $1 \leq i, j \leq k, 1 \leq t \leq n^2$. The line on which it lies can be described by linear equation $y = -x + (t + (i+j)(n^2+1))$, thus these segments are in fact right-diagonal.

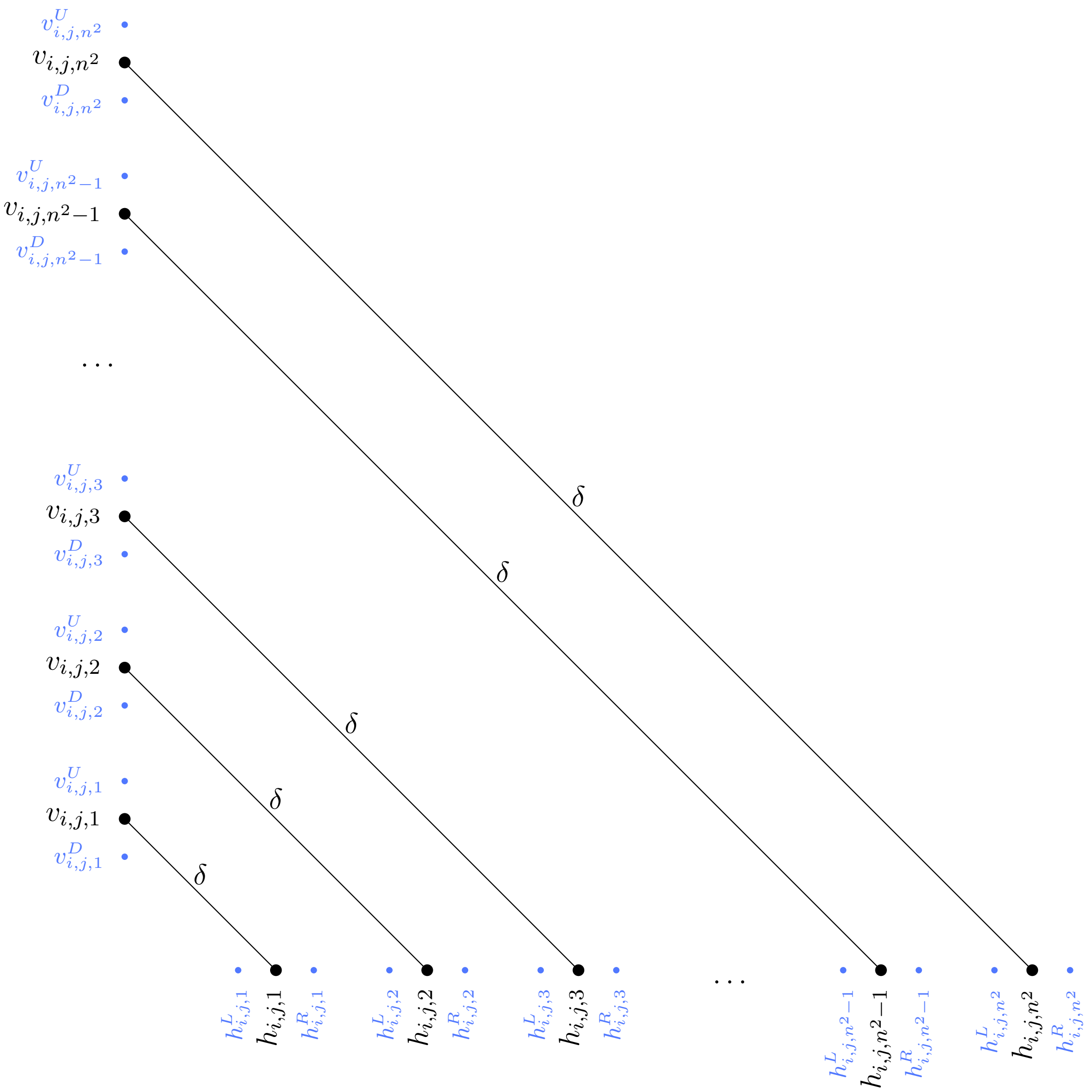

Figure 5.2: **Vertices and segments in DIAG.**
This is an example of constructed points any $1 \leq i, j \leq k$. Points from H and V are marked in black, their guards are marked in blue. You can also see segments from DIAG with their weights (equal to δ).

