A Polynomial-Time Approximation Scheme for Traveling Salesman Problem with Neighborhoods Over Parallel Line Segments of Similar Length

by

Benyamin Ghaseminia

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science

University of Alberta

Abstract

In this thesis, we consider the Travelling Salesman Problem with Neighbourhoods (TSPN) on the Euclidean plane and present a Polynomial-Time Approximation Scheme (PTAS) when the neighborhoods are parallel line segments with lengths between $[1, \lambda]$ for any constant value λ. In TSPN (which generalizes classic TSP) each client represents a set (or neighbourhood) of points in a metric and the goal is to find a minimum cost TSP tour that visits at least one point from each client set. In the Euclidean setting, each neighbourhood is a region on the plane. TSPN is significantly more difficult than classic TSP even in the Euclidean setting. A notable case of TSPN is when each neighbourhood is a line segment. Although there are PTAS's for when neighbourhoods are fat objects (with limited overlap), TSPN over line segments is **APX**-hard even if all the line segments have unit length. For parallel (unit) line segments, the best approximation factor is $3\sqrt{2}$ from 20 years ago [10]. The PTAS we present in this thesis settles the approximability of this case of the problem. Our algorithm finds a $(1 + \varepsilon)$-factor approximation for an instance of the problem with n segments with lengths in $[1, \lambda]$ in time $n^{O(\lambda/\varepsilon^3)}$.

Acknowledgements

I want to express my deepest gratitude to my supervisor, Mohammad R. Salavatipour for trusting me throughout my graduate program. Mohammad dedicated a lot of his time and energy to helping me with this thesis. Without his knowledge and experience, this result wouldn't have been possible. I am thankful for his invaluable mentorship in the past two years.

I would also like to thank my lab-mates Kinter Ren and Mohsen Mohammadi in the Algorithmics Group for our time together and wish them the best. I appreciate the guidance of my two other lab-mates, Ismail Naderi and Ramin Mousavi who helped me settle into this new environment.

I am always grateful for my parents and their unwavering love and support towards me. Without their emotional and financial help throughout my life, I wouldn't have been where I am.

Last but certainly not least, I wish for the freedom of my people back in Iran. I hope we all see the good day we have always deserved.

Contents

List of Figures

Chapter 1

Introduction

The Traveling Salesman Problem (TSP) is one of the most fundamental and well-studied problems in Theoretical Computer Science with various applications such as in microchip wiring and vehicle navigation for package delivery. In TSP, one is given a set of points (which we refer to as *clients*) in a metric and the goal is to find a tour of minimum cost visiting all the points. We study this problem in the two dimensional Euclidean plane, meaning the points are given on the plane and metric is the Euclidean distance. TSP even in the Euclidean Setting is an **NP**-hard problem. The topic of this thesis is in Approximation Algorithms. In the field of Approximation Algorithms, we study **NP**-hard problems while leaning towards the assumption that $\mathbf{P} \neq \mathbf{NP}$; so instead of looking for polynomial-time exact solutions for **NP**-hard problems, we take compromises to guarantee a *near-optimum* solution. In Subsection 1.1.2, we properly describe Approximation Algorithms. Given an instance of an **NP**-hard optimization problem, we guarantee a solution within an α-factor of the optimum solution along with specific runtime guarantees (usually polynomial time in the size of the input and possibly α). Parameter α can be a constant value in $\mathbb{R}^+$, or a function of the size of the input.

In Section 1.1, we introduce some terminologies and proper definitions used throughout this thesis. In Section 1.2 we give examples of some variations of TSP other than the problem we considered, then mention some related work. We then summarize our results in Section 1.3.

1.1 Preliminaries

Some definitions are borrowed from the books of Vazirani [24] and Williamson & Shmoys [26] on Approximation Algorithms.

1.1.1 Graphs and Metrics

Whenever we use the notion of a graph, we use the same definitions found in West's book on Graphs [25]; meaning a graph G is defined by a vertex set $V(G)$, and an edge set $E(G)$. Throughout this thesis, the only graphs we consider are simple graphs.

Metrics

A function $d : V \times V \to \mathbb{R}^{\geq 0}$ over a vertex set V is a *metric* if the following properties hold:

1. For all $v \in V,\ d(v, v) = 0$
2. For all $u, v \in V : d(u, v) = d(v, u)$
3. For all $u \neq v,\ d(u, v) > 0$
4. For all $u, v, w \in V : d(u, v) \leq d(u, w) + d(w, v)$. This is referred to as the *triangle inequality*.

Consider a graph G with vertex set V such that any $v \in V$ corresponds to a point (v_x, v_y) on the two dimensional plane. Define $d : V \times V \to \mathbb{R}^{\geq 0}$ on this vertex set such that for $u, v \in V : d(u, v) = \sqrt{|u_x - v_x|^2 + |u_y - v_y|^2}$; in other words, $d(u, v)$ is the Euclidean distance of the points corresponding to these two vertices. It can be seen that d is a metric and we refer to it as the *Euclidean metric*. We use the notation $||pq||$ to denote the Euclidean distance between any two points p and q on the plane.

1.1.2 Optimization Problems and Approximation Algorithms

Optimization Problem

An **NP**-optimization problem Π is defined by quadruple $(\mathcal{I}, D_\Pi, S_\Pi, \text{obj}_\Pi)$ where:

- Set $\mathcal{I}$ is the set of *instances*.
- D_Π is the set of *valid instances*, and given any instance $I \in \mathcal{I}$, we can check in polynomial time in $|I|$ whether or not I is a valid instance.
- For any valid instance $I \in D_\Pi$, the set $S_\Pi(I)$ is the set of *feasible solutions* of I, where $S_\Pi(I) \neq \emptyset$ and each $s \in S_\Pi(I)$ has a length polynomially bounded in $|I|$. We can, in polynomial time in $|I|$, decide whether or not any given solution s is a feasible solution for I.
- obj_Π is a polynomial-time computable *objective function* that given a valid instance I and a feasible solution $s \in S_\Pi(I)$, assigns a non-negative rational value to the pair (I, s).

Problem Π is specified to either be a minimization problem or a maximization problem. The goal for a minimization problem Π given any $I \in D_\Pi$, is to find a solution $s \in S_\Pi(I)$ that minimizes the objective function between all pairs (I, s); meaning $s = \arg\min_{s' \in S_\Pi(I)} \text{obj}_\Pi(I, s')$. The formulation for a maximization problems is analogous. We denote such a desired solution s as the *optimum solution* for instance I and write it as $\text{OPT}_\Pi(I)$, and denote the value of the objective function for this solution as $\text{opt}_\Pi(I)$. In this thesis, when context is clear, we simplify these notations and only use OPT to refer to an optimum solution, and opt to refer to the value of the objective function for that optimum.

TSP is an example of an optimization (minimization) problem that is shown to be **NP**-hard even in the Euclidean metric. The set of valid instances for TSP are any number of points on the Euclidean plane, the set of feasible solutions are any tour that intersects all the points in the given instance, and the objective function is the total length of the given tour.

Approximation Algorithms

Informally, in approximation algorithms we are given a valid instance I of some **NP**-hard optimization problem (say a minimization problem). The goal is to present a feasible solution s' such that the value of the objective function for s' is at most an α-factor larger than the objective function value for the optimum solution of instance I. A formal definition is as follows:

Consider a minimization problem Π and a function $\alpha : \mathbb{Z}^+ \rightarrow \mathbb{Q}^{\geq 1}$. An algorithm $\mathcal{A}$ is an *α-approximation* (or *factor α approximation*) for Π if for any valid instance I, $\mathcal{A}$ outputs a feasible solution s for which $\mathrm{obj}_\Pi(I, s) \leq \alpha(|I|) \cdot \mathrm{opt}_\Pi(I)$, and that the running time of $\mathcal{A}$ for instance I is polynomial in $|I|$. We refer to α as the *approximation ratio* of $\mathcal{A}$. Analogous definition holds for maximization problems. Since TSP is a minimization problem, we focus on minimization problems for the definitions from now on.

Unless **P** = **NP**, we will not be able to find any algorithm that is a 1-approximation for any **NP**-hard problem. So in this line of research, the goal is to find algorithms with approximation factors as close to 1 as possible, and with the most efficient running times.

Polynomial-Time Approximation Scheme

One special case of approximation algorithms are those with a $(1 + \varepsilon)$-factor approximation for any given real number $\varepsilon > 0$. If $\mathcal{A}$ is a $(1 + \varepsilon)$-factor approximation for a minimization problem Π that runs in poly-time in the size of the input, then $\mathcal{A}$ is called a *Polynomial-Time Approximation Scheme (PTAS)* of Π. Sometimes finding such $\mathcal{A}$ that runs in poly-time will prove to be difficult and there might be some relaxations in the running time; those relaxations, however, are not needed in this thesis as we will present a PTAS at the end. A special case of a PTAS is when the running time is not only polynomial in the size of the input, but also poly-time in $1/\varepsilon$; these algorithms are called *Fully Polynomial-Time Approximation Schemes (FPTAS)*.

PTAS-reduction

Given two optimization (minimization) problems Π and Π', we say Π is *PTAS-reducible* to Π' [7], and use the notation $\Pi \leq_{\text{PTAS}} \Pi'$ if there exist a triplet (f, g, c) of functions such that:

1. For any $I \in D_\Pi$ and any rational $\varepsilon > 1$, $f(I, \varepsilon) \in D_{\Pi'}$ and f is computable in poly-time with respect to $|I|$.

2. For any $I \in D_\Pi$, for any rational $\varepsilon > 1$ and $s' \in S_{\Pi'}(f(I, \varepsilon))$, $g(I, s', \varepsilon) \in S_\Pi(I)$, and g is computable in poly-time in respect to both $|I|$ and $|s'|$.

3. $c : \mathbb{R}^{>1} \to \mathbb{R}^{>1}$ is computable and invertible.

4. For any $I \in D_\Pi$, for any rational $\varepsilon > 1$ and $s' \in S_{\Pi'}(f(I, \varepsilon))$, if

$$\text{obj}_{\Pi'}(f(I, \varepsilon), s') \leq c(\varepsilon) \cdot \text{OPT}_{\Pi'}(f(I, \varepsilon)),$$

then

$$\text{obj}_\Pi(I, g(I, s', \varepsilon)) \leq \varepsilon \cdot \text{OPT}_\Pi(I).$$

Approximation Classes

An optimization problem Π is said to be in the class **APX** if there are any approximation algorithms for it with a constant approximation ratio. Π is said to be in the class **PTAS** or **FPTAS** if respective approximation schemes exist for it.
An optimization problem Π is said to be ***APX**-hard* if for any $\Pi' \in$ **APX**, $\Pi' \leq_{\text{PTAS}} \Pi$. If for an **APX**-hard problem Π we have $\Pi \in$ **APX**, then Π is said to be ***APX**-complete*.

Hardness of Approximation

A *hardness proof* consists of showing that there cannot be an approximation algorithm for a given optimization problem with a ratio better than some threshold, assuming some specific complexity theory assumptions. As an example, for Max-3SAT it is shown that there exist some $\varepsilon_0 > 0$ such that finding an approximation algorithm for this problem with ratio better than $(1 + \varepsilon_0)$ is

NP-hard. Since it is also shown that MAX-3SAT is **APX**-complete, then this implies that any **APX**-hard problem Π will not have a PTAS unless $\mathbf{P} = \mathbf{NP}$.

1.1.3 Randomized Algorithms and Derandomization

In some approximation algorithms, ours included, the provided solution is a randomized solution with a guaranteed expected value (i.e. a guaranteed expected ratio) approximation. These randomized algorithms can usually be derandomized. We will explain the process in which this derandomization happens, after mentioning Markov's inequality:

Concentration Bounds

When discussing the Dynamic Program for our algorithm, we will use Markov's Inequality, stated as follows [21]: If X is a non-negative random variable, then for all $a > 0$, $\Pr[X \geq a] \leq \mathbb{E}(X)/a$.

We will use Markov's inequality in this context: For any $a \in \mathbb{R}^+$, if $\mathbb{E}[X] \leq a/2$, then with probability of at least $1/2$ we have $X < a$.

Proof. Using Markov's inequality, we have $\Pr[X \geq a] \leq \mathbb{E}[X]/a \leq 1/2$, implying that with probability at least $1/2$, we have $X < a$. ∎

Derandomization

Suppose we provide a proof with parameter j that is uniformly at random chosen from $\{1, 2, \ldots, h\}$ for some integer h, such that if X is the expected increase in the value of the objective function (compared to an optimum solution OPT with value opt), then with probability of at least $1/2$, $\mathbb{E}[X] \leq \frac{\text{opt}}{2}$.

Using the linearity of expectation, we have

$$\mathbb{E}[X] = \sum_{k=1}^{h} \mathbb{E}[X \mid j = k] \Pr[j = k] = \frac{1}{h} \sum_{k=1}^{h} \mathbb{E}[X \mid j = k].$$

Therefore, if $j^* = \arg\min_{1 \leq k \leq h} \mathbb{E}[X \mid j = k]$, we have $\mathbb{E}[X] \geq \mathbb{E}[X \mid j = j^*]$.

The same argument holds for cases that there are variables $x_1, x_2, \ldots, x_m$, where each of which are independently chosen uniformly at random from

$\{1, 2, \ldots, h\}$. Starting by $x_1^* = \arg\min_{1 \leq k \leq h} \mathbb{E}[X \mid x_1 = k]$, for each $i = 2, \ldots, n$ iteratively set

$$x_i^* = \arg\min_{1 \leq k \leq h} \mathbb{E}[X \mid (x_1 = x_1^*) \wedge (x_2 = x_2^*) \wedge \cdots \wedge (x_{i-1} = x_{i-1}^*) \wedge (x_i = k)].$$

Similar to before, using the linearity of expectation it can be seen the above value is not larger than $\mathbb{E}[X]$; notice that if we continue this process until $i = n$, then the expected value above is equal to the actual value of the final solution for $(x_1, x_2, \ldots, x_n) = (x_1^*, x_2^*, \ldots, x_n^*)$, which is now a deterministic solution.

In our case, the randomization that we have is based on only two variables (that correspond to two lines parallel to the axis on the plane); we can simply try out all the possible choices of those two variables and be sure that the minimized (expected) cost we get, is not worse than the expected value that we calculate using randomized parameters.

1.2 Related Work and Other Generalizations of TSP

There is a wide variety of generalizations for TSP such as different metrics, dimensions, or clients with special properties. The generalization that we consider in this paper is the Euclidean TSP with clients are parallel line segments with similar size, and the goal is to find a minimum cost tour that intersects with each segment at least once.

For several decades, the classic algorithm by Christofides [6] and independently by Serdyukov [23] that implies a $\frac{3}{2}$-approximation was the best known approximation for TSP until a recent result by Karlin et al. [17] that shows a slight improvement. Several generalizations (or special cases) of TSP have been studied as well. Perhaps the most notable special case is when the points are given in fixed dimensional Euclidean space. Arora and Mitchell [4], [20] presented different PTAS's for Euclidean TSP. There have been many papers that have extended these results. Arkin and Hassin [3] introduced the notion of TSP with Neighborhoods (TSPN). Notice that if every region is a single point, this problem reduces to the vanilla TSP. An instance of TSPN is a set of neighbourhoods or regions given in a metric space and the goal is to find a minimum cost tour that visits all these regions. Each region can be a single point or could be defined by a subset of points of the plane. They gave several constant-factor approximations for the geometric settings where each regions is some well-defined shape on the plane, such as disks, parallel unit length segments, and more generally, for regions which have diameter segments that are parallel to a common direction, and have bounded ratio of the largest to smallest diameter. Several papers have studied TSPN for various classes of objects (neighborhoods) and under different metrics.

TSPN is much more difficult than TSP in general and in special cases, just as group Steiner tree is much more difficult than Steiner tree (one can consider each neighborhood as a group/set from which at least one point needs to be visited). In group Steiner tree or group TSP, one is given a metric along with groups of terminals. The goal is to find a minimum cost Steiner tree (or a tour) that contains (or visits) at least one terminal from each group. Using the result of Halperin and Krauthgamer [15] for hardness of group Steiner tree, it follows that general TSPN is hard to approximate within a factor better than $\Omega(\log^{2-\varepsilon} n)$ for any $\varepsilon > 0$ even on tree metrics. The algorithms for group Steiner tree on trees by Garg et al. [14] and embedding of metrics onto tree metrics by Fakcharoenphol et al. [13], imply an $O(\log^3 n)$-approximation for TSPN in general metrics. Unlike Euclidean TSP (which has a PTAS), TSPN is **APX**-hard on the Euclidean plane as shown by Berg et al. [5]. The special case when each region is an arbitrary finite set of points in the Euclidean plane (also known as Group TSP) has no constant approximation [22] and the problem remains **APX**-hard even when each region consists of exactly two points [9].

Focusing on Euclidean metrics, most of the earlier work have studied the cases where the regions (or objects) are *fat*. Roughly speaking, it usually means the ratio of the smallest enclosing circle to the largest circle fitting inside the object is bounded. There are some work on when regions are *not* fat, most notably when the regions are (infinite) lines or line segments or in higher dimensions when they are hyperplanes. For the case of infinite line segments in $\mathbb{R}^2$, the problem for n lines can be solved exactly in $O(n^4 \log n)$ time by a reduction to the Shortest Watchman Route Problem (see [8], [16]). For the same setting, Dumitrescu and Mitchell [10] presented a linear time $\frac{\pi}{2}$-approximation which was later improved to $\sqrt{2}$ by Jonsson [16] (again in linear time). For infinite lines in higher dimensions (i.e. dimension $d \geq 3$), the problem is proved to be **APX**-hard (see Antoniadis et al. [2] and references there). For neighborhoods being hyperplanes and dimension being $d \geq 3$, Dumitrescu and Tóth [11] present a constant factor approximation (which grows exponentially with d). For arbitrary d, they present an $O(\log^3 n)$-approximation. For any fixed $d \geq 3$, Antoniadis et al. [1] present a PTAS.

For parallel (unit) line segments on the plane Arkin and Hassin [3] presented a $(3\sqrt{2}+1)$-factor approximation which was improved to $3\sqrt{2}$ by [10] and it remains the best known approximation for this case as far as we know for over two decades. Elbassioni et al. [12] proved that TSPN for unit line segments (in arbitrary orientation) is **APX**-hard.

In this thesis, we settle the approximability of TSPN when regions are parallel line segments of similar length (which includes unit length as a special case) and present a PTAS for it. As mentioned above, the best known approximation for unit length parallel segments has ratio $3\sqrt{2}$ [10]. We first focus on the case of unit line segments and show how our result extends to when line segments have bounded length ratio. This is in contrast with the **APX**-hardness of [12] when we have unit line segments with arbitrary orientation. Our result also implies a $(2+\varepsilon)$-approximation for the case where we have axis-parallel similar size line segments.

1.3 Our Results

We prove the following theorem in this thesis:

Theorem 1 *Given a set of n parallel line segments with lengths in $[1, \lambda]$ for a fixed λ as an instance of TSPN, there is an algorithm that finds a $(1+\varepsilon)$-approximation solution in time $n^{O(\lambda/\varepsilon^3)}$.*

The algorithm we present is randomized but can be easily derandomized (see Subsection 1.1.3). To simplify the presentation, we give the proof for the case of unit line segments, and then explain how the result can be extended to the case where the aspect ratio is bounded by λ at the end of the thesis.

This problem generalizes the classic (point) TSP (at a loss of $(1+\varepsilon)$ factor). To see this, note that for the special case of line TSP where the line segments are far apart, i.e. the diameter of the minimal bounding box is at least $\Omega(n/\varepsilon)$, scaling the plane by a factor of ε yields an instance where the line segments have length equal to ε and the diameter is $\Omega(n)$. Since in this case the optimum is at least $\Omega(n)$, replacing each line segment with a point and solving (point) TSP implies that the solutions for both instance (the line version and point version) are within $(1+\varepsilon)$-factor of each other.

With some modifications, we follow the paradigm of Arora [4] for designing a PTAS for classic Euclidean TSP, specifically for dissecting the problem into smaller problems and recursively solving them using Dynamic Programming (DP). The reader is encouraged to familiarize themself with that solution, as explaining all the details of that solution are outside of the scope of this thesis.

The difficult cases are when the line segments are not too far apart (for e.g. they can be packed in a box of size $O(\sqrt{n})$ or smaller). There are two key ingredients to our proof that we explain here. One may try to adapt the hierarchical decomposition of Arora [4] for the PTAS for classic (point) TSP (which works by dissecting the plane into squares and making the tours portal respecting and using DP to combine the solutions), to this setting. Following that hierarchical decomposition, the first issue is that some line segments might be crossing the horizontal dissecting lines and so we don't have independent sub-instances and it is not immediately clear in which subproblem these crossing segments must be covered. Note that the line segments might be spread in a grid fashion (e.g. $\sqrt{n}$ segments spaced equally over each of $\sqrt{n}$ many horizontal lines). So the number of line segments crossing a dissecting line can be large. Our first insight is the following:

Insight 1: *At a loss of* $(1+\varepsilon)$, *we can drop the line segments crossing horizontal* *dissecting lines and instead requiring a subset of portals of each square to be visited, provided we continue the quad-tree decomposition until each square has size* $\Theta(1/\varepsilon)$.

In other words, assuming all the squares in the decomposition have height at least $\Omega(1/\varepsilon)$, then at a small loss we can show a solution for the modified instance where line segments on the boundary of the squares are dropped, can be extended to a solution for the original instance. So proving this property allows us to work with the hierarchical (quad-tree) decomposition until squares of size $\Theta(1/\varepsilon)$. This can be proved by a proper packing argument. But then we need to be able to solve instances where the height is bounded by $O(1/\varepsilon)$. Let's define the notion of shadow of a solution (or in general, shadow of a collection of paths on the plane) as the maximum number of times a vertical line Γ intersects any of these paths. Our second insight is the following:

Insight 2: *If we consider a window that is a horizontal strip of height* $O(h)$ *and move this window vertically anywhere over an optimum solution, then the shadow of the parts of optimum visible in this strip is at most* $O(h)$.

In other words, one expects that in the base case of the decomposition (where squares have height $\Theta(1/\varepsilon)$) the shadow is bounded by $O(1/\varepsilon)$. Despite our efforts, proving this appears to be more difficult than thought and it seems there are examples where even in the unit length segments, the shadow may be large (see Figure 1.1).

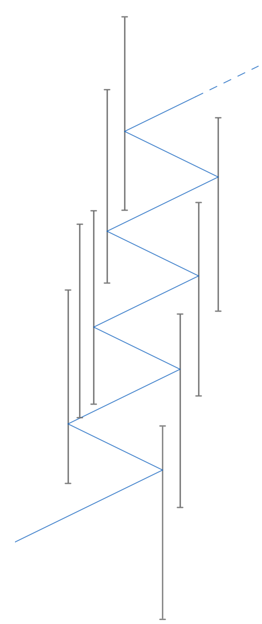


Figure 1.1: A potential arrangement of line segments where the solution has a large shadow

However, we are able to prove the following slightly weaker version that still allows us to prove the final result:

(Revised) Insight 2: *There is a* $(1+\varepsilon)$*-approximate solution such that the* *shadow of any strip of height h over that solution is bounded by $O(f(\varepsilon) \cdot h)$. for some function $f(\cdot)$.*

The proof of this insight forms bulk the of this thesis. To prove this, we characterize specific structures that would be responsible for having a large shadow in a solution and show how we can modify the solution so that for each of these structures the shadow is bounded by $O(1/\varepsilon)$ while increasing the cost by a $(1+\varepsilon)$ factor at most. This is formulated in the Theorem 2. For an instance with line segments of length 1, suppose opt is the cost of an optimum solution.

Theorem 2 *Given any $\varepsilon > 0$, there is a solution $\mathcal{O}'$ of cost at most $(1+\varepsilon) \cdot opt$ such that in any strip of height 1, the shadow of $\mathcal{O}'$ is $O(1/\varepsilon)$.*

We will show that this near optimum solution has further structural properties that allows us to solve the bounded height cases at the base cases of the hierarchical decomposition using a DP (later on, referred to as the inner DP). Proof of this theorem is fairly long and involves multiple steps that gradually proves structural properties for specific configurations.

Organization of the thesis: In Chapter 2, we define the problem and describe how to make changes to an optimum solution for a given instance of the problem to obtain specific structural properties. We start by proving some structural properties of an optimum and then a near-optimum solution in Sections 2.3 and 2.4; and finally prove Theorem 2 in Section 2.6. We describe the main algorithm in Chapter 3, which includes the outer DP (responsible for breaking the instance of a problem into smaller subproblems, then combining the answers) and inner DP (responsible for "solving" the base case subproblems). In Chapter 4, we summarize our results and mention some further problems one can consider next.

Chapter 2

Properties of a Structured Near-Optimum Solution

2.1 Problem Specification and Parameters

Suppose we are given n vertical line segments $s_1, \dots, s_n$ of length in the range $[1, \lambda]$, where the top and bottom points of each s_i are denoted by s_i^t and s_i^b, respectively. These end-points are also called *tips* of the segment. For any point p, let $x(p)$ and $y(p)$ denote the x and y-coordinates of p, respectively. Similarly, for any segment or vertical line s, let $x(s)$ denote its x-coordinate. For two points p, q, we use $||pq||$ to denote the Euclidean distance between them. A TSP tour on the plane is specified by a sequence of points where each of these points is on one of the segments of the instance such that each line segment has at least one such point, and the tour visits these points consecutively using straight lines. The line that connects two consecutive points in a tour is called a *leg* of the tour. In our problem, the goal is to find a TSP tour of minimum total length that touches (i.e. has an intersection with) each of these line segments. As mentioned earlier, we focus on the case where all line segments have length 1 and then show how the proof easily extends to the setting where they have lengths in $[1, \lambda]$. So from now on, all line segments are assumed to be unit length. Fix an optimum solution, which we refer to by OPT and use opt to refer to its cost. Our goal is to show the existence of a near-optimum (i.e. $(1 + \varepsilon)$-approximate) structured solution that allows us to find it using dynamic programming. We will state and prove a series of properties for OPT and later show how we can modify OPT to a near optimum solution with further structures.

First we show at a small loss we can assume all the line segments have different x-coordinates. We assume that the minimal bounding box of these line segments has length L and height H. For now, assume $H > 3$ (see Theorem 3). Let $B = \max\{L, H - 2\}$. So opt $\geq 2B$; we can also assume $B \leq \frac{n}{\varepsilon}$, because otherwise opt $\geq 2n/\varepsilon$ and if we consider an arbitrary point on each line segment (say the lower tip) and solve the classic TSP (using a PTAS) for these points, then it will be a PTAS for our original instance as well; that is because we pay at most an extra $+2$ for each line for a total of $2n$ which is $O(\varepsilon \cdot \text{opt})$. For a given $\epsilon > 0$, consider a grid on the plane with side length $\frac{\epsilon B}{n^2}$. Now move each line segment (parallel to the y-axis) so that the lower tip of each s_i is moved to the nearest grid point where there is no other line segment s_i with that x-coordinate. By doing this, all segments will have different x-coordinates and each segment would move at most $\frac{\sqrt{2}}{2} \cdot \frac{\varepsilon B}{n} < \frac{\epsilon B}{n}$, and in total, all segments would move at most a distance of ϵB. So the optimum value of the new instance has cost at most $(1 + \varepsilon) \cdot$ opt. For simplicity of notations, from now on we assume the original instance has this property and let OPT (and opt) refer to an optimum (and its value) of this modified instance.

As mentioned before, let the length of the sides of the minimal bounding box of an instance of the problem be $L \times H$. The following theorem holds:

Theorem 3 *If $H \leq 3$, then the shadow of an optimum solution is at most 2.*

We will not prove this theorem just yet, as we need some definitions and properties before we can prove it. In Section 2.5, we will prove this theorem. For now, assume that $H > 3$ for the lemmas and definitions in the following sections.

2.2 Structure Theorem

Our main goal is to prove Theorem 2. First we start by stating several properties for an optimum and later for a near-optimum solution.

