Lecture notes: “Algebra and Polyhedral Geometry”
May 30, 2014
Anders Nedergaard Jensen

Preface

This course is inspired by the book “Gröbner bases and convex polytopes” by Bernd Sturmfels [13]. Motivated by algorithmic problems for multivariate polynomial rings and polynomial equations we study Gröbner bases and their term orderings. Buchberger’s algorithm, convexity and Newton polytopes play important roles. The theory is applied to toric ideals and integer programming. Unexpectedly the combinatorial space of regular triangulations of a vector configuration naturally appears in this algebraic setting. The class will end with a brief introduction to *tropical geometry* where all the theory is combined.

While Sturmfels’ book can be difficult to read for a beginner, the course notes “Computational Algebra and Combinatorics of Toric Ideals“ by Diane Maclagan and Rekha Thomas [11] are more accessible. Those notes are recommended when the lecture notes for our class are too brief.

We begin with an introduction to Gröbner bases since not everybody took an introductory algebra class based on Lauritzen’s book [10].

These notes, which will keep growing through the semester, can be found at
`http://home.imf.au.dk/jensen/teaching/2014AlgebraAndPolyhedralGeometry/notes.pdf`
while a version from the same class in 2012 can be found here
`http://home.imf.au.dk/jensen/teaching/2012AlgebraAndPolyhedralGeometry/notes.pdf`
The old version is complete, while the new version contains fewer mistakes.

Contents

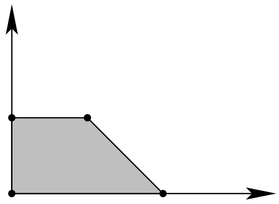

Figure 1: The Newton polytope of the polynomial in Example 1.1.3.

1 Gröbner bases

In this section we define Gröbner bases and start discussing their relation to convex geometry. On our way we will have to define term orders, initial terms and the division algorithm.

1.1 The polynomial ring

We let k be a field and $n \in \mathbb{N} := \{0, 1, 2, \dots\}$ and consider the ring $S := k[x_1, \dots, x_n]$ of polynomials in the variables $x_1, \dots, x_n$ with coefficients in k. In examples we will often use letters as variable names, and for example consider the ring $\mathbb{Q}[x, y, z]$.

Definition 1.1.1 A vector $u \in \mathbb{N}^n$ defines the *monomial* $x^u := x_1^{u_1} \cdots x_n^{u_n}$. The vector u is called an *exponent vector*. By a *term* we mean a polynomial in $k[x_1, \dots, x_n]$ of the form cx^u with $c \in k \setminus \{0\}$.

If we require the exponent vectors to be distinct then a polynomial can be written uniquely as a sum of terms.

Definition 1.1.2 The *support* $\text{supp}(f)$ of a polynomial $f \in k[x_1, \dots, x_n]$ is the set of exponent vectors in f (in its unique representation). The *Newton polytope* $\text{NP}(f)$ is the convex hull of $\text{supp}(f)$ in $\mathbb{R}^n$.

For the precise definition of convex hull see Definition 2.2.2.

Example 1.1.3 The polynomial $f = (x^3 + y + xy) - (1 + x^3 + x^2) = y + xy - 1 - x^2 \in \mathbb{Q}[x, y]$ has $\text{supp}(f) = \{(0, 1), (1, 1), (0, 0), (2, 0)\}$. Its Newton polytope is shown in Figure 1.

Definition 1.1.4 For polynomials $f, g \in k[x_1, \dots, x_n]$ we say that f divides g and write $f|g$ if there exists $h \in k[x_1, \dots, x_n]$ such that $fh = g$. We let $g/f := h$.

We will be interested in ideals in the polynomial ring S (nonempty subsets of S which are closed under (1) addition ($f, g \in I \Rightarrow f + g \in I$), and (2) multiplication by elements in S ($f \in I \wedge g \in S \Rightarrow fg \in I$)). Considering these sets as equations, they define subsets of k^n called *varieties*:

Definition 1.1.5 Let $I \subseteq k[x_1, \dots, x_n]$ be an ideal. The *variety* $V(I)$ defined by I is

$$V(I) := \{a \in k^n : \forall f \in I : f(a) = 0\}.$$

One way to get an ideal is to take a finite set of polynomials $f_1, \dots, f_m$ and look at the set they generate: $\langle f_1, \dots, f_m \rangle := \{\sum_i g_i f_i : g_i \in S\}$. This is an ideal. Even if we allow a possible infinite *generating set* of polynomials $F \subseteq k[x_1, \dots, x_n]$ the set they generate $\langle F \rangle = \{\sum_{i=0}^m g_i f_i : m \in \mathbb{N} \wedge g_i \in S \wedge f_i \in F\}$ is an ideal. Hilbert's basis theorem, which will follow from Proposition 1.6.7, says that a finite set of generators suffices:

Theorem 1.1.6 (Hilbert's Basis Theorem) *Let k be a field, $n \in \mathbb{N}$ and I an ideal in $k[x_1, \dots, x_n]$. Then there exists a finite set $f_1, \dots, f_m$ of polynomials such that $I = \langle f_1, \dots, f_m \rangle$.*

Lemma 1.1.7 *Let R be a commutative ring, and $F \subseteq R$ a generating set for an ideal $I := \langle F \rangle$. If I has a finite generating set G, then there is a finite subset $F' \subseteq F$ such that $I := \langle F' \rangle$.*

Proof. Each element in G can be written as $\sum_{i=1}^m g_i f_i$ for some $m \in \mathbb{N}$, $g_i \in R$, and $f_i \in F$. We now take F' to be the finite set of all f_i appearing when expressing all elements of G in this way. Then $I = \langle G \rangle \subseteq \langle F' \rangle \subseteq \langle F \rangle = I$. □

Recall that the *quotient ring* $k[x_1, \dots, x_n]/I$ consists of elements of the form $[f] := f + I = \{f + h : h \in I\}$ where $f \in k[x_1, \dots, x_n]$. The element $[f]$ is called a *coset* and together the cosets form a ring with operations $[f] + [g] := [f + g]$ and $[f][g] := [fg]$. Furthermore, $[f] = [g]$ if and only if $f - g \in I$.

We are interested in computational tools for the following problems:

- Finding all points in the variety $V(I)$.
- Doing computations in the quotient ring $k[x_1, \dots, x_n]/I$ – In particular testing ideal membership: Given $f \in S$ and generators for an ideal $I \subseteq S$, decide if $f \in I$.

Gröbner bases will help us solve these problems. Furthermore, the existence of Gröbner bases will prove Hilbert's basis theorem.

1.2 Monomial ideals and Dickson's Lemma

In this subsection we consider the special case of monomial ideals.

Definition 1.2.1 An ideal $I \subseteq k[x_1, \dots, x_n]$ is called a *monomial ideal* if it is generated by (possibly infinitely many) monomials.

We observe that a polynomial belongs to a monomial ideal if and only if each of its terms does. Furthermore, a monomial ideal is determined by the set of monomials it contains (because these generate the ideal). This makes it possible to draw monomial ideals by drawing the exponents vectors of their generators in $\mathbb{R}^n$.

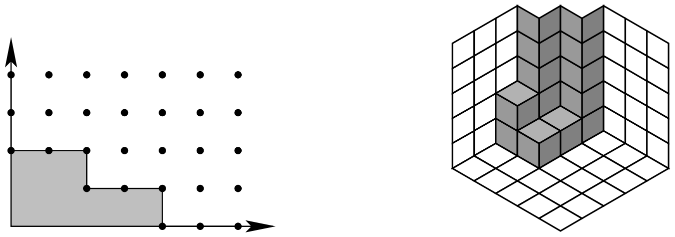

Figure 2: Staircase diagrams of the ideals in Example 1.2.2.

Observe that $x^v | x^u$ if and only if $\forall i : v_i \leq u_i$. Furthermore if M is a set a monomials then $x^u \in \langle M \rangle \Leftrightarrow \exists x^v \in M : x^v | x^u$. See Exercise 10, Sheet 1.

Example 1.2.2 *Staircase diagrams* of the monomial ideals $I := \langle x^4, x^2y, y^2 \rangle \subseteq k[x, y]$ and $J := \langle x^2, y^3, y^2z^2, xyz \rangle \subseteq k[x, y, z]$ are shown in Figure 1.2. The second picture is drawn without perspective, and can therefore be interpreted in two ways. Most likely your mind will see the grey cubes with coordinates being vectors not among the exponent vectors of monomials in J.

A generating set $F \subseteq k[x_1, \dots, x_n]$ for an ideal is called *minimal* if for every $f \in F : \langle F \rangle \neq \langle F \setminus \{f\} \rangle$.

Lemma 1.2.3 *Every monomial ideal $I \subseteq k[x_1, \dots, x_n]$ has a unique minimal monomial generating set.*

Proof. Consider the set $F := \{x^u \in I : \forall x^v \in I \setminus \{x^u\} : x^v \nmid x^u\}$. We first prove that F generates I by showing that every monomial $x^w \in I$ is divisible by some element of F. If $x^w \in F$ then indeed $x^w \in F$ divides x^w. If $x^w \notin F$ then there exists $x^{w'} \in I \setminus \{x^w\}$ such that $x^{w'} | x^w$. If $x^{w'} \in F$ then we are done. If $x^{w'} \notin F$ then there exists $x^{w''} \in I \setminus \{x^{w'}\}$ such that $x^{w''} | x^{w'} | x^w$. If $x^{w''} \in F$ then we are done. We continue in this way, but the process must eventually stop since the integer entries of the exponent vectors become smaller and smaller. Hence there exists $x^u \in F$ such that $x^u | x^w$.

We now argue that F is contained in any monomial generating set for I. But this is indeed the case because no other generator can divide these elements. This shows that F is minimal and unique. □

We prove Hilbert's basis theorem in the monomial case:

Lemma 1.2.4 (Dickson's Lemma) *Every monomial ideal $I \subseteq k[x_1, \dots, x_n]$ has a finite monomial generating set.*

Proof. Induction. For $n = 0$ the ideal is either $\{0\}$ or k. In the first case the empty set $\emptyset$ is a finite generating set. In the second case $\{1\}$ is.

For $n > 0$ we let $\pi : \mathbb{N}^n \to \mathbb{N}^{n-1}$ denote the projection which forgets the last coordinate. Define $E := \pi(\{v \in \mathbb{N}^n : x^v \in I\})$. By the induction hypothesis $J := \langle x^u : u \in E \rangle \subseteq k[x_1, \dots, x_{n-1}]$ has a finite generating set and by Lemma 1.1.7 there exists a finite subset $F \subseteq E$ such that $J = \langle x^u : u \in F\rangle$. Each $u \in F$ has some *lift* $v \in \mathbb{N}^n$ such that $\pi(v) = u$ and $x^v \in I$ with v_n minimal. We let G denote the set of these lifts. We now take $m = \max_{v\in G} v_n$. If $x^w \in I$ with $w_n > m$ then the there is some $u \in F$ such that $x^u | x^{\pi(w)}$. Since $w_n > m$ the lift v of u satisfies $x^v | x^w$. Now for $j = 0, \dots, m$ we consider the ideal $J_j := \langle x^u : u \in \mathbb{N}^{n-1} \text{ and } x^u x_n^j \in I\rangle \subseteq k[x_1, \dots, x_{n-1}]$. Geometrically J_j is a slice of (the complement of) the staircase diagram of I where $u_n = j$. By induction each J_j has a finite monomial generating set G_j. The set $\{x^v : v \in G\} \cup \bigcup_{j=0}^m \{x^u x_n^j : x^u \in G_j\}$ is a finite generating set of I. □

Corollary 1.2.5 *Let $M_1 \subseteq M_2 \subseteq M_3 \subseteq \cdots$ be monomial ideals in $k[x_1, \dots, x_n]$. For some $j \in \mathbb{N}$ we must have $M_j = M_{j+1} = M_{j+2} = \cdots$.*

Proof. We consider the ideal $M := \bigcup_i M_i$ generated by all monomials in all M_i. By Lemma 1.2.4 it has a finite generating set F. For each $f_i \in F$ there exists a $j_i \in \mathbb{N}$ such that $f_i \in M_{j_i}$. For $j := \max_i(j_i)$ we have $F \subseteq M_j$, implying $M \subseteq M_j$. Since $M_i \subseteq M$ for all i we have $M = M_j = M_{j+1} = \cdots$. □

A ring for which the above corollary holds for inclusions of any ideals $I_1 \subseteq I_2 \subseteq \cdots$ (not necessarily monomial ideals) is called a *Noetherian ring.* We will prove later that $k[x_1, \dots, x_n]$ is Noetherian.

1.3 Term orderings

Recall that a *total ordering* $\leq$ on a set X is a relation satisfying for all $a, b, c \in X$:

Antisymmetry: $a \leq b \wedge b \leq a$ implies $a = b$.

Transitivity: $a \leq b \wedge b \leq c$ implies $a \leq c$.

Totality: $a \leq b \vee b \leq a$.

Just like [11] and [13] we will be sloppy and sometimes forget the horisontal bar when writing $\leq$. For example when we say "Let $\prec$ be a total order(ing)" we really mean that $\preceq$ should be the total ordering, and $\prec$ is then defined by $a \prec b \Leftrightarrow a \preceq b \wedge a \neq b$.

Definition 1.3.1 A *term ordering* (or a *monomial ordering*) $\preceq$ on $k[x_1, \dots, x_n]$ is an total ordering on the monomials in $k[x_1, \dots, x_n]$ such that:

- $x^a \preceq x^b$ implies $x^a x^c \preceq x^b x^c$ for $a, b, c \in \mathbb{N}^n$.
- $1 = x^0 \preceq x^a$ for all $a \in \mathbb{N}^n$.

Since term orders are orderings on *monomials*, it would be more correct to call them *monomial orders* and some people do that. However, as we shall see later, we most often use orderings to order the *terms* of a polynomial.

We give two examples of term orderings:

Example 1.3.2 We define the *lexicographic term ordering* $\preceq_{\text{lex}}$ on $k[x_1, \dots, x_n]$ as follows. For $a, b \in \mathbb{N}^n$ we let $x^a \prec_{\text{lex}} x^b \Leftrightarrow a_1 < b_1 \vee a_1 = b_1 \wedge (a_2 < b_2 \vee a_2 = b_2 \wedge (\dots (a_n < b_n) \dots)))$. Or, more precisely, $x^a \prec_{\text{lex}} x^b \Leftrightarrow \exists j \leq n : a_1 = b_1 \wedge a_2 = b_2 \wedge \dots \wedge a_{j-1} = b_{j-1} \wedge a_j < b_j$.

Example 1.3.3 In $\mathbb{Q}[x, y, z]$ we have $1 \prec_{\text{lex}} z \prec_{\text{lex}} z^2 \prec_{\text{lex}} z^9 \prec_{\text{lex}} y \prec_{\text{lex}} yz^2 \prec_{\text{lex}} y^5 \prec_{\text{lex}} x^2y^2z \prec_{\text{lex}} x^3$.

Remark 1.3.4 For $a, b \in \mathbb{N}^n$, $x^a \preceq_{\text{lex}} x^b$ if and only if $a - b = 0$ or the first non-zero entry of $a - b$ is negative.

Lemma 1.3.5 *The lexicographic ordering* $\prec_{\text{lex}}$ *is a term ordering.*

Proof. **Antisymmetry:** We have $a, b \in \mathbb{N}^n$ such that $x^a \preceq_{\text{lex}} x^b$ and $x^a \preceq_{\text{lex}} x^b$. Suppose $a \neq b$. Then Remark 1.3.4 says that the first non-zero entry of $a - b$ is negative and the first non-zero entry of $b - a$ is negative. This is a contradiction. Hence $x^a = x^b$.

Transitivity: Suppose $x^a \preceq_{\text{lex}} x^b$ and $x^b \preceq_{\text{lex}} x^c$. If $a = b$ or $b = c$ then we conclude $x^a \preceq_{\text{lex}} x^b$. If both $a \neq b$ and $b \neq c$ then by Remark 1.3.4 the first non-zero entry of $a - b$ is negative. So is the first non-zero entry of $b - c$. We conclude that the first non-zero entry of the sum $(a-b)+(b-c) = a-c$ is negative, implying $x^a \preceq_{\text{lex}} x^c$.

Totality: We have $a, b \in \mathbb{N}^n$. If $a = b$ then $x^a \preceq_{\text{lex}} x^b$. Assume $a \neq b$ then the first non-zero entry of $a - b$ is either positive or negative. In the last case $x^a \preceq_{\text{lex}} x^b$. In the first the first case the first non-zero entry of $b - a$ is negative, implying $x^b \preceq_{\text{lex}} x^a$.

Multiplication respected: By Remark 1.3.4, $x^a \preceq_{\text{lex}} x^b$ is a condition on $a - b$. Furthermore, $x^{a+c} \preceq_{\text{lex}} x^{b+c}$ is the same condition on $(a + c) - (b + c) = a - b$.

1 is smallest: $x^0 \preceq_{\text{lex}} x^b$ since for $b \in \mathbb{N}^n$, either $0 - b = 0$ or the first nonzero entry of $0 - b$ is negative.

□

Example 1.3.6 We define the *graded (or degree) reverse lexicographic term ordering* $\prec_{\text{grlex}}$ on $k[x_1, \dots, x_n]$ as follows. For $a, b \in \mathbb{N}^n$ we let $x^a \prec_{\text{grlex}} x^b \Leftrightarrow \sum_i a_i < \sum_i b_i \vee \sum_i a_i = \sum_i b_i \wedge \exists j : a_j > b_j \wedge a_{j+1} = b_{j+1} \wedge \dots \wedge a_n = b_n$.

Lemma 1.3.7 *Every term ordering* $\prec$ *on* $k[x_1, \dots, x_n]$ *is a well ordering.*

Proof. Let X be a set of monomials in $k[x_1, \dots, x_n]$. We must show that X contains a smallest element. By Lemma 1.2.4 and Lemma 1.1.7 the ideal $\langle X \rangle$ has a finite monomial generating set $Y \subseteq X$. Let x^a be the smallest term in the finite set Y. We claim that x^a is a smallest element of X. Let x^b be any term in X. Then $x^b \in \langle X \rangle = \langle Y \rangle$. Hence some $x^c \in Y$ divides x^b. That is $x^b = x^c x^d$ for some $d \in \mathbb{N}^n$. By Definition 1.3.1 we have $1 \preceq x^d$, implying $x^c \preceq x^c x^d = x^b$. We also have $x^a \preceq x^c$ since $x^c \in Y$. Hence $x^a \preceq x^c \preceq x^b$ as desired. □

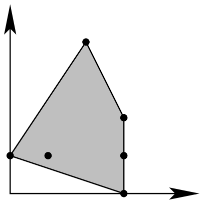

Figure 3: The Newton polytope of f in Example 1.4.2.

1.4 Initial terms and initial forms

Definition 1.4.1 Let $\prec$ be a term ordering, $\omega \in \mathbb{R}^n$ and $f = \sum_{u \in U} c_u x^u \in k[x_1, \ldots, x_n]$ a polynomial with support $U \subseteq \mathbb{N}^n$, $c_u \neq 0$. If $f \neq 0$ we define the *initial term* $\text{in}_\prec(f)$ of f to be $c_u x^u$ with x^u being largest with respect to $\prec$ among the monomials of f. For any $f = \sum_{u \in U} c_u x^u$ the *initial form* $\text{in}_\omega(f)$ is the sum of all $c_u x^u$ such that $\omega \cdot u = \max_{v \in U}(\omega \cdot v)$. We call $\max_{v \in U}(\omega \cdot v)$ the *ω-degree* of f.

When finding initial forms of f it is advantageous to draw $NP(f)$.

Example 1.4.2 Let $f = x^3 - x^3 y + 3x^3 y^2 + 7x^2 y^4 - xy + y \in \mathbb{Q}[x, y]$. Then

- $\text{in}_{\prec_{\text{lex}}}(f) = 3x^3 y^2$,
- $\text{in}_{(1,0)}(f) = x^3 - x^3 y + 3x^3 y^2$,
- $\text{in}_{(100,1)}(f) = 3x^3 y^2$,
- $\text{in}_{\prec_{grevlex}}(f) = 7x^2 y^4$, and
- $\text{in}_{(1,1)}(f) = 7x^2 y^4$.

See Figure 3.

Lemma 1.4.3 *Let $\prec$ be a term ordering, $\omega \in \mathbb{R}^n$ and $f, g \in k[x_1, \ldots, x_n]$. Then*

- $\text{in}_\omega(fg) = \text{in}_\omega(f)\text{in}_\omega(g)$, *and*
- *if $f \neq 0 \neq g$ then* $\text{in}_\prec(fg) = \text{in}_\prec(f)\text{in}_\prec(g)$.

Proof. Left to the reader. □

1.5 The division algorithm

If $n = 1$ and we have only one generator for the ideal $I = \langle g \rangle$, then we can check if a given polynomial f is in I by running the well-known polynomial division algorithm on f, dividing by g. The remainder is 0 if and only if $f \in I$.

In this section we generalize the division algorithm to more variables and more polynomials. Unfortunately, doing so, we loose the above important property. We can get a non-zero remainder even if f is I.

Algorithm 1.5.1 (Polynomial Division)
Input: *A polynomial* $f \in k[x_1, \dots, x_n]$ *and a list of polynomials* $\{f_1, \dots, f_s\}$ *with* $f_i \in k[x_1, \dots, x_n] \setminus \{0\}$ *and a term order* $\prec$*.*
Output: *A* remainder $r \in k[x_1, \dots, x_n]$ *and* $a_1, \dots, a_s \in k[x_1, \dots, x_n]$ *such that* $f = r + \sum_i a_i f_i$ *with no term of* r *divisible by any of* $\text{in}_\prec(f_1), \dots, \text{in}_\prec(f_s)$*. Furthermore, if* $f \neq 0$ *then every term* A *of* a_i *satisfies* $\text{in}_\prec(Af_i) \preceq \text{in}_\prec(f)$*.*

```
• For $i = 1, \dots, s$ let $a_i := 0$.
• Let $r := 0$ and $p := f$.
• While($p \neq 0$)
    – Choose a term $P$ from $p$. (For example $P := \text{in}_\prec(p)$.)
    – If there exists $i$ such that $\text{in}_\prec(f_i)|P$ then
        ∗ $a_i := a_i + P/\text{in}_\prec(f_i)$
        ∗ $p := p - (P/\text{in}_\prec(f_i))f_i$
    – else
        ∗ $r := r + P$
        ∗ $p := p - P$
• Return $r, a_1, \dots, a_s$.
```

We notice that the division algorithm is *non-deterministic* since there may be more possible choices of P and i and the algorithm can choose as it likes. In particular the output of the algorithm is not unique. Making the suggested choice $P := \text{in}_\prec(p)$ often makes the algorithm terminate sooner.

Proof. We prove *correctness* and *termination.* To prove that the algorithm is correct we must show that the output satisfies the specifications. We notice that the equation $f = p + r + \sum_i a_i f_i$ is satisfied at the beginning and after every iteration of the loop. At the end $p = 0$ and the equation $f = r + \sum_i a_i f_i$ follows. We also notice that only terms which are not divisible by any $\text{in}_\prec(f_i)$ are appended to r. Finally, notice that $\text{in}_\prec(p)$ never gets $\prec$-larger during the algorithm: In the case where the condition of the if statement is true because $\text{in}_\prec(P/\text{in}_\prec(f_i)f_i) = \text{in}_\prec(P/\text{in}_\prec(f_i))\text{in}_\prec(f_i) = (P/\text{in}_\prec(f_i))\text{in}_\prec(f_i) = \text{in}_\prec(P) \preceq \text{in}_\prec(p)$. In the second case because a term is removed from p. Consequently, any term $P/\text{in}_\prec(f_i)$ introduced to a_i satisfies $\text{in}_\prec((P/\text{in}_\prec(f_i))f_i) = \text{in}_\prec(P) \preceq \text{in}_\prec(p) \preceq \text{in}_\prec(f)$. Thus the output satisfies the specifications.

To prove that the algorithm *terminates* we observe that if we always make the choice $P := \text{in}_\prec(p)$, then at each iteration the initial term $\text{in}_\prec(p)$ keeps getting strictly smaller in the $\prec$ ordering: either because $-\text{in}_\prec(P/\text{in}_\prec(f_i)f_i) = -P$ cancels with $P = \text{in}_\prec(p)$ or because $P = \text{in}_\prec(p)$ is moved from p to r. The set of $\text{in}_\prec(p)$ appearing during a run of the algorithm must have a smallest element by Lemma 1.3.7. Hence the algorithm cannot continue forever.

If we do not consistently make the choice $P := \text{in}_\prec(p)$ then the proof is trickier: We will first assume that f is a single term. We let P_i denote the value of P in the ith iteration, starting at $i = 1, 2, \dots$. We now define a tree on the set of is appearing, namely we connect i to j if P_j was introduced to p when processing P_i. To be precise, we only connect i to j if the monomial of P_j was not present in p immediately before processing P_i. This ensures that j has just a single parent and that we therefore build a tree. We notice that $P_j \prec P_i$ if (i, j) is an edge. By Lemma 1.3.7 every path starting at the root must be finite. By Lemma 1.5.2 below we get that the tree is finite. Hence the algorithm has to terminate. If f was not a single polynomial, then the argument still works by adding f as an artificial vertex 0 of the tree, and adding an edge from 0 to i if i has no parent. □

Lemma 1.5.2 *Let T be a tree with the property that any vertex v has only finitely many child vertices. Suppose that T does not contain an infinite path starting at the root. Then T has only finitely many vertices.*

Proof. Suppose that T had an infinite number of vertices. We will construct an infinite path in T starting at the root v_0. The root v_0 has only finitely many children, so one of its children must have infinitely many vertices below it. Let's call that child v_1. We repeat the process with v_1. Since there are infinitely many vertices below it, one of the children v_2 has infinitely many vertices. The path $v_0, v_1, v_2, v_2, \dots$ constructed in this way is infinite. This is a contradiction. □

Example 1.5.3 Let $\prec = \prec_{\text{lex}}, f = x^2y^3 - 2y, f_1 = \underline{xy} - y, f_2 = \underline{x^2y^2} - x - 1, f_3 = \underline{x} - 2y + 1$. Here the initial terms have been underlined. We list some possible runs of the division algorithm. We keep track of the values p. A $\rightarrow$ means reducing by the subscript. A $\downarrow$ means moving the subscript to the remainder.

- $x^2y^3 - 2y \rightarrow_{f_1} xy^3 - 2y \rightarrow_{f_1} y^3 - 2y \downarrow_{y^3} -2y \downarrow_{-2y} 0 \qquad r = y^3 - 2y$
- $x^2y^3 - 2y \rightarrow_{f_2} xy + y - 2y = xy - y \rightarrow_{f_1} 0 \qquad r = 0$
- $x^2y^3 - 2y \downarrow_{-2y} x^2y^3 \rightarrow_{f_3} 2xy^4 - xy^3 \rightarrow_{f_1} 2y^4 - xy^3 \rightarrow_{f_1} 2y^4 - y^3 \downarrow_{-y^3}$ $2y^4 \downarrow_{2y^4} 0 \qquad r = 2y^4 - y^3 - 2y.$

If we keep track of the coefficient polynomials a_i in the second run, then we get the identity $x^2y^3 - 2y = y(x^2y^2 - x - 1) + 1(xy - y)$ proving that $x^2y^2 - 2y \in \langle xy - y, x^2y^2 - x - 1, x - 2y + 1\rangle$.

As the example shows, whether the remainder of the division is zero depends on the actual choices made in the algorithm. We would like to have a notion of "reduces to zero" which is independent of the division algorithm:

Definition 1.5.4 Let $f, f_1, f_2, \dots, f_s \in k[x_1, \dots, x_n]$ be polynomials and $\prec$ a term ordering. We say that f *reduces to zero* modulo $f_1, \dots, f_s$ if there exists $a_1, \dots, a_s$ such that $f = \sum_i a_i f_i$ and $\text{in}_\prec(f_i)\text{in}_\prec(a_i) \preceq \text{in}_\prec(f)$ for all i with $a_i f_i \neq 0$.

Lemma 1.5.5 *If the remainder produced by some run of the division algorithm on $f, f_1, \dots, f_s$ is 0 then f reduces to zero modulo $f_1, \dots, f_s$.*

Proof. Algorithm 1.5.1 produces the desired expression because $f = 0 + \sum_i a_i f_i$. All we need to check is that for $a_i \neq 0$ we have $\text{in}_\prec(f_i)\text{in}_\prec(a_i) \preceq \text{in}_\prec(f)$. But this also follows from the specifications of the algorithm. □

1.6 Gröbner bases

Example 1.5.3 showed that the output of the division algorithm does not always have the desired properties. In this section we introduce the notion of Gröbner bases. We will see in Lemma 1.6.6 that Algorithm 1.5.1 is well-behaved if run with a Gröbner basis $\{f_1, \ldots, f_s\}$.

Definition 1.6.1 Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal. Let $\prec$ be a term ordering and $\omega \in \mathbb{R}^n$. We define the *initial ideals* of I:

- $\text{in}_\prec(I) := \langle \text{in}_\prec(f) : f \in I \setminus \{0\}\rangle$ and
- $\text{in}_\omega(I) := \langle \text{in}_\omega(f) : f \in I\rangle$.

We observe that $\text{in}_\prec(I)$ is always a monomial ideal, while $\text{in}_\omega(I)$ might not be:

Example 1.6.2 Let $I := \langle x^2+y^2+x^2y, x^2+xy+x^2y\rangle \subseteq \mathbb{Q}[x,y]$ and $\omega = (1,1)$. Then it is easy to see that x^2y is an initial form of an element of I and must be in $\text{in}_\omega(I)$. But actually x^2y is not enough to generate $\text{in}_\omega(I)$. For example $\text{in}_\omega((x^2+y^2+x^2y)-(x^2+xy+x^2y)) = \text{in}_\omega(y^2-xy) = y^2-xy$. In fact we claim (without proof) that $\text{in}_\omega(I) = \langle y^3, xy-y^2, x^3\rangle$. We also have $\text{in}_{\prec_{\text{grlex}}}(I) = \langle y^3, xy, x^3\rangle$.

As the example shows, it is not always easy to find the initial ideal. Later we will see how to do this for term orders (Algorithm 1.7.3) and vectors (Corollary 4.4.4).

Definition 1.6.3 Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $\prec$ be a term ordering. A finite set $\{f_1, \ldots, f_s\} \subseteq I$ is called a *Gröbner basis* for I with respect to $\prec$ if $\langle \text{in}_\prec(f_1), \ldots, \text{in}_\prec(f_s)\rangle = \text{in}_\prec(I)$.

Example 1.6.4 The set $\{x^2+y^2+x^2y, x^2+xy+x^2y\}$ is not a Gröbner basis for the ideal I in Example 1.6.2 with respect to $\prec_{\text{grlex}}$ since the initial forms of elements in the set are $x^2y = x^2y$. Which do not generate $\text{in}_{\prec_{\text{grlex}}}(I) = \langle y^3, xy, x^3\rangle$. The set $\{y^3+y^2+x^2, xy-y^2, x^3+y^2+x^2\} \subseteq I$ is a Gröbner basis for I since its initial terms generate $\text{in}_{\prec_{\text{grlex}}}(I) = \langle y^3, xy, x^3\rangle$.