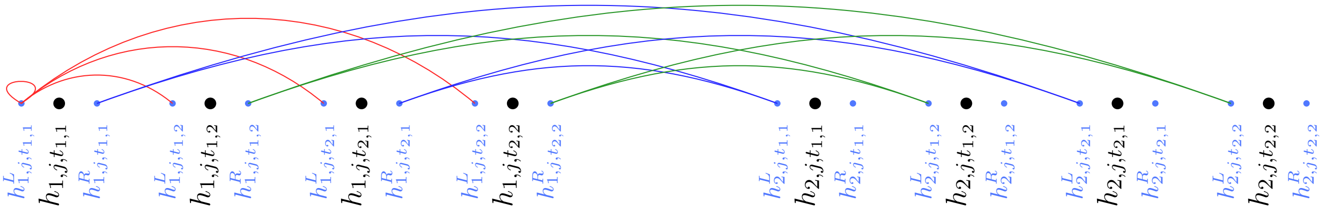

Figure 5.3: **Vertices and segments in** HOR**.**
This is an example for $n = 2$ and any $1 \leq j \leq k$. Points from H are marked in black, their guards are marked in blue. $t_{i,j}$ is a notation that we use for $\mathsf{order}^{-1}(i,j)$. Segments are represented as arcs between endpoints. You can see $\mathsf{horBeg}_{j,t}$ segments in red. $\mathsf{horBeg}_{j,1}$ is degenerated to a single point at $h^L_{1,1,t_{1,1}}$. Segments $\mathsf{hor}_{i,j,t_{x_1,y},t_{x_2,y}}$ are marked in blue and green. Blue segments connect $t_{x_1,y}$ and $t_{x_2,y}$ such that they share y-coordinate equal to 1, for green segments it is equal to 2.

The constructed segment set is defined as:

$$\mathcal{P} := \mathsf{HOR} \cup \mathsf{VER} \cup \mathsf{DIAG}.$$

The weight of each segment in $\mathsf{HOR} \cup \mathsf{VER}$ is equal to its length, while every segment in DIAG has weight $\delta := \frac{1}{4k^4}$.

$$w(s) = \begin{cases} length(s) & \text{if } s \in \mathsf{HOR} \cup \mathsf{VER} \\ \delta & \text{if } s \in \mathsf{DIAG} \end{cases}$$

Now, we prove that the constructed instance of geometric set cover with weighted segments is indeed a correct and sound reduction of the grid tiling problem. Lemma 5.1 proves that if the solution of the instance of the grid tiling instance exists, then there exists a solution with bounded size and weight of the constructed instance of geometric set cover problem.

Then Lemma 5.5 proves that if the solution of the geometric set cover instance with bounded weight exists, then there exists a solution to the original grid tiling instance.

Lemma 5.1. *If there exists a solution of the grid tiling instance* $(f_{i,j})$*, then there exists a solution of the instance* $(\mathcal{C}, \mathcal{P}, w, 3k^2 + 2k)$ *of geometric set cover with weight* $2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$*.*

Proof. Suppose there exists a solution x, y of the instance $(f_{i,j})$ of the grid tiling problem.

We define the proposed solution $\mathcal{R} \subset \mathcal{P}$ of the instance of geometric set cover in three parts $D \subset \mathsf{DIAG}$, $A \subset \mathsf{HOR}$ and $B \subset \mathsf{VER}$:

$$\begin{aligned}
D &:= \{(v_{i,j,t}, h_{i,j,t}) : 1 \leq i,j \leq k, t = \mathsf{order}^{-1}(x(i), y(j))\}, \\
A &:= \{\mathsf{horBeg}_{i,\mathsf{order}^{-1}(x(1),y(i))} : 1 \leq i \leq k\} \\
&\cup \{\mathsf{horEnd}_{i,\mathsf{order}^{-1}(x(k),y(i))} : 1 \leq i \leq k\} \\
&\cup \{\mathsf{hor}_{i,j,\mathsf{order}^{-1}(x(i),y(j)),\mathsf{order}^{-1}(x(i+1),y(j))} : 1 \leq i < k, 1 \leq j \leq k\}, \\
B &:= \{\mathsf{verBeg}_{i,\mathsf{order}^{-1}(x(i),y(1))} : 1 \leq i \leq k\} \\
&\cup \{\mathsf{verEnd}_{i,\mathsf{order}^{-1}(x(i),y(k))} : 1 \leq i \leq k\} \\
&\cup \{\mathsf{ver}_{i,j,\mathsf{order}^{-1}(x(i),y(j)),\mathsf{order}^{-1}(x(i),y(j+1))} : 1 \leq i \leq k, 1 \leq j < k\},
\end{aligned}$$

$$\mathcal{R} := D \cup A \cup B.$$

Since $\mathcal{C} = H \cup V$, we show that $\mathcal{R}$ covers the whole set H, proof for V is analogous.

Take any $1 \le j \le k$ and define $t_i := \mathsf{order}^{-1}(x(i), y(j))$. The two leftmost segments in A for this j are $\mathsf{horBeg}_{j,t_1} = (h^L_{1,j,1}, h^L_{1,j,t_1})$ and $\mathsf{hor}_{1,j,t_1,t_2} = (h^R_{1,j,t_1}, h^L_{2,j,t_2})$. Therefore points $h_{1,j,x}, h^L_{1,j,x}$ and $h^R_{1,j,x}$ for all $1 \le x \le n^2$ ale covered by horBeg_{j,t_1} and $\mathsf{hor}_{1,j,t_1,t_2}$, excluding point h_{1,j,t_1}.

Analogously for $2 \le i \le k-1$ for two consecutive segments $\mathsf{hor}_{i-1,j,t_{i-1},t_i}$ and $\mathsf{hor}_{i,j,t_i,t_{i+1}}$ we prove that all points $h_{i,j,x}, h^L_{i,j,x}$ and $h^R_{i,j,x}$ for all $1 \le x \le n^2$ ale covered by these segments excluding point h_{i,j,t_i}.

Finally $\mathsf{hor}_{k-1,j,t_{k-1},t_k}$ and horEnd_{j,t_k} cover all points $h_{k,j,x}, h^L_{k,j,x}$ and $h^R_{k,j,x}$ for all $1 \le x \le n^2$ excluding point h_{k,j,t_k}.

D covers all points h_{i,j,t_i} and v_{i,j,t_i}, therefore all points in H are covered.

Size of this proposed solution is:

$$|\mathcal{R}| = |D| + |A| + |B| = k^2 + (k+1)k + (k+1)k = 3k^2 + 2k.$$

Then, we need to compute the total weight of the solution $\mathcal{R}$. First, we compute the sum of weights of segments in A. Fix $1 \le j \le k$ and compute segments collinear with the j-th line. All points $h_{i,j,t}$, $h^L_{i,j,t}$ and $h^R_{i,j,t}$ for every $1 \le i \le k$ and $1 \le t \le n^2$ are covered by A excluding points $h_{i,j,\mathsf{order}^{-1}(x(i),y(j))}$. Every such point leaves a gap of length 2ϵ between $h^L_{i,j,\mathsf{order}^{-1}(x(i),y(j))}$ and $h^R_{i,j,\mathsf{order}^{-1}(x(i),y(j))}$. Therefore, the total weight of segments in A that lie on the line in question equals the length of the segment $(h^L_{i,1,1}, h^R_{i,k,n^2})$ minus $2\epsilon k$, which is $k(n^2+1) - 2(1-\epsilon) - 2k\epsilon$. We need to multiply that by k, as we consider all possible values of j.