In Section 2.3, we show the properties of an optimum solution; Subsection 2.3.1 includes the main property we want to prove. The lemma in Subsection 2.3.1, essentially proves that if we focus on a connected subpath of OPT in a bounded-height strip, then that subpath can be partitioned into disjoint parts made from structures called **sinks** and **zig-zags** (see Definition 8). We will leverage this property along with the properties proved its following subsection to find a structured solution with low complexity. Our notion of complexity is referred to as the **shadow** of the solution (see Definition 1). The shadow of a solution directly affects the size of the DP-table in Chapter 3.

In Section 2.4, we prove three main lemmas. The lemma in Subsection 2.4.1 shows that the aforementioned partitions in Subsection 2.3.1 (namely sinks and zig-zags) can be altered to a near-optimum solution such that each of them have a low complexity (more precisely, a constant shadow). In Subsection 2.4.3, we show that with some alterations, the number of subpaths of OPT that vertically overlap with each other (this is formally defined later on) can be bounded (at no extra cost) to a constant integer. All the alterations will lead to a solution with at most an $O(\varepsilon)$-factor increase in cost for the given $\varepsilon > 0$ and all the constant bounds are $O(1/\varepsilon)$ at worst. The three lemmas we mentioned so far, are alone enough to prove Theorem 2. The proof of that Theorem is in Section 2.6. There is an additional lemma in Subsection 2.4.2 that we later use in the DP for the problem. That lemma ensures that the number of "guesses" we need to take in our subproblems of the DP will be polynomially bounded. The near-optimum solution we provide will satisfy all these lemmas we mentioned. We will also prove Theorem 3 in Section 2.5.

2.3 Properties of an optimum Solution

We start by stating some lemmas that give a better understanding of the geometrical properties of an optimum solution, and later build up the proof of the lemma in subsection 2.3.1 from these properties.

One special instance of the problem is when there is a horizontal line that crosses all the input segments. This special case can be detected and solved easily. Otherwise, any optimum solution will visit at least 3 points that are not colinear. In such cases, like in the classic (point) TSP [4], we can assume the optimum does not cross itself, i.e. there are no two legs of optimum ℓ (between points p, q) and ℓ' (between points p', q') that intersect, as otherwise removing these two and adding the pair of $pq', p'q$ or pp', qq' will be a feasible solution of smaller cost.

Observation 1 *OPT is not self-crossing.*

Definition 1 *Given a collection $\mathcal{P}$ of paths on the plane and a vertical line at point $x_0 \in \mathbb{R}$, the* ***shadow*** *at x_0 is the number of legs of the paths in $\mathcal{P}$ that have an intersection with the vertical line at x_0. The shadow of a given range $[a, b]$ is defined to be the maximum shadow of any of values $x_0 \in [a, b]$.*

Note that if a solution is self-crossing, the operation of uncrossing (which reduces the cost) does not increase the shadow. Suppose the sequence of points of OPT is $p_1, p_2, \ldots, p_\sigma$ and the straight lines connecting these points (i.e. legs of OPT) are $\ell_1, \ell_2, \ldots, \ell_\sigma$ where ℓ_i connects two points p_i, p_{i+1} (with $p_{\sigma+1} = p_1$), and each s_i has at least one point p_j on it. We consider OPT oriented in this order, i.e. going from p_i to p_{i+1}. Since all segments have distinct x-coordinates, we can assume:

Observation 2 *No two consecutive points p_i, p_{i+1} can be on the same line segment of the instance (or else we can short-cut them), all points p_i on different line segments have distinct x-coordinates, and no leg ℓ_i is vertical.*

Definition 2 *Given a segment s of the problem (or any vertical line s) and a leg ℓ touching it (i.e. incident to a point on s), we say ℓ is to the* ***left*** *of s if ℓ is entirely in the subplane $x \leq x(s)$; and ℓ is to the* ***right*** *of s if ℓ is entirely in the subplane $x \geq x(s)$.*

Since there are no vertical legs, there is no leg that is both to the left and to the right of a segment of the instance at the same time.

Consider any segment s_i of the problem, and suppose that ℓ_j, ℓ_{j+1} are the two legs of OPT with common end-point p_j that is on s_i. Let s_i^t and s_i^b denote the top and the bottom tips of s_i. We consider 3 possible cases for the location of p_j and the arrangement of ℓ_j, ℓ_{j+1}. Informally, one possibility is that the two legs ℓ_j, ℓ_{j+1} form a straight line that crosses s_i at p_j; one possibility is that the two legs are touching s_i at one of its tips (i.e. $p_j = s_i^t$ or $p_j = s_i^b$) such that one is to the left and one is to the right of s_i and they don't make a straight line, and the third possibility is that the two legs ℓ_j, ℓ_{j+1} are on the same side (both left or both right) of s_i.

Observation 3 *Consider any segment s_i of the problem, and suppose that ℓ_j, ℓ_{j+1} are the two legs of OPT with common end-point p_j that is on s_i. Let s_i^t and s_i^b denote the top and the bottom points of s_i. Then either:*

- *(Straight point): the subpath of OPT going through p_{j-1}, p_j, p_{j+1} forms a straight line and ℓ_j and ℓ_{j+1} are on two sides (left/right) of s_i and $\angle \ell_i p_j \ell_{i+1} = \pi$; in this case p_j is called a* straight point*, or*
- *(Break point): p_j is a tip of s_i (i.e. $p_j = s_i^t$ or $p_j = s_i^b$), $\angle \ell_i p_j \ell_{i+1} \neq \pi$ and ℓ_j and ℓ_{j+1} are on two sides of s_i (one left and one right); in this case p_j is called a* break point*, or*
- *(Reflection point): both ℓ_j, ℓ_{j+1} are on the left or both are on the right of s_i; in this case p_j is called a* reflection point.

For the case of a reflection point p_j with two legs ℓ_j, ℓ_{j+1}, if both legs are to the left of the segment it is called a *left reflection* point and otherwise it is a *right reflection* point.

Also note that if ℓ_j, ℓ_{j+1} are on the two sides of s_i and $\angle \ell_i p_j \ell_{i+1} \neq \pi$, then p_j must be a tip or else we could move p_j slightly up or down and reduce the length of OPT (see Figure 2.1).

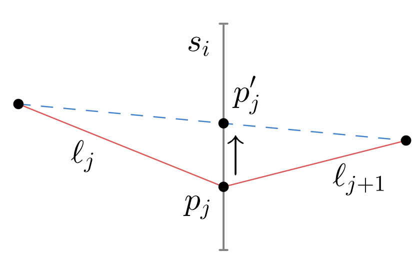


Figure 2.1: If p_j isn't a tip of s_i, then ℓ_j, ℓ_{j+1} must be collinear

We now state several lemmas about the structure of OPT.

Lemma 1 *If P is a subpath of OPT with end-points p, q where both are to the right of a vertical line Γ, and if P crosses Γ, then the left-most point on P to the left of Γ is a right reflection point (symmetric statement holds for opposite directions).*

Proof. Let r (on segment s) be the left-most point P visits, so both subpaths $P_{pr} = p \to r$ and $P_{qr} = q \to r$ are entirely to the right of r, in particular the two legs ℓ^- and ℓ^+ of P incident to r (which are the last two legs of the subpaths P_{pr}, P_{qr}) must be on the right of s which implies that r is a right reflection point. ∎

Definition 3 *Consider an arbitrary reflection point r on a segment s. Let the two legs of OPT incident to r visited before and after r (on the orientation of OPT) be ℓ^- and ℓ^+, respectively. ℓ^- is said to be* ***on top of*** *ℓ^+ if all the points of ℓ^- have larger y-coordinate than all of points of ℓ^+. In this case we also call ℓ^- the* upper leg *and ℓ^+ the* lower leg. *Also, in this case r is called a* ***descending*** *reflection point. If ℓ^+ is on top of ℓ^-, then r is called an* ***ascending*** *reflection point.*

Definition 4 *If ℓ_j, ℓ_{j+1} are two legs incident to a reflection point p on a segment s, if the angle between ℓ_j and s is the same as the angle between ℓ_{j+1} and s (i.e. ℓ_{j+1} is like the reflection of ray ℓ_j on mirror s), then p is called a* pure reflection *point.*

Lemma 2 *Any reflection point that is not a tip of a segment is a pure reflection point.*

Proof. Suppose p_j is a reflection point on s_i and is not a tip of it. If the two legs ℓ_j, ℓ_{j+1} don't have the same angle with s_i, then we can move p_j along s_i slightly up or down and one of the moves will decrease the cost of OPT, a contradiction. ∎

Lemma 3 *If a sweeping vertical line Γ moves left to right on the x-axis, the only values of x for which the shadow at Γ changes will be when Γ hits a reflection point on that x-coordinate. Specifically, this means that any subpath of OPT that doesn't contain a reflection, must have a shadow of 1 throughout its length. We say a subpath contains a reflection point p_j if p_j is not at the start or the end of the subpath (i.e. both legs of incident to p_j belong to that subpath.)*

Proof. According to Observation 3, we can see that straight points or break points will always contribute 1 to the shadow of Γ. But reflection points, depending on which direction the sweeping line moves, will either increase or decrease the shadow by 2. If a path doesn't contain any reflections, it means that it can only contain straight points or break points, meaning its shadow throughout its length will be equal to 1. ∎

Definition 5 *Let P_1 and P_2 be any two subpaths of OPT. We say P_1 is* ***above*** *P_2 in range $I = [x_0, x_1]$ if for every vertical line Γ with $x(\Gamma) \in I$, the top-most intersection of Γ with these two paths is a point on P_1. We say P_2 is* ***below*** *P_1 if the bottom-most intersection of Γ with P_1, P_2 is a point on P_2. Similarly, we say L_1 is to the* ***left*** *of L_2 in range $I' = [y_0, y_1]$ if for every horizontal line Λ with $y(\Lambda) \in I'$, the left-most intersection point of Λ with L_1, L_2 (i.e. one with the least x value) always belongs to L_1. We say L_2 is to the* ***right*** *of L_1 if the right-most intersection of Λ is with L_2.*

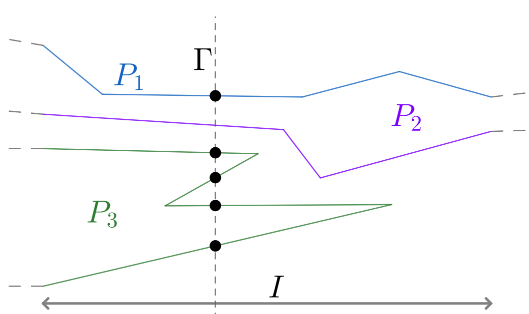


Figure 2.2: In range I, P_1 is above P_2, P_3, and P_2 is above P_3.

Lemma 4 *For any distinct points p_j and $p_{j'}$ on OPT, following OPT according to its orientation, either the path from p_j to $p_{j'}$ or the path from $p_{j'}$ to p_j must contain at least one reflection point.*

Proof. Without loss of generality, assume $x(p_j) < x(p_{j'})$, and following the orientation of OPT starting from p_j, suppose the path from p_j to $p_{j'}$ does not contain any reflection points (or the statement of lemma holds). According to Observation 3, the x-coordinate of points on OPT will not decrease if and only if the path contains only straight points or break points. The path from $p_{j'}$ to p_j has to have a decrease in the x-coordinate, due to $x(p_{j'}) > x(p_j)$, which is only possible if there is a reflection in this part of the path. ■

Lemma 5 *Let r_j be any reflection point on OPT, say it is a right reflection point, with incident legs ℓ_i, ℓ_{i+1}. Without loss of generality, assume that ℓ_i is above ℓ_{i+1}. Take any two subpaths P_1 and P_2 of OPT both starting at r_j with shadow of 1 such that $\ell_i \in P_1$ and $\ell_{i+1} \in P_2$. If there is a vertical line Γ with $x(\Gamma) > x(r_j)$ that intersects with both P_1 and P_2, then P_1 will be above P_2 in range $I = [x(r_j), x(\Gamma)]$.*

Proof. Note that for any vertical line Γ' with $x(\Gamma') \in I$, both P_1 and P_2 will intersect with it. Now assume the contrary, that P_1 is not above P_2. This means for some vertical line Γ' with $x(\Gamma') \in I$, there are points p_1 and p_2 on Γ' such that $p_1 \in P_1$, $p_2 \in P_2$, and $y(p_2) > y(p_1)$. Since both P_1 and P_2 have a shadow of 1, then using Lemma 3, we get that neither of them have a reflection point; this implies that the value of the x-coordinate on both P_1 and P_2 is monotone (or else there must be a reflection point). Since P_1 travels from r_j to p_1 and P_2 travels from r_j to p_2, both are crossing the same vertical lines (at $x = x(r_j)$ and Γ'). Now, because ℓ_i is above ℓ_{i+1} but p_1 is below p_2, we conclude that P_1 and P_2 will intersect with each other in the area between the vertical lines Γ' and $x = x(r_j)$. This is a contradiction, hence the lemma.
■

Lemma 6 *Among the set of points visited by OPT following its orientation, suppose $p_j, p_{j'}$, $j < j'$ (on segments s_i, $s_{i'}$, respectively) are two consecutive* *reflection points (i.e. no other reflection point exists in between them). Then p_j and $p_{j'}$ cannot be both left or both right reflection points. Furthermore, if s_i is to the left of $s_{i'}$ then p_j is a right reflection and $p_{j'}$ is a left reflection (the opposite holds if $s_{i'}$ is to the left of s_i).*

Proof. Without loss of generality, assume that s_i is to the left of $s_{i'}$, meaning $x(p_j) < x(p_{j'})$. By way of contradiction, first suppose both p_j and $p_{j'}$ are right reflection points, i.e. the two legs incident to p_j (ℓ_j, ℓ_{j+1}) and the two legs incident to $p_{j'}$ $(\ell_{j'}, \ell_{j'+1})$ are on the right of s_i and right of $s_{i'}$, respectively. This means following the orientation on OPT, along ℓ_j we have a decrease in x-coordinate, then following ℓ_{j+1} have an increase, then again following $\ell_{j'}$ have a decrease and following $\ell_{j'+1}$ have an increase. So the value of the x-coordinate isn't monotone in the subpath of OPT from p_j to $p_{j'}$ (excluding these two points themselves), because the legs ℓ_{j+1} and $\ell_{j'}$ are visited in this path in this order. Similar to the proof in Lemma 4, we see that this is only possible if there is a reflection point on this subpath, which contradicts the assumption that $p_j, p_{j'}$ are consecutive. Similar argument implies that we cannot have both $p_j, p_{j'}$ being left reflections or p_j being a left reflection and $p_{j'}$ being a right reflection; otherwise the leg after visiting p_j will have decreasing x-value while it will have to visit $p_{j'}$ eventually, which has a larger x-value. So the path from p_j to $p_{j'}$ must include another reflection point, again a contradiction. ■

Corollary 1 *Consecutive reflection points in OPT alternate between left reflections and right reflections.*

Lemma 7 *If segment s_i has a reflection point p_j on it, then it cannot have any other intersections with OPT (i.e. no other point p'_j of OPT can be on s_i).*

Proof. Assume otherwise, that a segment s_i contains a reflection point p_j with legs ℓ_j and ℓ_{j+1}, and another point $p_{j'}$ on s_i. We can by-pass p_j locally and reduce the length of OPT which would be a contradiction. More specifically, let $R^- \in \ell_j$ and $R^+ \in \ell_{j+1}$ be points on the legs that have a distance of $\delta > 0$

from p_j. By replacing the subpath $R^- \to p_j \to R^+$ with $R^- \to R^+$, the total cost of OPT will decrease, which gives us a contradiction. ■

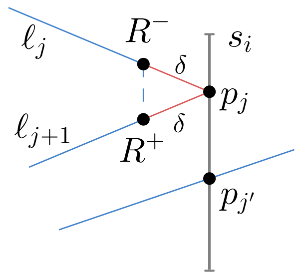


Figure 2.3: There can't be another $p_{j'} \in s_i$ if $p_j \in s_i$ is a reflection.

We decompose the problem into horizontal *strips* by considering some horizontal lines. Starting from the bottom tip of the top-most segment, draw horizontal lines that are 1-unit apart, these are called *cover-lines*. Each input segment is considered "covered" by the top-most (i.e. the first in this process) cover-line that intersects with it. Let's call these cover-lines $C_1, C_2, \ldots$ and so on.

Definition 6 (strip, top/bottom segments) *The region of the plane between two consecutive cover-lines $C_\tau, C_{\tau+1}$ is called a* strip *and denoted by S_τ. We consider $C_\tau, C_{\tau+1}$ to be parts of S_τ as well. The input line segments that are intersecting the top cover-line of S_τ (i.e. C_τ) are called* top segments *and the segments covered by the bottom cover-line (i.e. $C_{\tau+1}$) are called* bottom segments *of the strip.*

We show the near-optimum solution guaranteed by Theorem 2 has more structural properties that will be defined later. Note that once we prove this theorem, it follows that if we restrict a solution to $h > 1$ many strips, then the shadow is bounded by $O(h/\varepsilon)$ as well.

For now, let us focus on an (arbitrary) strip S_τ and imagine we cut the plane along $C_\tau, C_{\tau+1}$ and look at the pieces of line segments of the instance left inside this strip, along with pieces of OPT inside S_τ. Each top segment is now a partial segment in S_τ that has one end on C_τ and each bottom segment has one end on $C_{\tau+1}$. Let OPT_τ be the restriction of OPT to S_τ.

For each leg of OPT that intersects C_τ or $C_{\tau+1}$, we add a dummy point at the intersection(s) of that leg with C_τ and $C_{\tau+1}$ (so that the components of OPT_τ become consistent with our definition of legs). So OPT_τ can be seen as a collection of subpaths within S_τ (possibly along C_τ or $C_{\tau+1}$); following the orientation of OPT, each subpath of OPT_τ is when it intersects with S_τ, travels within S_τ (possibly along one of the cover-lines) until it exits S_τ. Using the dummy points added, each path in OPT_τ is a subpath of OPT that is between two points on cover-lines (these are called the entry points of the path with the strip. A formal definition is provided later on).

Recall Definition 5 of paths being above or below each other. Having the definition of top/bottom segments, we get the following:

Observation 4 *Consider OPT_τ, the restriction of OPT to any strip S_τ. Take any two subpaths of OPT_τ like P_1 and P_2 such that P_1 is above P_2 in some range I. If s_t is any top segment in range I that P_2 intersects with, then P_1 will also intersect with it. Similar statement holds for bottom segments if P_2 is below P_1.*

Definition 7 (entry points, loops, ladders) *For each subpath P_j of OPT_τ, let e_j and o_j be the first and last intersections of P_j with the interior of S_τ. Points e_j and o_j are called the* entry points *of P_j.*
If both e_j and o_j lie on the same cover-line (either C_τ or $C_{\tau+1}$), then P_j is called a loop*, otherwise it's called a* ladder*. If a subpath of OPT_τ enters S_τ at e_j on a cover-line and follows on that cover-line to point o_j and exits the strip, it is a special case of loop that we refer to as a* cover-line loop*.*

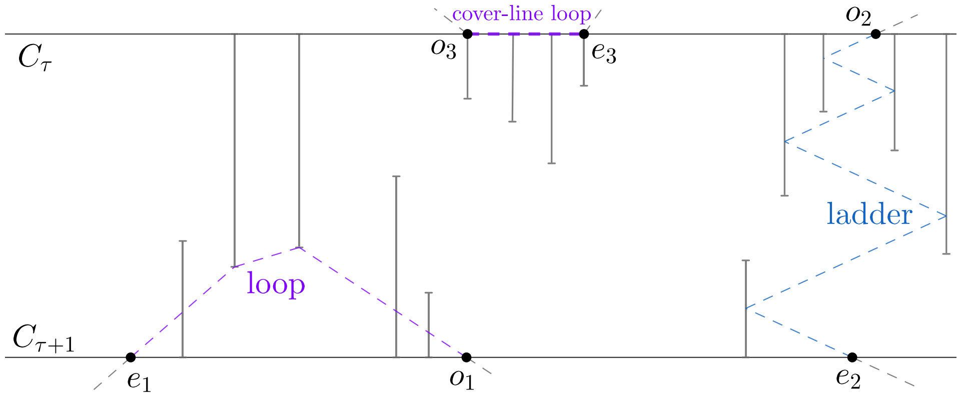


Figure 2.4: An example of loops and ladders in a strip S_τ

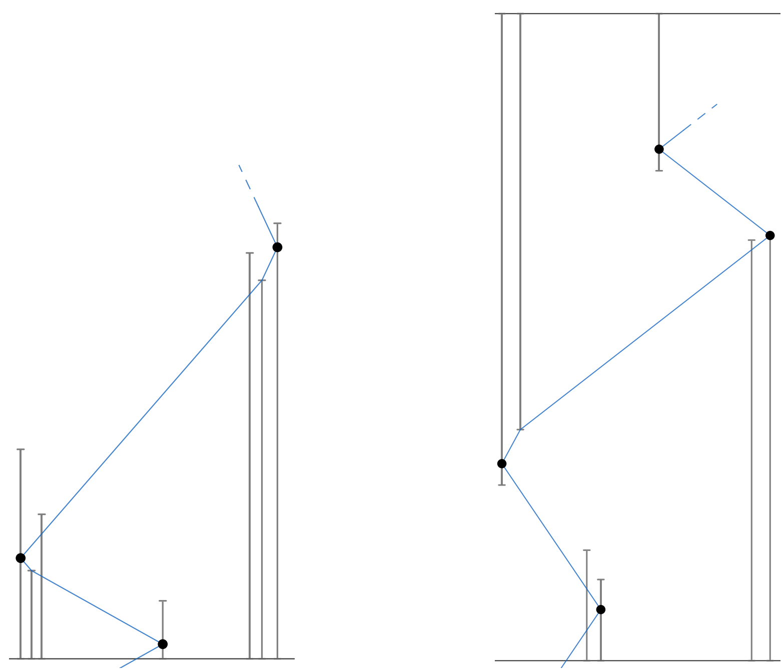
Since we're assuming $H > 3$ (see Theorem 3), we can assume that OPT is not limited to a single strip, and that it has to actually enter and exit any given strip that it intersects with (i.e. there is no strip that OPT completely lies inside it).

Note that if a path of OPT_τ is a cover-line loop, i.e a section of the line C_τ or $C_{\tau+1}$, then the entry points of that path must be the two end-points of this section. In other words, if for a cover-line loop of OPT_τ the first point is e_j on (say) C_τ, and the last point is o_j on C_τ, then this subpath must be traveling straight from e_j to o_j without any change of direction. This is true because otherwise, that cover-line loop would have to go back and forth on some portion on a cover-line, which is only possible if it's self-intersecting; but this is against our assumption that OPT is not self-crossing.

The two structures defined below (called a zig-zag and a sink) are the two configurations that can cause a large shadow.

Definition 8 (Zig-zag/Sink) *Consider any loop or ladder of OPT_τ, call it P. Let $\mathcal{R} = r_1, r_2, \ldots, r_m$ be the sequence of points of P that are reflection points (indexed by the order they're visited). Consider any maximal sub-sequence $r_j, r_{j+1}, \ldots, r_q$ of $\mathcal{R}$ with $q \geq 2$ such that the segments of the reflection points alternate between top and bottom segments and all are ascending or all are descending, then the subpath P that starts at r_j and ends at r_q is called a* ***zig-zag****.*

If $r_j, r_{j+1}, \ldots, r_q$ is a maximal sub-sequence of $\mathcal{R}$ that all belong to top segments *or all belong to bottom segments and are all ascending or all descending. The subpath P that starts at r_j and ends at r_q is called a* ***sink*** *(see Figure 2.5).*



(a) A sink (b) A zig-zag

Figure 2.5: Examples of sinks and zig-zags. The bold black dots represent the reflection points along these paths.

Using Corollary 1, the reflection points in a zig-zag or sink should alternate between left and right reflections.

Lemma 14 in Section 2.3.1 is used critically to show that very specific structures (made by zig-zags and sinks) are responsible for having large shadow along a ladder or loop in OPT_τ. And we can partition each ladder or loop into parts (subpaths), such that the shadow of the ladder/loop is equal to the maximum shadow among these parts; and that each part is a path consisting of up to three sinks and/or zig-zags. So the shadow of a loop/ladder is with $O(1)$ of the maximum shadow of zig-zag/sinks along that.

Before getting to the proof of Lemma 14, we still need to state some further lemmas and definitions.

Definition 9 *Let OPT_τ be the restriction of OPT to any strip S_τ. We say a segment $s \in S_\tau$ is* exclusively covered *by some path $P \in OPT_\tau$ if P covers s*

but no other subpath of OPT_τ intersects with s, i.e. OPT_τ/P doesn't intersect with it.

Lemma 8 *Each loop with entry points on $C_{\tau+1}$ in OPT_τ (i.e. bottom cover-line of S_τ) must exclusively cover a top segment, or else it must be a cover-line loop. Analogous argument holds for loops that have entry points on C_τ.*

Proof. Suppose P is a loop with entry points e_1, o_1 on $C_{\tau+1}$ that does not exclusively cover a point on a top segment. This implies if we change it to cover only bottom segments in S_τ, then the solution remains feasible. Let s_ℓ and s_r be the left-most and right-most bottom segments that P covers, let q_ℓ, q_r be intersections of s_ℓ and s_r with $C_{\tau+1}$, respectively. Replace P with e_1, q_ℓ, q_r, o_1 and then short-cut e_1, o_1 like the way we argued for cover-line loops after Definition 7. So we obtain a path that is shorter than the original, but is a cover-line loop and covers all the (bottom) segments P was covering. ■

Corollary 2 *If P is a non-cover-line loop with entry points on the bottom cover-line of some strip S_τ, then P has to exclusively cover some top segment in S_τ. Similar argument holds for bottom segments and non-cover-line loops with entry points on the top cover-line.*

Lemma 9 *Suppose that OPT_τ is crossing a vertical line Γ at least two times. Let p_1, p_2 be two such crossings and, L_1 be a subpath of OPT_τ from p_1 to p_2 with no other crossings with Γ. Then there cannot be any other crossings of OPT_τ with Γ on the section p_1p_2 of Γ.*

Proof. Without loss of generality, since L_1 doesn't intersect with Γ other than at points p_1 and p_2, assume that L_1 is on the left of Γ. By way of contradiction, suppose q_1 is another crossing of OPT_τ with Γ such that $y(p_1) < y(q_1) < y(p_2)$. This implies that there is a subpath of OPT_τ inside the region $A = L_1 \cup p_1p_2$ with one end-point being q_1. So there must be another crossing of OPT_τ with the region $A = L_1 \cup p_1p_2$; and since OPT_τ is not self-crossing, that other crossing point with A must be on p_1p_2, call it q_2. Let us denote the subpath of OPT_τ inside A with end-points q_1, q_2 by L_2. Let r_1 be the left-most point on

L_1. Since L_1 is a path from a point on Γ to the left of Γ and back to a point on Γ, using Lemma 1, r_1 must be a right reflection point. Similarly, if r_2 is the left-most point on L_2 then r_2 must be a right reflection point, say on segment s_{r_2} (see Figure 2.6). But since r_2 is inside A, then regardless of whether s_{r_2} is a top segment or a bottom segment it will intersect with L_1, contradicting Lemma 7. ■

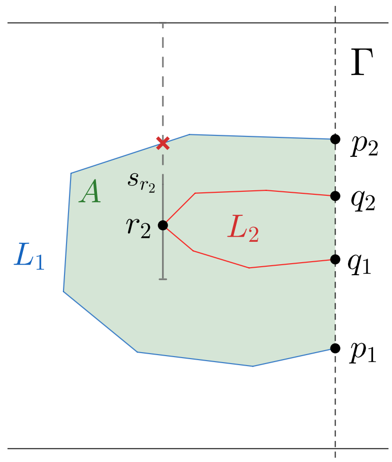


Figure 2.6: Configuration for Lemma 9

The following lemma is a special case of Lemma 9 but since it is used frequently, we state it as a separate lemma.