Lemma 1.6.5 *If $\{f_1, \ldots, f_s\}$ is a Gröbner basis for an ideal $I \subseteq k[x_1, \ldots, x_n]$ with respect to a term order $\prec$ then $I = \langle f_1, \ldots, f_s\rangle$.*

Proof. We need to show that $I \subseteq \langle f_1, \ldots, f_s\rangle$, so we pick $f \in I$. Let r be the remainder produced by a run of the division algorithm (Algorithm 1.5.1). Notice that $r \in I$. Suppose that $r \neq 0$. Then the term $\text{in}_\prec(r) \in \text{in}_\prec(I) = \langle \text{in}_\prec(f_1), \ldots, \text{in}_\prec(f_s)\rangle$. This means that some $\text{in}_\prec(f_i)$ divides $\text{in}_\prec(r)$. This contradicts the properties of Algorithm 1.5.1. Hence $r = 0$, which implies that the polynomials produced in the algorithm satisfy $f = r + \sum_i a_i f_i = \sum_i a_i f_i \in \langle f_1, \ldots, f_s\rangle$. □

Lemma 1.6.6 *Let $\{f_1, \ldots, f_s\}$ be a Gröbner basis for an ideal $I \subseteq k[x_1, \ldots, x_n]$ with respect to a term ordering $\prec$. The remainder produced by the division algorithm (Algorithm 1.5.1) when run on a polynomial f is independent of the choices performed in the run.*

Proof. Suppose that one run gave r and another gave r'. Then $r + \sum_i a_i f_i = f = r' + \sum_i a'_i f_i$ imply $r - r' \in I$. If $r \neq r'$ then there would be a leading term $\text{in}_\prec(r - r') \in \text{in}_\prec(I)$ which is not divisible by any $\text{in}_\prec(f_i)$. This contradicts $\langle \text{in}_\prec(f_1), \ldots, \text{in}_\prec(f_s)\rangle = \text{in}_\prec(I)$. □

Gröbner bases have the properties we want. We first give a non-constructive proof of their existence. In the next section we present a concrete algorithm.

Proposition 1.6.7 *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $\prec$ a term ordering on $k[x_1, \ldots, x_n]$. Then I has a Gröbner basis with respect to $\prec$.*

Proof. The ideal $\text{in}_\prec(I)$ is a monomial ideal. By Dickson's Lemma 1.2.4 it has the form $\langle x^{u_1}, \ldots, x^{u_s}\rangle$. By Exercise 10 of Sheet 1, for every i there exists $f_i \in I$ such that $\text{in}_\prec(f_i) = x^{u_i}$. The set $\{f_1, \ldots, f_s\} \subseteq I$ is a Gröbner basis of I w.r.t. $\prec$ because $\text{in}_\prec(I) = \langle x^{u_1}, \ldots, x^{u_s}\rangle = \langle \text{in}_\prec(f_1), \ldots, \text{in}_\prec(f_s)\rangle$. □

In particular we have proved Hilbert's Basis Theorem 1.1.6. Furthermore:

Corollary 1.6.8 *For a field k the polynomial ring $k[x_1, \ldots, x_n]$ is Noetherian. That is if $I_1 \subseteq I_2 \subseteq I_3 \ldots$ are ideals in $k[x_1, \ldots, x_n]$ then there exists j such that $I_j = I_{j+1} = I_{j+2} = \cdots$.*

Proof. We use the argument of the proof of Corollary 1.2.5. □

Definition 1.6.9 Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $\prec$ a term ordering. A monomial $x^u \notin \text{in}_\prec(I)$ is called a *standard monomial* (w.r.t. I and $\prec$). We let $\text{std}_\prec(I)$ denote the set of all standard monomials.

If we have a Gröbner basis for an ideal I one of the interpretations of the division algorithm is that it writes a polynomial f as a linear combination of standard monomials modulo I. The remainder is called the *normal form* of f.

Lemma 1.6.10 *The cosets of the standard monomials $\text{std}_\prec(I)$ form a k-vector space basis $\{[x^u] : x^u \in \text{std}_\prec(I)\}$ of the quotient ring $k[x_1, \ldots, x_n]/I$.*

Proof. Let $S = k[x_1, \ldots, x_n]$. To prove that the set spans S/I, take a vector $[f] \in S/I$ with $f \in S$. The Division Algorithm 1.5.1 gives an expression $f = r + \sum_{i=1}^s a_i f_i$ with $r = \sum_{x^u \in \text{std}_\prec(I)} c_u x^u$ and $c_u \in k$, implying $f - \sum_{i=1}^s a_i f_i = \sum_{x^u \in \text{std}_\prec(I)} c_u x^u$. Therefore $[f] = [f - \sum_{i=1}^s a_i f_i] = [\sum_{x^u \in \text{std}_\prec(I)} c_u x^u] = \sum_{x^u \in \text{std}_\prec(I)} c_u [x^u]$. This proves that $\{[x^u] : x^u \in \text{std}_\prec(I)\}$ spans S/I.

To prove independence of the set $\{[x^u] : x^u \in \text{std}_\prec(I)\}$, suppose that we had $\sum_{x^u \in \text{std}_\prec(I)} c_u [x^u] = [0]$ with $c_u \in k$. Then $\sum_{x^u \in \text{std}_\prec(I)} c_u x^u \in I$. If some c_u was non-zero, then taking initial term we get a standard monomial in the initial ideal: $\text{in}_\prec(\sum_{x^u \in \text{std}_\prec(I)} c_u x^u) = c_v x^v \in \text{in}_\prec(I)$ for some v — a contradiction. Therefore $c_u = 0$ for all u and the vectors must be independent. □

Corollary 1.6.11 *Let $\{f_1, \dots, f_s\}$ be a Gröbner basis for an ideal $I \subseteq k[x_1, \dots, x_n]$ with respect to a term ordering $\prec$. A polynomial f belongs to I if and only if the remainder produced by the division algorithm is* 0.

Proof. If the remainder is 0, then we have $f = 0 + \sum_i a_i f_i \in I$. On the other hand, if $f \in I$ then the remainder r produced by Algorithm 1.5.1 is a linear combination $r = \sum_{a \in \mathrm{std}_\prec(I)} c_a a$ with $c_a \in k$ and we have $[0] = [f] = [r] = [\sum_{a \in \mathrm{std}_\prec(I)} c_a a] = \sum_{a \in \mathrm{std}_\prec(I)} c_a [a]$ in $k[x_1, \dots, x_n]/I$. By Lemma 1.6.10 the standard monomials are independent, which shows $c_a = 0$ for all a. Hence $r = 0$. □

An ideal can have many Gröbner bases with respect to the same ordering as the following example shows.

Example 1.6.12 The ideal I of Example 1.6.2 has

$$\{\underline{y^3} + y^2 + x^2, \underline{2xy} - 2y^2, \underline{x^3} - x^2y, \underline{x^2y} + x^2 + y^2\} \subseteq I$$

as a Gröbner basis w.r.t. $\prec_{\mathrm{grlex}}$ because the initial terms generate $\mathrm{in}_{\prec_{\mathrm{grlex}}}(I) = \langle x^3, xy, y^3 \rangle$. Because $\mathrm{in}_{\prec_{\mathrm{grlex}}}(I)$ is generated by just the three monomials x^3, xy and y^3, we can leave out $x^2y + x^2 + y^2$ and will still have a Gröbner basis:

$$\{\underline{y^3} + y^2 + x^2, \underline{2xy} - 2y^2, \underline{x^3} - x^2y\} \subseteq I.$$

This basis is called a *minimal* Gröbner basis. We may also scale the polynomials to make the coefficients of the initial terms 1:

$$\{\underline{y^3} + y^2 + x^2, \underline{xy} - y^2, \underline{x^3} - x^2y\} \subseteq I.$$

To get an even nicer Gröbner basis, we observe that the *tail* of $\underline{x^3} - x^2y$, being $-x^2y$, contains a monomial which is divisible by an initial term. We perform division of $-x^2y$ modulo the polynomials and get the unique remainder $x^2 + y^2$. We now substitute this tail, and get the *reduced* Gröbner basis

$$\{\underline{y^3} + y^2 + x^2, \underline{xy} - y^2, \underline{x^3} + x^2 + y^2\} \subseteq I.$$

The precise definition of minimal and reduced follows below.

Definition 1.6.13 The Gröbner basis of Definition 1.6.3 is called *minimal* if if $\{\mathrm{in}_\prec(f_1), \dots, \mathrm{in}_\prec(f_s)\}$ is a minimal generating set for $\mathrm{in}_\prec(I)$. That is, no element can be left out. If furthermore, for every i no term of $f_i - \mathrm{in}_\prec(f_i)$ is divisible by any $\mathrm{in}_\prec(f_j)$ and $\mathrm{in}_\prec(f_i)$ has coefficient 1 then $\{f_1, \dots, f_s\}$ is called a *reduced* Gröbner basis.

Example 1.6.12 shows how to turn a Gröbner basis into a reduced one. We will later state these processes as Algorithms 1.7.8 and 1.7.9.

Proposition 1.6.14 *Every ideal has at most one reduced Gröbner basis with respect to a given term order* $\prec$.

Proof. By Lemma 1.2.3 the initial ideal $\text{in}_\prec(I)$ has a unique minimal monomial generating set $\{x^{u_1}, \dots, x^{u_s}\}$. Therefore every reduced Gröbner basis w.r.t. $\prec$ must consist of $f_1, \dots, f_s$ where $\text{in}_\prec(f_i) = x^{u_i}$ and all other monomials of f_i belong to $\text{std}_\prec(I)$. Suppose there were two polynomials f_i and f_i' in I with $\text{in}_\prec(f_i) = x^{u_i} = \text{in}_\prec(f_i')$ and all other monomials in $\text{std}_\prec(I)$. If $f_i - f_i'$ is non-zero, the monomial of $\text{in}_\prec(f - f')$ is in $\text{std}_\prec(I)$ which is a contradiction. Therefore there is only one possible choice of f_i. □

The unique reduced Gröbner basis of I with respect to $\prec$ is denoted $\mathcal{G}_\prec(I)$.

1.7 Buchberger's Algorithm

Proposition 1.6.7 says that every ideal ideal $I \subseteq k[x_1, \dots, x_n]$ has a Gröbner basis with respect to every term order. In this section we will show how to construct such a Gröbner basis given generators for I.

Definition 1.7.1 Let $\prec$ be a term order and f, g be two non-zero polynomials in $k[x_1, \dots, x_n]$. We define the S-polynomial of f and g:

$$S_\prec(f, g) = \frac{\text{lcm}(\text{in}_\prec(f), \text{in}_\prec(g))}{\text{in}_\prec(f)} f - \frac{\text{lcm}(\text{in}_\prec(f), \text{in}_\prec(g))}{\text{in}_\prec(g)} g$$

where $\text{lcm}(cx^u, c'x^v) := x^{\max(u,v)}$ (maximum is taken coordinate-wise).

We observe that the leading terms of the two parts of the S-polynomial cancel. In particular, every term of $S_\prec(f, g)$ is $\prec$-smaller than $\text{lcm}(\text{in}_\prec(f), \text{in}_\prec(g))$.

Theorem 1.7.2 *Let $G = \{g_1, \dots, g_s\} \subseteq k[x_1, \dots, x_n] \setminus \{0\}$ and $\prec$ be a term order. If for all i, j the polynomial $S_\prec(g_i, g_j)$ reduces to zero modulo G, then G is a Gröbner basis for $I := \langle G \rangle$.*

Proof. Suppose G was not a Gröbner basis. Then there exists $x^u \in \text{in}_\prec(I) \setminus \langle \text{in}_\prec(g) : g \in G \rangle$. By Exercise 10 of Sheet 1 there exists $f \in \langle G \rangle$ with $x^u = \text{in}_\prec(f)$. We may express f as a finite sum $\sum_i a_i g_i$ with a_i being a term and the g_i's being (possibly repeated) elements of G. But let us not just pick an arbitrary such expression, but one where the largest $\text{in}_\prec(a_i g_i)$ appearing is smallest possible. This can be done since $\prec$ is a well-order (Lemma 1.3.7). Now consider a $\prec$-largest term $cx^v = \text{in}_\prec(a_j g_j)$ appearing in $\sum_i a_i g_i$ before summing up. This term must cancel since otherwise $x^u = x^v \in \langle \text{in}_\prec(g) : g \in G \rangle$. Hence we find j' with $c'x^v = \text{in}_\prec(a_{j'} g_{j'})$. That the cancellation occurs implies that $a_j g_j - \frac{c}{c'} a_{j'} g_{j'}$ is a multiple of $S_\prec(g_j, g_{j'})$ which reduces to zero, meaning that $a_j g_j - \frac{c}{c'} a_{j'} g_{j'} = \sum_l d_l f_l$ for some $f_l \in G$ and d_l with $\text{in}_\prec(f_l d_l) \prec x^v$. In the sum $\sum_i a_i g_i$ we now replace $a_j g_j$ by $\sum_l d_l f_l$ and add $\frac{c}{c'} a_{j'}$ to the coefficient of $g_{j'}$ (possibly making this summand disappear). This removes at least one appearance of x^v, and only introduces $\prec$-smaller terms. We repeat this process until no more x^v appear. We now have a contradiction since the expression $\sum_i a_i g_i$ for f has the largest terms $\prec$-smallest, but we have an expression with smaller largest terms. Consequently, G is a Gröbner basis with respect to $\prec$. □

Algorithm 1.7.3 (Buchberger's Algorithm)
Input: *A generating set* $F = \{f_1, \dots, f_t\} \subseteq k[x_1, \dots, x_n] \setminus \{0\}$ *for an ideal* I *and a term order* $\prec$.
Output: *A Gröbner basis for* I *with respect to* $\prec$.

```
• $G := F$
• While $\exists g, h \in G$ such that $S_\prec(g, h)$ does not reduce to zero modulo $G$.
    – Let $r$ be a remainder produced by the division algorithm (Algorithm 1.5.1)
      run on $S_\prec(g, h)$ and $G$
    – Let $G := G \cup \{r\}$.
```

Proof. To guarantee that $S_\prec(g, h)$ reduces to zero modulo G we can use the Division Algorithm 1.5.1 and Lemma 1.5.5. (A technical remark: If the remainder is non-zero then it is not clear that $S_\prec(g, h)$ does not reduce to zero modulo G. However, it is clear that G is not yet a Gröbner basis (Corollary 1.6.11) and it is safe to add the remainder to G, ensuring that $S_\prec(g, h)$ now reduces to zero.)

If the algorithm terminates, then by Theorem 1.7.2 the set G is a Gröbner basis for $\langle G \rangle$. Furthermore $\langle G \rangle = I$ since we only add elements of I to G. To show that the algorithm terminates we observe that in every step the monomial ideal $\langle \text{in}_\prec(g) : g \in G \rangle$ keeps getting strictly larger because $\text{in}_\prec(r)$ is produced from the division algorithm with the property that no $\text{in}_\prec(g)$ divides it. By Corollary 1.2.5 this cannot go on forever. □

Example 1.7.4 We continue Example 1.6.2, but starting with the generating set $\{g_1, g_2\} = \{\underline{xy} - y^2, \underline{y^3} + x^2 + y^2\}$ for I and with $\prec$ being the degree reverse lexicographic ordering. To compute a Gröbner basis we first reduce $S_\prec(g_1, g_2) = \underline{-y^4} - x^3 - xy^2$ modulo $\{g_1, g_2\}$ using the division algorithm and get remainder $\underline{-x^3} + y^3 =: g_3$. Since the remainder is not zero, we add it to the generating set. We now check that $S_\prec(g_1, g_3) = \underline{-x^2y^2} + y^4$ gives remainder zero modulo $\{g_1, \dots, g_3\}$. Finally we check that $S_\prec(g_2, g_3)$ reduces to zero (possibly using Lemma 1.7.6 below). We conclude that $\{g_1, \dots, g_3\} = \{\underline{xy} - y^2, \underline{y^3} + x^2 + y^2, \underline{-x^3} + y^3\}$ is a Gröbner basis. In particular $\text{in}_\prec(I) = \langle xy, y^3, x^3 \rangle$.

Remark 1.7.5 From the proof it follows that if we for some reason know that $S_\prec(g, h)$ reduces to zero in the sense of Definition 1.5.4 then we can simply ignore that S-polynomial in the algorithm. The following lemma becomes useful.

Lemma 1.7.6 *Let* $f, g \in k[x_1, \dots, x_n] \setminus \{0\}$ *and* $\prec$ *a term ordering. If for all* i:$x_i \nmid \text{in}_\prec(f) \vee x_i \nmid \text{in}_\prec(g)$ *then* $S_\prec(f, g)$ *reduces to zero modulo* f *and* g.

Proof. We observe that $S_\prec(sf, tg) = S_\prec(f, g)$ for $s, t \in k \setminus \{0\}$. Hence, we may assume that the coefficients of $\text{in}_\prec(f)$ and $\text{in}_\prec(g)$ are both 1. We then have

$$S_\prec(f, g) = \frac{\text{lcm}(\text{in}_\prec(f), \text{in}_\prec(g))}{\text{in}_\prec(f)} f - \frac{\text{lcm}(\text{in}_\prec(f), \text{in}_\prec(g))}{\text{in}_\prec(g)} g$$

$$= \frac{\text{in}_\prec(f)\text{in}_\prec(g)}{\text{in}_\prec(f)} f - \frac{\text{in}_\prec(f)\text{in}_\prec(g)}{\text{in}_\prec(g)} g = \text{in}_\prec(g) f - \text{in}_\prec(f) g$$

$$= (\text{in}_\prec(g)f - gf) - (\text{in}_\prec(f)g - gf) = (f - \text{in}_\prec(f))g - (g - \text{in}_\prec(g))f.$$

By to Definition 1.5.4 we are done if f or g is a single term. If not it suffices argue that $\text{in}_\prec((f-\text{in}_\prec(f))g)$ and $\text{in}_\prec((g-\text{in}_\prec(g))f)$ are smaller than or equal to $\text{in}_\prec(S_\prec(f,g))$ in the $\prec$ ordering. If the exponents of $\text{in}_\prec((f-\text{in}_\prec(f))g) = \text{in}_\prec(f-\text{in}_\prec(f))\text{in}_\prec(g)$ and $\text{in}_\prec((g-\text{in}_\prec(g))f) = \text{in}_\prec(g-\text{in}_\prec(g))\text{in}_\prec(f)$ are equal, then we conclude (using the assumption that $\text{in}_\prec(f)$ and $\text{in}_\prec(g)$ have no common monomial factor) that $\text{in}_\prec(f)|\text{in}_\prec(f-\text{in}_\prec(f))$. This contradicts the properties of $\prec$ being a term order. Hence $\text{in}_\prec((f-\text{in}_\prec(f))g)$ and $\text{in}_\prec((g-\text{in}_\prec(g))f)$ have different exponent vectors and the largest of these cannot cancel when subtracting. Therefore the largest term also appears in $S_\prec(f,g)$. □

Example 1.7.7 Using Lemma 1.7.6 it is easy to check that $\{\underline{x^2}+2xy+y^3, \underline{3y^2}+3x+5\}$ is a Gröbner basis with respect to $\prec_{(5,3)^t}$.

It is common to extend Buchberger's algorithm with the following two steps to compute the reduced Gröbner basis $\mathcal{G}_\prec(I)$, thereby making the output unique.

Algorithm 1.7.8 (Minimizing a Gröbner basis)
Input: *A Gröbner basis $G \subseteq k[x_1,\dots,x_n]$ w.r.t. some term order $\prec$.*
Output: *A minimal Gröbner basis G' for $\langle G\rangle$ w.r.t. $\prec$.*

- $G' := G$
- *While it is possible to remove a $g \in G'$ from G', and still keep the equality $\langle \text{in}_\prec(g) : g \in G\rangle = \langle \text{in}_\prec(g) : g \in G'\rangle$, do so.*

Proof. The set remains a Gröbner basis for $\langle G\rangle$ since $\langle \text{in}_\prec(g) : g \in G'\rangle = \text{in}_\prec\langle G\rangle$. It is minimal since no further g can be deleted. □

Algorithm 1.7.9 (Autoreducing a Gröbner basis)
Input: *A minimal Gröbner basis $G' \subseteq k[x_1,\dots,x_n]$ w.r.t. some term order $\prec$.*
Output: *The reduced Gröbner basis $\mathcal{G}_\prec(\langle G'\rangle)$.*

- *Substitute each $g \in G'$ by $\text{in}_\prec(g) + r$, where r is the unique remainder produced by Algorithm 1.5.1 when run on the tail $g - \text{in}_\prec(g)$ and G'.*

1.8 Elimination

In Section 1.1 we stated three problems for polynomial rings which can be solved using Gröbner bases. We have already proved Hilbert's Basis Theorem 1.1.6 and shown how Gröbner bases can be used to compute in the quotient ring $k[x_1,\dots,x_n]/I$ (Corollary 1.6.11 and Exercise 5 on Sheet 2). We will now discuss how to solve polynomial equations. The technique presented works particularly well if the equations have only finitely many solutions over $\mathbb{C}$.

Proposition 1.8.1 *Let $I \subseteq k[x_1,\dots,x_n]$ be an ideal. Let G be a Gröbner basis of I with respect to $\prec_{\text{lex}}$. For $l = 1,\dots,n$ we have $G' := G \cap k[x_l,\dots,x_n]$ is a Gröbner basis for the* elimination ideal $I \cap k[x_l,\dots,x_n] \subseteq k[x_l,\dots,x_n]$.

Proof. Clearly, $G' \subseteq I \cap k[x_l, \ldots, x_n]$ and $\langle \mathrm{in}_{\prec_{\mathrm{lex}}}(g) : g \in G' \rangle \subseteq \mathrm{in}_{\prec_{\mathrm{lex}}}(I \cap k[x_l, \ldots, x_n])$. It remains to show that $\langle \mathrm{in}_{\prec_{\mathrm{lex}}}(g) : g \in G' \rangle \supseteq \mathrm{in}_{\prec_{\mathrm{lex}}}(I \cap k[x_l, \ldots, x_n])$. Let x^u be a monomial in $\mathrm{in}_{\prec_{\mathrm{lex}}}(I \cap k[x_l, \ldots, x_n])$. Then $x^u \in \mathrm{in}_{\prec_{\mathrm{lex}}}(I)$. Since G is a Gröbner basis, there must exist $g \in G$ such that $\mathrm{in}_{\prec_{\mathrm{lex}}(g)}|x^u$. Since x^u contains no x_j with $j < l$, this must also be the case for $\mathrm{in}_{\prec_{\mathrm{lex}}(g)}$. By the properties of the term order, no term from g can contain an x_j with $j < l$. Hence $g \in G'$, proving $x^u \in \langle \mathrm{in}_{\prec_{\mathrm{lex}}}(g) : g \in G' \rangle$. □

We can use Gröbner bases for solving polynomial equations:

Example 1.8.2 We wish to compute the solutions to the system $x^2 + y^2 = 1$ and $x^2 + y^2 - x - y = 2$. Let $I = \langle x^2 + y^2 - 1, x^2 + y^2 - x - y - 2 \rangle \subseteq \mathbb{C}[x, y]$. We compute the lexicographic Gröbner basis

$$\{y^2 + y, x + 1 + y\}$$

(which is an equivalent system of equations) and conclude that $I \cap \mathbb{C}[y] = \langle y^2 + y \rangle$. From this we conclude that $y = 0$ or $y = -1$. Substituting we get

$$V(I) = \{(-1, 0), (0, -1)\}.$$

Why did every solution of the elimination ideal extend to a solution of the ideal?

We show two examples where this is not the case:

Example 1.8.3 The set $\{y^2 - y, xy - y, x^2 + 1 - 2y\}$ is a lexicographic Gröbner basis for an ideal $I \subseteq \mathbb{R}[x, y]$. We solve $y^2 - y = 0$ and see that $y = 0$ and $y = 1$ are solutions. The point $(1, 1)$ is in $V(I) \subseteq \mathbb{R}^2$. However, there is no solution with $y = 0$ over the real numbers.

Example 1.8.4 Let $I = \langle xy - 1 \rangle \subseteq \mathbb{C}[x, y]$. The generator is already a lexicographic Gröbner basis. We conclude that $I \cap \mathbb{C}[y] = \langle \emptyset \rangle = \{0\}$. Any choice of y gives a solution to the elimination ideal. If we choose a value for y then the equation $xy - 1 = 0$ tells us the value of x. However, if $y = 0$ was chosen there is no solution for x.

The first example shows that the ideal must be algebraically closed for all solutions to extend, while the second shows that it is possible that not every point lifts in the case where we have more solutions than a finite set of points.

In the rest of this subsection we use the complex numbers $\mathbb{C}$, but any algebraically closed field will suffice. We will use the following classic result without proof:

Theorem 1.8.5 (Hilbert's Nullstellensatz) *Let $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal. If $f \in \mathbb{C}[x_1, \ldots, x_n]$ is zero on all points in $V(I)$ then there exists $N \in \mathbb{N}$ such that $f^N \in I$.*

Corollary 1.8.6 *Let $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal and $\prec$ a term ordering. Then $V(I) \subseteq \mathbb{C}^n$ is a finite set $\Leftrightarrow \dim_{\mathbb{C}}(\mathbb{C}[x_1, \ldots, x_n]/I) < \infty \Leftrightarrow |\mathrm{std}_\prec(I)| < \infty$.*

Proof. The last two statements are equivalent because the standard monomials form a vector space basis of $\mathbb{C}[x_1, \dots, x_n]/I$ by Lemma 1.6.10. If $V(I) \subseteq \mathbb{C}^n$ is infinite and $\dim_{\mathbb{C}}(\mathbb{C}[x_1, \dots, x_n]/I) =: d$ finite then we choose $d + 1$ point in $V(I)$ and for each point p_i we construct a polynomial $f_i \in \mathbb{C}[x_1, \dots, x_n]$ which take the value 1 at p_i and zero on all other chosen points. These $d + 1$ polynomials are linearly independent in $\mathbb{C}[x_1, \dots, x_n]/I$ since all $f \in I$ vanishes at the points. This contradicts the space having dimension d.

On the other hand suppose $V(I) \subseteq \mathbb{C}^n$ is finite. For each coordinate direction x_i we choose a polynomial $f_i \in \mathbb{C}[x_i]$ being zero on the projection of $V(I)$ to that coordinate. We also have that f_i is zero on $V(I)$. By Hilbert's Nullstellensatz there exists $N_i \in \mathbb{N}$ such that $f_i^{N_i} \in I$. The term $\text{in}_{\prec}(f)$ only involves the variable x_i. Therefore, the ith exponent of standard monomial in $\text{std}_{\prec}(I)$ is bounded. Since this holds on all coordinates x_i, there can be only finitely many standard monomial. □

Corollary 1.8.7 *Let $I \subseteq \mathbb{C}[x_1, \dots, x_n]$ be an ideal with $\dim_{\mathbb{C}}(\mathbb{C}[x_1, \dots, x_n]/I) < \infty$ and $J = I \cap \mathbb{C}[x_n]$. If $a_n \in V(J) \subseteq \mathbb{C}^1$ then there exists $a_1, \dots, a_{n-1} \in \mathbb{C}$ such that $(a_1, \dots, a_n) \in V(I) \subseteq \mathbb{C}^n$.*

Proof. The variety $V(I)$ is finite set of points, and so is the projection of these points onto the last coordinate. Let $p_1, \dots, p_m$ be these projected points. The polynomial $f = \prod_{i=1}^{m}(x_n - p_i)$ is zero on the projected points. If a_n does not lift, then f is non-zero on a_n. The polynomial is zero on all points in $V(I)$ and by Hilbert's Nullstellensatz there exists $N \in \mathbb{N}$ such that $f^N \in I$. It follows that $f^N \in J$. But $f^N(a_n) \neq 0$. This contradicts that $a_n \in V(J)$. □

In general the elimination ideal defines the "Zariski closure" of the projection of $V(I)$. Even with the limitations described above, lexicographic Gröbner bases are the first choice of tool for solving polynomial systems algebraically.

2 Lattices, convexity and Robbiano's Theorem

We have seen in Exercise 7 of Sheet 1 that vectors can be used to construct termorders. In this section we will prove Theorem 2.3.5 which says that every termorder can be represented by a matrix.

2.1 Lattices

In this section we introduce lattices. They will be important later for toric ideals, lattice ideals, integer programming and Robbiano's characterisation of term orders. Typically a lattice will sit inside some $\mathbb{R}^n$ as a subgroup.

Definition 2.1.1 A group L is called a *lattice* if it is isomorphic to the group $\mathbb{Z}^m$ for some $m \in \mathbb{N}$. The *rank* of the lattice is the number m. A set $\{b_1, \dots, b_m\} \subseteq L$ is called a *(lattice) basis* for L if $L = \{\sum_i a_i b_i : a_i \in \mathbb{Z}\}$.

Given a subset $B \subseteq \mathbb{Z}^n$ we let $\langle B \rangle$ denote the smallest subgroup of $\mathbb{Z}^n$ containing B. We call B a *generating set* for $\langle B \rangle$. We will prove that $\langle B \rangle$ is a lattice (Theorem 2.1.2) and see how to compute a lattice basis in case B is finite (Algorithm 2.1.6).

Theorem 2.1.2 *Every subgroup $G \subseteq \mathbb{Z}^n$ is a lattice of rank at most n.*

Proof. Let for $i = 1, \dots, m$ the function $\pi_i : \mathbb{Z}^n \to \mathbb{Z}$ be the projection on the ith coordinate and $S_i := \pi_i(G \cap (\{0\}^{i-1} \times \mathbb{Z}^{n-i+1}))$. For $i = 1, \dots, m$ if $S_i \neq \{0\}$ we construct a b_i as follows and collect these in a set B. The group $S_i \subseteq \mathbb{Z}$ is generated by one element $a \in \mathbb{Z}$ (as a ring $\mathbb{Z}$ is a *principal ideal domain*). We choose $b_i \in G \cap (\{0\}^{i-1} \times \mathbb{Z}^{n-i+1})$ such that $\pi_i(b_i) = a$. In this way we construct at most n vectors b_i. We claim that $\langle B \rangle = G$ and the vectors are independent. This would prove that G is isomorphic to $\mathbb{Z}^m$ with $m = |B| \leq n$.

To show that B is independent, let $\sum_{j \in J} c_j b_j = 0$ be a dependency with $c_j \in \mathbb{Z} \setminus \{0\}$ and $J \neq \emptyset$ minimal. Let j' be the smallest element of J. We have that the j'th coordinate of $b_{j'}$ is non-zero by construction. But this is a contradiction, since all other b_j in the sum are zero on the j'th coordinate.

To show that B generates G, suppose not and let i be the largest index such that $G \cap (\{0\}^{i-1} \times \mathbb{Z}^{n-i+1}) \setminus \langle B \rangle \neq \emptyset$, and pick an element v in the difference. Since i is largest $v_i \neq 0$ and therefore $S_i \neq \{0\}$ meaning that we have introduced a b_i to B. We now subtract b_i a suitable number of times from v to get v' with $v'_i = 0$ and $v' \in (\{0\}^{i-1} \times \mathbb{Z}^{n-i+1}) \setminus B$ contradicting the maximality of i. □

Remark 2.1.3 For the experts: Every submodule of a free module of rank n over a principal ideal domain is free with rank at most n. The proof is a straight forward generalization of the proof above. Important is that the ring is a principal ideal domain. In the following we will need that the ring is a Euclidean domain which is a slightly stronger condition.

In the rest of this section we explain the theorem by doing computations with lattices using the Gauss elimination algorithm over the integers.