Calculation for vertical segments is analogous and has the same result. Every segment in D has weight δ, therefore the sum of all weights is equal to:

$$2k(k(n^2+1) - 2(1-\epsilon) - 2k\epsilon) + k^2\delta = 2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$$

$\square$

Claim 5.1. *In any solution of the instance $(\mathcal{C}, \mathcal{P}, w, 3k^2+2k)$:*

- *left and right guards of points in H (points in $\{p^L : p \in H\} \cup \{p^R : p \in H\}$) have to be covered with segments from* HOR*,*
- *lower and upper guards of points in V (points in $\{p^D : p \in V\} \cup \{p^U : p \in V\}$) have to be covered with segments from* VER*.*

Proof. We prove the claim for the points from H as the proof for points from V is analogous.

Every segment in VER is vertical and has x-coordinate equal to $i(n^2+1)$ for some $1 \le i \le k$, so they all have different x-coordinate than any left or right guard of points in H.

Every point x, which is a left or right guard of points in H have kn^2 segments from DIAG that intersect with the horizontal line that goes through x. All of these segments intersect with this line in points from set H, therefore none of them cover any of the guards.

Therefore none of the segments from VER or DIAG cover any of the guards of the points in H. $\square$

Now we present a few additional properties of the constructed instance of the geometric set cover that help us to prove Lemma 5.5.

Claim 5.2. *For any $1 \leq i, j \leq n$ and any solution of the instance $(\mathcal{C}, \mathcal{P}, w, 3k^2 + 2k)$ all, but at most one point $h_{i,j,t}$ and at most one point $v_{i,j,t}$ for $1 \leq t \leq n^2$ must be covered with segments from* HOR *or* VER*.*

Proof. We prove the claim for horizontal segments, as the proof for vertical segments is analoguous.

We prove this by contradiction. Assume that we have two points h_{i,j,t_1}, h_{i,j,t_2} such that they are not covered with segments from HOR for any $1 \leq t_1 < t2 \leq n^2$.

Point h^R_{i,j,t_1} has to be covered with HOR by Claim 5.1. Every segment in HOR covering h^R_{i,j,t_1}, but not h_{i,j,t_1} must start at h^R_{i,j,t_1} and all such segments cover also h_{i,j,t_2}. This contradicts the assumption, which concludes the proof. ☐

Lemma 5.2. *For every solution of the instance $(\mathcal{C}, \mathcal{P}, w, 3k^2 + 2k)$, the sum of weights of segments chosen from sets* HOR *and* VER *is at least $2k^2(n^2 + 1) - 4k^2\epsilon - 4k(1 - \epsilon)$.*

Proof. We prove the lemma for vertical lines, as the proof for horizontal segments is analogous.

Let us fix $1 \leq i \leq k$.

We provide a lower bound for the sum of lengths of vertical segments from $\mathcal{R} \cap$ VER. This bound is the same for each i and is the same for horizontal lines, thus we need to multiply such bound by $2k$.

(1) The total length between $v^D_{i,1,1}$ and v^U_{i,k,n^2} is:

$$(k(n^2 + 1) + n^2 + \epsilon) - ((n^2 + 1) + 1 - \epsilon) = k(n^2 + 1) - 2(1 - \epsilon).$$

(2) For every $1 \leq j \leq k$ there exists at most one $1 \leq t \leq n^2$ such that $v_{i,j,t}$ is not covered by segments from VER (Claim 5.2). Its guards (see Definition 5.4) $v^U_{i,j,t}$ and $v^D_{i,j,t}$ have to be covered in VER (Claim 5.1). Therefore, at most k spaces of length 2ϵ can be left not covered by segments from VER between $v^D_{i,1,1}$ and v^U_{i,k,n^2}.

The sum of these lower bounds for vertical and horizontal lines is:

$$2k(k(n^2 + 1) - 2k\epsilon - 2(1 - \epsilon)) = 2k^2(n^2 + 1) - 4k^2\epsilon - 4k(1 - \epsilon)$$

☐

Let us name the bound from the previous lemma as $W_{hv} := 2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon)$ for future reference.

