

DISS. ETH No. 15179, 2003

LP-related Properties of Polytopes with Few Facets

A dissertation submitted to the
Swiss Federal Institute of Technology, ETH Zürich
for the degree of Doctor of Technical Sciences

presented by
Falk Tschirschnitz
MMath (Oxford)
born 18.10.1973, Erfurt, Germany

accepted on the recommendation of
Prof. Dr. Emo Welzl, ETH Zürich, examiner
Prof. Dr. Walter Morris, George Mason University, Fairfax, co-examiner
Dr. Bernd Gärtner, ETH Zürich, co-examiner

Abstract

Linear Programming is the problem of maximizing a linear function in d variables subject to n linear constraints. Its relevance arises from the huge number of optimization problems that can be described as linear programs. An algorithm that solves any linear program has been known almost as long as linear programming: it is the so-called Simplex Method.

This method performs extremely well in practice, but a satisfactory explanation why has eluded researchers.

In this thesis, we address the special class of linear programs where n , the number of the linear constraints, is less or equal $d + 3$. The set of feasible solutions is then a d -polyhedron with at most $d + 3$ facets.

Essentially two different approaches can be pursued: we distinguish between the *geometric* and the *combinatorial view*. While the first considers actual linear programs, the latter concentrates on determining and using purely combinatorial properties.

In the first part, we analyse the RANDOM-EDGE simplex algorithm by taking the geometric view. The crucial prerequisite for this is that we can apply the extended Gale transform that maps a linear program with d variables and $d + k$ constraints to a k -dimensional configuration of one line and $d + k$ points. The main results are a tight upper bound for the case of d variables and $d + 2$ constraints, and a lower bound if the number of constraints is $d + 3$.

In the second part, we take the combinatorial view and study the orientations of the vertex-edge graph of d -polytopes with $d + 2$ facets that are defined by Holt-Klee functions. (The orientations induced by Holt-Klee functions satisfy all currently known conditions necessary to be induced by a linear function; we show that these conditions are not sufficient.)

Again, we analyse the RANDOM-EDGE simplex algorithm, this time viewing it as a random walk on oriented graphs that satisfy the Holt-Klee axioms.

Finally, we prove that each such orientation determines a partial chirotope that is completable. For general partial chirotopes we show that to decide their completness is NP-complete.

Zusammenfassung

Der Simplex-Algorithmus zur Lösung linearer Programme ist ein klassisches Beispiel für einen Algorithmus, der zwar in der Praxis sehr effizient ist, sich einer theoretischen Erklärung seines Verhaltens bisher jedoch hartnäckig verschließt.

In dieser Arbeit untersuchen wir spezielle lineare Programme, bei denen die Anzahl der linearen Nebenbedingungen nur wenig mehr als die Anzahl der Variablen beträgt. Die Menge zulässiger Lösungen solcher linearen Programme definiert ein Polytop mit wenigen Facetten.

Unser Hauptergebnis ist die Analyse des Simplex-Algorithmus unter der Verwendung der Pivot-Regel RANDOM EDGE auf solch speziellen linearen Programmen. Hierbei folgen wir zwei unterschiedlichen Ansätzen, einem geometrischen und einem kombinatorischen. Während wir bei ersterem konkrete lineare Programme betrachten, abstrahieren wir bei letzterem und beschränken uns auf rein kombinatorische Eigenschaften.

Der geometrische Ansatz wird durch die Anwendung der erweiterten Gale-Transformation möglich gemacht. Diese bildet ein lineares Programm mit wenigen Nebenbedingungen auf eine niedrig-dimensionale Konfiguration ab, bestehend aus einer Geraden und n Punkten. Für solche Konfigurationen können wir einen Algorithmus definieren und analysieren, der das Verhalten von RANDOM EDGE auf dem ursprünglichen linearen Programm simuliert.

Der zweite Teil der Arbeit ist dem kombinatorischen Ansatz gewidmet. Hier konzentrieren wir uns auf lineare Programme, deren zulässige Lösungsmenge einem d -Polytop mit $d+2$ Facetten entspricht. Die orientierten Ecken-Kanten-Graphen dieser Polytope haben spezielle kombinatorische Eigenschaften: Sie gehören zur Klasse der ‘zulässigen Gitterorientierungen’. Für solche linearen Programme können wir RANDOM EDGE direkt analysieren: als Random Walk auf einer zulässigen Gitterorientierung.

Den Schlusspunkt dieser Arbeit bilden Betrachtungen zur Erweiterbarkeit partieller Chirotope. Wir beweisen, dass jede zulässige Gitterorientierung ein erweiterbares partielles Chiropot definiert. Wir können jedoch auch zeigen, dass das Problem NP-vollständig ist, für ein beliebiges partielles Chiropot zu entscheiden, ob es erweiterbar ist.

Acknowledgements

I would like to thank all the people without whom this thesis would not have been written.

Special thanks to:

Emo Welzl – it was a privilege to be part of his research group and to profit from his insights.

Bernd Gärtner who helped me shape my final results and for his open-door policy.

Walter Morris for those useful email conversations and for pointing out little known related work.

Jürgen Richter-Gebert, Jed Mihalisin and József Solymosi for stimulating discussions.

All the people from our group for a very collegial working atmosphere, in particular: Udo Adamy, Christoph Ambühl, AleXX Below, Jochen Giesen, Michael Hoffmann, Matthias John, Ingo Schurr, Tibor Szabo, Uli Wagner.

My friends in Zürich and elsewhere who kept in touch even if I didn't.

Ramy, Rita, Barbara and Peter – thanks for sharing a roof with me!

And finally my family, especially my wife for always being near me – even when she wasn't.

Contents

1	Introduction	1
1.1	The Combinatorial View	4
1.2	The Geometric View	5
1.3	Our Results and Outline of this Thesis	6
2	The Basic Concepts	9
2.1	Linear Programming & RANDOM EDGE	10
2.2	One Line and n Points	12
2.3	The Extended Gale Transform	15
2.4	Dantzig's Column Geometry	22
2.5	Admissible Grid Orientations	24
2.6	Partial Chirotopes	28
3	Analysing the <i>Fast Process</i>	31
3.1	The Trivial Case and a Recurrence with Influence	31
3.2	The Upper Bound for Dimension 2	33
3.3	The Slow Process	40
3.4	The Lower Bound for Dimension 2	42
3.4.1	The Instance that prepares us for the Third Dimension	43
3.4.2	An Instance with an Extraordinary Property	48
3.5	A Lower Bound for Dimension 3	52
3.5.1	Placing some Points	52
3.5.2	Finding the Place for the Remaining Points	55

3.5.3	The Analysis	59
4	Admissible Grid Orientations	67
4.1	Notation and Terminology	67
4.2	The Path Condition	69
4.3	A Random Walk on Admissible Grid Orientations	74
4.3.1	The Refined Index	75
4.3.2	Basic Properties	76
4.3.3	The Analysis	79
4.4	Pseudo Realizability of Admissible Grid Orientations	84
4.5	What about ‘Proper’ Realizability?	95
5	On the Completeness of Partial Chirotopes	101
5.1	A Variant of 3-SAT	103
5.2	The Frame of Reference	103
5.3	Adding Pseudolines	104
5.4	The Construction	107
Bibliography		113
Curriculum Vitae		119

Chapter 1

Introduction

Even though convex polytopes are surely among the oldest subjects of mathematical research, it is only since the middle of the last century that they re-established themselves as being at the heart of many fundamental mathematical problems. Their renaissance is closely (but, of course, not exclusively) connected to the advent of Linear Programming (LP).

Generally speaking, linear programming is the problem to maximize a linear function in d variables subject to n linear constraints. Systems of linear equalities have been studied as early as in the middle of the 19th century. But even so, the birth of linear programming as a mathematical discipline is widely considered to be in 1947, when G. B. Dantzig was the first to develop an algorithm to solve linear programs efficiently, the *Simplex method* [Dan63]. To be precise, we refer to a whole family of algorithms, each of which is characterized by its *pivot rule*.

We may look at linear programming as a geometric problem: under certain assumptions (which we might make) the solution space of LP defines a polytope; the objective function induces an orientation on the vertex-edge graph of this polytope; the Simplex method follows a path on this oriented graph, by choosing among all neighbours of the current vertex one that improves the objective function value by a given rule, the *pivot rule*.

Two other algorithms for linear programming that have been invented deserve mention: Khachiyan [Kha80] applied the *ellipsoid method* to linear programming and proved that it always converges in polynomial time. Four years later, the fact that LP belongs to the complexity class **P** could be re-established by

Karmarkar [Kar84] who introduced an *interior point method* with that property. Unlike Khachiyan's algorithm, it also performed well in practice.

In spite of these efforts, the Simplex method remains the most widely used algorithm. Several approaches have been followed to find a satisfactory theoretical explanation for its excellent performance.

Borgwardt studied a deterministic pivot rule, SHADOW VERTEX, and showed that on polytopes whose constraints are drawn independently from spherically symmetric distributions (e.g. Gaussian distribution centered at the origin), the Simplex algorithm runs in expected polynomial time [Bor87]. However, such randomly generated matrices have, in fact, quite special properties.

This was the motivation for Spielman and Teng [ST01] to introduce the *smoothed analysis*. Here, the coefficients of an arbitrary polytope are perturbed by adding independently chosen Gaussian random variables. If the instances that cause the bad worst case behaviour of the Simplex method are isolated, then perturbing the coefficients will suffice to speed up the performance. And, indeed, (also) using the SHADOW VERTEX pivot rule, they proved that the Simplex Method has polynomial *smoothed complexity*.

However, in this thesis, we take the ‘classic’ point of view, confining randomness to the pivot rule.

A captivating approach is very closely linked to the Hirsch conjecture. It states that any two vertices of a d -polytope with n facets are connected by a path of length at most $n - d$. Fritzsche, Holt and Klee [HK98a], [FH99] constructed ‘many’ polytopes which meet this bound. On the other hand, the best known upper bound is $n^{\log d + 2}$ and is due to Kalai and Kleitman [KK92]. Of course, the existence of a short path alone would tell us nothing about whether the Simplex method would find it.

As the Simplex method is a discrete method — unlike the ellipsoid and the interior point method — the natural measure of complexity for the Simplex Method is the number of pivots taken (while using a certain pivot rule). More formally, one wants to determine the combinatorial complexity in the unit cost (RAM) model, where one assumes that all arithmetic operations incur unit cost.

For a selection of possible pivot-rules see, for instance, [AZ99], we make do with just two:

GREATEST INCREASE: Choose the vertex v' that gives the greatest increase in the objective function, that is, such that $c^T v' - c^T v$ is maximal.

RANDOM EDGE: Among all neighbouring vertices v' of v for which $c^T v' > c^T v$ choose one uniformly at random.

An algorithm based on a deterministic pivot rule (like GREATEST INCREASE) will, when repeatedly started at a given vertex, always give us the same sequence of vertices leading us to the optimum. This property was used to construct for almost all known deterministic rules specific LPs for which the number of steps taken by the Simplex method is exponential in the dimension or in the number of constraints.

The research in this field was initiated by Klee and Minty [KM72] — who used their famous construction, the so-called “Klee-Minty cubes”, to prove that the pivot rule originally proposed by Dantzig [Dan63] leads to an exponential number of steps. In the subsequent years, similar constructions were developed for all sorts of deterministic pivot rules. These efforts finally culminated in the work by Amenta and Ziegler [AZ99]. They realized, that all previous constructions were just special cases of *deformed products* of polytopes.

On the other hand, RANDOM EDGE is a classic example for the class of *randomized* pivot rules, that have become fashionable only recently. Using coin flips to decide the next move, such algorithms cannot so easily be fooled into taking a long detour.

However, despite its virtual simplicity, RANDOM EDGE is notoriously difficult to analyse. A more complicated pivot rule that works in a recursive fashion has turned out to be more accessible: RANDOM FACET.

By analysing RANDOM FACET, a major breakthrough was achieved, independently, by Kalai [Kal92] and by Matoušek, Sharir and Welzl [MSW92] as both could prove subexponential bounds, that is, the logarithm of the expected number of pivots is sublinear. These bounds are still the best known when no further assumptions on the linear programs are taken.

Interestingly, this research focused on the combinatorial properties of linear programs, defining generalizations of linear programming to more abstract settings. Kalai introduced *abstract objective function*, Matoušek, Sharir and Welzl *LP-type problems*. Later it turned out that both approaches were essentially dual to each other.

In his thesis [Gär95], Gärtner studied *abstract optimization problems* which are a further generalization of LP-type problems. He realized that even for this more general setting, subexponential bounds can be proven.

The discussion, so far, has already indicated that it makes sense to distinguish between two points of view. On the one hand, one hopes to understand the Simplex method on actual polytopes. This is what we call the geometric view. On the other hand, we have seen that substantial progress could be made by

abstracting from linear programs to more general settings – on which one can still run the Simplex method. We say that the research falling in this category is taking the combinatorial view.

1.1 The Combinatorial View

Kalai proved the subexponential bound for RANDOM FACET mentioned above for what he called *abstract objective functions* (AOF):

Definition 1.1 Consider an edge orientation of the vertex-edge graph of the polytope \mathcal{P} with vertex set \mathcal{V} with the following properties:

1. Every nonempty face F of \mathcal{P} has a unique sink.
2. The orientation is acyclic.

Furthermore, consider a function $\phi : \mathcal{V} \rightarrow \mathbb{R}$ with the property that $\phi(v) > \phi(v')$ whenever the edge $\{v, v'\}$ is oriented from v to v' . Then we call ϕ an abstract objective function on \mathcal{P} .

Most interestingly, under this general framework, the subexponential analysis is tight. Matoušek [Mat94] described a class of AOFs on which Kalai's algorithm RANDOM FACET is subexponentially slow. Restricted to the LP-instances of the class however, the algorithm becomes polynomial, as proven by Gärtner [Gär02]. What makes the algorithm on LP-instances provably faster?

This question is closely connected to another question: given an AOF on a particular polytope, can one decide whether there exists a linear function inducing the AOF?

Holt and Klee [HK98a] showed that abstract objective functions are missing a crucial combinatorial feature implied by linear functions: whenever the graph of a simple d -polytope is oriented by a generic linear function, there will be d vertex-disjoint directed paths from the unique source to the unique sink. Imposing this path condition in addition to the axioms of abstract objective functions, we arrive at the class of *Holt-Klee* functions.

In the 3-dimensional case, Holt-Klee functions are exactly the linearly irreducible functions, as proven by Mihalisin and Klee [MK00]. In fact, they gave a full characterization of 3-polytopal digraphs: a digraph G is the vertex-edge graph of some 3-polytope oriented by means of a linear function if and

only if it has a unique sink, there are 3 vertex-disjoint paths from source to sink, and the underlying graph is simple, planar and 3-connected.

In general, though, linearly inducible functions form a proper subset of Holt-Klee functions, as first shown by Gärtner et al. [GST⁺01], [GT03]. Their function is defined on a 7-dimensional polytope with 9 facets. But already Holt-Klee functions on 4-dimensional polytopes are not necessarily linearly inducible, as Morris showed by defining such a function on a 4-cube [Mor02a]. His approach via *P-cubes* is arguably the most promising candidate to lead to further conditions, necessary for a Holt-Klee function to be linearly inducible. And finally, an asymptotic result is due to Develin [Dev02]: as the dimension d grows, the percentage of realizable d -cubes among those satisfying the axioms of abstract objective functions and the path condition tends to 0.

A further generalization are the *unique sink orientations* (USOs) [SW01], we get them by dropping condition 2 in the definition of the AOFs. They have mostly been studied on cubes. As these orientations can have cycles, they may fool even randomized pivot rules: Morris analysed a class of cubes, on which the expected number of pivot steps of RANDOM EDGE is bounded below by $\frac{d-1}{2}!$, [Mor02b]. Nonetheless, USOs are important, as they arise when studying *linear complementary problems* and some *quadratic optimization problems*.

1.2 The Geometric View

Because of the exponential worst case behaviour of deterministic pivot rules, the hopes lie on randomized pivot rules. They have received an increasing amount of attention ([BDF⁺95], [Kal97], [MR95, Section 9.10]), but their analysis turned out to be intrinsically difficult. Only very recently could some progress be made. Initially, the focus was put on the special class of Klee-Minty cubes, the prime example of linear programs for which the performance of deterministic pivot rules is exponential. As Gärtner et al. [GHZ98] showed, on Klee-Minty cubes the pivot rules RANDOM EDGE and RANDOM FACET are both essentially quadratic in the dimension. Joswig and Kaibel [JK99] proposed and analysed two new rules, RECURSIVE RANDOM EDGE (which also is quadratic in the dimension) and RANDOM MAJORITY (which turned out to be optimal, as it finds the shortest path).

For fixed dimension, Megiddo [Meg84] demonstrated that the linear programming problem can be solved in linear time. Nonetheless, in the light of the remarks above, the analysis of RANDOM EDGE in dimension 3 by Kaibel et al. [KMSZ02] was, therefore, well worth the considerable effort.

1.3 Our Results and Outline of this Thesis

In this thesis, we focus exclusively on LP-related properties of polytopes with few facets.

Our geometric point of view relies crucially on the *extended Gale transform* as introduced by Welzl [Wel01]. Let us first recall the standard Gale transform. (For a proper exposition of the Gale transform (that was developed by Perles after ideas by Gale [Gal56]) see Grünbaum [Grü67].) It maps a sequence of the $d + k$ vertices of some d -polytope to a dual sequence of $d + k$ points in dimension \mathbb{R}^{k-1} . The properties of the original polytope are preserved and can be read off from the dual. This allowed the detailed study of all d -polytopes with only few vertices as they could be described in terms of lower dimensional point configurations, [Grü67, Chapter 6] and [Zie94, Section 6.5].

In comparison, this is the *extended Gale transform* in a nutshell: it maps a d -polytope with $d + k$ facets which is given together with a linear function $f : x \mapsto c^T x$ to a k -dimensional configuration of one line (ℓ , say) and $d + k$ points. We call a $(k - 1)$ -simplex spanned by k of these points and intersected by the line an ℓ -stabbed simplex. Then there is a 1-1 correspondence between the vertices of \mathcal{P} and the ℓ -stabbed simplices. Furthermore, the order on the vertices of \mathcal{P} induced by f is equal to the order in which the line ℓ intersects the ℓ -stabbed simplices.

In fact, we can model the behaviour of the algorithm RANDOM EDGE on d polytopes with $d + k$ vertices by a randomized process involving one line and $d + k$ points in dimension k .

Chapter 2 is an introductory chapter, introducing and explaining in detail all the concepts used later. In particular, we would like to point the reader to **Section 2.3** where we give a detailed account of the extended Gale transform.

Throughout **Chapter 3** we then use this method. First, we prove a tight bound of $\Theta(\log^2 d)$ for the number of pivot steps needed by RANDOM EDGE on a d -polytope with $d + 2$ facets (**Section 3.2**). This is an exponential improvement over the previously best bound.

Remarkably, we prove the lower bound on an instance for which there exists a sequence of pivots visiting all vertices. This is **Section 3.4**.

Also, in **Section 3.5** we examine a lower bound construction for the case of $k = 3$. Here, RANDOM EDGE admits the lower bound $\Omega(\log^3 d)$.

The combinatorial point of view is taken in **Chapter 4**, where we study the graphs of simple d polytopes with $d + 2$ facets. For orientations satisfying the Holt-Klee axioms, we proof a tight bound of $O(\log^2 d)$ for the Simplex method under the pivot rule RANDOM EDGE, where the analysis crucially depends on the path condition (**Section 4.3**).

One may ask how close Holt-Klee functions are to linear functions. At least for the case of simple $(d, d + 2)$ -polytopes, we give a complete answer in **Section 4.4**. On the one hand, we show that not all Holt-Klee functions are linear, on the other hand we prove that every Holt-Klee function induces a *partial chirotope* of rank 3 which is completable. (Note that we get back to partial chirotopes in Chapter 5.) This means that the combinatorial essence of linear objective functions captured by the Holt-Klee functions is at par with the combinatorial essence of point configurations captured by chirotopes.

In **Section 4.2**, we exhibit another striking property of Holt-Klee functions which abstract objective functions do not share: consider the graph of a simple $(d, d + 2)$ -polytope, oriented by means of a Holt-Klee function. It is what we call an *admissible grid orientation*. Whenever two vertices are connected by a directed path, they are connected by a directed path of length at most three. In contrast, for abstract objective functions, there is no constant bound on the path length.

The thesis is rounded off with a chapter on *partial chirotopes*: in **Chapter 5** we provide proof that, in general, it is NP-complete to test whether a partial chirotope is completable.

Acknowledgments: As mentioned above, the extended Gale transform is due to Emo Welzl. Section 3.2 appeared first in [GST⁺01]. Some of the other results are joint work with Bernd Gärtner. Section 3.5 was inspired by discussions with Bernd Gärtner and József Solymosi.

To avoid the conflict of conscience as far as the spelling is concerned the author put himself under the authority of the *Oxford English Dictionary*^{1,2}.

¹<http://www.oed.com>.

²That is, we use the British spelling and suffix -ize.

Chapter 2

The Basic Concepts

This chapter comprises the definitions and concepts fundamental to our subsequent discussion.

We first recall the well-known linear programming problem (Section 2.1) and introduce special configurations that we usually refer to as ‘One line and n points’ (Section 2.2). There, we also specify an algorithm, the so-called *Fast Process*, that models the behaviour of the Simplex method using the pivot rule RANDOM EDGE on certain linear programs. The magic behind this is the extended Gale transform, defined in Section 2.3. The fact that each configuration of one line and n points actually describes a linear program whose feasible solution space is a polytope is derived in Section 2.4.

In the final two sections we lay the foundations for the combinatorial view. In Section 2.5 we introduce the class of admissible grid orientations. Essentially, they comprise all orientations that satisfy the known necessary conditions to be isomorphic to the vertex-edge digraph of a d -polytope with $d + 2$ facets whose orientation is induced by a linear function.

These necessary conditions are not sufficient. To prove this and to study admissible grid orientations in more detail, we need the concept of *partial chirotopes* which we define in Section 2.6.

2.1 Linear Programming & RANDOM EDGE

A linear program in d variables and n constraints can be written as

$$\begin{aligned} (\text{LP}) \quad & \text{maximize} && c^T x \\ & \text{subject to} && Ax \leq b. \end{aligned} \tag{2.1}$$

where

$$\begin{aligned} x &= (x_1, \dots, x_d)^T, \\ c &= (c_1, \dots, c_d)^T, & c, x \in \mathbb{R}^d, \\ b &= (b_1, \dots, b_n)^T, & b \in \mathbb{R}^n, \\ A &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nd} \end{pmatrix}, & A \in \mathbb{R}^{n,d}. \end{aligned}$$

Examine the set of the so-called feasible solutions of the LP,

$$\{x \in \mathbb{R}^d \mid Ax \leq b\}.$$

We make the following assumptions:

1. The set of feasible solutions is bounded and non-empty,

so it can be considered as a polytope \mathcal{P} that is defined as the intersection of $n \geq d + 1$ halfspaces ([Zie94]),

$$\mathcal{P} = \bigcap_j \{x \in \mathbb{R}^d \mid \sum_{i=1}^d a_{ji}x_i \leq b_j\}.$$

2. The affine span of \mathcal{P} is the whole space \mathbb{R}^d .

3. \mathcal{P} is simple.

Each polytope given as the bounded intersection of halfspaces can also be described as the convex hull of a finite point set ([Zie94]).

Observe that the maximum of the (linear) function $f_c : x \mapsto c^T x$ is achieved on a face of \mathcal{P} . In particular, there is a vertex $v \in \mathcal{P}$ such that $c^T v$ is maximal. This gives us an idea how to solve this problem: starting from an arbitrary

vertex v , we apply some clever rule that takes us to a neighbouring vertex v' with $c^T v' > c^T v$. Repeating the process, we will eventually reach a vertex achieving the optimum. The decisive rule is called *pivot rule*. Moreover, all algorithms proceeding this way are collectively referred to as the *Simplex Method*.

We distinguish between *deterministic* and *randomized* pivot rules. The latter may use the outcome of a random experiment, like a coin flip, to determine the vertex we pivot to.

In this thesis, we focus our attention on the prime example for randomized pivot rules, **RANDOM EDGE**. No other is as easy to describe: among all neighbouring vertices v' of v for which $f_c(v') > f_c(v)$ choose one uniformly at random. See Algorithm 2.1 for a formal description.¹

With every pivot we move to a vertex with bigger objective function value. Hence, we do not visit a vertex twice and the algorithm terminates in a finite number of steps.

Algorithm 2.1 RANDOM EDGE $\{\mathcal{P}, f_c, v\}$

```

1    $V \leftarrow \{v' \in N(v) \mid f_c(v') > f_c(v)\};$ 
2   while  $V \neq \emptyset$ 
3   do
4        $v \leftarrow_{\text{random}} V;$ 
5        $V \leftarrow \{v' \in N(v) \mid f_c(v') > f_c(v)\};$ 
6   return  $v$ .

```

Before we move on to the next section, we first take a closer look at some properties which the polytope \mathcal{P} (that is, the matrix A and the vector b defining it) either has or can be assumed to have.

Lemma 2.2 *If $\{x \mid Ax \leq b\}$ is bounded and feasible, then there exists a strictly positive n -vector \tilde{y} such that $\tilde{y}^T A = 0$.*

Proof Consider the linear program

$$\begin{aligned}
 (\text{LP}') \quad & \text{maximize} \quad \mu_1 + \cdots + \mu_n \\
 & \text{subject to} \quad Ax \leq b - \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \\
 & \quad \mu_i \geq 0, \quad i = 1, \dots, n.
 \end{aligned} \tag{2.2}$$

¹ $N(v)$ denotes the set of neighbouring vertices of v . “ $v' \leftarrow_{\text{random}} V$;” means v' is chosen uniformly at random from the set V .

Writing the linear constraints as $Ax + I\mu \leq b$ (where I is the identity matrix), it is easy to see that the dual problem is of the form

$$\begin{aligned} (\text{LP}'^\Delta) \quad & \text{minimize} && b^T y \\ & \text{subject to} && A^T y = 0, \\ & && y_i \geq 1, \quad i = 1, \dots, n. \end{aligned} \tag{2.3}$$

From the fact that $\{Ax \leq b\}$ is bounded, it follows that (LP') is bounded; because it is also feasible (set $\mu_i = 0, i = 1, \dots, n$), it has an optimal solution.

Then, by the LP duality theorem (*cf.* [Chv83, Chapter 9]), the dual (LP'^Δ) has an optimal solution, too, in particular a feasible solution. Any such feasible solution \tilde{y} is a vector with the required properties. \square

Assumption 2.3 *For A, b in (2.1) such that $\{x \mid Ax \leq b\}$ is simple, bounded and contains an interior point, we can assume the following without loss of generality (which for our purposes means, without changing the behaviour of RANDOM EDGE or RANDOM FACET on (2.1)):*

(i) *Any column of A sums up to 0, i.e. $\sum_{i=1}^n a_{ij} = 0$ for all j .*

This is achieved via Lemma 2.2, by a suitable scaling of the constraints in (2.1) by positive multiples.²

(ii) *Any set of d rows of A is linearly independent.*

For this, we apply a slight perturbation to A .

(iii) *$b > 0$.*

This is obtained by translating \mathcal{P} in such a way that 0 is an interior point.

2.2 One Line and n Points

Let S be a set of $n \geq d + 1$ points in general position in \mathbb{R}^d (i.e. no $d + 1$ on a common hyperplane), and let ℓ be a vertical line which is disjoint from S and from all intersections of hyperplanes spanned by points of S . Moreover, we assume that ℓ intersects the convex hull of S . We refer to such a configuration in general position whenever we use the phrase *One line and n points*. In short form, we write (S, ℓ) .

The convex hull of a i -tuple $s \in \binom{S}{i}$ defines a $(i - 1)$ -simplex; in a slight abuse of notation we will denote it also by s . If $i = d$, below(s) denotes the

²Lemma 2.2 even specifies suitable scaling factors: we may multiply the i th constraint by \tilde{g}_i .

set of points from S that lie below the hyperplane spanned by s . As we are mostly concerned with the d - or $(d - 1)$ -simplices intersected by ℓ , we give them a special name and call them ℓ -stabbed simplices.

Our goal is to find the unique ℓ -stabbed $(d - 1)$ -simplex that has no points below.

Each ℓ -stabbed d -simplex has two facets which are in turn ℓ -stabbed simplices but of dimension $d - 1$. They differ by exactly one vertex. Given an ℓ -stabbed $(d - 1)$ -simplex s and some point $s' \in \text{below}(s)$ there is, therefore, a unique point $s \in s$ such that $s \setminus \{s\} \cup \{s'\} =: s'$ is the facet of the d -simplex $s \cup \{s'\}$ through which ℓ is leaving. We say that we may *pivot* from s to s' and write $s' := \text{pivot}(s, s')$. Obviously, the point of intersection of s' and ℓ is below s . Pivoting from one ℓ -stabbed $(d - 1)$ -simplex to the next we will, eventually, reach our goal.

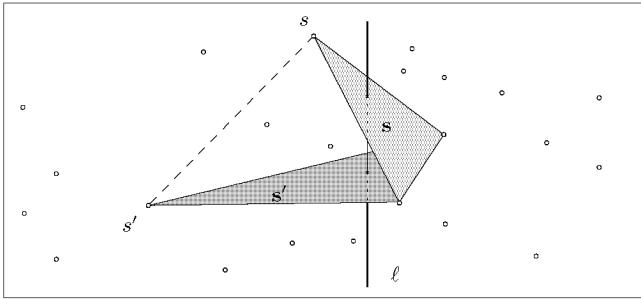


Figure 2.1: s and s' are ℓ -stabbed 2-simplices with $s' = \text{pivot}(s, s')$.

The method invoked to decide which point below the current simplex is chosen is called the *pivot rule*. (It is no coincidence that we use the same terminology as above when we were discussing the Simplex method. We will find this justified in Section 2.3 below.)

The pivot rule we will (exclusively) deal with is arguably also the one that is easiest to describe: among all points below the current ℓ -stabbed simplex s choose one uniformly at random, s' say, and move to the unique ℓ -stabbed simplex $s' := \text{pivot}(s, s')$.

For reasons which become clear later, we call the associated randomized process *Fast Process*. (For a formal definition see Algorithm 2.4). Given some initial ℓ -stabbed simplex $s = \{s_1, \dots, s_d\}$ it finds the unique ℓ -stabbed simplex \tilde{s} with $\text{below}(\tilde{s}) = \emptyset$.

Algorithm 2.4 *Fast Process on $\{S, \ell, s\}$*

```

1 while below( $s$ )  $\neq \emptyset$ 
2 do  $s' \leftarrow_{\text{random}} \text{below}(s)$ ;
3      $s \leftarrow \text{pivot}(s, s')$ ;
4 return  $s$ .

```

Frequently, we will not specify the dimension of an ℓ -stabbed simplex, whenever it is clear from the context.

We will be particularly interested in the 2-dimensional case. Here, the location of the points with respect to ℓ partitions S naturally into two subsets: the set S_L of all the points left of ℓ , and $S_R := S \setminus S_L$. Accordingly, let $L = |S_L|$ be the number of points left of ℓ and $R := n - L$. To indicate that a particular point lies left (right, respectively) of ℓ we denote it by p (q , respectively) instead of s . An ℓ -stabbed 2-simplex is then a pair of points $\{p, q\} = e \in \binom{S}{2}$; to simplify matters we will call it ℓ -edge. Given an ℓ -edge e and a point $s' \in \text{below}(e)$, $\text{pivot}(e, s')$ denotes, as before, the unique ℓ -edge $\{s_e, s'\}$, $s_e \in e$ (see Figure 2.2(a)).

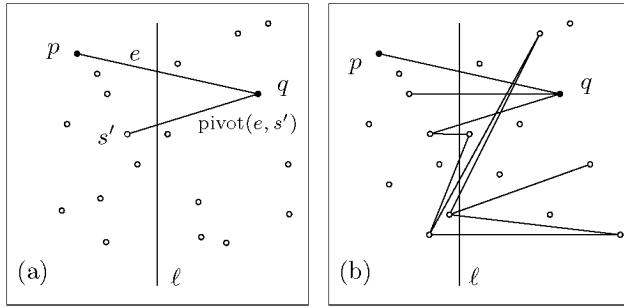


Figure 2.2: The setup (a) and a pivoting sequence (b) in \mathbb{R}^2 .

The 2-dimensional variant of Algorithm 2.4 is Algorithm 2.5:

Algorithm 2.5 *Fast Process on $\{S, \ell, e\}$ in \mathbb{R}^2*

```

1 while below( $e$ )  $\neq \emptyset$ 
2 do  $s' \leftarrow_{\text{random}} \text{below}(e)$ ;
3      $e \leftarrow \text{pivot}(e, s')$ ;
4 return  $e$ .

```

The significance of these processes becomes apparent once we have established the link between *Linear Programming* and *One line and n points*. This is the topic of the next section, Section 2.3.