Let us consider a matrix $A \in \mathbb{Z}^{d\times n}$ and let a and b be two non-zero entries of the same column. Recall that the Euclidean algorithm run on two integers $a, b \in \mathbb{Z}$ produces $\gcd(a, b)$ and an invertible matrix $M \in \mathbb{Z}^{2\times 2}$ such that $\begin{bmatrix}\gcd(a,b)\\ 0\end{bmatrix} = M\begin{bmatrix}a\\ b\end{bmatrix}$. Actually the Euclidean algorithm applies row operations to $\begin{bmatrix}a\\ b\end{bmatrix}$ to obtain the gcd and a zero entry. For example

$$\begin{bmatrix}54\\ 21\end{bmatrix} \sim \begin{bmatrix}12\\ 21\end{bmatrix} \sim \begin{bmatrix}12\\ 9\end{bmatrix} \sim \begin{bmatrix}3\\ 9\end{bmatrix} \sim \begin{bmatrix}3\\ 0\end{bmatrix}$$

Applying the same operations to A we obtain a zero entry in A. Doing this systematically we get an integer matrix in row echelon form. We have described an algorithm with the following specification:

Algorithm 2.1.4 (Matrix reduction over $\mathbb{Z}$)
Input: *A matrix $A \in \mathbb{Z}^{d\times n}$.*
Output: *A matrix $A' \in \mathbb{Z}^{d\times n}$ and an invertible (over $\mathbb{Z}$) matrix $U \in \mathbb{Z}^{d\times d}$ such that $A' = UA$ and A' is in row echelon form.*

Example 2.1.5 We do the reduction:

$$\left[\begin{array}{cc|cccc} 6 & 6 & 1 & 0 & 0 & 0\\ 4 & -6 & 0 & 1 & 0 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 9 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cccc} 2 & 12 & 1 & -1 & 0 & 0\\ 4 & -6 & 0 & 1 & 0 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 9 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cccc} 2 & 12 & 1 & -1 & 0 & 0\\ 0 & -30 & -2 & 3 & 0 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 9 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \sim$$

$$\left[\begin{array}{cc|cccc} 2 & 12 & 1 & -1 & 0 & 0\\ 0 & -30 & -2 & 3 & 0 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 1 & -49 & -4 & 4 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cccc} 0 & 110 & 9 & -9 & 0 & -2\\ 0 & -30 & -2 & 3 & 0 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 1 & -49 & -4 & 4 & 0 & 1 \end{array}\right] \sim$$

$$\left[\begin{array}{cc|cccc} 1 & -49 & -4 & 4 & 0 & 1\\ 0 & -30 & -2 & 3 & 0 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 0 & 110 & 9 & -9 & 0 & -2 \end{array}\right] \sim \left[\begin{array}{cc|cccc} 1 & -49 & -4 & 4 & 0 & 1\\ 0 & 30 & 2 & -3 & 0 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 0 & 110 & 9 & -9 & 0 & -2 \end{array}\right] \sim$$

$$\left[\begin{array}{cc|cccc} 1 & -49 & -4 & 4 & 0 & 1\\ 0 & 0 & 2 & -3 & -3 & 0\\ 0 & 10 & 0 & 0 & 1 & 0\\ 0 & 110 & 9 & -9 & 0 & -2 \end{array}\right] \sim \left[\begin{array}{cc|cccc} 1 & -49 & -4 & 4 & 0 & 1\\ 0 & 10 & 0 & 0 & 1 & 0\\ 0 & 0 & 2 & -3 & -3 & 0\\ 0 & 0 & 9 & -9 & -11 & -2 \end{array}\right]$$

to get $\begin{bmatrix}1 & -49\\ 0 & 10\\ 0 & 0\\ 0 & 0\end{bmatrix} = \begin{bmatrix}-4 & 4 & 0 & 1\\ 0 & 0 & 1 & 0\\ 2 & -3 & -3 & 0\\ 9 & -9 & -11 & -2\end{bmatrix}\begin{bmatrix}6 & 6\\ 4 & -6\\ 0 & 10\\ 9 & -1\end{bmatrix}$ with $\det\begin{pmatrix}-4 & 4 & 0 & 1\\ 0 & 0 & 1 & 0\\ 2 & -3 & -3 & 0\\ 9 & -9 & -11 & -2\end{pmatrix} = \pm 1.$

Algorithm 2.1.6 (Lattice basis)
Input: *A finite generating set $\{v_1, \ldots, v_s\}$ for a subgroup $G \subseteq \mathbb{Z}^n$.*
Output: *A lattice basis $\{b_1, \ldots, b_m\}$ for G.*

- *Write the generating set in the rows of an $s \times n$ matrix A.*
- *Let $\{b_1, \ldots, b_m\}$ be the non-zero rows of A' computed by Algorithm 2.1.4.*

Proof. Since $\{b_1, \ldots, b_m\}$ was obtained from $\{v_1, \ldots, v_s\}$ by multiplying with an invertible $\mathbb{Z}$-matrix, these sets generate exactly the same group. Furthermore, coming from a reduced matrix the elements $\{b_1, \ldots, b_m\}$ are linearly independent over $\mathbb{Q}$ and must therefore generate a lattice of rank m. □

The following corollary follows immediately from the algorithm:

Corollary 2.1.7 *Every finitely generated subgroup of $\mathbb{Z}^n$ is a lattice.*

Example 2.1.8 It follows from the computation in Example 2.1.5 that the group $G \subseteq \mathbb{Z}^2$ generated by $\{(6,6),(4,-6),(0,10),(9,-1)\}$ has $\{(1,-49),(0,10)\}$ as a lattice basis.

Theorem 2.1.2 now follows as a corollary:

Corollary 2.1.9 *Every subgroup $G \subseteq \mathbb{Z}^n$ is a lattice.*

Proof. Choose an $\mathbb{R}$ vector basis of $B = \{b_1,\dots,b_r\} \subseteq G$ for $\text{span}_{\mathbb{R}}(G)$. Every $g \in G$ can now be written as an $\mathbb{R}$-linear combination of vectors in B. Separating the fractional parts from the integral parts of the coefficients g can be written as a sum of an element in $\text{span}_{\mathbb{Z}}(B)$ and an element in $\{\sum_i c_i b_i : c_i \in [0,1)\} \cap G$. The last set is finite because $\{\sum_i c_i b_i : c_i \in [0,1)\}$ is bounded on all coordinates. Hence G is a subgroup of $\mathbb{Z}^n$ generated by the finite set $(\{\sum_i c_i b_i : c_i \in [0,1)\} \cap G) \cup B$. By Corollary 2.1.9 G is a lattice. □

Corollary 2.1.10 *For a matrix $A \in \mathbb{R}^{d \times n}$ the intersection* $\ker(A) \cap \mathbb{Z}^n$ *is a lattice (called the* lattice kernel *of A).*

Proof. We observe that $\ker(A) \cap \mathbb{Z}^n$ is subgroup of $\mathbb{Z}^n$ and apply Corollary 2.1.9. □

If the matrix has entries in $\mathbb{Z}$ we can even compute a basis for $\ker(A) \cap \mathbb{Z}^n$:

Algorithm 2.1.11 (Lattice kernel)
Input: *A matrix $A \in \mathbb{Z}^{d \times n}$.*
Output: *A lattice basis $\{b_1,\dots,b_m\}$ for* $\ker(A) \cap \mathbb{Z}^n$.

- *Let $B := A^T$.*
- *Compute B' and U as in Algorithm 2.1.4.*
- *Let $\{b_1,\dots,b_m\}$ be the last $n - \text{rank}(B)$ rows of U.*

Proof. We first observe that the rank of B equals the number of non-zero rows in B' because these rows are independent. The zero rows of B' are gotten by multiplying the last $n - \text{rank}(B)$ rows of U with B. This shows that these rows of U are indeed in $\ker(A) \cap \mathbb{Z}^n$. They are also independent. It remains to show that any element v of $\ker(A) \cap \mathbb{Z}^n$ can be written in this basis. This is indeed the case since any $v \in \mathbb{Z}^n$ can be written as a $\mathbb{Z}$-linear combination of all rows of U, since U is invertible over $\mathbb{Z}$. Furthermore if a row outside $\ker(A)$ was used, then v would have been outside as well, since the rows of U form a basis of $\mathbb{R}^n$. □

Example 2.1.12 The computation in Example 2.1.5 shows that $\{(2,-3,-3,0),$ $(9,-9,-11,-2)\}$ is a lattice basis of the lattice kernel of $\begin{bmatrix} 6 & 4 & 0 & 9 \\ 6 & -6 & 10 & -1 \end{bmatrix}$.

Definition 2.1.13 A lattice $L \subseteq \mathbb{Z}^n$ is called *saturated* if $L = \mathbb{Z}^n \cap \operatorname{span}_{\mathbb{R}}(L)$.

Lemma 2.1.14 *Any lattice basis of a saturated lattice $L \subseteq \mathbb{Z}^n$ can be completed to a basis of $\mathbb{Z}^n$. In particular the quotient group $\mathbb{Z}^n/L$ is a lattice.*

Proof. We first compute a vector space basis for $\operatorname{span}_{\mathbb{R}}(L)^{\perp}$ of integer vectors either by using Algorithm 2.1.11 or standard Gauss elimination scaling the result until it becomes integral. We write the results as rows in a matrix A. We then use Algorithm 2.1.11 to compute a lattice basis of $\ker(A) \cap \mathbb{Z}^n = \mathbb{Z}^n \cap \operatorname{span}_{\mathbb{R}}(L) = L$. We did this by picking rows of U. We observe that the remaining rows of U complete the obtained basis to a lattice basis of $\mathbb{Z}^n$. The same set of vectors will complete the original lattice basis to a lattice basis of $\mathbb{Z}^n$. Their cosets are a lattice basis for the quotient group $\mathbb{Z}^n/L$. □

Example 2.1.15 Let's complete $\{(2,-3,-3,0),(1,3,1,-2)\}$ (which generate a saturated lattice L) to a basis of $\mathbb{Z}^4$. We compute the generators $\{(2,4/3,0,3), (1,-1,5/3,-1/6)\}$ for the orthogonal complement $L^{\perp}$. After clearing denominators we get $\{(6,4,0,9),(6,-6,10,-1)\}$. The computation in Example 2.1.5 shows that the last two rows of the computed 4×4 matrix is a lattice basis of $\operatorname{span}_{\mathbb{Z}}\{(2,-3,-3,0),(1,3,1,-2)\}$. The remaining two rows complete this to a basis of $\mathbb{Z}^4$. Hence $\{(2,-3,-3,0),(1,3,1,-2),(-4,4,0,1),(0,0,1,0)\}$ is also a lattice basis of $\mathbb{Z}^4$.

2.2 Convexity

Definition 2.2.1 A set $X \subseteq \mathbb{R}^n$ is called *convex* if for every choice of $x, y \in X$ the line segment $\{tx + (1-t)y : t \in [0,1]\}$ between x and y is contained in X.

We note that an intersection of convex sets is convex.

Definition 2.2.2 Let $X \subseteq \mathbb{R}^n$. The *convex hull* of X is defined as the intersection of all convex sets containing X. That is, it is the smallest convex set containing X. We denote it by $\operatorname{conv}(X)$.

The following *separation* theorem is intuitively "obvious" but not so easy to prove.

Theorem 2.2.3 *Let A, B be convex subsets of $\mathbb{R}^n$. If $A \cap B = \emptyset$ then there exists a hyperplane H dividing $\mathbb{R}^n$ into two pieces H_- and H_+ such that $A \subseteq H_+ \cup H$ and $B \subseteq H_- \cup H$.*

We shall prove the theorem in the special case of Proposition 2.2.5 below. For the proof we need the two dimensional case as stated in Lemma 2.2.4.

Lemma 2.2.4 *Let $X \subseteq \mathbb{R}^2$ be a convex subset not containing any points on $\{(x,0) : x < 0\}$. Then there exists $N \in \mathbb{R}^2 \setminus \{0\}$ such that $\forall x \in X : N \cdot x \leq 0$.*

Proof. Without loss of generality we may assume that X is invariant under scaling by a positive number. That is $X = \{sx : x \in X, s \in \mathbb{R}_{>0}\}$. Identifying $\mathbb{R}^2$ with the complex plane in the standard way, the principal argument function $\text{Arg} : X \setminus \{0\} \to (-\pi, \pi)$ is a well-defined continuous function on $X \setminus \{0\}$. Let $r = \inf(\text{Arg}(X \setminus \{0\}))$ and $R = \sup(\text{Arg}(X \setminus \{0\}))$. If $R > \pi + r$ then the four values $(R + r)/2 \pm \pi/2 \pm \varepsilon$ are all in the interval $(r, R) \subseteq \text{Arg}(X \setminus \{0\})$ for $\varepsilon > 0$ sufficiently small. Hence the values are attained by four points on the unit circle. The convex hull of these four points is a rectangle contained in X with 0 in its interior. This contradict $\{(x, 0) : x < 0\} \cap X = \emptyset$. Hence $R \leq \pi + r$. We now choose $N = (-\cos(\frac{r+R}{2}), -\sin(\frac{r+R}{2}))$. It is now an easy trigonometric exercise to check the statement $\forall x \in X : N \cdot x \leq 0$. A drawing makes the situation clear. □

Proposition 2.2.5 *Let $X \subseteq \mathbb{R}^n$ be a convex set and let $v \in \mathbb{R}^n \setminus \{0\}$ be a generator of the (relatively open) half-line $h = \{tv | t \in \mathbb{R}_{>0}\}$. If $X \cap h = \emptyset$ then there exists an $N \in \mathbb{R}^n \setminus \{0\}$ such that $\forall x \in X : N \cdot x \leq 0$.*

Proof. For $n = 0$ the preassumptions cannot be fulfilled, so the theorem is trivially true. For $n = 1$ we may choose $N = v$. For $n = 2$ we may, after a linear change of coordinates, assume that $h = -\mathbb{R}_{<0} \times \{0\}$ and apply Lemma 2.2.4.

For $n \geq 3$ the proof goes by induction. We choose a two-dimensional plane H containing h. Now $H \cap X$ is convex in H and does not intersect $h \subseteq H$, so we may apply the proposition to $H \cap X$ with $n = 2$ and get an $N' \in H$ with $N' \cdot x \leq 0$ for all $x \in H \cap X$. We let l be the line in H perpendicular to N' and observe that the two-dimensional set $C := \{x \in H : N' \cdot x > 0\}$ does not intersect X. We now consider the image $\pi(X)$ of X under the orthogonal projection $\pi : \mathbb{R}^n \to l^\perp$. The image $\pi(X)$ is convex. Moreover, since $C \cap X = \emptyset$, C projects to a half line in $l^\perp$ not intersecting $\pi(X)$. We apply the induction hypothesis and get an $N \in l^\perp \setminus \{0\}$ with the property $N \cdot x \leq 0$ for every $x \in \pi(X)$. Since N is perpendicular to l this also holds for every $x \in X$. □

We will use Proposition 2.2.5 in the next section. Except from the next section most of our convex sets will be polyhedra. For polyhedra we can get away with many arguments without using analysis. We will see this in Section 3.

2.3 Robbiano's characterization of term orders

Let $A \in \mathbb{R}^{d \times n}$ be a matrix with $\ker(A) \cap \mathbb{Z}^n = \{0\}$ and the first non-zero entry of every column being positive. We define the *matrix term ordering* $\prec_A$ by $x^u \prec_A x^v \Leftrightarrow Au <_{\text{lex}} Av$, where $<_{\text{lex}}$ is the lexicographic ordering on $\mathbb{R}^d$.

That is, to compare x^u and x^v we first compare with respect to the first row of A. If we have a tie we continue to the second row and so on. It is left to the reader to show that this is in fact a term order.

Lemma 2.3.1 *A matrix term ordering is a term ordering.*

We notice that the two conditions on A (on lattice kernel and positivity of first entry of any column) are necessary to present a term ordering since term orders are antisymmetric and have $1 \preceq x^u$.

Example 2.3.2 The lexicographic term ordering $\prec_{\text{lex}}$ is defined by the identity matrix I. That is, $\prec_{\text{lex}}=\prec_I$.

Example 2.3.3 The graded reverse lexicographic term ordering on $k[x_1,\dots,x_4]$ is represented by the matrix:

$$\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 0 & 0 & -1\\ 0 & 0 & -1 & 0\\ 0 & -1 & 0 & 0\end{bmatrix}$$

Example 2.3.4 Let $A \in \mathbb{R}^{d\times n}$ represent a term ordering $\prec_A$ and let $\omega \in \mathbb{R}^n_{\geq 0}$ be a vector. Let $A' \in \mathbb{R}^{(d+1)\times n}$ be the matrix gotten by prepending ω as a first row to A. Then, using the definition from Sheet 1, Exercise 7, we have $(\prec_A)_\omega = \prec_{A'}$.

Theorem 2.3.5 below says that every term ordering can be gotten as a matrix term ordering. Robbiano was not the first person to prove this theorem but it has become known as Robbiano's characterization of term orders in the computational algebra community. Robbiano's rediscovery emphasizes the importance of convex geometry and is the justification for most of the material in this course.

Theorem 2.3.5 *Every term order $\prec$ on $k[x_1,\dots,x_n]$ is a matrix term order. That is, there exists a matrix $A \in \mathbb{R}^{d\times n}$ such that $\prec_A=\prec$.*

By a *Laurent monomial* we mean a monomial of the form x^v where $v \in \mathbb{Z}^n$ with possibly negative exponents. For the purpose of proving the theorem we need to extend our notion of orderings and identify Laurent monomials with the points of a lattice.

Definition 2.3.6 An *admissible* ordering $\preceq$ on a group L is a total ordering on L respecting addition. An *admissible* ordering $\preceq$ on the Laurent monomials (with fixed n) is a total ordering respecting multiplication.

Lemma 2.3.7 *Let $a, b \in \mathbb{Z}^n$ and $\prec$ be an admissible ordering. Suppose $x^a \prec 1$ and $x^b \prec 1$. Then $x^{a+b} \prec 1$.*

Proof. We have $x^{a+b} = x^a x^b \prec x^a \cdot 1 = x^a \prec 1$. □

We notice that every term ordering extends uniquely to an admissible ordering on Laurent monomials by $x^{a-b} \prec x^{c-d} \Leftrightarrow x^{a+d} \prec x^{c+b}$ for $a, b, c, d \in \mathbb{N}^n$.

Proof of Theorem 2.3.5 We prove the theorem under the more general assumption that $\prec$ is an admissible ordering of Laurent monomials. The proof goes by induction. For $n = 1$ we choose either $A = [1]$ or $A = [-1]$ depending on whether $1 \prec x_1$ or not.

For $n \geq 2$ we consider the cone $X \subseteq \mathbb{R}^n$ of points spanned positively by the set $G = \{u \in \mathbb{Z}^n : x^u \prec 1\}$, meaning $X := \{\sum_{i=1}^m a_i g_i : m \in$

$\mathbb{N} \setminus \{0\}, a_i \in \mathbb{R}_{>0}, g_i \in G\}$. This set is clearly convex. We now argue that $0 \notin X$. Suppose we could find a finite subset of $\{g_1, \ldots, g_m\} \subseteq G$ and coefficients $c_1, \cdots, c_m \in \mathbb{R}_{>0}$ with $\sum_i g_i c_i = 0$ and $m > 0$. Let H be the matrix with columns $g_1, \ldots, g_m$. Using linear algebra it is possible to find a parametrisation $\varphi : x \mapsto Cx$ of Nullspace(H), with C having entries from $\mathbb{Q}$. Since $(c_1, \cdots, c_m) \in$ Image(φ), perturbing the parameters slightly we can get a rational point $(c'_1, \ldots, c'_m) \in \mathbb{Q}^m_{>0} \cap$ Nullspace(H). Scaling we get a point $(c''_1, \ldots, c''_m) \in \mathbb{N}^m_{>0} \cap$ Nullspace(H) and Applying Lemma 2.3.7 repeatedly, we get for every i that $x^{c''_i g_i} \prec 1$. Applying it further we get $x^{\sum_i g_i c''_i} \prec 1$. But this is a contradiction since $\sum_i g_i c''_i = 0$. We conclude that $0 \notin X$.

Pick a line passing through the origin. Since X is convex with $0 \notin X$ one of its two half lines does not intersect X. We apply Proposition 2.2.5 and get a vector $N \neq 0$ such that $\forall x \in X : N \cdot x \leq 0$. We use N as the first row of the matrix A. To get the remaining rows we notice that $N^{\perp} \cap \mathbb{Z}^n$ is a lattice and choose a basis for it $\{b_1, \ldots, b_r\}$ with $(r < n)$ (Corollary 2.1.10). Using Lemma 2.1.14 we extend the basis to a basis for $\mathbb{Z}^n$ and write it in the columns of a matrix $B \in \mathbb{Z}^{n \times n}$. The order $\prec$ induces an admissible ordering on $\mathbb{Z}^r$. By the induction hypothesis there exists some matrix A' representing the restricted order (in the basis $\{b_1, \ldots, b_r\}$). Let B' be the first r rows of B^{-1}. To compare vectors in $N^{\perp}$ we multiply them with $A'B'$ and compare the result lexicographically. We let $A'B'$ be the remaining rows of A. To prove that $\prec_A$ equals $\prec$ let $u \in \mathbb{Z}^n$. It suffices to show $x^u \prec 1 \Rightarrow x^u \prec_A 1$. This is true since $x^u \prec 1 \Rightarrow u \in G \subseteq X \Rightarrow N \cdot u \leq 0$. If $N \cdot u < 0$ then $x^u \prec_A 1$. If $N \cdot u = 0$ then this also holds since by induction $\prec_{A'}$ equals $\prec$ in the sublattice $N^{\perp} \cap \mathbb{Z}^n$. □

3 Polyhedral geometry

3.1 Polyhedra

Definition 3.1.1 Let $u \in \mathbb{R}^n$ and $r \in \mathbb{R}$. We define the following sets:

- $H_{u,r} := \{x \in \mathbb{R}^n : u \cdot x = r\}$.
- $H^{\leq}_{u,r} := \{x \in \mathbb{R}^n : u \cdot x \leq r\}$.
- $H^{<}_{u,r} := \{x \in \mathbb{R}^n : u \cdot x < r\}$.

We also define $H^{\geq}_{u,r} := H^{\leq}_{-u,-r}$ and $H^{>}_{u,r} = H^{<}_{-u,-r}$. If $u \neq 0$ then $H_{u,r}$ is called a *hyperplane*, $H^{\leq}_{u,r}$ a closed half space and $H^{<}_{u,r}$ an open halfspace.

If we equip $\mathbb{R}^n$ with its usual Euclidean topology, then an open halfspace is open since every point in the halfspace has an ε-neighbourhood contained in the half space. Furthermore, a closed halfspace is the complement of an open halfspace and is therefore closed.

Definition 3.1.2 A subset $P \subseteq \mathbb{R}^n$ is called a polyhedron if it is an intersection of finitely many closed half spaces. For a matrix $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ we use the notation: $P_{A,b} := \bigcap_{i=1}^m H^{\leq}_{A_{i\cdot},b_i} = \{x \in \mathbb{R}^n : Ax \leq b\}$. If b is the zero vector, then $P_{A,b}$ is called a polyhedral cone.

In particular, if we intersect one closed half space we get that every closed halfspace is a polyhedron. If we intersect zero closed halfspaces we get that $\mathbb{R}^n$ is a polyhedron (by convention). If we intersect halfspaces with empty intersection we get that $\emptyset$ is polyhedron. Finally, any (affine) linear subspace of $\mathbb{R}^n$ is a polyhedron. Notice that A is allowed having rows being zero.

We observe that polyhedra are convex sets since they are intersections of convex sets.

Algorithm 3.1.3 (Fourier-Motzkin)
Input: *A matrix* $A \in \mathbb{R}^{m \times n}$ *and* $b \in \mathbb{R}^m$ *defining a polyhedron* $P_{A,b}$.
Output: *A matrix* $A' \in \mathbb{R}^{m' \times (n-1)}$ *and* $b' \in \mathbb{R}^{m'}$ *such that* $\pi(P_{A,b}) = P_{A',b'}$ *where* $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ *is the projection* $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{n-1})$.

```
• Rescale, without changing $P_{A,b}$, the rows of $A$ and entries of $b$ positively
  so that the $n$th entry of each row of $A$ is either $0, -1$ or $1$.
• Let $A'$ be a matrix without rows and $b'$ a vector with no entries.
• For $i = 1, \ldots, m$
    – If($A_{in} = 0$)
        ∗ Append the row $\pi(A_{i\cdot})$ to $A'$ and append the entry $b_i$ to $b'$.
    – else
        ∗ For $j = i + 1, \ldots, m$
```

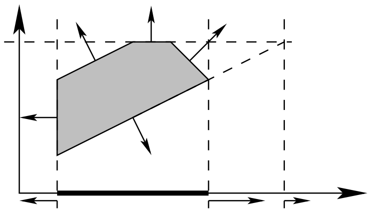

Figure 4: The polyhedra, and constructed inequalities of Example 3.1.4.

- *If($A_{in} + A_{jn} = 0$) append $\pi(A_{i\cdot} + A_{j\cdot})$ to A' and $b_i + b_j$ to b'.*

Proof. The algorithm clearly terminates. We only need to prove that the output satisfies the specifications.

We first argue that $\pi(P_{A,b}) \subseteq P_{A',b'}$. We must show that the projection of a point $x \in P_{A,b}$ satisfies the new constraints. This is true since the equation $(A_{i\cdot} + A_{j\cdot}) \cdot x \leq b_i + b_j$ is implied by the equations of $P_{A,b}$ and it does not involve x_n. Similarly for the other type of equation introduced to A' and b'.

Conversely, let $(x_1, \ldots, x_{n-1}) \in P_{A',b'}$. We wish to argue that we can find x_n such that $(x_1, \ldots, x_n) \in P_{A,b}$. After the first rewrite, there are three types of equations described by A and b. The first type does not involve x_n, so they are satisfied for every choice of x_n. The second type says that x_n should be $\geq$ some values $M \subseteq \mathbb{R}$ depending on $(x_1, \ldots, x_{n-1})$. The last that x_n should be $\leq$ some values $M' \subseteq \mathbb{R}$. Having a valid choice for x_n is exactly possible when $\forall (a, b) \in M \times M' : a \leq b$. This translates into the $|M| \cdot |M'|$ inequality conditions on $x_1, \ldots, x_{n-1}$ that we appended to A' and b'. Therefore x_n can be chosen. □

Example 3.1.4 Using Algorithm 3.1.3 we project the polyhedron given by:

$$\begin{bmatrix} -1 & 0 \\ -1/2 & 1 \\ 0 & 1 \\ 1 & 1 \\ 1/2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} -1 \\ 5/2 \\ 4 \\ 8 \\ -1/2 \end{bmatrix}$$

and get the following system describing the projection:

$$\begin{bmatrix} -1 \\ 0 \\ 1/2 \\ 3/2 \end{bmatrix} \begin{bmatrix} x_1 \end{bmatrix} \leq \begin{bmatrix} -1 \\ 2 \\ 7/2 \\ 15/2 \end{bmatrix}.$$

The first inequality was transferred directly, while the other three came from combining the 5th inequality with inequality 2-4. See Figure 4.

Corollary 3.1.5 *Let $P_{A,b} \subseteq \mathbb{R}^n$ be a polyhedron. The projection of $P_{A,b}$ to any subset of its coordinates is a polyhedron. If $P_{A,b}$ is a polyhedral cone then so is its projection.*

Proof. The corollary follows immediately from Algorithm 3.1.3. If $P_{A,b}$ is a cone we may without loss of generality assume that $b = 0$. We observe by inspecting the algorithm that the b' produced is also 0 and the projection is a cone. □

Theorem 3.1.6 *The image of a polyhedron $P_{A,b} \subseteq \mathbb{R}^n$ under an affine transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is a polyhedron. If $P_{A,b}$ is a polyhedral cone and φ a linear transformation then $\varphi(P_{A,b})$ is a polyhedral cone.*

Proof. Suppose φ is given by $x \mapsto Cx + d$ for some matrix $C \in \mathbb{R}^{m \times n}$ and $d \in \mathbb{R}^m$. Consider the graph of the function φ over $P_{A,b}$:

$$\{(x, y) \in \mathbb{R}^{n+m} : \varphi(x) = y \wedge x \in P_{A,b}\} = \{(x, y) \in \mathbb{R}^{n+m} : Cx-y = -d \wedge Ax \leq b\}.$$

This is clearly a polyhedron. Its projection onto the last m coordinates is the image $\varphi(P_{A,b})$. By Corollary 3.1.5 this projection is a polyhedron.

If $P_{A,b}$ is a polyhedral cone and φ a linear transformation we may without loss of generality assume that $b = 0$. Since also $d = 0$, the graph we are projecting is a polyhedral cone. Its projection is a polyhedral cone by Corollary 3.1.5. □

Definition 3.1.7 Let $V \subseteq \mathbb{R}^n$. We define the cone spanned by V as follows:

$$\text{cone}(V) = \{\sum_{i=1}^{d} c_i v_i : d \in \mathbb{N}, v_i \in V, c_i \in \mathbb{R}_{\geq 0}\}.$$

Definition 3.1.8 A bounded polyhedron P is called a *polytope*.

Lemma 3.1.9 *Let $X \subseteq \mathbb{R}^n$ be a finite set of points. Then*

$$\text{conv}(X) = \{\sum_{p \in X} c_p p : c_p \in \mathbb{R}_{\geq 0}, \sum_{p \in X} c_p = 1\}.$$

Proof. Left to the reader. □

Theorem 3.1.10 *The convex hull of a finite set $X \subseteq \mathbb{R}^m$ of points is a polytope. The cone spanned by a finite set $V \subseteq \mathbb{R}^m$ of vectors is a polyhedral cone.*

Proof. By Lemma 3.1.9 conv(X) is the image of the polyhedron $\{c \in \mathbb{R}^{|X|}_{\geq 0} : \sum_{i=1}^{|X|} c_i = 1\}$ under the linear map defined by the elements in X. By Theorem 3.1.6 this image is a polyhedron. To prove that it is bounded, observe that X is bounded and that a ball containing X will also contain conv(X).

Similarly, cone(V) is the image of the polyhedral cone $\mathbb{R}^{|V|}_{\geq 0}$ under a linear transformation and therefore a polyhedral cone by Theorem 3.1.6. □

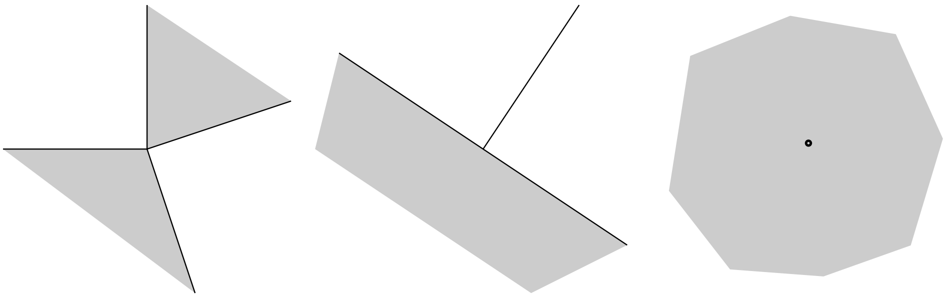

Figure 5: Three polyhedral cones and their duals, see Example 3.1.4.