Lemma 10 *Consider a strip S_τ and OPT_τ (the restriction of OPT within this strip). Let s be any segment in this strip which has a reflection point p_j on it. Without loss of generality, assume s is a top segment and p_j is a left reflection point. Let ℓ_u and ℓ_l be the upper and lower legs of OPT_τ incident with p_j. Then the subpath of OPT_τ starting at p_j and traveling on ℓ_u, will not reach to the right side of s.*

Proof. Suppose the subpath of OPT_τ starting at p_j and traveling along ℓ_u, call it P_u, reaches the right side of s while entirely within strip S_τ. So P_u crosses the vertical line $x = x(s)$ at a point p inside S_τ (different from p_j). This path will be L_1 in the setting of Lemma 9 and p_j, p will be p_1, p_2 of the lemma. Consider the subpath P_l of OPT_τ starting at p_j and following ℓ_l. This subpath is in the region defined by P_u and the vertical line at $x = x(s)$. Since

OPT is non-self-crossing, P_l has to exit this area between the lower tip of s and point p. But this will violate Lemma 9. This contradiction results in the statement of the lemma. ■

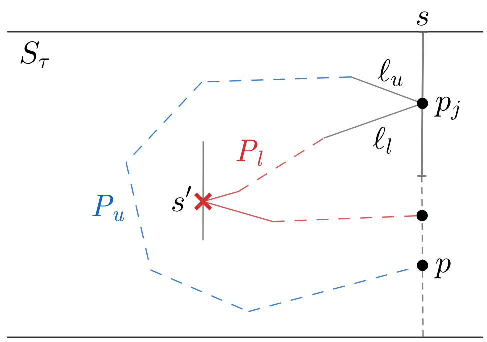


Figure 2.7: In a strip S_τ, the path from the upper leg of a left reflection on a top segment, can't reach to the right of that segment.

Lemma 11 *Suppose P_1 and P_2 are two ladders of OPT_τ in S_τ with entry points e_1 and e_2 on the bottom cover-line and entry points o_1, o_2 on the top cover-line, respectively, such that $x(e_1) < x(e_2)$, $x(o_1) < x(o_2)$ and both intersect a vertical line Γ to the right of e_1, e_2. Then P_1 is above P_2 to the left of Γ. (symmetric arguments apply to the top cover-line as well as entry points to the right of Γ)*

Proof. By way of contradiction, suppose P_1 is not above P_2 on the left of Γ, so there is a vertical line Γ' to the left of Γ whose top-most intersection is with P_2, say point p on Γ'. Consider the (vertical) segment of Γ' from p to the top cover-line, call it Γ'' and let the subpath of P_2 from e_2 to p be called P_2'. If we cut the strip S_τ along $P_2' \cup \Gamma''$, then e_1 is on one side, and o_1 on the other, which implies L_1 must be crossing $P_2' \cup \Gamma''$, which would be a contradiction (as it would have an intersection point on Γ' higher than p or has to cross P_2'). ■

Lemma 12 *Let P be any ladder or loop of OPT_τ in strip S_τ. Let r_{i_1} (on segment s_{m_1}) and r_{i_2} (on segment s_{m_2}) and r_{i_3} (on segment s_{m_3}) be any three consecutive reflections in the orientation of OPT_τ in that order. If $x(r_{i_2}) < x(r_{i_1}) < x(r_{i_3})$ and r_{i_2} is an ascending reflection, then s_{m_1} is a bottom segment and r_{i_1} is an ascending reflection. Symmetric argument applies for r_{i_2} being a* *descending reflection (for which case s_{m_1} will be a top segment and r_{i_1} will be descending).*

Proof. See Figure 2.8. According to Lemma 6, since r_{i_1} and r_{i_2} are consecutive reflections with $x(r_{i_2}) < x(r_{i_1})$, then r_{i_1} is a left reflection and r_{i_2} is a right reflection. Let $P_{1,2}$ be the subpath of P from r_{i_1} to r_{i_2}, and $P_{2,3}$ be the subpath of P from r_{i_2} to r_{i_3}. Since r_{i_2} is an ascending reflection, then $P_{1,2}$ contains the lower leg of r_{i_2}, and $P_{2,3}$ contains the upper leg of r_{i_2}.
Since r_{i_1} and r_{i_2} are two consecutive reflections with $x(r_{i_1}) > x(r_{i_2})$, this means that $P_{1,2}$ cannot reach to the left of r_{i_2} or to the right of r_{i_1}; because otherwise due to the difference in the x-coordinates, $P_{1,2}$ would require an additional reflection between r_{i_1} and r_{i_2}, which isn't possible.

This implies that the entirety of $P_{1,2}$, and specifically r_{i_1}, are in the region defined by $x = x(r_{i_2})$, $x = x(r_{i_1})$, and the path $P_{2,3}$. So $P_{1,2}$ is below $P_{2,3}$ in $I = [x(r_{i_2}), x(r_{i_1})]$.

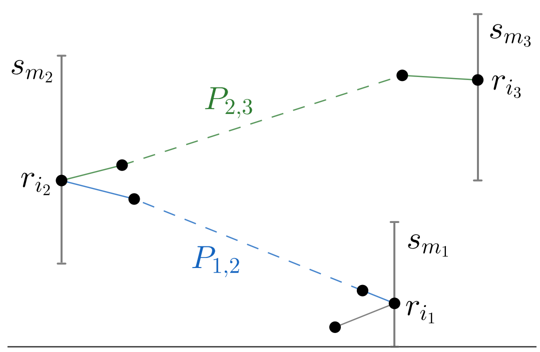


Figure 2.8: Valid arrangement of three consecutive reflections provided the x-coordinate of r_{i_1} is between the x-coordinates of r_{i_2} and r_{i_3}, and r_{i_2} is an ascending reflection. Segments s_{m_2} and s_{m_3} could either be top or bottom segments in this strip, but s_{m_1} must be a bottom segment.

Since $x(r_{i_2}) < x(r_{i_3})$ and $P_{2,3}$ is a path between these two reflections, we get that for any $x_0 \in [x(r_{i_2}), x(r_{i_3})]$, there is an intersection between $x = x_0$ and $P_{2,3}$. Now for the sake of contradiction, assume s_{m_1} is a top segment. Since r_{i_1} is below $P_{2,3}$, this would imply that s_{m_1} is intersecting with $P_{2,3}$. But this is in violation with Lemma 7. Thus, s_{m_1} must be a bottom segment. According to Lemma 10, r_{i_1} cannot be a descending reflection, because otherwise, $P_{1,2}$

would contain the lower leg of r_{i_1}; therefore, the path $P_{1,2} \cup P_{2,3}$ is a path that contains the lower leg of r_{i_1} and reaches to the right of segment s_{m_1}, which isn't possible. So we conclude that s_{m_1} is a bottom segment and furthermore, r_{i_1} is an ascending reflection. ∎

Lemma 13 *Suppose P is a loop or ladder of OPT_τ for a strip S_τ and $r_{i_1}, r_{i_2}, r_{i_3}$ are three reflection points visited in this order but not necessarily consecutively (following orientation of OPT), all are ascending (or all are descending) and are on segments $s_{m_1}, s_{m_2}, s_{m_3}$, respectively. Assume that r_{i_1}, r_{i_3} are left reflections and r_{i_2} is a right reflection and r_{i_2} is to the left of both r_{i_1} and r_{i_3}, i.e. $x(s_{m_2}) < x(s_{m_1})$ and $x(s_{m_2}) < x(s_{m_3})$.*
Let $P_{0,1}$ be the subpath of P up to r_{i_1}, $P_{1,2}$ be the subpath of P from r_{i_1} to r_{i_2}, $P_{2,3}$ be the subpath of P from r_{i_2} to r_{i_3}, and $P_{3,4}$ be the subpath of P from r_{i_3} to the end of P. Then we cannot have both $P_{0,1}$ and $P_{3,4}$ reach to the left of $x(s_{m_2})$.

Proof. Each of $P_{1,2}$ and $P_{2,3}$ include a leg of r_{i_2}; Without loss of generality, assume that the lower leg of r_{i_2} is in $P_{1,2}$, and its upper leg is in $P_{2,3}$ (i.e. assume that r_{i_2} is an ascending reflection). We take two cases based on whether s_{m_2} is a top segment or a bottom segment:

- s_{m_2} **is a top segment:** Path $P_2^u = P_{2,3} \cup P_{3,4}$ includes the upper leg of r_{i_2} (a right reflection) on a top segment s_{m_2}, so we can use the result of Lemma 10 to conclude that P_2^u and particularly $P_{3,4}$ can't reach to the left of s_{m_2}.

- s_{m_2} **is a bottom segment:** Path $P_2^l = P_{0,1} \cup P_{1,2}$ includes the lower leg of r_{i_2} (a right reflection) on a bottom segment s_{m_2}. Again, using Lemma 10, we get the same result that P_2^l and consequently $P_{0,1}$ can't reach to the left of s_{m_2}.

∎

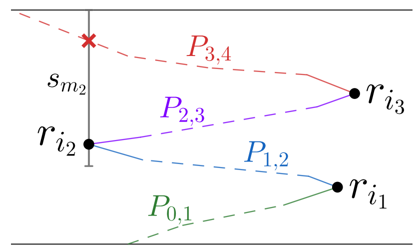


Figure 2.9: Configuration of Lemma 13 when s_{m_2} is a top segment

2.3.1 Vertically partitioning the solution in each strip

This subsection is dedicated to the proof of the following:

Lemma 14 *Consider any strip S_τ and any ladder or loop $P \in OPT_\tau$ within S_τ. Suppose the sequence of reflection points of P is $r_1, \ldots, r_q$. These reflection points can be partitioned into disjoint parts, say part i consists of reflection points $r_{a_i}, r_{a_i+1}, \ldots, r_{a_j}$, where the subpath of P from r_{a_i} to r_{a_j} is concatenation of up to three sections in the following order:*

a) A sink

b) A zig-zag

c) A sink

where any of these three sections can possibly be empty, and the last reflection of a section is common with the first reflection of the next section. Furthermore, for any vertical line Γ, there is at most one of these parts (of the partition) that intersects with it, i.e. the shadow of the ladder/loop is the maximum shadow among the parts plus 2.

The proof of this lemma is rather involved. To give an overview of the proof, we essentially show that for any loop or ladder in any strip, the vertical line at which the largest shadow for that loop or ladder happens, can intersect with at most two sinks and a zig-zag. So the shadow of a loop or ladder is $O(1)$ of the maximum shadow of the zig-zags and sinks along that.

If $q \leq 2$ then the correctness of lemma follows easily; so let's assume $q > 2$. Starting from $i = 1$, find the largest j such that the sequence $r_i, \ldots, r_j$ are all monotone, i.e. are all ascending or all are descending reflections. This will be the first part. We set $i = j+1$ and again, find the largest j such that $r_i, \ldots, r_j$ are monotone; this becomes the 2nd part. We repeat this procedure. So we find a partition into maximal sub-sequences of consecutive reflection points r_a's such that each sub-sequence contains only ascending or only descending reflections; each part might have just one point and the subpath path between the last reflection point of a part and the first reflection point of the next part has shadow 1 (since change in shadow can only happen if there is a reflection point by Lemma 3). The proof has two parts, we first show that each part can only have up to two sinks and possibly a zig-zag in between them, and then we show that for any vertical line, there is at most one part intersecting with it. Since the subpath from one part to another is a path between two consecutive reflections (and has shadow 1) the statement of the lemma will follow. We will use the following claim throughout this proof:

Claim 1 *For any subpath P of OPT_τ that is either a loop or a ladder, let r_j (on segment s_m) and $r_{j'}$ (on segment $s_{m'}$) be any two consecutive reflections along P. Without loss of generality assume that r_j is an ascending left reflection and in the orientation of OPT_τ, r_j comes before $r_{j'}$. If both s_m and $s_{m'}$ are bottom segments, then the subpath of P up to r_j (which includes the lower leg ℓ_j) can only contain reflection points lying on bottom segments. Analogous statement holds when both $s_m, s_{m'}$ are top segments.*

Proof. According to Lemma 6, $r_{j'}$ is a right reflection and $s_{m'}$ is to the left of s_m. Let P_j be the subpath of P from r_j to $r_{j'}$. Refer to the area of S_τ enclosed by s_m and $s_{m'}$ and below P_j by A_j; then ℓ_j lies inside A_j (see Figure 2.10). Let the subpath of P ending with leg ℓ_j be called P'_j. So P'_j is entirely within A_j as it cannot intersect with either of $s_m, s_{m'}$ (due to Lemma 7, since they both have reflection points) and P'_j cannot intersect P_j other than at r_j (since the solution is not self-crossing). So P'_j is below P_j within A_j. This implies any top segment that intersects P'_j must also intersect P_j due to Observation 4. So P'_j cannot have a reflection on a top segment by Lemma 7. So P'_j can have reflection points only on bottom segments. ■

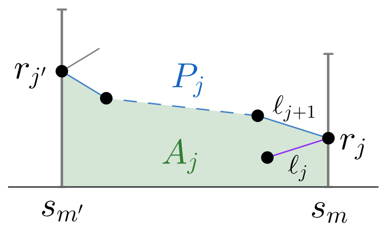


Figure 2.10: A_j is the area in S_τ surrounded by segments $s_m, s_{m'}$ and the subpath P_j

Now back to the proof of the lemma, we prove the following two parts:

1. Each partition can have up to two sinks and a zig-zag

Consider any part in the partition we defined before, which is a maximal subsequence of consecutive reflections that are all ascending or all descending. Our goal is to show this part is concatenation of a sink (possibly empty), followed by a zig-zag (possibly empty), followed by a sink (possibly empty), where the last point of the first sink is common with the first point of the zig-zag, and the last point of the zig-zag is common with the first point of the last sink. For simplicity, suppose the sequence of this part is $\mathcal{R} = r_1, \ldots, r_k$. Without loss of generality, assume $\mathcal{R}$ contains only ascending reflections and that the first one, r_1, is on a bottom segment. If all r_i's belong to bottom segments, then $\mathcal{R}$ is a sink and we are done. Otherwise, let j be the first index such that r_j is on a top segment (i.e. $r_1, \ldots, r_{j-1}$ are all on bottom segments). If $j = 2$, i.e. r_1 was a bottom and r_2 is a top segment, then the first sink is empty and this part starts with a zig-zag. If $j > 2$, then $r_1, \ldots, r_{j-1}$ is a sink. We argue that starting at $j' = \max\{1, j-1\}$, we can form a zig-zag. Let m be the largest index such that $r_{j'}, r_{j'+1}, \ldots, r_m$ is a zig-zag, i.e. the reflection points alternate between top and bottom segments. If no such m exists, it means $r_{j'}, \ldots, r_k$ all belong to top segments, giving us a sink; so this together with the first possible sink gives us two sinks at most, concluding the lemma. If $m = k$, then the partition has (up to) a single sink followed by a zig-zag, and we're done. Otherwise, $m < k$, meaning $r_{m+1} \in \mathcal{R}$. Since any zig-zag has at least 2 reflections, we have $m \geq 2$, meaning $r_{m-1} \in \mathcal{R}$ and since alternation between top and bottom ends at r_m, it means r_m and r_{m+1} are both either on top segments or both on bottom segments. We show that they can't be both on bottom segments. For the sake of contradiction, assume otherwise, i.e. both r_m and r_{m+1} are on bottom segments (and we assumed they are ascending). Also we know r_{m-1} is on a top segment (as it must be different from r_m). This violates Claim 1; because r_{m-1} is on a top segment and is on the subpath of OPT_τ reaching r_m. Thus both r_m, r_{m+1} are on top segments. Without loss of generality, assume that r_m is a left reflection; Lemma 7 implies r_{m+1} is right reflection with $x(r_{m+1}) < x(r_m)$. Let the path from r_m to r_{m+1} be P_m. Let A_m denote the area (of S_τ) bounded by P_m and between the segments containing r_m and r_{m+1}. If $\ell_{m''}, \ell_{m''+1}$ are the legs incident to r_{m+1} in the orientation of OPT, then $\ell_{m''+1}$ lies inside A_m (see Figure 2.11). Once again using Claim 1, we get that there can't be any reflections in the subpath in P starting at r_{m+1} through $\ell_{m''+1}$ that lie on a bottom segment. This implies all of $r_m, r_{m+1}, \ldots, r_k$ lie on top segments. Since all the remaining reflections are on top segments and all are ascending, this by definition means they form a sink. Thus, in total, we have up to a (bottom) sink, a zig-zag, and a (top) sink in this partition, concluding the first part of the proof.

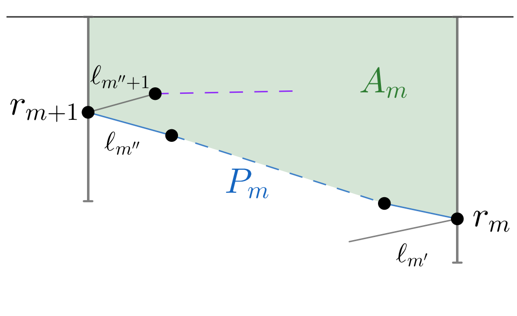


Figure 2.11: The upper leg of r_{m+1} lies inside A_m, and therefore, so does the rest of the path of OPT_τ until r_k.

2. Any vertical line can intersect at most one part

Recall that $r_1, r_2, \ldots, r_q$ denotes the sequence of *all* the reflection points on P (in strip S_τ). Let's call this $\mathcal{R}$. If $\mathcal{R}$ is made of only ascending or only descending reflections, we are done as it will have only one part in the partition. Otherwise, there must be two consecutive reflections $r_i, r_{i+1} \in \mathcal{R}$ such that one is an ascending reflection, but the other is descending. Suppose i is the first index that this happens. So the subpath from r_1 to r_i is one part, and r_{i+1} is the start point of another part. Note that the path from r_i to r_{i+1} has no reflection points; hence has shadow 1 because of Lemma 3. Without loss of generality, assume r_i is a right reflection and is an ascending reflection. Then r_{i+1} is descending and according to Lemma 6, it must be a left reflection as well with $x(r_i) < x(r_{i+1})$. Let P_i be the subpath of P from r_i to r_{i+1}. Since $q > 2$, we either have $i > 1$ or (if $i = 1$ then) $i + 1 < q$; meaning $r_{i-1} \in \mathcal{R}$ or $r_{i+2} \in \mathcal{R}$. In other words, there either is a reflection in $\mathcal{R}$ before r_i, or there is a reflection after r_{i+1}. Assume the first case holds, similar argument applies to the second one. Since r_i is a right reflection, using Lemma 6, we get that r_{i-1} is a left reflection with $x(r_i) < x(r_{i-1})$. We claim that we must have $x(r_{i-1}) < x(r_{i+1})$. For the sake of contradiction, assume otherwise. This means we have $x(r_i) < x(r_{i+1}) < x(r_{i-1})$. So we can use the result of Lemma 12 with parameters being $i_1 = i + 1,\ i_2 = i,\ i_3 = i - 1$ and following the points in reverse order of orientation, i.e. $r_{i+1} \to r_i \to r_{i-1}$ (this is the mirrored setting of Lemma 12). This implies r_{i+1} must be a descending reflection in the reverse orientation, which means it must be ascending in the original orientation (that ravels r_i to r_{i+1}). But we assumed r_{i+1} is descending. This contradicts Lemma 12 and proves our initial claim that $x(r_{i-1}) < x(r_{i+1})$. (see Figure 2.12). Thus,

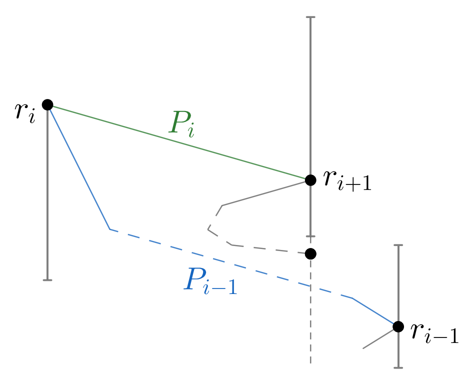


Figure 2.12: If between r_i and r_{i+1} one is ascending and the other is descending, then r_{i-1} (or r_{i+2}) must have an x-coordinate between $x(r_i)$ and $x(r_{i+1})$

$x(r_i) < x(r_{i-1}) < x(r_{i+1})$. Let s_j be the segment of the instance that r_{i-1} lies on. Once again, using Lemma 12, we get that r_{i-1} is an ascending reflection and s_j is a bottom segment (see Figure 2.13).

According to Lemma 10, we get that the subpath of P from r_1 to r_{i-1} (since it contains the lower leg of r_{i-1} due to it being an ascending reflection) can't reach to the right of s_j.

We will show that the subpath of P from r_{i+1} to r_k will not reach to the left of s_j either. This implies no vertical line can at the same time cross the first part that ends at r_i and the other parts starting at r_{i+1} onward. Repeating this argument implies no vertical line can intersect two parts as wanted. Consider the area surrounded by the line $x = x(s_j)$ and $P_{i-1} \cup P_i$, and refer to it by A_j. Now consider the subpath of P from r_{i+1} to r_k and refer to it as $\mathcal{P}_{i+1}$. Similar to the proof of Lemma 10, $\mathcal{P}_{i+1}$ can't enter A_j, because in order to exit from A_j, it has to reflect at some point inside A_j. But for such a reflection point to exists, there has to be a segment containing it, and that segment will intersect with P_{i-1} or P_i, which contradicts Lemma 7. So we conclude that if $\mathcal{P}_{i+1}$ were to go to the left of s_j, it has to do so from outside of A_j, i.e. from above P_i (since P_i is the upper hull of A_j).

Take two cases based on whether the segment $s_{j'}$ that contains r_{i+1} is a top segment or a bottom segment:

- $s_{j'}$ **is a top segment**:

  The area of S_τ is cut into two parts by $P_{i-1} \cup P_i \cup s_j \cup s_{j'}$. Since r_{i+1} is a descending reflection, then the lower leg of r_{i+1} is in the same part as the bottom tip of $s_{j'}$; refer to this part by A_1 and let A_2 be the other area. Since $\mathcal{P}_{i+1}$ includes this leg, it means that if $\mathcal{P}_{i+1}$ is going to reach to the left of s_j, it has to reach from A_1 to A_2. This would require it to either intersect with P_i or with $s_{j'}$. The former isn't possible because it would make OPT self-crossing, and the latter isn't possible because of Lemma 7.

- $s_{j'}$ **is a bottom segment**:

  The lower leg of r_{i+1} is in the area $A_{j'}$ surrounded by $P_{i-1} \cup P_i \cup s_j \cup s_{j'}$.

Since we mentioned $\mathcal{P}_{i+1}$ can't reach inside of A_j, then it needs to exit $A_{j'}$ and go over P_i. This means $\mathcal{P}_{i+1}$ has to either intersect with P_i or $s_{j'}$, which gives us the same contradictions as above.

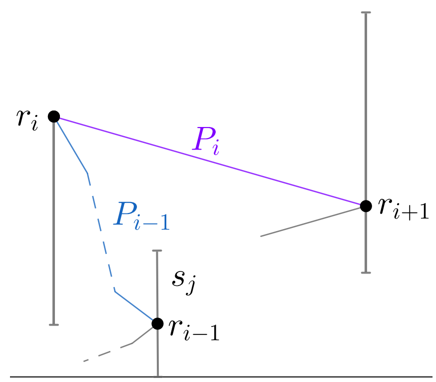


Figure 2.13: If r_i is an ascending reflection and r_{i+1} is a descending reflection, then r_{i-1} must be an ascending reflection lying on a bottom segment. r_i and r_{i+1} can be either on top or bottom segments.

So we conclude that there is no vertical line Γ that intersects both subpath $\mathcal{P}_i^- = \bigcup_{u=1}^{i} P_u$ and subpath $\mathcal{P}_{i+1} = \bigcup_{u=i+1}^{k} P_u$. Thus, $r_1, \ldots, r_i$ gives us a partition as desired. By continuing this process for the rest of the reflections, we get that no vertical line can intersect two parts. Since the path between two conseutive parts (last reflection of one part and the first reflection of the next part) has shadow 1, this completes the proof of the last part of Lemma 14. ∎

2.4 Properties of a Near optimum Solution

As mentioned before, we prove three main lemmas in subsections 2.4.1, 2.4.2, and 2.4.3. In this section, we make alterations to an assumed optimum solution such that some new structural properties hold; we ensure that the alterations have a limited additional cost.

Before getting to the main lemmas of this section, we need a few more definitions and lemmas. Note that some of the lemmas we prove here apply to any optimum solution, but we put them in this section (rather than Section 2.3) due to their connection to the near-optimum configuration.

Lemma 15 *Let $\mathcal{R} = r_1, r_2, \ldots, r_k$ denote the reflection points for any zig-zag or sink in a strip S_τ where $k \geq 3$. Without loss of generality, assume r_1 is on a bottom segment and is an ascending left reflection point. Then:*

- *If $\mathcal{R}$ is a zig-zag, then $x(r_1) < x(r_3) < \cdots < x(r_{2i-1}) < \ldots$ and $x(r_2) < x(r_4) < \cdots < x(r_{2i}) \ldots$.*
  All the inequalities hold in the other direction if r_1 is a right reflection point.

- *If $\mathcal{R}$ is a sink then $x(r_1) < x(r_3) < \cdots < x(r_{2i-1}) < \ldots$ and $x(r_2) > x(r_4) > \cdots > x(r_{2i}) \ldots$.*
  Again, all the inequalities hold in the other direction if r_1 is a right reflection point.

Proof. Assume $r_1, \ldots, r_k$ lie on segments $s_{i_1}, s_{i_2}, \ldots, s_{i_k}$, respectively. Also let's denote the path (following the orientation of OPT) from r_m to r_{m+1} by P_m. By definition, all P_m's are monotone in the x-coordinates (see Lemma 3).

First, consider the case that $\mathcal{R}$ is a zig-zag. Since r_1 is a left reflection and s_{i_1} is a bottom segment, and all reflection points are ascending, it means r_2 is a right reflection to the left of r_1 (because of Lemma 6), and s_{i_2} is a top segment. This implies P_1 is a decreasing path in the x-coordinate. Once again using Lemma 6, since r_2 is a right reflection, we have $x(r_3) > x(r_2)$. We claim that $x(r_3) > x(r_1)$. If this is not the case, then we have $x(r_2) < x(r_3) < x(r_1)$. Using Lemma 12 for parameters $r_{i_1} = r_3, r_{i_2} = r_2$, and $r_{i_3} = r_1$ in the order $r_3 \to r_2 \to r_1$ (which makes these reflections descending), implies that s_{i_3} must be a top segment, which is a contradiction. So we get $x(r_3) > x(r_1)$. Analogous argument shows that we must have $x(r_2) < x(r_4)$. Iteratively applying this argument establishes the inequalities.