There, we derive a correspondence under which for each vertex v of \mathcal{P} there is a corresponding ℓ -stabbed simplex $s(v)$. Even more, the expected runtime (number of pivot steps) of RANDOM EDGE, starting at some vertex \tilde{v} , will equal the expected length (number of pivot steps) of the *Fast Process*, starting from the ℓ -stabbed simplex $s(\tilde{v})$.

In this context it becomes clear why we will be mainly interested in computing (or estimating) the expected number of **while**-loops the *Fast Process* enters. If the goal would be to implement these algorithms, we would need to specify how to sample from $\text{below}(s)$. At least for the 2-dimensional case we can (and will) address this issue, *cf.* Section 3.3, where we study a modification of the *Fast Process*, called the *Slow Process*. Here, each round chooses a point among *all points* and only performs a pivot step if the chosen point is below the current edge.

Algorithm 2.6 *Slow Process on $\{S, \ell, e\}$ in \mathbb{R}^2*

```

1  while  $\text{below}(e) \neq \emptyset$ 
2  do  $s' \leftarrow_{\text{random}} S$ ;
3    if  $s' \in \text{below}(e)$ 
4      then  $e \leftarrow \text{pivot}(e, s')$ ;
5  return  $e$ .

```

2.3 The Extended Gale Transform

We want to show that there is a one-to-one correspondence between the vertices of a simple d -polytope with $d + k$ facets and the ℓ -stabbed $(k - 1)$ -simplices of a suitable configuration of one line and $n := d + k$ points in \mathbb{R}^k , with the properties that

1. the order of the vertices (according to a given generic linear function³) matches the order of the corresponding ℓ -stabbed simplices along ℓ .
2. any pair of adjacent vertices corresponds to a pair of ℓ -stabbed simplices that share $d - 1$ points.

³We call a linear function generic if it is not constant on edges.

The *extended Gale transform*, as defined in Definition 2.7 below, establishes such a correspondence. Theorem 2.8 tells us that it always exists; its constructive proof also points out a method to compute it.

Under this correspondence, a RANDOM EDGE pivot step on the polytope becomes a pivot step in the fast process applied to the point configuration corresponding to the polytope, and vice versa. Consequently, the expected runtime (number of pivot steps) of RANDOM EDGE, starting at some vertex v , equals the expected length of the fast process, starting from the ℓ -stabbed simplex corresponding to v .

As before, cf. Section 2.1, we suppose that a simple $(d, d+k)$ -polytope \mathcal{P} is given by $\mathcal{P} = \{x \in \mathbb{R}^d \mid Ax \leq b\}$, where $A \in \mathbb{R}^{d+k, d}$ and $b \in \mathbb{R}^{d+k}$. And we have a generic linear function $f(x) = c^T x$, $c, x \in \mathbb{R}^d$, that induces an orientation on the vertex-edge graph of \mathcal{P} .

On the other hand, let (S, ℓ) be a k -dimensional configuration of one line and n points as introduced in Section 2.2, with $n = d+k$.

To define the transform we need some terminology.

For an index set $I \subset [d+k]$, let $\bar{I} := [d+k] \setminus I$. Given a matrix M with $d+k$ rows, we write M_I to denote the submatrix consisting of all rows of M with indices in I . Analogously, given an $d+k$ -dimensional vector m , m_I denotes the subvector consisting of all entries of m with indices in I . Finally, given a set $S = \{s_1, \dots, s_{d+k}\}$, let $\Pi_I := \{s_i \mid i \in I\}$.

Definition 2.7 We call (S, ℓ) the *extended Gale transform of (\mathcal{P}, f)* (and (\mathcal{P}, f) the *extended Gale transform of (S, ℓ)*) whenever, for any index set I of size d , the following two statements are equivalent:

- (i) $A_I^{-1}b_I =: \tilde{v} \in \mathbb{R}^d$ is a vertex of \mathcal{P} with objective function value γ .⁴
- (ii) $\Pi_{\bar{I}}$ determines an ℓ -stabbed simplex that intersects the line ℓ (given as $\ell = \{ts_\ell \mid t \in \mathbb{R}\}$ for some point s_ℓ) in a unique point, at value $\tilde{t} = \gamma / \sum_i b_i$.⁵

Theorem 2.8 For any simple $(d, d+k)$ -polytope \mathcal{P} and a generic linear function $f : \mathbb{R}^d \mapsto \mathbb{R}$ there exists a k -dimensional configuration of one line ℓ and $d+k$ points S , such that (S, ℓ) is the extended Gale transform of (\mathcal{P}, f) .

Providing us with the means to construct the extended Gale transform of a given configuration, the lemma below lies at the heart of Theorem 2.8, whose proof will follow subsequently.

⁴By Assumption 2.3.(ii) A_I is of full rank and hence invertible.

⁵By Assumption 2.3.(iii) $\sum_i b_i > 0$ and \tilde{t} is well-defined.

Lemma 2.9 Consider the $(n+1) \times (d+1)$ -matrix

$$H := \begin{pmatrix} A & -b \\ c^T & 0 \end{pmatrix}.$$

There exists an $(n-d) \times (n+1)$ -matrix B of full row rank with $BH = 0$, such that the columns of B are in general position when interpreted as points in \mathbb{R}^{n-d} . (We will specify our general position requirements below.)

Proof The columns of H are vectors in \mathbb{R}^{n+1} , spanning a subspace of dimension at most $d+1$. The orthogonal dual of this subspace has therefore dimension at least $n-d$. Choose $n-d$ linearly independent vectors in the orthogonal dual to obtain the rows of B . The prior perturbation of A (we may even perturb H) also lets us choose B in such a way that it assumes any desired general position. Lemma 2.9 \square

Proof Theorem 2.8 For a configuration (\mathcal{P}, f) the lemma above gives us a configuration of one line and n points when we interpret the i th column of the matrix B as the point s_i in \mathbb{R}^{n-d} . The set S is then given as the set of points $\{s_1, \dots, s_n\}$, and ℓ is the line spanned by s_{n+1} , i.e. $\ell := \{ts_{n+1} \mid t \in \mathbb{R}\}$.

Note, that the matrix B and hence (S, ℓ) are determined uniquely up to linear isomorphisms. (We shall discuss this in more detail as Observation 2.10 below.)

To establish the theorem we need to show that for (\mathcal{P}, f) and (S, ℓ) as just described the conditions (i) and (ii) in Definition 2.7 are indeed equivalent.

Let $\tilde{v} := A_I^{-1}b_I \in \mathbb{R}^d$. Using Lemma 2.9, we can argue that

$$\begin{aligned} 0 &= BH \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} \\ &= \underbrace{B^I(A_I| - b_I)}_0 \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} + B^{\bar{I}}(A_{\bar{I}}| - b_{\bar{I}}) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} \\ &\quad + B^{\{n+1\}}(c^T| 0) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} \\ &= B^{\bar{I}}(A_{\bar{I}}| - b_{\bar{I}}) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} + c^T \tilde{v} s_{n+1} \end{aligned}$$

and, therefore,

$$c^T \tilde{v} s_{n+1} = -B^{\bar{I}}(A_{\bar{I}}| - b_{\bar{I}}) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix}. \quad (2.4)$$

With

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} := -(A| - b) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} \quad (2.5)$$

Equation (2.4) becomes

$$c^T \tilde{v} s_{n+1} = \sum_{i \in \bar{I}} \mu_i s_i. \quad (2.6)$$

(Note that $\mu_i = 0$ for $i \in I$ by definition of \tilde{v} .)

Using Assumptions 2.3(i) and 2.3(iii) we derive

$$\begin{aligned} \sum_{i \in \bar{I}} \mu_i &= \sum_{i=1}^n \mu_i \stackrel{(2.5)}{=} (-1, \dots, -1)(A| - b) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} \\ &\stackrel{(i)}{=} (\mathbf{0} | \sum_j b_j) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} \\ &= \sum_j b_j \stackrel{(iii)}{>} 0. \end{aligned}$$

So, we may use $1/\sum_j b_j$ as a scaling factor in Equation 2.6. Setting $\lambda_i := \mu_i / \sum_j b_j$, we obtain

$$\frac{c^T \tilde{v}}{\sum_i b_i} s_{n+1} = \sum_{i \in \bar{I}} \lambda_i s_i, \quad \sum_{i \in \bar{I}} \lambda_i = 1. \quad (2.7)$$

If statement (i) of the theorem holds, i.e. \tilde{v} is a vertex with $c^T \tilde{v} = \gamma$, then

$$(A_{\bar{I}}| - b_{\bar{I}}) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} < \mathbf{0}$$

by simplicity of \mathcal{P} , which implies $\mu_i > 0$ for $i \in \bar{I}$ in (2.5), (2.6) and $\lambda_i > 0$ in (2.7). Hence, the point

$$\frac{c^T \tilde{v}}{\sum_i b_i} s_{n+1} = \frac{\gamma}{\sum_i b_i} s_{n+1}$$

is a convex combination of the points in $\Pi_{\bar{I}}$. Statement (ii) follows when we assume that $s_{n+1} \neq 0$ and that the line $\{ts_{n+1}\}$ is disjoint from all affine spaces spanned by less than $n - d$ of the points in $\Pi_{[n]}$. In fact, these are requirements on the ‘general position’ in Lemma 2.9 and on configurations of one line and n points, as introduced in Section 2.2.

Now assume that statement (ii) holds and that the set $\Pi_{\bar{I}}$ is affinely independent (our final general position requirement). In this case, there are unique values $\lambda_i > 0$ such that

$$\frac{\gamma}{\sum_i b_i} s_{n+1} = \sum_{i \in \bar{I}} \lambda_i s_i, \quad \sum_{i \in \bar{I}} \lambda_i = 1, \text{ for some value } \gamma.$$

Because the line spanned by s_{n+1} intersects $\text{aff}\Pi_{\bar{I}}$ in a single point, there is no other value of γ for which the previous equation can be satisfied. Therefore, there are unique values γ, μ_i with

$$\gamma s_{n+1} = \sum_{i \in \bar{I}} \mu_i s_i, \quad \sum_{i \in \bar{I}} \mu_i = \sum_i b_i. \quad (2.8)$$

On the other hand, by the above computations, the values

$$\begin{aligned} \gamma &:= c^T \tilde{v}, \\ \mu_i &:= -(A_i| - b_i) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} \end{aligned}$$

satisfy equation (2.8), so they are the desired unique values. Because $\{ts_{n+1}\}$ intersects $\text{conv}\Pi_{\bar{I}}$, it follows that

$$(A_{\bar{I}}| - b_{\bar{I}}) \begin{pmatrix} \tilde{v} \\ 1 \end{pmatrix} < 0,$$

and \tilde{v} is a vertex with $c^T \tilde{v} = \gamma$, proving (i).

Theorem 2.8 \square

The reader who is familiar with the Gale transform⁶ will have noticed that instead of invoking Lemma 2.9, we could have applied the well-known Gale transform (*cf.* [Mat02], [Wel01]) to the vertices of the polar dual of the polytope \mathcal{P} adjoined by the vector c defining the objective function.

We will not elaborate on this further, but rather make a couple of important observations:

⁶For an excellent introduction we recommend Matoušek [Mat02, Section 5].

Observation 2.10 *The extended Gale transform is determined up to linear isomorphism.*

The rows of the matrix B in Lemma 2.9 are an arbitrary basis of the orthogonal dual to the vector space spanned by the column vectors of H in \mathbb{R}^{n+1} . Choosing a different basis corresponds to multiplying the matrix B from the left by a non-singular $(n-d) \times (n-d)$ -matrix, and this means transforming configuration (S, ℓ) by a linear transformation of \mathbb{R}^{n-d} .

In particular, this means that we can always find a representative of the extended Gale transform with ℓ being the vertical line through the origin.⁷

Observation 2.11 *Each planar configuration (S, ℓ) defines a directed graph $G(S, \ell)$: each ℓ -edge defines a node in the graph, and we have a directed edge from s to s' if one can pivot directly from s to s' .*

Consider (S, ℓ) as the extended Gale transform of some (\mathcal{P}, f) . Then $G(S, \ell)$ is isomorphic to the vertex-edge graph of \mathcal{P} oriented by means of the linear function f .

This follows directly from the properties of the extended Gale transform and comes as no surprise. Its usefulness will become clear as soon as we have introduced admissible grid orientations. But before we move on to them, an example shall demonstrate the extended Gale transform ‘in action’.

Example

Let \mathcal{P} be the 3-polytope given as the set of feasible solutions of the LP

$$\begin{array}{lllll} \text{maximize} & 5x_1 & + & 6x_2 & + & 4x_3 \\ \text{subject to} & 3x_1 & & & & \leq 3 \\ & 3x_2 & & & & \leq 3 \\ & & x_3 & & & \leq 1 \\ & -2x_1 & - & x_2 & + & x_3 \leq 1 \\ & -x_1 & - & 2x_2 & - & 2x_3 \leq 2 \end{array}$$

\mathcal{P} is a 3-dimensional polytope with 5 facets, a prism over a triangle. Figure 2.3 shows this configuration (\mathcal{P}, f) where the orientation of the 1-skeleton as induced by the linear function is visualized by arrows along the edges, pointing to the vertex with higher objective value.

⁷In \mathbb{R}^d , we call a line *vertical*, whenever it contains the unique vector e_d .

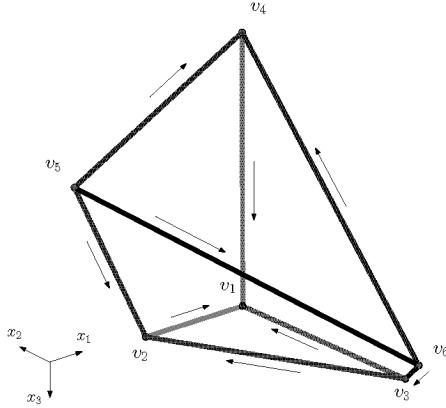


Figure 2.3: A configuration (\mathcal{P}, f) ...

Having formed the matrix H ,

$$H = \begin{pmatrix} A & -b \\ c^T & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 & -3 \\ 0 & 3 & 0 & -3 \\ 0 & 0 & 1 & -1 \\ -2 & -1 & 1 & -1 \\ -1 & -2 & -2 & -2 \\ 5 & 6 & 4 & 0 \end{pmatrix},$$

we may compute the extended Gale dual of (\mathcal{P}, f) : a possible choice for the matrix B would be

$$B = \begin{pmatrix} 1.1 & 0.1 & -3.9 & 2.1 & -0.9 & 0 \\ \frac{5}{6} & 0.5 & 2.5 & -1.5 & -1.5 & -1 \end{pmatrix}$$

giving us the configuration (S, ℓ) with the points

$$\begin{aligned} s_1 &= \begin{pmatrix} 1.1 \\ \frac{5}{6} \end{pmatrix}, & s_2 &= \begin{pmatrix} 0.1 \\ 0.5 \end{pmatrix}, & s_3 &= \begin{pmatrix} -3.9 \\ 2.5 \end{pmatrix}, & s_4 &= \begin{pmatrix} 2.1 \\ -1.5 \end{pmatrix}, \\ s_5 &= \begin{pmatrix} -0.9 \\ -1.5 \end{pmatrix} & \text{and} & s_6 &= \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \end{aligned}$$

ℓ just being the y -axis (since $s_6 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$), see Figure 2.3.

Each point in (S, ℓ) represents a facet of the triangular prism \mathcal{P} . The ℓ -edge connecting two points $s_i, s_j \in S$ represents the unique vertex v that is in neither of the two facets represented by s_i and s_j . Here is a list of the relations.

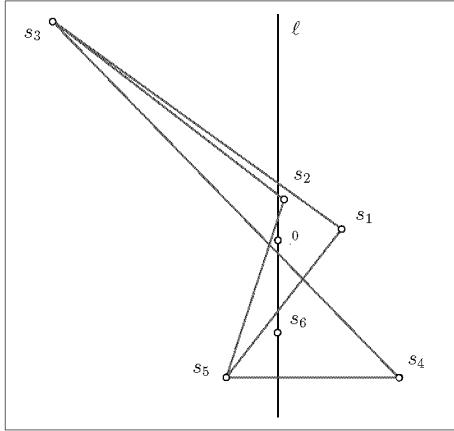


Figure 2.4: ... and its extended Gale dual (S, ℓ) .

Facets of \mathcal{P}	Points in S	Vertices of \mathcal{P}	ℓ -edges in (S, ℓ)
$v_1v_3v_6v_4$	s_1	v_1	s_4s_5
$v_1v_4v_5v_2$	s_2	v_2	s_1s_5
$v_1v_2v_3$	s_3	v_3	s_2s_5
$v_2v_3v_6v_5$	s_4	v_4	s_4s_3
$v_4v_5v_6$	s_5	v_5	s_1s_3
		v_6	s_2s_3

Using this example, we can now demonstrate that the extended Gale transform has the wanted properties we formulated at the beginning of this section.

Consider, for instance, the vertices v_2 and v_6 . The corresponding ℓ -edges are s_1s_5 and s_2s_3 . In the oriented vertex-edge graph of \mathcal{P} , there is a path from v_2 to v_6 (via v_3) — in the extended Gale transform (S, ℓ) the line ℓ intersects the ℓ -edge s_1s_5 above its point of intersection with s_2s_3 . Furthermore, we can pivot from s_1s_5 to s_2s_3 via s_2s_5 which is the ℓ -edge corresponding to v_3 .

2.4 Dantzig's Column Geometry

We have seen that each bounded linear program on d variables and $d + k$ constraints that satisfies certain non-degeneracy conditions can be transformed into a k -dimensional configuration of one line and $n := d + k$ points, such that the latter is the extended Gale transform of the linear program. Our proof

already strongly indicates that the converse is equally true, that is, for each configuration (S, ℓ) there is a polytope \mathcal{P} and a linear function f such that (\mathcal{P}, f) is the extended Gale transform of (S, ℓ) .

For our purposes, it shall suffice to point out that we can describe each configuration of *One line and n points* as a linear program. Most interestingly, this was already observed by the ‘father’ of Linear Programming Theory, George B. Dantzig ([Dan63]). But let us start from the beginning.

Equation (2.1) is not the only ‘standard’ form for linear programs; Dantzig, for instance, preferred to use the following version⁸:

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b, \\ & && x_i \geq 0, \quad i = 1, \dots, n, \\ &\text{where} && c, x \in \mathbb{R}^n, b \in \mathbb{R}^k \text{ and } A \in \mathbb{R}^{k,n} \end{aligned} \tag{2.9}$$

We may assume that

$$(A|b) = \begin{pmatrix} 1 & \dots & 1 & 1 \\ A' & & & b' \end{pmatrix} \text{ for some } A' \in \mathbb{R}^{k-1,n}, b' \in \mathbb{R}^{k-1}.$$

Now, consider the matrix $\bar{A} = \begin{pmatrix} A' \\ c \end{pmatrix}$, $\bar{A} \in \mathbb{R}^{k,n}$. Each of the n columns of \bar{A} defines a point $s_i = (a'_{i1}, a'_{i2}, \dots, a'_{i,k-1}, c_i)$, $i = 1, \dots, n$, in \mathbb{R}^k . Furthermore, let ℓ be the so-called *requirement line*, that is, the line given as $\{\binom{b'}{\mu} \mid \mu \in \mathbb{R}\}$. This configuration is sometimes referred to as the *column geometry*, an interesting application can be found in [Lee97].

To solve the LP (2.9) means therefore, in geometric terms, to find the lowest point on the requirement line that can be given as convex combination of the points s_1, \dots, s_n — which is exactly the objective of the *Fast* and the *Slow Process* (Algorithms 2.4 and 2.6).

In fact, the *Fast Process* is the geometric visualization of RANDOM EDGE on a k -dimensional linear program as given in (2.9) since the basic feasible solutions of the linear program correspond to the ℓ -stabbed $(k-1)$ -simplices in the configuration of one line and n points.

This geometric interpretation of the Simplex method was actually the motivation by Dantzig to give it that name.

We do not want to go into further details, but there is time for one small observation: the pivot rule originally suggested by Dantzig chooses the point below

⁸It is not difficult to see that each linear program given as in (2.1) can be written in the form (2.9) and vice versa.

the current ℓ -stabbed simplex s which has the maximal vertical distance from the hyperplane s .

Finally, we would like to point out that there is yet another way to describe configurations of *One line and n points* as a linear program. Let the coordinates of the n points be given as $s_i = (x_1^{(i)}, x_2^{(i)}, \dots, x_k^{(i)}), i = 1, \dots, n$, let the vertical line by given as $\{(r_1, r_2, \dots, r_{k-1}, \lambda) \mid \lambda \in \mathbb{R}\}$. Then the aim is to find values $\tilde{\mu}_j, j = 1, \dots, k$ such that the hyperplane $x_k = \sum_{j=1}^{k-1} \tilde{\mu}_j x_j + \tilde{\mu}_k$ carries the lowest ℓ -stabbed simplex:

$$\begin{aligned} & \text{maximize} && \mu_k + \sum_{j=1}^{k-1} \mu_j r_j \\ & \text{subject to} && x_k^{(i)} \geq \sum_{j=1}^{k-1} \mu_j x_j^{(i)} + \mu_k, \quad i = 1, \dots, n. \end{aligned} \quad (2.10)$$

But note that this LP is just the dual of LP (2.9)!

2.5 Admissible Grid Orientations

All d -polytopes with $d+1$ facets are combinatorially equivalent to the standard d -simplex Δ_d , defined as

$$\Delta_d = \{x \in \mathbb{R}^{d+1} \mid x_1 + x_2 + \dots + x_{d+1} = 1, x_i \geq 0\}.$$

Having just one extra facet compared to the simplices of the same dimension, the $(d, d+2)$ -polytopes still have a simple, though non-trivial, structure:

Lemma 2.12 *The d -polytopes with $d+2$ facets are combinatorially equivalent to the products of simplices $\mathcal{P} = \Delta_{L-1} \times \Delta_{R-1}$, where $L+R = d+2$, $L, R > 1$.*

Proof We use a well-known theorem by Grünbaum [Grü67, Result 5.1.1]: *Every d -polytope with $f \geq d+1$ facets is the intersection of an $(f-1)$ -simplex with some flat.* Applying it, we only need to show that if the intersection of a $(d+1)$ -simplex with a hyperplane has $d+2$ facets, then it is combinatorially equivalent to the product of two simplices.

Consider the standard $d+1$ -simplex,

$$\Delta_{d+1} = \{x \in \mathbb{R}^{d+2} \mid x_1 + x_2 + \dots + x_{d+2} = 1, x_i \geq 0\}.$$

Let h be the intersecting hyperplane. As we only consider simple polytopes, we may assume that none of the vertices of Δ_{d+1} lies in h . So, combinatorially, h is given by the sets of vertices of Δ_{d+1} on its positive (and its negative)

side, their cardinality shall be denoted by L (and R). Clearly, $L + R = d + 2$. The combinatorial type of the intersection of Δ_{d+1} with h only depends on the numbers L and R . So, we can write h (w.l.o.g.) as

$$h = \{x \in \mathbb{R}^{d+2} \mid x_1 + x_2 + \cdots + x_L - x_{L+1} - x_{L+2} - \cdots - x_{L+R} = 0\}.$$

This implies that

$$\Delta_{d+1} \cap h = \{x \in \mathbb{R}^{d+2} \mid x_1 + \cdots + x_L = \frac{1}{2} = x_{L+1} + \cdots + x_R, x_i \geq 0\}$$

— which is just $\Delta_{L-1} \times \Delta_{R-1}$. \square

As an immediate corollary, we can deduce that the vertex-edge-graph of a $(d, d + 2)$ -polytope is what we will call a *grid graph* with L rows and R columns, see Figure 2.5(a). The vertices that belong to some subset of rows and some subset of columns (plus the connecting edges) define a *subgrid*. Each subgrid corresponds to a non-empty face of the polytope. (Again, we refer to Lemma 2.12.)

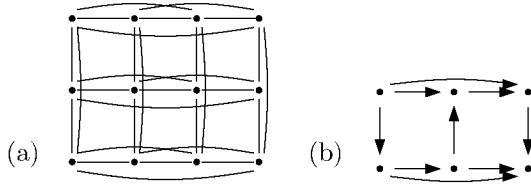


Figure 2.5: (a) A grid graph with 3 rows and 4 columns and (b) the forbidden subgraph.

Any generic linear function induces an orientation on the vertex-edge-graph of the polytope and hence an orientation on the underlying grid graph.

We observe that if such a grid orientation is induced, it has the following three properties, the first two being obvious:

1. The orientation is acyclic.
2. Every nonempty subgrid has a unique sink.
3. No subgrid is isomorphic to the ‘forbidden subgrid’ depicted in Figure 2.5(b).

We call a grid orientation with these properties *admissible*.

In our specific scenario, Properties 1 and 2 boil down to the axioms of orientations induced by *abstract objective functions* (AOF)([Kal92], [Wil88]). They imply that every nonempty subgrid also has a unique source you find the proof at the end of this section.

Lemma 2.13 *Consider a grid orientation. If every nonempty subgrid has a unique sink, then the orientation also has a unique source.*

As proven in Lemma 4.1, Property 3 can be obtained by observing that this condition is equivalent to the so-called *path condition* found by Holt and Klee [HK98b]: Source and sink of the 1-skeleton of a d -polytope oriented by a generic linear function can be connected by d vertex-disjoint monotone paths. This condition is necessary for an orientation to be induced by a linear objective function. We call all AOF that induce orientations satisfying Holt and Klee's path condition *Holt-Klee functions*.

It is a very interesting fact that, for *unique sink orientations* of $(d, d + 2)$ -polytopes i.e. orientations satisfying Property 2 the path condition is sufficient for acyclicity. In other words, Properties 2 and 3, in the definition of admissible grid orientations, already imply Property 1. This is Theorem 4.2. For $(d, d + 3)$ -polytopes, this statement is not true anymore. Consider, for instance, the 3-dimensional cube orientation depicted in Figure 2.6. Obviously, with respect to d , this is also the smallest possible example.

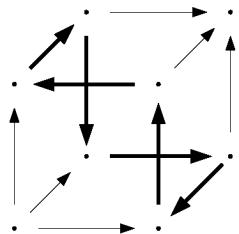


Figure 2.6: A cyclic orientation of a $(3, 6)$ -polytope (a 3-cube) that satisfies the path condition.

A natural question is whether any admissible grid orientation is actually induced by some $(d, d + 2)$ -polytope together with a linear function. (We call these orientations *realizable*.)

As we show in Section 4.5, this is not the case. This implies that Holt and Klee's path condition is not sufficient for an AOF to be induced by a linear objective function.

For the proof we will recall Observation 2.11: If an admissible grid orientation is induced by some configuration (\mathcal{P}, f) then it is also the digraph of its extended Gale transform (S, ℓ) . In fact, we will consider admissible grid orientations as partial chirotopes and look at the question of realizability by examining the completness of these partial chirotopes. These notions are formally introduced in Section 2.6 below.

We finish this current section by giving the proof for what we have claimed above: every acyclic grid orientation has a unique source whenever every nonempty subgrid has a unique sink.

Proof of Lemma 2.13 We follow Kalai [Kal97]: Suppose we are given a grid orientation $G = G_{L \times R}$. The underlying grid graph can be viewed as the 1-skeleton of the $(d, d+2)$ -polytope \mathcal{P} with $d = L + R - 2$, by Lemma 2.12. Impose the given orientation onto the 1-skeleton. Define h_k to be the number of vertices of \mathcal{P} with exactly k incoming edges. Since we have a unique sink orientation, we already know $h_d = 1$. In fact, what we claim is $h_0 = 1$.

Consider some vertex v with r incoming edges. The number of k -faces with v as their unique sink is $\binom{r}{k}$, as any k edges incident to v span a k -face. If f_k denotes the number of k -faces of \mathcal{P} , we, therefore, get the equation:

$$f_k = \sum_{r=0}^d h_r \binom{r}{k}, \quad k = 0, 1, \dots, d.$$

This equation can be inverted (we omit the proof) to

$$h_k = \sum_{r=0}^d (-1)^{r-k} f_r \binom{r}{k}, \quad k = 0, 1, \dots, d.$$

Therefore, the h -numbers h_k are in fact linear combinations of the face numbers f_k , in particular, they do not depend on the specific grid orientation. So, since any orientation induced by some generic linear function will have a unique source, *all* unique sink grid orientations satisfy the lemma. \square

As we may apply this lemma to any nonempty subgrid in its own right, we may immediately deduce that, in fact, every nonempty subgrid has a unique source in a unique sink orientation.

To determine the h -vector depending on L and R , we choose the simplest configuration (S, ℓ) we can think of: Let ℓ be the y -axis, set $p_i = (-1, i)$, $i = 1, \dots, L$ and $q_j = (1, j)$, $j = 1, \dots, R$. Then it is not hard to see that

$$h = (1, 2, 3, \dots, \underbrace{\min(L, R), \dots, \min(L, R)}_{|L-R|+1}, \dots, 3, 2, 1).$$

	ℓ
.	.
.	.
.	.
.	.

2.6 Partial Chirotopes

The final section of this chapter we use exclusively to introduce the notions of chirotopes, partial chirotopes and completable. Their relevance in the context of this thesis will become clear especially in Section 4.4 where we will show that each admissible grid orientation defines a completable partial chiotope.

Definition 2.14 *Let E_n be the index set of a finite set of elements. A chiotope of rank 3 on E is a map $\chi : E^3 \rightarrow \{-1, 0, +1\}$ that satisfies the axioms*

1. *χ is an alternating sign map, i.e. for any permutation $\sigma \in S_3$ and $(\alpha, \beta, \gamma) \in E^3$ we have $\chi(\sigma(\alpha), \sigma(\beta), \sigma(\gamma)) = \text{sgn}(\sigma)\chi(\alpha, \beta, \gamma)$.*
2. *For pairwise different elements $\alpha, \beta, \gamma, \delta, \epsilon \in E_n$ the set $\{\chi(\alpha, \beta, \gamma) \cdot \chi(\alpha, \delta, \epsilon), -\chi(\alpha, \beta, \delta) \cdot \chi(\alpha, \gamma, \epsilon), \chi(\alpha, \beta, \epsilon) \cdot \chi(\alpha, \gamma, \delta)\}$ either equals 0 or contains $\{-1, +1\}$.
(These are the so-called Grassmann-Plücker-Relations.)*
3. *For each element α there is at least one pair $\beta, \gamma \in E_n$ with $\chi(\alpha, \beta, \gamma) \neq 0$.*

If the alternating sign map $\chi : E^3 \rightarrow \{-1, 0, +1\}$ is only partially defined, and condition 2. above holds whenever χ is defined on all six triples, then χ is a partial chiotope of rank 3 on E_n .

We call χ uniform if $0 \notin \chi(\binom{E}{3})$, i.e. $\chi(\alpha, \beta, \gamma) \neq 0$ for all pairwise different elements $\alpha, \beta, \gamma \in E_n$.

Finally, we call a partial chiotope χ' of rank 3 on a set E *completable*, if there exists a chiotope χ of rank 3 on E and for any $\alpha, \beta, \gamma \in E$, $\chi(\alpha, \beta, \gamma) = \chi'(\alpha, \beta, \gamma)$ holds whenever $\chi'(\alpha, \beta, \gamma)$ is defined. We say that χ is a *completion* of χ' .

In Chapter 5 we will discuss the following problem:

COMPLETABILITY OF PARTIAL CHIROTOTES (CPC)

Given: A partial chiotope χ' of rank 3 on a set E .

Question: Is there a chiotope χ on E such that χ is a completion of χ' ?

Chapter 3

Analysing the *Fast Process*

We analyse here the expected number of pivots taken by the Algorithms 2.4, 2.5 and 2.6 which we introduced in Section 2.2. We first take a brief look at the trivial, 1-dimensional case (Section 3.1). It will not only serve as the perfect warm-up, but also set the foundation for the higher dimensional lower bound constructions.

High emphasis is placed on the planar case. Section 3.2 is reserved for the proof of the upper bound. (Later, in Section 4.3, we will revisit this proof—but translated into a strictly combinatorial framework.) The upper bound is shown to be tight in Section 3.4, where we investigate several lower bound constructions meeting it. Embedded into these two sections, there is Section 3.3, dedicated to the interesting Algorithm 2.6, the *Slow Process*.

The final section is devoted to the 3-dimensional case. Here, we give a lower bound construction (Section 3.5).

3.1 The Trivial Case and a Recurrence with Influence

The 1-dimensional version of the *Fast Process* models the behaviour of RANDOM EDGE on n -facet polytopes in dimension $d = n - 1$, that is, on simplices. The underlying vertex-edge-graph is the complete graph with $d + 1$ vertices. All orderings on the vertices that are induced by some generic linear function are isomorphic to each other. The outdegree of a vertex tells us the position

of this vertex in the ordering: from any vertex, we can reach exactly those vertices which have smaller outdegree.

So, for this very special case, it is not difficult to compute the expected number of pivots directly. Let this number be denoted by t_k if the starting vertex v has outdegree k . Then

$$t_k = 1 + \frac{1}{k} \sum_{i=0}^{k-1} t_i, \quad k > 0, \quad t_0 = 0, \quad (3.1)$$

implying

$$t_k = \sum_{i=1}^k \frac{1}{i} = H_k. \quad (3.2)$$

To be able to use the extended Gale transform on a d -simplex and a linear function we need to drop one of the non-degeneracy conditions for (S, ℓ) : all $n (= d + 1)$ points have to lie *on* a line. Each point represents a vertex of the original simplex. The order in which the line passes through the points is identical to the order imposed on the vertices by the linear function.