Definition 3.1.11 Let P_1 and P_2 be polyhedra in $\mathbb{R}^n$. The *Minkowski sum* of P_1 and P_2 is defined as

$$P_1 + P_2 := \{p_1 + p_2 : p_1 \in P_1, p_2 \in P_2\}.$$

Proposition 3.1.12 *The Minkowski sum of two polyhedra P_1 and P_2 in $\mathbb{R}^n$ is a polyhedron. If P_1 and P_2 are polyhedral cones, then so is $P_1 + P_2$.*

Proof. The cartesian product $P_1 \times P_2$ in $\mathbb{R}^n \times \mathbb{R}^n$ is a polyhedron. The function $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ given by $(x, y) \mapsto x + y$ is linear. The image $\varphi(P_1 \times P_2)$ equals $P_1 + P_2$. This image is a polyhedron by Theorem 3.1.6. Furthermore, if P_1 and P_2 are polyhedral cones, then so is the image. □

3.2 Cone duality

The purpose of this subsection is to prove the converse of Theorem 3.1.10.

Definition 3.2.1 Let $C \subseteq \mathbb{R}^n$ be a polyhedral cone. We define its *dual cone* to be:

$$C^\vee := \{y \in \mathbb{R}^n : \forall x \in C : x \cdot y \leq 0\}.$$

Example 3.2.2 Three examples of a polyhedral cone $C \subseteq \mathbb{R}^3$ and its dual $C^\vee$ are shown in Figure 5. In the last example $C = \{0\}$ and $C^\vee = \mathbb{R}^2$

Proposition 3.2.3 *Let $C = P_{A,0}$ be a polyhedral cone for some matrix $A \in \mathbb{R}^{m\times n}$. Then*

$$C^\vee = \mathrm{cone}(A_{1\cdot}, \dots, A_{m\cdot}).$$

In particular the dual cone $C^\vee$ is a polyhedral cone.

Proof. Let $y \in \mathrm{cone}(A_{1\cdot}, \dots, A_{m\cdot})$ and $x \in C$. Since $A_{i\cdot} \cdot x \leq 0$ for all i then also $y \cdot x \leq 0$ since y is a non-negative linear combination of $A_{1\cdot}, \dots, A_{m\cdot}$. Since this holds for all $x \in C$ we conclude that $y \in C^\vee$.

We observe by Theorem 3.1.10 that the right hand side is a polyhedral cone $P_{B,0}$ for some matrix B. If $y \in \mathbb{R}^n$ is not in the right hand side, then one of the rows $B_{j\cdot}$ would have $B_{j\cdot} \cdot y > 0$ and $B_{j\cdot} \cdot A_{i\cdot} \leq 0$. The latter implies that $B_{j\cdot} \in C$. We conclude, since $B_{j\cdot} \cdot y > 0$, that $y \notin C^\vee$. □

Proposition 3.2.4 *Let $C \subseteq \mathbb{R}^n$ be a polyhedral cone. Then*

$$(C^\vee)^\vee = C.$$

Proof. The inclusion $\supseteq$ is clear since every vector in C indeed does have non-positive dot product with vectors of $C^\vee$. Conversely, if $C = P_{A,0}$ for some matrix $A \in \mathbb{R}^{m\times n}$ then by Proposition 3.2.3 $C^\vee = \text{cone}(A_{1\cdot}, \dots, A_{m\cdot})$. If $z \in (C^\vee)^\vee$ then $A_{i\cdot} \cdot z \leq 0$ for all i. This proves $z \in P_{A,0} = C$. □

Corollary 3.2.5 *Any polyhedral cone $C \subseteq \mathbb{R}^n$ is of the form $C = \text{cone}(v_1, \dots, v_m)$ for a finite list of vectors $v_1, v_2, \dots, v_m \in \mathbb{R}^n$.*

Proof. By Proposition 3.2.3 the dual cone $C^\vee$ is a polyhedron of the form $C^\vee = P_{A,0}$ for some matrix $A \in \mathbb{R}^{m\times n}$. By Proposition 3.2.4 $C = (C^\vee)^\vee$ and $(C^\vee)^\vee = \text{cone}(A_{1\cdot}, \dots, A_{m\cdot})$ by Proposition 3.2.3. □

Theorem 3.2.6 *Every polyhedron $P \subseteq \mathbb{R}^n$ has the form*

$$P = \text{conv}(u_1, \dots, u_r) + \text{cone}(v_1, \dots, v_s).$$

In particular, P is the Minkowski sum of a polytope and a polyhedral cone.

Proof. We start by forming a cone $C \subseteq \mathbb{R}^{n+1}$ with the property that $C \cap \mathbb{R}^n \times \{1\} = P \times \{1\}$. To be precise, if $P = P_{A,b}$ with $A \in \mathbb{R}^{m\times n}$ we let $C = P_{A',0}$ where A' is an $\mathbb{R}^{m\times(n+1)}$ with the first n columns equal to the columns of A, and the last column equal to $-b$. It is straight forward to check that $C \cap \mathbb{R}^n \times \{1\} = P \times \{1\}$. We add an additional constraint $C' = C \cap H^{\geq}_{e_{n+1},0}$, such that every point in C' has last coordinate non-negative.

By Corollary 3.2.5 we have $C' = \text{cone}(w_1, \dots, w_t)$ for some $w_i \in \mathbb{R}^{n+1}$. Without loss of generality we may assume that the last coordinate of each of these vectors is either 0 or 1. Let $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ be the projection onto the first n coordinates. We now construct the u_i and v_i vectors as follows. If w_i has last coordinate 1 then we construct a u vector $\pi(w_i)$. If the last coordinate was 0 we construct the v vector $\pi(w_i)$. In total we have constructed $t = r + s$ vectors. We now prove the equality

$$P = \text{conv}(u_1, \dots, u_r) + \text{cone}(v_1, \dots, v_s).$$

To prove the inclusion $\supseteq$ let p be obtained as a valid non-negative combination of $u_1, \dots, u_r$ and $v_1, \dots, v_s$. Using the same coefficients for a combination of $w_1, \dots, w_t$ we get $(p_1, \dots, p_n, 1)$ in $C \cap \mathbb{R}^n \times \{1\} = P \times \{1\}$. This proves $p \in P$. On the other hand if $p \in P$ then $(p_1, \dots, p_n, 1) \in C$ is a non-negative linear combination of $w_1, \dots, w_t$. We use the same coefficients on $u_1, \dots, u_r$ and $v_1, \dots, v_s$ and get p as a non-negative combination. Furthermore, the coefficients of $u_1, \dots, u_s$ sum to 1. This proves $p \in \text{conv}(u_1, \dots, u_r) + \text{cone}(v_1, \dots, v_s)$. Finally, Theorem 3.1.10 says that we have written P as the sum of a polytope and a polyhedral cone. □

Remark 3.2.7 While the decomposition of Theorem 3.2.6 is not unique in general, the cone part is. Namely, it is gotten by intersection C in the proof with the hyperplane $H_{e_{n+1},0}$ and projecting away the last coordinate. Equivalently, with the notation of the proof it is $P_{A,0}$. We call this cone the *recession cone* $R(P_{A,b})$ of $P_{A,b}$.

Corollary 3.2.8 *Every polytope is the convex hull of a finite set of points.*

Proof. We use Theorem 3.2.6 and observe that if the cone constructed this way is not $\{0\}$ then the Minkowski sum would be unbounded, which is a contradiction. □

3.3 Dimension and faces

Definition 3.3.1 Let $P_{A,b} \subseteq \mathbb{R}^n$ be a non-empty polyhedron. We define its *lineality space* $L(P_{A,b})$ to be $\ker(A)$.

That the lineality space is well-defined follows from the following lemma.

Lemma 3.3.2 *A non-empty polyhedron $P_{A,b} \subseteq \mathbb{R}^n$ is invariant under translation by exactly the vectors in $L(P_{A,b})$.*

Proof. If $y \in L(P_{A,b})$ and $x \in P_{A,b}$ then $Ax \leq b$ and $Ay = 0$, implying $A(x+y) \leq b$. We conclude that $x + y \in P_{A,b}$ and that $P_{A,b}$ is invariant under translation by y, meaning $P_{A,b}+y = P_{A,b}$. On the other hand suppose $P_{A,b}+y = P_{A,b}$ for some $y \in \mathbb{R}^n$ and $x \in P_{A,b}$. Then for all $s \in \mathbb{R}$ we have $A(x + sy) \leq b$. If Ay was non-zero, we could make the left hand side arbitrarily large. We conclude that $y \in \ker(A)$. □

Example 3.3.3 The cone $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R} \subseteq \mathbb{R}^3$ has the one dimensional lineality space $\{0\} \times \{0\} \times \mathbb{R}$.

Definition 3.3.4 We say that a polyhedral cone C is *pointed* if $\dim(L(C)) = 0$.

Definition 3.3.5 The dimension of a non-empty polyhedron $P \subseteq \mathbb{R}^n$ is the dimension of the smallest affine subspace of $\mathbb{R}^n$ containing it.

Lemma 3.3.6 (Farkas' Lemma) *Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ then $P_{A,b} = \emptyset$ if and only if there exists a row vector $y \in \mathbb{R}^m_{\geq 0}$ such that $yA = 0$ and $yb = -1$.*

Proof. The "if" direction is clear because the non-negative y tells us how to combine the equations $Ax \leq b$ to the impossible equation $0 \leq -1$. Conversely, if $P_{A,b} = \emptyset$ we can, as in the proof of Theorem 3.2.6 consider the matrix $A' \in \mathbb{R}^{m\times(n+1)}$ whose first n columns are the columns from A and whose last column is $-b$. By the argument in the proof of Theorem 3.2.6 $P_{A',0}$ cannot contain any point with last coordinate positive (because $P_{A,b} = \emptyset$). Hence $e_{n+1} \in P^\vee_{A',0} = \text{cone}(A'_{1\cdot}, \dots, A'_{m\cdot})$. Hence we can find $y \in \mathbb{R}^m_{\geq 0}$ with $yA = 0$ and $y(-b) = 1$. □

Lemma 3.3.7 *Let $P \subseteq \mathbb{R}^n$ be a non-empty polyhedron and $\omega \in \mathbb{R}^n$. Then $\max_{y\in P}(\omega \cdot y)$ is attained if and only if ω is bounded from above on P.*

Proof. Define the projection $\pi : \mathbb{R}^n \to \mathbb{R}$ by $x \mapsto \omega \cdot x$. By Corollary 3.1.5 $\pi(P)$ is a non-empty, (from above) bounded polyhedron in $\mathbb{R}$ and therefore a closed interval with an upper end point $y \in \mathbb{R}$. Hence ω attains its maximum in the preimage $P \cap \pi^{-1}(y)$. □

Lemma 3.3.8 *Let $P \subseteq \mathbb{R}^n$ be a non-empty polyhedron and $\omega \in \mathbb{R}^n$. Then $\max_{y\in P}(\omega \cdot y)$ is attained if and only if $\omega \in R(P)^\vee$ (where $R(P)$ is the recession cone of P).*

Proof. Let $P = P_{A,b}$ for some $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Suppose $\omega \in R(P)^\vee$. By Proposition 3.2.3 there exists a row vector $y \in \mathbb{R}^m_{\geq 0}$ such that $\omega^T = yA$. If $x \in P$ then $Ax \leq b$, which implies by non-negativity of entries of y that $\omega \cdot x = yAx \leq yb$. The right hand side is independent of x which means that the linear form ω is bounded from above on P.

Conversely, suppose ω is bounded over P. Then there exists $h \in \mathbb{R}$ such that $H^{\geq}_{\omega,h} \cap P = H^{\leq}_{-\omega,-h} \cap P_{A,b} = \emptyset$. By Farkas' Lemma there exists y such that $y\begin{bmatrix} -\omega^T & -h \\ A & b \end{bmatrix} = \begin{bmatrix} 0 \cdots 0 & -1 \end{bmatrix}$. The first coordinate of y cannot be zero because $P \neq \emptyset$. We conclude that $\omega \in \mathrm{cone}(A_{1\cdot}, \dots, A_{m\cdot}) = (P_{A,0})^\vee = R(P)^\vee$. □

Definition 3.3.9 Let $P \subseteq \mathbb{R}^n$ be a polyhedron and $\omega \in \mathbb{R}^n$. If $\max_{y\in P}(\omega \cdot y)$ is attained, the set

$$\mathrm{face}_\omega(P) := \{x \in P : \omega \cdot x = \max_{y\in P}(\omega \cdot y)\}$$

is called a *face* of P. The hyperplane $H_{\omega,\max_{y\in P}(\omega\cdot y)}$ is called a *supporting hyperplane* for P.

We observe that if $\max_{y\in P}(\omega \cdot y)$ is attained, then $P \subseteq H^{\leq}_{\omega,\max_{y\in P}(\omega\cdot y)}$ and $\mathrm{face}_\omega(P) = P \cap H_{\omega,\max_{y\in P}(\omega\cdot y)}$. Consequently, $\mathrm{face}_\omega(P)$ is a polyhedron.

Remark 3.3.10 Most people also call the empty set $\emptyset$ a face, and give it the name "the empty face". We will try not to do so in these notes.

Definition 3.3.11 We define the following terms:

- A *vertex* of a polyhedron $P \subseteq \mathbb{R}^n$ is a face of P of dimension 0.
- A *facet* of P is a face of dimension $\dim(P) - 1$.
- A *ray* of a pointed polyhedral cone $C \subseteq \mathbb{R}^n$ is a face of C of dimension 1.

Proposition 3.3.12 *Let $P \subseteq \mathbb{R}^n$ be a polyhedron. Let $\omega, \omega' \in R(P)\vee$ such that the faces $A := \mathrm{face}_\omega(P)$ and $B := \mathrm{face}_{\omega'}(P)$ are well-defined. Then $\omega' \in R(A)^\vee$. If $A \cap B \neq \emptyset$ then $A \cap B = \mathrm{face}_{\omega'}(A)$.*

Proof. By Lemma 3.3.8, the maximum of the linear form ω' is attained over P and therefore also over A. Consequently, $\omega' \in R(A)^\vee$. We start by observing that since $A \cap B \neq \emptyset$ the hyperplane H with normal ω' and $B = P \cap H$ is a supporting hyperplane for A (because the maximal value of ω' over P is the same as over A). It now follows that $\mathrm{face}_{\omega'}(A) = A \cap H = (A \cap P) \cap H = A \cap (P \cap H) = A \cap B$. □

Proposition 3.3.13 *Let $u_1, \dots, u_r, v_1, \dots, v_s \in \mathbb{R}^n$ with $r \geq 1$, let*

$$P = \mathrm{conv}(u_1, \dots, u_r) + \mathrm{cone}(v_1, \dots, v_s)$$

be a polyhedron and let $\omega \in R(P)^\vee$ (with $R(P)$ being the recession cone). Then

$$\mathrm{face}_\omega(P) = \mathrm{conv}_{\omega \cdot u_i = U}(u_i) + \mathrm{cone}_{\omega \cdot v_i = 0}(v_i)$$

where $U = \max_i(\omega \cdot u_i)$. Furthermore, if we have some other $\omega' \in R(P)^\vee$ with $\mathrm{face}_{\omega'}(P) = \mathrm{face}_\omega(P)$ then $\{i : \omega \cdot u_i = U\} = \{i : \omega' \cdot u_i = \max_j \omega' \cdot u_j\}$ and $\{i : \omega \cdot v_i = 0\} = \{i : \omega' \cdot v_i = 0\}$.

Proof. Let $p \in \mathrm{face}_\omega(P)$. Then for some non-negative choice of coefficients $p = \sum_i a_i u_i + \sum_i b_i v_i$ with $\sum_i a_i = 1$. We wish to prove that the coefficients are zero for the vectors not mentioned in the right hand side. If for some i $a_i > 0$, but with $\omega \cdot u_i \neq U$ then we could decrease a_i and increase another coefficient to reach point $p' \in P$ with bigger dot product with ω. This would be a contradiction. Similarly, first notice that for all i:$\omega \cdot v_i \leq 0$ because ω is bounded from above on P. If $b_i > 0$ with $\omega \cdot v_i < 0$, then we could again increase the dot product with ω by choosing $b_i = 0$ instead. That would be a contradiction. Hence

$$\mathrm{face}_\omega(P) \subseteq \mathrm{conv}_{\omega \cdot u_i = U}(u_i) + \mathrm{cone}_{\omega \cdot v_i = 0}(v_i).$$

Conversely, let now $p \in \mathrm{conv}_{\omega \cdot u_i = U}(u_i) + \mathrm{cone}_{\omega \cdot v_i = 0}(v_i)$ with according choice of coefficients $p = \sum_i a_i u_i + \sum_i b_i v_i$ and $\sum_i a_i = 1$. Then $\omega \cdot p = \sum_i a_i \omega \cdot u_i + \sum_i b_i \omega \cdot v_i = \sum_i a_i U + \sum_i b_i 0 = 1U + 0 = U = \max_i(\omega \cdot u_i) = \max_{x \in \mathrm{conv}(u_1, \dots, u_r)}(\omega \cdot x)$. This is the maximum of ω over $P = \mathrm{conv}(u_1, \dots, u_r) + \mathrm{cone}(v_1, \dots, v_s)$ because for all i:$\omega \cdot v_i \leq 0$.

For the second claim, suppose $\mathrm{face}_{\omega'}(P) = \mathrm{face}_\omega(P)$. Because $\omega \cdot v_j \leq 0$ for all j we have $u_i \in \mathrm{face}_\omega(P)$ iff $\omega \cdot u_i = \max_j(\omega \cdot u_j)$. Similarly, $u_i \in \mathrm{face}_{\omega'}(P)$ iff $\omega \cdot u_i = \max_j(\omega' \cdot u_j)$. This prove the first equality. Let $p \in \mathrm{face}_{\omega'}(P)$ and suppose that some v_i is perpendicular to ω but not ω'. Then $\omega' \cdot v_i < 0$, preventing $p + tv_i \in \mathrm{face}_\omega(P)$ from being in $\mathrm{face}_{\omega'}(P)$ for t big – a contradiction. Similarly for $v_i \cdot \omega' = 0 \neq v_i \cdot \omega$. This proves the last equality. □

Corollary 3.3.14 *A polyhedron has only finitely many faces.*

Proof. By Theorem 3.2.6 every polyhedron has the form of Proposition 3.3.13. In Proposition 3.3.13 there is only a finite number of subsets of $\{u_1, \dots, u_r\}$ and $\{v_1, \dots, v_s\}$ leading to only finitely many possible faces. □

Lemma 3.3.15 *Let $a_1, \dots, a_r, b_1, \dots, b_r \in \mathbb{R}$ and $A = \{i : a_i = \max_j(a_j)\}$ and $B = \{i \in A : b_i = \max_{j \in A}(b_j)\}$. There exists $\varepsilon \in \mathbb{R}_{>0}$ such that $B = \{i : a_i + \varepsilon b_i = \max_j(a_j + \varepsilon b_j)\}$.*

Proof. Let $\alpha = \max_j(a_j)$ and $\beta = \max_{j \in A}(b_j)$. We choose $\varepsilon > 0$ such that for all i we have $(\alpha - a_i) > \varepsilon(b_i - \beta)$ when ever $(a_i, b_i) \neq (\alpha, \beta)$. This is possible since either $\alpha - a_i > 0$, or when not we have $\alpha - a_i = 0$ and $b_i - \beta < 0$. We now observe that the right hand side $\{i : a_i + \varepsilon b_i = \max_j(a_j + \varepsilon b_j)\}$ is the set of indices such that $(1, \varepsilon)$ is maximized over $P := \{(a_1, b_1), \ldots, (a_r, b_r)\}$. The indices of B is exactly those i for which $a_i = \alpha$ and $b_i = \beta$. Hence it suffices to show that (α, β) is the unique optimum of $(1, \varepsilon)$ over P. First of all $(\alpha, \beta) \in P$. Let i be given. We would like to prove that if $(a_i, b_i) \neq (\alpha, \beta)$ we have $(1, \varepsilon) \cdot (\alpha, \beta) < (1, \varepsilon) \cdot (a_i, b_i)$. But this follows from the choice of ε. □

The following corollary says that the face of a face is a face.

Corollary 3.3.16 *Let $P \subseteq \mathbb{R}^n$ be a polyhedron. Let $\omega \in R(P)^\vee$. Let $\omega' \in R(\text{face}_\omega(P))^\vee$. Then $F := \text{face}_{\omega'}(\text{face}_\omega(P))$ is a face of P.*

Proof. Using Theorem 3.2.6 we know that P has the form

$$P = \text{conv}(u_1, \ldots, u_r) + \text{cone}(v_1, \ldots, v_s)$$

and by Proposition 3.3.13 we have

$$\text{face}_\omega(P) = \text{conv}_{\omega \cdot u_i = U}(u_i) + \text{cone}_{\omega \cdot v_i = 0}(v_i) \text{ and}$$

$$\text{face}_{\omega'}(\text{face}_\omega(P)) = \text{conv}_{\omega \cdot u_i = U \wedge \omega' \cdot u_i = U'}(u_i) + \text{cone}_{\omega \cdot v_i = 0 \wedge \omega' \cdot v_i = 0}(v_i)$$

where $U = \max_i(\omega \cdot u_i)$ and $U' = \max_{i : w \cdot u_i = U}(\omega' \cdot u_i)$.

We wish to choose $\varepsilon \in \mathbb{R}_{>0}$ such that $\omega_\varepsilon := \omega + \varepsilon \omega' \in R(\text{face}_\omega(P))^\vee$ and $F = \text{face}_{\omega_\varepsilon}(P)$. That $\omega \in R(P)^\vee$ simply means that for all i:$\omega \cdot v_i \leq 0$ and that $\omega' \in R(\text{face}_\omega(P))^\vee$ means that whenever $\omega \cdot v_i = 0$ then $\omega' \cdot v_i \leq 0$. We conclude that for $\varepsilon > 0$ sufficiently small we have $\omega_\varepsilon \cdot v_i \leq 0$ for all i. Hence $\omega_\varepsilon \in R(\text{face}_\omega(P))^\vee$.

It suffices to prove that for $\varepsilon > 0$ sufficiently small

$$\{i : \omega \cdot u_i = U \wedge \omega' \cdot u_i = U'\} = \{i : \omega_\varepsilon \cdot u_i = \max_i(\omega_\varepsilon \cdot u_i)\}$$

and

$$\{i : \omega \cdot v_i = 0 \wedge \omega' \cdot v_i = 0\} = \{i : \omega_\varepsilon \cdot v_i = 0\}.$$

The first equality follows from Lemma 3.3.15 for small $\varepsilon > 0$. To prove the second, first observe that we have the inclusion "$\subseteq$". To prove "$\supseteq$", we choose $\varepsilon > 0$ such that $\varepsilon < -\frac{\omega \cdot v_i}{\omega' \cdot v_i}$ whenever $\omega' \cdot v_i > 0$ and $\omega \cdot v_i < 0$. Suppose that $\omega_\varepsilon \cdot v_i = 0$. We know that $\omega \cdot v_i \leq 0$ because $\omega \in R(P)^\vee$ but suppose for contradiction that $w \cdot v_i < 0$ then $\omega' \cdot v_i > 0$. Now $\omega_\varepsilon \cdot v_i = \omega \cdot v_i + \varepsilon \omega' \cdot v_i < \omega \cdot v_i - \omega \cdot v_i = 0$ by the choice of ε, which is a contradiction. Hence $\omega \cdot v_i = 0$. □

3.4 Polyhedral complexes and fans

Definition 3.4.1 A collection Σ of polyhedra in $\mathbb{R}^n$ is called a *polyhedral complex* if it satisfies:

- for $A \in \Sigma$ every face of A is in Σ, and
- for every $A, B \in \Sigma$, if $A \cap B \neq \emptyset$ then $A \cap B$ is a face of A (and of B).

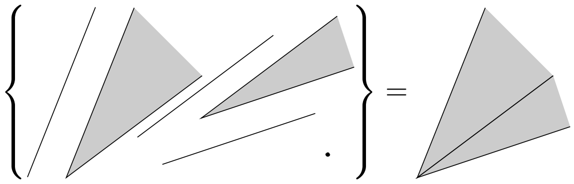

Figure 6: The polyhedral fan mentioned in Example 3.4.2.

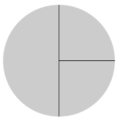

Figure 7: The collection of cones in Example 3.4.3 not being a fan.

The *support* of Σ is $\text{supp}(\Sigma) := \bigcup_{A \in \Sigma} A$. A polyhedral complex is called *complete* if its support is $\mathbb{R}^n$. A polyhedral complex consisting only of polyhedral cones is called a *polyhedral fan.*

Example 3.4.2 The fan in Figure 6 consists of two 2-dimensional cones, three 1-dimensional cones and one zero-dimensional cone. To actually see the cones we need to pull them apart when drawing. We check by inspection that the properties for being a polyhedral complex are satisfied.

Example 3.4.3 The three two-dimensional cones in Figure 7 cannot be part of the same polyhedral complex, since the big cone intersected with one of the small cones gives a cone which is not a face of the big cone.

Proposition 3.4.4 *Let $P \subseteq \mathbb{R}^n$ be a polyhedron. The set of faces of P is a polyhedral complex.*

Proof. To see this recall that by Corollary 3.3.16 every face of a face A of P is a face of P. Furthermore, by Proposition 3.3.12 if A and B are faces of P with non-empty intersection then $A \cap B$ is a face of A. □

Definition 3.4.5 Let $P \subseteq \mathbb{R}^n$ be a polyhedron and F a face of P. We define the *normal cone* of F to be

$$N_P(F) := \overline{\{\omega \in \mathbb{R}^n : \text{face}_\omega(P) = F\}}$$

where the closure is taken in the usual topology on $\mathbb{R}^n$.

Lemma 3.4.6 *Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{m' \times n}$ such that $C := \{x \in \mathbb{R}^n : Ax \leq 0 \text{ and } Bx < 0\}$ is non-empty, then $\overline{C} = \{x \in \mathbb{R}^n : Ax \leq 0 \text{ and } Bx \leq 0\}$.*

Proof. Let $p \in C$. To prove $\supseteq$, let $x \in \mathbb{R}^n \setminus \{p\}$ such that $Ax \leq 0$ and $Bx \leq 0$. The open line segment from p to x is contained in C. Therefore $x \in \overline{C}$. For the other inclusion we observe that the right hand side contains C and is closed. □

Proposition 3.4.7 *Every normal cone $N_P(F)$ is a polyhedral cone.*

Proof. By second part of Proposition 3.3.13, a vector picks out the face F if it satisfies a certain set of linear inequalities. Since F is a face such a vector exists and Lemma 3.4.6 tells us that the closure of these vectors is a polyhedral cone. □

Definition 3.4.8 The normal fan $NF(P)$ of a polyhedron $P \subseteq \mathbb{R}^n$ is the collection of all normal cones of faces of P.

We state the following proposition without proof.

Proposition 3.4.9 *The normal fan of a polyhedron is a polyhedral fan.*

Lemma 3.4.10 *For a polyhedron $P \subseteq \mathbb{R}^n$ we have* $\mathrm{supp}(NF(P)) = R(P)^\vee$.

Proof. Every ω in $R(P)^\vee$ defines a face of P, which proves the inclusion $\supseteq$. To prove $\subseteq$, we observe that the vectors introduced on the left hand side when taking normal cones are in the closure of the vectors defining faces. The inclusion follows since the right hand side is closed. □

Definition 3.4.11 Let Σ_1 be a polyhedral complex in $\mathbb{R}^n$ and Σ_2 be a polyhedral complex in $\mathbb{R}^{n'}$. The *product complex* is defined as

$$\Sigma_1 \times \Sigma_2 := \{P_1 \times P_2 : P_1 \in \Sigma_1, P_2 \in \Sigma_2\}$$

and consists of polyhedra in $\mathbb{R}^{n+n'}$.

Lemma 3.4.12 *Let $P_1 \subseteq \mathbb{R}^n$ and $P_2 \subseteq \mathbb{R}^{n'}$ be polyhedra. The faces of the polyhedron $P_1 \times P_2 \subseteq \mathbb{R}^{n+n'}$ are exactly the products $F_1 \times F_2$ where F_1 is a face of P_1 and F_2 a face of P_2.*

Proof. A vector $\omega = \omega_1 \times \omega_2 \in \mathbb{R}^n \times \mathbb{R}^{n'}$ attains its maximum over $P_1 \times P_2$ in a point $p_1 \times p_2$ if and only if ω_1 attains its maximum over P_1 in p_1 and ω_2 its maximum over P_2 in p_2. This proves that the faces of $P_1 \times P_2$ are exactly the product of faces of P_1 and P_2. □

Lemma 3.4.13 *The product complex is a polyhedral complex.*

Proof. By Lemma 3.4.12 the face of a polyhedron $P_1 \times P_2$ in $\Sigma_1 \times \Sigma_2$ is of the form $F_1 \times F_2$ with F_i face of P_i. Because Σ_1 and Σ_2 are polyhedral complexes, $F_1 \times F_2$ is in $\Sigma_1 \times \Sigma_2$. Suppose now that we have $A_1 \times A_2 \cap B_1 \times B_2$ non-empty. Then $A_1 \cap B_1$ is non-empty and it must be a face of A_1. Similarly, $A_2 \cap B_2$ is a face of A_2. By Lemma 3.4.12 the product $(A_1 \cap B_1) \times (A_2 \cap B_2) = A_1 \times A_2 \cap B_1 \times B_2$ is a face of $A_1 \times A_2$. □

Lemma 3.4.14 *Let $P_{A,b} \subseteq \mathbb{R}^n$ be a polyhedron. Let $\omega \in R(P_{A,b})^\vee$ and $F =$* $\mathrm{face}_\omega(P_{A,b})$. *Let $M \subseteq \{1, \ldots, m\}$ be the subset of indices i such that $F \subseteq H_{A_i \cdot, b_i}$. Then $F = P_{A,b} \cap \bigcap_{i \in M} H_{A_i \cdot, b_i}$.*

Proof. The inclusion $\subseteq$ is clear. For $\supseteq$ we consider the situation in the subspace $\bigcap_{i\in M} H_{A_{i\cdot},b_i}$. Here the face is full dimensional. Let p be an interior point. Assume that x is on the right hand side. We need to prove that ω is maximal on x. Suppose not and consider the line segment between x and p. Continuing it slightly, we stay in side F because p is in the interior. Using that the ω-value is not maximal in x but is in p we get even larger ω-values on the extended line segment. This is a contradiction. □

Proposition 3.4.15 *Let Σ be a polyhedral complex in $\mathbb{R}^n$ with $n \geq 1$. Consider the hyperplane $H := H_{e_1,0} \subseteq \mathbb{R}^n$. The set $\Sigma' := \{P \cap H : P \in \Sigma \text{ and } P \cap H \neq \emptyset\}$ is a polyhedral complex.*

Proof. Let F be a face of $P \cap H$ where $P \in \Sigma$. We wish to show that $F \in \Sigma'$. The polyhedron P is described by list of inequalities and H by two. By Lemma 3.4.14 F is gotten by turning some of these inequalities into equations. The changed inequalities define supporting hyperplanes for P and therefore faces of P. By Proposition 3.4.4 the set of faces of P is a complex. Therefore the intersection of the faces in question is a face F' of P. We now have $F = F' \cap H$ as desired.