Now consider the case that $\mathcal{R}$ is a sink. The argument is very similar to the case of zig-zag. Note that in this case, all the segments $s_{i_1}, \ldots, s_{i_k}$ are now bottom segments, all the reflection points are ascending and they must alternate between left and right reflection points. Since r_1 is a left reflection, r_2 is a right reflection with $x(r_2) < x(r_1)$ (due to Lemma 6). We again have

$x(r_3) > x(r_2)$ because of Lemma 6. Again, if we have $x(r_3) < x(r_1)$, then we have $x(r_2) < x(r_3) < x(r_2)$. Using Lemma 12 for parameters $r_{i_1} = r_3, r_{i_2} = r_2$, and $r_{i_3} = r_1$ in the order $r_3 \to r_2 \to r_1$ (which means the reflections are descending) implies s_{i_3} is a top segment, a contradiction. Thus, we get $x(r_3) > x(r_1)$. Similar argument shows that $x(r_2) > x(r_4)$, otherwise by an application of Lemma 6, s_{i_4} must be a top segment, which contradicts the assumption of a sink. By iteratively applying the same argument, we obtain the inequalities stated. ■

Lemma 16 *If $p_j, p_{j'}$ are consecutive reflection points in OPT, and both are pure reflections and all the other points of OPT in between them (if any) are straight points, then either both $p_j, p_{j'}$ are ascending or both are descending.*

Proof. By way of contradiction, suppose p_j is ascending and $p_{j'}$ is descending. Note that one is a left reflection and the other is a right reflection (as reflection points must alternate). Suppose p_j is a point on segment s_i, and $p_{j'}$ is on segment $s_{i'}$. From the assumption, the path from p_j to $p_{j'}$ is a straight line. Let ℓ_j, ℓ_{j+1} be the two legs incident to p_j and $\ell_{j'}, \ell_{j'+1}$ be the two legs incident to $p_{j'}$. From the definition of pure reflection, we need to have the angle between ℓ_j and s_i and the angle between ℓ_{j+1} and s_i be the same, and the angle between $\ell_{j'}$ and $s_{i'}$ and the angle between $\ell_{j'+1}$ and $s_{i'}$ be the same. The only way this is possible is when $\ell_j, \ell_{j+1}, \ell_{j'+1}$ are all horizontal but this means OPT is self-crossing. This contradiction yeilds the result of the lemma. ■

Lemma 17 *Let $\mathcal{R} = r_0, r_1, \ldots, r_k$ be any sequence of reflections that form a sink or zig-zag in a strip S_τ. For $1 \leq j \leq k$ let P_j be the subpath of $\mathcal{R}$ between r_{j-1} to r_j. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_k\}$. Take any vertical line Γ and let $P_\Gamma = \{P_{j_1}, P_{j_2}, \ldots, P_{j_m}\}$ $(j_1 < j_2 < \cdots < j_m)$ be the maximal subset of $\mathcal{P}$ that each $P_j \in P_\Gamma$ intersect with Γ. Then P_Γ must be a consecutive subset of $\mathcal{P}$. In other words, $P_\Gamma = \{P_{j_1}, P_{j_1+1}, P_{j_1+2}, \ldots, P_{j_1+m-1}\}$*

Proof. We say a reflection r_j is *included* in P_Γ if $P_j \in P_\Gamma$ or $P_{j+1} \in P_\Gamma$. We prove the following claim to use throughout this proof:

Claim 2 *There are no included left reflections to the left of* Γ *(and similarly no included right reflections to the right of it).*

Proof of Claim. Assume the contrary, that there is some left reflection r_{j_i} included in P_Γ that is to the left of Γ. Without loss of generality, assume that $P_{j_i} \in P_\Gamma$. Similar to the proof of Lemma 4, the points on P_{j_i} are monotone in the x-coordinate. This means the path from r_{j_i} on P_{j_i}, is decreasing in the x-coordinate (because r_{j_i} has its legs facing left), implying P_{j_i} is completely to the left of r_{j_i}. Since Γ is to the right of r_{j_i}, this means that P_{j_i} can't intersect with Γ, contradicting the assumption that $P_{j_i} \in P_\Gamma$. This proves the claim. ■

Now back to the statement of the lemma; Without loss of generality, assume that reflections in $\mathcal{R}$ are all ascending. For the sake of contradiction, assume that there is an index $1 \leq a < m$ for which P_{j_a} and $P_{j_{a+1}}$ aren't consecutive. This means $j_a < j_{a+1} - 1$, and so we conclude that the subpath $P' = \bigcup_{j'=j_a+1}^{j_{a+1}-1} P_{j'}$ of $\mathcal{R}$ from r_{j_a} to $r_{j_{a+1}} - 1$ is on one side of Γ (or else there will be another $P_{j'} \in P_\Gamma$ with $j_a < j' < j_{a+1}$). So there is at least one reflection point from $\mathcal{R}$ that is in P'. Let r_i be the first reflection on P' after r_{j_a}. Without loss of generality, assume that r_{j_a} (and therefore the entirety of P') is on the right side of Γ. So r_{j_a} is a left reflection because of Claim 2. By Lemma 6, both r_i and r_{j_a-1} (the reflection in $\mathcal{R}$ before r_{j_a}) are right reflections.

Let r_q be the end-point of $P_{j_{a+1}}$ that is to the left of Γ (either $r_q = r_{j_{a+1}}$ or $r_q = r_{j_{a+1}-1}$). Once again, using Claim 2, we get that r_q is a right reflection. So we have three right reflections r_{j_a-1}, r_i, and r_q such that $x(r_i) \geq x(\Gamma) \geq \{x(r_{j_a-1}), x(r_q)\}$ and the order they're visited in $\mathcal{R}$ is r_{j_a-1}, then r_i, and then r_q. According to Lemma 15, based on whether $\mathcal{R}$ is a sink or a zig-zag, we either must have $x(r_{j_a-1}) < x(r_i) < x(r_q)$ or the reversed inequality; which neither are the case here. This contradiction implies the statement of the lemma. ■

2.4.1 Bounding the Shadow of each Sink/Zig-zag

In this section, we prove the following lemma:

Lemma 18 *Consider OPT_τ for an arbitrary strip S_τ and let opt_τ be the total cost of OPT_τ. Given any $\varepsilon > 0$, we can change OPT_τ to a solution of cost at most $(1 + O(\varepsilon))opt_\tau$ where the shadow of each zig-zag and sink is at most $O(1/\varepsilon)$.*

Let $\sigma = \lceil 1/\varepsilon \rceil + 1$ and consider any loop or ladder $P \in \mathrm{OPT}_\tau$ and let $\mathcal{R}$ be an arbitrary zig-zag/sink along P with shadow larger than σ at some vertical line $x = x_0$. Without loss of generality, assume that following the orientation of OPT along P, reflection points on $\mathcal{R}$ are ascending. Suppose that the subpath of P following reflection points $r_j, r_{j+1}, \ldots, r_k$ of $\mathcal{R}$ is crossing $x = x_0$ (note that using Lemma 17, the reflection points must be consecutive). Let this subpath of P starting at r_j and ending at r_k be $\mathcal{R}'$ and let $s_{a_j}, s_{a_{j+1}}, \ldots, s_{a_k}$ denote the segments that contain reflections $r_j, r_{j+1}, \ldots, r_{j+k}$, respectively. Also let P_i (for $1 \leq i \leq k - j$) be the subpath of $\mathcal{R}'$ from r_{j+i-1} to r_{j+i}. Note that there might be several straight points or break points between $r_{j'}, r_{j'+1}$ on P (for each j'); the segments of these points are all covered by the shadow one path (due to Lemma 3) from $r_{j'}$ to $r_{j'+1}$. According to Lemma 5, since all reflections are ascending, if $m_1 < m_2$, then P_{m_1} is below P_{m_2} (in the range that P_{m_1} is defined on the x-axis). So this specifically implies P_1 is below any other P_m (in the range that P_1 is defined), and similarly, P_k is above any other P_m (in the range that P_k is defined). Note that $\mathcal{R}'$ is part of a zig-zag/sink itself (the only difference with the definition of zig-zag/sink is that $\mathcal{R}'$ is no longer necessarily maximal in OPT_τ). Also, note that for each path P_i, the x value of the points it visits between the two reflection points r_{j+i-1} to r_{j+i} are monotone increasing or decreasing (see Lemma 3). Let Ψ denote the cost of legs of $\mathcal{R}'$. It follows that $r_j, r_{j+2}, r_{j+4}, \ldots$ are on one side of x_0 (say to the right) and $r_{j+1}, r_{j+3}, \ldots$ are on the other side (say left of x_0). Since the number of reflections to the right of $x = x_0$ differs from the number of reflections to the left of $x = x_0$ by at most 1, then on each side of $x = x_0$ we have at least $(\sigma - 1)/2 = \lceil 1/2\varepsilon \rceil$ reflections. Let $\sigma' = \lceil 1/2\varepsilon \rceil$. The idea of the proof is to show that aside from the $2\sigma'$ reflections at the end of of $\mathcal{R}'$ (i.e. the last $2\sigma'$ paths P_j), we can replace the paths between the rest of the reflection points so that it reduces the shadow of the entire $\mathcal{R}'$ to $O(1/\varepsilon)$ while increasing the cost of the path by at most $O(\varepsilon \cdot \Psi)$.

Note that using Lemma 3, each subpath P_j of $\mathcal{R}'$ is between two consecutive reflection points and so has a shadow of 1. This implies that $P_{k-2\sigma'} \cup \ldots \cup P_{k-1}$ has a shadow of $O(\frac{1}{\varepsilon})$ as it has $2\sigma'$ consecutive reflections. We replace the rest of $\mathcal{R}$ (as we describe below) with a new path of a shadow of $O(1)$; this will yield the result of the lemma.

When $\mathcal{R}'$ is a part of a zig-zag

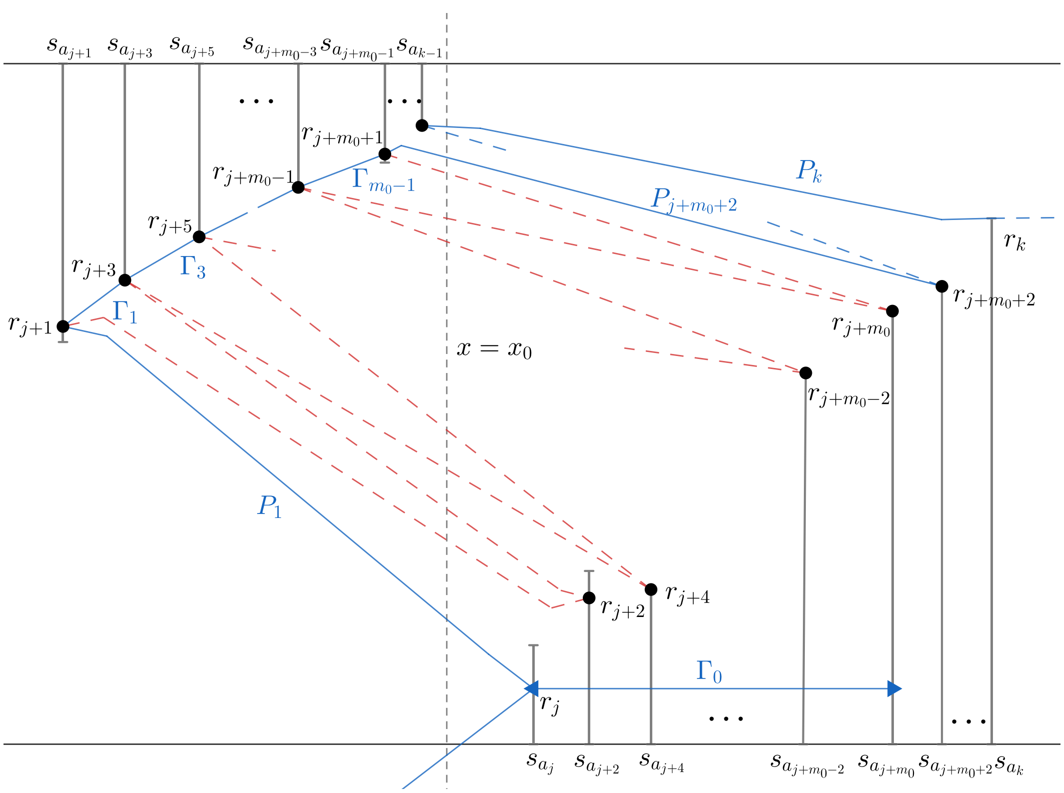


Figure 2.14: Alternative path for a zig-zag; The red parts are discarded. There are further details about Γ_m's that are explained throughout the proof of Lemma 18.

Without loss of generality, assume that $s_{a_j}, s_{a_{j+2}}, \ldots$ are bottom segments and to the right of $x = x_0$, and consequently, $s_{a_{j+1}}, s_{a_{j+3}}, \ldots$ are top segments and to the left of $x = x_0$. Let d_m for $m = 0, 1, \ldots, k-j$ denote $|x(r_{j+m}) - x_0|$. Using Lemma 15, we have $d_0 < d_2 < \cdots$ and $d_1 > d_3 > \cdots$. Let's focus on the right side of $x = x_0$ (where the bottom segments are), so $r_j, r_{j+2}, \ldots, r_{j+2q}$ are all the reflections of $\mathcal{R}'$ on this side where $q \geq \sigma' - 1$. We claim that except for at most the σ' largest values in $d_2, d_4, d_6, \ldots$, all other values of d_{2m}'s are at most $\varepsilon \cdot \Psi$ and this is done by an averaging argument. More specifically, we show that the largest integer $m_0 \in \{0, 2, 4, \ldots, 2q\}$ for which we have $d_{j+m_0} \leq \varepsilon \cdot \Psi$, has value $m_0 \geq 2(q - (\sigma' - 1))$. To see why this is the case, assume otherwise, that for all even integers $m \geq 2(q - (\sigma' - 1))$, we have $d_{j+m} > \varepsilon \cdot \Psi$. Adding these inequalities for $m = 2(q - (\sigma' - 1)),\ 2(q - (\sigma' - 2)),\ \ldots,\ 2q$ give us

$$\begin{aligned} d_{j+2(q-(\sigma'-1))} + d_{j+2(q-(\sigma'-2))} + \cdots + d_{j+2(q)} &> \sigma' \cdot (\varepsilon \cdot \Psi) \\ &= \lceil 1/2\varepsilon \rceil \cdot \varepsilon \cdot \Psi \\ &\geq \Psi/2, \end{aligned}$$

which clearly isn't possible, due to $2 \sum_{m \in \{2(q-(\sigma'-1)), \ldots, 2q\}} d_{j+m} \leq \Psi$; this inequality holds because paths P_{j+m}, P_{j+m+1} ($m \in \{2(q-(\sigma'-1)), \ldots, 2q\}$) have to travel the x-distance from x_0 to $s_{a_{j+m}}$ to the reflection points r_{j+m}, and all these paths are part of $\mathcal{R}'$. This contradiction shows our initial claim, that for some $m_0 \geq 2(q - (\sigma' - 1))$, we have all of $d_j, d_{j+2}, \ldots, d_{j+m_0} \leq \varepsilon \cdot \Psi$.

We are going to change $\mathcal{R}'$ from r_j up to r_{j+m_0}, but keep P_{j+m_0+1} and after; this change will result in another feasible solution with an $O(1)$ shadow up to r_{j+m_0}, and cost increase will be at most $O(\varepsilon \cdot \Psi)$. Our modification of $\mathcal{R}'$ is informally as follows (skipping some details to be explained soon). Starting at r_j instead of following P_1 to r_{j+1}, we first travel horizontally to the right until we hit $s_{a_{j+m_0}}$ (the bottom segment which r_{j+m_0} is located on), and travel back to r_j. Let's call this horizontal back and forth subpath Γ. This subpath Γ will ensure that all the bottom segments that $\mathcal{R}'$ covers between $x = x_0$ and $s_{a_{j+m_0}}$ are covered (we may need to deviate from Γ further down if $\mathcal{R}'$ goes further below Γ at some point; will formalize this soon). The shadow of Γ will easily be shown to be 2. Then from r_j, we follow P_1 and go to r_{j+1} which is the left-most reflection on a top segment (to the left of $x = x_0$). Now instead of following P_2 to go to r_{j+2} and then P_3 to go to r_{j+3}, we go straight from r_{j+1} to r_{j+3} (with some little details skipped here), then to r_{j+5} and so on until we get to r_{j+m_0+1}, and from there we follow $\mathcal{R}'$. One observation is that the shadow of the new path from r_{j+1} to r_{j+m_0+1} is also 1 since it won't have any reflection points. The rest of the path from r_{j+m_0+1} to r_k that follows $\mathcal{R}'$ has at most $O(\sigma')$ reflection points and hence the shadow is $O(1/\varepsilon)$. We show that the new path hits all the segments $\mathcal{R}'$ was hitting; and so we still have a feasible solution where the overall increase in the cost is at most $O(d_{j+m_0})$, which is bounded by $O(\varepsilon \cdot \Psi)$. Hence we find a modification of the path $\mathcal{R}'$ with shadow bounded by $O(1/\varepsilon)$, and cost increase is at most $O(\varepsilon \cdot \Psi)$. Note that any bottom segment (if any) to the left of $x = x_0$ that was covered by $\mathcal{R}'$, must intersect P_1; as P_1 is below the rest of $\mathcal{R}'$ to the left of $x = x_0$. Thus, any bottom segment to the left of x_0 that is covered by any of the $P_{b>1}$, is also covered by P_1. There are some details missing in this informal description that are explained below.

We will introduce a new subpath Γ_0, responsible for covering all bottom segments in $\mathcal{R}'$ to the right of $x = x_0$ until s_{j+m_0}; and we introduce a collection of subpaths Γ_m for odd m in $\{1, \ldots, m_0\}$ for covering the top segments to the left of $x = x_0$. All of Γ_m's, will have a shadow of 1. We ensure that any bottom segment hit by $\mathcal{R}'$ between s_{j_0} and s_{j+m_0}, is also hit by Γ_0 between $x = x(s_j)$ and $x = x(s_{j+m_0})$; and also any top segments that $\mathcal{R}'$ was hitting in the range that each Γ_m is defined, is hit by Γ_m, for $m \geq 1$.

Consider the horizontal line $y = y(r_j)$ from s_{a_j} to $s_{a_{j+m_0}}$. Refer to this horizontal portion as Γ. For reflection points $r_j, r_{j+2}, \ldots, r_{j+m_0}$ (on bottom segments $s_{a_j}, s_{a_{j+2}}, \ldots, s_{a_{j+m_0}}$), the two paths that contain a leg incident to r_{j+m} are P_{j+m} and P_{j+m+1} for each $0 \leq m \leq m_0$. Recall that using Lemma 5, P_{j+m} is below P_{j+m+1} between s_{a_j+m-1} and s_{a_j+m}. Consider the area A_Γ of the strip bounded by $\Gamma \cup s_{a_j} \cup s_{a_{j+m_0}}$. Then $\mathcal{R}' \cap A_\Gamma$ are (possibly empty) subpaths that start and end at Γ. These subpaths form the lower-envelope of $\mathcal{R}' \cup \Gamma$ in A_Γ (for e.g. in Figure 2.15, paths P_4, P_5 that reach r_{j+4}, cross Γ at points q_1^4, q_2^4.).

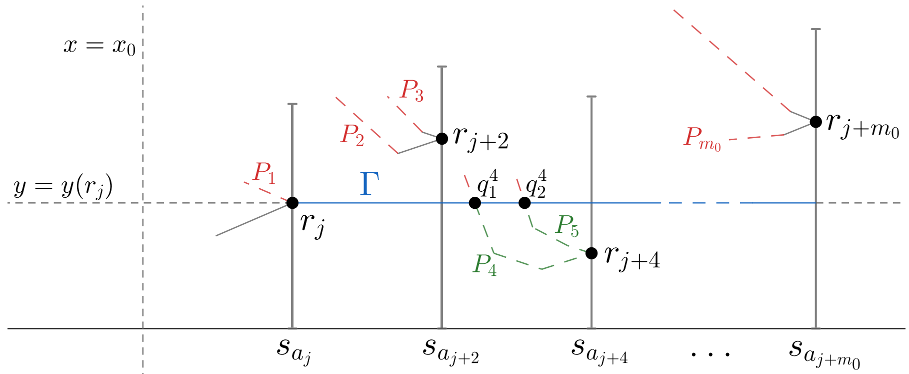


Figure 2.15: Γ_0 is the lower envelope of the blue line (the segment Γ) together with the green parts (portions of OPT_τ that go below Γ)

We define Γ_0 to be a path starting at r_j that travels right along Γ and the lower envelope of $\mathcal{R}'$ in this area, i.e. whenever traveling right on Γ, if we arrive at an intersection of $\mathcal{R}'$ with Γ (say a path P_{j+m}) then we travel along P_{j+m} inside A_Γ until we hit back at Γ, and then continue traveling right. For instance, in Figure 2.15, when traveling on Γ from r_j to right, once we arrive at q_1^4, we follow P_4 to r_{j+4}, then follow P_5 to q_2^4, and then continue right on Γ. Once we arrive at $s_{a_{j+m_0}}$, we travel Γ horizontally back to r_j. The length of Γ_0 can be bounded by the length of $\mathcal{R}' \cap A_\Gamma$ plus $2d_{j+m_0} \leq 2\varepsilon\Psi$. Also, it can be seen that any bottom segment that was covered by $\mathcal{R}'$ in between $x(r_j)$ and $x(r_{j+m_0})$, is covered by Γ_0 (since we travel the lower envelope of $\mathcal{R}' \cup A_\Gamma$ in the range we're defining Γ_0). Any top segment that is covered by $\mathcal{R}'$ within $[x(r_j), x(r_{j+m_0})]$, must be also covered by P_{m_0+1}; as that path is above all other P_m's in the range of $[x_0, x(r_{j+m_0})]$. After traveling Γ_0, we travel along P_1 to r_{j+1}. Now we're going to define Γ_m for odd $1 \leq m \leq m_0$. Each Γ_m goes from r_{j+m} to r_{j+m+2} until we arrive at r_{j+m_0+1}; after which we follow along $\mathcal{R}'$ (i.e. P_{j+m_0+2}, then P_{j+m_0+3} and so on). Path Γ_1 will replace $P_2 + P_3$, Γ_3 will replace $P_4 + P_5$, and so on. Note that Γ_m's are all to the left of $x = x_0$. For any two reflections r_{j+m} and r_{j+m+2} that lie on top segments $s_{a_{j+m}}$ and $s_{a_{j+m+2}}$, let γ_m be the subpaths of $\mathcal{R}'$ restricted to the area of the strip cut by segment $r_{j+m}r_{j+m+2}$ and $s_{a_{j+m}}$ and $s_{a_{j+m+2}}$ (i.e. the area between $s_{a_{j+m}}$ and $s_{a_{j+m+2}}$ and above $r_{j+m}r_{j+m+2}$). Path Γ_m is obtained by starting at r_{j+m} and following line $r_{j+m}r_{j+m+2}$ and whenever we hit $\mathcal{R}'$, i.e. a subpath of γ_m (see Figure 2.16) we follow that subpath until we arrive back to $r_{j+m}r_{j+m+2}$ again; we continue until we reach r_{j+m+2}. In other words, we follow the upper envelope of $r_{j+m}r_{j+m+2} \cup \mathcal{R}'$ between r_{j+m} and r_{j+m+2}. If

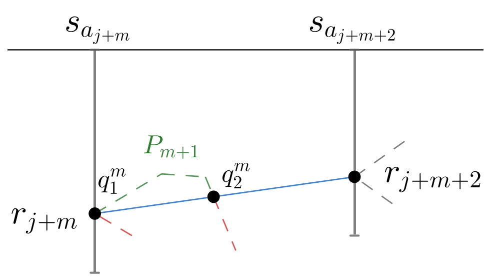


Figure 2.16: Γ_m is traveling the upper envelope of the blue line (the segment $r_{j+m}\, r_{j+m+2}$) and the green parts (portions of OPT_τ that go above $r_{j+m}\, r_{j+m+2}$)

we define $q_1^m, \ldots, q_{2i_m}^m$ to be the intersections of $\mathcal{R}'$ with Γ_m, ordered in the direction of $r_{j+m} \rightarrow r_{j+m+2}$, then if $\mathcal{R}'$ intersects with $r_{j+m}r_{j+m+2}$ and goes above it, it has to be at a point q_u^m where u is odd, and otherwise u has to be even.

Overall, we have changed the subpaths of $\mathcal{R}'$ from r_j to r_{j+m_0} as follows (see Figure 2.14):

- Follow Γ from r_j towards $s_{a_{j+m_0}}$, such that every time an intersection point with $\mathcal{R}'$ (say point q_{2u-1}^0) is reached, then follow along $\mathcal{R}'$ until the next intersection of $\mathcal{R}'$ with Γ (say point q_{2u}^0) is reached; then continue along Γ. Repeat this process until we reach $s_{a_{j+m_0}}$, then follow along Γ from right to left directly back to r_j; this is subpath Γ_0

- From r_j, follow P_1 to reach r_{j+1}.

- From r_{j+m} (initially $m = 1$), follow Γ_m similar to the first step; meaning follow the segment $r_{j+m}\, r_{j+m+2}$, and when an intersection point q_{2u-1}^m with $\mathcal{R}'$ is reached, follow $\mathcal{R}'$ instead, until you reach the next intersection point q_{2u}^m on $r_{j+m}r_{j+m+2}$. Repeat this process (for $m = 1, 3, \ldots$) until r_{j+m_0+1} is reached.

- From r_{j+m_0+1}, follow P_{j+m_0+2} and the rest of $\mathcal{R}'$ to the end.

First, we show that we still have a feasible solution, i.e. every segment that $\mathcal{R}'$ used to cover, will have an intersection with the new solution. To see why this is the case, first note that Γ_0 by definition is always on or below $\mathcal{R}'$ in the ranges it's defined; this means that (using Observation 4) Γ_0 covers any bottom segment that $\mathcal{R}'$ covers between s_{a_j} to $s_{a_{j+m_0}}$. Also all the top segments in this range will be intersecting the path P_{m_0+1}, since P_{m_0+1} is above all $P_{\leq m_0}$ in this range. Similarly, each Γ_m is on or above $\mathcal{R}'$ in the range $[x(r_{j+m}), x(r_{j+m+2})]$, meaning they cover all the top segments that $\mathcal{R}'$ used to cover between $s_{a_{j+1}}$ to $s_{a_{j+m_0+1}}$. Also, any bottom segment that was covered in this range is covered by P_1.

Next note that from r_j to r_{j+m_0+1}, the shadow is at most 3. That is because shadow of Γ_0 is 2, shadow of each Γ_m ($1 \leq m$) is 1, and shadow of P_1 is 1.

Now we show the new solution has an additional cost of at most $3\varepsilon\Psi$. All parts of Γ_0 and the rest of Γ_m's that used portions of $\mathcal{R}'$ can be charged onto $\mathcal{R}'$ itself. So we only have to properly charge the line segment Γ along with its duplicate (part of Γ_0) and line segments $r_{j+m}r_{j+m+2}$ (part of Γ_m). We know that $||\Gamma|| = x(r_{j+m_0}) - x(r_j) < x(r_{j+m_0}) - x_0 = d_{m_0} \leq \varepsilon \cdot \Psi$. So we pay at most $2\varepsilon\Psi$ extra (compared to OPT_τ) for traveling Γ_0. We consider one additional copy of Γ for the extra cost we pay elsewhere in Γ_m ($m \geq 1$), and we are going to use this for our charging scheme. So at the end, the total extra cost is going to be bounded by $3\varepsilon\Psi$.

For each two reflections r_{j+m} and r_{j+m+2} that lie on top segments, note that $\mathcal{R}'$ had two subpaths paths P_{j+m+1}, P_{j+m+2} whose concatenation makes a path from r_{j+m} to r_{j+m+2}; but now, it is possible that some portions of P_{j+m+1}, P_{j+m+2} are used in Γ_0 during our alternate solution (those that belonged to A_Γ). But having that additional copy of Γ that we accounted for, we can use it to short-cut the missing parts of $P_{j+m+1} \cup P_{j+m+2}$ to again make a path from r_{j+m} to r_{j+m+2}. Overall, the total length of $P_1 + \sum_{m \geq 0} \Gamma_m$ that is replacing $P_1 + P_2 + \ldots + P_{j+m_0+1}$ is at most $3\varepsilon\Psi$ larger than length of $P_1 + P_2 + \ldots + P_{j+m_0+1}$. Thus, we conclude the lemma for case of zig-zags.

When $\mathcal{R}'$ is a part of a sink

The proof is analogous to the case of zig-zags. Without loss of generality, assume all reflections in $\mathcal{R}'$ are on bottom segments. Define $d_m = |x(r_{j+m}) - x_0|$ like before. If $r_j, r_{j+2}, \ldots, r_{j+2q}$ are all the reflections to the right of $x = x_0$, then with the same arguments as the case of zig-zags, we will find an integer $m_0 \geq 2(q - (\sigma' - 1))$ ($\sigma' = \lceil \frac{1}{2\varepsilon} \rceil$) for which we have $d_{m_0} \leq \varepsilon \cdot \Psi$.