We can view S as a set of real numbers. Then the *Fast Process* becomes Algorithm 3.1, called with an arbitrary element s of S :

Algorithm 3.1 *The Fast Process on $\{S, s\}$ in \mathbb{R}*

```

1  while  $s \neq \min S$ 
2  do
3       $s \leftarrow_{\text{random}} \{s' \in S \mid s' < s\};$ 
4  return  $s$ .
```

We will see later that the good lower bounds we can prove for certain 2- and 3-dimensional instances rely on the repeated call of 1-dimensional subprocesses. However, these differ in the sense that they are joined by ρ additional *exit*-points, located below the vertex that used to be the lowest one, see Figure 3.1.

We want to know the expected number of steps that such an extended 1-dimensional process takes until one of the ρ points is hit. Starting at the point with $k + \rho$ points below, this number shall be denoted by $t_{k,\rho}$. Naturally, the recurrence is a close relative of Equation (3.1):

$$t_{k,\rho} = 1 + \frac{1}{k + \rho} \sum_{i=0}^{k-1} t_{i,\rho}, \quad \rho > 0. \quad (3.3)$$

It is not hard to see that this implies

$$\begin{aligned}
 t_{k,\rho} &= \frac{1}{k+\rho} + \frac{1}{k-1+\rho} + \cdots + \underbrace{\frac{1}{\rho+1}}_1 + t_{0,\rho} \\
 &= \sum_{i=\rho+1}^{k+\rho} \frac{1}{i} + 1 \\
 &= H_{k+\rho} - H_\rho + 1, \quad \rho > 0
 \end{aligned} \tag{3.4}$$

Defining $t_{k,0} := t_k$, we will not need to consider the case $\rho = 0$ separately.

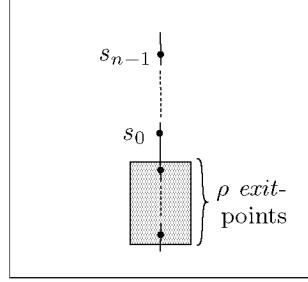


Figure 3.1: Augmented one-dimensional process.

3.2 The Upper Bound for Dimension 2

The 1-dimensional scenario revisited. To get the spirit of the later analysis we have yet another look at the 1-dimensional *Fast Process* as given with Algorithm 3.1. This time, we only want to provide a rough estimate. Let X_i , $i \in \mathbb{N}_0$, be the random variable for the number of iterations of the while-loop with

$$2^i \leq \#\{s \in S \mid s < \tilde{s}\} < 2^{i+1} \tag{3.5}$$

for \tilde{s} the value of s at the beginning of the respective iteration of the while-loop. For $k \geq 2$, the random variable $Z = \sum_{i=0}^{\lfloor \log_2(k-1) \rfloor} X_i$ gives the overall number of executions of the while-loop. $E(X_i) \leq 2$ for all $i \in \mathbb{N}_0$, since, whenever (3.5) holds, we have a chance of at least $\frac{1}{2}$ to choose an element of

rank 2^i or smaller.

Hence,

$$E(Z) \leq 2(1 + \lfloor \log_2(k-1) \rfloor) = O(\log k).$$

The obvious extension of that analysis to the 2-dimensional process fails, since the number of points below edges appearing in the process oscillates. In fact, the number of points below the current edge is no measure of progress at all. This number may be 1, we pivot, and the number becomes as large as $n - 3$ (see Figure 3.2).

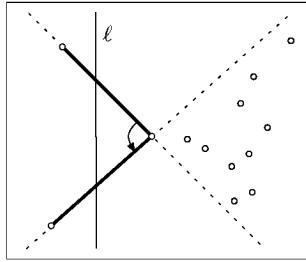


Figure 3.2: We thought we were so close!

One word to the notation used in this Section. We follow Section 2.2 with one exception: p and q still denote points on opposite sides of ℓ but not necessarily with p on the left.

k -Lines as Milestones. Here is the crucial definition that will allow us to measure progress. Given $k \in \mathbb{N}_0$, a non-vertical line λ is called a k -line of S and ℓ if on both sides of ℓ there are exactly k points from S below λ . It is easy to see that every point $x \in \ell$ is contained in a k -line for some $k \in \mathbb{N}_0$, as long as x is disjoint from all segments connecting two points in S . Start with a line through x that has large slope so that all points on the right side of ℓ are below, and all on the left side are above. Now rotate the line by decreasing its slope. Eventually, we will reach the situation opposite to what we started with: no points below to the right, all below to the left. All transitions in between change the number of points below on exactly one side by ± 1 . Somewhere in between we must have had a transition where the numbers of points below were the same on both sides.

A k -line disjoint from S exists for all k , $0 \leq k \leq m := \min\{L, R\}$. This is not difficult to see: each point x on ℓ lies on several k -lines but $k = k(x)$ is

unique. Traversing ℓ , the current k changes by 1 whenever we cross an ℓ -edge. (Here we need our assumption of general position.)

For each i , $0 \leq i \leq \lfloor \log_2 m \rfloor$, fix some 2^i -line λ_i disjoint from S . Line λ_0 has to be chosen, so that the only edge intersecting ℓ below λ_0 is the edge \hat{e} with $\text{below}(\hat{e}) = \emptyset$. Moreover, let $\lambda_{\lfloor \log_2 m \rfloor + 1}$ be some m -line that intersects ℓ above all ℓ -edges (and above¹ $\lambda_{\lfloor \log_2 m \rfloor}$). The line λ_i intersects ℓ below λ_j for $0 \leq i < j \leq \lfloor \log_2 m \rfloor + 1$.

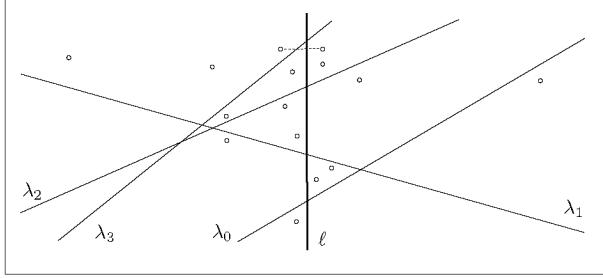


Figure 3.3: Setting milestones.

We define the random variable X_i , $i = 0, 1, \dots, \lfloor \log_2 m \rfloor$, as the number of executions of the while-loop (in Algorithm 2.5) where the current ℓ -edge e intersects ℓ below λ_{i+1} but not below λ_i . The sequence of these executions we call *phase i of the process*². The careful choice of λ_0 ensures that completion of phase 0 entails completion of the whole process. Hence, $Z = \sum_{i=0}^{\lfloor \log_2 m \rfloor} X_i$ is the random variable whose expectation we want to analyse.

We will show that $E(X_i) = O(\log n)$ for all i and, hence,

$$E(Z) = O((\log n)(1 + \log m)) = O(\log^2 n).$$

Analysis of a Single Phase. Fix some i , $0 \leq i \leq \lfloor \log_2 m \rfloor$, set $k = 2^i$, $\lambda' = \lambda_i$ and $\lambda = \lambda_{i+1}$. So λ' is a k -line, and there are at most $4k$ points below λ (actually exactly, unless $i = \lfloor \log_2 m \rfloor$). We have an edge intersecting ℓ not below λ' but below λ , and the phase starts. The phase ends whenever we reach an edge that intersects ℓ below λ' . Note that for every edge occurring in the phase, one endpoint has to be below λ (since the edge intersects ℓ below λ) and there is an endpoint above λ' (since otherwise, we are already in a new phase).

¹This is automatically satisfied, unless m is a power of 2.

²Note that phases count *down* during the process.

A few words on what we are heading for. We further split phase i into *strokes*. A stroke starts after we have sampled a point in $\text{below}(\lambda) \cup \text{below}(\lambda')$ (or at the very beginning of the phase) and it finishes after another point in $\text{below}(\lambda) \cup \text{below}(\lambda')$ is chosen (this includes the event that the phase ends); thus, any stroke in the phase terminates with a point in $\text{below}(\lambda) \cup \text{below}(\lambda')$. If N is the number of strokes, then we can write $X := X_i$ as

$$X = Y_1 + Y_2 + \cdots + Y_N$$

where Y_j is the number of iterations of the j th stroke. Note that N itself is a random variable. (For $j > N$ we set $Y_j = 0$.)

We will show that

- (i) $E(Y_j | j \leq N) = O(\log n)$ for all j , and
- (ii) $E(N) = O(1)$.

It follows that $E(X) = O(\log n)$:

$$\begin{aligned} E(X) &= \sum_{j=1}^{\infty} \overbrace{E(Y_j | j \leq N)}^{O(\log n)} \Pr(j \leq N) \\ &= O(\log n) \sum_{j=1}^{\infty} \Pr(j \leq N) \\ &= O(\log n) E(N). \end{aligned} \tag{3.6}$$

As for the points sampled from $\text{below}(\lambda) \cup \text{below}(\lambda')$ we distinguish several cases depending on where the respective new point pivoted into the current edge lies. We will see that each of these situations is more or less promising in our goal to escape this phase.

Here are the steps in our reasoning: *at any time during the phase*, the following four claims hold.

Claim 1 *The expected number of pivots until we sample a new point in $\text{below}(\lambda)$ is at most $2 \log_2 n$.*

Proof At least one of the two endpoints of the current edge has to be below λ . So in a contiguous subsequence where the new point is always chosen above

λ , the other endpoint below stays the same throughout this sequence. We denote this point by q . If we order the points on the other side of ℓ according to their visibility from q , we get almost the situation as in the one-dimensional process described as Algorithm 3.1. In fact, there are two differences which can only improve our expectations: we terminate not only in the lowest point but also in $k - 1$ other points. In each step we may also sample on q 's side of ℓ , in which case we immediately terminate (we have surely sampled below λ). Hence, the expected length of such a subsequence is at most³ $2 \log_2 n$. \square

Since any new point sampled in $\text{below}(\lambda)$ starts a new stroke, this also establishes our claim (i) from above: the expected number of iterations during a stroke is $O(\log n)$.

Claim 2 *Conditioned on the event that we sample a point in*

$$\text{below}(\lambda) \cup \text{below}(\lambda') ,$$

the point will be in

$$\text{below}(\lambda')$$

with probability at least $\frac{1}{5}$.

Proof Since all edges in this phase intersect ℓ not below λ' it follows: for one side of ℓ , all k points below λ' must also lie below the line through the current edge. That is, *at least* k points below λ' are also below the line through the current edge. On the other hand, *at most* $5k$ points are below λ or λ' . This holds, since $\#\text{below}(\lambda) \leq 4k$, $\#\text{below}(\lambda') = 2k$, and on one side of ℓ , all k points below λ' are also below λ . \square

Claims 1 and 2 combined assure that we reach a point below λ' within an expected number of at most $10 \log_2 n$ steps.

So what happens after we see such a point p below λ' ? Two cases have to be distinguished, depending on whether p is also below λ or not.

Claim 3 *If an endpoint p of the current edge is in*

$$\text{below}(\lambda') \setminus \text{below}(\lambda) ,$$

then the next point sampled below λ or λ' will be in

$$\text{below}(\lambda) \cap \text{below}(\lambda')$$

with probability at least $\frac{1}{5}$.

Proof Two relevant conclusions right away (see Figure 3.4): (i) Since p is not below λ and below λ' , while λ' intersects ℓ below λ , the lines λ and λ'

³Even H_{n-m-1} is true.

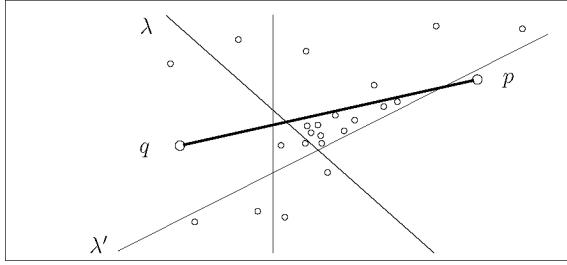


Figure 3.4: $p \in \text{below}(\lambda') \setminus \text{below}(\lambda)$.

must intersect on p 's side. (ii) Since p is not below λ , the other endpoint q of the current edge has to be below λ .

Before q can be substituted by a point not below λ , the other endpoint has to be below λ . That is, when we first sample a point below λ or λ' , point q is still in the edge. Therefore, the current edge connects q to a point below the edge $\{p, q\}$, above λ' . So the line carrying this edge must intersect λ' on p 's side. But then, on q 's side, all k points below λ' are also below the then current edge.

Moreover, on q 's side, all points below λ' are also below λ (since these lines intersect on the other side). So, summing up, the k points below λ' are both below the current edge and below λ , and they are at disposal, when we sample a point below λ or λ' . The claim follows. \square

Claim 4 *If an endpoint p of the current edge is in*

$$\text{below}(\lambda) \cap \text{below}(\lambda') ,$$

then the next point sampled below λ or λ' will be in

$$\text{below}(\lambda') \text{ on the side opposite to } p$$

with probability at least $\frac{1}{5}$.

Proof If p is substituted in a pivot, it must be substituted by a point below λ or λ' . This holds, since on p 's side of ℓ , everything below the current edge has to be below λ or λ' (see Figure 3.5). As a consequence, until the first pivot with a point below λ or λ' , point p is still an endpoint of the edge. But since p is below λ' , on the opposite side everything below λ' is also below the current edge. So there are at least k good choices, and at most $5k$ choices of points below λ or λ' . \square

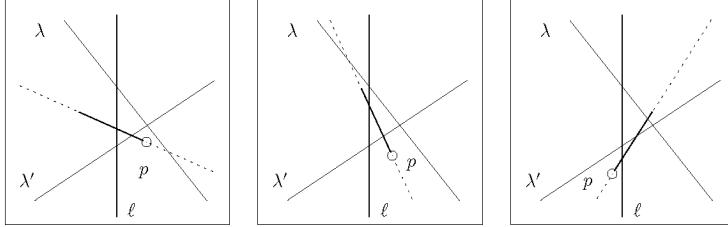


Figure 3.5: $p \in \text{below}(\lambda') \cap \text{below}(\lambda)$.

Claim 4 entails that once we have chosen a point below λ' and λ , then – with probability at least $\frac{1}{5}$ – the next point chosen below λ or λ' will terminate the phase.

To complete the argument, we look at the sequence of points from $\text{below}(\lambda) \cup \text{below}(\lambda')$ that are pivoted into the current edge. Recall that these are exactly the points that terminate the strokes of a phase (except for the last one). If we can show claim (ii) from above, i.e. that the expected length of this sequence is at most some constant c , then the expected length of the whole sequence is at most $2c \log_2 n$ due to Equation 3.6. Each point in this sequence is classified depending on whether it lies in

Class 0: $\text{below}(\lambda) \setminus \text{below}(\lambda')$

Class 1: $\text{below}(\lambda') \setminus \text{below}(\lambda)$

Class 2: $\text{below}(\lambda') \cap \text{below}(\lambda)$

Every point in the sequence considered is in Class 0, 1, or 2. If we have a point in Class 0, the next will be in Class 1 or 2 with probability at least $\frac{1}{5}$ (by Claim 2). If we have a point in Class 1, the next will be in Class 2 with probability at least $\frac{1}{5}$ (by Claim 3). (All of this of course conditioned on the event that a next point exists at all, i.e. the phase hasn't stopped already.) Finally, if we are in Class 2, it is the last point in the sequence with probability at least $\frac{1}{5}$ (by Claim 4).

Now we estimate the expected length of the sequence by the Markov chain⁴ depicted in Figure 3.6, with four states

`start = 0, 1, 2, and 3 = stop,`

and the indicated transition probabilities. On one hand, it is easy to calculate

⁴It simulates a biased coin with success probability $\frac{1}{5}$ and counts the number of experiments until we have three consecutive successes.

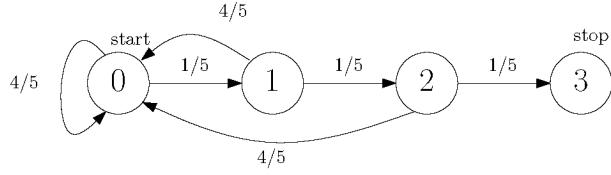


Figure 3.6: A pessimistic Markov chain.

that the expected number of steps from start to stop is 155. On the other hand, the chain and our sequence can be coupled so that whenever the chain is in state $s \in \{0, 1, 2, 3\}$, then the corresponding point in the sequence in Class $t \geq s$, or the sequence has ended already. Hence, we have shown that the expected number of pivots in a single phase is bounded by $310 \log_2 n$, and the theorem follows.

Theorem 3.2 *The expected number of pivots in the process defined as Algorithm 2.5 is at most*

$$O((\log n)(1 + \log m)) = O(\log^2 n),$$

where n is the number of points, and m is the smaller of the numbers of points on the two sides of the line.

3.3 The Slow Process

If ℓ is the y -axis, both the *Fast Process* and the *Slow Process* find the lowest ℓ -edge determined by points $s_i = (x_i, y_i)$, $i = 1, \dots, n$.

In fact, they find the solution of the following linear program (LP) in two variables μ_1, μ_2 , cf. 2.4:

$$\begin{aligned} &\text{maximize} && \mu_2 \\ &\text{subject to} && y_i \geq \mu_1 x_i + \mu_2, \quad i = 1 \dots n. \end{aligned} \tag{3.7}$$

We have already mentioned that to implement and use the *Fast Process* (Algorithm 2.4), we need to specify how to sample from below(s). If we can sample efficiently (say in time logarithmic in n), then the process gives a polynomial time algorithm even for exponential size sets. If we sample in the obvious way in $O(n)$ time, then this gives us an $O(n(\log n)^2)$ algorithm.

On that we can improve by looking at the alternative *Slow Process* (Algorithm 2.6).⁵ The number of pivots has the same distribution as the *Fast Process*. Again, we analyse the number of iterations of the while-loop. As before, we split the process into *phases* that are formed by successive pivots. Recall that in phase i , the current edge intersects ℓ not below the 2^i -line λ_i .

Claim 5 *For each i , the expected number of iterations in phase i is at most*

$$O\left(\frac{n}{2^i}\right).$$

Proof Divide phase i into strokes as previously done. That is, a stroke is ended, whenever we sample a point in

$$(\text{below}(\lambda') \cup \text{below}(\lambda)) \cap \text{below}(\tilde{e}),$$

(\tilde{e} being the current ℓ -edge), or when the phase ends.

In phase i , there are always at least 2^i points from $\text{below}(\lambda')$ that lie below the current edge (on some side of ℓ , all points below λ_i are also below the current edge). That is, at any point, we sample a point resulting in the termination of the stroke with probability at least $\frac{2^i}{n}$. Therefore, the expected number of iterations in a stroke is at most $\frac{n}{2^i}$. The number of strokes is, of course, the same as in the *Slow Process*; its expectation is constant. The claim follows. \square

Theorem 3.3 *The expected number of iterations of Algorithm 2.6 is*

$$\Theta(n)$$

where n is the number of points, unless the starting edge $\{p, q\}$ is already the lowest ℓ -edge.

Proof $\sum_{i=0}^{\lfloor \log_2 m \rfloor} \frac{n}{2^i} < 2n$ and so the upper bound follows from Claim 5. If $\{p, q\}$ is disjoint from the lowest ℓ -edge, a lower bound of $\frac{3}{2}n$ is obvious, since on the average it takes that long until we have sampled both endpoints of the lowest edge at least once. Even if $\{p, q\}$ contains exactly one of the two endpoints of the lowest ℓ -edge, we still need n steps on the average before we meet the other endpoint for the first time. The lower bound follows. \square

The coupon collector analysis (*cf.* [GS92, Exercise 3.13]) tells us that it takes $\Theta(n \log n)$ iterations until we expect to have sampled each point at least once.

⁵The ‘ $\text{below}(e) \neq \emptyset$ ’-test can be made once in n rounds only, thus causing amortized constant cost. Or, after every pivot, we can go through all points in random order (without replacement) until we find the first point in $\text{below}(e)$; if no such point is found, we are done. Compared to the ‘pure version’, this can only speed up the procedure.

Thus, the *Slow Process* finds the lowest ℓ -edge long before all points have been seen at least once.

3.4 The Lower Bound for Dimension 2

We are presenting two instances of *One line and n points*, that share the property that the expected number of pivots of the *Fast Process* on them is $\Omega(\log^2 n)$. The instance we discuss first is already based on the ideas that later motivate the 3-dimensional lower bound construction. The second instance shines with a special property: there exists a sequence of pivots which visits all possible ℓ -edges.

Before we start, let us recall a useful fact as well as prove two easy lemmata that will turn out to be useful for the actual analysis.

Lemma 3.4 (*[GKP94, Equation 6.60]*)

$$\ln n < H_n < \ln n + 1, \text{ for } n > 1, \quad (3.8)$$

Lemma 3.5

$$\sum_{i=1}^{k-1} \frac{H_i}{i+1} = \frac{1}{2} \left(H_k^2 - \sum_{i=1}^k \frac{1}{i^2} \right). \quad (3.9)$$

Proof

$$\begin{aligned} \sum_{i=1}^{k-1} \frac{H_i}{i+1} &= \sum_{i=1}^{k-1} \frac{H_{i+1}}{i+1} - \sum_{i=1}^{k-1} \frac{1}{(i+1)^2} = \sum_{i=1}^k \frac{H_i}{i} - \sum_{i=1}^k \frac{1}{i^2} \\ &= \sum_{1 \leq j \leq i \leq k} \frac{1}{i \cdot j} - \sum_{i=1}^k \frac{1}{i^2} \\ &= \frac{1}{2} \left(\left(\sum_{i=1}^k \frac{1}{i} \right)^2 + \sum_{i=1}^k \frac{1}{i^2} \right) - \sum_{i=1}^k \frac{1}{i^2} \\ &= \frac{1}{2} \left(H_k^2 - \sum_{i=1}^k \frac{1}{i^2} \right). \end{aligned} \quad \square$$

Lemma 3.6

$$\sum_{i=1}^{k-1} \frac{H_{2i}}{i+1} \leq \frac{H_{2k-1}^2}{2}. \quad (3.10)$$

Proof

$$\begin{aligned} \sum_{i=1}^{k-1} \frac{H_{2i}}{i+1} &\leq \sum_{i=1}^{k-1} \left(\frac{H_{2i}}{2i} + \frac{H_{2i+1}}{2i+1} \right) = \sum_{i=1}^{2k-1} \frac{H_i}{i} - 1 \\ &= \frac{1}{2} \left(H_{2k-1}^2 + \sum_{i=1}^{2k-1} \frac{1}{i^2} \right) - 1 \leq \frac{H_{2k-1}^2}{2}. \end{aligned} \quad \square$$

3.4.1 The Instance that prepares us for the Third Dimension

Suppose we have a configuration with n_1 points left of ℓ and only one single point, q say, on ℓ 's right hand side. Then q must be an element of all ℓ -edges. The *Fast Process* on this 2-dimensional configuration is essentially just a fancy version of the 1-dimensional *Fast Process* on n_1 points and takes, therefore, time $\Theta(\log n_1)$ (Equation (3.2)).

Suppose further that we add ρ points ($\{q_0, \dots, q_{\rho-1}\}$) to our configuration that lie below all existing ℓ -edges and on the line parallel to ℓ through $q_\rho := q$. Then we can run the 1-dimensional process augmented by j exit-points when starting with an arbitrary ℓ -edge incident to q_j .

But we can also run the 2-dimensional *Fast Process* on this configuration. By saying that we are in phase i as long as q_i is in the current edge, we can view it as a succession of 1-dimensional subprocesses with i exit-points.

Unfortunately, the size of the 1-dimensional subprocesses is getting smaller with each new phase. Only if the number of exit points is in comparison also sufficiently small (e.g. of order $O(\sqrt{m})$, where m is the number of points currently available for the 1-dimensional subprocess), will the subprocess still asymptotically take at least $\Omega(\log m)$ number of pivots by Equation (3.4).

The crucial idea is to add auxiliary points such that whenever a new phase starts the n_1 points on the left of ℓ lie – with constant probability – below the current edge. This will suffice to show that during the *Fast Process* sufficiently many of the subphases take a sufficiently high expected number of pivot steps.

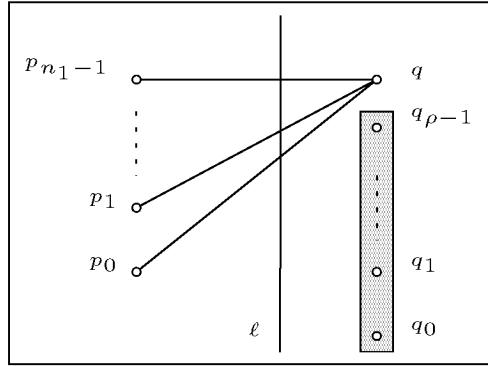


Figure 3.7: Running the 1-dimensional process with ρ exit points on a 2-dimensional configuration.

The actual configuration will have $\rho + n_1$ points on the left and $\rho + n_2$ points on the right of ℓ . (So, $n = 2\rho + n_1 + n_2$ — the exact values of n_1, n_2 and ρ shall be determined later.) It is defined as follows, see also Figure 3.8:

ℓ is the y -axis,

$$p_i = \begin{cases} \binom{-1}{i} & \text{for } i = 0, \dots, \rho - 1 \\ \binom{-2n}{i} & \text{for } i = \rho, \dots, \rho + n_1 - 1 \end{cases}$$

and

$$q_i = \begin{cases} \binom{1}{i} & \text{for } i = 0, \dots, \rho - 1 \\ \binom{2n}{i} & \text{for } i = \rho, \dots, \rho + n_2 - 1 \end{cases}.$$

Finally, to get general position, we slightly perturb the points. This construction has the following properties:

Observation 3.7

- (i) $\{p_0, \dots, p_{i-1}\} \cup \{q_0, \dots, q_{j-1}\} \subseteq \text{below}(\{p_i, q_j\})$.
- (ii) $\{p_\rho, \dots, p_{\rho+n_1-1}\} \subseteq \text{below}(\{p_i, q_j\})$ whenever $j < i < \rho$
and $\{q_\rho, \dots, q_{\rho+n_2-1}\} \cap \text{below}(\{p_i, q_j\}) = \emptyset$ whenever $j \leq i < \rho$.
- (iii) $\{q_\rho, \dots, q_{\rho+n_2-1}\} \subseteq \text{below}(\{p_i, q_j\})$ whenever $i < j < \rho$
and $\{p_\rho, \dots, p_{\rho+n_1-1}\} \cap \text{below}(\{p_i, q_j\}) = \emptyset$ whenever $i \leq j < \rho$.

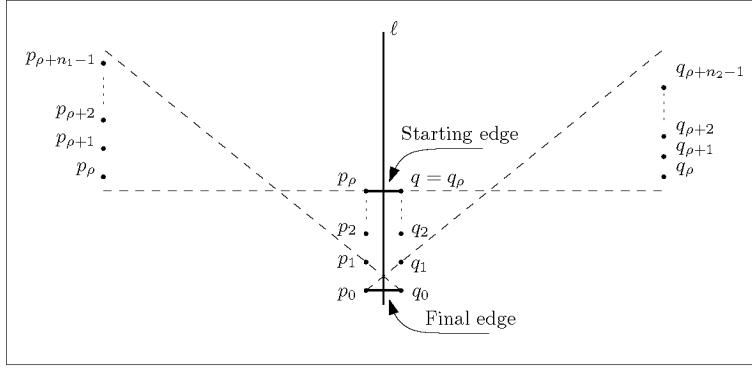


Figure 3.8: Instance for the lower bound.

For this configuration of one line and $n_1 + n_2 + 2\rho$ points we will now derive a lower bound of the *Fast Process* when started from the edge $\{p_{\rho-1}, q_{\rho-1}\}$. In a final step, we will determine those values for n_1 , n_2 and ρ that give us the best lower bound.

As we have already indicated our analysis is again based on the idea to divide the sequence of pivots into distinct phases. We say we are in phase a as long as $\min(i, j) = a$, where i, j are the indices of the points defining the current edge e . In other words, a new phase is entered whenever we choose a point whose index sets a new minimum. This definition makes sense by Observation 3.7: each phase entered is succeeded by a phase with smaller index. We say that the phase is on the left (right, respectively) of ℓ whenever its defining minimum point is on the left (right, respectively).

We define the random variables $X_i, i = 0, \dots, \rho - 1$ as the number of pivots during phase i ; Z denotes the total number of pivots under the assumption that we start with the edge $\{p_{\rho-1}, q_{\rho-1}\}$. We aim to bound the expected value of Z from below, but a direct approach $E(Z) = \sum_{i=0}^{\rho-1} E(X_i)$ seems not suitable.

Conditioned on the event of choosing a point among

$\{p_0, \dots, p_i\} \cup \{q_0, \dots, q_i\}$,
 p_i and q_i are equally likely chosen with probability $\frac{1}{2(i+1)}$. This implies that

$$\Pr(\text{Phase } i \text{ is entered}) = \begin{cases} \frac{1}{i+1} & \text{for all } i, 0 \leq i < n-1 \\ 1 & \text{for } i = n-1. \end{cases} \quad (3.11)$$

Thus, denoting the event that phase i is entered by Ph_i , we get

$$\begin{aligned} \mathbb{E}(Z) &= \sum_{i=0}^{\rho-1} \mathbb{E}(X_i) = \sum_{i=0}^{\rho-1} \mathbb{E}(X_i \mid \text{Ph}_i) \Pr(\text{Ph}_i) \\ &= \mathbb{E}(X_{\rho-1} \mid \text{Ph}_{\rho-1}) + \sum_{i=0}^{\rho-2} \frac{\mathbb{E}(X_i \mid \text{Ph}_i)}{i+1}. \end{aligned} \quad (3.12)$$

To estimate the expectations of the X_i conditioned on the event that phase i is actually entered we distinguish two cases. If phase i is on the same side as the previous phase, we have to assume the worst: possibly, it is over after just one flip! But if the minimum switched to the other side with the beginning of phase i then either the n_1 points $\{p_\rho, \dots, p_{\rho+n_1-1}\}$ or the n_2 points $\{q_\rho, \dots, q_{\rho+n_2-1}\}$ are put back into the game.

Using the same argument as for Equation (3.11), we see that with probability $\frac{1}{2}$ phase i does not lie on the same side with respect to ℓ as the previous phase.

Let r and s denote the number of points on the left and on the right of ℓ below the current edge; it is easy to see that we are in phase $\min(r, s)$. The missing piece is to compute $\tau_{r,s}$, the expected number of flips until we leave this phase. Clearly, $\tau_{r,s} = \tau_{s,r}$, so w.l.o.g. assume $s \leq r$. $\tau_{r,s}$ is monotone in r . For $s < r$ (i.e. in particular, at the beginning of a new phase), we have at least $s + n$ points below the current edge on the side of r .

The simple recursion

$$\tau_{0,0} := 0, \quad \tau_{r,s} = 1 + \frac{1}{r+s} \sum_{r'=s}^{r-1} \tau_{r',s} \quad (r+s > 0)$$

gives us

$$\tau_{r,s} = \begin{cases} H_r & \text{whenever } s = 0 \\ H_{r+s} - H_{2s} + 1 & \text{otherwise.} \end{cases} \quad (3.13)$$

Applying these observations to Equation (3.12) we can therefore deduce that

$$\begin{aligned}
E(Z) &\geq \sum_{i=0}^{\rho-2} \frac{E(X_i \mid \text{Ph}_i)}{i+1} \\
&\geq \sum_{i=0}^{\rho-2} \frac{E\left(X_i \mid \text{Ph}_i \wedge \begin{array}{c} \text{side of minimum} \\ \text{changed} \end{array}\right)}{i+1} \Pr\left(\begin{array}{c} \text{side of minimum} \\ \text{changed} \end{array}\right) \\
&\geq \sum_{i=0}^{\rho-2} \frac{\frac{1}{2}(\min_{j \geq 0}(\tau_{i+n_1+j,i}) + \min_{j \geq 0}(\tau_{i+n_2+j,i}))}{2(i+1)} \\
&= \sum_{i=0}^{\rho-2} \frac{\tau_{i+n_1,i}}{4(i+1)} + \sum_{i=0}^{\rho-2} \frac{\tau_{i+n_2,i}}{4(i+1)}
\end{aligned}$$

since we are equally likely to encounter a left or a right phase.