Let $A \cap H$ and $B \cap H$ be elements in Σ' with $A, B \in \Sigma$. Suppose that some $p \in (A \cap H) \cap (B \cap H) \neq \emptyset$ then also $A \cap B \neq \emptyset$. This means that $A \cap B$ is a face of A. There exists ω such that $\text{face}_\omega(A) = A \cap B$. We claim that $\text{face}_\omega(A \cap H) = \text{face}_\omega(A) \cap H$. To see this it suffices to prove that ω has the same maximum over $A \cap H$ and A. But this holds since $p \in A \cap H$. □

Definition 3.4.16 Let Σ_1 and Σ_2 be polyhedral complexes in $\mathbb{R}^n$. We define their *common refinement* as follows:

$$\Sigma_1 \wedge \Sigma_2 := \{P_1 \cap P_2 : P_1 \in \Sigma_1, P_2 \in \Sigma_2, P_1 \cap P_2 \neq \emptyset\}.$$

Proposition 3.4.17 *The common refinement of two polyhedral complexes Σ_1 and Σ_2 in $\mathbb{R}^n$ is a polyhedral complex with*

$$\text{supp}(\Sigma_1 \wedge \Sigma_2) = \text{supp}(\Sigma_1) \cap \text{supp}(\Sigma_2).$$

Proof. By Lemma 3.4.13 $\Sigma_1 \times \Sigma_2$ is a complex in $\mathbb{R}^{n+n}$. Let $\Delta := \{(x,x) : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^{n+n}$ be the "diagonal". After a linear transformation, Proposition 3.4.15 tells us that

$$A := \{P \cap \Delta : P \in \Sigma_1 \times \Sigma_2, P \cap \Delta \neq \emptyset\}$$

is a polyhedral complex. It is also the common refinement of $\Sigma_1 \times \Sigma_2$ and $\{\Delta\}$. Let $\pi : \Delta \to \mathbb{R}^n$ be the bijective projection on the first copy of $\mathbb{R}^n$. We observe that $\{\pi(P) : P \in A\} = \Sigma_1 \wedge \Sigma_2$. This proves that $\Sigma_1 \wedge \Sigma_2$ is a polyhedral complex. The statements about the supports follows from the definition of the common refinement and the distributive rule for union and intersection. □

Proposition 3.4.18 *Let P_1 and P_2 be polyhedra in $\mathbb{R}^n$. Then*

$$NF(P_1 + P_2) = NF(P_1) \wedge NF(P_2).$$

Proof. By Theorem 3.2.6 we may write P_1 and P_2 in the form

$$P_i = \mathrm{conv}(U_i) + \mathrm{cone}(V_i)$$

for finite sets $U_i, V_i \subseteq \mathbb{R}^n$. Hence

$$P = \mathrm{conv}(u_1 + u_2 : (u_1, u_2) \in U_1 \times U_2) + \mathrm{cone}(V_1 \cup V_2).$$

A vector ω attains its maximum over P_1+P_2 if and only if it attains its maximum over P_1 and over P_2. Let ω pick the face $\mathrm{face}_\omega(P)$ of P. By Proposition 3.3.13

$$\mathrm{face}_\omega(P) = \mathrm{conv}(u_1 + u_2 : (u_1, u_2) \in U) + \mathrm{cone}(V)$$

for subsets $U \subseteq U_1 \times U_2$ and $V \subseteq V_1 \cup V_2$. Here $V = (V_1 \cup V_2) \cap \omega^\perp = V_1 \cap \omega^\perp \cup V_2 \cap \omega^\perp$ and U is the set of pairs maximizing ω, and thus equal to $U_1' \times U_2'$ with $U_i' \subseteq U_i$ maximizing ω. By Lemma 3.4.6 the condition for a vector ω' to be in $N_P(\mathrm{face}_\omega(P))$ is that $\omega' \cdot (u_1' + u_2') \geq \omega' \cdot (u_1 + u_2)$ for $(u_1', u_2') \in U$ and $(u_1, u_2) \in U_1 \times U_2$, furthermore that $\omega' \in V^\perp$. Applying Proposition 3.3.13 and Lemma 3.4.6 again to P_1, P_2 and ω we see that this is exactly the condition for being in both $N_{P_1}(\mathrm{face}_\omega(P))$ and $N_{P_2}(\mathrm{face}_\omega(P))$.

To prove that the right hand side is contained in the left hand side we use Theorem 3.4.19 below. The supports of the two fans are equal because $\mathrm{supp}(NF(P_1 + P_2)) = R(P_1 + P_2)^\vee = (R(P_1) + R(P_2))^\vee = R(P_1)^\vee \cap R(P_2)^\vee = \mathrm{supp}(NF(P_1) \cap \mathrm{supp}(NF(P_2)) = \mathrm{supp}(NF(P_1) \wedge NF(P_2))$. □

We present the following theorem without proof.

Theorem 3.4.19 *Let Σ_1 and Σ_2 be polyhedral complexes with $\Sigma_1 \subseteq \Sigma_2$ and* $\mathrm{supp}(\Sigma_1) = \mathrm{supp}(\Sigma_2)$. *Then $\Sigma_1 = \Sigma_2$.*

4 The Gröbner fan of an ideal

For a polynomial $f \in k[x_1, \dots, x_n]$ we have

$$\text{face}_\omega(NP(f)) = NP(\text{in}_\omega(f)).$$

We conclude that ω and ω' pick out the same initial form of f if and only if $\text{face}_\omega(NP(f)) = \text{face}_{\omega'}(NP(f))$. Therefore, which initial form ω picks from f depends on which normal cones of $NP(f)$ the vector ω belongs to. See Example 4.0.20 below.

In this section we will generalize the concept of a normal fan of a Newton polytope of a polynomial, namely *we will fix an ideal* $I \subseteq k[x_1, \dots, x_n]$ and define its *Gröbner fan*. First we consider the equivalence relation on $\mathbb{R}^n$:

$$u \sim v \Leftrightarrow \text{in}_u(I) = \text{in}_v(I) \tag{1}$$

with initial ideals defined as in Definition 1.6.1. In particular, for a vector $v \in \mathbb{R}^n$ and a term ordering $\prec$ we consider the closure of the equivalence classes:

$$C_\prec(I) := \overline{\{u \in \mathbb{R}^n : \text{in}_u = \text{in}_\prec(I)\}} \text{ and}$$
$$C_v(I) := \overline{\{u \in \mathbb{R}^n : \text{in}_u = \text{in}_v(I)\}}.$$

We will prove the following:

- The set $\{u \in \mathbb{R}^n : \text{in}_u(I) = \text{in}_\prec(I)\}$ is indeed an equivalence class. (That is, $\text{in}_\prec(I)$ is of the form $\text{in}_v(I)$ for some $v \in \mathbb{R}^n$.)
- There are only finitely many initial ideals of the form $\text{in}_\prec(I)$ and of the form $\text{in}_v(I)$.
- For $v \in \mathbb{R}^n_{>0}$ the set $C_v(I)$ is a polyhedral cone and every face of $C_v(I)$ is of the form $C_u(I)$ for some u.
- We can choose a set of cones $C_v(I)$ which cover $\mathbb{R}^n_{\geq 0}$ and form a polyhedral fan.

The argument below for finiteness (Proposition 4.1.1) was presented by Sturmfels [13], while the structure of the proof that the Gröbner fan is a fan (and its construction) comes from [7]. The original construction of the Gröbner fan is by Mora and Robbiano [12].

Example 4.0.20 Consider the principal ideal $I = \langle x + y + 2xy^2 + 3x^2y\rangle \subseteq k[x, y]$. This ideal has 9 initial ideals giving rise to 9 polyhedral cones forming a fan as shown in Figure 8.

Example 4.0.21 Consider the ideal $I = \langle x - 1, y - 1\rangle$. The ideal has 5 initial ideals. The vectors $(-1, 3)$ and $(3, -1)$ pick out the same initial ideal $\text{in}_{(-1,3)}(I) = \text{in}_{(3,-1)}(I) = \langle 1\rangle$. The equivalence class is not convex since $\frac{1}{2}((-1, 3)+(3, -1)) = (1, 1)$ which picks out the initial ideal $\langle x, y\rangle$. See Figure 8.

Example 4.0.22 [13, Example 3.9] The ideal $I = \langle x^5 - 1 + z^2 + y^3, y^2 - 1 + z + x^2, z^3 - 1 + y^5 + x^6\rangle \subseteq \mathbb{Q}[x, y, z]$ has 360 initial ideals of the form $\text{in}_\prec(I)$. The cones $C_\prec(I)$ of these together with their faces form a fan shown in Figure 9.

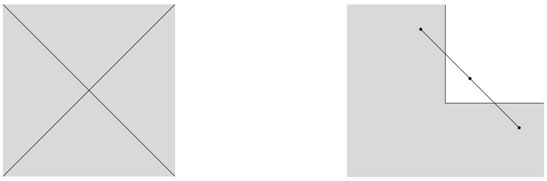

Figure 8: The closures of equivalence classes in Example 4.0.20 and Example 4.0.21.

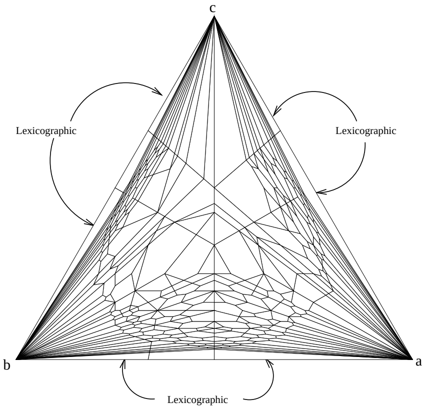

Figure 9: The intersection of the triangle with corners $(1,0,0),(0,1,0)$ and $(0,0,1)$ with the Gröbner fan of the ideal I of Example 4.0.22 as defined in Definition 4.3.1.

4.1 Finiteness

We have seen in Exercise 8 on the first sheet that for $n > 1$ the polynomial ring $k[x_1, \ldots, x_n]$ has infinitely (in fact uncountably) many term orders. The following shows that these only define finitely many initial ideals of a fixed ideal.

Proposition 4.1.1 *Let $I \subseteq k[x_1, \ldots, x_n]$ be a polynomial ideal. Then I has only finitely many initial ideals of the form $\text{in}_\prec(I)$ where $\prec$ is a term ordering.*

Proof. By contradiction. Let Σ_0 be the set of initial ideals of I and suppose that $|\Sigma_0| = \infty$. In particular, $I \neq \langle 0 \rangle$ and we may choose a non-zero $f_1 \in I$.

Each $M \in \Sigma_0$ contains a term of f_1. Hence we may choose a term m_1 of f_1 which is contained in infinitely many $M \in \Sigma_0$. We let $J_1 = \langle m_1 \rangle$ and $\Sigma_1 := \{M \in \Sigma_0 : J_1 \subseteq M\}$. Since Σ_1 is infinite, there is some $M_1 \in \Sigma_1$ with $J_1 \subset M_1$ (strictly). By Proposition 1.6.10 the monomials outside M_1 form a k-vector basis for $k[x_1, \ldots, x_n]/I$. Therefore, since $J_1 \subset M_1$, the set of all monomials outside J_1 must be dependent modulo I. Consequently, there exists a non-zero $f_2 \in I$ with all terms of f_2 outside J_1.

Each $M \in \Sigma_1$ contains a term of f_2. Hence we may choose a term m_2 of f_2 which is contained in infinitely many $M \in \Sigma_1$. We let $J_2 = \langle m_1, m_2 \rangle$ and $\Sigma_2 := \{M \in \Sigma_1 : J_2 \subseteq M\}$. Since Σ_2 is infinite, there is some $M_2 \in \Sigma_2$ with $J_2 \subset M_2$ (strictly). By Proposition 1.6.10 the monomials outside M_2 form a k-vector basis for $k[x_1, \ldots, x_n]/I$. Therefore, since $J_2 \subset M_2$, the set of monomials outside J_2 must be dependent modulo I. Consequently, there exists a non-zero $f_3 \in I$ with all terms of f_3 outside J_2.

Continuing like this we construct an infinite sequence of strict inclusions:

$$J_1 \subset J_2 \subset J_3 \ldots$$

This contradicts Corollary 1.2.5. □

4.2 Every $C_\prec(I)$ is of the form $C_v(I)$

Recall that we use the notation $\mathcal{G}_\prec(I)$ for the reduced Gröbner basis of I with respect to $\prec$. Furthermore, for $f = \sum_{u \in U} c_u x^u$ with support U we call $\max_{u \in U}(v \cdot u)$ the *v-degree* of f.

Lemma 4.2.1 *[7, Lemma 2.10] Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal, $\prec$ a term ordering and $v \in \mathbb{R}^n$. Then*

$$\text{in}_v(I) = \text{in}_\prec(I) \Leftrightarrow \forall g \in \mathcal{G}_\prec(I) : \text{in}_v(g) = \text{in}_\prec(g).$$

Proof. $\Rightarrow$: Let $g \in \mathcal{G}_\prec(I)$. Since $\mathcal{G}_\prec(I)$ is reduced only one term of g is in $\text{in}_\prec(I)$, namely $\text{in}_\prec(g)$. The initial form $\text{in}_v(g)$ is in the monomial ideal $\text{in}_v(I) = \text{in}_\prec(I)$. Hence every term of $\text{in}_v(g)$ is in $\text{in}_\prec(I)$. We conclude that $\text{in}_v(g) = \text{in}_\prec(g)$.

$\Leftarrow$: To show $\text{in}_v(I) \supseteq \text{in}_\prec(I)$ we use that $\text{in}_\prec(I)$ is generated by $\text{in}_\prec(g)$ with $g \in \mathcal{G}_\prec(I)$ because $\mathcal{G}_\prec(I)$ is a Gröbner basis. Since $\text{in}_\prec(g) = \text{in}_v(g) \in \text{in}_v(I)$ the inclusion follows.

To show $\text{in}_v(I) \subseteq \text{in}_\prec(I)$ we let $f \in I \backslash \{0\}$ and wish to show $\text{in}_v(f) \in \text{in}_\prec(I)$. Using the division algorithm with $\prec$ we may write $f = \sum_j m_j g_j + 0$ where m_j is a term and $g_j \in \mathcal{G}_\prec(I)$. By Lemma 1.5.5 we have $\text{in}_\prec(m_j g_j) \preceq \text{in}_\prec(f)$ but the same argument shows that since the v-degree of p in Algorithm 1.5.1 is non-strictly decreasing and $\text{in}_\prec(g_j)$ has maximal v-degree among terms of g_j we have that the v-degree of f is non-strictly larger than the v-degree of each $m_j g_j$. (Left to the reader.) Therefore $\text{in}_v(f) = \sum_{j \in J} \text{in}_v(m_j g_j) = \sum_{j \in J} m_j \text{in}_v(g_j)$ for a suitable index set J. This proves $\text{in}_v(f) \in \text{in}_v(I)$. □

Lemma 4.2.2 *Let $A \in \mathbb{R}^{d \times n}$ define a matrix term ordering $\prec_A$ and $u, v \in \mathbb{N}^n$. If $x^u \prec_A x^v$ then for $\varepsilon \in \mathbb{R}_{>0}$ sufficiently small*

$$(\varepsilon^0 A_{1\cdot} + \varepsilon^1 A_{2\cdot} + \cdots + \varepsilon^{d-1} A_{d\cdot}) \cdot (u - v) < 0.$$

Proof. We wish to prove $(\varepsilon^0, \varepsilon^1, \ldots, \varepsilon^{d-1}) A(u - v) < 0$ for $\varepsilon > 0$ sufficiently small. The first non-zero entry of the vector $A(u-v)$ is negative by assumption. Call it $-a$. Let $M > 0$ be a bound on the absolute value of the other entries. For $0 < \varepsilon < \min(1, (a/(Mn)))$ we have $\varepsilon^i < a/(Mn)$ for $i \geq 1$. This strictly bounds the contribution of the positive terms to the dot product above by a. □

Proposition 4.2.3 *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $\prec$ a term ordering. There exists a vector $\omega \in \mathbb{R}^n_{>0}$ such that $\text{in}_\omega(I) = \text{in}_\prec(I)$.*

Proof. By Theorem 2.3.5 there exists a matrix $A \in \mathbb{R}^{d \times n}$ representing $\prec$ such that $\prec_A = \prec$. Applying Lemma 4.2.2 several times we can find a ε such that $\omega := (\varepsilon^0 A_{1\cdot} + \varepsilon^1 A_{2\cdot} + \cdots + \varepsilon^{d-1} A_{d\cdot})$ such that $\text{in}_\omega(g) = \text{in}_{\prec_A}(g)$ for all g in the finite set $\mathcal{G}_\prec(I)$. By Lemma 4.2.1 $\text{in}_\omega(I) = \text{in}_\prec(I)$. □

We conclude that every $C_\prec(I)$ is of the form $C_v(I)$ for some $v \in \mathbb{R}^n$.

Corollary 4.2.4 *[7, Corollary 2.11] Let $\prec$ be a term order and $v \in \mathbb{R}^n$. Then*

$$v \in C_\prec(I) \Leftrightarrow \forall g \in \mathcal{G}_\prec(I) : \text{in}_\prec(g) = \text{in}_\prec(\text{in}_v(g)).$$

Proof. By Lemma 4.2.1 $C_\prec(I)$ is the closure of the set of

$$\{v \in \mathbb{R}^n : \forall g \in \mathcal{G}_\prec(I) : \text{in}_v(g) = \text{in}_\prec(g)\}.$$

These conditions translate into a set of linear inequalities on v. By Proposition 4.2.3 the set is non-empty and by Lemma 3.4.6 its closure is gotten by making the strict inequalities non-strict. This translates into the short expressions "$\text{in}_\prec(g) = \text{in}_\prec(\text{in}_v(g))$". (Left to the reader.) □

Since the condition $\text{in}_\prec(g) = \text{in}_\prec(\text{in}_v(g))$ translates into a set of non-strict linear inequalities on v, we conclude that the set $C_\prec(I)$ is a polyhedral cone.

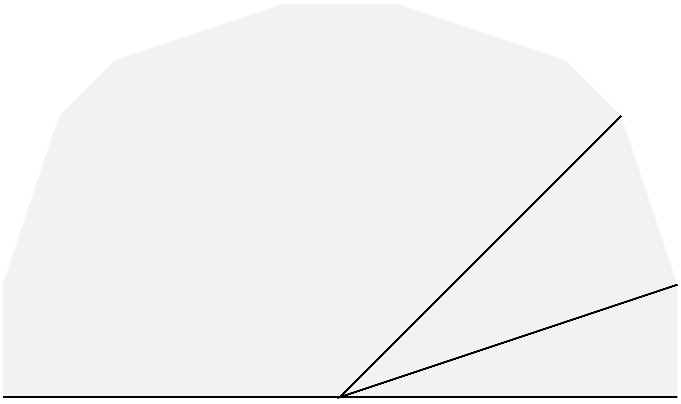

Figure 10: The Gröbner fan in Example 4.3.2. The Gröbner region is $\mathbb{R} \times \mathbb{R}_{\geq 0}$.

4.3 Definition of the Gröbner fan

Now that we know that $C_{\prec}(I)$ is a polyhedral cone we are ready to define the Gröbner fan of an ideal.

Definition 4.3.1 Let $I \subseteq k[x_1, \dots, x_n]$ be an ideal. We define the *Gröbner fan* of I to be the set of cones of the form $C_{\prec}(I)$ together with all their faces. We denote it by $\mathrm{Gfan}(I)$. The support of $\mathrm{Gfan}(I)$ is called the *Gröbner region* of I.

Example 4.3.2 Consider $I = \langle y^2 + 1 + xy, x^2 \rangle \subseteq \mathbb{Q}[x, y]$. This ideal happens to have three initial ideals of the form $\mathrm{in}_{\prec}(I)$: $\langle y^4, x \rangle$, $\langle y^3, xy, x^2 \rangle$, and $\langle y^2, x^2 \rangle$. For each of these ideals we find a term order and a Gröbner bases as in Corollary 4.2.4. For each Gröbner bases we underline the initial terms with respect to the ordering:

- $\{\underline{y^3} + y - x, \underline{xy} + 1 + y^2, \underline{x^2}\}$
- $\{\underline{y^2} + 1 + xy, \underline{x^2}\}$
- $\{\underline{y^4} + 1 + 2y^2, \underline{x} - y - y^3\}$

We apply Corollary 4.2.4. For the first Gröbner basis the condition $\mathrm{in}_{\prec}(\mathrm{in}_v(y^3 + y - x)) = y^3$ translates into $3v_2 \geq v_2$ (because y^3 must be preferred over y) and $3v_2 \geq v_1$ (because y^3 is preferred over x). The Gröbner basis element $\underline{xy} + 1 + y^2$ gives inequalities $v_1 + v_2 \geq 0$ and $v_1 + v_2 \geq 2v_2$. The third Gröbner basis element $\underline{x^2}$ gives no condition. In total we have four inequalities, but two are redundant. The cone is the middle cone of Figure 10. The same procedure gives the other two full-dimensional Gröbner cones of the Gröbner fan. Taking all faces we get all cones of the Gröbner fan. It consists of $3 + 4 + 1 = 8$ cones. We notice that the cones cover more than $\mathbb{R}^2_{\geq 0}$ and less than $\mathbb{R}^2$.

Our first observation is that since I has only finitely many initial ideals of the form $\mathrm{in}_{\prec}(I)$, there are only a finite set of cones $C_{\prec}(I)$. Each of these has only finitely many faces (Corollary 3.3.14). Hence Gfan(I) is a finite set of cones. We still need to show that it actually is a fan.

4.4 Equivalence classes are relatively open cones

When we say "polyhedral cone" we have so far always meant closed polyhedral cones, and unless stated otherwise we still mean that polyhedral cones are closed. By an *open polyhedral cone* in $\mathbb{R}^n$ we mean a finite intersection of open halfspaces passing through the origin. If $C \subseteq \mathbb{R}^n$ is a set with the property that there exists a linear subspace $L \subseteq \mathbb{R}^n$ with $C \subseteq L$ and C considered as a subset of L (in the topology induced on L) then C is called a *relatively open polyhedral cone.* Alternatively, if we do not wish to use topological notions, we can define a relatively open polyhedral cone to be the intersection of an open polyhedral cone and a linear subspace of $\mathbb{R}^n$.

Example 4.4.1 The sets

- $\mathbb{R}^3_{>0} \subseteq \mathbb{R}^3$
- $\mathbb{R}^2_{>0} \times \{(0)\} \subseteq \mathbb{R}^3$
- $\{(0,0,0)\} \subseteq \mathbb{R}^3$

are relatively open cones.

In this subsection we will prove that for fixed ideal I the equivalence relation defined in Equation 1, page 40 gives rise to equivalence classes which are relatively open cones - almost. Example 4.0.21 showed that equivalence classes may be non-convex, so we need to be careful when stating our result.

In [13] proofs of some of the following statements are given, but there $v \in \mathbb{R}^n$ is assumed to have non-negative entries. In the following we need to be more careful. The vector v may have negative entries.

Lemma 4.4.2 *[7, Lemma 2.12] Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal, and $v \in \mathbb{R}^n$. If $f \in \mathrm{in}_v(I)$ then we may write f as a sum $\sum_i \mathrm{in}_v(c_i)$ with $c_i \in I$ and each $\mathrm{in}_v(c_i)$ having different v-degree.*

Proof. By definition of the initial ideal we may write f as $\sum_i a_i \mathrm{in}_v(p_i)$ where $p_i \in I$ and a_i are polynomials. In fact we may assume that a_i are single terms. We rewrite $f = \sum_i a_i \mathrm{in}_v(p_i) = \sum_i \mathrm{in}_v(a_i p_i)$. So in fact, we may assume that $a_i = 1$. All terms of each summand $\mathrm{in}_v(p_i)$ have the same v-degree. Suppose $\mathrm{in}_v(p_i)$ and $\mathrm{in}_v(p_j)$ have the same v-degree. Then either $\mathrm{in}_v(p_i) + \mathrm{in}_v(p_j) = 0$ or $\mathrm{in}_v(p_i) + \mathrm{in}_v(p_j) = \mathrm{in}(p_i + p_j)$. In the sum $\sum_i \mathrm{in}_v(p_i)$ we may therefore group polynomials together, possibly removing summands, until the sum involves only summands with different v-degrees. □

Lemma 4.4.3 *[7, Lemma 2.13] Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $\prec$ a term ordering. If a vector v is in $C_\prec(I)$ then $\mathrm{in}_\prec(\mathrm{in}_v(I)) = \mathrm{in}_\prec(I)$.*

Proof. To prove "$\supseteq$" we look at a generator of $\text{in}_\prec(f)$ where $f \in \mathcal{G}_\prec(I)$ and observe that by Corollary 4.2.4 $\text{in}_\prec(f) = \text{in}_\prec(\text{in}_v(f))$. This proves $\text{in}_\prec(f) \in \text{in}_\prec(\text{in}_v(I))$.

To prove "$\subseteq$" we let $f \in \text{in}_v(I)$. By the Lemma above we may write $f = \sum_i \text{in}_v(c_i)$ with $c_i \in I$ and c_i having different v-degree. There must exist j such that $\text{in}_\prec(f) = \text{in}_\prec(\text{in}_v(c_j))$. We want to prove that this is in $\text{in}_\prec(I)$. Using the division algorithm with $\mathcal{G}_\prec(I)$ and $\prec$ we write

$$c_j = m_1 g_1 + \cdots + m_r g_r$$

with m_i being terms and $g_i \in \mathcal{G}_\prec(I)$. We now argue that the v-degree of c_j is $\geq$ the v-degree of any $m_i g_i$. The reason for this is that the v-degree of p in the algorithm is non-strictly decreasing. Namely we cancel a term of p with the largest term of (a multiple of) an $f_i \in \mathcal{G}_\prec(I)$ in the algorithm, possibly introducing new terms and canceling other. The terms that are introduced cannot have larger v-degree than the term P we wish to cancel because $\text{in}_\prec(\text{in}_v(f_i)) = \text{in}_\prec(f_i)$ by Corollary 4.2.4. The terms $m_i g_i$ are produced by the algorithm when the a_i are modified. At this point the v-degree is bounded by the degree of p. We conclude that for all terms of $m_i g_i$ have smaller degree than c_j. Hence

$$\text{in}_v(c_j) = \sum_{i \in J} \text{in}_v(m_i g_i)$$

for some index set J. In the "$P := \text{in}_\prec(p)$" variant of the division algorithm $\text{in}_\prec(p)$ gets $\prec$ smaller in each iteration. Therefore $\text{in}_\prec(m_1 g_1), \ldots, \text{in}_\prec(m_r g_r)$ are all different. By Corollary 4.2.4 they equal $\text{in}_\prec(\text{in}_v(m_1 g_1)), \ldots, \text{in}_\prec(\text{in}_v(m_r g_r))$. The initial term $\text{in}_\prec(\text{in}_v(c_j)) = \text{in}_\prec(\sum_{i \in J} \text{in}_v(m_i g_i))$ must equal $\text{in}_\prec(\text{in}_v(m_i g_i)) = \text{in}_\prec(m_i g_i)$ for some i. This proves $\text{in}_\prec(\text{in}_v(c_j))$ is in $\text{in}_\prec(I)$. □

Corollary 4.4.4 *[7, Corollary 2.14] Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal, $\prec$ a term ordering and $v \in C_\prec(I)$. Then $\mathcal{G}_\prec(\text{in}_v(I)) = \{\text{in}_v(f) : f \in \mathcal{G}_\prec(I)\}$.*

Proof. To show that we have a Gröbner basis we must prove that $\text{in}_\prec(\text{in}_v(I)) = \langle \text{in}_\prec(\text{in}_v(f)) : f \in \mathcal{G}_\prec(I) \rangle$. By the lemma $\text{in}_\prec(\text{in}_v(I)) = \text{in}_\prec(I)$. By Corollary 4.2.4 $\{\text{in}_\prec(\text{in}_v(f)) : f \in \mathcal{G}_\prec(I)\} = \{\text{in}_\prec(f) : f \in \mathcal{G}_\prec(I)\}$, which equals $\text{in}_\prec(I)$ since $\mathcal{G}_\prec(I)$ is a Gröbner basis. The set $\{\text{in}_\prec(\text{in}_v(f)) : f \in \mathcal{G}_\prec(I)\}$ is a minimal generating set since $\{\text{in}_\prec(f) : f \in \mathcal{G}_\prec(I)\}$ is. The set $\{\text{in}_v(f) : f \in \mathcal{G}_\prec(I)\}$ is a reduced Gröbner basis because no non-initial terms of elements in $\mathcal{G}_\prec(I)$ are in $\text{in}_\prec(I) = \text{in}_\prec(\text{in}_v(I))$. □

Algorithm 4.4.5
Input: *An ideal $I \subseteq k[x_1, \ldots, x_n]$ and a vector $v \in \mathbb{R}^n_{\geq 0}$ and a term order $\prec$.*
Output: *A reduced Gröbner basis for $\text{in}_v(I)$ with respect to $\prec$.*

- *Compute $\mathcal{G}_{\prec_v}(I)$.*
- *Return $\{\text{in}_v(f) : f \in \mathcal{G}_{\prec_v}(I)\}$.*

Proof. We first prove that $v \in C_{\prec_v}(I)$ using Corollary 4.2.4. For g in the Gröbner basis of the corollary we have to check that $\text{in}_{\prec_v}(g) = \text{in}_{\prec_v}(\text{in}_v(g))$, but this holds for any polynomial g. The correctness now follows from Corollary 4.4.4. □

Example 4.4.6 Let $I = \langle x + y + z^2, y^2 - z^3 \rangle \subseteq \mathbb{Q}[x, y, z]$ and $v = (0, 1, 1)$. To compute $\text{in}_v(I)$ we first compute

$$\begin{gathered}\mathcal{G}_{\prec_v}(I) = \{z^2 + y + x,\\ yz + y^2 + xz,\\ y^3 + y^2 + 2xy - xy^2 + x^2 - x^2z\}\end{gathered}$$

where $\prec$ is some term ordering. We now take initial forms:

$$\mathcal{G}_{\prec}(\text{in}_v(I)) = \{z^2, yz + y^2, y^3\}.$$

Lemma 4.4.7 *If* $\text{in}_{\prec}(I) = \text{in}_{\prec'}(I)$ *then* $\mathcal{G}_{\prec}(I) = \mathcal{G}_{\prec'}(I)$.

Proof. Exercise! Hint: the proof of Proposition 1.6.14. □

Proposition 4.4.8 *Let* $I \subseteq k[x_1, \ldots, x_n]$ *be an ideal. There exists only finitely many initial ideals of the form* $\text{in}_v(I)$ *where* $v \in \mathbb{R}^n_{\geq 0}$.

Proof. By Proposition 4.1.1 there are only finitely many initial ideals of the form $\text{in}(I)_{\prec}$. It follows from Lemma 4.4.7 above that there are only a finite number of reduced Gröbner bases of I. By Corollary 4.4.4 a generating set for any initial ideal is gotten by taking initial forms of one of these Gröbner bases. There are only a finite number of choices for what terms a vector v will pick. □

Proposition 4.4.9 *[7, Proposition 2.6] Let* $I \subseteq k[x_1, \ldots, x_n]$ *be an ideal and* $\prec$ *a term order and* $v \in C_{\prec}(I)$. *For* $u \in \mathbb{R}^n$ *we have:*

$$\text{in}_u(I) = \text{in}_v(I) \Leftrightarrow \forall g \in \mathcal{G}_{\prec}(I) : \text{in}_u(g) = \text{in}_v(g).$$

Proof. $\Leftarrow$: Since $v \in C_{\prec}(I)$, Corollary 4.2.4 shows that $\text{in}_{\prec}(g) = \text{in}_{\prec}(\text{in}_v(g)) = \text{in}_{\prec}(\text{in}_u(g))$. Applying the corollary again we get $u \in C_{\prec}(I)$. By Corollary 4.4.4 $\mathcal{G}_{\prec}(\text{in}_v(I)) = \{\text{in}_v(f) : f \in \mathcal{G}_{\prec}(I)\} = \{\text{in}_u(f) : f \in \mathcal{G}_{\prec}(I)\} = \mathcal{G}_{\prec}(\text{in}_u(I))$. Since the Gröbner bases are the same, $\text{in}_v(I) = \text{in}_u(I)$.