We will replace the subpath of $\mathcal{R}'$ from r_j to r_{j+m_0} in the same fashion as before. Let Γ be the segment on the line $y = y(r_j)$ in the range $[x(r_j), x(r_{j+m_0})]$. Define Γ_0 to be the union of Γ with the portions of $\mathcal{R}'$ that go below it. Define each Γ_m for a reflection r_{j+m} to the left of $x = x_0$ to be the union of segment $r_{j+m}\, r_{j+m+2}$ with the portions of $\mathcal{R}'$ that go below it.

The same arguments as before hold, that each of Γ or $r_{j+m}\, r_{j+m+2}$ that we defined above, will have an even number of intersections with $\mathcal{R}'$. Define the new path between these reflections in the same way as we did for zig-zags.

The cost arguments still hold, implying that the new path has an additional cost of $O(\varepsilon \cdot \Psi)$. Also, the new path is always on or below $\mathcal{R}'$, so it covers all the bottom segments that $\mathcal{R}'$ used to cover previously. But one can see that P_{m_0} is above all of $\mathcal{R}'$ in the path between r_j to r_{j+m_0}. Thus, P_{m_0} alone will cover all top segments that $\mathcal{R}'$ used to cover. Once again, the new solution has a shadow of at most 3 in the subpath between r_j and r_{j+m_0} and shadow $O(1/\varepsilon)$ afterwards. This concludes the proof for the case of sinks and the proof of Lemma 18. ■

The following corollary immediately follows form Lemmas 14 and 18:

Corollary 3 *There is a $(1 + \varepsilon)$-approximate solution in which any loop or ladder has shadow $O(1/\varepsilon)$.*

The following definition is used in Lemma 19 that is later on applied in our main algorithm:

Definition 10 *Let $\mathcal{R} = p_i, p_{i+1}, \ldots, p_q$ be any sequence of consecutive points in OPT such that p_i and p_q are reflection points. If none of p_j's in $\mathcal{R}$ is a tip of a segment, then $\mathcal{R}$ is called a **pure reflection sequence**.*

So each point in $\mathcal{R}$ is either a straight point or a pure reflection according to Lemma 2.

2.4.2 Bounding the Size of Pure Reflection Sequences

In this section, we will prove the following lemma:

Lemma 19 *Consider OPT_τ for an arbitrary strip S_τ and suppose the total length of legs of OPT_τ is opt_τ. Given $\varepsilon > 0$, we can change OPT_τ to a solution of cost at most $(1+\varepsilon)opt_\tau$ in which the size of any pure reflection sequence is bounded by $O(\frac{1}{\epsilon})$.*

We prove this by showing how to change each ladder or loop (i.e. any path of OPT_τ that starts and ends on one of the cover-lines) so that the size of each pure reflection sub-sequence is bounded without increasing the cost by more than $(1+\varepsilon)$ factor. Consider any loop or ladder $P \in \text{OPT}_\tau$ and any maximal pure reflection sequence $P' = r_0, r_2, \ldots, r_k$ in P where $k > \frac{1}{\epsilon}$. Let Ψ be the length of subpath of OPT_τ from r_0 to r_k. We show how we can modify this subpath to another one whose length is at most $(1 + O(\varepsilon))\Psi$ such that the length of each pure reflection sub-sequence is bounded by $O(1/\varepsilon)$. Note that using Lemma 16, all r_i's are ascending or all are descending. This also implies that the y-coordinates of r_i's are monotone. i.e. either $y(r_0) \leq y(r_1) \leq \cdots \leq y(r_k)$ or the other way around. Without loss of generality, assume it is the former case and so all are ascending reflection points. Proof of Lemma 14, shows if we have a maximal monotone (i.e. all ascending or all descending) sequence of reflection points, then it consists of at most a sink followed by a zig-zag, followed by a sink. Therefore, it suffices to bound the size of pure reflection sequence in a single sink or a zig-zag alone as a function of $1/\varepsilon$. So let's assume all r_i's form a single sink or all form a single zig-zag.

Recall from the definition of pure reflection sequence that there might be straight points in P between two consecutive reflection points. For any reflection point r_i on a segment s, let d_i^+ and d_i^- be the distances of r_i to the top and bottom tips of s, respectively. By the definition of a pure reflection sequence, $d_i^- > 0$ and $d_i^+ > 0$ for all $0 \leq i \leq k$ (because the reflections are not at the tips). With $\sigma = \lceil \frac{1}{\epsilon} \rceil$, we break P' into m subpaths $G_1, \ldots, G_m$ where G_j is the subpath of P' from $r_{j-1\sigma}$ to $r_{j\sigma}$, except that the last group ends at r_k. Note that the concatenation of these paths is P', and each subpath has at most $\sigma + 1$ reflections. Consider any group G_j and let $\mathcal{G}_j$ be the cost of the legs of P between the reflection points of G_j and let D_j be the smallest value among minimum of d_a^+, d_a^- among all reflection points $r_a \in G_j$, i.e. $D_j = \min_{(j-1)\sigma \leq a \leq j\sigma}\{d_a^+, d_a^-\}$.

Claim 3 *For each* $1 \leq j \leq m$*:* $D_j \leq \frac{2\varepsilon}{1-\varepsilon} \cdot \mathcal{G}_j$.

Proof of Claim. For simplicity of notation of indices, we prove this for $j = 1$, i.e. $G_1 = r_0, \ldots, r_\sigma$. As mentioned above, it suffices to show the claim for G_1 being part of a a sink or a zig-zag.

• G_1 **is part of a sink**:

Without loss of generality, assume all reflections in G_1 are on bottom segments. Consider the three consecutive reflection points r_0, r_1, and r_2. Using Lemma 7, we get that subpath $r_1 \to r_2$ (which according to the definition of pure reflection sequence, is a straight line) does not intersect with the segment containing r_0. Since we assumed the y-coordinates of r_i's are increasing, this implies that $r_1 r_2$ and consequently r_2 lie above the segment containing r_0. This along with triangle inequality yields us $d_0^+ \leq y(r_2) - y(r_0) \leq ||r_2 r_0|| \leq ||r_0 r_1|| + ||r_1 r_2||$. With the same argument, we get $d_i^+ \leq y(r_{i+2}) - y(r_i) \leq ||r_i r_{i+1}|| + ||r_{i+1} r_{i+2}||$ for all $0 \leq i \leq \sigma - 2$ (see Figure 2.17). Considering these inequalities for different r_i's ($1 \leq i \leq \sigma - 2$) and summing them up for all r_i's in group G_1, using the fact that $D_1 \leq d_i^+$, we obtain

$$(\sigma - 1) \cdot D_1 \leq \sum_{i=1}^{\sigma-2} d_i^+ \leq \sum_{i=1}^{\sigma-2} (||r_i r_{i+1}|| + ||r_{i+1} r_{i+2}||) \leq 2 \cdot \mathcal{G}_1$$
$$\implies D_1 \leq \frac{2}{\sigma - 1} \cdot \mathcal{G}_1.$$

This implies in a sink, $D_1 \leq \frac{2}{1/\varepsilon - 1} \cdot \mathcal{G}_1 = \frac{2\varepsilon}{1-\varepsilon} \cdot \mathcal{G}_1$ and in general, $D_j \leq \frac{2\varepsilon}{1-\varepsilon} \cdot \mathcal{G}_j$ for all $1 \leq j \leq m$.

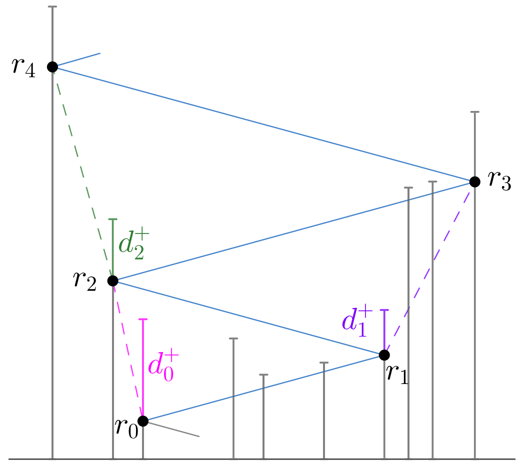


Figure 2.17: Each d_i^+ ($i \leq \sigma - 2$) can be charged into the lines reaching the next two reflections

• G_1 **is part of a zig-zag**:

The inequalities are almost analogous, but there are two of them. Without loss of generality, assume that $r_1, r_3, \ldots$ are on bottom segments, and therefore $r_0, r_2, r_4, \ldots$ are on top segments. We give inequalities for d_i^+ on bottom segments, and for d_i^- on top segments. For $i = 1, 3, \ldots$ with $i \leq \sigma - 2$, similar to the case of G_1 being a sink, we have $d_i^+ \leq y(r_{i+2}) - y(r_i) \leq ||r_i r_{i+1}|| + ||r_{i+1} r_{i+2}||$. For $i = 2, 4, 6, \ldots,$ we have $d_i^- \leq y(r_i) - y(r_{i-2}) \leq ||r_i r_{i-1}|| + ||r_{i-1} r_{i-2}||$ (see Figure 2.18).

Now if we add these inequalities (with proper selection between d_i^+ and $d_{i'}^-$) we get

$$
\begin{aligned}
(\sigma - 1) \cdot D_1 &\leq (d_1^+ + d_3^+ + \cdots) + (d_2^- + d_4^- + \cdots) \\
&\leq \sum_{i \text{ is odd}, \, i \geq 1} (||r_i r_{i+1}|| + ||r_{i+1} r_{i+2}||) + \sum_{i \text{ is even}, \, i \geq 2} (||r_i r_{i-1}|| + ||r_{i-1} r_{i-2}||) \\
&\leq 2 \cdot G_1 \\
\implies D_1 &\leq \frac{2}{\sigma - 1} \cdot \mathcal{G}_1
\end{aligned}
$$

And like before, this implies that in a zig-zag, $D_1 \leq \frac{2\varepsilon}{1-\varepsilon} \cdot \mathcal{G}_1$, and in general, $D_j = \frac{2\varepsilon}{1-\varepsilon} \cdot \mathcal{G}_j$.

So we see that the claim holds for loops and ladders. ■

Also note that if we have any three consecutive points p, r_j, q on OPT_τ where r_j is a reflection on segment s (with s^t being its top tip), then $||pr_j|| +$

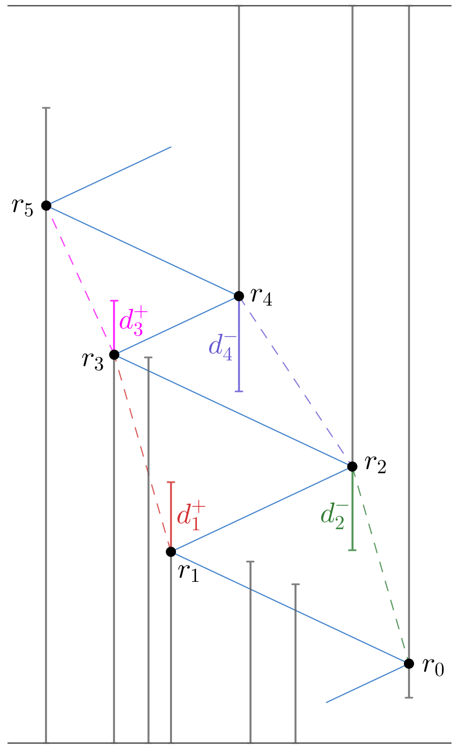


Figure 2.18: d_i^+ ($i \leq \sigma - 2$) on bottom segments and $d_{i'}^-$ ($i' \geq 2$) on top segments can be charged into the lines reaching the next/previous two reflections

$||r_j q|| + 2d_j^+ \geq ||ps^t|| + ||s^t q||$ using triangle inequality. Now consider any group G_j ($1 \leq j \leq m$) and assume r_{j^*} is a reflection point in G_j that lies on segment s for which $d_{j^*}^+ = D_j$ (if $d_{j^*}^- = D_j$, then consider the bottom tip, s^b instead). Consider the two legs of OPT_τ incident to r_{j^*}, namely ℓ_{j^*-1} and ℓ_{j^*}. Let $\ell_{j^*-1} = p_a r_{j^*}$ and $\ell_{j^*} = r_{j^*} p_b$ where p_a and p_b are points on OPT_τ. Suppose we move r_{j^*} from its current location to s^t, i.e. replace the two legs with $p_a s^t$ and $s^t p_b$. Note that this will remain a feasible solution as ℓ_{j^*-1}, ℓ_{j^*} have no other intersections with any other segment (as the definition of legs). The new cost is upper bounded by $||p_a r_{j^*}|| + ||r_{j^*} p_b|| + 2d_j^+$, which means the increase is bounded by $2d_j^+ = 2D_i \leq \frac{4\varepsilon}{1-\varepsilon} \cdot \mathcal{G}_i$.

Note that in this new solution in each group G_i, one of the points is moved to be a tip of the segment it lies on. This implies the maximum size of a pure reflection sequence is now bounded by $2\sigma = 2\lfloor 1/\varepsilon \rfloor$ and the total increase in the cost (over all groups) is bounded by $\sum_j \frac{4\varepsilon}{1-\varepsilon} \cdot \mathcal{G}_j = O(\varepsilon \cdot \sum_j \mathcal{G}_j) = O(\varepsilon \Psi)$. This completes the proof of Lemma 19. ■

Our next goal, in Lemma 20, is to show that for any vertical line it can intersect at most $O(1)$ many loops or ladders of OPT_τ in a strip S_τ. This together with corollary 3 implies that there is a near optimum solution that the shadow in each strip S_τ is bounded by $O(1/\varepsilon)$. The following definition formalizes what we mean by overlapping paths:

Definition 11 *A collection of loops and or ladders are said to be* overlapping *with each other if there is a vertical line that intersects all of them.*

2.4.3 Bounding the Number of Overlapping Loops or Ladders

This section is dedicated to the proof of the following lemma:

Lemma 20 *Consider OPT_τ, the restriction of OPT to any strip S_τ. We can modify the solution (without increasing the shadow or the cost) such that there are at most $O(1)$ loops or ladders in OPT_τ that all are overlapping with each other.*

We will show that there are at most 12 overlapping loops, and at most 7 overlapping ladders in OPT_τ. Suppose there is a vertical line Γ and a number of loops and ladders are all crossing Γ. We bound the number of loops separately from the number of ladders.

Overlapping Loops

For each of the cover-lines of S_τ, we will show that there are at most 6 overlapping loops that have both their entry points on that cover-line. This will imply that there are at most 12 overlapping loops in total. So from this point onward, let's focus on all overlapping loops on the bottom cover-line. This holds for all the claims and proofs that we introduce in this subsection, unless stated otherwise.

Recall Observation 1 that OPT is not self-crossing, so it cannot have two overlapping cover-line loops. We say a loop L_1 with entry points e_1, o_1 is *nested* over loop L_2 with entry points e_2, o_2 if both e_2, o_2 are between e_1, o_1.

Lemma 21 *Let L_1 and L_2 be any two loops such that L_1 is nested over L_2. Let p_r^2 and p_l^2 be the right-most and left-most points on L_2, respectively. Then in the range $I = [x(p_l^2), x(p_r^2)]$, L_1 is above L_2. For simplicity, in this case we say L_1 is above L_2.*

Proof. For $L_j, j = 1, 2$, let e_j and o_j be its entry points and without loss of generality, assume that $x(e_1) \leq x(e_2) \leq x(o_2) \leq x(o_1)$. So L_1 is a path from e_1 to o_1; meaning it crosses the vertical lines $x = x(e_2)$ and $x = x(o_2)$ at some point. This implies if $\mathcal{L}_1$ is the area of strip S_τ bounded by L_1 and the bottom cover-line, then L_2 is entirely inside $\mathcal{L}_1$. This means if the left-most and right-most points on L_1 are p_l^1 and p_r^1, then $x(p_l^1) \leq x(p_l^2)$ and $x(p_r^1) \geq x(p_r^2)$. So we conclude that L_1 is defined in the range $I' = [x(p_l^1), x(p_r^1)]$ and that $I \subseteq I'$. Therefore in particular, L_1 is defined in the range I and is above L_2. ∎

Lemma 22 *Suppose L_1 with entry points e_1, o_1 and L_2 with entry points e_2, o_2 are overlapping such that $x(e_1) < x(e_2) < x(o_1)$. Then L_1 must be nested over L_2 and L_2 is a cover-line loop.*

Proof. If L_1, L_2 are not nested (i.e. $x(e_1) < x(e_2) < x(o_1) < x(o_2)$) and none is a cover-line loop, then they are intersecting inside S_τ, a contradiction. If they are not nested and one (say L_2) is a cover-line loop, then again they are intersecting at one of the entry points. So they must be nested, say $x(e_1) < x(e_2) < x(o_2) < x(o_1)$. Thus, using Lemma 21, L_1 is above L_2; and if L_2 intersects with any top segment, L_1 would already be intersecting with it because of Observation 4. So L_2 should only cover bottom segments, which means it must be a cover-line loop by Lemma 8. ∎

Using these lemmas it follows that there are at most 2 overlapping loops with entry points on opposite sides of Γ. Furthermore, if there are two such loops, then one of them is a cover-line loop.

We will finally show that there are at most 2 overlapping loops that have both their entry points on the same side, say left of Γ. This will imply the result of the lemma for loops, because on each of the cover-lines, there are at most 2 loops with entry points on the left of Γ, 2 with entry points on the right, and 2 with entry points on the opposite sides. Between the loops with both entry points to the left of Γ, none can be a cover-line loop because such a loop cannot intersect Γ (Γ needs to be between the two entry points of a cover-line loop). We will show that there will be at most 2 (non-cover-line) overlapping loops with entry points to the left of Γ.

For the sake of contradiction, assume that there are at least 3 loops with entry points on the left of Γ that none are cover-line loops and all cross Γ. Let L_1, L_2, and L_3 be any 3 consecutive loops with this property. Without loss of generality let $x(e_1) \leq x(o_1)$, $x(e_2) \leq x(o_2)$, and $x(e_3) \leq x(o_3)$, and assume an order for the entry points of L_m's, say $x(e_1) \leq x(e_2) \leq x(e_3)$. We must have $x(e_2) \geq x(o_1)$; or else L_1, L_2 must be nested by Lemma 22, implying L_2 should be a cover-line loop which contradicts the assumption. So we get that $x(e_2) \geq x(o_1)$. Similarly, we have $x(e_3) \geq x(o_2)$. These imply that $e_1, o_1, e_2, o_2, e_3, o_3$ appear in this order on the bottom cover-line. Corollary 2 implies each of L_1, L_2, and L_3 must exclusively cover some top segment. Let r_1, r_2 be the right-most point on L_1, L_2, respectively. Since each L_1, L_2 starts and ends on the left of Γ and travels to the right of Γ, by Lemma 1, the right-most point on each is a reflection point, which implies it must be exclusively covered by using Lemma 7. Let s_{i_1} be the segment that reflection point r_1 lies on, and similarly s_{i_2} the segment for r_2 (see Figure 2.19).

Lemma 23 *s_{i_1}, s_{i_2} are top segments and $x(s_{i_1}) < x(s_{i_2})$*

Proof. By way of contradiction, assume s_{i_1} is a bottom segment. Consider the two subpaths of L_1 between the entry points e_1, o_1 and r_1, let us denote them by $P_r^1 : e_1 \rightarrow r_1$ and $P_l^1 : r_1 \rightarrow o_2$. L_2 (starting at e_2) is in the region bounded by $P_r^1 \cup s_r^1$ and the bottom cover-line, which means L_1 will intersect any top segment L_2 intersects with (i.e. L_2 cannot exclusively cover any top segment), which implies L_2 is a cover-line loop, a contradiction. This implies that s_{i_1} is a top segment. Similar argument (for L_2, L_3) implies s_{i_2} is a top segment.

We show that $x(s_{i_1}) \leq x(s_{i_2})$. Similar to before, define the subpath P_r^1 of L_1 that goes from e_1 to r_1 and P_r^2 from e_2 to r_2. Considering the two areas of strip S_τ separated by $P_r^1 \cup s_{i_1}$, if segment s_{i_2} is on one side and the entry points o_1, e_2 on the other side, then path P_r^2 must either intersect P_r^1 or s_{i_1}, which is not possible (due to Lemma 7). So s_{i_2} and e_2, o_2 are on the same part of S_τ cut by $P_r^1 \cup s_{i_1}$. This implies s_{i_2} is to the right of s_{i_1} i.e $x(s_{i_1}) \leq x(s_{i_2})$. ■

We can reuse the same arguments in the second part of the proof to conclude the following lemma:

Lemma 24 *Neither of L_1 or L_2 exclusively cover a bottom segment on the right of Γ.*

Proof. Let L_j be either one of L_1 or L_2. Assume the contrary, that there is some bottom segment s_j to the right of Γ that L_j exclusively covers. So L_3 does not intersect with this segment. Let p_j be the last intersection point of L_j with s_j. Consider the subpath $P_j : p_j \to o_j$ on L_j. Similar to the proof of Lemma 23, we get that both entry points of L_3 are surrounded by $P_j \cup s_j$ from the right or above; which means any top segment that L_3 intersects with, is already intersecting with P_j. This requires L_3 to be a cover-line loop, giving us a contradiction. ■

Now we define an alternate path that replaces L_1 and L_2 with two new loops that no longer overlap at Γ, and overall the shadow does not increase but also costs less than the cost of current solution. The idea of this change (which will be made precise soon) is to follow L_1 from e_1 to the right-most point on L_1 (which must be a reflection on s_{i_1}), then from that point follow a horizontal line until it hits s_{i_2}; if there are portions of L_2 that are above this horizontal line, we follow the upper envlope of those portions of L_2 and the horizontal line (similar to how we reduced the shadow in the case of zig-zag or sink), and then from the intersection point on s_{i_2}, follow the horizontal line back to the right-most reflection on L_1 and continue to follow L_1 to o_1; L_2 is going to be simply replaced with a smaller subset of its projection on the bottom cover-line. We show we will have a cheaper feasible solution with smaller shadow at Γ, a contradiction. Now we describe this more precisely.

Again, let r_1, r_2 be the right-most points on L_1, L_2, respectively. Lemma

23 and 1 imply that they are reflection points on segments s_{i_1}, s_{i_2}, respectively, which both are top segments. Consider the horizontal line $y = y(r_1)$, and let q be the intersection point of this line with the vertical line $x = x(s_{i_2})$. Define the subpath P_r^2 on L_2 as $P_r^2 : e_2 \to r_2$ (assuming that e_2 is to the left of o_2), In other words, between the two paths from r_2 to the two entry points of L_2, P_r^2 is the one that is above the other. Let U_2 be the portion of P_r^2 in the region bounded by lines $s_{i_1} \cup (y = y(r_1)) \cup s_{i_2}$, and the top cover-line. So these are the portions of P_r^2 that go above the line segment r_1q (see Figure 2.19). Let L_1'' be the upper envelope of $U_2 \cup r_1q$ plus the line r_1q.

So L_1'' consists of a path that goes on the upper envelope of $U_2 \cup r_1q$ from r_1 to q and then goes straight back to r_1. We now define the replacements for L_1 and L_2.

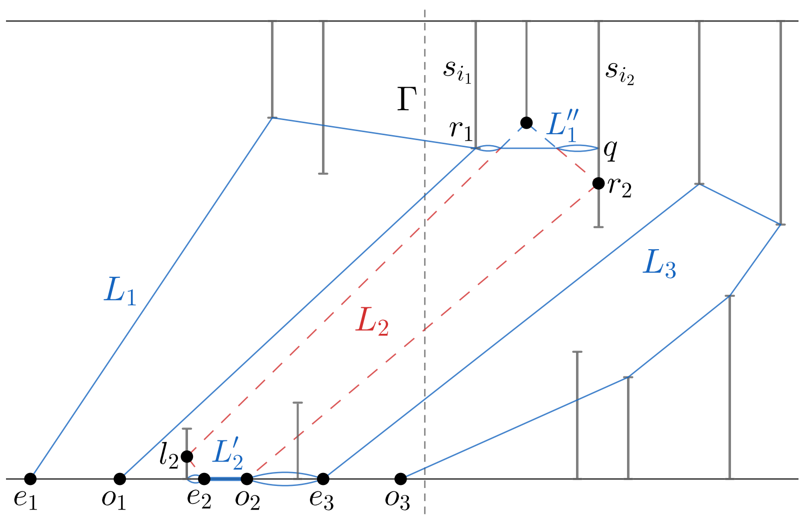


Figure 2.19: Alternative solution for 3 overlapping (non-cover-line) loops. Pairs of arcs represent doubled segments.

We replace L_1 with L_1' as follows:

- Take the subpath $P_r^1 : e_1 \to r_1$ on L_1.
- From r_1, follow L_1'' and thus, get back to r_1.
- From r_1, follow the rest of L_1 to o_1.

So L_1' is obtained by adding L_1'' to L_1 at r_1. If l_2 is the left-most point that L_2 travels, then let L_2' be a cover-line loop that travels from e_2 left to

$x(l_2)$, then right to e_3 and then back to o_2 (this is essentially the projection of the portions of L_2 to the left of e_3 and hence to the left of Γ on the bottom cover-line); recall that we can reduce L'_2 to remove the possible overlapping of its legs. Now replace L_2 with L'_2. We will show that these two loops in total cost strictly less than L_1 and L_2, the shadow does not increase (and in fact shadow decreases at Γ) and we still have a feasible solution. It's clear to see that between the loops L'_1, L'_2, and L_3, only L'_1 and L_3 overlap at Γ. So we decreased the number of overlapping loops at Γ by at least one.

To prove all segments are still covered, note that L'_1 includes the entirety of L_1, and thus covers all the segments that L_1 used to cover. In order to show that all the segments that L_2 covered, are still covered, we only need to show that the segments that L_2 exclusively covered, are still covered. That is because in the new configuration we still have all the parts of L_1 and L_3. According to Lemma 24, there are no bottom segments that L_2 exclusively covers to the right of Γ. Also, it is easy to see that any bottom segment that was exclusively covered by L_2 to the left of Γ must have an x-coordinate between $x(l_2)$ and $x(e_3)$. All of those bottom segments are now covered by L'_2. Finally, for the top segments that L_2 exclusively covers, with the same arguments as in the second part of the proof in Lemma 24, we get that there are no such segments to the left of s_{i_1}. So it suffices to show that only the top segments that L_2 covers to the right of s_{i_1}, are covered. This is easy to see, because L''_1 includes the entire U_2; and it is always on or above L_2 in the range between s_{i_1} and s_{i_2}. Thus, L''_1 will cover all the top segments that L_2 exclusively covers in that range.

Also, the shadow does not increase: the shadow of L''_1 from r_1 to r_2 can be charged to the sections of L_2 between $x = x(r_1)$ and $x = x(r_2)$ and hence is no more than that; note that this portion is entirely to the right of Γ. The cover-line loop L'_2 is entirely to the left of Γ and its shadow can be charged to the shadow of L_2 to the left of Γ in the range $[x(l_2), x(e_3)]$.

Now let's prove that the new cost is decreased compared to L_1 and L_2. L'_1 includes L_1, so we set aside those parts and charge them on L_1. So it suffices to show that L''_1 along with L'_2 can be charged into L_2. Note that L'_2 is part of the projection of L_2 on the bottom cover-line to the left of e_3. So the cost of L'_2 is strictly less than the cost of L_2 to the left of Γ, since L'_2 extends at most to e_3 which is to the left of Γ. As for L''_1, note that L_2 travels back and forth between $x(s_{i_1}), x(s_{i_2})$; so L''_1 can be charged to these two sections of L_2 between $x(s_{i_1})$ and $x(s_{i_2})$.

So at the end, we found a new solution with 1 fewer overlapping loops at Γ, no increase of shadow elsewhere, and with a strictly less cost than OPT_τ. Applying this argument implies that at most two overlapping non-cover-line loops can exist to the left of Γ. So in total on each of the cover-lines of S_τ, there are at most 2 non-cover-line loops to the left of Γ, similarly 2 to the right, plus at most 2 with entry points to opposite sides of Γ. In total, there are at most $6 \times 2 = 12$ overlapping loops at Γ.