Let us first estimate the expected number of steps of all the left phases:

$$\begin{aligned}
\sum_{i=0}^{\rho-2} \frac{\tau_{i+n_1,i}}{4(i+1)} &\stackrel{(3.13)}{=} \frac{H_{n_1}}{4} + \sum_{i=1}^{\rho-2} \frac{H_{n_1+i} - H_{2i} + 1}{4(i+1)} \\
&\geq \frac{1}{4} (H_{n_1} + (H_{\rho-1} - 1)(H_{n_1} + 1)) - \frac{1}{4} \sum_{i=1}^{\rho-2} \frac{H_{2i}}{i+1} \\
&\stackrel{(3.10)}{\geq} \frac{1}{4} \left(H_{\rho-1} H_{n_1} - \frac{1}{2} H_{2\rho-3}^2 \right) \\
&\stackrel{(3.8)}{\geq} \ln \rho (\ln n_1 - \ln \rho) + O(\log n). \tag{3.14}
\end{aligned}$$

We may assume that ρ as well as n_1 are multiples of powers of n , $\rho = an^\alpha$ and $n_1 = bn^\beta$, say, $a, b > 0$, $\alpha, \beta \leq 1$. Then Equation (3.14) becomes

$$\sum_{i=0}^{\rho-2} \frac{\tau_{i+n_1,i}}{4(i+1)} \geq \alpha \left(\beta - \frac{\alpha}{2} \right) \ln^2 n + O(\log n) \tag{3.15}$$

Clearly, we achieve the best bound for $\alpha = \beta = 1$.

As we may estimate the number of pivots during right phases analogously, we may deduce that for the considered configuration we get the best lower bound for the expected number of pivot steps by choosing $n_1 = n_2 = \rho/2 = n/4$; then $E(Z) \geq \frac{1}{4} \ln^2 n + O(\log n)$.

Thus, we have proved:

Theorem 3.8 *There exist instances of one line and n points in general position in the plane such that the expected number of pivot steps for the Fast Process 2.4 satisfies the following bound:*

$$E(Z) \geq \frac{1}{4} \ln^2 n + O(\log n).$$

If we set $n_1 = n - 2\sqrt{n}$, $n_2 = 0$ and $\rho = \sqrt{n}$, we get $E(Z) \geq \frac{3}{8} \ln^2 n + O(\log n)$. This instance will form the 2-dimensional building block of our 3-dimensional lower bound construction in Section 3.5.

3.4.2 An Instance with an Extraordinary Property

As before, let ℓ be the y-axis, and assume that the number of points in the set S is a multiple of 4, $n = 4k$ for some $k \in \mathbb{N}$. We place the points s_0, s_1, \dots, s_{n-1} onto the graph of the function $y = f(x) = \ln|x|$ by the following procedure:

Let $s_{n-1} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, $s_{n-2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $s_{n-3} = \begin{pmatrix} -1+\epsilon \\ \ln(1-\epsilon) \end{pmatrix}$. Now consider the line spanned by the point s_{n-1} and s_{n-3} . It intersects the graph of the function f in a third point on the opposite side of ℓ . Place the point s_{n-4} onto the graph of f and just below this point of intersection, cf. Figure 3.9.

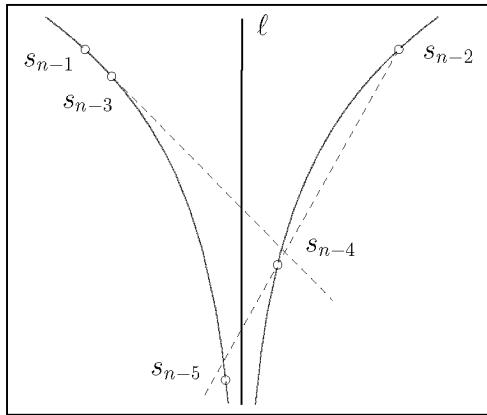


Figure 3.9: Placing the first five points.

For the remaining $n - 4$ points we repeat the procedure just described: Each point s_i , for $i = n - 5, \dots, 0$ is placed just below the intersection point of the graph of f with the line spanned by the points s_{i+1} and s_{i+3} .

The configuration constructed according to these rules has the following properties, cf. Figure 3.10:

Observation 3.9

- (i) $x_{s_{2j+1}} < x_{s_{2i+1}} < 0 < x_{s_{2i}} < x_{s_{2j}}$ for any $0 \leq i < j < n/2$,
- (ii) $\{s_{i-j'} \mid 0 < j' \leq i\}$
 $\subseteq \text{below}(\{s_i, s_{i+1+2j}\})$ for $0 \leq i < n - 1$ and $0 \leq j < \lfloor \frac{n-i}{2} \rfloor$.
- (iii) $\{s_{i+1+2j} \mid 0 < j < \lfloor \frac{n-i}{2} \rfloor\}$
 $\subseteq \text{below}(\{s_{i-2j'}, s_{i+1}\})$ for $0 \leq i < n - 1$ and $0 \leq j' < i$.

Proof (i) and (ii) follow directly from the construction while (iii) is less obvious. Recall, that we place s_i below the line spanned by s_{i+1} and s_{i+3} . This implies that s_{i+3} lies below the ℓ -edge $\{s_i, s_{i+1}\}$. By concavity of the function $f(x) = \ln|x|$, the points s_{i+3+2j} ($j > 0$) also lie below $\{s_i, s_{i+1}\}$. In fact, these points lie below any ℓ -edge $\{s_{i-2j'}, s_{i+1}\}$, $0 \leq j' \leq i$. The observation follows. \square

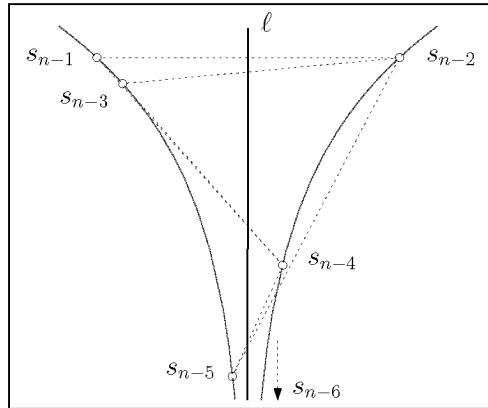


Figure 3.10: There exists a sequence of pivots which visits all possible ℓ -edges.

Corollary 3.10 In the configuration (S, ℓ) as constructed above and shown in Figure 3.10, there exists a sequence of pivots which visits all possible ℓ -edges.

Proof We can rewrite Observation 3.9.(iii) as

$$\{s_{i+1+2(j+j')} \mid 0 < 2j < n - 2j' - i - 1\} \subseteq \text{below}(\{s_i, s_{i+1+2j'}\})$$

for $0 \leq i < n - 1$ and $0 \leq j' < \lfloor \frac{n-i}{2} \rfloor$. This implies that starting from some ℓ -edge $\{s_i, s_{i+1}\}$ we can pivot through all ℓ -edges $\{s_i, s_{i+1+2j'}\}$, $0 \leq j' < \lfloor \frac{n-i}{2} \rfloor$, until we reach $\{s_i, s_{n-1}\}$ (if i is even) or $\{s_i, s_{n-2}\}$ (otherwise).

From there, we pivot to $\{s_i, s_{i-1}\}$ — this is possible by Observation 3.9.(ii) — and repeat the procedure. Having started from the top-most ℓ -edge $\{s_{n-1}, s_{n-2}\}$, we will, eventually, have pivoted through all possible ℓ -edges to the bottom-most ℓ -edge: $\{s_{n-1}, s_0\}$. \square

This implies that there are simple $(n - 2)$ -polytopes with n facets where we can pivot through all vertices in a monotone fashion (w.r.t. some linear function).

The analysis is based on the same principles that proved successful for the previous instance in Section 3.4.

Again, we divide the process into distinct phases, using the following criteria: phase a comprises all visited ℓ -edges $e = \{s_i, s_j\}$ for which $\min(i, j) = a$. More specifically, a new phase is entered whenever we choose a point whose index sets a new minimum. We refer to this points as *phase-point*; clearly, s_a is the phase-point of phase a . Observe, that the phase-point is endpoint of all edges belonging to that phase.

We use the notation we are already familiar with from Section 3.4.1: the random variables $X_i, i = 0, \dots, 2k - 1$ denote the number of pivots during phase i ; Z is the total number of pivots under the assumption that we start with the edge $\{s_{2k-1}, s_{2k}\}$; Ph_i denotes the event that phase i is entered.

We aim to bound the expected value of Z from below.

Conditioned on the event of choosing the first point in

$\{p_0, \dots, p_i\}$
 p_i is chosen with probability $\frac{1}{2(i+1)}$. This implies that

$$\Pr(\text{Phase } i \text{ is entered}) = \begin{cases} \frac{1}{i+1} & \text{for all } i, 0 \leq i < 2k - 1 \\ 1 & \text{for } i = 2k - 1. \end{cases} \quad (3.16)$$

and

$$\mathbb{E}(Z) = \mathbb{E}(X_{2k-1} \mid \text{Ph}_{2k-1}) + \sum_{i=0}^{2k-2} \frac{\mathbb{E}(X_i \mid \text{Ph}_i)}{i+1}. \quad (3.17)$$

We can get a good estimate for the expectations of the X_i (conditioned on the event that phase i is actually entered) if s_i and the phase-point preceding it lie on different sides of ℓ . If that is the case, we say that phase i is *special*. Then at least k points (either $s_{2k}, s_{2k+2}, \dots, s_{4k-2}$ or $s_{2k+1}, s_{2k+3}, \dots, s_{4k-1}$) are once again below the current ℓ -edge by Observation 3.9.

During phase i , there are exactly i points below the current edge which – when chosen – would start a new phase. $\lfloor \frac{i}{2} \rfloor$ of these points are on the same side of ℓ as s_j , $\lceil \frac{i}{2} \rceil$ are on the other side. So, the probability, that the phase-points of two successively entered phases are separated by ℓ is $\geq \frac{1}{2}$.

Let τ_i denote the expected length of phase i (that is, the expected number of steps until a new phase starts), under the assumption that i is *special*, i.e. that the previous phase was on the other side of ℓ .

A special phase i can be considered as an extended 1-dimensional process on k points and i exit-points, cf. Section 3.1. Using Equations (3.4), we get

$$\tau_i \geq t_{k,i} = \begin{cases} H_{k+i} - H_i + 1 & \text{whenever } i > 0, \\ H_k & \text{whenever } i = 0. \end{cases} \quad (3.18)$$

To derive the lower bound we shall use this in Equation (3.17):

$$\begin{aligned} \mathbb{E}(Z) &\geq \sum_{i=0}^{2k-2} \frac{\mathbb{E}(X_i \mid \text{Ph}_i)}{i+1} \\ &\geq \sum_{i=0}^{2k-2} \frac{\mathbb{E}\left(X_i \mid \text{Ph}_i \wedge \begin{array}{c} \text{side of minimum} \\ \text{changed} \end{array}\right)}{i+1} \Pr\left(\begin{array}{c} \text{side of minimum} \\ \text{changed} \end{array}\right) \\ &\stackrel{(3.18)}{\geq} \sum_{i=0}^{2k-2} \frac{\tau_i}{2(i+1)} = \frac{H_k}{2} + \sum_{i=1}^{2k-2} \frac{H_{k+i} - H_i + 1}{2(i+1)} \\ &\geq \frac{1}{2} \left(\sum_{i=0}^{2k-2} \frac{H_k}{i+1} - \sum_{i=1}^{2k-2} \frac{H_i}{i+1} \right), \end{aligned}$$

giving us

$$\begin{aligned} E(Z) &\stackrel{(3.9)}{\geq} \frac{1}{2} \left(H_k H_{2k-1} - \frac{1}{2} H_{2k-1}^2 \right) \\ &\stackrel{(3.8)}{\geq} \frac{1}{2} \left(\ln k \ln(2k-1) - \frac{1}{2} \ln^2(2k-1) \right) + O(1) \\ &\geq \frac{1}{4} \ln^2 n + O(\log n), \end{aligned}$$

recalling that $n = 4k$.

Thus, we have proven:

Theorem 3.11 *There exist simple $(n-2, n)$ -polytopes where we can pivot through all vertices in a monotone fashion (w.r.t. some linear function) and the expected number of pivot steps for RANDOM EDGE satisfies the following bound:*

$$E(Z) \geq \frac{1}{4} \ln^2 n + O(\log n).$$

3.5 A Lower Bound for Dimension 3

As already indicated in Section 3.4 above, we view the 3-dimensional process as the succession of 2-dimensional subprocesses; this will be the key to analyse RANDOM EDGE on $(d, d+3)$ -polytopes.

In fact, we can partition the set of points S into three pairwise disjoint sets: the set Π_1 , consisting of the points that form the 1-dimensional subgame, Π_2 , the additional points that (together with Π_1) form the 2d-subgame, and Π_3 , the remaining ones.

We will commence with placing the points forming the set Π_3 . These will induce several conditions on the location of the other points. Having identified them, we can allocate their coordinates accordingly. Finally, having done all the prerequisites, we may do the analysis.

3.5.1 Placing some Points

Suppose we have $n = m + 2\sqrt{m} + 3\sqrt[4]{m}$ points at our disposal.⁶ For our convenience, we group them into six point sets P, Q, R, U, V and W , pairwise

⁶We choose m such that $m = (2\tilde{m})^4$ for some $\tilde{m} \in \mathbb{N}$.

disjoint, such that $|P| = |Q| = |R| = \sqrt[4]{m}$, $|U| = |V| = \sqrt{m}$ and $|W| = m$. Using the classification from above, we let $\Pi_1 := W$, $\Pi_2 := U \cup V$ and $\Pi_3 := P \cup Q \cup R$. As before, we will use the small, indexed letter s to denote points. We denote points by the same letter as the set they belong to, if we need to indicate this property. For example, we may write p_i for some point in P , v_j for a point in V .

Let ℓ be the z -axis. We place the points $p_0, \dots, p_{\sqrt[4]{m}-1}$ (i.e. the set P) onto the line $\{(x, y, z)^T \in \mathbb{R}^3 \mid x = -1, y = 0\}$, we place $q_0, \dots, q_{\sqrt[4]{m}-1}$ (i.e. the set Q) onto $\{(x, y, z)^T \in \mathbb{R}^3 \mid x = 1, y = -1\}$, and $r_0, \dots, r_{\sqrt[4]{m}-1}$ (i.e. R) onto the line $\{(x, y, z)^T \in \mathbb{R}^3 \mid x = 1, y = 1\}$ such that p_i, q_i, r_i lie on the plane $\{(x, y, z) \in \mathbb{R}^3 \mid z = i\}$. To keep the size of the indices on a manageable level, we set $m' := \sqrt{m} - 1$, $m'' := \sqrt[4]{m} - 1$, and write, for instance, $p_{m''}$ when referring to $p_{\sqrt[4]{m}-1}$.

Consider now some plane e given by the equation $x = C$, where $C \gg 1$ is a large constant, depending on n . This plane is parallel to ℓ . The intersection of e with the plane spanned by P and ℓ shall be denoted by ℓ' . Running the 2-dimensional process of k points and the line ℓ' on the plane e is then identical to running the 3-dimensional process of $k + 1$ points and the line ℓ where the additional point is some fixed $p \in P$, and vice versa.

This motivates placing the remaining point sets onto e as depicted in Figures 3.12 and 3.11, *cf.* our 2-dimensional construction of Section 3.4.1 in Figure 3.7: U and $V \cup W$ are separated by ℓ' , U lying in the halfspace $y < 0$.

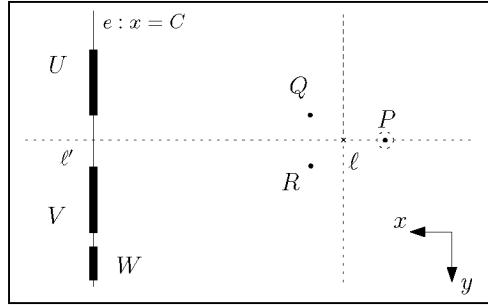


Figure 3.11: Outlook where we hope to place the points – from above.

Before we can allocate specific coordinates to each point, however, we first need to understand some properties that turn out to be crucial for the successful analysis.

Recall, that we call any triple of points in $\{s_1, s_2, s_3\} \in S$ an ℓ -stabbed simplex, if the line ℓ intersects their convex hull. We will write (s_1, s_2, s_3) whenever we refer to the plane spanned by the simplex $\{s_1, s_2, s_3\}$. Furthermore, (s_1, s_2) will denote the line through the points s_1 and s_2 .

The following Lemma is easily observed, as illustrated by Figure 3.11:

Lemma 3.12 *The vertex set of any ℓ -stabbed simplex must contain a point in P , a point in $Q \cup U$ and a point of the set $R \cup V \cup W$.*

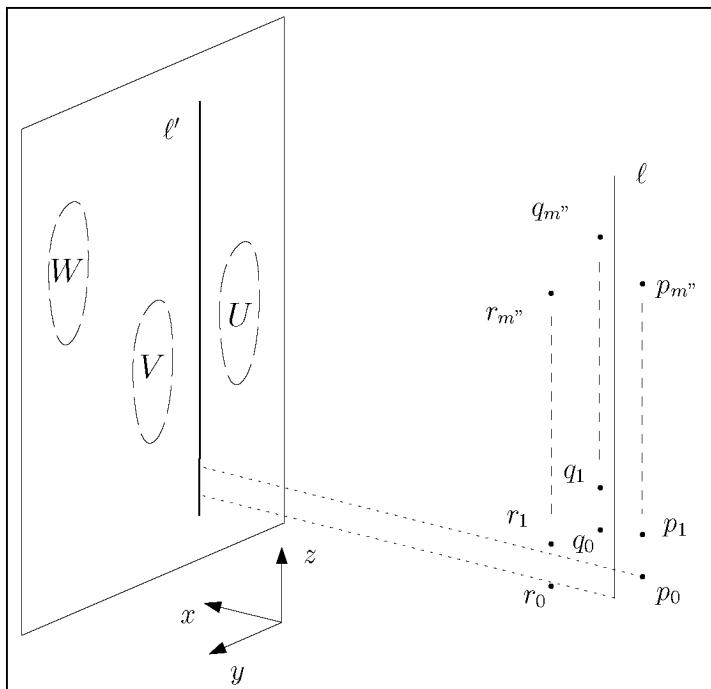


Figure 3.12: Outlook where we hope to place the points.

3.5.2 Finding the Place for the Remaining Points

We want to analyse the number of pivots given that the starting ℓ -stabbed simplex is $\{p_{m''}, q_{m''}, r_{m''}\}$. If we denote this number by G then we are interested in a lower bound for the value of $E(G)$.

We divide the process into distinct phases. Let I be the index set of those vertices of the current ℓ -stabbed simplex which are in $\Pi_3 = \{P \cup Q \cup R\}$. Then we say that we are in phase i as long as $\min I = i$. We will ensure that

- (i) $\min I$ can only decrease during the entire process.

By Lemma 3.12, I is nonempty. Any point whose index sets a new minimum starts a new phase. We call such a point a *phase-point*. Consequently, we call a phase p -phase whenever its phase-point belongs to the set P . Similarly, we may refer to q - and r -phases.

As we will see later, we may estimate the length of a p -phase if two further conditions are met:

- (ii) The vertex set of the first ℓ -stabbed simplex of the p -phase consists of points in Π_3 only, and
- (iii) the entire 2-dimensional subgame is below the first ℓ -stabbed simplex of the p -phase.

Our goal is, therefore, to ensure that these requirements are met for any p -phase with some positive probability.

This can be achieved by imposing the following conditions:

Condition 1 *The phase-point of a given phase is part of all ℓ -simplices belonging to that phase. This follows if,*

- U is above any plane $\left\{ \begin{array}{l} (p_i, q_j, r_k), \quad j \leq i, k, \\ (p_i, q_j, s_k), \quad j \leq i, \quad s_k \in V \cup W, \end{array} \right.$
(that is, above any plane encountered during a q -phase),
- $V \cup W$ is above any plane $\left\{ \begin{array}{l} (p_i, q_j, r_k), \quad k \leq i, j, \\ (p_i, s_j, r_k), \quad k \leq i, \quad s_j \in U. \end{array} \right.$
(that is, above any plane encountered during an r -phase).

Condition 2 Given that the vertex set of the first ℓ -stabbed simplex of the current p -phase is a subset of Π_3 , the entire 2-dimensional subgame lies below the plane spanned by this simplex. This follows if,

- $U \cup V \cup W$ is below any plane (p_i, q_j, r_k) , $i < j, k$.

While Condition 2 just repeats prerequisite (iii) from above, the link between Condition 1 and (i) is not as obvious. Here, we ensure that the phase-point does not leave the phase it has initiated. Therefore, this phase-point still has to be present at the beginning of the succeeding phase — unless both phase-points belong to the same set (P , Q or R). In any case, the value of $\min I$ always goes down.

As we will see later, requirement (ii) cannot always be met. But we shall prove the following implication of Condition 1: the first ℓ -stabbed simplex of some p -phase which succeeds consecutive q - and r -phases will, *with positive, constant probability*, be spanned solely by points of the set Π_3 .

But first let us impose the conditions intrinsic to the 2-dimensional subprocess:

Condition 3 The set W lies below the plane (p_i, u_j, v_k) whenever $j < k$. If $j \geq k$ then W lies above (p_i, u_j, v_k) .

As we hope to place all the points of the 2-dimensional subgame onto the plane e we translate these 3-dimensional conditions into 2-dimensional relationships valid on e .

To determine the line of intersection of the plane e with some ℓ -stabbed simplex (p_i, q_j, r_k) we may consider the two points of intersection of the line (p_i, q_j) with e and of (p_i, r_k) with e , denoted by p_{iqj} and p_{irk} , respectively. (p_{iqj}, p_{irk}) is then the line we were looking for.

By construction, the points of the sets P , Q and R lie on three parallel, vertical lines. Therefore, $P \cup Q$ and $P \cup R$ both affinely span a plane whose intersection with the plane e is again a vertical line. More precisely, the points of intersections of e with some line (p_i, q_j) (and (p_i, r_k) , respectively) lie on the line $\{(x, y, z)^T \in \mathbb{R}^3 \mid x = C, y = -\frac{1}{2}(C + 1)\}$ (and $\{(x, y, z)^T \in \mathbb{R}^3 \mid x = C, y = \frac{1}{2}(C + 1)\}$, respectively). The potential location of these points of intersection can be narrowed down even more. Choosing C sufficiently large ($C \gg n$), the points p_{iqj} have a z -coordinate that has its extrema at $p_m q_0$ and $p_0 q_m$, the points p_{irk} lie on a line segment bounded by $p_m r_0$ and $p_0 r_m$. This fact can be easily deduced from the actual coordinates of the

points of intersection:

$$p_i q_j = \begin{pmatrix} C \\ -\frac{C+1}{2} \\ \frac{C}{2}(j-i) + \frac{1}{2}(j+i) \end{pmatrix}$$

and

$$p_i r_k = \begin{pmatrix} C \\ \frac{C+1}{2} \\ \frac{C}{2}(k-i) + \frac{1}{2}(k+i) \end{pmatrix}.$$

Let us consider the constraints for U . By Condition 1, U has to be above any plane (p_i, q_j, r_k) with $j \leq i, k$. On the plane e , therefore, all points in U need to be above any line given by two points of intersection $p_i q_j$ and $p_i r_k$ with $j \leq i, k$. This is certainly the case if we enforce U to be above any line through points $p_i q_j$ and $p_{i'} r_k$, where $i \geq j$ and without further constraints on i', k . Thus, it is sufficient to ensure that U lies above the two ‘extremal’ lines $(p_m q_m, p_m r_0)$ and $(p_m q_m, p_0 r_m)$.

On the other hand, Condition 2 says that U is below any plane (p_i, q_j, r_k) , $i < j, k$. This is the case, if on the plane e , all elements of the set U lie below any line $(p_i q_j, p_{i'} r_k)$ with $i < j, i' < k$. Again, there are essentially only two lines we need to consider: if U lies below $(p_0 q_1, p_0 r_1)$ and below $(p_0 q_1, p_0 r_m)$ then U satisfies Condition 2.

By symmetry, the conditions for $V \cup W$ can be derived similarly. We call potential regions to place the remaining points ‘good’. In Figure 3.13 they are shaded accordingly.

Having determined the potential regions for the elements of U , V and W we realize that the part of Condition 1 that we have not considered so far is automatically fulfilled: U is above any plane (p_i, q_j, s_k) , $j \leq i, s_k \in V \cup W$, $V \cup W$ is above any plane (p_i, s_j, r_k) , $k \leq i, s_j \in U$. Our next step is to find the exact location of the points in U , V and W .

The vertical line segments $(p_0 r_1, p_m r_m)$ and $(p_0 q_1, p_m q_m)$ (with the exception of their endpoints) are ‘good’. So, they are the natural location to place the points of U and V . After placing these points equidistantly, i.e. such that $\|u_i - u_{i+1}\| = c_1$ and $\|v_i - v_{i+1}\| = c_2$ for two constants $c_1 > c_2$, and any i , $0 \leq i < \sqrt{m} - 1$, we may apply the theorem of intersecting lines: the lines (u_i, v_i) , $i = 0, \dots, m'$ intersect in one point (we call it A), as do the lines (u_i, v_{i+1}) , $i = 0, \dots, m' - 1$ (we call it B). Furthermore, the line through A and B is vertical, too.

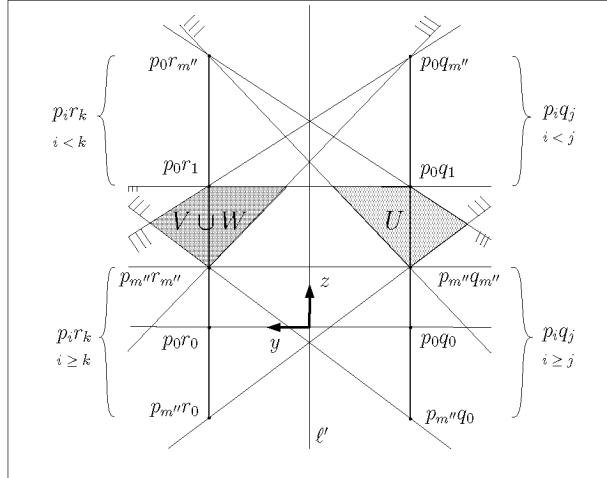


Figure 3.13: The sets U and $V \cup W$ need to be located in the shaded areas.

Recall Condition 3: the points in the set W are required to lie above all lines (u_j, v_k) (whenever $j \geq k$) and below the lines (u_j, v_k) with $j < k$. Placing W onto the segment (A, B) , this condition will be fulfilled.

Finally, by choosing suitable constants c_1 and c_2 , we may move the points A and B arbitrarily close to V 's line segment, cf. Figure 3.14. In particular, A and B and hence the point set W may lie in the ‘good’ region while Conditions 1—3 are fulfilled.

We finish this subsection by giving specific coordinates as promised. Setting $C := 2m^2 - 1$, $c_1 = 1$ and $c_2 = \frac{1}{\sqrt{m}}$, we let:

$$u_i := \begin{pmatrix} 2m^2 - 1 \\ -m^2 \\ m - \sqrt{m} + i \end{pmatrix}, \quad v_i := \begin{pmatrix} 2m^2 - 1 \\ m^2 \\ m - 1 + \frac{i}{\sqrt{m}} \end{pmatrix} \quad \text{for } i = 0, \dots, \sqrt{m} - 1,$$

$$\text{and } w_j := \begin{pmatrix} 2m^2 - 1 \\ m^2(1 + \frac{2}{\sqrt{m}-1}) \\ m + \frac{j+1}{m^2} \end{pmatrix} \quad \text{for } j = 0, \dots, m - 1.$$

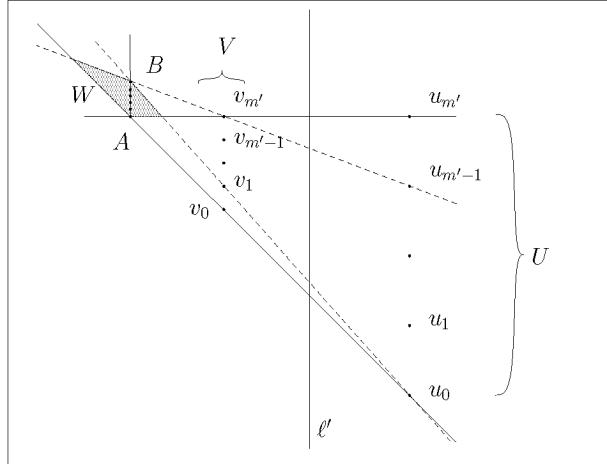


Figure 3.14: W can be placed arbitrarily close to V .

Lemma 3.13 *The given construction satisfies the Conditions 1—3.*

3.5.3 The Analysis

We define the random variables $Z_i, i = 0, \dots, \sqrt[4]{m} - 1$ as the length of phase i , that is, the number of pivots *leading to* ℓ -simplices belonging to the phase i .⁷

During phase i , let J be the index set of those vertices of the current ℓ -stabbed simplex which are in $\Pi_2 = U \cup V$. We say we are in subphase $j^{(i)}$ if $\min J = j$. Phase i has a subphase $\sqrt{m}^{(i)}$ if J is empty and the plane spanned by s is above the points in Π_2 . We do not count any other pivots of the current phase where $J = \emptyset$.

By Condition 1, the point starting a new subphase will remain as a vertex in all ℓ -stabbed simplices belonging to that subphase. We call it *the subphase-point*, and refer to a subphase initiated by some point in U (or V) as u - (or v -)subphase.

⁷Note that this definition implies that the phase $\sqrt[4]{m} - 1$ has length 0. Therefore, we can safely ignore this phase in all our considerations below. The first phase will be the phase we enter with the first pivot.

The random variables $Y_j^{(i)}, j = 0, \dots, \sqrt{m}$ shall denote the number of pivots during the subphase j in phase i .

Finally, let us even divide the subphases. The subsubphase $k^{(i,j)}$ consists just of the single pivot which puts w_k into the current ℓ -stabbed simplex. This pivot is succeeded by either a new subsubphase $k'^{(i,j)}$ (if $w_{k'}$ is the next point) or the beginning of a new (sub)phase. So defining the random variables $X_k^{(i,j)}$ as the length of subsubphase k of subphase j in phase i , we have

$$\mathbb{E}(X_k^{(i,j)} \mid \text{subsubphase } k^{(i,j)} \text{ is entered}) = 1. \quad (3.19)$$

This was easy, but we also need to know how likely it is to actually enter $k^{(i,j)}$, $j^{(i)}$ or i , and need to estimate their expected duration. We shall first determine candidates for ‘long’ phases and estimate the probability of their occurrence. Then we do the same for subphases. With these results, the actual analysis is reduced to a technical computation.

Definition 3.14 *We call a p -phase special whenever its initial ℓ -stabbed simplex s contains only points in Π_3 , and the entire 2-dimensional subgame, i.e. $\Pi_1 \cup \Pi_2$, lies below the plane spanned by s .*

We make the following observation that helps us to estimate the likelihood of a p -phase to be special:

Observation 3.15 *A p -phase i is special if it directly succeeds a q - and an r -phase (both appearing in any order) and the phase-points of these two phases are still in the ℓ -stabbed simplex initiating i .*

So, our requirement (ii) from above is met whenever we encounter a special p -phase.

Lemma 3.16 *Suppose that we are in phase i_1 , let i_2 denote the index of the succeeding phase. Then $\Pr(i_2 \geq \alpha) = 1 - \frac{\lfloor \alpha \rfloor}{i_1}$ for any α , $0 \leq \alpha < i_1$.*

Proof Conditioned on the event of choosing a point with index smaller than i_1 for the first time, all such points are equally likely chosen, with probability $\frac{1}{i_1}$. $i_1 - \lfloor \alpha \rfloor$ of these points have index bigger or equal than α . \square

We also recall the following lemma:

Lemma 3.17 *For any events A, B, C ,*

$$\begin{aligned}\Pr(A) &= \Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c) \\ &\geq \Pr(A|B)\Pr(B),\end{aligned}$$

and

$$\begin{aligned}\Pr(A|C) &= \Pr(A|C \cap B)\Pr(B) + \Pr(A|C \cap B^c)\Pr(B^c) \\ &\geq \Pr(A|C \cap B)\Pr(B),\end{aligned}$$

where B^c denotes the complement of B .

Proof $A = (A \cap B) \cup (A \cap B^c)$. This is a disjoint union and so

$$\begin{aligned}\Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ &= \Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c).\end{aligned}$$

Replacing A by $(A|C)$, the second (in)equality follows as well. \square

Thus armed, we can provide proof of the following lemma:

Lemma 3.18 *The probability that some entered phase i is a special p -phase is at least $\frac{1}{45}$.*

Proof We want to estimate the probability, that the initial ℓ -stabbed simplex of phase i is of the form $\{p_i, q_j, r_k\}$ with $i < j, k$. By Observation 3.15, this can be bounded from below by the probability of the following event: phase i is a p -phase that was preceded by a q - and an r -phase, where the respective phase-points q_j and r_k are still in the simplex. (Note that the very first or the second phase will be special whenever they are p -phases. The general estimate will also be valid for these early phases.)

The probability that a p -phase is preceded by a q - and an r -phase (successively, but not necessarily in this order) is at least $\frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$, where no assumptions are made about the fate of the phase-point of the first phase. By the symmetry of the Conditions 1 and 2, we may assume (without loss of generality) that the q -phase succeeds an r -phase. Then by Condition 1 the first phase-point, r_k , is still in the current ℓ -stabbed simplex when the q -phase starts. At this moment our simplex is of the form $\{p_{i'}, q_j, r_k\}$ with $i' \geq k > j$. It is crucial to begin the p -phase before a point in $V \cup W$ may replace r_k – which may happen as

soon as some $p_{i''}$, $k > i'' > j$ has entered the ℓ -stabbed simplex (during the q -phase).⁸

We now consider the event that conditioned on the event of choosing a point in $\{p_0, \dots, p_{k-1}\} \cup \{q_0, \dots, q_{j-1}\} \cup \{r_0, \dots, r_{k-1}\}$ the next point will be in $\{p_0, \dots, p_{j-1}\}$. Calling this event F , we need to show that the probability of F is greater or equal some positive constant. The situation is illustrated in Figure 3.15 below.