Lemma 5.3. *Let $\mathcal{R}$ be a solution of a constructed instance $(\mathcal{C}, \mathcal{P}, w, 3k^2 + 2k)$ with weight at most $2k^2(n^2 + 1) - 4k^2\epsilon - 4k(1 - \epsilon) + k^2\delta$. Then for every $1 \leq i, j \leq k$ there exists such $1 \leq t \leq n^2$ that:*

(1) $v_{i,j,t}, h_{i,j,t}$ are not covered by segments from VER *or* HOR*;*

(2) segment $(v_{i,j,t}, h_{i,j,t})$ is in solution $\mathcal{R}$;

(3) $\mathsf{order}(t) \in f(i, j)$*, that is,* $\mathsf{order}(t)$ *is an allowed tile for (i, j);*

(4) for every $1 \leq s \leq n^2$, $s \neq t$, $v_{i,j,s}$ is covered in VER*;*

(5) for every $1 \leq s \leq n^2$, $s \neq t$, $h_{i,j,s}$ is covered in HOR*.*

Proof. At most one of points $\{h_{i,j,t_x} : 1 \le t_x \le n^2\}$ and one of points $\{v_{i,j,t_y} : 1 \le t_y \le n^2\}$ is covered with DIAG (Claim 5.2).

Moreover, exactly one such point h_{i,j,t_x} and one such point v_{i,j,t_y} is covered with DIAG, because if none of them were covered, then the solution would have to have weight at least $W_{hv} + 2\epsilon$ (Lemma 5.2), which is more than $2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$.

We observe that points h_{i,j,t_x} and v_{i,j,t_y} have to be covered with the same segment from DIAG. Indeed we need to use at least k^2 of them to use exactly one DIAG segment for every pair of $1 \le i,j \le k$, if we used 2 segments from DIAG for one pair (i,j), then we would have used $W_{hv} + k^2\delta + \delta$ (Lemma 5.2), which if more than $2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$. Since points h_{i,j,t_x} and v_{i,j,t_y} are covered by a single segment from DIAG, we have $t_x = t_y$.

Therefore $t_x = t_y$ and $\mathsf{order}(t_x)$ is an allowed tile for (i,j) because the corresponding segment is in DIAG. ☐

We refer to the function mapping $1 \le x \le k$ to t_x from Lemma 5.3 as $\mathsf{diagonal} : \{1 \dots k\} \times \{1 \dots k\} \to \{1 \dots n^2\}$.

Lemma 5.4. *For any solution $\mathcal{R}$ of a constructed instance $(\mathcal{C}, \mathcal{P}, w, 3k^2 + 2k)$ with weight at most $2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$:*

1. *for any $1 \le i < k, 1 \le j \le k$, $\mathsf{match}_h(\mathsf{diagonal}(i,j), \mathsf{diagonal}(i+1,j))$ is `true`;*
2. *for any $1 \le i \le k, 1 \le j < k$, $\mathsf{match}_v(\mathsf{diagonal}(i,j), \mathsf{diagonal}(i,j+1))$ is `true`.*

Proof. We prove (1) by contradiction, the proof of (2) is analogous.

Let us take any $1 \le i < k, 1 \le j \le k$ and name $t_1 = \mathsf{diagonal}(i,j)$ and $t_2 = \mathsf{diagonal}(i+1,j)$. We also assume that $\mathsf{match}_h(t_1,t_2)$ is `false`, which is equivalent to the fact that segment $(h^R_{i,j,t_1}, h^L_{i+1,j,t_2})$ is not in set HOR.

Therefore h_{i,j,t_1} and h_{i+1,j,t_2} are not covered by segments from HOR (Lemma 5.3), while h^R_{i,j,t_1} and h^L_{i+1,j,t_2} have to be covered by segments from HOR (Claim 5.1).

Every segment from HOR starts at point h^R_{x,y,z_1} and ends at point h^L_{x+1,y,z_2} for some $1 \le x < k$,$1 \le y \le k$ and $1 \le z_1, z_2 \le n^2$. All of the points between h^R_{i,j,t_1} and h^L_{i+1,j,t_2} are covered by segments in HOR and there is no segment $(h^R_{i,j,t_1}, h^L_{i+1,j,t_2})$ in HOR. Hence, there are at least two different segments covering them. One of them must begin at h^R_{i,j,t_1} and end at h^L_{i+1,j,z_2} and there must be other one that begins at h^R_{i,j,z_1} and ends at h^L_{i+1,j,t_2} for some $1 \le z_1, z_2 \le n^2$.