Overlapping Ladders

Recall that by Definition 7, ladders are subpaths of OPT_τ (in strip S_τ) that have one entry point on the bottom cover-line of S_τ, and one on the top cover-line. Depending on their orientation compared to Γ, there are two types of ladders (see Figure 2.20):

- **Type 1 Ladder:** Has both its entry points on the same side of Γ.
- **Type 2 Ladder:** Has its entry points on opposite sides of Γ.

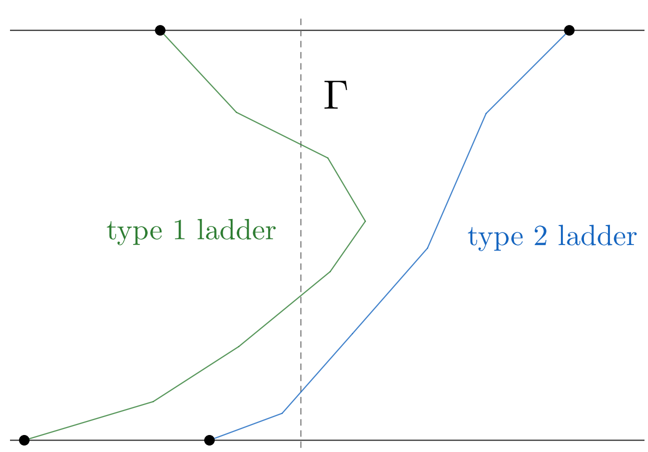


Figure 2.20: An example of type 1 and type 2 ladders.

We will prove that there are at most 2 overlapping Type 1 ladders, and at most 5 overlapping type 2 ladders.

Type 1 Ladders

In particular, we show that there is at most one Type 1 ladder with entry points to the left of Γ, and one with entry points to its right. To prove this, assume the contrary, that there are at least 2 overlapping Type 1 ladders with entry points to the same side, say right of Γ. Let L_1 and L_2 be two such ladders.

Let $(b_m, t_m),\ m = 1, 2$ be the entry points of L_m on the bottom cover-line and the top cover-line, respectively. Without loss of generality, assume that t_1 is to the left of t_2. This implies b_1 is also to the left of b_2 (or else L_1 and L_2 intersect inside S_τ). So if we consider cutting S_τ along L_2, L_1 is entirely in one of the two regions created, namely the one that contains b_1, t_1. Since both L_1 and L_2 overlap at Γ and are Type 1, and they both have their entry points on the same side of Γ, say left, this means that both have to reach to the right of Γ. First we show that the top-most and bottom-most intersection point of L_1, L_2 with Γ must be on L_2. By way of contradiction suppose p is a point on L_1 and is the bottom-most intersection of these two ladders on Γ. Consider the subpath of L_1 from b_1 to p, call it L_1' and consider the region bounded by $L_1' \cup \Gamma$ and the bottom cover-line, call it A. Since L_2 starts at b_2 inside A and t_2 is outside A, L_2 must either cross Γ at a point lower than p, or cross L_1', both of which are contradictions. Similar argument shows the top-most intersection point on Γ is with L_2.

Consider any two consecutive crossing of L_1 with Γ, say p_1, p_2, where the subpath of L_1 from p_1 to p_2 (denoted by L_1') is to the right of Γ. Since L_2 crosses Γ both above and below p_1, p_2 (the lowest and highst intersection points on Γ are with L_2), there is a subpath of L_2 with end-points q_1, q_2 on Γ with q_1 below p_1, p_2, and with q_2 above them, call it L_2'. We consider two cases based on whether L_2' is on the left or right of Γ, and derive contradictions in each case. If L_2' is on the right (like L_1') then L_1' is inside the region bounded by $L_2' \cup q_1q_2$ and this violates Lemma 9. So let us assume L_2' is on the left of Γ.

Since p_1, p_2 are between q_1, q_2 there is subpath of L_1 starting from p_1 inside the region $L'_2 \cup q_1q_2$ that crosses q_1q_2. This subpath with L'_2 violates Lemma 9 again. Thus, we conclude that there can be at most 1 Type 1 ladder with entry points to the right of Γ, and similarly, at most 1 with entry points to the left of Γ.

Type 2 Ladders

For each Type 2 ladder L_m with entry points (b_m, t_m) on bottom and top cover-lines, there are two cases:

- b_m is to the left of Γ, therefore t_m is to the right of Γ. We say L_m is a *top-right/bottom-left* ladder.
- b_m is to the right of Γ, therefore t_m is to the left of Γ. We say L_m is a *top-left/bottom-right* ladder.

There can't be two overlapping ladders that one is a top-right/bottom-left ladder, and the other is a top-left/bottom-right ladder (or else they intersect). So if we have a collection of Type 2 overlapping ladders they are all either top-right/bottom-left or all top-left/bottom-right. We show we can have at most 5 Type 2 overlapping ladders. For the sake of contradiction, assume there is a maximal set $\mathcal{L} = \{L_1, L_2, \dots, L_k\}$ of Type 2 ladders that all overlap at some vertical line Γ with $k \geq 6$ and all are top-right/bottom-left. Let (b_m, t_m), $1 \leq m \leq k$ denote the bottom and top entry points of ladder L_m. Without loss of generality, assume that $x(b_1) \leq x(b_2) \leq \cdots \leq x(b_k)$, which also implies $x(t_1) \leq x(t_2) \leq \cdots \leq x(t_k)$ (or else the ladders will be intersecting each other). Let L^l_m be the subpath of of L_m from b_m to the first intersection of L_m with Γ (so L^l_m is to the left of Γ), and L^r_m be the subpath of L_m from its last intersection with Γ to t_m (so it is to the right of Γ). Note that if $m < m'$ then L^l_m is above $L^l_{m'}$ (in the range that L^l_m is defined) and $L^r_{m'}$ is below L^r_m (in the range that $L^r_{m'}$ is defined) due to Lemma 11. Using Observation 4, this implies L^l_1 covers all the top segments that $L^l_2, L^l_3, \dots, L^l_k$ cover to the left of Γ and similarly, L^r_k covers all the bottom segments that $L^r_1, L^r_2, \dots, L^r_{k-1}$ cover to the right of Γ (we will use this fact shortly).

We will introduce an alternate set of ladders (and loops) that cover all the segments the ladders in $\mathcal{L}$ cover without increasing the shadow anywhere, with a cost strictly smaller cost, and with a smaller shadow at Γ. The set of ladders we introduce differ based on the parity of k. For odd k we keep L_1, L_{k-1}, L_k, and for even k we keep L_1, L_3, L_{k-2}, L_k. We also add some cover-line loops (possibly two copies) to make sure we still have a tour that visits all the points b_j, t_j and all the top and bottom segments that $L_1, \ldots, L_k$ covered remain covered in the new solution [1].

Imagine a graph $G(V, E)$ where V consists of all b_j, t_j's and there are edges between two vertices if there is a subpath in OPT between them without visiting any vertex (so we have direct edge between b_j, t_j and also an edge between b_j, t_i if there is a path in OPT between them outside the strip S_τ). Note that G is simply a cycle. In the new alternative solution, we keep L_1, L_k and either L_{k-1} or both L_3, L_{k-2} (depending on the parity of k) and add cover-line loops between some consecutive b_j's and consecutive t_j's such that the resulting graph G' defined based on these new paths still forms an Eulearian (connected) graph on V, all the segments covered in S_τ by $L_1, \ldots, L_k$ are covered. Let b'_2 be the projection of the left-most point on L_2 on the bottom cover-line, and let t'_{k-1} be the projection of the right-most point on L_{k-1} on the top cover-line. We add doubled segment $b_2 b'_2$ and $t_{k-1} t'_{k-1}$. These intend to cover any bottom segment exclusively covered by L_2 to the left of b_2, and any top segment exclusively covered by L_{k-1} to the right of t_{k-1}. The doubled segments $b_2 b'_2$ and $t_{k-1} t'_{k-1}$ fully appear in the projection of L_2 and L_{k-1} on those cover-lines; meaning that they can be charged onto L_2 and L_{k-1} that travel left (and right) to those segments, respectively. Add each of these two segments twice to the solution. Since we're adding these segments twice, the parity of the degree of nodes in G' won't change. We keep L_1, L_k from $\mathcal{L}$ and add the following segments and ladders as well to the alternative solution (see Figure 2.21):

- If $k = 2m$ for some integer $m \geq 3$, then include L_3 and L_{k-2}. We also

[1] This change is somewhat similar to the proof of patching lemma used in the PTAS for Euclidean TSP that reduces the number of crossings into a region.

add the following cover-line loops:

- $b_2b_3,\ b_{k-2}b_{k-1},\ b_{2q-1}b_{2q}$ $(2 \le q \le m-1)$
- $t_2t_3,\ t_{k-2}t_{k-1},\ t_{2q-1}t_{2q}$ $(2 \le q \le m-1)$

We add double the following cover-line loops (i.e. a path back and forth on the same pair of points):

- $b_{k-1}b_k, b_{2q}b_{2q+1}$ $(2 \le q \le m-2)$
- $t_1t_2,\ t_{2q}t_{2q+1}$ $(2 \le q \le m-2)$

- If $k = 2m+1$ for some integer $m \ge 3$, then we include L_{k-2} and also the following cover-line loops:
  - $b_{k-2}b_{k-1},\ b_{2q}b_{2q+1}$ $(1 \le q \le m-2)$
  - $t_{2q}t_{2q+1}$ $(1 \le q \le m-2)$

  We add double the following segments:
  - $b_{k-1}b_k,\ b_{2q-1}b_{2q}$ $(2 \le q \le m-1)$
  - $t_{2q-1}t_{2q}$ $(1 \le q \le m)$

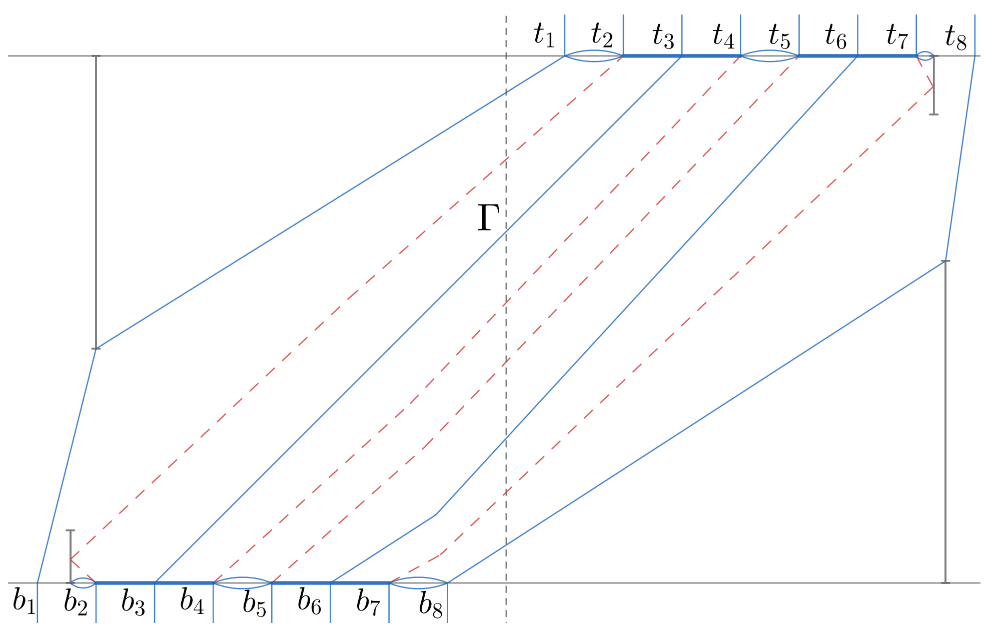


Figure 2.21: Alternative solution for 8 overlapping (bottom-left/top-right) ladders. Red dashed lines are discarded. The arcs represent the doubled segments.

It can be seen that with the above additions, if we build the graph G' based on the new paths it is an Eulerian graph as each b_j, t_j has even degree; also G' remains connected since all the $t_1, \ldots, t_{k-1}$ are connected via cover-line loops added at the top and $b_2, \ldots, b_k$ are connected via cover-line loops at the bottom and we have L_1, L_k and there is a path from b_1 to at least one of $b_2, \ldots, b_k$ in outside the strip, and similarly a path from t_k to one of $t_1, \ldots, t_{k-1}$. Thus in the new solution we visit all the points b_j, t_j and this tour can be short-cut over repeated points to obtain a new solution that visits all the b_j, t_j's and covers all the segments outside the strip S_τ.

Next we show all the segments that $L_1, \ldots, L_k$ were covering, remain covered. Recall that the portion of L_1 to the left of Γ covers all the top segments that were covered by these paths to the left of Γ, and similarly L_k covers all the bottom segments that were covered to the right of Γ. The bottom segments covered to the left of Γ are covered by the new cover-line loops added and similarly, the top segments covered to the right of Γ are covered by the cover-line loops added. So the new solution remains feasible.

Now we are going to bound the total cost of the new solution. We charge all the new parts that we added to some portion of the ladders that we have discarded. Note that in every case, L_2, L_4, L_{k-3}, and L_{k-1} are discarded. We will use only these ladders to charge the new parts to. The doubled segments b_2b_2' and $t_{k-1}t_{k-1}'$ are already charged to the portion of L_2 traveling in the interval $[x(b_2'), x(b_2)]$ and the portion of L_{k-1} traveling in $[x(t_{k-1}), x(t_{k-1}')]$. Now consider the ranges $\beta_j = [x(b_{j-1}), x(b_j)]$, $3 \leq j \leq k$ and $\theta_j = [x(t_{j-1}), x(t_j)]$, $2 \leq j \leq k-1$. These are disjoint and all β_j's lie under L_2^l and L_4^l, while all θ_j's lie above L_{k-3}^r and L_{k-1}^r. Each of the new included segments (doubled or not) can be charged to one or two of the ladders $L_2, L_4, L_{k-3}, L_{k-1}$.

It can be seen that in the new configuration, there are at most 4 overlapping ladders and 2 overlapping loops (doubled segment loops that we added). This concludes the case for ladders.

In general, when given a collection of loops and ladders, we first alter the ladders as described above (and might get some new cover-line loops in the process), then we apply the alteration on the loops. The statement of lemma

20 follows from this. ■

We can now get to the proof of Theorem 3, then Theorem 2.

2.5 Proof of Theorem 3

We reiterate the theorem for convenience, then prove it:

Theorem 3 *If $H \leq 3$, then the shadow of an optimum solution is at most 2.*

Proof. As defined before, let $C_1, \ldots, C_\sigma$ be the cover-lines for an instance of the problem with $H \leq 3$. It can be seen that $\sigma \leq 2$; in other words, all the segments of the instance can be covered with only at most 2 cover-lines. If $H \leq 2$, then the number of cover-lines is 1, and similar to the special case that we discussed at the start of Section 2.3, the portion on that cover-line itself (doubled from the left-most segment to the right-most segment) is an optimum solution. So let's assume $2 < H \leq 3$, therefore $\sigma = 2$, and that we have a single strip, S_1. Furthermore, there must be both top segments and bottom segments in S_1 (otherwise one of the cover-lines would intersect with all segments). We will essentially prove that the optimum solution must be a bitonic tour.

Take any optimum solution OPT for this instance of the problem, and let p^l and p^r be the left-most and right-most points on it, respectively. There is a path P_1 from p^l to p^r, and there is a path P_2 in the other way. Since OPT is not self-intersecting, and since both P_1 and P_2 cover the range $I = [x(p^l), x(p^r)]$, then for any vertical line Γ with $x(\Gamma) \in I$, they both will intersect with it at distinct points. We can use Lemma 9 (for the concatenation of P_1 and P_2 restricted to the left of Γ) to get that p^l is a right reflection. Similarly, p^r is a left reflection.

Without loss of generality, assume that P_1 includes the upper leg of p^l, and thus P_2 includes its lower leg. Using Lemma 5 for the reflection point p^l and the vertical line $x = x(p^r)$, we get that P_1 is above P_2 in range I, which is the entirety of OPT. Observation 4 implies that all the top segments are covered by P_1, while all the bottom segments are covered by P_2. We claim that there are no reflection points other than p^r and p^l in OPT. To see why this is the case, assume the contrary, that there is some reflection point r on OPT other than those two points.

Without loss of generality, assume $r \in P_1$, and assume that r is the first such reflection point on P_1 after p^l. According to Lemma 6, r is a right reflection. Let s be the segment of the instance that r lies on. If s is a bottom segment, then P_2 will be intersecting with it, and we get a violation of Lemma 7. Thus, s is a top segment.

Now, let $\mathcal{P}_1$ be the concatenation of P_1 (restricted to the subpath from p^l to r) along with the entirety of P_2. $\mathcal{P}_1$ is a path that goes from r (a left reflection on a top segment s) and reaches to the right of s. The rest of the path of P_1 (from r to p^r), refer to it as $\mathcal{P}_2$, is another path that goes from r and reaches to its right. Depending on whether the top leg of r belongs to $\mathcal{P}_1$ or $\mathcal{P}_2$, we get a violation of Lemma 23. This contradiction shows that such r cannot exist, and that both P_1 and P_2 are monotone paths with shadow 1, due to Lemma 3. So in total, OPT has a shadow of 2, as was to be shown. ■

Note. It can be shown that in these special cases, we can find an exact solution in poly-time. But since we made some assumptions about the x-coordinates of the segments of the instance, we have to undo those assumptions to prove this claim. The resulting algorithm will be somewhat detailed for such a limited special case of the problem, because we have to cover cases such as vertical legs in an optimum solution. So we only settled on showing that an optimum solution has a constant shadow instead, as it's enough for the purposes of our main algorithm in this thesis.

2.6 Proof of Theorem 2

For convenience, we re-estate the theorem, which is our main structure theorem for a near-optimum solution:

Theorem 2 *Given any $\varepsilon > 0$, there is a solution $\mathcal{O}'$ of cost at most $(1+\varepsilon) \cdot opt$ such that in any strip of height 1, the shadow of $\mathcal{O}'$ is $O(1/\varepsilon)$ (where opt is* *the cost of an optimum solution).*

Proof. If the height of the bounding box is at most 2, refer to Theorem 3. Consider any strip S_τ (to be more precise, S_τ can be any arbitrary strip of height 1 in the plane). Using Lemma 18 for parameter $\varepsilon_1 = \frac{\varepsilon}{2}$, there is a solution $\mathcal{O}''$ of cost at most $(1 + \frac{\varepsilon}{2}) \cdot \text{opt}$ where the shadow of each sink and zig-zag is bounded by $O(\frac{1}{\varepsilon/2}) = O(1/\varepsilon)$. By Lemma 14, each loop or ladder in S_τ has a shadow that is at most 3 times the maximum shadow of a sink or zig-zag in it, plus two. So each loop or ladder has shadow $O(1/\varepsilon)$. Finally, Lemma 20 shows that there can be at most $O(1)$ overlapping loops or ladders in a strip. Thus, the overall shadow of $\mathcal{O}''$ in S_τ is bounded by $O(1/\varepsilon)$. Furthermore, we apply Lemma 19 on O'' for parameter $\varepsilon_2 = \frac{\varepsilon}{\varepsilon+2}$ to get a solution $\mathcal{O}'$. This new solution has the property that with an additional cost of factor $(1+\frac{\varepsilon}{\varepsilon+2})$ compared to $\mathcal{O}''$, the size of any pure reflection sequence is bounded by $O(\frac{\varepsilon+2}{\varepsilon}) = O(1/\varepsilon)$. The total cost of $\mathcal{O}''$ is at most

$$
\begin{aligned}
(1+\varepsilon_1)\cdot(1+\varepsilon_2)\cdot \text{opt} &= (1+\tfrac{\varepsilon}{2})\cdot(1+\tfrac{\varepsilon}{\varepsilon+2})\cdot \text{opt} \\
&= (1+\tfrac{\varepsilon}{2}+\tfrac{\varepsilon}{\varepsilon+2}+\tfrac{\varepsilon^2}{2(\varepsilon+2)})\cdot \text{opt} \\
&= (1+\varepsilon(\tfrac{1}{2}+\tfrac{1}{\varepsilon+2}+\tfrac{\varepsilon}{2(\varepsilon+2)}))\cdot \text{opt} \\
&= (1+\varepsilon)\cdot \text{opt},
\end{aligned}
$$

resulting in the statement of the theorem. ■

Chapter 3

Dynamic Program and the Main Algorithm

As mentioned in the introduction, we follow the paradigm of Arora [4] for designing a PTAS for classic Euclidean TSP with some modifications. We focus more on defining the modifications that we need to make to that algorithm.

In this chapter, we describe the main algorithm and how it reduces the problem into a collection of instances with a constant-height bounding box. We show how those instances can be solved using DP (referred to as the inner DP), and how we can combine the solutions for them using another DP (referred to as the outer DP) to find a near optimum solution of the original instance. Recall that in Section 2.1, we assumed the minimal bounding box of the instance has length L and height H and we defined $B = \max\{L, H-2\}$, which gives opt $\geq 2B$ and also we can assume that $B \leq \frac{n}{\varepsilon}$. Also, recall that we moved each line segment to be aligned with a grid point with side length $\frac{\epsilon B}{n^2}$, while making sure all line segments have distinct x-coordinates. By doing this, we obtain an instance whose optimum is within a $(1+\varepsilon)$-factor of the optimum of the original instance. Now, we scale the grid (as well as the line segments of the instance) by a factor of $\rho = \frac{4n^2}{\epsilon B}$ so that each grid cell has size 4. We obtain an instance where each line segment has size ρ, all have even integer coordinates, any two segments are at least 4 units apart, and the bounding box has size $N = O(n^2/\varepsilon)$. Let this new instance be $\mathcal{I}$. Note that if we define cover-lines as before but with a spacing of ρ, all the arguments for the existence of a near-optimum solution with a bounded shadow in any strip (the area between two consecutive cover-lines) still hold. We will present a PTAS for this instance. It can be seen that this implies a PTAS for the original instance of the problem. From now on, we use OPT to refer to an optimum solution of instance $\mathcal{I}$, and opt to refer to its value. Note that since the bounding box has side length N, then opt $\geq 2N$.

3.1 Dissecting the Original Instance into Smaller Subproblems

Similar to Arora's approach, we do the hierarchical dissectioning of the instance into nested squares using random axis-parallel dissectioning lines, and put portals at these dissecting lines. We continue this dissectioning process until the distances between horizontal (and so vertical) dissecting lines is $h \cdot \rho$ for $h = \lceil 1/\varepsilon \rceil$. So at the leaf nodes of our recursive decomposition quad-tree, each square is $(h \cdot \rho) \times (h \cdot \rho)$, and the height of the decomposition is $\log(N/\rho h) = O(\log n)$ since $B \leq \frac{n}{\varepsilon}$. We choose vertical dissecting lines only at odd x-coordinates so no line segment of the instance will be on a vertical dissecting line.

We define our cover-lines C_τ based on these horizontal dissecting lines carefully. Consider the first (horizontal) dissecting line we choose, this will be a cover-line, and then moving in both up and down directions from this line, we draw horizontal lines that are ρ apart. These will be all the cover-lines. Label the cover-lines from the top to bottom by $C_1, C_2, \ldots, C_\sigma$ in that order. As before, and the smallest index τ such that C_τ hits a line segment is the cover-line that "covers" that line segment. We partition the cover-lines into h groups based on their indices: Group G_j contains all those cover-lines with index τ where $j = \tau \pmod h$. Let G_{j*} be the group of cover-lines that includes the first horizontal dissecting line, and hence all the other horizontal dissecting lines as well.

The arguments in Chapter 2 for the case of unit-length line segments that show there is a near optimum solution in which the shadow in each strip of height 1 is $O(1/\varepsilon)$ (Theorem 2), also imply the same for the scaled instance $\mathcal{I}$;

i.e. there is a near optimum solution with shadow $O(1/\varepsilon)$ in each strip between two consecutive cover-lines. Furthermore, if we consider h consecutive strips, i.e. the area between two consecutive cover-lines in the same group G_j, then there is a near optimum solution that has shadow $O(h/\varepsilon) = O(1/\varepsilon^2)$.

Our goal is to show that, at a small loss in approximation, we can simply drop the line segments that are intersecting the horizontal dissecting lines (i.e. all those intersecting cover-lines in G_{j^*}) with appropriate consideration of portals (to be described). Removing the line segments that cross the dissecting lines allows us to decompose the instance into "independent" instances that interact only via portals.

For each cover-line C_τ, we define a set B_τ of disjoint *intervals* of length ρ placed on it so that each line segment covered by C_τ, is intersecting one of these interval. On C_τ, from left to right, start by placing the left corner of the first interval of B_τ on it at the intersection of the left-most segment covered by C_τ; all the segments covered by C_τ intersecting this interval are considered "covered" by this interval. Next, pick the first segment to the right of the latest interval that is intersecting C_τ, but not intersecting (and so not covered by) the previous intervals, and place the left point of the next interval of B_τ at that intersection (all the segments intersecting C_τ and this interval are now covered by this interval). Continue this process until all segments on C_τ are covered by an interval (see Figure 3.1). Let $\mathcal{B} = \cup_{\tau=1}^{\sigma} B_\tau$.

Observation 5 *A segment covered by an interval of cover-line C_τ and another segment covered by an interval of cover-line $C_{\tau+2}$ are at least ρ apart ($\tau \leq \sigma - 2$).*

Lemma 25 *$opt \geq \dfrac{\rho \cdot |\mathcal{B}|}{6}$.*

Proof. For each B_τ, let $i_1, i_2, \ldots, i_\eta$ be the intervals on C_τ ordered from left to right. Now partition B_τ into $O_\tau \cup E_\tau$ where O_τ consists of intervals i_q with an odd q, and E_τ consists of those with even q's. We also partition C_τ's into 3 groups based on the value of $\tau \pmod 3$. We get a partition of all intervals into 6 groups based on: Whether an interval on C_τ is in O_τ or E_τ (two choices), and what $\tau \pmod 3$ is (three choices). Let N_j's ($1 \leq j \leq 6$) be the total number of intervals in these 6 parts. Note that $\sum_{j=1}^{6} N_j = |\mathcal{B}|$, and any two segments s, s' covered by intervals from different groups are at least ρ apart (if they there are covered by intervals in the same cover-line then they are ρ apart horizontally and if they covered by intervals in different cover-lines then by Observation 5 they are at least ρ apart). So an optimum solution for the instance that only contains segments covered by intervals of part N_j, must have cost at least ρN_j as it must have at least N_j legs of size at least ρ. Since one of these parts has size at least $|\mathcal{B}|/6$, the statement follows. ■

Lemma 26 *For a j chosen randomly from $[1..h]$, we have*

$$\mathbb{E}[\ \rho \sum_{C_\tau \in G_j} |B_\tau|\ ] = O(\varepsilon \cdot opt).$$

Proof. For each $1 \leq j \leq h$, let $\mathcal{B}_j = \bigcup_{C_\tau \in G_j} B_\tau$. Using Lemma 25, we have $\sum_{j=1}^{h} |\mathcal{B}_j| = |\mathcal{B}| \leq 6 \cdot opt/\rho$. Now we obtain

$$\begin{aligned}
\mathbb{E}[\rho \sum_{C_\tau \in G_j} |B_\tau|] &= \rho \cdot \mathbb{E}[\sum_{C_\tau \in G_j} |B_\tau|] \\
&= \rho \cdot \mathbb{E}[|\mathcal{B}_j|] \\
&= \frac{\rho}{h} \cdot |\mathcal{B}| \\
&\leq \frac{\rho}{h} \cdot \frac{6 \cdot opt}{\rho} = O(\mathrm{opt}/h) = O(\varepsilon \cdot \mathrm{opt}).
\end{aligned}$$

■

Similar to Arora's scheme for TSP, for $m = O(\frac{1}{\varepsilon} \log(N/\rho h))$, we place portals at all 4 corners of a square in the decomposition, plus an additional $m-1$ equally distanced portals along each side (so a total of $4m$ portals on the perimeter of a square of the dissection). For simplicity, we assume m is a power of 2 and at least $\frac{4}{\varepsilon} \log(N/\rho h)$. We say a tour is *portal respecting* if it crosses between two squares in our decomposition only via portals of the squares. A tour is *r-light* if it crosses the portals on each side of a square of the dissection at most r times. For classic (point) TSP, it can be shown that there is a near-optimum solution that is portal respecting and r-light for $r = O(1/\varepsilon)$. Our goal is to show a similar statement, except that we want the restriction of the tour to each "base" square of side length $O(h \cdot \rho)$ to have bounded (by $O(h/\varepsilon) = O(1/\varepsilon^2)$) shadow as well. We then show that we can find an optimum solution with a bounded shadow for the base cases using a DP. This will be our *inner DP*. We then show how the solutions of for the 4 sub-squares of a square in our decomposition can be combined into a solution for the bigger subproblem (like in the case of TSP) using another DP, which will be our *outer DP*.