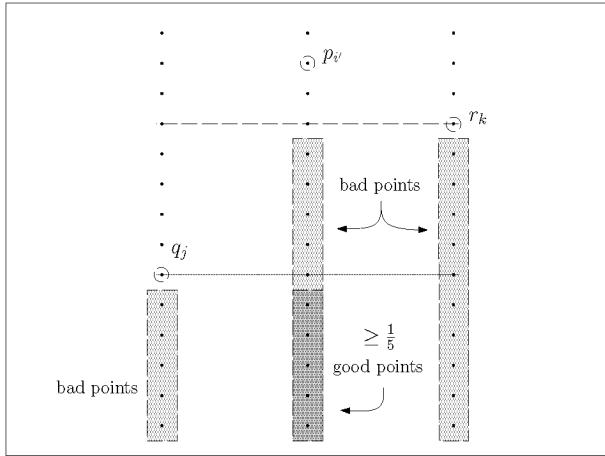


Figure 3.15: The new p -phase preceded by a q - and an r -phase is special with probability $\geq \frac{1}{2} \cdot \frac{1}{5} = \frac{1}{10}$.

Using Lemma 3.17 we can bound $\Pr(F)$ from below:

$$\Pr(F) \geq \Pr(F | j \geq \alpha) \Pr(j \geq \alpha).$$

By Lemma 3.16 we know $\Pr(j \geq \alpha) = 1 - \frac{\lfloor \alpha \rfloor}{k}$. On the other hand, it is not hard to see that $\Pr(F | j \geq \alpha) \geq \frac{\lfloor \alpha \rfloor}{\lfloor \alpha \rfloor + 2k}$.

A convenient choice for α is $\frac{k}{2}$, giving us

$$\Pr(F | j \geq \alpha) \geq \frac{1}{5} \text{ and } \Pr(j \geq \alpha) \geq \frac{1}{2}.$$

⁸Our conditions are not strong enough to prevent this possibility. In fact, one can show that imposing an additional condition (*the set $V \cup W$ lies above any plane (p_i, q_j, r_k) , $k > i > j$*) would reduce the potential region where we can place the points in V and W to zero.

(Note that $\lfloor \alpha \rfloor = \frac{k-1}{2}$ if k is odd.) Hence,

$$\Pr(F) \geq \frac{1}{5} \cdot \frac{1}{2} = \frac{1}{10}.$$

Putting everything together (again using Lemma 3.17) we have

$$\Pr(i \text{ is special } p\text{-phase} \mid \text{Ph}_i) \geq \frac{2}{9} \cdot \frac{1}{2} \cdot \frac{1}{5} = \frac{1}{45}. \quad (3.20)$$

□

We commence with the discussion of subphases:

Definition 3.19 We call a u -subphase special if (i) it belongs to a special p -phase, and (ii) the entire 1-dimensional subgame, i.e. Π_1 , lies below the plane spanned by s_u , where s_u is the first ℓ -stabbed simplex of the subphase.

Observation 3.20 A u -subphase j is special if it is part of a special p -phase and succeeds a v -subphase.

Lemma 3.21 The probability that some subphase j of a special p -phase is a special u -subphase (under the condition that it is entered) is at least $\frac{1}{4}$.

Proof Each subphase is equally likely a u - or a v -subphase. Since by the observation above, a u -subphase is special if it was preceded by a v -subphase, the result follows. (Note that the very first u -subphase of some special phase is always special. But the lower bound holds in this case as well.) □

Let us introduce some terminology. We denote the event that phase i is entered by Ph_i , that i is special under the condition that it is entered by ${}^s\text{Ph}_i$. ${}^s\text{Ph}_j^{(i)}$ (${}^s\text{Ph}_j^{(i)}$) will denote that during phase i subphase j is entered (and special) under the assumption that i is special. Analogously, ${}^s\text{Ph}_k^{(i,j)}$ is the event that subsubphase k is entered while the game is in the special subphase j of the special phase i .

Conditioned on the event of choosing a point in $\{p_0, \dots, p_i\} \cup \{q_0, \dots, q_i\} \cup \{r_0, \dots, r_i\}$ for the first time, p_i, q_i and r_i are equally likely chosen with probability $\frac{1}{3(i+1)}$. This implies that

$$\begin{aligned} \Pr(\text{Phase } i \text{ is entered}) &= \Pr(\text{Ph}_i) \\ &= \begin{cases} \frac{1}{i+1} & \text{for all } i, 0 \leq i < \sqrt[4]{m} - 1 \\ 1 & \text{for } i = \sqrt[4]{m} - 1. \end{cases} \end{aligned} \quad (3.21)$$

Assume that we are currently in the special phase i . Conditioned on the event of choosing the first point in $\{u_0, \dots, u_j\} \cup \{v_0, \dots, v_j\} \cup \{p_0, \dots, p_{i-1}\} \cup \{q_0, \dots, q_{i-1}\} \cup \{r_0, \dots, r_{i-1}\}$, u_j and v_j are equally likely chosen with probability $1/(2(j+1) + 3i)$. (Note that a replacement of r_k by some w_l during the special phase keeps the set $\{u_0, \dots, u_j\} \cup \{v_0, \dots, v_j\}$ below the plane spanned by the current simplex, cf. Figure 3.13.) This implies that

$$\Pr(\text{sPh}_j^{(i)}) = \frac{2}{2(j+1) + 3i} \geq \frac{2}{2(j+1) + 3\sqrt[4]{m}}. \quad (3.22)$$

Analogously, being in the special subphase j of the special phase i , w_k is chosen with probability $1/((k+1) + 2j + 3i)$ as the first point in $\{w_0, \dots, w_k\} \cup \{u_0, \dots, u_{j-1}\} \cup \{v_0, \dots, v_{j-1}\} \cup \{p_0, \dots, p_{i-1}\} \cup \{q_0, \dots, q_{i-1}\} \cup \{r_0, \dots, r_{i-1}\}$. We have

$$\begin{aligned} \Pr(\text{ssPh}_k^{(i,j)}) &= \frac{1}{(k+1) + 2j + 3i} \\ &\geq \frac{1}{(k+1) + 2\sqrt{m} + 3\sqrt[4]{m}}. \end{aligned} \quad (3.23)$$

We also recall the Lemmas 3.18 and 3.21 that gave us a lower bound for the probability that some phase i is a special p -phase under the condition that it is entered and that some subphase j is a special u -subphase under the condition that it is entered and belongs to a special i -phase.

A sibling of Lemma 3.17 is the following lemma, whose proof is very similar and will, therefore, be omitted:

Lemma 3.22 *For any events A, B, C ,*

$$\begin{aligned} \mathbb{E}(A) &= \mathbb{E}(A|B)\Pr(B) + \mathbb{E}(A|B^c)\Pr(B^c) \\ &\geq \mathbb{E}(A|B)\Pr(B), \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}(A|C) &= \mathbb{E}(A|C \cap B)\Pr(B) + \mathbb{E}(A|C \cap B^c)\Pr(B^c) \\ &\geq \mathbb{E}(A|C \cap B)\Pr(B). \end{aligned}$$

We are now ready to find a lower bound estimate for the expected number of pivots given that the starting ℓ -stabbed simplex is $\{p_{m''}, q_{m''}, r_{m''}\}$, $\mathbb{E}(G)$.

Recall that we chose m such that $m = (2\tilde{m})^4$ for some $\tilde{m} \in \mathbb{N}$, and that we denoted the number of pivots during phase i by Z_i .

$$\begin{aligned} E(G) &= \sum_{i=0}^{\sqrt[4]{m}-2} E(Z_i) \geq \sum_{i=0}^{\sqrt[4]{m}-2} E(Z_i \mid {}^s\text{Ph}_i) \Pr({}^s\text{Ph}_i) \\ &\geq \sum_{i=0}^{\sqrt[4]{m}-2} E(Z_i \mid {}^s\text{Ph}_i) \overbrace{\Pr(i \text{ is special} \mid \text{Ph}_i)}^{\geq \frac{1}{45}} \Pr(\text{Ph}_i) \\ &\stackrel{(3.21)}{=} \frac{1}{45} \sum_{i=0}^{\sqrt[4]{m}-2} \frac{E(Z_i \mid {}^s\text{Ph}_i)}{i+1}. \end{aligned}$$

We bound the length of each special phase by the length of some of its (long) subphases:

$$\begin{aligned} E(G) &\geq \frac{1}{45} \sum_{i=0}^{\sqrt[4]{m}-2} \frac{1}{i+1} \sum_{j=0}^{\sqrt{m}} E(Y_j^{(i)} \mid {}^s\text{sPh}_j^{(i)}) \Pr({}^s\text{sPh}_j^{(i)}) \\ &\geq \frac{1}{45} \sum_{i=0}^{\sqrt[4]{m}-2} \frac{1}{i+1} \sum_{j=0}^{\sqrt{m}} \left(E(Y_j^{(i)} \mid {}^s\text{sPh}_j^{(i)}) \right. \\ &\quad \left. \cdot \overbrace{\Pr(j \text{ is special} \mid {}^s\text{sPh}_j^{(i)})}^{\geq \frac{1}{4}} \Pr({}^s\text{sPh}_j^{(i)}) \right) \\ &\stackrel{(3.22)}{\geq} \frac{1}{180} \sum_{i=0}^{\sqrt[4]{m}-2} \frac{1}{i+1} \sum_{j=0}^{\sqrt{m}-1} \left(\frac{E(Y_j^{(i)} \mid {}^s\text{sPh}_j^{(i)})}{(j+1) + \frac{3}{2}\sqrt[4]{m}} \right). \end{aligned}$$

The final step involves the splitting of the subphases into subsubphases. Their length of just one step is known, so, we can do the adding-up:

$$\begin{aligned} E(G) &\geq \frac{1}{180} \sum_{i=0}^{\sqrt[4]{m}-2} \frac{1}{i+1} \sum_{j=0}^{\sqrt{m}-1} \left(\frac{1}{(j+1) + \frac{3}{2}\sqrt[4]{m}} \right. \\ &\quad \left. \cdot \sum_{k=0}^{m-1} E(X_k^{(i,j)} \mid \text{ssPh}_k^{(i,j)}) \Pr(\text{ssPh}_k^{(i,j)}) \right) \end{aligned}$$

$$\begin{aligned}
E(G) &\stackrel{(3.23, 3.19)}{\geq} \frac{1}{180} \sum_{i=0}^{\sqrt[4]{m}-2} \frac{1}{i+1} \sum_{j=0}^{\sqrt[4]{m}-1} \left(\frac{1}{(j+1) + \frac{3}{2} \sqrt[4]{m}} \right. \\
&\quad \cdot \left. \sum_{k=0}^{m-1} \frac{1}{(k+1) + 2\sqrt{m} + 3\sqrt[4]{m}} \right) \\
&\geq \frac{1}{180} H_{\sqrt[4]{m}-1} \cdot (H_{\sqrt{m} + \frac{3}{2} \sqrt[4]{m}} - H_{\frac{3}{2} \sqrt[4]{m}}) \\
&\quad \cdot (H_{m+2\sqrt{m}+3\sqrt[4]{m}} - H_{2\sqrt{m}+3\sqrt[4]{m}}).
\end{aligned}$$

Using Equation (3.8) ($\ln n < H_n < \ln n + 1$, for $n > 1$), this gives us

$$\begin{aligned}
E(G) &\geq \frac{1}{180} \ln(\sqrt[4]{m} - 1) \cdot \left(\ln\left(\sqrt{m} + \frac{3}{2} \sqrt[4]{m}\right) - \ln\left(\frac{3}{2} \sqrt[4]{m}\right) - 1 \right) \\
&\quad \cdot (\ln(m + 2\sqrt{m} + 3\sqrt[4]{m}) - \ln(2\sqrt{m} + 3\sqrt[4]{m}) - 1) \\
&\geq \frac{1}{180} \ln(\sqrt[4]{m} - 1) \cdot \left(\ln\left(\frac{2}{3} \sqrt[4]{m} + 1\right) - 1 \right) \\
&\quad \cdot \left(\ln\left(\frac{1}{2} \sqrt{m} - \frac{3}{4} \sqrt[4]{m} + \frac{17}{8} + \frac{75}{8} \sqrt[4]{m}\right) - 1 \right) \\
&\geq \frac{1}{180} \ln\left(\sqrt[4]{m} \left(1 - \frac{1}{\sqrt[4]{m}}\right)\right) \cdot \left(\frac{1}{4} \ln m + \ln 2 - \ln 3 - 1 \right) \\
&\quad \cdot \left(\ln\left(\sqrt{m} \left(\frac{1}{2} - \frac{3}{4\sqrt[4]{m}}\right)\right) - 1 \right) \\
&\geq \frac{1}{5760} \ln^3 m + O(\ln^2 m).
\end{aligned}$$

Recalling that $n = m + 2\sqrt{m} + 3\sqrt[4]{m}$ we get our theorem:

Theorem 3.23 *There exist instances of one line and n points in general position in \mathbb{R}^3 such that the expected number G of pivot steps for the Fast Process described as Algorithm 2.4 satisfies the following bound:*

$$E(G) \geq \frac{1}{5760} \ln^3 n + O(\log^2 n).$$

Chapter 4

Admissible Grid Orientations

4.1 Notation and Terminology

Talking about oriented graphs is bound to be a technical affair. To clarify the discussion, we therefore introduce several notions which will help us to identify and use properties of the graph.

Let G be an admissible grid orientation with L rows and R columns, w.l.o.g. $L \geq R$, say. We use capital letters to denote the nodes. Each node can be further identified by its coordinates, that is a pair (u, v) , $0 \leq u < L$, $0 \leq v < R$, where u is the number of the row and v is the number of the column the node $P = (u, v)$ belongs to.

Two nodes P, Q in the same row or column are connected by an edge. We write $P \rightarrow Q$ to indicate that this edge is oriented from P to Q . We say that there is a *path* (of length $k - 1$) from P to Q , $(P \rightarrow Q)$, if there is a sequence of k vertices $P = P_1, P_2, \dots, P_k = Q$, such that $P_i \rightarrow P_{i+1}$ for all $i = 1, \dots, k - 1$. The length of the shortest path between P and Q is called the *distance* of P and Q , denoted by $|P \rightarrow Q|$. Note that not all pairs of points are actually linked by a path.

We define the set

$$\text{reach}(P) = \{Q \mid P \rightarrow Q\}$$

which comprises all nodes Q that are accessible from P by some path. To indicate that some node $Q \notin \text{reach}(P)$, we write $P \not\rightarrow Q$. Note that there might be nodes P, Q such that $P \not\rightarrow Q$ and $Q \not\rightarrow P$. Clearly, $P^*Q \implies P \rightarrow Q$ and, for P and Q in the same row or column, $P \rightarrow Q \implies P^*Q$.

Furthermore, let $\langle P, Q \rangle$ denote the smallest subgrid spanned by P and Q . (So, with $P = (u, v)$ and $Q = (r, s)$, $\langle P, Q \rangle$ consists of the nodes $(u, v), (u, s), (r, v)$ and (r, s) , and their connecting edges.) Analogously, let $\langle P, Q, R \rangle$ denote the smallest subgrid spanned by P, Q and R .

The remaining terminology we need for Section 4.3 exclusively.

For each node $P = (u, v)$ we define the set of rows

$$\text{R-below}(P) := \{u' \mid (u, v) \rightarrow (u', v)\}.$$

Analogously, we define the set

$$\text{C-below}(P) := \{v' \mid (u, v) \rightarrow (u, v')\},$$

comprising of all columns accessible directly from the node P . Finally, we call a node Q *below* P , if there is an edge from P taking us into the same row or in the same column as Q . We define the set $\text{below}(P)$ accordingly:

$$\text{below}(P) := \{(u', v') \mid (u' \in \text{R-below}(P)) \vee (v' \in \text{C-below}(P))\}.$$

Note that a node below P is not necessarily in the reach of P !

By this definition, each node $Q \in \text{below}(P)$ has a row in $\text{R-below}(P)$ or a column in $\text{C-below}(P)$, possibly both. Accordingly, we call Q 's row or column *witness for being below* P .

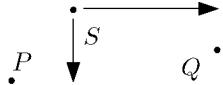


Figure 4.1: $P, Q \in \text{below}(S)$: with $P = (u, v)$ and $Q = (r, s)$, we have $u \in \text{R-below}(S)$ and $s \in \text{C-below}(S)$, respectively. u is P 's witness, s is Q 's witness for being below S .

4.2 The Path Condition

Lemma 4.1 *Given some grid orientation G of size $L \times R$, there are $L + R - 2$ vertex-disjoint paths from source to sink if no subgrid is isomorphic to the ‘forbidden subgrid’ depicted in Figure 2.5(b).*

Proof Let P be the source and Q be the sink of the oriented graph G . If P and Q share the same column then, just by the unique sink property, there are $R + L - 2$ vertex-disjoint paths from P to Q : $R - 1$ along the column of P and Q , and exactly one path along each of the $L - 1$ remaining columns, see Figure 4.2(a). Of course, the same is true if P and Q lie in the same row.

So, suppose that P and Q do not share the same row or column of G and there are less than $L + R - 2$ vertex-disjoint paths from source P to sink Q .

Then P and Q span a 2×2 -grid $PP'QQ'$. As we assumed that the path condition is not fulfilled, there must be vertices A and B such that $P \rightarrow A$ and $B \rightarrow Q$ but $B \rightarrow A$, cf. Figure 4.2(b). W.l.o.g. we may assume that A (B , respectively) shares the row with P and P' (Q and Q' , respectively). The unique sink property forces $A \rightarrow P'$ and $Q' \rightarrow B$, hence the presence of the forbidden subgraph. \square

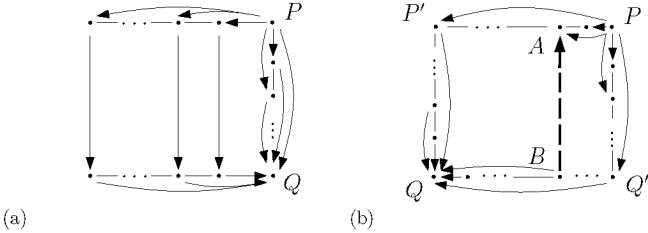


Figure 4.2: In a unique sink grid orientation, the path condition is equivalent to the nonexistence of the forbidden subgrid.

Theorem 4.2 *A grid orientation satisfying the unique sink condition is acyclic whenever it satisfies the path condition.*

The proof of the theorem is easy once we have established the Claims 1–3:

Claim 1 *A unique sink grid orientation cannot contain a cycle as in Figure 4.3 below.*

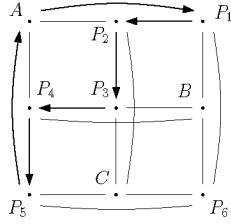


Figure 4.3: No unique sink grid orientation contains this cycle.

Proof Claim 1 To have a unique sink in the face A, P_4, P_5 , we need the edge $P_4 \rightarrow A$. But this leaves no way to orient the edge connecting A with P_2 . Either the unique sink property is violated in face A, P_1, P_2 or in A, P_2, P_3, P_4 . \square

Claim 2 A unique sink grid orientation cannot contain a cycle as in Figure 4.4 below.

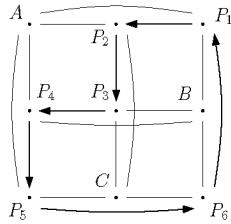
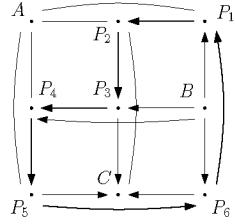


Figure 4.4: No unique sink grid orientation contains this cycle.

Proof Claim 2 This one is a little bit trickier.

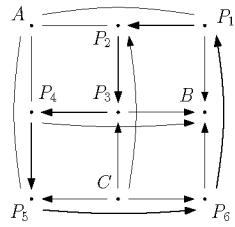
Suppose $P_5 \rightarrow C$. By the unique sink property of the various suborientations, this implies $P_3 \rightarrow C$, $P_2 \rightarrow C$, $P_6 \rightarrow C$, which means that C is, in this case, the unique sink. Assuming $P_3 \rightarrow B$ gives a contradiction as it implies $P_1 \rightarrow B$ leaving no choice for $B \rightarrow P_6$. (Either there would be no sink on the face P_1, P_6, B or two sinks on C, B, P_3, P_6 .) So, we must have $B \rightarrow P_3$, leading to $B \rightarrow P_4$, $B \rightarrow P_6$ and $B \rightarrow P_1$; see Figure 4.5 for the current state of affairs.

We still have all options for the edges incident to A . But $P_1 \rightarrow A$ implies $P_5 \rightarrow A$, hence $P_4 \rightarrow A$ and $P_2 \rightarrow A$, a contradiction to having the unique sink at C already. So, we must have $A \rightarrow P_1$. This, in turn, implies $P_4 \rightarrow A$. No matter how we now orient the edge connecting A and P_2 , we will always get a subgraph not satisfying the unique sink property.

**Figure 4.5:** Assuming $P_5 \rightarrow C$.

So $P_5 \rightarrow C$ leads to a contradiction.

Therefore, suppose $C \rightarrow P_5$. Then, we must have $C \rightarrow P_6$, $C \rightarrow P_2$ and $C \rightarrow P_3$. Since $B \rightarrow P_3$ makes it impossible to orient the edge between B and P_6 , we may assume $P_3 \rightarrow B$. Therefore, $P_1 \rightarrow B$, $P_6 \rightarrow B$ and $P_4 \rightarrow B$, leading us to an orientation as in Figure 4.6: C is the source, B the unique sink.

**Figure 4.6:** Assuming $C \rightarrow P_5$.

Now, $P_5 \rightarrow A$ will lead to $P_4 \rightarrow A$, $P_2 \rightarrow A$ and $P_1 \rightarrow A$, making A another sink which is impossible. Hence, we must have $A \rightarrow P_5$. But this would imply $P_2 \rightarrow A$, leaving no way to orient the edge connecting A with P_1 without causing a violation of the unique sink condition – giving us the contradiction which establishes the claim. \square

Claim 3 Given a grid orientation satisfying the unique sink property¹ as well as the path condition. For any pair of vertices P, Q , connected by some path $P \rightarrow Q$, we have $|P \rightarrow Q| \leq 3$.

Proof Claim 3 Suppose we have found two vertices P and Q with $|P \rightarrow Q| \geq 4$, we may assume $|P \rightarrow Q| = 4$. As we could replace two consecutive edges

¹By Lemma 2.13, such an orientation has also a unique source.

along the same row/column by one edge, the shortest path will be zig-zagging and of the form as in Figure 4.7 below; there we only consider the 3×3 -subgraph spanned by this path. Denote this subgraph by G_{PQ} . Note that by Claims 1 and 2 there cannot be a path $Q \rightarrow P$ in G_{PQ} . (Of course, these two claims do not cover all possible options. But the others are easily observed.)

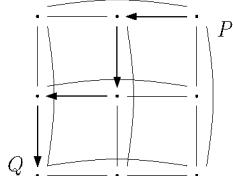


Figure 4.7: A shortest connecting path $P \rightarrow Q$ of length 4.

Neither P nor Q can be the source or the sink of G_{PQ} as otherwise we would either have a shorter path $P \rightarrow Q$ or some path $Q \rightarrow P$, something we have already ruled out.

So, for the 3×3 subgraph induced by P and Q we have essentially two options, as drawn in Figure 4.8 below. We see that A has to be either the unique sink or the unique source.

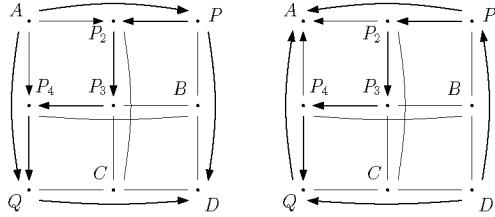


Figure 4.8: A shortest connecting path $P \rightarrow Q$ of length 4.

Case 1: A is unique source. $C \rightarrow Q$ would imply $P_2 \rightarrow C$ leading to a path of length 3 connecting P and Q , in contradiction to our assumption. So, we have $Q \rightarrow C$, which implies $P_3 \rightarrow C$ and hence $P_2 \rightarrow C$, nonetheless. But now we have a violation of the path condition, as it creates the forbidden subgraph of Figure 2.5(b).

Case 2: A is unique sink. $P \rightarrow B$ would imply $B \rightarrow P_4$ leading again to a path of length 3, this time connecting P and Q , in contradiction to our assumption. But $B \rightarrow P$ implies $B \rightarrow P_3$ and $B \rightarrow P_4$, creating the forbidden subgraph of Figure 2.5(b), thus, violating the path condition. \square

Proof Theorem 4.2 Claim 3 above tells us that the maximal length of the shortest cycle is 4 as there is an obvious path from each point to any other point on the same cycle. So, this cycle must form a 2×2 -subgrid, which is a contradiction. \square

Claim 3 established the fact, that the maximal length of the shortest path between two vertices P, Q in an admissible grid orientation is at most 3. For the proof, the validity of the path condition was crucial. In fact, dropping the path condition, we may construct a unique sink orientation of size $L \times R$ with a shortest path as long as $2(m - 1)$, $m := \min\{L + R\}$, see Figure 4.9.

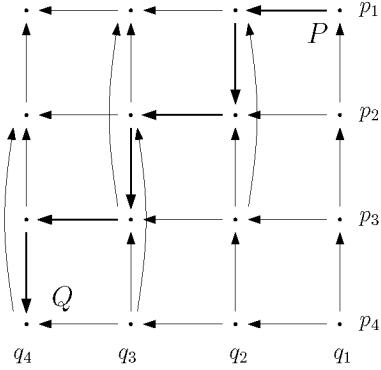


Figure 4.9: A shortest connecting path $P \rightarrow Q$ of length 6 in a 4×4 -grid.

In general, the grid orientation is defined as follows. Given the rows $\{p_1, \dots, p_L\}$ and columns $\{q_1, \dots, q_R\}$ which define the vertices (p_i, q_j) , $1 \leq i \leq L, 1 \leq j \leq R$, we have the edges

- $(p_i, q_{j_1}) \rightarrow (p_i, q_{j_2}), \quad \text{for all } 1 \leq i \leq m, 1 \leq j_1 < j_2 \leq m$
(horizontal edges)
- $(p_i, q_{i+1}) \rightarrow (p_{i+1}, q_{i+1}), \quad \text{for all } i, 1 \leq i \leq m - 1$
(vertical edges on the path)
- $(p_{i_1}, q_j) \rightarrow (p_{i_2}, q_j), \quad \text{for all } 1 \leq j \leq m, 1 \leq i_2 < i_1 \leq m,$
whenever $i_1 \neq i_2 + 1$ or $j \neq i_1$.
(remaining vertical edges)

It is not hard to see that this defines a unique sink orientation.

We claim that the shortest path from $P = (p_1, q_1)$ to $Q = (p_m, q_m)$ has length $2(m - 1)$. Observe that only $m - 1$ column-edges lead from some row p_i to row p_j with $i < j$. In fact, in all $m - 1$ cases, we have $j = i + 1$. Furthermore, no two of them lie in the same column. So, in order to get from P (in row p_1) to Q (in row p_m) we need to make use of all of them plus the $m - 1$ row-edges $(p_i, q_i) \rightarrow (p_i, q_i)$ that take us from the end of one possible column edge to the starting point of the next. Hence, $|P \rightarrow Q| = 2(m - 1)$, as claimed.

Observe, that for the acyclic case this is best possible: For a shortest path we may use only one edge from each row or column.

4.3 A Random Walk on Admissible Grid Orientations

Our successful analysis of RANDOM EDGE on $(d, d + 2)$ -polytopes in Chapter 3 relied on the ability to model its behaviour as a random process on a planar configuration of one line and n points.

In this section we will see that one can analyse RANDOM EDGE directly, that is, as a *Random Walk* on an admissible grid orientation.

With the terminology introduced at the end of Section 4.1, we define this random walk formally as Algorithm 4.3.

Algorithm 4.3 RANDOM WALK
on an admissible grid orientation G

```

1   $P \leftarrow$  arbitrary node of  $G$ ;
2   $\mathcal{N} \leftarrow \{Q \mid P \rightarrow Q\}$ ;
3  while  $\mathcal{N} \neq \emptyset$ 
4    do
5       $P \leftarrow_{\text{random}} \mathcal{N}$ ;
6       $\mathcal{N} \leftarrow \{Q \mid P \rightarrow Q\}$ ;
7  return  $P$ .
```

We will prove the following theorem:

Theorem 4.4 *Starting at an arbitrary node of an $L \times R$ admissible grid orientation, the expected number of steps needed to reach the sink by means of a random walk on the graph is at most*

$$O((\log n)(1 + \log m)) = O(\log^2 n),$$

where $n = L + R$ and $m = \min\{L, R\}$.

Note the similarity to Theorem 3.2. In the light of Observation 2.11 this is no coincidence. In fact, the analysis of Algorithm 4.3 can be based on the Analysis of the *Fast Process* in Section 3.2. The challenge lies in extracting and translating the decisive geometric properties there into properties of an admissible graph orientation here. Most interestingly, the analysis will depend crucially on the validity of the path condition. Practical experiments suggest that the upper bound of Theorem 4.4 holds for *any* unique sink orientations (not necessarily fulfilling the path condition). From our discussion below it becomes clear that in order to prove this conjecture an entirely different approach would be needed.

4.3.1 The Refined Index

There are $L - 1$ row-edges and $R - 1$ column-edges incident to each node P . We say that P has *refined index* $[r, s]$ ($\text{index}(P) = [r, s]$) if P has r outgoing row-edges and s outgoing column-edges.

Lemma 4.5 *Given some unique sink grid orientation G of size $L \times R$, the function*

$$\text{index} : G \rightarrow \{[r, s] \mid 0 \leq r < L, 0 \leq s < R\}$$

is a bijection between the coordinates of the vertices and their indices.

Proof We have to show that for each refined index there is exactly one node carrying it. Suppose otherwise. Then there are two nodes $P = (p_1, p_2)$ and $Q = (q_1, q_2)$, both with refined index $[r, s]$. Clearly, P and Q are neither lying in the same row nor in the same column, i.e. $p_i \neq q_i$, $i = 1, 2$. As $\langle P, Q \rangle$ must be a unique sink orientation we may assume, w.l.o.g., $P^{\rightarrow}(q_1, p_2)$ and $(p_1, q_2)^{\rightarrow}Q$, cf. Figure 4.10(a). As both P and Q have the same number of outgoing column-edges, there must be some row r , say, that is only reachable from Q but not from P , i.e. $Q^{\rightarrow}(r, q_2)$ and $(r, p_2)^{\rightarrow}P$. But this implies $(r, p_2)^{\rightarrow}(q_1, p_2)$, and the edges in the rows p_1, q_1 and r that link the columns p_2 and q_2 must all have the same orientation, cf. Figure 4.10(a).

Case 1: $(x, q_2)^{\rightarrow}(x, p_2)$ for $x = p_1, q_1, r$.

As P and Q have the same column index, there must be a column c , say, such that $P^{\rightarrow}(p_1, c)$ and $(q_1, c)^{\rightarrow}Q$. Figure 4.10(b) demonstrates the situation. Observe, that this results in a cycle isomorphic to the one in Figure 4.4. With Claim 2 in Section 4.2 we have proven that this is in contradiction to the unique sink property.

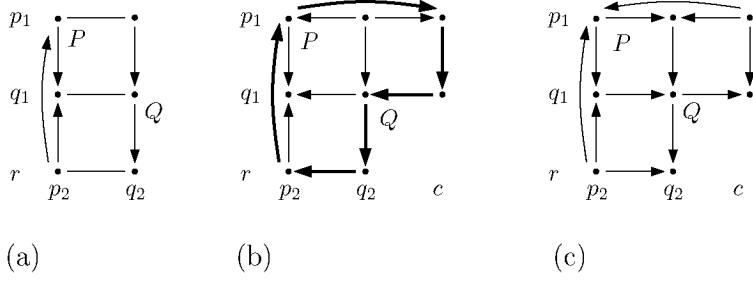


Figure 4.10: Suppose, there are nodes P and Q with equal refined index.

Case 2: $(x, p_2) \rightarrow (x, q_2)$ for $x = p_1, q_1, r$.

Again, there must be a column c , this time such that there are edges $(p_1, c) \rightarrow P$ and $Q \rightarrow (q_1, c)$. Two edge orientations follow due to the unique sink property: $(p_1, c) \rightarrow (q_1, c)$, $(p_1, c) \rightarrow (p_1, q_2)$, cf. Figure 4.10(c). Now, $(q_1, c) \rightarrow (r, c)$ leads to a contradiction as we cannot orient the edge connecting (r, c) and (r, p_2) to our satisfaction. Either $\langle (q_1, p_2), (r, c) \rangle$ or $\langle P, (r, c) \rangle$ will not have a unique sink. But setting $(r, c) \rightarrow (q_1, c)$ would imply $(r, q_2) \rightarrow (r, c)$, making it impossible to orient the edge connecting (r, c) and (p_1, c) without violating the unique sink condition either in $\langle P, (r, c) \rangle$ or in $\langle (p_1, q_2), (r, c) \rangle$. \square

Note that the refined index determines the h -vector (cf. Lemma 2.13): h_k is the number of nodes with exactly k incoming (i.e. $L + R - 2 - k$ outgoing) edges.