$\Rightarrow$: Let $g \in \mathcal{G}_{\prec}(I)$. By Corollary 4.2.4 we have $\text{in}_{\prec}(\text{in}_v(g)) = \text{in}_{\prec}(g)$. By Lemma 4.4.3, since $v \in C_{\prec}(I)$, $\text{in}_{\prec}(\text{in}_u(g)) \in \text{in}_{\prec}(\text{in}_u(I)) = \text{in}_{\prec}(\text{in}_v(I)) = \text{in}_{\prec}(I)$. Since g comes from a reduced basis only one term is in $\text{in}_{\prec}(I)$. Hence $\text{in}_{\prec}(\text{in}_u(g)) = \text{in}_{\prec}(g) = \text{in}_{\prec}(\text{in}_v(g))$. Now we cancel out the term by subtracting. Suppose $\text{in}_u(g) - \text{in}_v(g) \in \text{in}_u(I) = \text{in}_v(I)$ is non-zero. Then $\text{in}_{\prec}(\text{in}_u(g) - \text{in}_v(g)) \in \text{in}_{\prec}(\text{in}_v(I)) = \text{in}_{\prec}(I)$, contradicting that g contains only one term from $\text{in}_{\prec}(I)$.□

Example 4.4.10 Write up inequality system for having

Lemma 4.4.11 *Given a polynomial $g \in k[x_1, \dots, x_n]$ and a vector $v \in \mathbb{R}^n$. The set of vectors u such that $\text{in}_u(g) = \text{in}_v(g)$ is a relatively open polyhedral cone.*

Proof. It is straight forward to see that $\text{in}_u(g) = \text{in}_v(g)$ translates into equations (arising from having all terms in $\text{in}_v(g)$ having the same u-degree) and *strict* inequalities (arising from having all other terms of g u-degree less than the u-degree of $\text{in}_v(g)$). Hence the set is a relatively open polyhedral cone. □

We conclude, using the proposition, that every equivalence class described by the proposition is a finite intersections relatively open polyhedral cones and therefore a relatively open polyhedral cone.

4.5 The relative interior of a cone is an equivalence class

This subsection and the rest of Chapter 4 have not been covered in the lectures in 2014. However, Proposition 4.5.2 and Theorem 4.6.2 were presented without proof.
By the affine span of a set we mean the smallest affine subspace containing it.

Definition 4.5.1 Let $P \subseteq \mathbb{R}^n$ be a d-dimensional polyhedron and $L : \text{affinespan}(P) \rightarrow \mathbb{R}^d$ a bijective affine transformation. The relative interior of P is $L^{-1}(\text{int}(L(P)))$, where $\text{int}(L(P))$ denotes the interior of $L(P)$.

Proposition 4.5.2 *Every polyhedron in $P \subseteq \mathbb{R}^n$ is the the disjoint union of the relative interiors of its faces.*

Proof. SKETCH □

We state the following Lemma without proof.

Lemma 4.5.3 *Define the polyhedron $P := \{x \in \mathbb{R}^n : Ax = 0 \wedge Bx \leq 0\}$ for some $A \in \mathbb{R}^{d \times n}$ and $B \in \mathbb{R}^{d' \times n}$. The relative interior of P is $\{x \in \mathbb{R}^n : A'x = 0 \wedge B'x < 0\}$ where A' consists of all rows of A together with the rows of B with the property that $P \subseteq H_{B_{i\cdot}}$ and where B' consists of the remaining rows of B.*

NOT true that we always make all inequalities strict.

Example 4.5.4

Proposition 4.5.5 *[7, Lemma 2.16] The relative interior of a cone in the Gröbner fan of I is an equivalence class of the equivalence relation defined in Equation (1).*

Proof. By definition every cone in the Gröbner fan is a face C of $C_\prec(I)$ for some term order $\prec$.

Consider Corollary 4.2.4 again:

$$v \in C_\prec(I) \Leftrightarrow \forall g \in \mathcal{G}_\prec(I) : \text{in}_\prec(\text{in}_v(g)) = \text{in}_\prec(g).$$

The condition $\mathrm{in}_\prec(\mathrm{in}_v(g)) = \mathrm{in}_\prec(g)$ is equivalent to saying that the v-degree of every term of g is less than or equal to the v-degree of $\mathrm{in}_\prec(g)$. That is, $C_\prec(I)$ is given by non-strict linear inequalities - one for each term in the tail of every $g \in \mathcal{G}_\prec(I)$.

A face C of $C_\prec(I)$ is therefore described by the same system of inequalities, but where some of the inequalities are turned into equations (Lemma 3.4.14). To get the relative interior of C we make some of the inequalities strict and the other inequalities into equations (Lemma 4.5.3). We get a system of equations and inequalities:

$$u \in \text{rel int}C \Leftrightarrow Au = 0 \wedge A'u < 0$$

for some matrices A and A'. Whether u is in the relative interior of C is determined by the sign pattern of Au and $A'u$. There is one sign for each term in the tail of a polynomial in $\mathcal{G}_\prec(I)$. Now, suppose that u is in the relative interior of C and u' is such that the $(Au, A'u)$ and $(Au', A'u')$ have the same sign pattern. Then $u \in C_\prec(I)$ and therefore $\mathrm{in}_\prec(g)$ is a term of $\mathrm{in}_u(g)$. The sign pattern of Au and $A'u$ now tells which other terms appear in $\mathrm{in}_u(g)$. The vector u' gives the same pattern and therefore $\mathrm{in}_{u'}(g) = \mathrm{in}_u(g)$. This proves by Proposition 4.4.9 that $\mathrm{in}_{u'}(I) = \mathrm{in}_u(I)$. Therefore the relative interior of C is contained in the equivalence class of u.

On the other hand, suppose that u' was a vector such that $\mathrm{in}_{u'}(I) = \mathrm{in}_u(I)$. By Proposition 4.4.9 $\mathrm{in}_{u'}(g) = \mathrm{in}_u(g)$. This proves that the sign pattern of $(Au, A'u)$ and $(Au', A'u')$ are the same. Therefore, u' is also in the relative interior of C. □

4.6 The intersection of two Gröbner cones is a face of both

To prove that the Gröbner fan defined in Definition 4.3.1 is a fan, the only thing that remains to be shown is that cones intersect nicely (Theorem 4.6.1 below).

Theorem 4.6.1 *[7, Proposition 2.18] Let C_1 and C_2 be cones in the Gröbner fan of I then $C_1 \cap C_2$ is a face of C_1 (and C_2).*

Proof. The intersection $C_1 \cap C_2$ is a polyhedral cone. By Proposition 4.5.5, if u is in C_1 all of the equivalences class of u is in C_1. Similarly, for C_2. Therefore the equivalence class of $u \in C_1 \cap C_2$ is in $C_1 \cap C_2$. We conclude that $C_1 \cap C_2$ is a union of equivalence classes.

Let E be one such class and let $u \in E$. By Proposition 4.5.2 u is in the relative interior of a unique face of C_1 which is a cone in the Gröbner fan by definition. The set of relative interior points of this face is exactly the equivalence class E of u and $\overline{E}$ is the face. Hence $C_1 \cap C_2$ is a union of relative interiors of faces of C_1. Our goal is to show that $C_1 \cap C_2$ is the closure of a single such equivalence class.

A face F of C_1 is the intersection $\mathrm{span}_\mathbb{R}(F) \cap C_1$. Hence the span of a face is different for every face of C_1. Consider those spans where F is the closure of an equivalence class $E \subseteq C_1 \cap C_2$. We claim that there is a unique span of maximal dimension. Suppose there were two equivalence classes E_1 and $E_2 \subseteq C_1 \cap C_2$ with maximal dimensional span of dimension D, then $\mathrm{conv}(E_1 \cup E_2)$ is contained in $C_1 \cap C_2$, but with dimension higher than D since the spans of E_1 and E_2 are different. However, this is a contradiction since then $C_1 \cap C_2$ cannot be covered by finitely many relatively open cones of dimension at most D.

Let E be the equivalence class in $C_1 \cap C_2$ of highest dimension D. We claim that $\overline{E} = C_1 \cap C_2$. Since $C_1 \cap C_2$ is closed we get $\overline{E} \subseteq C_1 \cap C_2$. If there was $\omega \in C_1 \cap C_2 \setminus \overline{E}$ then $\mathrm{conv}(\omega \cup \overline{E})$ would have dimension D and be contained in $C_1 \cap C_2$. This is a contradiction, since $\mathrm{conv}(\omega \cup \overline{E})$ cannot be covered by finitely many equivalence classes of lower dimension. □

After a lot of hard work we have proved:

Theorem 4.6.2 *The Gröbner fan of an ideal $I \subseteq k[x_1, \ldots, x_n]$ is a polyhedral fan.*

5 Homogeneous ideals

A natural question to ask is whether the Gröbner fan of an ideal is the normal fan of a polyhedron. In the case of certain *homogeneous* ideals this is indeed the case. **We do not prove that this is the case in these notes.**

5.1 Semigroups and monoids

A *semigroup* $(G, +)$ is a set of elements G together with an operation $+$ which satisfies the associate law:

$$\forall a, b, c \in G : (a + b) + c = a + (b + c).$$

An *abelian semigroup* is one where $+$ is also commutative ($a + b = b + a$). If $(G, +)$ has a neutral element $0 \in G$ with respect to $+$ (meaning $a+0 = a = 0+a$) then $(G, +)$ is called a *monoid*. Monoids and semigroups are almost groups, but they miss inverse elements.

Example 5.1.1 $\mathbb{N} := \{0, 1, 2, 3, \dots\}$ is an abelian semigroup with the operation $+$. It is also an abelian monoid. The same holds for $\mathbb{N}^n$.

Definition 5.1.2 Let A and B be two monoids and $\phi : A \to B$ be a function. If $\phi(0) = 0$ and $\phi(x+y) = \phi(x)+\phi(y)$ then we call ϕ a *monoid homomorphism.*

5.2 The semigroup ring

Definition 5.2.1 For a commutative ring R and an abelian monoid S the semigroup ring $R[S]$ is the set of all functions $f : S \to R$ with $f(x) \neq 0$ for only finitely many $x \in S$. The addition of $f, g \in R[S]$ is given by addition in k:

$$(f + g)(x) := f(x) + g(x)$$

and multiplication is done by "convolution":

$$(f \cdot g)(z) := \sum_{(x,y)\in S\times S:z=x+y} f(x) \cdot f(y). \tag{2}$$

The zero function $S \to \{0\}$ is the neutral element for addition. The neutral element for multiplication is the function which takes $0 \in S$ to $1 \in R$ and all other elements to $0 \in R$.

We notice that there is only a finite number of $f(x) \cdot g(y)$ which can be non-zero in Equation 2 and therefore $(f \cdot g)(z)$ is non-zero for only finitely many z.

Lemma 5.2.2 *The semigroup ring is a commutative ring.*

Proof. Left to the reader. □

Example 5.2.3 Let $R = k$ and $S = \mathbb{N}^n$. Then the semigroup ring $R[S] = k[\mathbb{N}^n]$ is isomorphic to our polynomial ring $k[x_1, \dots, x_n]$. Namely an $f \in R[S]$ is mapped to the polynomial $g = \sum_{u\in\mathbb{N}^n:f(u)\neq 0} f(u)x^u$.

In this course we will always let $R = k$, where k is our field.

5.3 Gradings and homogeneity

In the following we will define gradings on $k[x_1,\ldots,x_n]$. These definitions can easily be generalized to any semigroup ring $R[S]$, but we stick to the polynomial ring to keep things simple.

Definition 5.3.1 By a *grading* on $k[x_1,\ldots,x_n]$ we mean a monoid homomorphism $\phi : \mathbb{N}^n \to S$ to a semigroup S. For $u \in \mathbb{N}^n$ the *ϕ-degree* of x^u and u is $\phi(u)$.

Definition 5.3.2 Let ϕ be a grading on $k[x_1,\ldots,x_n]$. A polynomial $f \in k[x_1,\ldots,x_n]$ is ϕ-homogeneous if every exponent of f has the same ϕ-degree. An ideal $I \subseteq k[x_1,\ldots,x_n]$ is ϕ-homogeneous if it has a generating set of ϕ-homogeneous polynomials.

Example 5.3.3 The *standard grading* on $k[x_1,\ldots,x_n]$ is defined as $\phi : \mathbb{N}^n \to \mathbb{N}$ with $\phi(u) = \sum_{i=1}^n u_i$ for $u \in \mathbb{N}^n$. We sometimes call $\phi(u)$ the *total degree* of x^u. A polynomial is homogeneous in the standard grading if every term has the same usual total degree.

- $x_1^3 - x_2^2x_3 + x_3^3$ is homogeneous in the standard grading.
- $x_1^2 - x_2^1$ is not homogeneous in the standard grading.
- $\langle x_1 - x_2, x_1^3 - x_2^2x_3 + x_3^3 + x_1 - x_2\rangle = \langle x_1 - x_2, x_1^3 - x_2^3x_3 + x_3^3\rangle$ is a homogeneous ideal in the standard grading.

Example 5.3.4 Let $\omega \in \mathbb{R}^n$. We may define the grading $\phi_\omega : \mathbb{N}^n \to \mathbb{R}$ by $\phi_\omega(u) := \omega \cdot u$. For a polynomial $f \in k[x_1,\ldots,x_n]$ the initial form $\text{in}_\omega(f)$ is ϕ_ω-homogeneous by Definition 1.4.1. For simplicity we sometimes just say *ω-homogeneous.* If $I \subseteq k[x_1,\ldots,x_n]$ is an ideal then the initial ideal $\text{in}_\omega(I)$ is ω-homogeneous by Definition 1.6.1.

Example 5.3.5 A matrix $A \in \mathbb{N}^{d\times n}$ defines a grading $\phi_A : \mathbb{N}^n \to \mathbb{N}^d$ by $\phi_A(u) := Au$. Notice that if ω is a a vector in the rowspace of A then any ϕ-homogeneous polynomial is also ω-homogeneous.

A priori, it is not clear that the last ideal in Example 5.3.3 is homogeneous. In the following we will find an algorithm for deciding if an ideal is homogeneous.

Lemma 5.3.6 *Let ϕ be a grading on $k[x_1,\ldots,x_n]$. Let $f, g \in k[x_1,\ldots,x_n]$ be ϕ-homogeneous and $h \in k[x_1,\ldots,x_n]$ a single term. The polynomial hf is ϕ-homogeneous. If the terms of f and g have the same ϕ-degree, then $f+g$ is ϕ-homogeneous.*

Proof. Let $h = cx^u$ and let $c'x^{u'}$ and $c''x^{u''}$ be two terms of f resulting in two terms $cc'x^{u+u'}$ and $cc''x^{u+u''}$. We check that they have the same ϕ-degree:

$$\phi(u+u') = \phi(u) + \phi(u') = \phi(u) + \phi(u'') = \phi(u+u'').$$

Hence hf is ϕ-homogeneous. It is clear from the definition that $f+g$ is homogeneous since all terms have the same ϕ-degree. □

Lemma 5.3.7 *Let $f, g \in k[x_1, \ldots, x_n] \setminus \{0\}$ be ϕ-homogeneous and $\prec$ a term ordering. Then the S-polynomial $S_\prec(f, g)$ is ϕ-homogeneous.*

Proof. The S-polynomial was defined as

$$S_\prec(f,g) = \frac{\operatorname{lcm}(\operatorname{in}_\prec(f), \operatorname{in}_\prec(g))}{\operatorname{in}_\prec(f)} f - \frac{\operatorname{lcm}(\operatorname{in}_\prec(f), \operatorname{in}_\prec(g))}{\operatorname{in}_\prec(g)} g$$

to carefully make two terms cancel - one from each summand. Each summand is ϕ-homogeneous by Lemma 5.3.6. Since two terms cancel they must have the same ϕ-degree. Therefore all terms of the S-polynomial have the same ϕ-degree. □

Lemma 5.3.8 *Let f and $f_1, \ldots, f_s$ be ϕ-homogeneous polynomials. The remainder r produced by the division algorithm (Algorithm 1.5.1) is ϕ-homogeneous.*

Proof. In the division algorithm f is assigned to p and p is adjusted until it eventually becomes zero and the algorithm terminates. At the beginning p is ϕ-homogeneous because f is. In each iteration the p remains ϕ-homogeneous because we only subtract ϕ-homogeneous polynomials from p of the same ϕ-degree. Therefore p remains ϕ-homogeneous of the same degree until it becomes 0. Terms are moved from p to the remainder r. Therefore the $r = 0$ when the algorithm terminates. □

Proposition 5.3.9 *Let ϕ be a grading on $k[x_1, \ldots, x_n]$. Let $I \subseteq k[x_1, \ldots, x_n]$ be a ϕ-homogeneous ideal and $\prec$ a term ordering. Then the reduced Gröbner basis $\mathcal{G}_\prec(I)$ is ϕ-homogeneous.*

Proof. By Hilbert's Basis Theorem 1.1.6 there exists a finite generating set $G \subseteq k[x_1, \ldots, x_n]$ for I. We also know that there exists a generating set $G' \subseteq k[x_1, \ldots, x_n]$ of ϕ-homogeneous polynomials since I is ϕ-homogeneous, which could be infinite. We now take every $g \in G$ and express it in terms of elements of G':

$$g = \sum_{i=1}^{m} f_i g_i$$

for some $m \in \mathbb{N}$, $f_i \in k[x_1, \ldots, x_n]$ and $g_i \in G'$. Doing so for all finitely many terms in g requires only a finite number of terms of the type g_i'. We let G'' denote this finite set of ϕ-homogeneous polynomials which generate I.

If we perform Buchberger's Algorithm on G'', the result is ϕ-homogeneous because the operations of taking S-polynomials and remainder preserves homogeneity (Lemma 5.3.7 and Lemma 5.3.8. The minimizing and autoreducing algorithms (Algorithm 1.7.8 and Algorithm 1.7.9) also preserve ϕ-homogeneity and together produce the unique reduced Gröbner basis of I with respect to $\prec$. We conclude that this Gröbner basis is ϕ-homogeneous. □

Knowing that the reduced Gröbner basis is always homogeneous can be used to compute a homogeneous generating set for a homogeneous ideal but also to check if and ideal is homogeneous.

Algorithm 5.3.10
Input: *A set $G \subseteq k[x_1, \dots, x_n]$ and a grading ϕ.*
Output: *"Yes" if the ideal $\langle G \rangle$ is ϕ-homogeneous and "No" otherwise.*

- *Compute the reduced Gröbner basis $H := \mathcal{G}_{\prec}(\langle G \rangle)$.*
- *Return "Yes" if all polynomials in H are ϕ-homogeneous and "No" otherwise.*

Example 5.3.11 The ideal $\langle x^2 + y - 3zx, y^2z + zx - 1 \rangle \subseteq \mathbb{Q}[x, y, z]$ is not homogeneous in the standard grading because its reduced Gröbner basis with respect to the lexicographic ordering (with $z \prec y \prec x$) is not homogeneous:

$$\{y^4z^2 + 1 - 3z^2 + yz^2 - 2y^2z + 3y^2z^3, x - 3z + yz - y^2 + 3y^2z^2 + y^4z\}.$$

Definition 5.3.12 By a *ϕ-homogeneous part* of a polynomial f we mean the sum of all terms in f of a particular ϕ-degree.

Lemma 5.3.13 *Let ϕ be a grading on $R := k[x_1, \dots, x_n]$. Let $I \subseteq R$ be a ϕ-homogeneous ideal. A polynomial $f \in I$ if and only if every ϕ-homogeneous part of f is in I.*

Proof. Clearly, if every ϕ-homogeneous part of f is in I then so is the sum, which equals f. On the other hand, let $f \in I$ be a polynomial, let h be a ϕ-homogeneous part of f, and let $\mathcal{G}_{\prec}(I)$ be the reduced Gröbner basis of I with respect to some term ordering $\prec$. The division algorithm will produce an expression $f = \sum_i a_i f_i$ with a_1 being polynomials and $f_i \in \mathcal{G}_{\prec}(I)$. By splitting each a_i into terms and multiplying out we get an expression $f = \sum_j b_j g_j$ where b_j is a single term and $g_j \in \mathcal{G}_{\prec}(I)$. Since each b_j is homogeneous, any homogeneous part of f is written as $f = \sum_{j \in J} b_j g_j$ for J chosen to give just the terms in the right ϕ-degree. Since $g_j \in I$ we conclude that the ϕ-homogeneous part is in I. □

Proposition 5.3.14 *Let $R := k[x_1, \dots, x_n]$ and $\phi : \mathbb{N}^n \to S$ a grading. As a k-vector space we may write R as a direct sum:*

$$R = \bigoplus_{m \in S} R_m$$

where R_m is the k-vector space of homogeneous polynomials of ϕ-degree m (together with 0). Furthermore, if I is a ϕ-homogeneous ideal then we can define I_m to be the set of ϕ-homogeneous polynomials in I of degree m (together with 0). As a vector space we have

$$I = \bigoplus_{m \in S} I_m.$$

Proof. Clearly, a polynomial can be uniquely be split into finitely many nonzero homogeneous parts of different ϕ-degree, which proves $R = \bigoplus_{m \in S} R_m$. A polynomial $f \in I$ is also in R and therefore splits into ϕ-homogeneous parts. Each of these is in I by Corollary 5.3.13 and therefore in an I_m. □

5.4 Hilbert functions

For an ideal I homogeneous with respect to a grading $\phi : \mathbb{N}^n \to S$ we would like to define the function:

$$H^I_\phi(m) := \dim_k(R_m/I_m)$$

with R_m and I_m defined as in Proposition 5.3.14. One problem is that this dimension might not be finite dimensional. We will therefore restrict to the case where I is A-homogeneous (Example 5.3.5) with $A \in \mathbb{R}^{d\times n}$ with a positive vector in its rowspace.

Lemma 5.4.1 *Let $I \subseteq k[x_1, \dots, x_n]$ be an A-homogeneous ideal with $A \in \mathbb{R}^{d\times n}$. Suppose that A has a positive vector in its rowspace. Then R_m is a finite dimensional k-vector space.*

Proof. Let $\omega = y^t A$ be a positive vector in the rowspace, with $y \in \mathbb{R}^d$, and let $b \in \mathbb{R}$ be the degree. There is only a finite number of monomials with ω-degree b. Namely, if the ith entry an exponent vector v is bigger than $\frac{b}{\omega_i}$ then the ω-degree of x^v must be bigger than b.

Now, let $b := y^t m$. If a monomial x^v has A-degree m, then $Av = m$, implying $yAv = y^t m$ and therefore x^v has ω-degree $y^t m$. Since there are only finitely many x^u of that ω-degree, there are only finitely many x^u of A-degree m. We conclude that R_m is a finite dimensional vector space. □

Definition 5.4.2 Let $I \subseteq k[x_1, \dots, x_n]$ be an A-homogeneous ideal with $A \in \mathbb{R}^{d\times n}$. Suppose that A has a positive vector in its rowspace. We define the A-graded Hilbert function $H^I_A : \mathbb{R}^d \to \mathbb{N}$ of I as follows:

$$H^I_A(m) := \dim_k(R_m/I_m).$$

The function is well defined since R_m is finite dimensional and I_m is a k-subspace of R_m.

In Lemma 1.6.10 we saw that the cosets of standard monomials form a k-vector space basis of R/I. From the direct sum of Proposition 5.3.14 we get that the cosets of the standard monomials of A-degree m form a k-vector space basis of R_m/I_m.

Lemma 5.4.3 *Let I and A be as in Definition 5.4.2. Let $\prec$ be a term ordering. Then*

$$H^I_A(m) = H^{\mathrm{in}_\prec(I)}_A(m).$$

Proof. To find $H^I_A(m)$ we count the monomials in $\mathrm{std}_\prec(I)$ of degree m. To find $H^{\mathrm{in}_\prec(I)}_A(m)$ we count the monomials in $\mathrm{std}_\prec(\mathrm{in}_\prec(I))$ of degree m. But these are the same since $\mathrm{in}_\prec(I) = \mathrm{in}_\prec(\mathrm{in}_\prec(I))$. (The initial ideal of a monomial ideal is the ideal itself). □

5.5 Homogeneity implies completeness

Proposition 5.5.1 *Let $u, v \in \mathbb{R}^n$ and let I be u-homogeneous polynomial ideal. Then for any $\lambda \in \mathbb{R}$ we have*

$$\text{in}_v(I) = \text{in}_{v+\lambda u}(I).$$

Proof. We only need to show the inclusion $\subseteq$ since the other would then follow by making suitable choices of u, v and λ. Let $f \in I$ and decompose it into its u-homogeneous parts: $f = \sum_i f_i$. By Lemma 5.3.13 $f_i \in I$ since I is u-homogeneous. Since there is no cancellation we get $\text{in}_v(f) = \sum_{i \in J} \text{in}_v(f_i)$ for a suitable subset of indices J. For every i, f_i is u-homogeneous, implying $\text{in}_v(f_i) = \text{in}_{v+\lambda u}(f_i) \in \text{in}_{v+\lambda u}(I)$. Therefore $\text{in}_v(f) = \sum_{i \in J} \text{in}_v(f_i) \in \text{in}_{v+\lambda u}(I)$ and we have proved the inclusion $\subseteq$ since the left hand side is generated by initial forms of the form $\text{in}_v(f)$ with $f \in I$. □

Corollary 5.5.2 *Let $\omega \in \mathbb{R}^n$ and $I \subseteq k[x_1, \ldots, x_n]$ be a ω-homogeneous ideal. The Gröbner cones of the form $C_u(I)$ and $C_\prec(I)$ are invariant under translation by ω.*

Proof. A set S is translation invariant under translation by ω iff $S = \{x+\omega : x \in S\}$. By Proposition 5.5.1 the equivalence classes of the relation in Equation 1 are translation invariant. Therefore the closures, $C_u(I)$, are also translation invariant. By Proposition 4.2.3, every cone of the form $C_\prec(I)$ is of the form $C_u(I)$ for some $u \in \mathbb{R}^n$. □

Proposition 5.5.3 *Let $\omega \in \mathbb{R}^n_{>0}$ and $I \subseteq k[x_1, \ldots, x_n]$ be an ω-homogeneous ideal. Then* Gfan(I) *is complete.*

Proof. Let u be any vector in $\mathbb{R}^n$. Choosing $\lambda \in \mathbb{R}$ sufficiently large, the vector $u + \lambda\omega$ becomes positive. Let $\prec$ be a term ordering. By Corollary 4.2.4, the vector $u + \lambda\omega$ is in the Gröbner cone $C_{\prec_{u+\lambda\omega}}(I)$. By the same corollary this Gröbner cone is translation invariant under ω (because the g mentioned in the corollary is ω-homogeneous by Proposition 5.3.9). Therefore, $u \in C_{\prec_{u+\lambda\omega}}(I)$. We conclude that every vector $u \in \mathbb{R}^n$ is in some Gröbner cone of the Gröbner fan of I and that the fan therefore must be complete. □

Remark 5.5.4 Ideals that are homogeneous in gradings induced by positive vectors are nice for several reasons:

- Their Gröbner fans are complete.
- Their Gröbner fans are normal fans of polytopes. **(We will not prove this.)**
- We may compute Gröbner bases with respect to orderings that do not satisfy the 1-is-smallest-property. Consider the "reverse lexicographic" ordering $\prec_{rlex}$, where for $a, b \in \mathbb{N}^n$ we let $x^a \prec_{rlex} x^b \Leftrightarrow \exists j : a_j > b_j \wedge a_{j+1} = b_{j+1} \wedge \cdots \wedge a_n = b_n$ which is not a term ordering. Suppose that we have an ω-homogeneous ideal I where $\omega \in \mathbb{R}^n_{>0}$ and an ω-homogeneous

generating set G for I. Then $\prec_{rlex_\omega}$ agree picks out the same initial terms as $\prec_{rlex}$ from polynomials in $\mathcal{G}$. When running Buchberger's algorithm on G all polynomials that appear will be ω-homogeneous (proof of Proposition 5.3.9) and therefore $\prec_{rlex_\omega}$ and $\prec_{rlex}$ pick the same initial terms.

- They define varieties in "projective space". See the explanation in the next section.

5.6 Homogenisation

Since homogeneous ideals have some nice properties, we my want make non-homogeneous ideals homogeneous by introducing a new variable. This idea is also used when moving a variety to *projective space*. We give a simple example for the reader who is unfamiliar with the notion of projective space:

Example 5.6.1 Every line in $\mathbb{R}^2$ is of the form $V(\langle f\rangle)$ where $f = ax+by+c \in \mathbb{R}[x,y]$ and with a and b are in $\mathbb{R}$, not both 0. Usually two lines will have a unique intersection point. But parallel lines will not. We fix this "problem" by introducing a *homogenising variable* t. Instead of $V(\langle f\rangle)$ we consider the variety of the homogeneous ideal $V(\langle f^h\rangle) \in \mathbb{R}^3$ where $f^h := ax + by + ct$. We want to identify points (x, y, t) with (x', y', t') if there exists $s \neq 0$ such that $s(x, y, t) = (x', y', t')$. In this case we write $(x, y, t) \sim (x', y', t')$. We check that (x, y, t) satisfies the equation if and only if $s(x, y, t)$ does:

$$f^h(sx, sy, st) = asx + bsy + zst = s(ax + by + cz) = sf^h(x, y, z).$$

The *projective plane* is now defined as $\mathbb{P}^2 := (\mathbb{R}^3 \setminus (0,0,0))/\sim$. Whether f^h is 0 in a point or not only depends on the equivalence class. We therefore define the *projective variety* defined by f^h to be the set of equivalence classes where f is 0. We get the original picture of $\mathbb{R}^2$ back by intersecting $\mathbb{R}^3$ with $H := \mathbb{R}^2 \times \{1\}$. The equivalence classes of points with last coordinate 0 do not intersect H. These are called *points at infinity.* In the projective space $\mathbb{P}^2$ any two different lines have exactly one common point.

We have earlier used the notion of Laurent monomials. By the Laurent polynomial ring we mean the semigroup ring $k[\mathbb{Z}^n]$. That is, polynomials where exponents may be negative.

Definition 5.6.2 Let $f \in k[x_1, \dots, x_n]$ be a polynomial. We define the *homogenisation* of f to be $f^h := x_0^d f(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0})$ where d is the maximal total degree of an term in f. If $f = 0$ we define $f^h = 0$. Let $I \subseteq k[x_1, \dots, x_n]$ be an ideal. The homogenisation of I is

$$I^h := \langle f^h : f \in I\rangle \subseteq k[x_0, \dots, x_n].$$

It is clear that $(p_1, \dots, p_n) \in k^n$ is in $V(I)$ if and only if $(1, p_1, \dots, p_n)$ is in $V(I^h)$. But as we have seen in Example 5.6.1, there may also be points in "at infinity" with the additional coordinate being 0. In the definition of I^h it is important to consider all elements in I.