We will show that at a small loss in approximation (i.e. $O(\varepsilon \cdot \text{opt})$), we can drop all the line segments of input that are intersecting the horizontal dissecting lines (i.e. covered by a cover-line in group G_{j^*}), solve appropriate subproblems, and then extend the solutions to cover those dropped segments. This modification requires certain portals of each square in the decomposition to be visited in the solution for that square. More precisely, we will remove all the segments crossing a horizontal dissecting line (i.e. those cover-lines in G_{j^*}), and instead consider some of the portals around each square to be *required* to be visited in a feasible solution. We show there is a feasible solution that visits all the remaining segments as well as the "required" portals, of total cost at most $(1+\varepsilon)\cdot\text{opt}$, and that such a solution can be extended to a feasible solution visiting all the segments of the original instance (i.e. including the ones that we dropped) at an extra cost of $O(\varepsilon \cdot \text{opt})$.

3.1.1 Dropping the segments intersecting horizontal dissecting lines

We say the edges of the bounding box are *level 0* dissecting lines, the first pair of dissecting lines are *level 1* dissecting lines, and so on.

Consider a square S in our hierarchical decomposition and suppose it is cut into four squares S_1, S_2, S_3, S_4 by two dissecting lines where the horizontal one, line Γ, is the cover-line C_τ from G_{j^*}, and is a level j dissecting line. Recall that we place a total of $2m$ portals along Γ inside S; m portals on the common sides of S_1, S_4 and m along the common side of S_2, S_3. Define $B_\tau(S)$ to be the set of intervals in B_τ (intervals of C_τ) that cover a segment that lies inside S (and so intersects with Γ) (see Figure 3.1).

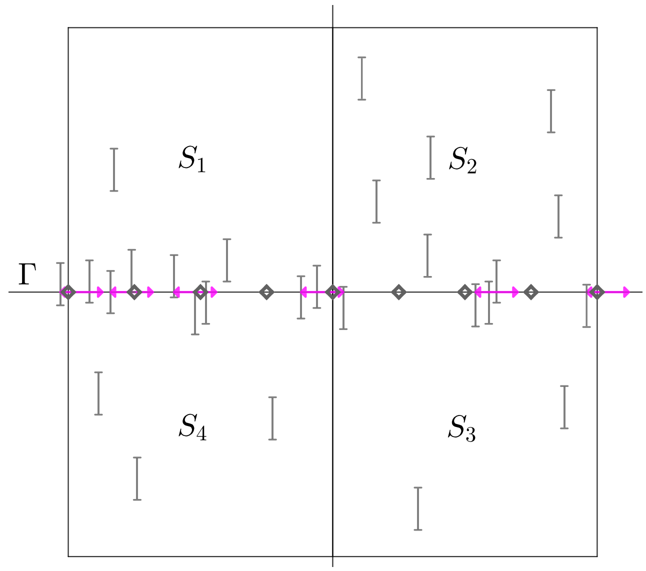


Figure 3.1: Breaking a square S into 4 smaller squares. The magenta parts on line Γ (i.e. the cover-line C_τ) show the interval set $B_\tau(S)$.

For each $b \in B_\tau(S)$, suppose $p(b)$ is the nearest portal to it in S among the portals on Γ, and let $s(b)$ be the left-most segment covered by b that is in S. We are going to modify OPT in the following way: Consider a point p_s on $s(b)$ visited by OPT. Insert the following "legs" to the path: travel from p_s vertically along $s(b)$ until you arrive at its intersection with Γ, i.e. arrive on interval b (this length is at most ρ), then travel along Γ to the right-most segment covered by b (this is also at most ρ), and then travel to $p(b)$, and then travel back to p_s. For every other segment s' covered by b in S, we are going to short-cut any point on s' that was visited by OPT as all these segments are now covered by the newly added legs (see Figure 3.2). We also short-cut the second visit to p_s.

Using triangle inequality, the expected length of the new legs will increase the cost of the solution by at most $2\rho + 2||p_s p(b)|| \leq 2(\rho + \frac{N}{2^j m})$. We do this for all the intervals on Γ and inside S, i.e. if OPT visits a segment covered by that interval b, we change OPT to make a detour to visit $p(b)$ as well. Note that each interval $b \in B_\tau$ can belong to at most two $B_\tau(S)$'s (two adjacent squares that b intersects with), and the intervals for which this modification can happen for, are at least $h \cdot \rho$ apart because that is the minimum size of a

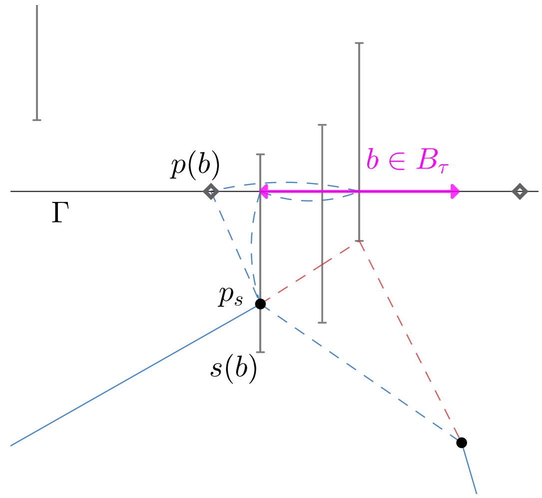


Figure 3.2: The modified solution for dropping segments crossing horizontal dissecting lines: follow the blue dashed lines from p_s; red dashed lines are parts of the original path.

square of the dissection.

Given the random choice of our dissecting lines, since dissecting lines are $h \cdot \rho$ apart, are randomly chosen, and each interval has length ρ, the probability that an interval $b \in B_\tau$ appears in two $B_\tau(S)$'s (i.e. cut by a dissecting line), is at most $1/h = \varepsilon$. Also, each cover-line in G_{j^*} is a level j dissecting line with probability $2^{j-1}/(N/\rho h)$. Thus, the expected increase in the cost by this modification for all the interval of C_τ is at most

$$\sum_{j=1}^{\log(N/\rho h)} \Pr[\Gamma \text{ is level } j] \cdot (1+\varepsilon) \cdot |B_\tau| \cdot 2(\rho + \frac{N}{2^j m})$$

$$\begin{aligned} &\leq 2(1+\varepsilon) \cdot |B_\tau| \cdot \sum_{j=1}^{\log(N/\rho h)} \frac{2^{j-1}}{N/\rho h} \cdot (\rho + \frac{N}{2^j m}) \\ &\leq 2(1+\varepsilon) \cdot |B_\tau| \cdot \frac{\rho h}{N} \cdot \left(\frac{N}{h} + \frac{N \log(N/\rho h)}{2m} \right) \\ &\leq (1+\varepsilon) \cdot |B_\tau| \cdot \rho \cdot (1 + \varepsilon h) \\ &\leq 4\rho \cdot |B_\tau|. \end{aligned}$$

Considering all cover-lines in G_{j^*}, this implies the total expected increase in the cost is at most $\sum_{C_\tau \in G_{j^*}} 4\rho |B_\tau|$, which combined with Lemma 26, implies with probability at least $1/2$, the increase in total cost is at most $O(\varepsilon \cdot \text{opt})$. Each portal p that is visited by a detour as described above is called a *required* *portal.*

In fact, we can short-cut more paths so that the number of detours to each portal is bounded by 2. Informally, only the left-most interval to the left of p that has made a detour to p, along with the right-most interval to the right of p that has made a detour to p are sufficient to cover all the segments of the intervals in between them. More specifically, consider a portal p on Γ and let $b_L(p)$ be the left-most interval in $B_\tau(S)$ to the left of p that covered a segment whose path was detoured to visit p (null if there is no such interval). Similarly, let $b_R(p)$ be the right-most interval among $B_\tau(S)$ to the right of p that covered a segment whose path was detoured to visit p (this too can be null if there is no such interval). In other words, there was a segment $s_L = s(b_L(p))$ and a segment $s_R = s(b_R(p))$ that were visited by OPT, and we made a detour to p when OPT visited s_L and s_R. The detour from s_L to p covers all the segments of intervals between $b_L(p)$ and p. Similarly, the detour from s_R to p covers all the segments of interval between p and s_R. Thus, for any interval b' between $b_L(p)$ and $b_R(p)$, all the segments covered by b' are also covered by the detours of s_L and s_R. This means for all those intervals b', we can short-cut the segments covered by them entirely (in particular, they don't need to make a detour to p). Therefore, at most two intervals will have detours to p, namely $b_L(p)$ and $b_R(p)$. And the detours to different portals are disjoint, so the added detours don't overlap on Γ, and since short-cutting doesn't increase the shadow, we only add a shadow of at most 2 per cover-line to the solution. This implies that if we focus on the modified solution restricted to the strip between two cover-lines in G_{j^*}, it still has a bounded shadow. These arguments imply the following:

Lemma 27 *Given instance $\mathcal{I}$, there is another instance $\mathcal{I}'$ that is obtained by removing all the segments that are crossing cover-lines in G_{j^*} (i.e. intersecting horizontal dissecting lines), and instead some of the portals around (more precisely, the top and bottom sides of) each square of quad-tree dissection are required to be covered (visited); such that there is a solution for $\mathcal{I}'$ of cost at most $(1+O(\varepsilon)) \cdot opt$, and such a solution can be extended to a feasible solution* *of $\mathcal{I}$ of cost at most $(1+O(\varepsilon))\cdot opt$. Furthermore, the shadow of the solution for $\mathcal{I}'$ between any two consecutive cover-lines in G_{j^*} is at most 4 more than the shadow of OPT between those two lines.*

3.2 Outer DP

The outer DP based on the quad-tree dissection is similar to the classic PTAS for Euclidean TSP. One can show that for $r = O(1/\varepsilon)$, there is an r-light portal respecting tour for $\mathcal{I}'$ with cost at most $(1+\varepsilon)\cdot \text{opt}'$, where opt' is the cost of an optimum solution for $\mathcal{I}'$. The base case of this DP will be instances with bounding box of size $\rho \cdot h$. For such instances, we solve the problem using an inner DP that is described in the Section 3.3.

We will use the "patching lemma" the same way it is described in Arora's approach. We show there is a near optimum solution for $\mathcal{I}'$ that is portal respecting and *r-light*, meaning each square in our quad-tree decomposition is crossed by the solution only r many times on each side for a parameter $r = O(1/\varepsilon)$. Then a DP similar to the point TSP (outer DP) will combine the solutions for the subproblems to find the solution for a bigger subproblem. Since we don't know which portals for each square are supposed to be "required" in $\mathcal{I}'$ (so that the solution can be extended to cover the dropped line segments), for each such square we "guess" the set of required portals in our DP; i.e. we will have an entry for each guessed set of portals on the horizontal sides of a square as the set of required portals in our DP. Since the number of portals is logarithmic, this guessing remains polynomially bounded. For now, assume that we know all the required portals, and hence, instance $\mathcal{I}'$ itself (even though $\mathcal{I}'$ is defined based on OPT which we don't know).

Consider instance $\mathcal{I}'$ and let OPT' be the optimum solution for it, and let the cost of that solution be opt'. For each dissecting line Γ (vertical or horizontal), let $t(\Gamma)$ be the number of intersections of OPT' with Γ and $T = \sum_{\Gamma} t(\Gamma)$.

Lemma 28 ([4]) *$T \leq 2 \cdot opt'/(\rho h)$.*

Proof. Let $\ell = (x_1, y_1) \to (x_2, y_2)$ be any leg of OPT$'$. Let $\Delta x = |x_1 - x_2|$ and $\Delta y = |y_1 - y_2|$. The contribution of ℓ to opt$'$ is its length, i.e. $L_\ell = \sqrt{(\Delta x)^2 + (\Delta y)^2}$. Note that due to the scaling, we have $L_\ell \geq 4$. Since the dissecting lines are ρh apart, there are at most $(\Delta x + \Delta y + 2)/(\rho h)$ dissecting lines that intersect with Γ; so the contribution of Γ to T is at most the same amount. Using the Cauchy-Schwarz inequality, we have $2((\Delta x)^2 + (\Delta y)^2) \geq (\Delta x + \Delta y)^2$, which implies

$$(\Delta x + \Delta y + 2)/(\rho h) \leq (\sqrt{2((\Delta x)^2 + (\Delta y)^2)} + 2)/(\rho h) = (\sqrt{2} \cdot L_\ell + 2)/(\rho h).$$

It suffices to show $\sqrt{2} \cdot L_\ell + 2 \leq 2L_\ell$; this is easily seen to be true because $L_\ell \geq 4$. Therefore, if we add these inequalities for all legs ℓ of OPT$'$, we get the lemma's statement as the result. ■

The following lemma is essentially the same as the one in the case of point TSP (except we have different stopping points):

Lemma 29 ([4]) *Considering the randomness of the dissecting lines, with probability of at least $\frac{1}{2}$, there exists a portal-respecting solution for $\mathcal{I}'$ with cost at most $(1+\varepsilon) \cdot opt'$ for portal parameter $m = O(\frac{1}{\varepsilon} \cdot \log \frac{N}{\rho h})$*

Proof. The proof is similar to the survey in [24]. Consider any dissecting line Γ of level j and focus on the intersections of OPT$'$ with that line. Consider any leg $\ell = ab$ of OPT$'$ which intersects Γ, say at a point q and suppose p is the nearest portal of Γ to q. Replace ℓ with with two new "legs" $\ell_1 = ap$ and $\ell_2 = pb$. Let d be the distance of q to p. Using triangle inequality, it can be seen that $\ell_1 + \ell_2 \leq \ell + 2d$; meaning the additional cost for going through portal p is at most $2d$. The distances between the portals on level j line Γ are $d_j = \frac{N}{2^j m}$, and clearly $d \leq d_j$. Recall that OPT$'$ intersects with Γ, $t(\Gamma)$ times. Thus, the expected increase in cost for any dissecting line Γ is at most

$$\begin{aligned}\sum_{j=1}^{\log N/\rho h} \Pr[\Gamma \text{ is level } j] \cdot t(\Gamma) \cdot 2 \cdot \frac{N}{2^j m} &\leq \sum_{j=1}^{\log N/\rho h} \frac{2^{j-1}}{N/\rho h} \cdot t(\Gamma) \cdot 2 \cdot \frac{N}{2^j m} \\ &= \frac{\rho h}{m} \cdot \sum_{i=1}^{\log N/\rho h} t(\Gamma) \\ &= \frac{\rho h}{m} \cdot \log \frac{N}{\rho h} t(\Gamma).\end{aligned}$$

For $m \geq \frac{4}{\varepsilon} \log \frac{N}{\rho h}$, the last value above is at most $\frac{\varepsilon \rho h}{4} \cdot t(\Gamma)$. Adding all these inequalities over different Γ's gives us $\frac{\varepsilon \rho h}{4} \cdot T$, which according to Lemma 28 is at most $\frac{\varepsilon}{2} \cdot \mathrm{opt}'$. Using Markov's inequality the statement of the lemma follows.

■

The patching Lemma (stated below) for classic Euclidean TSP holds in our setting as well.

Lemma 30 (The patching Lemma [4]) *For any dissecting line segment τ with length L_τ, if a tour crosses τ more than twice, it can be altered to still contain the original tour, but intersect with τ at most twice with an additional cost not greater than $6L_\tau$.*

Proof. The same proof as in [4] applies here. ■

Observation 6 *A single point can be seen as a 0-length segment. By using Lemma 30, we get that at no additional cost (i.e. extra cost of 6×0), each portal is visited at most twice.*

The next lemma shows the existence of a near-optimum solution that is r-light and portal respecting for $r = O(1/\varepsilon)$:

Lemma 31 *Given the randomness in picking the dissecting lines, with probability at least $\frac{1}{2}$, there is an r-light portal respecting tour for $\mathcal{I}'$ with cost at most $(1+\varepsilon) \cdot opt'$ for $r = O(\frac{1}{\varepsilon})$.*

Proof. This is implied by the *Structure Theorem* in [4], and the similar proof works here.

■

3.2.1 DP Table and Time Complexity

The outer DP is similar to the DP for classic Euclidean TSP except that we need to take care of required portals that are going to be guessed and passed down to the subproblems. Note that there are $O(n)$ subproblems in each level of the dissection tree, and so a total of $O(n \log n)$ squares to consider. For each square S with $4m$ portals around it, we guess a subset of portals on the horizontal sides of S to be required. The number of such guesses is 2^{2m} where $m = O(\frac{1}{\varepsilon} \cdot \log \frac{N}{\rho h}) = O(\log n/\varepsilon)$. There are $(4m+1)^{4r}$ guesses for up to $4r$ portals to be chosen for an r-light portal respecting, and at most $(4r)!$ for the pairings of these portals. So the size of the DP table is at most $O(n \log n \cdot 2^{2m} \cdot (4m+1)^{4r} \cdot (4r)!) = O(n \log^{O(r)} n)$.

The DP table is filled bottom up. The base cases are when we have a square of side length $\rho \cdot h$. These subproblems are solved using the inner DP described in the next section. For every other square S that is broken into 4 squares $S_1, \ldots, S_4$, we solve the subproblem of S after we have solved all subproblems for $S_1, \ldots, S_4$. The way we combine the solutions from those of the sub-squares to obtain the solution for S is very much like the classic point TSP. However, we have to extend the solutions so that the line segments that were intersecting the horizontal dissecting line that split S, are now fully covered by the guessed required portals for $S_1, \ldots, S_4$. More specifically, suppose Γ is the horizontal dissecting line that corresponds to a cover-line C_τ from group G_{j^*} (and hence we removed all the segments crossing C_τ and instead made some of the portals along C_τ as required). We add those segments of the instance back, and we extend the solutions from the require portals to travel left and right to cover these segments. Similar to the classic TSP, the total time to fill in the outer DP table is $O(n \log^{O(r)} n)$.

3.3 Inner DP

Recall that each base case of the quad-tree decomposition is a subproblem defined on a square S with size $\rho h \times \rho h$, and has $4m$ portals around it. Since we assume the solution we are looking for is r-light, it means the instance defined by S has also a set P of size at most $4r$ of portal pairs (where $r = O(1/\varepsilon)$). Each pair $(p_i, q_i) \in P$ specifies that the solution restricted to S, has a p_i, q_i-path. We are also given a guessed subset Q of the portals around S (specifically on the top and bottom side of S) as the required portals. The goal is to find a minimum cost collection of paths that start/end at the given set of portal pairs P that cover all the line segments in S, as well as visit all the required portals in Q. Let us denote this instance by (S, P, Q). Note that by Theorem 2, Lemma 27, and Lemma 31, there is a near-optimum solution such that it is r-light for each square of the dissection, is portal respecting, covers all the required portals, and has shadow bounded by $O(1/\varepsilon^2)$. Also using Lemma 19, the length of any pure reflection sequence in it is bounded by $O(1/\varepsilon)$. We describe the inner DP to find an optimum solution with bounded shadow (and pure reflection sequence bounded to $O(1/\varepsilon)$ elements) restricted to subproblem (S, P, Q). For square S, let us use OPT_S and opt_S to denote such a bounded shadow optimum solution and its value, respectively.

Informally, the DP is a (nontrivial) generalization of the DP for the classic (and textbook example) bitonic TSP in which the shadow is 2. In our case, the shadow is $O(1/\varepsilon^2)$. We are going to consider a sweeping vertical line Γ in S (that moves left to right) and "guess" the intersections of OPT_S with it.

We define an *event point set* in the following way:

Definition 12 (Event Point) *Given a subproblem triplet (S, P, Q), each line segment in S is in the event point set. Also, each portal that is on a horizontal side of S and is either in Q, or participates in a pair of P, is also in the event point set.*

We consider an ordering of all the elements in the event point set from left to right (i.e. increasing x-coordinate), say $v_1, v_2, \dots, v_{n_S}$, where n_S is the number of event points; note that $n_S = O(n)$. There are $n_S - 1$ equivalent classes for positions of Γ, where each class corresponds to when Γ is located between v_i, v_{i+1}. A sweep line between v_i, v_{i+1} is denoted by Γ_i. Since the shadow of OPT_S is bounded, the intersection of Γ_i with OPT_S has a low complexity. We will give a more concrete explanation of that complexity below.

Recall Observation 3 and the types of points in a solution (straight point, break point, or reflection point). Also recall the definition of a pure reflection point (a reflection point that is not at a tip of a segment of the instance). Consider the global optimum solution that is r-light and portal respecting with bounded shadow and bounded pure reflection sequence that also covers the required portals of each square. Suppose $p_{a_1}, p_{a_2}, \dots, p_{a_k}$ is the sequence of points in S visited by OPT_S in this order that are *not* a straight point nor a pure reflection point; so each of them is a break point (tip of a segment) or perhaps a required portal in Q, or a portal in P (i.e. is an entry or exit point in some pair belonging to P). So any point visited by OPT_S between $p_{a_i}, p_{a_{i+1}}$ (if there is any) is either a straight point or a pure reflection point. We define subpaths of OPT_S named *large legs* as follows:

Definition 13 (Large Leg) *The path of OPT_S from p_{a_i} to $p_{a_{i+1}}$ is a large leg. Each large leg starts and ends from a portal or a tip of a segment, and all the points in between are either straight points or pure reflection points.*

It follows from Lemma 19 that the number of pure reflection points in each large leg is bounded by $O(1/\varepsilon)$. Each large leg can be guessed by making at most $O(1/\varepsilon)$ guesses for segments or points: guess the two end-points of the large leg (which are either portals or tips of segments), then guess at most $O(1/\varepsilon)$ segments that have pure reflection points on them; once we guess the two end-points and the segments for pure reflections, the pure reflection points are uniquely determined. Since there are $O(n^2)$ choices for the end-points and $O(n^{1/\varepsilon})$ choices for the segments of pure reflection points, the total number of possible large legs is bounded by $n^{O(1/\varepsilon)}$. Now since we assume OPT_S has bounded shadow of $O(1/\varepsilon^2)$, for any sweep line Γ_i, there are at most $O(1/\varepsilon^2)$ large legs of OPT_S that can cross Γ_i.

So for a fixed i (and sweep the line Γ_i), let $\mathcal{L}_i = L_1, \ldots, L_\sigma$ be the sequence of large legs ($\sigma = O(1/\varepsilon^2)$) of OPT_S that cross Γ_i; where each large leg is specified by the end-points as well as the intermediate segments for pure reflections (if there are any). Then the number of possible choices for $\mathcal{L}_i$ is $n^{O(1/\varepsilon^3)}$. Given i and $\mathcal{L}_i$, let S_i^L, S_i^R be the left and right part of S (cut by Γ_i). If we ignore the segments covered by $\mathcal{L}_i$ in S_i^L, and consider the end-points of each L_j as portals too, then the restriction of OPT_S to S_i^L is a collection of paths that start/end at portals of P in S_i^L or end-points of L_j's in S_i^L that cover all the segments in S_i^L not already covered by $\mathcal{L}_i$, as well as points in $Q \cap S_i^L$. More specifically, each part of OPT_S in S_i^L is a path that starts at a p_j for a pair $(p_j, q_j) \in P$, or at an end-point of L_k that is in S_i^L and ends at a point $p_{j'}$ (or $q_{j'}$) of another pair in P that is also in S_i^L, or at another end-point of some $L_{k'}$ that is in S_i^L. So this induces some pairs of points, denoted by P_i^L:

Definition 14 (Path-wise Pairing P_i^L) *Set P_i^L of pairs of points is said to be the path-wise pairing for S_i^L, if there is a path in S_i^L between the two points of any given pair $(a, b) \in P_i^L$. Furthermore, each point in a pair $(a, b) \in P_i^L$ is either a portal in S_i^L that is part of a pair in P, or is an end-point of a large leg L_j that is in S_i^L.*

For any such point in S_i^L, say p, there must be a pair in P_i^L containing that point. We also assume $(p, p) \in P_i^L$, and if p is an end point of a large leg in $\mathcal{L}_i$ or S_i^L, and if q is the other end point of that large leg, then $(p, q) \in P_i^L$.

We say a set of pairs P_i^L is not *promising* if given $\mathcal{L}_i$, there is no feasible solution in the entire S whose restriction to S_i^L defines subpaths consistent with P_i^L (i.e. they start and end on the same pairs as specified by P_i^L). Otherwise, we consider it promising. For example if $(p_j, q_j) \in P$, both p_j, q_j belong to S_i^L, and if $(p_j, u), (q_j, v) \in P_i^L$ where u is one end of a long leg L_1 and v is one end of a long leg L_2, it must be the case that it is possible to have a path from the other end of L_1 to the other end of L_2. This would be impossible if, for instance, those other ends of L_1, L_2 are paired up with other portals in P_i^L. Note that since there are at most $4r$ pairs in P and $O(1/\varepsilon^2)$ end-points in $\mathcal{L}_i$, the number of possible choices for P_i^L is $(1/\varepsilon)^{O(1/\varepsilon)}$. Also, a given P_i^L (together with $\mathcal{L}_i$), it can be checked if P_i^L is promising or not in poly-time in n.

This suggests how we can break the instance (S, P, Q) into polynomially many sub-instances. For a fixed i, guess $\mathcal{L}_i$ among all those with shadow $O(1/\varepsilon^2)$, break S into S_i^L, S_i^R, let $Q_i^L = Q \cap S_i^L$, and guess the new pairs P_i^L (for S_i^L) that are promising. We solve (S_i^L, P_i^L, Q_i^L) for each S_i^L, P_i^L, Q_i^L obtained this way. We can solve each such subproblem assuming we have solved all subproblems defined by each Γ_j for $j < i$. So formally, let us define a *configuration*:

Definition 15 (Configuration) *A configuration is a vector $(i, \mathcal{L}_i, P_i^L)$ where* *the components are:*

- *i (indicating Γ_i and defining S_i^L),*
- *The large legs of OPT_S crossing Γ_i, denoted by $\mathcal{L}_i$, $|\mathcal{L}_i| = O(1/\varepsilon^2)$,*
- *The pairing P_i^L defined by $\mathcal{L}_i$, P, and the restriction of OPT_S to S_i^L.*

This configuration (see Figure 3.3), defines a subproblem: Suppose $\mathcal{L}_i$ is a given set of large legs crossing Γ_i. Find a collection of paths in S_i^L such that P_i^L specifies the start/end of these paths (and is promising), such that these paths cover all the segments in S_i^L (excluding those already covered by $\mathcal{L}_i$), and also cover all the points in $Q \cap S_i^L$, with shadow at most $O(1/\varepsilon^2)$.

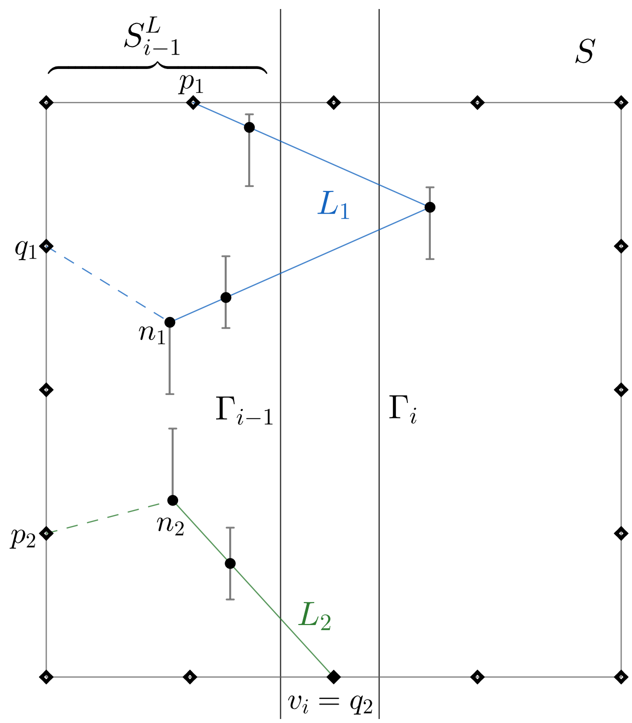


Figure 3.3: An example of an event point v_i and vertical lines Γ_{i-1}, Γ_i from two consecutive equivalent classes in square S. In this figure, $\mathcal{L}_{i-1} = L_1, L_2$ and $\mathcal{L}_i = L_1$; plus, it is the case that $(p_2, n_2) \in P_{i-1}^L, (q_1, n_1) \in P_{i-1}^L$, and $(q_1, n_1) \in P_i^L$.