A corollary of Lemma 4.5 is, therefore, that the h -vector of unique sink orientations on grid graphs of size $L \times R$ is a function of L and R only.

4.3.2 Basic Properties

Lemma 4.6 For any two nodes $P = (p_1, p_2)$, $Q = (q_1, q_2)$ in a unique sink grid orientation G , we observe:

- (i) $q_1 \in (R\text{-below}(P) \cup \{p_1\}) \quad \wedge \quad q_2 \in (C\text{-below}(P) \cup \{p_2\}) \implies Q \in \text{reach}(P)$,
- (ii) If G is acyclic then $Q \in \text{reach}(P) \implies Q \in \text{below}(P)$
- (iii) $Q \notin \text{below}(P) \implies P \in \text{reach}(Q)$.

Proof All three are easy. Here is the proof for (ii): suppose, $Q \notin \text{below}(P)$, that is, $q_1 \notin \text{R}-\text{below}(P)$ and $q_2 \notin \text{C}-\text{below}(P)$. Then P is the sink of $\langle P, Q \rangle$. Hence, from any node in $\langle P, Q \rangle$, there is a path to P , in particular, $Q \rightarrow P$. This implies $Q \notin \text{reach}(P)$ by acyclicity. \square

Lemma 4.7 Given nodes $P = (p_1, p_2)$ and $Q = (q_1, q_2)$ with indices $\text{index}(P) = [r, s]$, $\text{index}(Q) = [x, y]$ where $r \geq x$, $s \geq y$. Then $Q \not\rightarrow P$.

Proof Suppose $q_1 \in (\text{R}-\text{below}(P) \cup \{p_1\})$ and $q_2 \in (\text{C}-\text{below}(P) \cup \{p_2\})$. By Lemma 4.6(i), then $P \rightarrow Q$ and the claim holds thanks to acyclicity.

Therefore, we may assume that P and Q neither share the same row nor the same column and that, w.l.o.g., $q_2 \notin \text{C}-\text{below}(P)$, i.e. $(p_1, q_2) \rightarrow P$.

We distinguish between two cases:

Case 1: $p_2 \notin \text{C}-\text{below}(Q)$, cf. Figure 4.11(a) with the thick edges determined. Either $P \rightarrow Q$ and the claim holds. Or we have edges $(q_1, p_2) \rightarrow P$ and $Q \rightarrow (p_1, q_2)$. Since P has at least as many outgoing edges to other rows as Q , there is a row r_1 with $P \rightarrow (r_1, p_2)$ and $(r_1, q_2) \rightarrow Q$. In this row, $(r_1, q_2) \rightarrow (r_1, p_2)$ by the unique sink property in the subgrid $\langle P, (r_1, q_2) \rangle$. But no matter how we now orient the edge between the nodes (r_1, p_2) and (q_1, p_2) , the unique sink property is always violated.

Case 2: $p_2 \in \text{C}-\text{below}(Q)$, cf. Figure 4.11(b), the thick edges being given. Since P has at least as many outgoing edges to other columns as Q , there must be a column c that is in $\text{C}-\text{below}(P) \setminus \text{C}-\text{below}(Q)$ i.e. we have some c with $P \rightarrow (p_1, c)$ and $(q_1, c) \rightarrow Q$. Either $P \rightarrow Q$ and the claim holds, or $(q_1, c) \rightarrow (p_1, c)$. By the unique sink property we have $(q_1, p_2) \rightarrow P$, $(p_1, q_2) \rightarrow (p_1, c)$, $Q \rightarrow (p_1, q_2)$. That is, $p_1 \in \text{R}-\text{below}(Q)$ and $q_1 \notin \text{R}-\text{below}(P)$. P has at least as many outgoing edges to other rows as Q , there must be (as in case 1) a row r_2 that is in $\text{R}-\text{below}(P) \setminus \text{R}-\text{below}(Q)$. But then the 3×3 -subgrid $\langle P, Q, (r_2, c) \rangle$ contains two nodes with refined index $[1, 1]$, namely P and Q . Thus, by Lemma 4.5, a contradiction. \square

Lemma 4.8 Given nodes $P = (p_1, p_2)$ and $Q = (q_1, q_2)$ with $Q \not\rightarrow P$. Then $\text{C}-\text{below}(Q) \subset \text{C}-\text{below}(P)$ or $\text{R}-\text{below}(Q) \subset \text{R}-\text{below}(P)$.

Proof We may assume that P and Q share neither a row nor a column, for otherwise $Q \not\rightarrow P$ implies $P \rightarrow Q$ and the lemma clearly holds. Suppose there is a node $R = (r_1, r_2)$ with $r_1 \in \text{R}-\text{below}(Q) \setminus \text{R}-\text{below}(P)$ and $r_2 \in \text{C}-\text{below}(Q) \setminus \text{C}-\text{below}(P)$. $Q \not\rightarrow P$ implies that we have $(p_1, r_2) \rightarrow (q_1, r_2)$

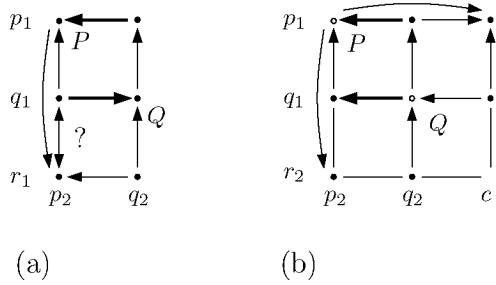


Figure 4.11: Consider nodes $P, \text{index}(P) = [r, s]$ and $Q, \text{index}(Q) = [x, y]$ with $r \geq x, s \geq y$. $Q \rightarrow P$ leads to a contradiction.

and $(r_1, p_2) \rightarrow (r_1, q_2)$. Since by Lemma 4.6(i) there is a path from Q to R , $Q \not\rightarrow P$ further implies that there is no path from R to P . Thus, we have $(p_1, r_2) \rightarrow (r_1, r_2)$ and $(r_1, p_2) \rightarrow (r_1, r_2)$. The resulting configuration is depicted in Figure 4.12. Clearly, $\langle P, R \rangle$ does not have a unique sink. \square

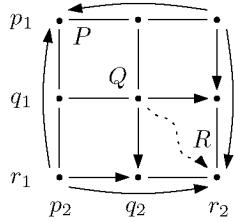


Figure 4.12: $Q \not\rightarrow P$ implies $C\text{-below}(Q) \subset C\text{-below}(P)$ or $R\text{-below}(Q) \subset R\text{-below}(P)$.

Corollary 4.9 Given nodes $P = (p_1, p_2)$ and $Q = (q_1, q_2)$ with $Q \not\rightarrow P$ and $p_1 \in R\text{-below}(Q)$ ($p_2 \in C\text{-below}(Q)$, respectively). Then $C\text{-below}(Q) \subset C\text{-below}(P)$ ($R\text{-below}(Q) \subset R\text{-below}(P)$, respectively).

Lemma 4.10 Given some grid orientation consisting of a single row or column with k nodes. Then a random walk (as defined in Algorithm 4.3) starting at an arbitrary node takes at most $\log_2 k$ steps.

Such a grid orientation is the complete graph with k nodes. We have studied this (trivial) case extensively in Section 3.1. So, there is nothing (left) to prove.

4.3.3 The Analysis

We follow closely the analysis of the *Fast Process* in Section 3.2, dividing the random walk into phases and estimate their number and their respective duration. Whenever possible, we will also use the same notation as in Section 3.2. In fact, some of the terminology we have introduced was motivated by the need to describe the (geometric) properties of the planar configuration of one line and n points as the combinatorial properties of an admissible grid orientation.

Recall Observation 2.11: each ℓ -edge corresponds to a vertex in the graph, there is an edge from one to another vertex if one can pivot directly from one to the other of the corresponding ℓ -edges. So, the rows of the grid can be identified with the points left of ℓ , the columns with the points on the right hand side of ℓ . In this light, the notions of R -below(P) and C -below(P) suddenly make sense: these denote those rows and columns that correspond to points lying below the ℓ -edge that is represented by the node P .

As before, we assume that G is an admissible grid orientation of size $L \times R$, $L + R = n$; assume further, w.l.o.g., that $L \geq R$.

For $0 < i \leq \lfloor \log_2 R \rfloor$, let $\lambda_i = (l_1^{(i)}, l_2^{(i)})$ denote the unique node with refined index $[2^{i-1}, 2^{i-1}]$, λ_0 is the sink $[0, 0]$. $\lambda_{\lfloor \log_2 R \rfloor + 1}$ is the source $[L-1, R-1]$. (So, if $R = L$ and equals a power of 2 then $\lambda_{\lfloor \log_2 R \rfloor + 1} = \lambda_{\lfloor \log_2 R \rfloor}$.) We say that a node P belongs to *phase i of the random walk* if $\lambda_{i+1} \rightarrow P$ and $\lambda_j \not\rightarrow P$, for all $j \leq i$ or if $\lambda_i = P$. We define the random variable X_i , $i = 0, \dots, \lfloor \log_2 R \rfloor$, as the number of nodes visited during phase i . As λ_0 is the sink of the graph, completion of phase 0 entails completion of the whole random walk. Hence, $Z = \sum_{i=0}^{\lfloor \log_2 R \rfloor} X_i$ is the random variable whose expectation we want to analyse.

We will show that $E(X_i) = O(\log n)$ for all i and, hence,

$$E(Z) = O((\log n)(1 + \log R)) = O(\log^2 n).$$

Analysis of a Single Phase Fix some i , $0 \leq i \leq \lfloor \log_2 R \rfloor$, set $k = 2^{i-1}$. We have node λ_{i+1} or some node in $\text{reach}(\lambda_{i+1}) \setminus \text{reach}(\lambda_i)$ and the phase starts. The phase ends whenever we reach the node λ_i or some node which is in the reach of λ_j , $j \leq i$.

Recalling Lemma 4.6 we see that for any node $P = (p_1, p_2)$ in phase i we have $(p_1, p_2) \in \text{below}(\lambda_{i+1})$ while $p_1 \notin \text{R-below}(\lambda_i)$ or $p_2 \notin \text{C-below}(\lambda_i)$. (For otherwise it would already belong to a lower phase.)

We further split phase i into *strokes*. A stroke starts after we have sampled a node with a new witness for being below λ_i or λ_{i+1} , or at the beginning of a new phase. The start of a new stroke also terminates the previous one.

In other words, a new stroke starts whenever we have moved *along a row* to a column in $\text{C-below}(\lambda_i) \cup \text{C-below}(\lambda_{i+1})$ or *along a column* to a row in $\text{R-below}(\lambda_i) \cup \text{R-below}(\lambda_{i+1})$ (or at the beginning of a phase).

If N is the number of strokes, then we can write $X := X_i$ as

$$X = Y_1 + Y_2 + \cdots + Y_N$$

where Y_j is the number of nodes visited during the j th stroke. Note that N itself is again a random variable. (For $j > N$ we set $Y_j = 0$.)

We will show that

- (i) $E(Y_j | j \leq N) = O(\log n)$ for all j , and
- (ii) $E(N) = O(1)$.

It follows that $E(X) = O(\log n)$:

$$\begin{aligned} E(X) &= \sum_{j=1}^{\infty} \overbrace{E(Y_j | j \leq N)}^{O(\log n)} \Pr(j \leq N) \\ &= O(\log n) \sum_{j=1}^{\infty} \Pr(j \leq N) \\ &= O(\log n) E(N). \end{aligned} \tag{4.1}$$

The nodes starting a new stroke can be grouped into different ‘categories’, depending on the associated witness. We will see that with a certain probability we move to a new category and come closer to our goal: to escape the current phase.

We argue as follows: *At any point of phase i , the following four claims hold:*

Claim 1 *The expected number of nodes visited until we reach either a new² row in R-below(λ_{i+1}) or a new column in C-below(λ_{i+1}) is at most $\log_2 n$.*

²We use ‘new’ as ‘new with respect to the current node’. With each pivot we either enter a new row or a new column —while either staying in the current column or in the current row.

Proof The current node $P = (p_1, p_2)$ is in $\text{reach}(\lambda_{i+1})$. By Lemma 4.6 we have $p_1 \in \text{R-below}(\lambda_{i+1})$ or $p_2 \in \text{C-below}(\lambda_{i+1})$. All nodes reachable from P must also be in $\text{reach}(\lambda_{i+1})$. W.l.o.g. assume, that we pivot to a point Q in the same row as P , $Q = (p_1, q_2)$. Then either $q_2 \in \text{C-below}(\lambda_{i+1})$ and we are done. Or we will reach a node with the desired property as soon as we have left the current row p_1 .

How long can we stay in p_1 ? If we have reached its local sink we have no alternative to choosing one of the outgoing column-edges. So the situation is almost like in the analysis of the graph with dimension $(1 \times n)$, see Lemma 4.10. In fact, there are two differences which can only improve our expectations: We terminate not only in the local sink but also when we reach a column which is in $\text{C-below}(\lambda_{i+1})$. Furthermore, in each step we choose out of all outgoing edges one at random, including the ones which take us away from the current row. Hence, the expected number of visited nodes is at most $\log n$. \square

Since any node sampled that has a new witness for being in $\text{below}(\lambda_i)$ starts a new stroke, this also establishes (i) from above: the expected number of iterations during a stroke is $O(\log n)$.

Claim 2 *Conditioned on the event that we sample a node in*

$$\text{below}(\lambda_{i+1}) \cup \text{below}(\lambda_i),$$

the node will have a witness for being in

$$\text{below}(\lambda_i)$$

with probability at least $\frac{1}{5}$.

Proof Suppose we are currently at node $P = (p_1, p_2)$ in phase i . We have $\lambda_i \not\rightarrow P$. By Lemma 4.8 and w.l.o.g. $\text{C-below}(\lambda_i) \subset \text{C-below}(P)$. Therefore, exactly $k = 2^{i-1}$ columns of those belonging to $\text{C-below}(P)$ are also in $\text{C-below}(\lambda_i)$. On the other hand, the sets $\text{R-below}(P) \cap (\text{R-below}(\lambda_{i+1}) \cup \text{R-below}(\lambda_i))$ and $\text{C-below}(P) \cap (\text{C-below}(\lambda_{i+1}) \cup \text{C-below}(\lambda_i))$ have combined at most $5k$ elements. Again, this holds by Lemma 4.8: since $\lambda_i \not\rightarrow \lambda_{i+1}$ by Lemma 4.7, we have $\text{R-below}(\lambda_i) \subset \text{R-below}(\lambda_{i+1})$ or $\text{C-below}(\lambda_i) \subset \text{C-below}(\lambda_{i+1})$. In addition, recall that $|\text{R-below}(\lambda_{i+1})| = |\text{C-below}(\lambda_{i+1})| = 2k$ and $|\text{C-below}(\lambda_i)| = |\text{R-below}(\lambda_i)| = k$. \square

Claims 1 and 2 combined assure that we reach a node below λ_i within an expected number of at most $10 \log_2 n$ steps.

So what happens after we see such a node $Q \in \text{below}(\lambda_i)$?

Two cases have to be distinguished, depending on whether Q 's witness for being below λ_i is also a witness for being below λ_{i+1} or not. First, we deal with the latter case:

Claim 3 *If $Q = (q_1, q_2)$ is the current node and*

$$\begin{aligned} \text{either } & q_1 \in R\text{-below}(\lambda_i) \setminus R\text{-below}(\lambda_{i+1}), q_2 \in C\text{-below}(\lambda_{i+1}) \\ \text{or } & q_2 \in C\text{-below}(\lambda_i) \setminus C\text{-below}(\lambda_{i+1}), q_1 \in R\text{-below}(\lambda_{i+1}), \end{aligned}$$

then the next node sampled which terminates the stroke will have a witness for being in

$$\text{below}(\lambda_{i+1}) \cap \text{below}(\lambda_i)$$

with probability at least $\frac{1}{5}$.

Proof Assume $q_1 \in R\text{-below}(\lambda_i) \setminus R\text{-below}(\lambda_{i+1})$; the other case can be discussed analogously. Since $\lambda_i \not\rightarrow Q$ we have $C\text{-below}(\lambda_i) \subset C\text{-below}(Q)$ (by Corollary 4.9).

Furthermore, we must have $C\text{-below}(Q) \subset C\text{-below}(\lambda_{i+1})$. For, suppose otherwise. Then there is a column c such that $Q \rightarrow (q_1, c)$ and $(l_1^{(i+1)}, c) \rightarrow \lambda_{i+1}$, implying the thick arrows in Figure 4.13.

Now $\lambda_{i+1} \rightarrow Q$ implies $Q \not\rightarrow \lambda_{i+1}$ and we must have $(l_1^{(i+1)}, c) \rightarrow (q_1, c)$. By the unique sink orientation in various subgrids, this further implies $(l_1^{(i+1)}, c) \rightarrow (l_1^{(i+1)}, q_2)$, $(l_1^{(i+1)}, q_2) \rightarrow Q$ and $(q_1, l_2^{(i+1)}) \rightarrow \lambda_{i+1}$ gives a contradiction to the unique sink property in the subgrid $\langle \lambda_{i+1}, (q_1, c) \rangle$, see Figure 4.13.

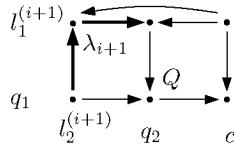


Figure 4.13: $C\text{-below}(Q) \subset C\text{-below}(\lambda_{i+1})$.

So, we have shown that $C\text{-below}(\lambda_i) \subset C\text{-below}(Q) \subset C\text{-below}(\lambda_{i+1})$.

When we sample the first node of a new stroke the current node must still be of the form (x, q_2) with $x \in R\text{-below}(\lambda_i) \setminus R\text{-below}(\lambda_{i+1})$. (Recall that we are in $\text{reach}(\lambda_{i+1})$. Staying in a row which is not in $R\text{-below}(\lambda_{i+1})$ we can only pivot towards columns in $C\text{-below}(\lambda_{i+1})$.)

Hence all columns in $C\text{-below}(\lambda_i)$ are at our disposal. The claim follows. \square

Claim 4 If $Q = (q_1, q_2)$ is the current node and

$$\begin{aligned} \text{either } & q_1 \in R\text{-below}(\lambda_i) \cap R\text{-below}(\lambda_{i+1}), \\ \text{or } & q_2 \in C\text{-below}(\lambda_i) \cap C\text{-below}(\lambda_{i+1}), \end{aligned}$$

then the next node sampled which terminates the stroke will have a new witness for being in

$$\text{below}(\lambda_i)$$

(that is, finish the phase) with probability at least $\frac{1}{5}$.

Proof W.l.o.g. assume $q_1 \in R\text{-below}(\lambda_i) \cap R\text{-below}(\lambda_{i+1})$. So, in particular, $q_1 \in R\text{-below}(\lambda_i)$ and $R\text{-below}(\lambda_i) \not\subset R\text{-below}(Q)$. Since $\lambda_i \not\rightarrow Q$ we may deduce $C\text{-below}(\lambda_i) \subset C\text{-below}(Q)$ by Lemma 4.8.

Suppose, there is some row r such that $Q^*(r, q_2)$ with $r \notin R\text{-below}(\lambda_i) \cup R\text{-below}(\lambda_{i+1})$. Since Q belongs to phase i there is neither a path from λ_i to Q nor from Q to λ_{i+1} , and we must have the situation as drawn in Figure 4.14.

Consider the subgrid $\langle Q, (r, l_2^{(i)}), (r, l_2^{(i+1)}) \rangle$: it is isomorphic to the forbidden subgraph from Figure 2.5(b) – a violation of Property 3 on page 25.³ Hence, we may deduce that $R\text{-below}(Q) \in R\text{-below}(\lambda_i) \cup R\text{-below}(\lambda_{i+1})$.

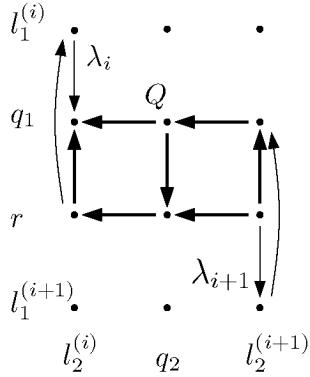


Figure 4.14: $R\text{-below}(Q) \setminus (R\text{-below}(\lambda_i) \cup R\text{-below}(\lambda_{i+1})) = \emptyset$

So, the first time we reach a new node that has a new witness for being below(λ_{i+1}) or below(λ_i) the preceding node must still have been in row q_1 .

³Note that this is the only time where this property is used in the entire analysis!

Since $C\text{-below}(\lambda_i) \subset C\text{-below}(Q)$ this implies that out of at most $5k$ choices of nodes that terminate the stroke we have at least k good ones that actually terminate the phase. \square

Claim 4 entails that once we have chosen a node with one witness for being below λ_i and λ_{i+1} , then – with probability at least $\frac{1}{5}$ – the next node that has (another) witness for being below λ_i or λ_{i+1} will terminate the phase.

This completes the argument as the remaining part is identical to the final part of the analysis of the *Fast Process* in Section 3.2 — apart from only slight differences that are irrelevant for the analysis: whenever we speak of sampling a new point there, we are referring to choosing the next node here. Furthermore, the nodes are classified depending on whether the new witness for being below λ_i or below λ_{i+1} lies in

- Class 0: $R\text{-below}(\lambda_{i+1}) \setminus R\text{-below}(\lambda_i)$ or $C\text{-below}(\lambda_{i+1}) \setminus C\text{-below}(\lambda_i)$,
- Class 1: $R\text{-below}(\lambda_i) \setminus R\text{-below}(\lambda_{i+1})$ or $C\text{-below}(\lambda_i) \setminus C\text{-below}(\lambda_{i+1})$,
- Class 2: $R\text{-below}(\lambda_i) \cap R\text{-below}(\lambda_{i+1})$ or $C\text{-below}(\lambda_i) \cap C\text{-below}(\lambda_{i+1})$.

Hence, we have shown that the expected number of pivots in a single phase is bounded by $310 \log_2 n$, and the theorem follows. \square Theorem 4.4 \square

4.4 Pseudo Realizability of Admissible Grid Orientations

A *uniform pseudoconfiguration of points* of rank 3 is a pair (\mathcal{A}, S) where S is a planar point set of size n , and \mathcal{A} is an arrangement of $\binom{n}{2}$ pseudolines $\lambda_{ss'}$ through all pairs of points $s, s' \in S$. Uniform means that no three points in S lie on the same pseudoline. Let ℓ be a vertical line which is disjoint from S and from all intersections of pseudolines. We refer to such a planar pseudoconfiguration as the configuration (\mathcal{A}, S, ℓ) .

Obviously, this is a generalization⁴ of the scenario *One line and n points* which was discussed in Section 2.2. We will, therefore, reuse the notation introduced there. Each configuration (\mathcal{A}, S, ℓ) still induces an admissible grid orientation, *cf.* Figure 4.15. (Recall Observation 2.11 and Section 2.5.) The orientation is determined by the relative order of pseudolines λ_{pq} along the vertical line ℓ .

We call an admissible grid orientation *pseudo realizable* whenever it is induced by some configuration (\mathcal{A}, S, ℓ) .

⁴Section 4.5 is dedicated to the proof, that it is a *proper* generalization.

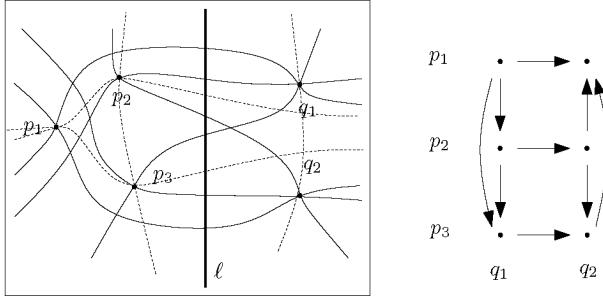


Figure 4.15: One line and a pseudoconfiguration of n points induces an admissible grid orientation.

Theorem 4.11 Every admissible grid orientation is pseudo realizable, i.e. induced by a line ℓ , n points, and pseudolines through the pairs of points whose relative orders along ℓ determine the orientation.

The proof of this theorem needs some preparation. We will not construct a configuration (\mathcal{A}, S, ℓ) directly; rather, we define so-called hyperline sequences, from which we can deduce the existence of suitable S , \mathcal{A} and ℓ by using a connection between hyperline sequences and oriented matroids via abstract determinant functions.

Hyperline sequences Let S be a set of n distinct points in general position, indexed by $E_n = \{1, \dots, n\}$. Adding to E_n the new elements $\{\bar{i} | i \in E_n\}$, we get the *signed index set* \bar{E}_n (with $\bar{s} = s, \forall s \in \bar{E}_n$). $\text{conv } S$ shall denote the convex hull of S .

Rotating an oriented line in counterclockwise order around $s_i, i \in E_n$, and looking at the successive positions where it coincides with lines defined by pairs of points (s_i, s_j) defines the *hyperline sequence* π_i over \bar{E}_n , cf. [BMS01]. If a point s_j is encountered by the rotating line in positive direction from s_i , it will be recorded as a positive index j , otherwise as a negative index \bar{j} . The whole sequence is recorded in the order induced by the rotating line, and an arbitrary half-period is chosen to represent it, cf. Figure 4.16 for a simple example.

Note that the hyperline sequence of a point which belongs to the convex hull of the point configuration has a half-period consisting only of positive indices. In other words, it is represented by some permutation of the other points.

Similarly, we can define for any additional point s' the hyperline sequence of s' with respect to S . If s' is outside $\text{conv } S$, and not on any of the lines defined by two points of S , this sequence is represented by a permutation $\sigma \in \mathcal{S}_n$.

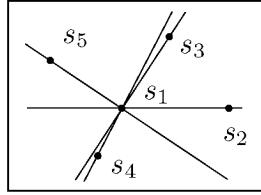


Figure 4.16: $\pi_1 = 2\bar{3}45$.

The rule to create a set of hyperline sequences for a point set can be carried over to pseudoconfiguration of points: The hyperline sequence π_i for a point $s_i, i \in E_n$ describes the order in which we encounter the pseudolines passing through s_i and some other point s_j in counterclockwise order around s_i .

Abstract determinant functions Given an abstract set of hyperline sequences (not necessarily coming from a point configuration with pseudolines), we want to deduce orientations for triples α, β, γ . For this, we define $\chi : \binom{E_n}{3} \rightarrow \{-1, 1\}$ (partially) by setting $\chi(\alpha, \beta, \gamma) := 1$ if β and γ occur in this order in π_α , and $\chi(\alpha, \beta, \gamma) := -1$ if β and γ occur in reverse order in π_α . We can write this as

$$\chi(\alpha, \beta, \gamma) := \begin{cases} 1, & \text{if } \alpha : ..\beta..\gamma.., \\ -1, & \text{if } \alpha : ..\gamma..\beta.. \end{cases}$$

We also define $\chi(\alpha, \gamma, \beta) = -\chi(\alpha, \beta, \gamma)$, obtaining values $\chi(\alpha, \beta, \gamma)$ for all distinct $\alpha, \beta, \gamma \in E_n$. We call χ consistent if $\chi(\alpha, \beta, \gamma) = \chi(\sigma(\alpha)\sigma(\beta)\sigma(\gamma))$ exactly for the even permutations $\sigma \in \mathcal{S}_3$. In this case, we call χ an *abstract determinant function* [BMS01], and we say that the hyperline sequence admits an abstract determinant function.

Proof Theorem 4.11 It has been shown in [BMS01] that an abstract determinant function fulfills the axioms of a *rank-3-chirotope*. Oriented matroid theory [BLW⁺93] then guarantees that this chirotope has a representation as a pseudoconfiguration of points which is consistent with the chirotope information (and hence with the hyperline sequence). This is, that there are points $S = \{s_\alpha, s_\beta, s_\gamma, \dots\}$ and oriented pseudolines \mathcal{A} through pairs of points such

that $\chi(\alpha, \beta, \gamma)$ is positive (negative) if and only if s_γ is to the left (to the right) of the pseudoline through s_α and s_β .

The challenge lies in defining an abstract determinant function for a given grid orientation G .

To simplify notation, we will assume that the grid graph has L rows, associated with indices $\mathcal{L} = \{i, j, k, \dots\}$, and R columns, indexed with capital letters $\mathcal{R} = \{I, J, K, \dots\}$. We use Greek letters $\alpha, \beta, \gamma, \dots$ for any other indices. Thus, nodes correspond to pairs (i, I) , and we write $v \rightarrow w$ to indicate the fact that there is a directed edge from node v to node w .

We shall construct point sets $S_L = \{p_i, p_j, p_k, \dots\}$ and $S_R = \{q_I, q_J, q_K, \dots\}$ ($S_L \cup S_R = S$ of sizes L and R ($L + R = n$), respectively, separated by a vertical line ℓ). Along with this we will have an arrangement \mathcal{A} of $\binom{n}{2}$ oriented pseudolines $\lambda_{ss'}$ through all pairs of points $s, s' \in S$.

But first, we derive a set of hyperline sequences. Those will then define an abstract determinant function whose values on ‘mixed’ triples are consistent with G , meaning that

$$\begin{aligned} (i, I) \rightarrow (i, J) &\iff \chi(i, J, I) > 0 \\ &\iff q_J \text{ is to the right of } \lambda_{p_i, q_I}, \end{aligned} \quad (4.2)$$

$$\begin{aligned} (i, I) \rightarrow (j, I) &\iff \chi(I, i, j) > 0 \\ &\iff p_j \text{ is to the right of } \lambda_{p_i, q_I}. \end{aligned} \quad (4.3)$$

This gives us the desired relation between the original grid orientation and the pair (S, \mathcal{A}) constructed from the chiotope χ on S .

The requirements (4.2) and (4.3) already determine the order in which the indices of \mathcal{R} must appear in the sequences $\pi_i, i \in \mathcal{L}$ (and vice versa).

Namely, for every index $i \in \mathcal{L}$, the graph gives us an ordering I_1, \dots, I_t of \mathcal{R} such that

$$(i, I_1) \rightarrow (i, I_2) \rightarrow \dots \rightarrow (i, I_R).$$

Restricted to \mathcal{R} , the sequence π_i must then be representable by the following half-period:

$$i : I_R I_{R-1} \dots I_1.$$

Similarly, for an index $I \in \mathcal{R}$ such that

$$(i_1, I) \rightarrow (i_2, I) \rightarrow \dots \rightarrow (i_L, I),$$

we obtain the restriction

$$I : i_1 i_2 \dots i_L.$$

It remains to insert the residual indices into the sequences in a consistent way. We describe how to do this for sequences $\pi_i, i \in \mathcal{L}$; the other case is symmetric.

Lemma 4.12 *Assume that $i : I_R I_{R-1} \dots I_1$ is the partial list for i , and for some $j \in \mathcal{L}, j \neq i$ we have $(i, I_R) \rightarrow (j, I_R)$. Then there is a unique index $\tau_{ij} \in \{1, \dots, R\}$ such that*

$$\begin{aligned} (i, I_t) &\rightarrow (j, I_t), & t \geq \tau_{ij}, \\ (i, I_t) &\leftarrow (j, I_t), & t < \tau_{ij}. \end{aligned}$$

Similarly, if $(i, I_R) \leftarrow (j, I_R)$, there is a unique index $\tau_{ij} \in \{1, \dots, R\}$ such that

$$\begin{aligned} (i, I_t) &\leftarrow (j, I_t), & t \geq \tau_{ij}, \\ (i, I_t) &\rightarrow (j, I_t), & t < \tau_{ij}. \end{aligned}$$

The lemma easily follows from the fact that no forbidden subgrid (as described in Figure 2.5(b)) exists and all 2×2 subgrids have unique sinks. This implies that the orientation of the edge connecting (i, I_t) and (j, I_t) can change at most once as we let t decrease from R to 1.

The lemma points out a canonical way to insert j into π_i : in the first case (i.e. $(i, I_R) \rightarrow (j, I_R)$), we get

$$i : I_R \dots I_\tau \bar{j} I_{\tau-1} \dots I_1,$$

in the second case we obtain

$$i : I_R \dots I_\tau j I_{\tau-1} \dots I_1.$$

Doing this for all j , we obtain the half-period $h(\pi_i)$ representing the hyperline sequence π_i which is complete and unique up to the order of elements j, k that give rise to the same value of $\tau = \tau_{ij} = \tau_{ik}$ in the lemma.

From now on, we will always refer to the half-period $h(\pi_i)$ even though we frequently just speak of ‘the half-period’ (of element i).

Recall that $\chi(i, j, k)$ can be read off from the order of the elements j, k in the sequence π_i . To prove that the partial map χ is extendible to an abstract determinant function we therefore need to establish the following claims:

Claim 1 Whenever the values are determined by the construction above, the orders of elements j, k in π_i , i, k in π_j , and i, j in π_k are consistent.