Example 5.6.3 Let $I = \langle x+y+1, x+y\rangle \subseteq \mathbb{Q}[x,y]$. Each of the two generators defines a line in the plane $\mathbb{R}^2$. These two lines are parallel and do not intersect. We notice that $1 \in I$ and therefore $1^h = 1 \in I^h$. If we just homogenise the generators of I we get $\langle x+y+t, x+y\rangle$, but this ideal does not contain 1. (To see this compute a Gröbner basis or observe that all terms in this ideal must have degree at least 1.) Hence we cannot get I^h just by homogenizing the generators.

The following theorem gives an algorithm for computing the homogenization.

Proposition 5.6.4 *[4, Theorem 4, page 397] Let $I \subseteq k[x_1,\dots,x_n]$ be an ideal. Let $\omega = (1,\dots,1) \in \mathbb{N}^n$ and let $\prec$ be a term ordering. Then $\{g^h : g \in \mathcal{G}_{\prec_\omega}(I)\}$ generates I^h.*

Proof. To prove $\langle g^h : g \in \mathcal{G}_{\prec_\omega}(I)\rangle = I^h$ we first observe that the inclusion $\subseteq$ is clear. Now, consider a generator f^h of I^h where $f \in I$. Using the division algorithm on f modulo $\mathcal{G}_{\prec_\omega}(I)$ we can write $f = \sum_i a_i g_i$ where a_i is a term, $g_i \in \mathcal{G}_{\prec_\omega}(I)$ and the total degree of each $a_i g_i$ is less than or equal to the total degree of f. Because of this we get $f^h = \sum_i a_i x_0^{d_i} g_i^h$ where $d_i = \deg(f) - \deg(g_i)$. This shows that f^h is in the left hand side. □

Example 5.6.5 Example 5.6.3 continued. Applying the proposition, the reduced Gröbner basis of $\mathcal{G}_{\prec_\omega}(I)$ becomes $\{1\}$. We now homogenise and get $\{1\}$ as a generating set for I^h.

Proposition 5.6.6 *[9, Proposition 5.2.3] Let $I \subseteq k[x_1,\dots,x_n]$ be an ideal, $\omega \in \mathbb{R}^n$ a vector and $\prec$ a term ordering on $k[x_1,\dots,x_n]$. Let $\prec'$ be the termordering on $k[x_0,\dots,x_n]$ defined by $x^u \prec' x^{u'}$ if and only if*

$$\sum_i u_i < \sum_i v_i \vee (\sum_i u_i = \sum_i v_i \wedge x^u_{|x_0=1} \prec x^v_{|x_0=1}).$$

The set $\mathcal{G} := \{\mathrm{in}_\omega(g_{|x_0=1}) : g \in \mathcal{G}_{\prec'_{(0,\omega)}}(I^h)\}$ is a Gröbner basis for $\mathrm{in}_\omega(I)$ with respect to $\prec$.

Notice that $\prec'_{(0,\omega)}$ might not be a term ordering. But, as explained in Remark 5.5.4, this is not a problem since I^h is homogeneous in the total grading, and we can talk about $\mathcal{G}_{\prec'_{(0,\omega)}}(I^h)$ by considering a term order which agrees with $\prec'_{(0,\omega)}$ on monomials of the same total degree.

Proof. It is straight forward to prove the containment $\mathcal{G} \subseteq \mathrm{in}_\omega(I)$. It remains to be proved that $\mathrm{in}_\prec(\mathrm{in}_\omega(I)) \subseteq \langle \mathrm{in}_\prec(g) : g \in \mathcal{G}\rangle$. The left hand side is generated by elements of the form $m = \mathrm{in}_\prec(\sum_i \mathrm{in}_\omega(f_i))$ where $f_i \in I$. We will show that any such m is on the right hand side. Without loss of generality we may assume that the f_i's have the same ω-degree as m. Hence $m = \mathrm{in}_\prec(\mathrm{in}_\omega \sum_i f_i)$. Let $f = \sum_i f_i \in I$. Then $f^h \in I^h$ and the initial term $\mathrm{in}_{\prec'(0,\omega)} f^h$ must be divisible by the initial term $\mathrm{in}_{\prec'_{(0,\omega)}}(g)$ of some Gröbner basis element $g \in$

$\mathcal{G}_{\prec'_{(0,\omega)}}(I^h)$. Consequently, $(\mathrm{in}_{\prec'_{(0,\omega)}}(g))_{|x_0=1}$ divides $(\mathrm{in}_{\prec'(0,\omega)} f^h)_{|x_0=1}$. Observe that $(\mathrm{in}_{\prec'_{(0,\omega)}}(g))_{|x_0=1} = \mathrm{in}_{\prec_\omega}(g_{|x_0=1}) = \mathrm{in}_{\prec}(\mathrm{in}_\omega(g_{|x_0=1}))$ since g is homogeneous. Similarly, $(\mathrm{in}_{\prec'_{(0,\omega)}}(f^h))_{|x_0=1} = \mathrm{in}_{\prec_\omega}(f^h_{|x_0=1}) = \mathrm{in}_{\prec}(\mathrm{in}_\omega(f))$. This proves that $\mathrm{in}_{\prec}(\mathrm{in}_\omega(g_{|x_0=1}))$ divides $m = \mathrm{in}_{\prec}(\mathrm{in}_\omega(f))$ as desired. □

We strengthen Proposition 4.4.8:

Corollary 5.6.7 *Every ideal $I \subseteq k[x_1, \dots, x_n]$ has only finitely many initial ideals of the form $\mathrm{in}_\omega(I)$ where $\omega \in \mathbb{R}^n$.*

Proof. By Proposition 4.1.1 the ideal I^h has only a finite number of initial ideals of form $\mathrm{in}_{\prec}(I^h)$. Lemma 4.4.7 shows that there can be at most one reduced Gröbner basis for each such $\mathrm{in}_{\prec}(I^h)$. Hence I^h has only finitely many reduced Gröbner bases. By Proposition 5.6.6 generators for initial ideals of I can be obtained by taking initial forms of elements of these reduced Gröbner bases after dehomogenisation. For each reduced Gröbner basis there is only a finite number of such ways that ω can pick generators for the initial ideal $\mathrm{in}_\omega(I)$. Hence there are only finitely many initial ideals of I. □

5.7 Links in Gröbner fans

Recall that $N_P(F)$ means the (outer) normal cone of a polyhedron P at the face F. In the following definition we implicitly use Proposition 4.5.2.

Definition 5.7.1 Let $P \subseteq \mathbb{R}^n$ be a polyhedron and $\omega \in P$. We define the *tangent cone* at ω to be dual cone $\mathrm{link}_\omega(P) := N_P(F)^\vee$ where F is the face of P containing ω in its relative interior. Since this does not depend on ω but only on P.

Why we chose the weird notation $\mathrm{link}_\omega(P)$ should become clear soon.

Example 5.7.2 Let $P = \mathrm{conv}((0,0),(0,1),(1,0),(1,1)) \subseteq \mathbb{R}^2$. The tangent cone at $(1,1)$ is the negative orthant $\mathbb{R}^2_{\leq 0}$.

Lemma 5.7.3 *Let $P \subseteq \mathbb{R}^n$ be a polyhedron and $\omega \in P$. Then $u \in \mathrm{link}_\omega(P)$ if and only if $\omega + \varepsilon u \in P$ for all $\varepsilon > 0$ sufficiently small. Furthermore, for $\varepsilon > 0$ sufficiently small*

$$\mathrm{link}_u(\mathrm{link}_v(P)) = \mathrm{link}_{v+\varepsilon u}(P).$$

Proof. Left to the reader. □

Example 5.7.4 Let $P = \mathrm{conv}((0,0),(0,1),(1,0),(1,1)) \subseteq \mathbb{R}^2$. Then

$$\mathrm{link}_{(\frac{9}{10},1)}(P) = \mathbb{R} \times \mathbb{R}_{\leq 0}$$

and

$$\mathrm{link}_{(-1,0)}(\mathrm{link}_{(1,1)}(P)) = \mathrm{link}_{(-1,0)}(\mathbb{R}^2_{\leq 0}) = \mathbb{R} \times \mathbb{R}_{\leq 0}.$$

This is also what the lemma states for $v = (1,1)$, $u = (-1,0)$ and $\varepsilon = \frac{1}{10}$.

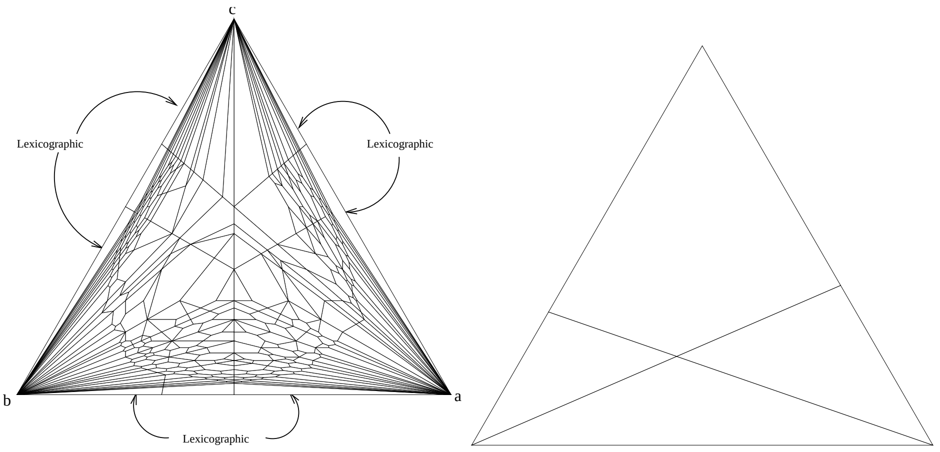

Figure 11: The intersection of the Gröbner fan of the ideal of Example 4.0.22 with the triangle $\text{conv}\{(1,0,0),(0,1,0),(0,0,1)\}$. The intersection of the link at the point $(3,4,2)$.

Definition 5.7.5 Let $\mathcal{F}$ be a polyhedral complex and $\omega \in \text{supp}(\mathcal{F})$. We define the link of $\mathcal{F}$ at ω:

$$\text{link}_\omega(F) = \{\text{link}_\omega(P) : \omega \in P \in \mathcal{F}\}.$$

Again, the link does not depend on ω but only the face containing ω in its relative interior.

Lemma 5.7.6 *The link of a polyhedral complex at a point is a polyhedral fan.*

We will not prove this lemma, but rather see that it is always true for Gröbner fans.

Example 5.7.7 The link of a fan in a point is shown in Figure 9.

Proposition 5.7.8 *[13, Proposition 1.13] Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $u, v \in \mathbb{R}^n$. Suppose that I is homogeneous or $u \in \mathbb{R}^n_{>0}$. Then for $\varepsilon > 0$ sufficiently small*

$$\text{in}_{u+\varepsilon v}(I) = \text{in}_v(\text{in}_u(I)).$$

Proof. Let $\prec$ be a term ordering. We claim that $u+\varepsilon v \in C_{(\prec_v)_u}(I)$. Notice that $(\prec_v)_u$ might not be a term ordering, but by our discussion in Remark 5.5.4 this is not a problem because $(\prec_v)_u$ agrees with some term ordering on a homogeneous generating set for I. We will use Corollary 4.2.4 to show $u + \varepsilon v \in C_{(\prec_v)_u}(I)$. Let $g \in \mathcal{G}_{(\prec_v)_u}(I)$. It suffices to prove that $\text{in}_{(\prec_v)_u}(g) = \text{in}_{(\prec_v)_u}(\text{in}_{u+\varepsilon v}(g))$. But this follows from $\text{in}_{(\prec_v)_u}(\text{in}_{u+\varepsilon v}(g)) = \text{in}_{(\prec_v)_u}(\text{in}_v(\text{in}_u(g)))$ for $\varepsilon > 0$ sufficiently small and $\text{in}_{(\prec_v)_u}(\text{in}_v(\text{in}_u(g))) = \text{in}_{\prec_v}(\text{in}_v(\text{in}_u(g))) = \text{in}_{\prec_v}(\text{in}_u(g)) = \text{in}_{(\prec_v)_u}(g)$. We apply Corollary 4.4.4 which says

$$\text{in}_{u+\varepsilon v}(I) = \langle \text{in}_{u+\varepsilon v}(f) : f \in \mathcal{G}_{(\prec_v)_u}(I)\rangle = \langle \text{in}_v(\text{in}_u(f)) : f \in \mathcal{G}_{(\prec_v)_u}(I)\rangle$$

$$= \langle \mathrm{in}_v(g) : g \in \mathcal{G}_{\prec_v}(\mathrm{in}_u(I))\rangle = \mathrm{in}_v(\mathrm{in}_u(I)).$$

Here the second equality is true when ε is sufficiently small and the third is obtained by applying Corollary 4.4.4 a second time using $u \in C_{(\prec_v)_u}(I)$. The last equality again follows from Corollary 4.4.4 using $v \in C_{\prec_v}(\mathrm{in}_u(I))$. □

Example 5.7.9 The following is a reduced Gröbner basis for the initial ideal $\mathrm{in}_{(3,4,2)}(I)$ of I of Example 4.0.22

$$\{c^7, bc^5, b^2, ac^6, abc^3 - \frac{1850}{19281}ac^5, a^2c^4, a^2bc^2, a^3c^2 - \frac{980}{19281}ac^5, a^3b - \frac{916}{19281}ac^5, a^4c, a^5\}$$

This was computed with Algorithm 4.4.5. The Gröbner fan of this ideal equals the link at the point $(3, 4, 2)$ of the Gröbner fan of I. It is shown in Figure 11.

Corollary 5.7.10 *Let $I \subseteq k[x_1, \dots, x_n]$ be an ideal and let $u \in \mathbb{R}^n_{>0}$. Then*

$$\mathrm{link}_u(\mathrm{Gfan}(I)) = \mathrm{Gfan}(\mathrm{in}_u(I)).$$

5.8 "Very homogeneous" ideals

Clearly, the initial ideal $\mathrm{in}_\omega(I)$ is ω-homogeneous, but if ω comes from a cone in the Gröbner fan which is not just a ray, the ideal would me homogeneous with respect to many more vectors.

Definition 5.8.1 Let $I \subseteq k[x_1, \dots, x_n]$ be an ideal. We call the set $\{\omega \in \mathbb{R}^n : \mathrm{in}_\omega(I) = I\}$ the *homogeneity space* of I.

Lemma 5.8.2 *The homogeneity space of an ideal $I \subseteq k[x_1, \dots, x_n]$ is a linear subspace of $\mathbb{R}^n$.*

Proof. We wish to apply Proposition 4.4.9. We choose $v = 0 \in \mathbb{R}^n$ and $\prec$ to be the lexicographic term ordering. Now the proposition tells us that $u \in \mathbb{R}^n$ is in the homogeneity space of I if and only if $\forall g \in \mathcal{G}_\prec(I) : \mathrm{in}_u(g) = g$. This is equivalent to saying that all terms of g have the same u-degree. This translates into a set of linear condition of u that must be satisfied. Therefore the homogeneity space is a subspace of $\mathbb{R}^n$. □

Since the homogeneity space is a linear subspace, it equals its closure. Therefore $C_0(I) = \overline{\{\omega \in \mathbb{R}^n : \mathrm{in}_\omega(I) = \mathrm{in}_0(I)\}} = \overline{\{\omega \in \mathbb{R}^n : \mathrm{in}_\omega(I) = I\}} = \{\omega \in \mathbb{R}^n : \mathrm{in}_\omega(I) = I\}$, which is exactly the homogeneity space. Therefore $C_0(I)$ is our notation for the homogeneity space of I.

Example 5.8.3 We wish to compute the homogeneity space of $I = \mathrm{in}_{(2,18,36)}(J)$, where J is the ideal in Example 4.0.22. We compute the following reduced Gröbner basis for the initial ideal $\{c^2, bc, b^2 + c, a^3c, a^9b, a^{18}\}$. By the argument of the lemma, the homogeneity space is all vectors which pick the same polynomials as initial forms. This translates just into the condition $\mathrm{in}_\omega(b^2+c) = b^2+c$. Which means $\omega \cdot (0, 2, 0)^t = \omega \cdot (0, 0, 1)^t$. Consequently the homogeneity space is the hyperplane passing through the origin with normal vector $(0, 2, -1)$.

In Definition 3.3.1 we defined the lineality space of a cone C. This is a face of C because (Proposition 3.3.12) it is the intersection of faces of C (every inequality $A_{i\cdot}$ gives rise to a face $\text{face}_{A_{i\cdot}}(C)$). By the lineality space of a fan we mean the intersection of all cones in the fan. This is the smallest cone in the fan. We notice that the lineality space of the Gröbner fan of an ideal I equals the homogeneity space of I. (Because the homogeneity space is a cone in the Gröbner fan and has no faces by Lemma 5.8.2.)

In the following we will be interested in ideals in $k[x_1, \ldots, x_n]$ with $n-1$-dimensional homogeneity space. Fix such an ideal I and call the homogeneity space L. Let g be an element of a reduced Gröbner basis of I. By Proposition 5.3.9 we know the g must be homogeneous in any grading given by a vector $\omega \in L$. Let cx^α and $c'x^\beta$ be two terms of g we conclude that $\omega \cdot \alpha = \omega \cdot \beta$ for all $\omega \in L$. That is $\alpha - \beta$ is in the orthogonal complement $L^\perp$. In other words the exponent vectors of g lie on a line, or equivalently, the Newton polytope of g is a line segment. We have proved the following lemma.

Lemma 5.8.4 *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal with a $n-1$-dimensional homogeneity space and $\prec$ a term ordering. The Newton polytope of any $g \in \mathcal{G}_\prec(I)$ is either a single point or a line segment. Furthermore, the line segments, as g runs through $\mathcal{G}_\prec(I)$, are parallel.*

Fix a generator $v \in \mathbb{R}^n$ for $L^\perp$. We wish to argue that our "very homogeneous" ideal has at most two reduced Gr̈bner bases. Let G be one reduced Gröbner basis of I and suppose we want to compute $\mathcal{G}_\prec(I)$ with respect to some term ordering $\prec$. Only one of two things can happen: $\prec$ will pick the terms with exponent in direction v or in direction $-v$. As we observed earlier (proof of Proposition 5.3.9) all intermediate polynomials in a run of Buchberger's algorithm on G will also be homogeneous and therefore line segments (or points). We have proved the following Proposition.

Proposition 5.8.5 *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal with a $n-1$-dimensional homogeneity space. Then I has only one or two reduced Gröbner bases.*

Example 5.8.6 The ideal $\langle x - y \rangle \subseteq \mathbb{Q}[x, y]$ is homogeneous in the standard grading and has the reduced Gröbner bases $\{\underline{x} - y\}$ and $\{\underline{y} - x\}$.

Example 5.8.7 The ideal $I := \langle xy - 1 \rangle \subseteq \mathbb{Q}[x, y]$ is homogeneous in the grading induced by the vector $(1, -1)$. The homogeneity space of I is $\text{span}\{(1, -1)\}$. The ideal has only on reduced Gröbner basis because xy has to be larger than -1 in every term ordering.

Let $A \in \mathbb{R}^{d \times n}$ be a matrix whose rows form a basis of the lineality space of I. Let's assume that the rowspace contains a positive vector. This matrix gives rise to an A-grading as in Section 5.4. Returning to our ideal I, we notice by Lemma 5.4.3 that its A-graded Hilbert function equals that of $\text{in}_\prec(I)$ for any $\prec$. Therefore, the two initial ideals of I have the same Hilbert functions.

Our final observation in this subsection is that the Hilbert function

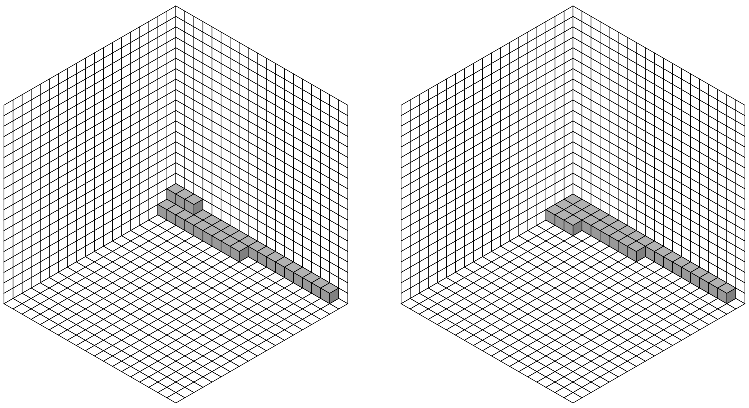

Figure 12: The two staircase diagrams of the initial ideals of Example 5.8.8.

Example 5.8.8 The "very homogeneous" ideal I in Example 5.8.3 has two reduced Gröbner bases:

$$\{c^2, bc, b^2+c, a^3c, a^9b, a^{18}\}$$

and

$$\{c+b^2, b^3, a^3b^2, a^9b, a^{18}\}$$

The corresponding staircase diagrams are shown in Figure 12. Let's pick $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}$ whose rowspace is the homogeneity space of I. The monomials of A-degree $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ are $\{a^2c, a^2b^2\}$. Looking at the first Gröbner basis w.r.t. $\prec$ (and the corresponding initial ideal $\text{in}_\prec(I) = \langle c^2, bc, b^2, a^3c, a^9b, a^{18}\rangle$) we get $H_A^I(\begin{bmatrix} 2 \\ 2 \end{bmatrix}) = H_A^{\text{in}_\prec(I)}(\begin{bmatrix} 2 \\ 2 \end{bmatrix}) = 2 - 1 = 1$ because there are two monomials of this A-degree but one is in the initial ideal. For the A-degree $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ we get the monomials $\{a^2c^2, a^2b^2a^2b^2c, a^2b^4\}$. But here all monomials are in the initial ideal(s) so the Hilbert function value is $3 - 3 = 0$.

5.9 The Gröbner walk

Our observations from the previous subsection can be used to convert Gröbner bases with respect to some ordering $\prec$ into a Gröbner basis with respect to some other ordering $\prec'$. To get a geometric sense of what is going on we pick a vector $\omega \in C_\prec(I)$ and let $A \in \mathbb{R}^{d \times n}$ be a matrix representation of $\prec'$. For $\varepsilon > 0$ very small $a_\varepsilon := A^t(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{d-1})^t$ is going to be in $C_{\prec'}(I)$. (See the argument in the proof of Proposition 4.2.3.) Pick a small $\varepsilon > 0$ such that this is the case. We now consider the line segment between ω and A_ε. This line segment passes through a set of full-dimensional Gröbner cones. We wish to compute the reduced Gröbner bases for these.

We now discuss how to come from one cone to the next. Suppose that we know a Gröbner basis $\mathcal{G}_{\prec}(I)$. Let $C = C_{\prec}(I)$ and suppose that the line segment from A_{ε} passes through a facet F of C. This facet F is also a facet of a different cone $C' \in$ Gfan. Let $N \in \mathbb{R}^n$ be the normal pointing in direction C'. This facet is in the Gröbner fan of I and it contains some ω in its relative interior. Let's assume that ω is positive. (If it is not positive to choose ω positive, then it is because F lies outside $\mathbb{R}^n_{>0}$ and we are not sure that there is a Gröbner cone on the other side of F.) For $\varepsilon' > 0$ sufficiently small the vector $\omega + \varepsilon' N$ is in C'. Our (sub)goal is to compute $\mathcal{G}_{\prec'_{\omega+\varepsilon' N}}(I)$.

We first notice that $\text{in}_{\prec'_{\omega+\varepsilon' N}}(I) = \text{in}_{\prec'}(\text{in}_{\omega+\varepsilon' N}(I)) = \text{in}_{\prec'}(\text{in}_N(\text{in}_\omega(I))) = \text{in}_{\prec'_N}(\text{in}_\omega(I))$. Using Algorithm 4.4.5 we compute:

$$\mathcal{G}_{\prec}(\text{in}_\omega(I)) = \{\text{in}_\omega(g) : g \in \mathcal{G}_{\prec}(I)\}$$

Because ω is positive, $\text{in}_\omega(I)$ has two reduced Gröbner bases as explained in Section 5.8. To compute the other we use Buchberger's Algorithm 1.7.3 with a term ordering $\prec'_N$ induced by N. Taking the initial terms of the computed Gröbner basis we get generators for the initial ideal $\text{in}_{\prec'_{\omega+\varepsilon' N}}(I)$. We are almost there - we know the initial terms of the elements in the Gröbner basis. We just need to find their tail.

To make things clear we present an example:

Example 5.9.1 Let $I = \langle x^2 - y, z^2 - xy + 2\rangle \subseteq \mathbb{Q}[x, y, z]$. We have the reduced Gröbner basis

$$\mathcal{G}_{\prec}(I) = \{y^2 - 2x - xz^2, xy - 2 - z^2, x^2 - y\}$$

The Gröbner cone $C_{\prec}(I)$ equals cone$((0, 0, -1), (2, 1, 0), (2, 4, 3))$. See Figure 13. An interior point of the Gröbner cone is $(5, 7, 3)$. We choose $F =$ cone$((2, 1, 0), (2, 4, 3))$. A normal vector for F is $N = (1, -2, 2)$. The vector $\omega = (4, 5, 3)$ is in the relative interior of F. We get that

$$\mathcal{G}_{\prec}(\text{in}_\omega(I)) = \{y^2 - xz^2, xy, x^2\}$$

The other reduced Gröbner basis of $\text{in}_\omega(I)$ is:

$$\{\underline{y^3}, \underline{xz^2} - y^2, \underline{xy}, \underline{x^2}\}$$

Hence we need to find four polynomials in I which have the above underlined initial terms with respect to

The following algorithm is useful:

Algorithm 5.9.2
Input: *A reduced Gröbner basis $\mathcal{G}_{\prec''}(I)$, a vector $\omega \in C_{\prec''}(I)$ and an ω-homogeneous polynomial $h \in \text{in}_\omega(I) \setminus \{0\}$.*
Output: *A polynomial $f \in I$ such that $\text{in}_\omega(f) = h$.*

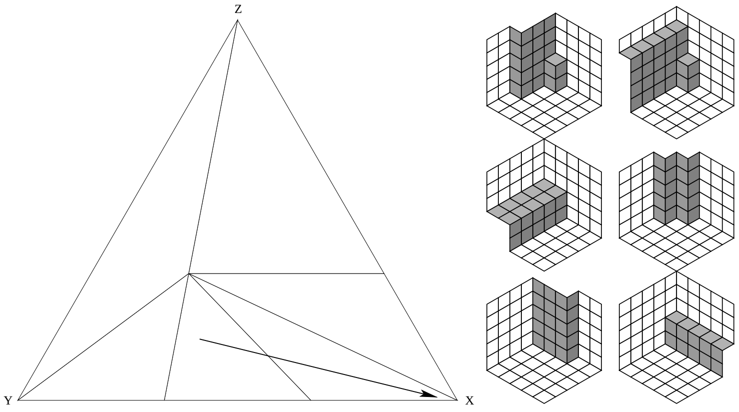

Figure 13: In Example 5.9.1 we walk from the point $(5, 7, 3)$ towards $(1, \varepsilon, \varepsilon^2)$. We need to cross just a single facet. The picture shows the part of the Gröbner fan which is in the positive orthant. On the right all the staircase diagrams of the monomial initial ideals are shown. In the example we walk from picture number 4 (center, right) to number 1 (top, left). The standard monomials are shuffled around. The top part of one pile in picture number 4 is translated along the vector $-N = (-1, 2, -2)$ to get picture number 1.

- *$f := h - r$, where r is the remainder of division of h with $\mathcal{G}_{\prec''}(I)$ using the term order $\prec''$.*

Proof. We first notice that the division algorithm will give remainder 0 if run on h and $\{\text{in}_\omega(g) : g \in \mathcal{G}_{\prec''}(I)\}$ because this set is a Gröbner basis. The division algorithm finds $a_i \in k[x_1, \ldots, x_n]$ and $g_i \in \mathcal{G}_{\prec''}(I)$ such that:

$$h = 0 + \sum_i a_i \text{in}_\omega(g_i),$$

Reducing h with $G_{\prec''}(I)$ we can make the same choices which would reduce m to 0 if the elements of $\mathcal{G}_{\prec''}(I)$ had only consisted of their ω-initial forms. That is after the first steps we have

$$h = 0 + \sum_i a_i g_i +'' \text{lower } \omega - \text{degree terms}'',$$

where the lower degree terms have ω-degree less than that of h. We continue the division on the lower degree terms and get

$$h = r + \sum_i a_i g_i + \sum_j b_j f_j,$$

with $b_j \in k[x_1, \ldots, x_n]$ and $f_j \in \mathcal{G}_{\prec''}(I)$ and r is the remainder. For every $f \in \mathcal{G}_{\prec''}(I)$ we have $\text{in}_{\prec''}(f) = \text{in}_{\prec''}(\text{in}_\omega(f))$. That is, the initial term has maximal ω-degree, showing that the ω-degree (of p in Algorithm 1.5.1) cannot increase during the division. Therefore all terms of r and $\sum_j b_j f_j$ have lower ω-degree than h. Subtracting r on both sides we get $\text{in}_\omega(h - r) = \sum_i a_i \text{in}_\omega(g_i) = h$. □

Example 5.9.3 Continued. We apply the division algorithm to the four polynomials

- $y^3 \to y^3 - y(y^2 - 2x - xz^2) = 2xy + xyz^2 \to 2xy + xyz^2 - 2(xy - 2 - z^2) = xyz^2 + 4 + 2z^2 \to xyz^2 + 4 + 2z^2 - z^2(xy - 2 - z^2) = 4 + 4z^2 + z^4$
- $xz^2 - y^2 \to xz^2 - y^2 + (y^2 - 2x - xz^2) = -2x$
- $xy \to 2 + z^2$
- $x^2 \to y$

We now subtract from the original terms and get $\{\underline{y^3} - 4 - 4z^2 - z^4, \underline{xz^2} - y^2 + 2x, \underline{xy} - 2 - z^2, \underline{x^2} - y\}$ which we know is a Gröbner basis with respect to $\prec'_{\omega+\varepsilon' N}$. The final step in the algorithm is to autoreduce the Gröbner basis using Algorithm 1.7.9. In our example the Gröbner basis is already a reduced Gröbner basis.

We describe the complete algorithm for walking through a facet.

Algorithm 5.9.4
Input: *A reduced Gröbner basis $\mathcal{G}_\prec(I)$ of an ideal $I \subseteq k[x_1, \dots, x_n]$ and a facet F of $C_\prec(I)$, with F containing at least one positive vector. Finally, an outer normal $N \in \mathbb{R}^n$ such that $\text{face}_N(C) = F$.*
Output: *The reduced Gröbner basis $\mathcal{G}_{\prec'}(I)$ with $C_{\prec'}(I)$ being the other full-dimensional Gröbner cone having F as a facet.*

- *Let $\omega \in \mathbb{R}^n_{>0} \cap F$.*
- *Compute $\mathcal{G}_\prec(\text{in}_\omega(I)) = \{\text{in}_\omega(g) : g \in \mathcal{G}_\prec(I)\}$.*
- *Compute the other Gröbner basis $\mathcal{G}_{\prec_N}(\text{in}_\omega(I))$ using Buchberger's algorithm.*
- *For each $h \in \mathcal{G}_{\prec_N}(\text{in}_\omega(I))$ apply Algorithm 5.9.2. Store the computed set of polynomial in G.*
- *Autoreduce G and output the result which is the desired $\mathcal{G}_{\prec'}(I)$.*

We notice that it is never necessary to know $\prec'$ in the algorithm.