Thus, the space between h^R_{i,j,z_1} and $h^L_{i,j+1,z_2}$ would be covered twice and is longer than ϵ. By Lemma 5.2, the lower bound for weight of such a solution is $W_{hv} + \epsilon$ which is more than $2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$.

Therefore h^R_{i,j,t_1} and h^L_{i+1,j,t_2} must be covered by one segment from HOR, $(h^R_{i,j,t_1}, h^L_{i+1,j,t_2})$ is a segment in HOR and $\mathsf{match}_h(t_1,t_2)$ is `true`. ☐

Lemma 5.5. *If there exists solution of instance $(\mathcal{C}, \mathcal{P}, w, 3k^2 + 2k)$ with weight at most $2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$, then there exists a solution for the grid tiling instance $(f_{i,j})$.*

Proof. Take diagonal function from Lemma 5.3.

To define the x funtion for every $1 \le i \le k$ set $x(i) := x_i$ where $(x_i, a) = \mathsf{order}(v_{i,1})$. Similarily, to define the y function, for every $1 \le i \le k$ set $y(i) := y_i$ where $(b, y_i) = \mathsf{order}(h_{1,i})$

To prove that it is a correct solution for grid tiling, we need to prove that for every $1 \le i,j \le k$ $(x(i), y(j))$ is in allowed tiles set $f(i,j)$.

Let us take any $1 \leq i, j \leq k$. By Lemma 5.4 and simple induction, we know that $\mathsf{match}_h(\mathsf{diagonal}(1, j), \mathsf{diagonal}(i, j))$ and $\mathsf{match}_v(\mathsf{diagonal}(i, 1), \mathsf{diagonal}(i, j))$ are `true`. Therefore $\mathsf{order}(\mathsf{diagonal}(i, j)) = (x(i), y(j))$. By Lemma 5.3 we know that $\mathsf{order}(\mathsf{diagonal}(i, j))$ is in $f(i, j)$. Therefore $(x(i), y(j))$ is in $f(i, j)$. $\square$

Proof of Theorem 5.1. Follows from Lemmas 5.1 and 5.5. $\square$

TODO: Add reference when known In proof of reduction we did not use the assumption that the solution is of bounded size. Thus this reduction proves that the problem is not only W[1]-hard, but assuming ETH there also does not exist permissive FPT algorithm for this problem.

Chapter 6

Geometric Set Cover with lines

6.1. Lines parallel to one of the axis

When $\mathcal{R}$ consists only of lines parallel to one of the axis, the problem can be solved in polynomial time.

We create bipartial graph G with node for every line on the input split into sets: H – horizontal lines and V – vertical lines. If any two lines cover the same point from $\mathcal{C}$, then we add edge between them.

Of course there will be no edges between nodes inside H, because all of them are pararell and if they share one point, they are the same lines. Similar argument for V. So the graph is bipartial.

Now Geometric Set Cover can be solved with Vertex Cover on graph G. Since Vertex Cover (even in weighted setting) on bipartial graphs can be solved in polynomial time.

Short note for myself just to remember how to this in polynomial time:

Non-weighted setting - Konig theorem + max matching

Weighted setting - Min cut in graph of $\neg A$ or $\neg B$ (edges directed from V to H)

6.2. FPT for arbitrary lines

You can find this is Platypus book. We will show FPT kernel of size at most k^2.

(Maybe we need to reduce lines with one point/points with one line).

For every line if there is more than k points on it, you have to take it. At the end, if there is more than k^2 points, return NO. Otherwise there is no more than k^4 lines.

In weighted settings among the same lines with different weights you leave the cheapest one and use the same algorithm.