Claim 2 Whenever elements j, k give rise to the same value of τ , $\tau = \tau_{ij} = \tau_{ik}$, their order in π_i can be determined in a consistent way.

Observe that $\tau_{ij} = \tau_{ji}, \forall i, j$. Furthermore, whenever an element j is included into the half-period of some element i as \bar{j} , the half-period of j will contain i and vice versa. More precisely, j is represented by a positive index in i 's half-period if and only if the sink of the $2 \times R$ subgrid consisting of the two rows i and j lies in row i . Extending this argument, we can deduce the following lemma:

Lemma 4.13 Consider the half-periods representing the hyperline sequences of the indices in \mathcal{L} . For any $\alpha, 0 \leq \alpha < |\mathcal{L}|$, there exists a unique index $i = i(\alpha) \in \mathcal{L}$ whose half-period contains exactly α positive indices of \mathcal{L} . (The analog property holds for the sequences of the indices \mathcal{R} .)

Proof Consider the global sink of the graph. All edges point to it, in particular, the ones connecting it to the nodes of the same column. Thus, the other indices in \mathcal{L} are represented by positive indices in the sink row's half-period. Deleting this row, we get an induced subgraph which again has a global sink. In the half-period of the index representing the according row, we find all the other indices of \mathcal{L} represented by positive indices – with the exception of the one corresponding to the original global sink. Reapplying this argument the lemma becomes obvious. \square

Proof Claim 1 Without loss of generality, we may assume that j and k are represented by positive indices in the half-period representing the hyperline sequence of i and that k appears as positive index in $h(\pi_j)$. (So, negative indices represent i and j in the half-periods of k and k in j 's half-period.) Furthermore, as the other case can be dealt with in very much the same way, we assume $\tau_{ij} > \tau_{ik}$, i.e. $\chi(i, j, k) = 1$. The half-period for i is then of the form: $i : A_{ij}B_ikC_i$, where

$$\begin{cases} A_i = \{I \in S_R | (j, I) \rightarrow (i, I), (k, I) \rightarrow (i, I)\}, \\ B_i = \{I \in S_R | (k, I) \rightarrow (i, I) \rightarrow (j, I)\}, \\ C_i = \{I \in S_R | (i, I) \rightarrow (j, I), (i, I) \rightarrow (k, I)\}, \end{cases} \quad (4.4)$$

and $S_R = A_i \cup B_i \cup C_i$. By our assumption, B_i cannot be empty. So let J be an index in B_i . We will show that whenever the order of i, k in π_j and

i, j in π_k is given, it is consistent with the order of j, k in π_i . This means $\chi(j, i, k) \neq 1$, $\chi(k, j, i) \neq 1$.

So, assume for a contradiction $\chi(j, i, k) = 1$, $j : A_j k B_j \bar{i} C_j$, and $S_R = A_j \cup B_j \cup C_j$. Hence, index J from above must be in one of the sets A_j, B_j or C_j . Since for any index $I \in A_j \cup B_j$ $(j, I) \rightarrow (i, I)$ (implying $I \in A_i$), and for any index $I \in C_j$ $(j, I) \rightarrow (k, I)$ (implying $I \notin B_i$) must hold, this gives a contradiction to B_i being non-empty.

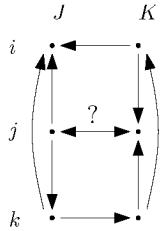


Figure 4.17: The case $\chi(k, j, i) = 1$.

Assuming $\chi(k, j, i) = 1$ also leads us to a contradiction: The half-period for k would be of the form $k : A_k \bar{j} B_k \bar{i} C_k$.

$S_R = A_k \cup B_k \cup C_k$ and

$$\begin{cases} A_k = \{I \in S_R | (k, I) \rightarrow (i, I), (k, I) \rightarrow (j, I)\}, \\ B_k = \{I \in S_R | (j, I) \rightarrow (k, I) \rightarrow (i, I)\}, \\ C_k = \{I \in S_R | (i, I) \rightarrow (k, I), (j, I) \rightarrow (k, I)\}. \end{cases}$$

Considering Equation 4.4, we see that $A_i \supseteq B_k = \{I \mid (j, I) \rightarrow (k, I) \rightarrow (i, I)\}$ and $A_k \supseteq B_i = \{I \mid (k, I) \rightarrow (i, I) \rightarrow (j, I)\}$. Taking a representative from each set, J from $B_k \subseteq A_i$ and K from $B_i \subseteq A_k$ say, we see – as shown in Figure 4.17 – that however the edge $(j, J) - (j, K)$ is directed, we always have a subgraph which does not have a unique sink. Hence, a contradiction.

Claim 1 \square

For the proof of Claim 2 we need yet another lemma:

Lemma 4.14 *Whenever elements j, k give rise to the same value of τ , $\tau_{ij} = \tau_{ik}$, then in both sequences π_j and π_k the respective order of i, k and i, j is either undetermined or determined. Moreover, in the latter case, the indices j and k are either both positive or both negative in the half-period of i .*

Proof If in all three sequences π_i , π_j and π_k the relative order of j, k, i, k and j, k is not determined there is nothing to prove.

So consider the case, where those values $\chi(i, j, k)$, $\chi(j, k, i)$, $\chi(k, i, j)$ that are determined all equal $+1$. (The discussion of the other case runs analogously.) Renaming i, j and k , if necessary, we may assume that their half-periods are of the form $i : A_{ij}B_{ik}C_i$, $j : A_{j\bar{i}}B_{jk}C_j$ and $k : A_{k\bar{i}}B_{kj}C_k$. The statement of the lemma can then be restated as: B_j is nonempty if and only if B_i or B_k is nonempty.

Suppose, there is some $I \in B_j$. Then either $(i, I) \rightarrow (k, I) \rightarrow (j, I)$ (implying $I \in B_k$) or $(k, I) \rightarrow (i, I) \rightarrow (j, I)$ (implying $I \in B_i$). The other direction is equally simple to see, as $B_i \subset B_j$ and $B_k \subset B_j$. \square

Proof Claim 2 As Lemma 4.14 tells us, we need to distinguish between two cases. Either, in all three hyperline sequences π_i , π_j and π_k the respective order of j, k, i, k , and i, j is not determined. (This case will be discussed later.) Or two such orders are induced by the according hyperline sequences but not the third. This latter sequence can only be one in which the indices concerned are either both positive or both negative, i.e. π_i or π_k .

We may assume that π_i is the sequence with $\tau_{ij} = \tau_{jk}$ and j, k are both positive. (By flipping all column edges we may change the sign of all row-indices \mathcal{L} .) We would like to impose the order of j and k in π_i (and thus to define the value $\chi(i, j, k)$) such that it is consistent with the values $\chi(j, k, i) = \chi(k, i, j)$ given by π_j and π_k .

If only two indices j, k are involved, this is certainly possible. However, in case three or more indices in π_i have the same τ we need to exclude the possibility that a cycle is induced.

Suppose that for the half-period $h(\pi_i)$ of the index i the elements k_0, \dots, k_{n-1} give rise to the same value of τ , $\tau = \tau_{ik_0} = \tau_{ik_1} = \dots = \tau_{ik_{n-1}}$ while the hyperline sequences of the other elements determine the values $\chi(k_0, k_1, i)$, $\chi(k_1, k_2, i), \dots, \chi(k_{n-1}, k_0, i)$. Then, by Lemma 4.14, we may assume that k_0, \dots, k_{n-1} are all positive in $h(\pi_i)$.

The values $\chi(k_j, k_{j+1}, i)$ determine an order on the pairs k_i, k_j in $h(\pi_i)$. We claim that this is a partial order whose extension we can use to define the values $\chi(i, k_j, k_{j+1})$ consistently.

Assume for a contradiction, that

$\chi(k_0, k_1, i) = \chi(k_1, k_2, i) = \dots = \chi(k_{n-1}, k_0, i) = +1$,
thus inducing the subsequences: $i : ..k_0k_1\dots k_{n-1}..$ and $i : ..k_{n-1}k_0..$ that

exclude each other. Furthermore assume that n is smallest possible. (Note that $n \geq 3$, by Claim 1.)

Then the set S_R can be partitioned into $S_R = A_i \cup C_i$ where

$$\begin{cases} A_i = \{I \in S_R \mid (k_j, I) \rightarrow (i, I) \forall j\}, \\ C_i = \{I \in S_R \mid (i, I) \rightarrow (k_j, I) \forall j\}. \end{cases} \quad (4.5)$$

This means that the vertices in row i are either the source or the sink with respect to the column they belong to.

By Lemma 4.13, index i is negative in all half-periods $h(\pi_{k_j})$, $j = 0, \dots, n-1$. To define the aforementioned values $\chi(k_j, k_{j+1}, i) = +1$, $h(\pi_{k_j})$ is therefore either of the form $k_j : \dots \bar{k}_j \dots$ or $k_j : \dots \bar{k}_{j+1} \dots \bar{i} \dots$. We say that $h(\pi_{k_j})$ is of type T_{\oplus} or T_{\ominus} , respectively.⁵

Suppose first that we have hyperline sequences of both types. Then there must be a j , such that $h(\pi_{k_j})$ is of type T_{\oplus} and $h(\pi_{k_{j+1}})$ is of type T_{\ominus} . More specifically, there is a j such that $k_j : A_{k_j} \bar{i} B_{k_j} k_{j+1} C_{k_j}$ and $k_{j+1} : A_{k_{j+1}} \bar{k}_{j+2} B_{k_{j+1}} \bar{i} C_{k_{j+1}}$, where

$$\begin{cases} A_{k_j} = \{I \in S_R \mid (k_{j+1}, I) \rightarrow (k_j, I) \rightarrow (i, I)\}, \\ B_{k_j} = \{I \in S_R \mid (i, I) \rightarrow (k_{j+1}, I) \rightarrow (k_j, I)\}, \\ C_{k_j} = \{I \in S_R \mid (i, I) \rightarrow (k_j, I) \rightarrow (k_{j+1}, I)\}, \end{cases} \quad (4.6)$$

$$\begin{cases} A_{k_{j+1}} = \{I \in S_R \mid (k_{j+1}, I) \rightarrow (k_{j+2}, I) \rightarrow (i, I)\}, \\ B_{k_{j+1}} = \{I \in S_R \mid (k_{j+2}, I) \rightarrow (k_{j+1}, I) \rightarrow (i, I)\}, \\ C_{k_{j+1}} = \{I \in S_R \mid (i, I) \rightarrow (k_{j+2}, I) \rightarrow (k_{j+1}, I)\}, \end{cases} \quad (4.7)$$

and $S_R = A_{k_j} \cup B_{k_j} \cup C_{k_j} = A_{k_{j+1}} \cup B_{k_{j+1}} \cup C_{k_{j+1}}$. Note that we already used the implications of Equations 4.5.

Equations 4.6 and 4.7 imply that $A_{k_j} = A_{k_{j+1}} \cup B_{k_{j+1}}$ and $B_{k_j} \cup C_{k_j} = C_{k_{j+1}}$. By our assumption, B_{k_j} and $B_{k_{j+1}}$ cannot be empty. So, there exist indices $J, K \in S_R$ with

$$J \in B_{k_{j+1}} \subset A_{k_j} \text{ and } (k_{j+2}, J) \rightarrow (k_{j+1}, J) \rightarrow (k_j, J) \rightarrow (i, J), \quad (4.8)$$

$$K \in B_{k_j} \subset C_{k_{j+1}} \text{ and } (i, K) \rightarrow (k_{j+2}, K) \rightarrow (k_{j+1}, K) \rightarrow (k_j, K). \quad (4.9)$$

In the half-period of the hyperline sequence π_{k_j} , where would we find the index k_{j+2} ? Requirement 4.9 implies that k_{j+2} comes after K whenever it is

⁵We compute the indices of k modulo n .

represented by a positive index. For Requirement 4.8 to hold, k_{j+2} needs to come before J whenever it is negative. This tells us that $h(\pi_{k_j})$ is either of the form $k_j : \dots k_{j+2} \dots J \dots \bar{i} \dots K \dots k_{j+1} \dots$, or of the form $k_j : \dots J \dots \bar{i} \dots K \dots \{k_{j+1} k_{j+2}\} \dots$ (where the relative order of k_{j+1} and k_{j+2} is still undefined). In either case, $\chi(k_j, k_{j+2}, i) = 1$. Therefore, we may delete column k_{j+1} and the hyperline sequences $\pi_{k_0}, \dots, \pi_{k_j}, \pi_{k_{j+2}}, \dots, \pi_{k_{n-1}}$ still induce a cycle on the order of the elements $k_0, \dots, k_j, k_{j+2}, \dots, k_{n-1}$ in π_i . As we chose n to be smallest possible, this implies that we must have had $n = 3$. So there are $j := k_{j_1}$ and $k := k_{j_2}$ with $\chi(j, k, i) = -\chi(k, i, j)$. A contradiction to Claim 1.

So we may assume that *all* half-periods $h(\pi_{k_0}), \dots, h(\pi_{k_{n-1}})$ are of the same type. Suppose they are of type T_\oplus , i.e. $k_j : A_{k_j} \bar{i} B_{k_j} k_{j+1} C_{k_j}$ with $S_R = A_{k_j} \cup B_{k_j} \cup C_{k_j}$. Equations 4.6 hold for all j . Therefore, all sets A_{k_j} are identical, let $A := A_{k_j}, \forall j$.

Recall that the representing half-periods were defined in such a way that the sets A_{k_j} are nonempty, cf. Lemma 4.12. So, there is an index $J \in A$ for which Equations 4.6 hold $\forall j$, that is, $(k_{n-1}, J) \rightarrow (k_{n-2}, J) \rightarrow \dots \rightarrow (k_1, J) \rightarrow (k_0, J) \rightarrow (k_{n-1}, J)$. But this is a cycle in row J — in contradiction to acyclicity of admissible grid orientations.

Analogously, the assumption that all half-periods are of type T_\ominus can be lead to a contradiction.

So we showed that if at least two of the values $\chi(i, j, k), \chi(j, k, i), \chi(k, i, j)$ are defined, the remaining one can be imposed such that they are all equal.

Finally, consider all triples i, j, k for which no constraint for the order of j and k in π_i is given. There, we impose the generic order. Claim 2 \square

We have extended the partial chirotope defined by the admissible grid orientation G to the rank-3-chirotope χ on S . Thus, it has a representation as a pseudoconfiguration of points (\mathcal{A}, S) , and S is of the form $S = S_L \dot{\cup} S_R$ where the points in S_L and in S_R represent the rows \mathcal{L} and the columns \mathcal{R} of G , respectively.

What remains to be shown is the existence of a vertical line ℓ separating S_L and S_R . For this, we view (\mathcal{A}, S) as the oriented matroid \mathcal{M} making us the standard machinery for oriented matroids available. In terminology and notation we follow [BLW⁺93]. Then a separating pseudoline, say ℓ' , exists if there is a covector $Y = (\{i, j, k, \dots\}, \{I, J, K, \dots\})$. By applying a suitable homeomorphic transformation we can map $\mathcal{A} \cup \{\ell'\}$ to an isomorphic arrangement $\mathcal{A}' \cup \{\ell\}$ where ℓ is a vertical line.

Let i be the unique element for which all other elements $k \in \mathcal{L}$, are represented by negative indices in the half-period $h(\pi_i)$. Let j be the first of those indices in $h(\pi_i)$.⁶ Let I and J be the analog indices in \mathcal{R} .

Then we have $\chi(i, j, k) = +$ for all $k \in \mathcal{L} \setminus \{i, j\}$ and $\chi(i, j, K) = -$ for all $K \in \mathcal{R}$, giving us the cocircuit $X_{ij} = (\{k, l, \dots\}, \{I, J, K, L, \dots\})$. Similarly, we may deduce $X_{IJ} = (\{i, j, k, l, \dots\}, \{K, L, \dots\})$. The composition of X_{ij} and X_{IJ} is the covector $X_{ij} \circ X_{IJ} = (\{i, j, k, \dots\}, \{I, J, K, \dots\})$, which is just the covector Y from above.

Thus, we have proved that there exists a representation (\mathcal{A}, S, ℓ) for the chirotope χ on S where χ is the extension of the partial chirotope induced by some admissible grid orientation G . Theorem 4.11 \square

Remark: With hindsight we see that the partial function χ as defined directly by the admissible grid orientation constituted a partial chirotope. Note that one can check this easily directly: We only need to verify the Grassmann-Plücker-Relations (Axiom 2 in Definition 2.14), as the other axioms are obviously fulfilled.

But if there are indices $\alpha, \beta, \gamma, \delta$ and ϵ such that

$$\begin{aligned} & \{\chi(\alpha, \beta, \gamma)\chi(\alpha, \delta, \epsilon), -\chi(\alpha, \beta, \delta)\chi(\alpha, \gamma, \epsilon), \chi(\alpha, \beta, \epsilon)\chi(\alpha, \gamma, \delta)\} \\ &= \{-1, +1\} \end{aligned}$$

then there exists a forbidden subgrid in G (or the unique sink property is violated), see Figure 4.18.

Equally, we see that if there is a forbidden subgrid, w.l.o.g. induced by rows i, j, k and columns J, K , then the Grassmann-Plücker-Relations for the elements i, j, k, I, J are not fulfilled.

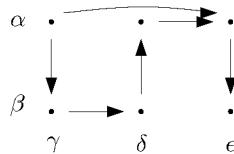


Figure 4.18: $\chi(\alpha, \beta, \gamma) = \chi(\alpha, \delta, \epsilon) = -\chi(\alpha, \beta, \delta) = \chi(\alpha, \gamma, \epsilon) = \chi(\alpha, \beta, \epsilon) = \chi(\alpha, \gamma, \delta)$ implies the existence of the forbidden subgrid.

⁶Note that, by Lemma 4.12, index i must exist and j will be very first index in $h(\pi_i)$.

4.5 What about ‘Proper’ Realizability?

Not all admissible grid orientations are ‘properly’ realizable. We will see below that only the structure of $2 \times m$ gridgraphs, where m is an arbitrary positive number, is simple enough to guarantee realizability.

Deriving a nonrealizable admissible grid orientation. It is a well-known fact that every oriented matroid consisting of less than 9 elements is realizable, *c.f.* [BLW⁺93]. Having just 9 elements, the example we are about to present is, therefore, smallest possible.⁷

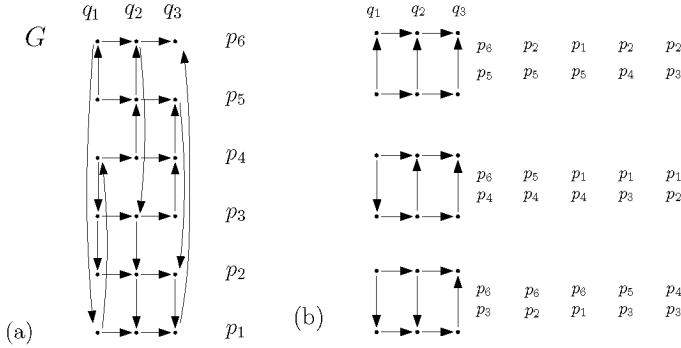


Figure 4.19: The 6×3 grid orientation G is admissible.

Consider the 6×3 grid orientation G of Figure 4.19(a).⁸ Obviously, it is an acyclic unique sink orientation. We only have to convince ourselves that it does not contain the forbidden subgrid. Clearly, it could only ‘hide’ in a 2×3 subgrid. But having examined all pairs of rows p_i, p_j in Figure 4.19(b) we know that this is not the case. So, G is admissible, hence pseudo realizable (by Theorem 4.11).

So assume further that it can be realized by a point configuration \mathcal{S} , $\mathcal{S} = \{p_1, \dots, p_6, q_1, q_2, q_3\}$ and a vertical line ℓ separating each pair p_i, q_j such that (\mathcal{S}, ℓ) induces G . Denote p_1, \dots, p_6 and q_1, q_2, q_3 by \mathcal{S}_L and \mathcal{S}_R , respectively, such that $\mathcal{S} = \mathcal{S}_L \cup \mathcal{S}_R$.

The following hyperline sequences are witnessed: $\pi_{q_1} : p_5 p_6 p_1 p_4 p_3 p_2$, $\pi_{q_2} : p_4 p_5 p_6 p_3 p_2 p_1$ and $\pi_{q_3} : p_3 p_4 p_5 p_2 p_1 p_6$. Points p_1, \dots, p_6 all witness the sequence $q_3 q_2 q_1$.

⁷We are defining a variant of the non-Pappus matroid, *c.f.* [Rin56], [BLW⁺93, Section 8.3].

⁸For better visibility, we omitted all directed edges whose direction is non-ambiguous.

Using the construction derived from Lemma 4.12, we can complete these sequences by including the points belonging to the same set as the observing point. We have:

$$\begin{aligned}\pi_{p_1} &: q_3 \bar{p}_6 q_2 p_2 p_3 p_4 q_1 p_5, & \pi_{p_2} &: q_3 \bar{p}_6 q_2 \bar{p}_1 q_1 p_3 p_4 p_5, \\ \pi_{p_3} &: q_3 \bar{p}_4 \bar{p}_5 \bar{p}_6 q_2 \bar{p}_1 q_1 \bar{p}_2, & \pi_{p_4} &: q_3 p_3 q_2 \bar{p}_5 \bar{p}_6 \bar{p}_1 q_1 \bar{p}_2, \\ \pi_{p_5} &: q_3 p_3 q_2 p_4 q_1 \bar{p}_6 \bar{p}_1 \bar{p}_2, & \pi_{p_6} &: q_3 p_1 p_2 p_3 q_2 p_4 q_1 p_5.\end{aligned}$$

Focusing our attention on \mathcal{S}_L , we are interested in the hyperline sequences for each point p_i with respect to \mathcal{S}_L . We get:

$$\begin{aligned}\pi'_{p_1} &: p_2 p_3 p_4 p_5 p_6, & \pi'_{p_2} &: p_3 p_4 p_5 p_6 p_1, & \pi'_{p_3} &: p_4 p_5 p_6 p_1 p_2, \\ \pi'_{p_4} &: p_5 p_6 p_1 p_2 p_3, & \pi'_{p_5} &: p_6 p_1 p_2 p_3 p_4, & \pi'_{p_6} &: p_1 p_2 p_3 p_4 p_5.\end{aligned}$$

Note that we were able to choose for each sequence a half-period consisting only of positive indices. This tells us, that the six points actually lie in convex position. \mathcal{S}_L is the hexagon $p_1 p_2 p_3 p_4 p_5 p_6$, listing the vertices counterclockwise.

Now we want to deduce the location of q_1 , q_2 and q_3 with respect to \mathcal{S}_L . For this, suppose that each point p_i is given in normalized homogeneous coordinates, i.e. in the form $p_i = (x_i \ y_i \ 1)^T$. We use the bracket notation $[p_i \ p_j \ p_k]$ as shorthand for the orientation determinant $\det |p_i \ p_j \ p_k|$ which equals the (oriented) area of the triangle spanned by the three points. Its sign indicates the orientation of the triple: It is positive if and only if the points appear in counterclockwise order. (This holds whenever the homogenizing coordinates of all three points are positive.)

Then $\pi_{q_1} : \dots p_i \dots p_j \dots$ implies $[q_1 \ p_i \ p_j] > 0$ and we say q_1 lies *left* of the line $p_i p_j$. Accordingly, q_1 lies *right* of the line $p_i p_j$ whenever $\pi_{q_1} : \dots p_j \dots p_i \dots$, and the orientation determinant $[q_1 \ p_1 \ p_j]$ is negative.

Specifically, $\pi_{q_1} : p_5 p_6 p_1 p_4 p_3 p_2$ implies (among other relations) that q_1 lies left of $p_1 p_4$, right of $p_2 p_3$ and right of $p_3 p_4$. This implies that the point of intersection of $p_2 p_3$ and $p_1 p_4$, call it M , must lie right of $p_3 p_4$ as well, cf. Figure 4.20.

q_2 lies right of $p_1 p_2$, left of $p_6 p_3$, and right of $p_2 p_3$. So $p_6 p_3$ and $p_1 p_2$ intersect right of $p_2 p_3$ in the point N , say. Finally, we deduce that q_3 lies right of $p_6 p_1$, left of $p_5 p_2$, and right of $p_1 p_2$ which implies that point of intersection of $p_6 p_1$ and $p_5 p_2$ lies right of $p_1 p_2$.

The combinatorial structure of the point configuration is, therefore, as in Figure 4.21.

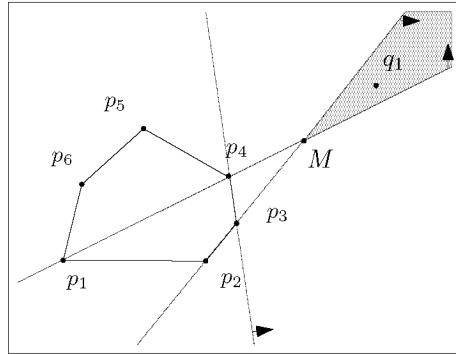


Figure 4.20: The ‘right’ point q_1 must lie in the shaded sector as determined by the location of the ‘left’ points p_1, \dots, p_6 .

We want to argue that it is impossible to draw this picture with straight lines connecting the points.

M can be expressed as:

$$M = [p_2 \ p_3 \ p_1] p_4 - [p_2 \ p_3 \ p_4] p_1,$$

since

$$\begin{aligned} & [p_2, p_3, [p_2 \ p_3 \ p_1] p_4 - [p_2 \ p_3 \ p_4] p_1] \\ &= [p_2 \ p_3 \ p_1] [p_2 \ p_3 \ p_4] - [p_2 \ p_3 \ p_4] [p_2 \ p_3 \ p_1] = 0. \end{aligned}$$

We even know that the homogenizing coordinate of M is positive in this representation, because $[p_2 \ p_3 \ p_1] > [p_2 \ p_3 \ p_4]$. This follows from the fact that p_1, p_4, M appear in this order along $p_1 p_4$. By the same reasoning, the homogenizing coordinates of

$$\begin{aligned} N &= [p_1 \ p_2 \ p_6] p_3 - [p_1 \ p_2 \ p_3] p_6 \\ \text{and } P &= [p_6 \ p_1 \ p_5] p_2 - [p_6 \ p_1 \ p_2] p_5 \end{aligned}$$

are positive too. As q_1 lies to the right of $p_6 p_5$ so does M ; equivalently, $[M \ p_5 \ p_6] > 0$ must hold. This can be rewritten as:

$$\begin{aligned} 0 &< [M \ p_5 \ p_6] \\ &= [([p_2 \ p_3 \ p_1] p_4 - [p_2 \ p_3 \ p_4] p_1) p_5 \ p_6] \\ &= [p_2 \ p_3 \ p_1] [p_4 \ p_5 \ p_6] - [p_2 \ p_3 \ p_4] [p_1 \ p_5 \ p_6]. \end{aligned} \tag{4.10}$$

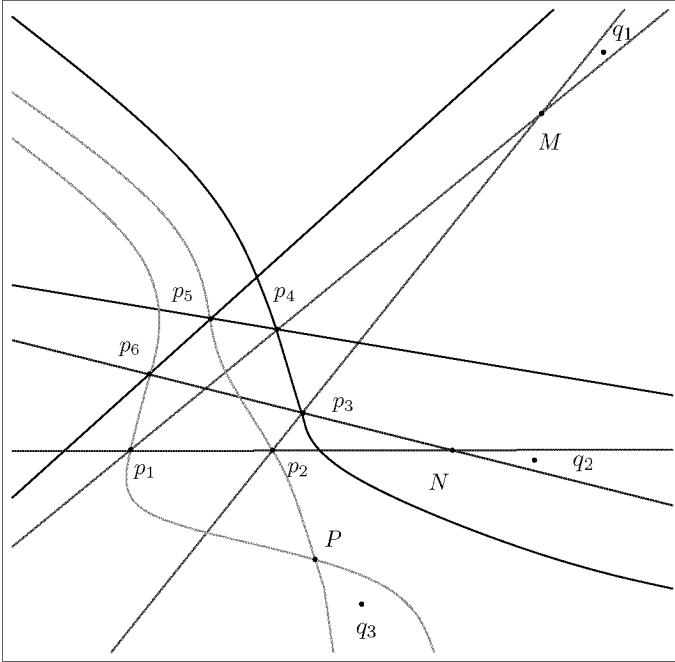


Figure 4.21: Constraints on the location of the 6 points.

Analogously, we see that $[N \ p_4 \ p_5] > 0$, which can be rewritten as:

$$[p_1 \ p_2 \ p_6] [p_3 \ p_4 \ p_5] - [p_1 \ p_2 \ p_3] [p_6 \ p_4 \ p_5] > 0, \quad (4.11)$$

and $[P \ p_3 \ p_4] > 0$, which yields:

$$[p_6 \ p_1 \ p_5] [p_2 \ p_3 \ p_4] - [p_6 \ p_1 \ p_2] [p_5 \ p_3 \ p_4] > 0. \quad (4.12)$$

Adding (4.11) and (4.12), we get

$$[p_6 \ p_1 \ p_5] [p_2 \ p_3 \ p_4] - [p_1 \ p_2 \ p_3] [p_6 \ p_4 \ p_5] > 0,$$

a contradiction to (4.10).

The case $2 \times m$. We are given an admissible grid orientation G of size $2 \times m$. We will show that it is induced by a configuration (S, ℓ) . The proof is by construction.

As before, we let $S = S_L \dot{\cup} S_R$, where $S_L = \{p_1, p_2\}$ are the two points left of ℓ and S_R is the set of m right points.

Then we can partition S_R into two sets, according to the order in which its elements see the points p_1 and p_2 . The crucial observation is that both p_1 and p_2 see first all points which see them in one order, and then all the others which see them in the other order. (For, assuming otherwise, the unique sink property would imply the existence of a forbidden subgraph.)

Therefore, each $2 \times n$ (sub-)graph is essentially of the form depicted in Figure 4.22.

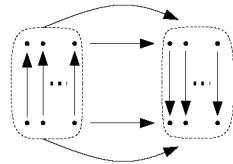


Figure 4.22: The combinatorial structure of a $2 \times m$ admissible grid orientation.

Figure 4.23 shall suffice as a demonstration of how a particular example can be realized.

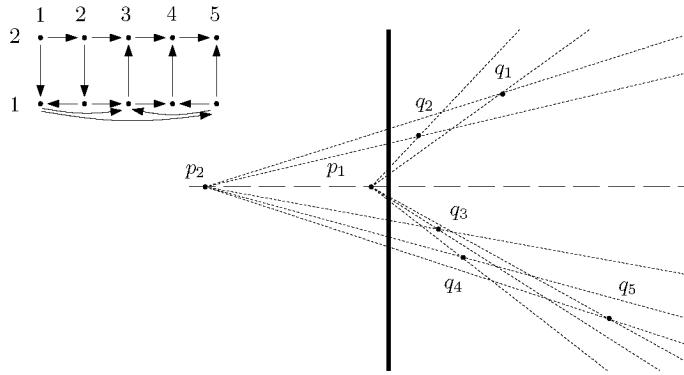


Figure 4.23: An admissible $2 \times m$ orientation and its realization.

Chapter 5

On the Completeness of Partial Chirotopes

By the Folkman-Lawrence representation theorem [FL78] every oriented matroid of rank 3 can be represented as an arrangement of oriented pseudolines. (By oriented pseudolines we mean pseudolines in the projective plane which are oriented with respect to a special line at infinity.) Uniformity means that no three pseudolines meet in a common point.

If we are given such a uniform pseudoline arrangement, it is simple to read off the chirotope: Any three pseudolines i, j, k , say, enclose a pseudotriangle. Denoting by $r \in \{1, 2, 3\}$ the number of lines whose positive halfplane contains the triangle we define $\chi(i, j, k)$ by $(-1)^r$ and $(-1)^{r+1}$, respectively, whenever i, j, k bound the pseudotriangle in clockwise and counterclockwise order, respectively. Figure 5.1 demonstrates how, given two pseudolines ℓ, ℓ' , the orientation and location of a third pseudoline p affects the value of $\chi(\ell, \ell', p)$.¹ It shall serve as an easy reference point for the reader when studying more complicated drawings later.

Thus the problem of completeness of uniform partial chirotopes of rank 3, CPC, is equivalent to the following problem:

¹From now on, we will use the bracket notation as abbreviation for χ , i.e. $\chi(i_1, \dots, i_r)$ becomes $[i_1 i_2 \dots i_r]$. +1 and -1 will also be abbreviated to + and -, respectively.

REALIZABILITY OF PARTIAL CHIROTOPES (RPC)

Given: A partial, uniform chirotope χ of rank 3 on a set E .

Question: Is there an arrangement (p_1, \dots, p_n) of oriented pseudolines such that $[i j k] = \chi(i, j, k) = +(-)$ whenever the i, j, k bound the triangle as described above and $\chi(i, j, k)$ is defined?

Since there exists a polynomial-time test for whether a chirotope satisfies the Grassmann-Plücker relations (Definition 2.14.2) this problem is in NP. Thus NP-hardness implies NP-completeness of RPC – which in turn implies NP-completeness of CPC.