The cost of this solution is defined to be the sum of the costs of all the edges that are entirely (i.e. both end-points) in S_i^L (including those legs of a large leg in $\mathcal{L}_i$ that are entirely in S_i^L, but not those that are crossing Γ_i). Entry

$A[i, \mathcal{L}_i, P_i^L]$ of the inner DP, stores the minimum cost of such a solution. Recall that there are $n_s = O(n)$ choices for i (and so for Γ_i), $n^{O(1/\varepsilon^3)}$ choices for $\mathcal{L}_i$, and $(1/\varepsilon)^{O(1/\varepsilon)}$ choices for P_i^L. So there are $n^{O(1/\varepsilon^3)}$ possible configurations, which is the size of our DP table as well.

We fill in the entries of this table $A[.,.,.]$ for increasing values of i. For $i = O(1)$, $A[i,.,.]$ can be computed exhaustively in $O(1)$ time.

For any other value of i, we compute $A[i, \mathcal{L}_i, P_i^L]$ by considering various subproblems $(i-1, \mathcal{L}_{i-1}, P_{i-1}^L)$ that are *consistent* (see Subsection 3.3.1) with $(i, \mathcal{L}_i, P_i^L)$. Consider event point v_{i-1}; it is either a segment or a portal that is between Γ_{i-1} and Γ_i; which means it does not belong to S_{i-1}^L, but belongs to S_i^L. Consider the solution for $(i, \mathcal{L}_i, P_i^L)$, and the legs (in that solution) that visit v_i. In case v_i is a start/end portal in P, there is one leg incident to v_i; if $v_i \in Q$ there are two legs incident to v_i, and if v_i is a segment, there are two legs that are incident to a point v_i' on that segment. If there is one leg only (v_i is a start/end portal), call that leg ℓ_i, and if there are two legs, call them ℓ_{i-1}, ℓ_i. Depending on whether these legs cross Γ_{i-1} or Γ_i, we have the following situations, which are the *consistent* outcomes:

1. v_i **is a start/end portal**, we consider 2 different subcases:

   (a) ℓ_i **crosses** Γ_{i-1} **but not** Γ_i: Say $\ell_i = v_i u$, where u is a point in S_{i-1}^L. In this case, there is a large leg $L \in \mathcal{L}_{i-1}$ with one end-point v_i. Then if L crosses Γ_i, it means L is a large leg in $\mathcal{L}_i$. If L does not cross Γ_i, then $\mathcal{L}_i = \mathcal{L}_{i-1} \setminus L$. We consider both possibilities and in each case, consider P_{i-1}^L's that are consistent with P_i^L and set $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}} \{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\} + ||\ell_i||$.

   (b) ℓ_i **crosses** Γ_i **but not** Γ_{i-1}**:** In this case, there is a large leg $L \in \mathcal{L}_i$ that starts with ℓ_i and does not cross Γ_{i-1}, so does not belong to $\mathcal{L}_{i-1}$. All the other large legs in $\mathcal{L}_{i-1}$ and $\mathcal{L}_i$ are the same (as there is no other event point between Γ_{i-1} and Γ_i), and P_i^L and P_{i-1}^L are consistent. Then $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}} \{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\}$.

2. $v_i \in Q$, we consider 3 different subcases:

(a) ℓ_{i-1}, ℓ_i **both cross** Γ_{i-1} **but not** Γ_i**:** In this case, there are two large legs $L, L' \in \mathcal{L}_{i-1}$ that both end at v_i, say L contains ℓ_i and L' contains ℓ_{i-1}. If L crosses Γ_i, then L is a large leg in $\mathcal{L}_i$ as well, similarly for L'. The other large legs of $\mathcal{L}_i$ and $\mathcal{L}_{i-1}$ are the same, and P_{i-1}^L is consistent with P_i^L. We set $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}}\{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\} + ||\ell_i|| + ||\ell_{i-1}||$.

(b) ℓ_{i-1}, ℓ_i **both cross** Γ_i **but not** Γ_{i-1}**:** This similar to the previous case. There are two legs $L, L' \in \mathcal{L}_i$ that both start at v_i, say L contains ℓ_i and L' contains ℓ_{i-1}. If L crosses Γ_{i-1}, then L is a large leg in $\mathcal{L}_{i-1}$ as well, similarly for L'. The other large legs of $\mathcal{L}_i$ and $\mathcal{L}_{i-1}$ are the same and P_{i-1}^L is consistent with P_i^L. In this case, $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}}\{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\}$.

(c) **Exactly one of** ℓ_{i-1}, ℓ_i **crosses** Γ_{i-1} **and one crosses** Γ_i**:** Say ℓ_{i-1} crosses Γ_{i-1}, and ℓ_i crosses Γ_i. So ℓ_{i-1} will be the last leg of a large leg $L \in \mathcal{L}_{i-1}$, and ℓ_i will be the first leg of a large leg $L' \in \mathcal{L}_i$. If L does not cross Γ_i, then L is not in $\mathcal{L}_i$ at all. Similarly, if L' doesn't cross Γ_{i-1}, then L' isn't a large leg in $\mathcal{L}_{i-1}$. We consider both possiblities (i.e. consider sets $\mathcal{L}_{i-1}$ that are consistent with one of these cases). $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}}\{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\} + ||\ell_{i-1}||$.

3. v_i **is a segment**: Subcases are similar to the previous case; let v_i' be the intersection point of OPT_S with v_i:

(a) ℓ_{i-1}, ℓ_i **both cross** Γ_{i-1} **but not** Γ_i**:** If v_i' is a tip, then ℓ_{i-1} is the last leg of a large leg $L \in \mathcal{L}_{i-1}$, and ℓ_i is the last leg of another large leg $L' \in \mathcal{L}_{i-1}$. Depending on whether L (L') crosses Γ_i, it can be a large leg in $\mathcal{L}_i$ or not. We consider both possibilities. If v_i' is not a tip, then it must be a pure reflection, so there must be a large leg $L \in \mathcal{L}_{i-1}$ that contains this as a pure reflection. That large leg may or may not belong to $\mathcal{L}_i$. We consider all these possibilities (i.e. those $\mathcal{L}_{i-1}$ consistent with these), and also for each case consider a P_{i-1}^L consistent with P_i^L. Then set $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}}\{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\} + ||\ell_{i-1}|| + ||\ell_i||$.

(b) ℓ_{i-1}, ℓ_i **both cross Γ_i but not Γ_{i-1}:** If v_i' is a tip, then ℓ_{i-1} is the first leg of a large leg $L \in \mathcal{L}_i$, and ℓ_i is the first leg of another large leg $L' \in \mathcal{L}_i$. Depending on whether L (L') crosses Γ_{i-1}, it can be a large leg in $\mathcal{L}_{i-1}$ or not. We consider both possibilities. If v_i' is not a tip, then it must be a pure reflection, so there must be a large leg $L \in \mathcal{L}_i$ that contains this as a pure reflection. That large leg may or may not belong to $\mathcal{L}_{i-1}$ depending on whether it crosses Γ_{i-1} or not. We consider all these possibilities, and also for each case consider a P_{i-1}^L consistent with P_i^L. Then set $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}}\{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\}$.

(c) **Exactly one of ℓ_{i-1}, ℓ_i crosses Γ_{i-1} and one crosses Γ_i:** In this case, v_i' must be a tip or a straight point. Say ℓ_{i-1} crosses Γ_{i-1}, and ℓ_i crosses Γ_i. If v_i' is a tip, then ℓ_{i-1} is the last leg of a large leg $L \in \mathcal{L}_{i-1}$, and ℓ_i is the first leg of a large leg $L' \in \mathcal{L}_i$. L may cross Γ_i (in which case it also belongs to $\mathcal{L}_i$), also L may cross Γ_{i-1} in which case belongs to $\mathcal{L}_{i-1}$. We consider these possibilities. If v_i' is a straight point, then both ℓ_{i-1}, ℓ_i are part of a large leg $L \in \mathcal{L}_{i-1}$, and L belongs to $\mathcal{L}_i$ as well. We consider all these cases and consistent P_{i-1}^L, P_i^L and set $A[i, \mathcal{L}_i, P_i^L] = \min_{P_{i-1}^L, \mathcal{L}_{i-1}}\{A[i-1, \mathcal{L}_{i-1}, P_{i-1}^L]\} + ||\ell_{i-1}||$.

3.3.1 Consistent Subproblems

The consistency of a subproblem by configuration $(i, \mathcal{L}_i, P_i^L)$, with a previous subproblem by configuration $(i-1, \mathcal{L}_{i-1}, P_{i-1}^L)$, comes down to one of the cases mentioned in the previous Section. In each subcase, we only need to define what we mean by consistent P_i^L and P_{i-1}^L.

We say P_i^L as a part of the configuration $(i, \mathcal{L}_i, P_i^L)$, and P_{i-1}^L as a part of the configuration $(i-1, \mathcal{L}_{i-1}, P_{i-1}^L)$ are consistent if for any pair $(a, b) \in P_i^L$:

- If both a, b are in S_{i-1}^L, then either:

– $(a, b) \in P^L_{i-1}$, or

– (When $v_i \in Q$ or when v_i is a segment containing a pure reflection) There is a large leg $L_j \in \mathcal{L}_{i-1} \cup \mathcal{L}_i$ with end points p_1, p_2 corresponding to (i.e. having an intersection with) the event point v_i, such that $(a, p_1), (b, p_2) \in P^L_{i-1}$, or

– (When $v_i \in P$ or v_i is a segment containing a non-pure reflection or a break point) There are two large legs (in $\mathcal{L}_{i-1} \cup \mathcal{L}_i$) that have v_i as an end point, and have another end point, say respectively p_1 and p_2, such that $(a, p_1), (b, p_2) \in P^L_{i-1}$.

- If both a, b are not in S^L_{i-1}, then it means that either a or b, say a, corresponds to the event point v_i. This means either a is a portal ($\in P \cup Q$) between Γ_{i-1} and Γ_i, or a is a tip of the segment corresponding to v_i. In either case, there is at least a large leg $L_j \in \mathcal{L}_{i-1} \cup \mathcal{L}_i$ that has a as one of its end points. There can be at most two such large legs; say p_1 and possibly p_2 are the other ends of these at most two large legs. Then it must be the case that (either) $(b, p_1) \in P^L_{i-1}$ (or $(b, p_2) \in P^L_{i-1}$).

3.4 Algorithm for Similar-Length Line Segments

We first finalize the proof of our algorithm for the case of unit length segments, then generalize the proof to the case of similar-length segments.

3.4.1 Unit-Length Line Segments

We prove the following theorem to finalize the proof for unit-length line segments:

Theorem 4 *There is a $(1 + \varepsilon)$-approximation algorithm for TSPN over n parallel unit-length line segments that runs in time $n^{O(1/\varepsilon^3)}$.*

Proof. Take any instance of the problem. As described at the beginning of this chapter, we first scale the instance (at a loss of $(1 + \varepsilon)$) so that all segments have integer coordinates. We employ the hierarchical decomposition of Arora using dissecting lines as described in Section 3.1, and drop the line segments crossing horizontal dissecting lines as described in Subsection 3.1.1. We require a subset of portals around each square S of the dissectioning to be covered in the subproblems as described in the outer DP in Section 3.2. Lemma 27 shows that we lose at most another $(1 + \varepsilon)$ factor in doing so. At the leaf level of our decomposition, we need to solve instances where each square has sides of length $\rho \cdot h$. Note that as discussed in the first paragraph of Subsection 3.3, for any base square of the dissection, using Theorem 2, Lemma 27, Lemma 31, and Lemma 19, there is a near-optimum solution such that it is portal respecting, r-light for $r = O(\frac{1}{\varepsilon})$, covers all the required portals, has a shadow bounded by $O(1/\varepsilon^2)$, and the length of any pure reflection sequence in it is bounded by $O(1/\varepsilon)$. The inner DP describes how to find such a solution. Note that the size of the inner DP table is $n^{O(1/\varepsilon^3)}$. To compute each entry, we may consider (at worst) all other entries, and so the time complexity of computing the table for each square S is at most $n^{O(1/\varepsilon^3)}$. Given that the number of squares at the leaf nodes of the decomposition is $O(n \log^{O(r)} n)$, the total time for the inner and outer DP is $n^{O(1/\varepsilon^3)}$. ■

3.4.2 Similar-Length Line Segments (Main Theorem)

We finally prove the main theorem in this thesis, which we reiterate here for convenience:

Theorem 1 *Given a set of n parallel line segments with lengths in $[1, \lambda]$ for a fixed λ as an instance of TSPN, there is an algorithm that finds a $(1 + \varepsilon)$-approximation solution in time $n^{O(\lambda/\varepsilon^3)}$.*

Proof. We discuss how the result presented for unit-length line segments in Theorem 2 can be extended to the case that line segments have length ratio $\lambda = O(1)$, and obtain a PTAS for it. In the case of segments with lengths in $[1, \lambda]$, for every strip of height 1, we still have some top and bottom segments and we might have some line segments that completely span the height of the strip. Let's call these segments *full segments* of a strip. We claim that whenever we change the solution in the proof of Theorem 2 to one that has a bounded shadow, the full segments of the strip remain covered. These changes are done in Lemmas 19 and 18. For each of these cases, any new subpath (with smaller shadow) that replaces a subpath of larger shadow, will travel the same interval in the x-coordinate, and hence any full segment covered by the original path, remains covered by the new path.

Next, when we scale the instance, we get line segments with length between $[\rho, \lambda\rho]$. Now we do our hierarchical decomposition until base squares have side length of $\lambda\rho h$, so the space between two cover-lines in the same group is $\lambda\rho h$ instead of ρh. Lemma 25 holds with bound opt $\geq \frac{\rho\cdot|\mathcal{B}|}{6\lambda}$. This implies Lemma 26 holds if j is chosen from $[1 \ldots h\lambda]$. It is straight-forward to check that Lemma 27 holds with the same ratio. For the inner DP, noting that the instance we start from has height $\rho\lambda h$, the shadow is bounded by $O(\lambda/\varepsilon^2)$. The same DP works but the runtime will be $n^{O(\lambda/\varepsilon^3)}$. This implies we get a PTAS with the same run time which completes the proof of Theorem 1. ■

Chapter 4

Conclusion, Further Extensions, and Open Problems

In this thesis, we proved Theorem 1, that there is a PTAS for parallel line segments with comparable sizes. Recall that in [12], it is shown that the Euclidean TSPN for segments of comparable sizes in arbitrary orientation is **APX**-hard. There are still a few extensions of our problem that one can consider:

1. Line segments of the instance have arbitrary sizes and they're all parallel to each other. Is there a PTAS for this case?

2. Line segments of the instance have comparable sizes, and they are parallel to the axes of the plane (so the slopes of the lines have two possible choices). Is there a PTAS for this setting?

   Note. If the segments are unit-length, we can apply our result in Theorem 1 for this case and obtain a $(2 + \varepsilon)$-approximation for this problem that runs in poly-time in the size of the input:

   Proof sketch. Split the segments into two groups based on them being horizontal or vertical. Let the minimal bounding box for the vertical segments have sides $L_v \times H_v$, and the one for horizontal segments have sides $L_h \times H_h$. Similar to what was mentioned at the start of Chapter 3, if opt is the cost of an optimum solution OPT, then $\text{opt}/2 \geq \max\{L_v, H_v - 2, L_h - 2, H_h\}$.

Consider the boxes of sizes $L_v \times (H_v - 2)$ and $(L_h - 2) \times H_h$ contained in the aforementioned minimal bounding boxes. Let B be the smallest bounding box that contains these two new boxes. So we get that opt is at least as large as any sides of B; and also it is the case that OPT lies completely inside of B.

The left side of B, refer to it as B_l either has a vertical segment on it (in the case when the left side of the $L_v \times (H_v - 2)$ box overlaps with B_l) or it has the right-most point of the left-most horizontal line (in the case when the left side of the $(L_h - 2) \times H_h$ box overlaps with B_l). The same argument holds for B_r, the right side of B. If neither of B_l and B_r are on a side of the $(L_h - 2) \times H_h$ box, then it means that all the horizontal segments of the problem have an intersection with the interior of B. In this case, take any horizontal segment s_h that has a length of l_h inside of B; we get that opt $\geq l_h$.

Assuming opt $\geq l_h$, take the portion of s_h lying inside B, and break it into $8/\varepsilon$ parts of size $l_h \cdot \varepsilon/8$. For each of these parts, consider their left-most points, and let them be $p_1, p_2, \ldots, p_{8/\varepsilon}$. OPT must intersect this segment at one of these parts. Assume that p_i is the left-most point of the part that OPT intersects with (we can check all the $8/\varepsilon$ cases).
Add p_i to the set of vertical segments, and apply the result of Theorem 1 for parameter $\varepsilon/4$ to get a solution covering all the vertical segments along with point p_i.
This solution is a lower bound for the restriction of OPT on s_h along with the vertical segments; with the exception that the intersection on s_h itself can add at most $l_h \cdot \varepsilon/4$ to the cost. So in total, this new solution along with a doubled copy of the part containing p_i, cost at most $(1 + \varepsilon/4) \cdot \text{opt} + l_h \cdot \varepsilon/4 \leq (1 + \varepsilon/2) \cdot \text{opt}$.
Do the same thing for horizontal segments, meaning find a solution covering p_i and all the other horizontal segments with parameter $\varepsilon/4$ in Theorem 1. We get another solution with cost at most $(1 + \varepsilon/2) \cdot \text{opt}$. The conjunction of these two solutions, make a feasible solution for the

main problem and cost at most $(2+\varepsilon)\cdot \text{opt}$, proving our claim.

If we don't have $\text{opt} \geq l_h$, it must be the case that there is a point on a horizontal segment on one of the vertical sides of B. This implies that OPT must specifically contain that point. Similar to above, we can add that point to the set of vertical segments and the set of horizontal segments separately; we then find solutions using Theorem 1 with parameter $\varepsilon/2$. Combining those two solutions will yield the same result. ■

3. Line segments of the instance have comparable sizes, and each segment has a slope equal to one of k possible choices, for some $k \in \mathbb{Z}^+$. What is the best approximation when:

   (a) k is a constant (specifically, is there a PTAS)?

   (b) k is any positive integer in general?

   Note. Using the result of Mitchell [19] that there exists a PTAS for TSPN over convex neighborhoods, there is a constant-factor approximation (with unspecified factor) in both cases.

4. Line segments with similar size that are at least δ apart from each other for some $\delta > 0$. Is there a PTAS for this setting?

   Answer. Yes.

   Proof sketch. We will use the result in [18] that there is PTAS for TSPN over disjoint fat objects. Assume we are given an instance $\mathcal{I}$ of the problem and there are n segments. We get that if opt is the cost of an optimum solution for I, then $\text{opt} \geq \delta \cdot n$. Take any $\varepsilon \in (0, 1)$. For each segment of length s_L, consider a $s_L \times \varepsilon\delta$ rectangle that has that segment as a side. By the assumption of the problem, it's implied that these rectangles are not intersecting each other. For small enough ε, it can be seen that all these rectangles are also "fat" aligning with the definition in [18]. Define an instance $\mathcal{I}'$ of TSPN where the neighborhoods are these rectangles we defined. Let opt' be the cost of an optimum solution for I'. Using [18], there's a PTAS for I'; we can now take a $(1+\varepsilon)$-factor solution for I', and extend each intersection with a rectangle to its corresponding segment at a total additional cost of at most $2n \cdot \varepsilon\delta$. Let this new extended solution be OPT$''$ and its cost be opt$''$. So OPT$''$ will be a feasible solution for I, and opt$'' \leq (1+\varepsilon) \cdot$ opt$' + 2n\varepsilon\delta$. Note that a feasible solution for I is also a feasible solution for I', thus opt$' \leq$ opt. So we get that opt$'' \leq (1+\varepsilon)\cdot$opt$+2n\varepsilon\delta \leq (1+\varepsilon)\cdot$opt$+2\varepsilon\cdot$opt $= (1+\varepsilon')\cdot$opt, giving us a PTAS for I. ■

Moving away from the case of neighborhoods being segments, the following open problems proposed in [18] remain:

5. Is there a PTAS for TSPN when neighborhoods are general connected shapes on the plane that don't overlap?

6. Is there a constant-factor approximation for TSPN when neighborhoods are connected general shapes on the plane?

References

[1] A. Antoniadis, K. Fleszar, R. Hoeksma, and K. Schewior, "A ptas for euclidean tsp with hyperplane neighborhoods," *ACM Trans. Algorithms*, vol. 16, no. 3, 2020, ISSN: 1549-6325. DOI: 10.1145/3383466.

[2] A. Antoniadis, S. Kisfaludi-Bak, B. Laekhanukit, and D. Vaz, "On the approximability of the traveling salesman problem with line neighborhoods," in *18th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2022, June 27-29, 2022, Tórshavn, Faroe Islands*, A. Czumaj and Q. Xin, Eds., ser. LIPIcs, vol. 227, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022, 10:1–10:21. DOI: 10.4230/LIPICS.SWAT.2022.10.

[3] E. M. Arkin and R. Hassin, "Approximation algorithms for the geometric covering salesman problem," *Discret. Appl. Math.*, vol. 55, no. 3, pp. 197–218, 1994. DOI: 10.1016/0166-218X(94)90008-6.

[4] S. Arora, "Polynomial time approximation schemes for euclidean traveling salesman and other geometric problems," *J. ACM*, vol. 45, no. 5, pp. 753–782, 1998. DOI: 10.1145/290179.290180. [Online]. Available: https://doi.org/10.1145/290179.290180.

[5] M. de Berg, J. Gudmundsson, M. J. Katz, C. Levcopoulos, M. H. Overmars, and A. F. van der Stappen, "TSP with neighborhoods of varying size," *J. Algorithms*, vol. 57, no. 1, pp. 22–36, 2005. DOI: 10.1016/J.JALGOR.2005.01.010.

[6] N. Christofides, "Worst-case analysis of a new heuristic for the travelling salesman problem," Graduate School of Industrial Administration, Carnegie Mellon University, Technical Report 388, 1976.

[7] P. Crescenzi and V. Kann, "Approximation on the web: A compendium of np optimization problems," in *Randomization and Approximation Techniques in Computer Science*, J. Rolim, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 111–118, ISBN: 978-3-540-69247-8.

[8] M. Dror, A. Efrat, A. Lubiw, and J. S. B. Mitchell, "Touring a sequence of polygons," in *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, L. L. Larmore and M. X. Goemans, Eds., ACM, 2003, pp. 473–482. DOI: 10.1145/780542.780612.

[9] M. Dror and J. B. Orlin, "Combinatorial optimization with explicit delineation of the ground set by a collection of subsets," *SIAM J. Discret. Math.*, vol. 21, no. 4, pp. 1019–1034, 2008. DOI: 10.1137/050636589.

[10] A. Dumitrescu and J. S. B. Mitchell, "Approximation algorithms for TSP with neighborhoods in the plane," *J. Algorithms*, vol. 48, no. 1, pp. 135–159, 2003. DOI: 10.1016/S0196-6774(03)00047-6.

[11] A. Dumitrescu and C. D. Tóth, "The traveling salesman problem for lines, balls, and planes," *ACM Trans. Algorithms*, vol. 12, no. 3, 2016, ISSN: 1549-6325. DOI: 10.1145/2850418.

[12] K. M. Elbassioni, A. V. Fishkin, and R. Sitters, "Approximation algorithms for the euclidean traveling salesman problem with discrete and continuous neighborhoods," *Int. J. Comput. Geom. Appl.*, vol. 19, no. 2, pp. 173–193, 2009. DOI: 10.1142/S0218195909002897.

[13] J. Fakcharoenphol, S. Rao, and K. Talwar, "A tight bound on approximating arbitrary metrics by tree metrics," *J. Comput. Syst. Sci.*, vol. 69, no. 3, pp. 485–497, 2004. DOI: 10.1016/J.JCSS.2004.04.011.

[14] N. Garg, G. Konjevod, and R. Ravi, "A polylogarithmic approximation algorithm for the group steiner tree problem," in *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, 25-27 January 1998, San Francisco, California, USA*, H. J. Karloff, Ed., ACM/SIAM, 1998, pp. 253–259. [Online]. Available: http://dl.acm.org/citation.cfm?id=314613.314712.

[15] E. Halperin and R. Krauthgamer, "Polylogarithmic inapproximability," in *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, L. L. Larmore and M. X. Goemans, Eds., ACM, 2003, pp. 585–594. DOI: 10.1145/780542.780628.

[16] H. Jonsson, "The traveling salesman problem for lines in the plane," *Inf. Process. Lett.*, vol. 82, no. 3, pp. 137–142, 2002. DOI: 10.1016/S0020-0190(01)00259-9.

[17] A. R. Karlin, N. Klein, and S. O. Gharan, "A (slightly) improved bound on the integrality gap of the subtour LP for TSP," in *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, IEEE, 2022, pp. 832–843. DOI: 10.1109/FOCS54457.2022.00084.

[18] J. S. B. Mitchell, "A PTAS for TSP with neighborhoods among fat regions in the plane," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, Louisiana, USA, January 7-9, 2007*, N. Bansal, K. Pruhs, and C. Stein, Eds., SIAM, 2007, pp. 11–18. [Online]. Available: http://dl.acm.org/citation.cfm?id=1283383.1283385.

[19] J. S. B. Mitchell, "A constant-factor approximation algorithm for TSP with pairwise-disjoint connected neighborhoods in the plane," in *Proceedings of the 26th ACM Symposium on Computational Geometry, Snowbird, Utah, USA, June 13-16, 2010*, D. G. Kirkpatrick and J. S. B. Mitchell, Eds., ACM, 2010, pp. 183–191. DOI: 10.1145/1810959.1810992.

[20] J. S. Mitchell, "Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric tsp, k-mst, and related problems," *SIAM Journal on computing*, vol. 28, no. 4, pp. 1298–1309, 1999.

[21] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis* (Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis). Cambridge University Press, 2017, ISBN: 9781107154889. [Online]. Available: https://books.google.ca/books?id=E9UlDwAAQBAJ.

[22] S. Safra and O. Schwartz, "On the complexity of approximating TSP with neighborhoods and related problems," in *Algorithms - ESA 2003, 11th Annual European Symposium, Budapest, Hungary, September 16-19, 2003, Proceedings*, G. D. Battista and U. Zwick, Eds., ser. Lecture Notes in Computer Science, vol. 2832, Springer, 2003, pp. 446–458. DOI: 10.1007/978-3-540-39658-1_41.

[23] A. I. Serdyukov, "O nekotorykh ekstremal'nykh obkhodakh v grafakh," *Upravlyaemye sistemy 17*, pp. 76–79, 1978. [Online]. Available: http://nas1.math.nsc.ru/aim/journals/us/us17/us17_007.pdf.

[24] V. V. Vazirani, *Approximation algorithms*. Berlin, Heidelberg: Springer-Verlag, 2001, ISBN: 3540653678.

[25] D. B. West, *Introduction to Graph Theory*, 2nd ed. Prentice Hall, 2000, ISBN: 0130144002.

[26] D. P. Williamson and D. B. Shmoys, *The Design of Approximation Algorithms*. Cambridge University Press, 2011.