6.3. APX-completeness for arbitrary lines

We will show a reduction from Vertex Cover problem. Let's take an instance of the Vertex Cover problem for graph G. We will create a set of $|V(G)|$ pairwise non-pararell lines, such that no three of them share a common point.

Then for every edge in $(v, w) \in E(G)$ we put a point on crossing of lines for vertices v and w. They are not pararell, so there exists exactly one such point and any other line do not cover this point (any three of them do not cross in the same point).

Solution of Geometric Set Cover for this instance would yield a sound solution of Vertex Cover for graph G. For every point (edge) we need to choose at least one of lines (vertices) v or w to cover this point.

Vertex Cover for arbitrary graph is APX-complete, so this problem in also APX-complete.

6.4. 2-approximation for arbitrary lines

Vertex Cover has an easy 2-approximation algorithm, but here very many lines can cross through the same point, so we can do d-approximation, where d is the biggest number of lines crossing through the same point. So for set where any 3 lines do not cross in the same point it yields 2-approximation.

The problematic cases are where through all points cross at least k points and all lines have at least k points on them. It can be created by casting k-grid in k-D space on 2D space.

Greedy algorithm yields $\log |\mathcal{R}|$-approximation, but I have example for this for bipartial graph and reduction with taking all lines crossing through some point (if there are no more than k) would solve this case. So maybe it works.

Unfortunaly I have not done this :(

I can link some papers telling it's hard to do.

6.5. Connection with general set cover

Problem with finite set of lines with more dimensions is equivalent to problem in 2D, because we can project lines on the plane which is not perpendicular to any plane created by pairs of (point from $\mathcal{C}$, line from $\mathcal{P}$).

Of course every two lines have at most one common point, so is every family of sets that have at most one point in common equivalent to some geometric set cover with lines?

No, because of Desargues's theorem. Have to write down exactly what configuration is banned.

Chapter 7

Geometric Set Cover with polygons

7.1. State of the art

Covering points with weighted discs admits PTAS [Li and Jin, 2015] and with fat polygons with δ-extensions with unit weights admits EPTAS [Har-Peled and Lee, 2009].

Although with thin objects, even if we allow δ-expansion, the Set Cover with rectangles is APX-complete (for $\delta = 1/2$), it follows from APX-completeness for segments with δ-expansion in Section **??**.

Covering points with squares is W[1]-hard [Marx, 2005]. It can be proven that assuming $SETH$, there is no $f(k) \cdot (|\mathcal{C}| + |\mathcal{P}|)^{k-\epsilon}$ time algorithm for any computable function f and $\epsilon > 0$ that decides if there are k polygons in $\mathcal{P}$ that together cover $\mathcal{C}$, *Theorem 1.9* in [Marx and Pilipczuk, 2015].

Chapter 8

Conclusions

We do not know FPT for axis-parallel segments without δ-extensions.

Bibliography

[Cygan et al., 2015] Cygan, M., Fomin, F. V., Kowalik, L., Lokshtanov, D., Marx, D., Pilipczuk, M., Pilipczuk, M., and Saurabh, S. (2015). *Parameterized Algorithms*. Springer.

[Har-Peled and Lee, 2009] Har-Peled, S. and Lee, M. (2009). Weighted geometric set cover problems revisited. *Journal of Computational Geometry*, 3.

[Håstad, 2001] Håstad, J. (2001). Some optimal inapproximability results. *J. ACM*, 48(4):798–859.

[Li and Jin, 2015] Li, J. and Jin, Y. (2015). A PTAS for the weighted unit disk cover problem. *CoRR*, abs/1502.04918.

[Marx, 2005] Marx, D. (2005). Efficient approximation schemes for geometric problems? In Brodal, G. S. and Leonardi, S., editors, *Algorithms – ESA 2005*, pages 448–459, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Marx, 2007] Marx, D. (2007). On the optimality of planar and geometric approximation schemes. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA, Proceedings*, pages 338–348. IEEE Computer Society.

[Marx and Pilipczuk, 2015] Marx, D. and Pilipczuk, M. (2015). Optimal parameterized algorithms for planar facility location problems using voronoi diagrams. *CoRR*, abs/1504.05476.