We will reduce RPC to a variant of 3-SAT. Our construction is inspired by Richter-Gebert's proof [RG99] that the problem of deciding whether a given matroid is orientable is NP-complete. However, having to avoid zeros meant that we required something more elaborate.

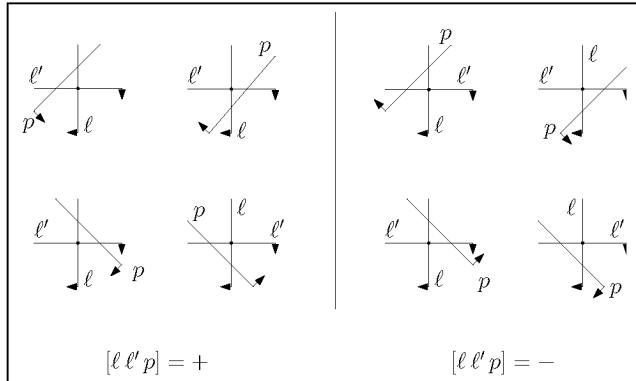


Figure 5.1: Pseudoline p added to the arrangement of lines ℓ and ℓ' .

5.1 A Variant of 3-SAT

NOT-ALTERNATING-3SAT (NA-3SAT)

Given:	An ordered set of boolean variables $X = (x_1, \dots, x_n)$ and a set \mathcal{C} of m three-clauses. partial, uniform chirotope χ of rank 3 on a set E .
Question:	Is there a truth assignment for the elements of X such that in none of the clauses the truth values of the three literals alternate, i.e. are $(\text{false}, \text{true}, \text{false})$ or $(\text{true}, \text{false}, \text{true})$?

This problem is known to be NP-complete (*cf.* [RG99], [GJ79]). Note that it is essential to have a total order on the indices of X , which induces an order on the literals of each clause. In our construction, each clause will correspond to a pair of pseudolines that have at least one crossing for each *false/true* (or *true/false*) transition in a clause. Thus, alternating clauses would force this pair of pseudolines to cross twice, something which is forbidden by the definition of pseudolines.

The next sections introduce the necessary configurations. We will later use these building blocks to actually do our construction.

5.2 The Frame of Reference

The frame of reference into which we embed our construction will be a rectangular grid $\mathcal{G}_{m,n}$ contained in an oriented matroid $\mathcal{F}_{m,n}$ as visualized in Figure 5.2: We shall have two ordered sets of lines, *vertical* and *horizontal* ones, and all lines belonging to the same set shall have the same orientation in the grid.

So, let $\mathcal{F}_{m,n}$ be the oriented matroid with elements $0, \dots, m$, the so-called ‘verticals’, $0', \dots, n'$, the so-called ‘horizontals’, and ω , the line at infinity. Impose the canonical order on these two sets: $0 < 1 < \dots < m$, and $0' < 1' < \dots < n'$, referring to their elements by using indices i, j, k , $i < j < k$ and i', j', k' , $i' < j' < k'$, respectively.

The oriented matroid $\mathcal{F}_{m,n}$ is given by the chirotope

$$[\omega ij] = [ijk] = [k'j'i'] = [ij'i'] = [\omega ii'] = [\omega j'i'] = [ij'j] = +,$$

where i, j, k are any three distinct ‘verticals’ with $i < j < k$, and i', j', k' are any three distinct ‘horizontals’ with $i' < j' < k'$; ω is the line at infinity.

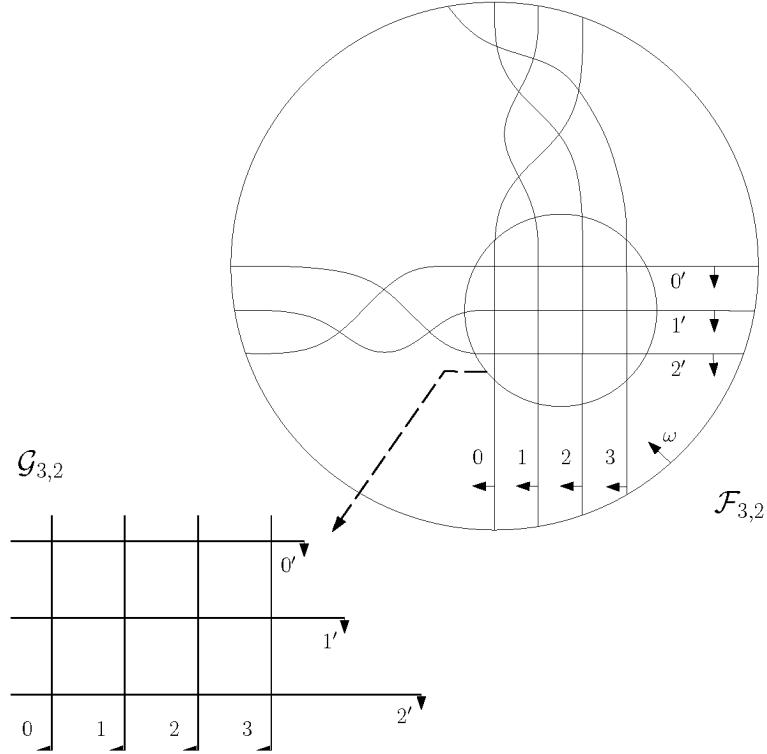


Figure 5.2: Instance of the 4×3 grid.

5.3 Adding Pseudolines

A given oriented matroid can be completed, e.g. by adding one element at a time, cf. [BLW⁺93]. From the chirotope information for the additional element, one can deduce the location of the corresponding pseudoline in the pseudoline arrangement. Conversely, having drawn an extra pseudoline, one can immediately read off the corresponding chirotope information.

We will now enlarge the oriented matroid defined in the previous section by those elements for which only partial information is given. The location of the corresponding pseudolines is, therefore, constrained, yet possesses a certain degree of freedom.

Note that each drawing satisfying all given constraints defines, in fact, a possible completion of the partial chirotope.

The crucial fact is that the partial information can be given in such a way that there are only a few (two, preferably) very specific options to place the corresponding pseudoline (i.e. to actually complete the chirotope). Hence, we are able to devise switches representing clauses and possible truth assignments. The feasibility of certain choices will only become clear once a whole clause is defined.

The actual construction will be the best explanation.

To simplify the description, we call the intersection point between a vertical i and a horizontal j' in our rectangular grid $\mathcal{G}_{m,n}$ vertex $i \wedge j'$ of $\mathcal{G}_{m,n}$. We write $[i', j']_i$ short for the segment from $i \wedge i'$ to $i \wedge j'$ on i ; $[i, j]_{j'}$ denotes the segment from $i \wedge i'$ to $j \wedge i'$ on j' , see Figure 5.3.

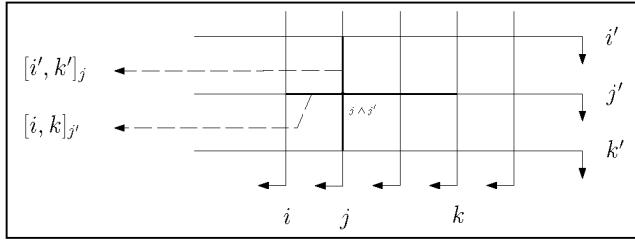


Figure 5.3: Vertices and segments in a rectangular grid.

The pseudolines $0, 0'$ and their intersection point $0 \wedge 0'$ will play a special role in our discussion as we will see shortly. Therefore, from now on, whenever we refer to pseudolines i, j and i', j' , we assume $i, j > 0, i', j' > 0'$.

Diagonal lines. Suppose we add just a single basis orientation for a new line p , $[i \ i' \ p] = -$, say. In Figure 5.1, we illustrated already the four options we have to draw this new pseudoline in our picture with respect to $i \wedge i'$.

Having two conditions for p , $[i \ i' \ p] = -, [j \ j' \ p] = +$, say, reduces these possibilities to exactly three, see Figure 5.4. In order to get rid of the ‘non-rigid’ situation drawn in the right picture of Figure 5.4, we impose a third condition, $[0 \ 0' \ p] = +$. It reduces the options further to two, as shown in Figure 5.5. This is what we will work with.

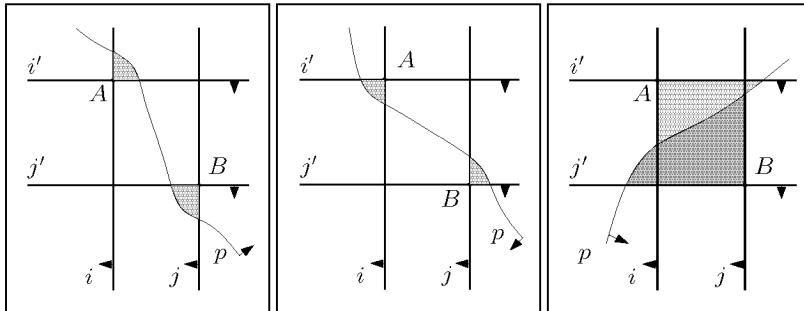


Figure 5.4: Three ways to draw p given $i < j$, $i' < j'$ and $[i i' p] = -$, $[j j' p] = +$.

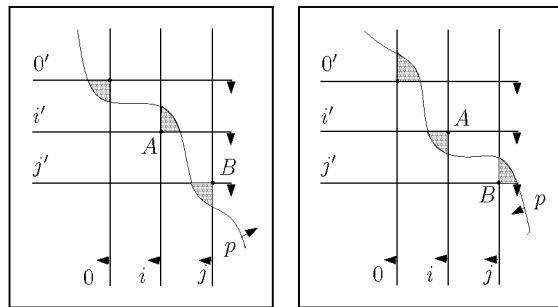


Figure 5.5: Two ways to draw p given $0 < i < j$, $0' < i' < j'$ and $[i i' p] = -$, $[j j' p] = +$, and $[0 0' p] = +$.

We call the added pseudoline p in Figure 5.5 the *diagonal p passing vertices A and B* . Later we will use this terminology as an abbreviation for ‘ p ’, with chirotope information $[i i' p] = -$, $[j j' p] = +$, $[0 0' p] = +$ where $0 < i < j$, $0' < i' < j'$, and $A = i \wedge i'$, $B = j \wedge j'$.

Parallel lines. We observed that there are essentially two ways to draw an additional pseudoline p which is ‘a diagonal passing $A = i \wedge i'$ and $B = j \wedge j'$ ’. Which one is chosen can be determined by defining further values of the chirotope function: Either, we set $[\omega 0 p] = +$ and get the picture on the right in Figure 5.5, or $[\omega 0 p] = -$ and p has to lie as shown on the left. (Recall, that ω denotes the line at infinity, cf. Figure 5.2.) By setting

just this one bracket value many others will be determined. However, this is of no concern to us – whenever we can draw a pseudoline such that all given chirotope information is correct, there obviously is a completion.

Of course, the location of p is still far from being non-ambiguous. But note that p either cuts the interior of the segments $[i, j]_{x'}$ (whenever $[\omega 0 p] = -$) or the interior of the segments $[i', j']_x$ (for $[\omega 0 p] = +$) for all $i < x < j$ or $i' < x' < j'$, respectively.

In the first case, therefore, we call p *vertical parallel* in $[i', j']_{[i, j]}$. Similarly, we call, in the latter case, p *horizontal parallel* in $[i, j]_{[i', j']}$, cf. Figure 5.6.

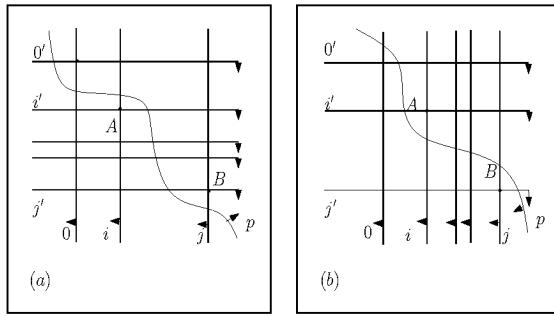


Figure 5.6: p as (a) vertical parallel in $[i, j]_{[i', j']}$ and as (b) horizontal parallel in $[i', j']_{[i, j]}$.

5.4 The Construction

We are now ready to encode an instance of the problem NA-3SAT into a completable problem. For this we need a switch which sets each variable true or false, and a structure to encode the actual clauses. In a final step, we will get a realization (and hence a full chirotope) if and only if NA-3SAT has an admissible assignment of boolean variables.

Let $X = (x_1, \dots, x_n)$ be the sequence of boolean variables, and let C_1, \dots, C_m be the set of clauses. The frame $\mathcal{F} := \mathcal{F}_{3m+3,3n}$ contains the rectangular grid $\mathcal{G}_{3m+3,3n}$. This is large enough for our construction: for each variable x_i from X , we reserve three consecutive horizontal lines (rows) a_i, b_i, c_i of \mathcal{G} . For each clause C_j , we reserve three vertical lines (columns) $1_j, 2_j, 3_j$ of \mathcal{G} . In addition, we reserve three vertical lines 1, 2, 3 for encoding

the switches to choose between values for the boolean variables. A horizontal line $0'$ and a vertical line 0 will be used as auxiliary pseudolines to define certain diagonals. Figure 5.7 sketches the global situation.

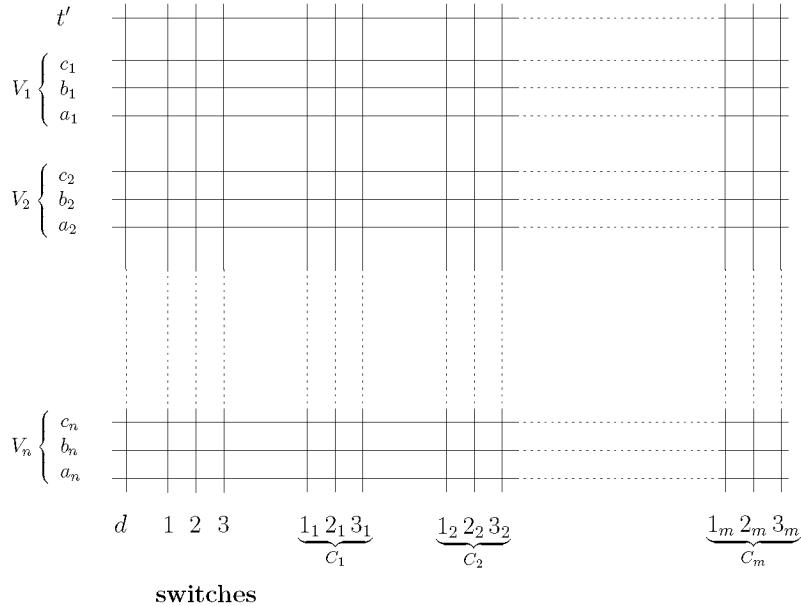


Figure 5.7: The structure to embed the NON-ALTERNATING-3SAT instance.

The switch. The idea is to give partial information such that exactly two well-defined configurations can occur. One will then represent the case ‘ x_i is true’, the other ‘ x_i is false’. You can follow the construction in Figure 5.8.

For every variable x_i we do the following:

First, we consider the rows a_i, b_i, c_i and add elements X_i and X'_i as *horizontal parallels* in $[a_i, c_i]_{[1, 3_m]}$, representing x_i and $\neg x_i$. This alone does not give us any control over the location of the intersection points of X_i, X'_i, b_i yet—we want them to intersect in the region between 1 and 3.

Secondly, for this very reason, we introduce two new *diagonals* U_i and W_i

passing $1 \wedge c_i$ and $2 \wedge a_i$, and $2 \wedge c_i$ and $3 \wedge a_i$, respectively. The information $[U_i a_i W_i] = [U_i b_i W_i] = [U_i c_i W_i] = -$ guarantees that they have opposite orientation and do not intersect inside our little gadget.

Finally, we force U_i and W_i to meet X'_i, b_i, X_i in the same order. Since U_i and W_i have opposite orientation, this will imply that X'_i, b_i, X_i pairwise intersect between U_i and W_i inside the 3×3 -grid-gadget. This can be guaranteed by the additional chirotope information $[U_i b_i X_i] = [W_i b_i X_i] = [U_i X'_i b_i] = [W_i X'_i b_i] = [U_i X'_i X_i] = [W_i X'_i X_i] = -$.

Note that these additional conditions also ensure that X_i, X'_i and b are pairwise non-identical.

The two possible situations in our realization are shown in Figure 5.8. We associate the situation on the right with $x_i = \text{false}$, and the situation on the left with $x_i = \text{true}$. (We omit the drawing of the pseudolines $0, 0'$ for the sake of lucidity.)

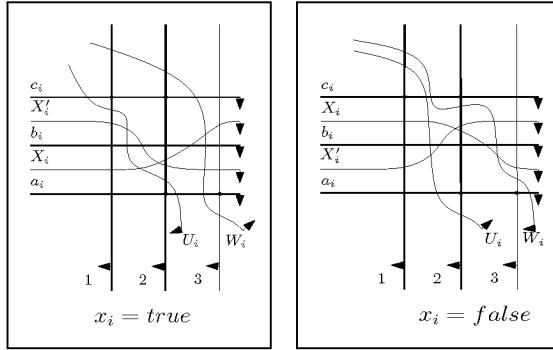


Figure 5.8: The two states of a switch.

We make the following crucial observation:

Lemma 5.1 *If we are in the situation $x_i = \text{true}$, then for all $3 \leq k \leq 3m + 3$ the line X'_i cuts the interval $[a_i, b_i]_k$ and the line X_i cuts the interval $[b_i, c_i]_k$. For $x_i = \text{false}$, the situation is reversed.*

So, the locally fixed choice in our switches is having an impact on the whole configuration, the information $x_i = \text{false}/\text{true}$ is ‘present’ not only in the switches. Thanks to this key property, the clauses (which we will describe next) can function.

The clauses.

Let ℓ_i denote the i^{th} literal, i.e. $\ell_i = x_i$ or $\neg x_i$; let $L_i = \begin{cases} X_i, & \ell_i = x_i \\ X'_i, & \ell_i = \neg x_i \end{cases}$.

Then Lemma 5.1 implies that the truth value of ℓ_i is *true* whenever L_i cuts $[b_i, c_i]_k$, $\ell_i = \text{false}$ whenever L_i cuts $[a_i, b_i]_k$, for all k , $3 \leq k \leq 3m + 3$.

Remember that, for each clause C_j , we reserved three vertical lines $1_j, 2_j, 3_j$. The clause C_j consists of three literals. For each literal ℓ_i (represented by L_i) that appears in the clause C_j , we enlarge our chirotope with a diagonal T_i^j passing $1_j \wedge c_i, 2_j \wedge L_i$ and $3_j \wedge a_i$, and satisfying $[T_i^j b_i L_i] = -$. The latter condition determines which of the two possible drawings for the diagonal T_i^j is chosen (recall Figure 5.5)—depending on the relative position of b_i and L_i . So, in accordance with the state of ℓ_i , we can only have one of the two situations shown in Figure 5.9.

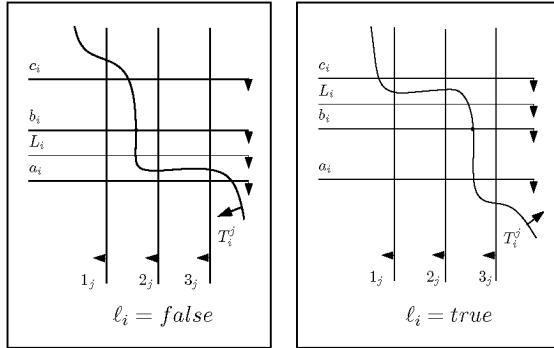


Figure 5.9: Connecting a switch and a clause.

The proof of the following lemma is again straightforward.

Lemma 5.2 *If we are in the situation $\ell_i = \text{false}$, then the line T_i^j cuts the interval $[1_j, 2_j]_{b_i}$. Otherwise T_i^j cuts the interval $[2_j, 3_j]_{b_i}$.*

Before we do the final step, let us look back for a moment at what we have got so far. For a given set of three-clauses over n boolean variables, we defined a partial chirotope² which has *inside the gadget* exactly 2^n possible completions — each one corresponding to a truth assignment.

²Note that there is no need to review the axioms of Definition 2.6 in order to see that the set of basic orientations given so far constitutes a partial chirotope. Instead, we have shown that we can construct arrangements of oriented pseudolines for which all the given basic orientations are

Our aim is therefore to add further chirotope information such that only truth assignments without alternating clauses are possible, while still having a partial chirotope.

For each clause C_j , consisting of literals coming from the variables x_{i_1}, x_{i_2} and x_{i_3} , say, we add Z_j , which is *vertical parallel* in $[1_j, 3_j]$ and for which $[T_{i_1}^j b_{i_1} Z_j] = [T_{i_2}^j b_{i_2} Z_j] = [T_{i_3}^j b_{i_3} Z_j] = +$ holds.

These values suffice to force the line Z_j to pass $T_{i_1}^j \wedge b_{i_1}$, $T_{i_2}^j \wedge b_{i_2}$, and $T_{i_3}^j \wedge b_{i_3}$ “away from” 2_j . It remains to be checked that they do not violate the chirotope axioms (Definition 2.14). But since we do not specify any further basic orientations involving the new elements Z_j , this is clearly the case.

Extending our arrangement of pseudolines by such lines Z_1, \dots, Z_m (and hence completion of the partial chirotope) is only possible if the corresponding literals in the clauses do not alternate. Alternating literals in clause C_j would force the line Z_j to cross the line 2_j twice which is forbidden by the axioms. In all other cases, the pseudolines are insertable. For an example, see Figure 5.10.

This completes the proof of our result: Given an ordered set of boolean variables $X = (x_1, \dots, x_n)$ and a set S of m three-clauses, we may define a partial chirotope which is completable if and only if there is a ‘non-alternating’ truth assignment for the elements of X , cf. NA-3SAT. Moreover, we can read off the admissible truth values for the boolean variables from the completion: Using our notation from above, we have

$$x_i = \begin{cases} \text{true} & \text{whenever } [1 X_i X'_i] = + \\ \text{false} & \text{whenever } [1 X_i X'_i] = -. \end{cases}$$

Conversely, every admissible assignment of truth values for \mathcal{C} corresponds to a chirotope (representable by a pseudoline arrangement) that can be considered as a completion of the partial chirotope induced by the set \mathcal{C} as described.

Thus we have proved:

Theorem 5.3 *The problem of testing completable of a partial chirotope is NP-complete.*

valid. Since each such arrangement defines a chirotope this implies that we have defined indeed a partial chirotope, and this partial chirotope is completable.

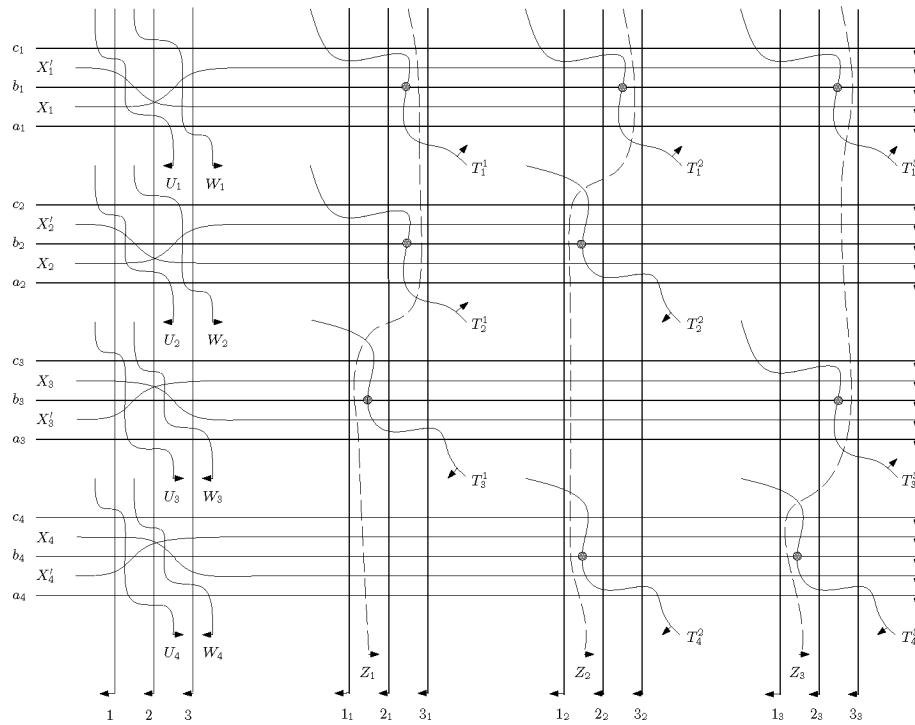


Figure 5.10: $(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \neg x_2 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$, setting $x_1 = x_2 = \text{true}$ and $x_3 = x_4 = \text{false}$.

Bibliography

- [AZ99] Nina Amenta and Günter M. Ziegler. Deformed products and maximal shadows of polytopes. In J. Chazelle, J.E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*, volume 223 of *Contemporary Mathematics*, pages 57–90. Amer. Math. Soc., 1999.
- [BB80] Gerd Blind and Roswitha Blind. Die primitiven Polytope des R^d mit $d + 2$ und mit $2d - 1$ Facetten. *Journal für die reine und angewandte Mathematik*, 320:127–136, 1980.
- [BDF⁺95] Andrei Z. Broder, Martin E. Dyer, Alan M. Frieze, Prabhakar Raghavan, and Eli Upfal. The worst-case running time of the random simplex algorithm is exponential in the height. *Information Processing Letters*, 56(2):79–81, 1995.
- [BLW⁺93] Anders Björner, Michel Las Vergnas, Neil White, Bernd Sturmfels, and Günter M. Ziegler. *Oriented Matroids*. Cambridge University Press, Cambridge, 1993.
- [BMS01] Jürgen Bokowski, Susanne Mock, and Ileanu Streinu. The Folkman-Lawrence Topological Representation Theorem for Oriented Matroids: an Elementary Proof in Rank 3. *European Journal of Combinatorics*, 22(5):601–615, 2001.
- [Bor87] Karl Heinz Borgwardt. *The Simplex Method: A Probabilistic Analysis*, volume 1 of *Algorithms and Combinatorics*. Springer-Verlag, New York, 1987.
- [Chv83] Vašek Chvátal. *Linear Programming*. W. H. Freeman, New York, NY, 1983.

- [Dan63] George B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey, 1963.
- [Dev02] Mike Develin. On LP-orientations of cubes and crosspolytopes. preprint, 2002.
- [FH99] Kerstin Fritzsche and Fred B. Holt. More polytopes meeting the conjectured Hirsch bound. *Discrete Mathematics*, 205:77–84, 1999.
- [FL78] Jon Folkman and Jim Lawrence. Oriented matroids. *Journal of Combinatorial Theory Ser. B*, 25:199–236, 1978.
- [Gal56] David Gale. Neighboring vertices on a convex polyhedron. In *Linear inequalities and related system*, volume 38 of *Annals of Mathematics Studies*, pages 255–263. Princeton University Press, Princeton, NJ, 1956.
- [Gär95] Bernd Gärtner. *Randomized Optimization by Simplex-Type Methods*. PhD thesis, Institut für Informatik, Freie Universität Berlin, 1995.
- [Gär02] Bernd Gärtner. The random-facet simplex algorithm on combinatorial cubes. *Random Structure and Algorithms*, 20(3):353–381, 2002.
- [GHZ98] Bernd Gärtner, Martin Henk, and Günter M. Ziegler. Randomized simplex algorithms on Klee-Minty cubes. *Combinatorica*, 18:349–372, 1998.
- [GJ79] Michael R. Garey and Daniel S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, New York, NY, 1979.
- [GKP94] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading, MA, second edition, 1994.
- [Grü67] Branko Grünbaum. *Convex Polytopes*. Interscience, London, 1967.
- [GS92] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and Random Processes*. Oxford University Press, New York, NY, second edition, 1992.

- [GST⁺01] Bernd Gärtner, József Solymosi, Falk Tschirschnitz, Pavel Valtr, and Emo Welzl. One line and n points. *Proc. 33rd Ann. ACM Symp. on the Theory of Computing (STOC)*, pages 306–315, 2001.
- [GT03] Bernd Gärtner and Falk Tschirschnitz. Partial chirotopes and the digraph of $(d, d + 2)$ -polytopes. submitted, 2003.
- [HK98a] Fred B. Holt and Victor Klee. Many polytopes meeting the conjectured hirsch bound. *Discrete & Computational Geometry*, 20:1–17, 1998.
- [HK98b] Fred B. Holt and Victor Klee. A proof of the strict monotone 4-step conjecture. In J. Chazelle, J.E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*, volume 223 of *Contemporary Mathematics*, pages 201–216. Amer. Math. Soc., 1998.
- [JK99] Michael Joswig and Volker Kaibel. Randomized simplex algorithms and random cubes (extended abstract). Technical report, TU Berlin, 1999.
- [Kal92] Gil Kalai. A subexponential randomized simplex algorithm. In *Proc. 24th Annu. ACM Sympos. Theory Comput.*, pages 475–482, 1992.
- [Kal97] Gil Kalai. Linear programming, the simplex algorithm and simple polytopes. *Mathematical Programming Ser. B*, 79:217–233, 1997.
- [Kar84] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.
- [Kha80] Leonid G. Khachiyan. Polynomial algorithms in linear programming. *USSR Computational Mathematics and Mathematical Physics*, 20:53–72, 1980. (Russian original in Журнал Вычислительной Математики и Математической Физики, 20:51–68).
- [KK92] Gil Kalai and Daniel J. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *BAMS: Bulletin of the American Mathematical Society*, 26:315–316, 1992.
- [KM72] Victor Klee and George J. Minty. How good is the simplex algorithm? In O. Shisha, editor, *Inequalities III*, pages 159–175. Academic Press, 1972.

- [KMSZ02] Volker Kaibel, Rafael Meichel, Micha Sharir, and Günter M. Ziegler. The RANDOM EDGE rule on three-dimensional linear programs. *preprint*, 2002.
- [Lee97] Jon Lee. Hoffman’s circle untangled. *SIAM Review*, 39(1):98–105, March 1997.
- [Loe96] Jesus De Loera. Nonregular triangulations of products of simplices. *Discrete & Computational Geometry*, 15:253–264, 1996.
- [Mat94] Jiří Matoušek. Lower bounds for a subexponential optimization algorithm. *Random Structures & Algorithms*, 5(4):591–607, 1994.
- [Mat02] Jiří Matoušek. *Lectures on Discrete Geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, 2002.
- [Meg84] Nimrod Megiddo. Linear programming in linear time when the dimension is fixed. *Journal of the ACM*, 31(1):114–127, 1984.
- [MK00] Jed Mihalisin and Victor Klee. Convex and linear orientations of polytopal graphs. *Discrete & Computational Geometry*, 24:421–435, 2000.
- [Mor02a] Walter D. Morris, Jr. Distinguishing cube orientations arising from linear programs. submitted, 2002.
- [Mor02b] Walter D. Morris, Jr. Randomized principal pivot algorithms for p -matrix linear complementarity problems. *Mathematical Programming Ser. A*, 92:285–296, 2002.
- [MR95] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, New York, NY, 1995.
- [MSW92] Jiří Matoušek, Micha Sharir, and Emo Welzl. A subexponential bound for linear programming. In *Proc. Eighth Ann. ACM Symp. on Computational Geometry*, pages 1–8, Berlin, Germany, 1992.
- [Pfe03] Julian Pfeifle. *Extremal Constructions for Polytopes and Spheres*. PhD thesis, Technische Universität Berlin, 2003.
- [RG99] Jürgen Richter-Gebert. Orientability of matroids is NP-complete. *Advances in Applied Mathematics (special issue: “in the honour of Henry Crapo”*, ed. J. Kung), 23:78–90, 1999.

- [Rin56] Gerhard Ringel. Teilung der Ebene durch Geraden oder topologische Geraden. *Mathematische Zeitschrift*, Bd. 20:79–102, 1956.
- [ST01] Daniel Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time. In ACM, editor, *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 296–305, New York, NY, USA, 2001. ACM Press.
- [SW01] Tibor Szabo and Emo Welzl. Unique sink orientations of cubes. In *42nd IEEE Symposium on Foundations of Computer Science*, pages 547–555. IEEE Computer Society Press, 2001.
- [Tsc02] Falk Tschir schnitz. Deciding the extendibility of partial chirotopes is NP-hard. *Proc. 13th Canadian Conference on Computational Geometry (CCCG)*, pages 165–168, 2002.
- [Wel01] Emo Welzl. Entering and leaving j -facets. *Discrete & Computational Geometry*, 25(3):351–364, 2001.
- [Wil88] Kathy Williamson Hoke. Completely unimodal numberings of a simple polytope. *Discrete Applied Mathematics*, 20:69–81, 1988.
- [Zie94] Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, 1994.

Curriculum Vitae

Falk Tschirschnitz

born on Oktober 18, 1973 in Erfurt, Germany

1988–1992 **Specialist Highschool “Albert Schweizer”
for Mathematics and Natural Sciences
in Erfurt, Germany**
Degree: Abitur

1992–1994 **Civil Service in Erfurt, Germany**

1994–1998 **studies at Oxford University, United Kingdom**
Degree: Master of Mathematics

1998–2003 **Ph.D. student at ETH Zürich, Switzerland**
Theoretical Computer Science