Sometimes walking through a single facet is not enough. Corollary 4.2.4 gives a way to find the inequalities for $C_{\prec'}(I)$. To find the facet to walk through, we find the first inequality which is violated when moving from a long the segment line from our starting vector $\omega \in C_\prec(I)$ towards our target a_ε. The process is repeated for the next cone until we find the cone containing a_ε. This is known as the Gröbner walk procedure. It is sometimes useful when we want to compute a Gröbner basis with respect to a difficult term ordering (such as the lexicographic) and know one for a "cheap" ordering (such as graded reverse lexicographic).

6 Toric ideals

We have seen in Example 5.8.3 how to compute the homogeneity space $C_0(I)$ of an ideal I. We can write a generating set of the homogeneity space as the rows of a $d \times n$ matrix. Then I will be A-homogeneous. In Section we will see how one can start with a matrix A and construct and A-homogeneous ideal. There are many such A-homogeneous ideals. Toric ideals is an interesting kind.

6.1 Saturation

This section is based on [8, Section 3.2].

Definition 6.1.1 Let R be a commutative ring and $I \subseteq R$ an ideal and $f \in R$. We define the *ideal quotients*

$$(I : f) = \{g \in R : gf \in I\} \text{ and}$$

$$(I : f^\infty) = \{g \in R : \exists n \in \mathbb{N} : gf^n \in I\}.$$

Notice that $(I : f) \supseteq I \subseteq (I : f^\infty)$. These sets are in fact ideals.

Proposition 6.1.2 *The sets $(I : f)$ and $(I : f^\infty)$ are ideals.*

Proof. We will only prove the case $(I : f^\infty)$. Let $g, g' \in (I : f^\infty)$. Then $gf^n \in I$ and $g'f^{n'} \in I$ for some n and n' in $\mathbb{N}$. This implies $gf^{n''}, g'f^{n''} \in I$ for $n'' = \max(n, n')$ and we conclude $(g + g')f^{n''}$ and $g + g' \in (I : f^\infty)$. Clearly multiplication by an element in $(I : f^\infty)$ gives a new element in $(I : f^\infty)$. □

The following lemma tells us how to compute ideal quotients in $k[x_1, \dots, x_n]$ with respect to one of the variables.

Proposition 6.1.3 *[13]Let $I \subseteq k[x_1, \dots, x_n]$ be a homogeneous ideal with respect to a grading induced by a positive vector $v \in \mathbb{R}^n_{>0}$. Let $\prec$ be a term ordering satisfying (for all v-homogeneous elements $f \in k[x_1, \dots, x_n]$):*

$$x_n | \mathrm{in}_\prec(f) \Rightarrow x_n | f.$$

If G is a Gröbner basis for $I \subseteq k[x_1, \dots, x_n]$ with respect to $\prec$ consisting of v-homogeneous elements then

$$G' := \{f \in G : x_n \not| f\} \cup \{f/x_n : f \in G, x_n | f\}$$

is a Gröbner basis for $(I : x_n)$ with respect to $\prec$ and

$$G'' := \{f/x_n^i : f \in G, x_n^i | f, x_n \not| f/x_n^i\}$$

is a Gröbner basis for $(I : x_n^\infty)$ with respect to $\prec$.

Proof. We will prove only the last claim. Clearly, $\mathcal{G}'' \subseteq (I : x_n^\infty)$. To prove $\text{in}_\prec(I : x_n^\infty) \subseteq \langle \text{in}_\prec(f) : f \in G'' \rangle$ let $g \in (I : x_n^\infty)$. We want to show that $\text{in}_\prec(g) \in \langle \text{in}_\prec(f) : f \in G'' \rangle$. There exists an r such that $gx_n^r \in I$. Since G is a Gröbner basis there exists an $f \in G$ such that $\text{in}_\prec(f) | \text{in}_\prec(gx_n^r) = \text{in}_\prec(g)x_n^r$. Let R be the number of times that x_n divides f. By the choice of term order this is also the number of times that x_n divides $\text{in}_\prec(f)$ since f is v-homogeneous. We have $\text{in}_\prec(f/x_n^R)x_n^R | \text{in}_\prec(g)x_n^r$. Since $\text{in}_\prec(f/x_n^R)$ does not contain any x_n we have $\text{in}_\prec(f/x_n^R) | \text{in}_\prec(g)$ and we are done since $f/x_n^R \in G''$. □

Remark 6.1.4 To apply the proposition we must be sure that a term ordering with the desired property exists. As mentioned in Remark 5.5.4 if we have an ideal homogeneous in a positive grading, then there exists a term ordering which agrees with the reverse lexicographic ordering (which is not a *term* ordering itself) on all homogeneous elements. We observe the the reverse lexicographic ordering has the property $x_n | \text{in}_\prec(f) \Rightarrow x_n | f$ for homogeneous f.

We introduce the concept of saturated ideals and show some basic properties.

Definition 6.1.5 Let $f \in k[x_1, \ldots, x_n]$. An ideal $I \subseteq k[x_1, \ldots, x_n]$ is called f-saturated if $(I : f^\infty) = I$.

Lemma 6.1.6 *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $f, g \in k[x_1, \ldots, x_n]$ then*

$$(I : fg^\infty) = ((I : f^\infty) : g^\infty).$$

Proof. To show the inclusion $\subseteq$, let $h \in (I : (fg)^\infty)$. Then $h(fg)^n \in I \Rightarrow hf^n g^n \in I \Rightarrow hg^n \in (I : f^\infty) \Rightarrow h \in ((I : f^\infty) : g^\infty)$ for some n.

To show the inclusion $\supseteq$, let $h \in ((I : f^\infty) : g^\infty)$. Then $hg^n \in (I : f^\infty) \Rightarrow hg^n f^m \in I \Rightarrow h(fg)^{\max(n,m)} \in I \Rightarrow h \in (I : (fg)^\infty)$ for some $n, m \in \mathbb{N}$. □

Corollary 6.1.7 *An ideal $I \subseteq k[x_1, \ldots, x_n]$ is (fg)-saturated if it is f-saturated and g-saturated.*

Proof. We know that $(I : f^\infty) = I$ and $(I : g^\infty) = I$. Hence $I = (I : f^\infty) = ((I : g^\infty) : f^\infty) = (I : (fg)^\infty)$. □

Remark 6.1.8 If an ideal $I \subseteq k[x_1, \ldots, x_n]$ is fg-saturated then it is f-saturated. This can be seen by using the lemma and the definition to get the inclusions:

$$(I : f^\infty) \supseteq I = (I : (fg)^\infty) = ((I : f^\infty) : g^\infty) \supseteq (I : f^\infty).$$

Lemma 6.1.9 *If I and J are ideals in $k[x_1, \ldots, x_n]$ satisfying $I \subseteq J \subseteq (I : f^\infty)$ then $(J : f^\infty) = (I : f^\infty)$.*

Proof. The inclusion $\supseteq$ follows from $I \subseteq J$. To prove the other inclusion, let $h \in (J : f^\infty)$. Then $hf^n \in J \subseteq (I : f^\infty)$ and h times f to some power is indeed in I. □

6.2 Lattice ideals

Definition 6.2.1 Let $u \in \mathbb{Z}^n$. We define the binomial $p_u := x^{u^+} - x^{u^-}$, where $u_i^+ = \max(u_i, 0)$ and $u_i^- = \max(-u_i, 0)$.

Definition 6.2.2 For $C \subseteq \mathbb{Z}^n$ we define:

$$J_C := \langle p_v : v \in C \rangle.$$

If C is a lattice, then J_C is called a *lattice ideal.*

Let $u \wedge v$ denote the coordinatewise minimum of u and v and $u \vee v$ the coordinatewise maximum. We notice that for $u, v \in \mathbb{N}^n$

$$x^u - x^v = x^{u \wedge v} p_{u-v} \tag{3}$$

by counting appearances of x_i for each of the cases $u_i < v_i$, $u_i > v_i$ and $u_i = v_i$.

Lemma 6.2.3 *Let $u, v \in \mathbb{Z}^n$ be vectors then*

$$\frac{x^{u^+ \vee v^+}}{x^{u^+}} p_u - \frac{x^{u^+ \vee v^+}}{x^{v^+}} p_v = x^w p_{u-v}$$

for some $w \in \mathbb{N}^n$. In particular, if $\prec$ is a term ordering on $k[x_1, \ldots, x_n]$ such that $\text{in}_\prec(p_u) = x^{u+}$ *and* $\text{in}_\prec(p_v) = x^{v+}$ *then*

$$S(p_u, p_v) = x^w p_{u-v}.$$

Proof. We compute (using Equation 3 for the third equality):

$$\begin{aligned}\frac{x^{u^+ \vee v^+}}{x^{u^+}} p_u - \frac{x^{u^+ \vee v^+}}{x^{v^+}} p_v &= -\frac{x^{u^+ \vee v^+}}{x^{u^+}} x^{u^-} + \frac{x^{u^+ \vee v^+}}{x^{v^+}} x^{v^-} \\ &= x^{(u^+ \vee u^+) - v} - x^{(u^+ \vee u^+) - u} \\ = x^{((u^+ \vee u^+) - v) \wedge ((u^+ \vee u^+) - u)} & p_{((u^+ \vee u^+) - v) - ((u^+ \vee u^+) - u)} = x^w p_{u-v}\end{aligned}$$

Here x^w is chosen to be the greatest common divisor of the two terms $x^{(u^+ \vee u^+) - v}$ and $-x^{(u^+ \vee u^+) - u}$.□

Lemma 6.2.4 *The elements of a reduced Gröbner basis of a lattice ideal J_L have the form $x^u - x^v$.*

Proof. The lattice ideal has a finite generating set. Each generator can be expressed using finitely many p_v with $v \in L$. In total we need only finitely many p_w to generate I_L. We now compute a Gröbner basis. The S-polynomial of two binomials of the desired form still has the desired form. The remainder of a binomial of the desired form by a set of binomials of the desired form gives a remainder which is again of the desired form. Minimizing and autoreducing the basis keeps this desired form. □

Lemma 6.2.5 *Let J_L be a lattice ideal. Then J_L contains no monomial. Furthermore, $x^u - x^v \in J_L \Leftrightarrow u - v \in L$.*

Proof. Each generator for J_L in Definition 6.2.2 has the property that evaluating it at the point $(1, 1, \dots, 1)$ gives value $0 \in k$. Therefore this holds for any polynomial in J_L. In particular, a polynomial in J_L with just one term would have to have coefficient zero – a contradiction.

Suppose $u - v \in L$. By Equation 3, $x^u - x^v = x^{u \wedge v} p_{u-v}$ which is in J_L because $p_{u-v} \in J_L$ by definition. This proves the "$\Leftarrow$" implication.

Consider the grading $\phi : \mathbb{N}^n \to \mathbb{Z}^n/L$ by letting $\phi(v) := v + L$. Since for every $v \in L$, $\phi(v^+) - \phi(v^-) = \phi(v^+ - v^-) = \phi(v) \in L$ the binomial p_v is ϕ-homogeneous. Since this holds for all $v \in L$, the lattice ideal J_L is ϕ-homogeneous. By Proposition 5.3.14, if $x^u - x^v \in J_L$ then the ϕ-homogeneous parts of this binomial are in J_L. Since J_L contains no monomials, $x^u - x^v$ must be a ϕ-homogeneous part. This proves $\phi(u) = \phi(v)$, implying that $u + L = v + L$ and therefore $u - v \in L$. □

Proposition 6.2.6 *The elements of $\mathcal{G}_\prec(J_L)$ for a lattice ideal J_L have the form $x^{w^+} - x^{w^-}$ with $w \in \mathbb{Z}^n$.*

Proof. By Lemma 6.2.4 we know that every reduced Gröbner basis element is of the form $x^u - x^v$. Consider such a binomial. By Lemma 6.2.5 we know that $u - v \in L$. Since $(u-v)^+ - (u-v)^- = u - v$ the lemma gives $x^{(u-v)^+} - x^{(u-v)^-} \in J_L$. Notice $x^u - x^v = x^{u \wedge v}(x^{(u-v)^+} - x^{(u-v)^-})$. Because $x^u = \text{in}_\prec(x^u - x^v)$, $\text{in}_\prec(x^{(u-v)^+} - x^{(u-v)^-}) = x^{(u-v)^+}$. If $x^{(u-v)^+}$ divides x^u strictly, then x^u cannot be one of the minimal generators of $\text{in}_\prec(I)$. Hence $u \wedge v = 0$ and $(u-v)^+ = u$. We also have $u - v = (u-v)^+ - (u-v)^-$, implying $v = (u-v)^-$. □

Another way of phrasing the proposition is by saying that monomials of a binomial of a reduced Gröbner basis of a lattice ideal have no common factors.

Lemma 6.2.7 *Every lattice ideal J_L is x_i saturated.*

Proof. For simplicity we shall prove this only in the case when J_L is homogeneous with respect to a positive grading. We need to check $(J_L : x_i^\infty) = J_L$. Without loss of generality we may assume $i = n$. Applying Proposition 6.1.3 to compute $(J_L : x_i^\infty)$ we get $G = G''$ because no Gröbner basis element is divisible by x_n (Proposition 6.2.6). This proves that $(J_L : x_i^\infty) = J_L$. □

Proposition 6.2.8 *Let $C \subseteq \mathbb{Z}^n$ be generators of a lattice L. Then $(J_C : (x_1 \cdots x_n)^\infty) = J_L$.*

Proof. We start by showing the "$\supseteq$" inclusion. If $p_u \in (J_C : (x_1 \cdots x_n)^\infty)$ then $p_{-u} \in (J_C : (x_1 \cdots x_n)^\infty)$. We observe that if p_u and $p_v \in (J_C : (x_1 \cdots x_n)^\infty)$, then by Lemma 6.2.3 $x^w p_{u-v} \in (J_C : (x_1 \cdots x_n)^\infty)$ for $w = u^+ \wedge v^+$. Since $(J_C : (x_1 \cdots x_n)^\infty) = (J_C : (x_1^2 \cdots x_n^2)^\infty) = ((J_C : (x_1 \cdots x_n)^\infty) : (x_1 \cdots x_n)^\infty)$ by Lemma 6.1.6, $(J_C : (x_1 \cdots x_n)^\infty)$ is $x_1 \cdots x_n$-saturated and $p_{u-v} \in (J_C : (x_1 \cdots x_n)^\infty)$. Consider a generator p_w for J_L with $w \in L$. We can write $w =$

$\sum_i v_i - \sum_i u_i$ where $u_i, v_i \in C$ and vectors possibly appear more than once in the sum. Applying our observations repeatedly we get $p_w \in (J_C : (x_1 \cdots x_n)^\infty)$.

To prove "$\subseteq$", observe that $J_C \subseteq J_L$ implies $(J_C : (x_1 \cdots x_n)^\infty) \subseteq (J_L : (x_1 \cdots x_n)^\infty)$ which again equals $((\cdots(J_L : x_1^\infty)\cdots) : x_n^\infty) = J_L$ by repeated use of Lemma 6.1.6 and Lemma 6.2.7. □

The proposition actually gives an algorithm for computing lattice ideals. We shall make this completely precise in the following section in the special case of toric ideals.

6.3 Toric ideals

In this section we study a special class of lattice ideals known as *toric ideals.*

Definition 6.3.1 Let $A \in \mathbb{Z}^{d \times n}$ be a matrix. We define the *toric ideal* $I_A := J_L$ where $L \subseteq \mathbb{Z}^n$ is the lattice kernel of A.

Notice that I_A is homogeneous with respect to the grading induced by A because if $Av = 0$ then $Av^+ = Av^-$ and $p_v = x^{v^+} - x^{v^-}$ is A-homogeneous. By Definition 6.2.2 J_L is A-homogeneous. The ideal is also homogeneous in any grading induced by a vector of rowspace(A). In particular, if the rowspace contains a positive vector, Proposition 6.1.3 applies. And the following algorithm can be used to compute I_A.

Algorithm 6.3.2
Input: *A matrix $A \in \mathbb{Z}^{d \times n}$ containing a positive vector in its rowspace.*
Output: *Generators for the toric ideal I_A.*

- *Compute a lattice basis $C \subseteq Z^n$ of the lattice kernel of A using Algorithm 2.1.11*
- *Construct the ideal J_C using Definition 6.2.2.*
- *Compute $I_A = (\cdots(J_C : x_1^\infty) : \cdots : x_n^\infty)$ using Algorithm 6.1.3.*

Proof. By the Definition 6.3.1 and Proposition 6.2.8 we know that $I_A = (J_C : (x_1 \cdots x_n)^\infty)$. Now we only need to observe that

$$(J_C : (x_1 \cdots x_n)^\infty) = (\cdots(J_C : x_1^\infty) : \cdots : x_n^\infty)$$

by applying Lemma 6.1.6 repeatedly. □

Alternatively, I_A can be computed via elimination as we will now explain. Given $A \in \mathbb{Z}^{d \times n}$ with columns $a_1, \ldots, a_n$, we define the ring homomorphism

$$\pi_A : k[x_1, \ldots, x_n] \to k[t_1^{\pm 1}, \ldots, t_d^{\pm 1}]$$
$$cx^v \mapsto ct^{Av}$$

Here $c \in k, v \in \mathbb{N}^n$ and $k[t_1^{\pm 1}, \ldots, t_d^{\pm 1}]$ is the Laurent polynomial ring in d variables. We only defined what π_A does to a single term, while the properties of a ring homomorphisms tells us how to extend π_A to $k[x_1, \ldots, x_n]$.

An other way of defining π_A is by saying that for all $i = 1, \ldots, n$ it substitutes x_i with t^{a_i}.

Example 6.3.3 Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$. The homomorphism $\pi_A : k[x_1, \dots, x_3] \to k[t_1^{\pm 1}, t_2^{\pm 1}]$ sends $x_1x_3 - x_2^2$ to $t^{A(1,0,1)} - t^{A(0,2,0)} = t_1^2 t_1^2 - t_1^2 t_2^2 = 0$. Equivalently, to $(t_1)(t_1 t_2^2) - (t_1 t_2)^2 = 0$.

Proposition 6.3.4 *Let $A \in \mathbb{Z}^{d \times n}$. Then $I_A = \ker(\pi_A)$.*

Proof. The inclusion $\subseteq$ is easy: By definition I_A is generated by binomials of the form $x^{u^+} - x^{u^-}$ with $u \in \text{latticekernel}(A)$. But $A(u^+ - u^-) = 0$ implies $Au^+ = Au^-$ which means that $\pi_A(x^{u^+} - x^{u^-}) = \pi_A(x^{u^+}) - \pi_A(x^{u^-}) = t^{Au^+} - t^{Au^-} = 0$.

Suppose that $\supseteq$ does not hold. Then there is an element $f \in \ker(\pi_A) \setminus I_A$. Choose f such that $\text{in}_{\prec_{\text{lex}}}(f)$ is smallest possible in the $\prec_{\text{lex}}$ ordering. (This can be done because $\prec_{\text{lex}}$ is a well ordering). Without loss of generality $\text{in}_{\prec}(f) = x^u$ for some $u \in \mathbb{N}^n$. Applying π_A to f, $\pi_A(x^u)$ has to cancel with some other term cx^v of f. This means $\pi_A(x^v) = \pi_A(x^u)$, implying $A(u - v) = 0$ and therefore $x^u - x^v \in I_A$. The polynomial $f - (x^u - x^v)$ has smaller initial term than f, and must therefore be in I_A by minimality of the choice of f. However, this implies $f - (x^u - x^v) + (x^u - x^v) = f$ is in I_A, which is a contradiction. □

Corollary 6.3.5 *Every toric ideal is a prime ideal.*

Proof. Let $f, g \in k[x_1, \dots, x_n]$ and suppose $fg \in I_A$. Then $0 = \pi_A(fg) = \pi_A(f)\pi_A(g)$. We proved in Exercise 9, Sheet 1 that $k[x_1, \dots, x_n]$ is an integral domain. The same argument shows that $k[t_1^{\pm 1}, \dots, t_d^{\pm 1}]$ is an integral domain. We conclude that either $\pi_A(f) = 0$ or $\pi_A(g) = 0$. Equivalently $f \in I_A$ or $g \in I_A$. □

We will see a second application of the proposition. But first we need a lemma.

Lemma 6.3.6 *Let $f_1, \dots, f_n \in k[t_1, \dots, t_d]$. Let $\pi : k[x_1, \dots, x_n] \to k[t_1, \dots, t_d]$ be the ring homomorphism defined by $x_i \mapsto f_i$. Define the ideal $J = \langle x_1 - f_1, \dots, x_n - f_n \rangle \subseteq k[t_1, \dots, t_d, x_1, \dots, x_n]$. Then $\ker(\pi) = J \cap k[x_1, \dots, x_n]$.*

Proof. To prove $\supseteq$ we extend π to $\hat{\pi} : k[t_1, \dots, t_d, x_1, \dots, x_n] \to k[t_1, \dots, t_d]$ by $t_i \mapsto t_i$. We observe that $\hat{\pi}(x_i - f_i) = 0$ and therefore $J \subseteq \ker(\hat{\pi})$. Therefore, if $g \in J \cap k[x_1, \dots, x_n]$ then $0 = \hat{\pi}(g) = \pi(g)$ since g contains no t_i's.

Let $g \in \ker(\pi)$. We now compute in $k[t_1, \dots, t_d, x_1, \dots, x_n]/J$. Notice $x_i + J = f_i + J$. Therefore we may do the substitutions performed by π on the representative without changing the coset. Hence $g + J = 0 + J$ and we conclude that $g \in J$. □

Algorithm 6.3.7
Input: *A matrix $A \in \mathbb{N}^{d \times n}$ with columns $a_1, \dots, a_n$.*
Output: *Generators for the lattice ideal I_A.*

- *Define the ideal $J = \langle x_1 - t^{a_1}, \dots, x_n - t^{a_n} \rangle \subseteq k[t_1, \dots, t_d, x_1, \dots, x_n]$.*
- *Compute a lexicographic Gröbner basis G for J with $t_1 \succ \cdots \succ t_d \succ x_1 \succ \cdots \succ x_n$.*

- *Return* $G \cap k[x_1, \ldots, x_n]$.

Proof. The correctness follows from Proposition 6.3.4, Lemma 6.3.6 and Proposition 1.8.1, which tells us how to compute the intersection of an ideal of a polynomial ring with a polynomial ring with fewer variables. □

Notice that the algorithm requires the entries of A to be non-negative. It is possible to extend the algorithm to also work when A has negative entries.

Another important difference between Algorithm 6.3.2 and Algorithm 6.3.7, is that Algorithm 6.3.2 does many reverse lexicographic Gröbner basis computations, while Algorithm 6.3.7 does a single lexicographic Gröbner basis computation. Most often Algorithm 6.3.2 will be fastest. Simply because lexicographic and term orderings are slow.

6.4 Fibers and integer programming

Let again $A \in \mathbb{N}^{d \times n}$. For convenience we assume that A has a positive vector in its rowspace. We are interested in the following optimisation problem:

$$\text{minimize } \omega \cdot v \tag{4}$$

$$\text{subject to } v \in \mathbb{N}^n \text{ and } Av = b$$

where $b \in \mathbb{N}^d$ and $\omega \in \mathbb{R}^n$. This means that we seek an ω-smallest vector v of natural numbers satisfying $Av = b$. Such vector v is called the *optimal point* and $\omega \cdot v$ is called the *optimal value*. The optimization problem is called an *integer programming* problem.

Example 6.4.1 We imagine a country with the following currency. There is a 3 unit coin, a 5 unit coin and a 7 unit coin. Suppose we want to pay the amount $b \in \mathbb{N}$ using coins as few coins as possible. Let $A = \begin{bmatrix} 3 & 5 & 7 \end{bmatrix}$. Then we want to find $v \in \mathbb{N}^3$ such that $Av = b$ and $(1, 1, 1)^t \cdot v$ is minimal. The v_1 is the number of 3 unit coins, v_2 the number of 5 unit coins and v_3 the number of 7 unit coins we pay. That is, we want to solve an integer programming problem.

Let's consider the linear map $A : \mathbb{N}^n \to \mathbb{N}^d$ with $v \mapsto Av$. For $b \in \mathbb{N}^d$ we call the preimage $A^{-1}(b) = \{v \in \mathbb{N}^n : Av = b\}$ the *fibre* of b and denote it $\text{Fiber}(A, b)$. We call the points in $\text{Fiber}(A, b)$ the *feasible* points of the optimization problem in Equation 4. Because A is assumed to have some positive vector $c^t = q^t A$ in its rowspace (with $q \in \mathbb{R}^d$), every point v in the fiber satisfies $c^t v = q^t A v = q^t b$. In particular, for all i we have $v_i \leq \frac{q^t b}{c_i}$. This proves that $|\text{Fiber}(A, b)| < \infty$. (This argument is the similar to that of Lemma 5.4.1.)

Example 6.4.2 Continued. If $b = 29$, the fiber $A^{-1}(b)$ is

$$\{(1, 1, 3), (0, 3, 2), (5, 0, 2), (4, 2, 1), (3, 4, 0), (8, 1, 0)\}.$$

These points are inside the 2-dimensional polytope $\{v \in \mathbb{R}^3_{\geq 0} : Av = 29\}$. The polytope with its lattice points are drawn in Figure 14.

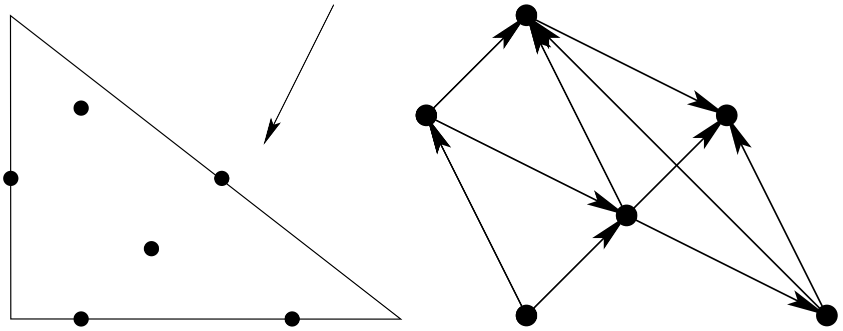

Figure 14: See Example 6.4.1. The 2-dimensional polytope defined by $Av = b, v \in \mathbb{R}^3_{\geq 0}$ and the 6 lattice points in the fiber are shown on the left. The first coordinate has been projected away to get a two-dimensional drawing. The arrow shows the direction we want to minimize. The picture on the right shows the directed graph $\mathrm{Fiber}_{\prec}(A, b)$, which has a unique sink. To get the optimization direction in the first picture we cannot simply project away the last coordinate of $(1, 1, 1)$. In particular it is not obvious that the minimization direction has been drawn correctly. The easiest way to see this is to evaluate $(1, 1, 1)^t \cdot v$ at each of the 6 lattice points.

We now pick a term ordering $\prec$ and compute the reduced Gröbner basis $\mathcal{G}_{\prec}(I_A)$. We identify the point v in the fibre with the monomial $x^u \in k[x_1, \dots, x_n]$.

Definition 6.4.3 Let $A \in \mathbb{N}^{d \times n}$, $b \in \mathbb{N}^d$ and $\prec$ be a term ordering. We define the directed graph $\mathrm{Fiber}_{\prec}(A, b)$ as follows. The vertices are $A^{-1}(b) \subseteq \mathbb{N}^n$. There is an outgoing edge from u for every $x^\alpha - x^\beta \in \mathcal{G}_{\prec}(I)$ with $x^\alpha | x^u$ and $u - \alpha + \beta \in \mathrm{Fiber}_{\prec}(A, b)$. The outgoing edge ends in $v = u - \alpha + \beta$.

Lemma 6.4.4 *There is an edge from u to v in the graph* $\mathrm{Fiber}_{\prec}(A, b)$ *if and only if it is possible to reduce x^u to x^v with one step of the division algorithm (Algorithm 1.5.1) where $\{f_1, \dots, f_s\} = \mathcal{G}_{\prec}(I_A)$ and we use the term order $\prec$.*

Proof. We only have to check that $x^u - (x^u / x^\alpha)(x^\alpha - x^\beta) = x^{u - \alpha + \beta}$. □

Definition 6.4.5 A *sink* in a directed graph is a vertex with no out-going edges.

Example 6.4.6 Continued. Let $\prec$ be the graded lexicographic term ordering. Using the variable names x, y, z we have $\mathcal{G}_{\prec}(I_A) =$

$$\{\underline{y^7} - z^5,$$

$$\underline{xz} - y^2,$$

$$\underline{xy^5} - z^4,$$

$$\underline{x^2y^3} - z^3,$$

$$\underline{x^3y} - z^2,$$

$$\underline{x^4} - yz\}$$

The graph $\mathrm{Fiber}_\prec(A,b)$ is shown on the right in Figure 14. The vertex $(0,3,2)$ is the unique sink because no initial term in $\mathcal{G}_\prec(I_A)$ divides y^3z^2.

Proposition 6.4.7 *The graph* $\mathrm{Fiber}_\prec(A,b)$ *has no cycles and a unique sink. In particular there is a directed path in* $\mathrm{Fiber}_\prec(A,b)$ *from any vertex to the sink.*

Proof. Let u and v be two different sinks in $\mathrm{Fiber}_\prec(A,b)$ then $Au = b = Av$ implying $x^u - x^v \in I_A$. There must exist an $f \in \mathcal{G}_\prec(I_A)$ such that $\mathrm{in}_\prec(f)$ divides $\mathrm{in}_\prec(x^u - x^v)$, which is either x^u or $-x^v$. But this proves that u and v cannot both be sinks.

If we have a cycle $u_0, \dots, u_s = u_0$ in the graph $\mathrm{Fiber}_\prec(A,b)$ then it would be possible for the division algorithm to do the step $x^{u_0} \to x^{u_1} \to x^{u_s} = x^{u_0}$ forever, which contradict that the division algorithm always terminates. □

We now turn to the problem of finding the optimal solution to the integer programming problem.

Lemma 6.4.8 *Let* $A \in \mathbb{N}^{d\times n}$, $b \in \mathbb{N}^d$, $\omega \in \mathbb{R}^n$ *and* $\prec$ *be a term ordering. Then the sink of* $\mathrm{Fiber}_{\prec_\omega}(A,b)$ *is an optimal solution to the minimization problem in Equation 4.*

Proof. Let u be the sink, and let $v \in \mathrm{Fiber}(A,b)\backslash\{u\}$. We must show that $\omega\cdot u \leq \omega\cdot v$. But suppose that $\omega\cdot u > \omega\cdot v$, then $\mathrm{in}_{\prec_\omega}(x^u - x^v) = x^u$. Since $x^u - x^v \in I_A$ we have $x^u \in \mathrm{in}_{\prec_\omega}(I_A)$ and therefore there exists a binomial $x^\alpha - x^\beta \in \mathcal{G}_{\prec_\omega}(I)$ such that $x^\alpha | x^u$ which means that u is not a sink. A contradiction. □

Example 6.4.9 Continued. Suppose we know the feasible point $(5,0,2)$ of the integer programming problem. We want to find an optimal point in the to the problem, that is a point which "minimizes ω". We find the sink by doing a polynomial division, starting with the monomial x^5z^2:

$$x^5z^2 \to xyz^3 \to y^3z^2$$

In the graph of Figure 14 we move from the left-most matrix to the upper-right vertex in two steps. The solution $(0,3,2)$ is optimal. That is, we need only 5 coins to pay the amount 29, we do it paying 3 five-unit coins and 2 seven-unit coins.

Notice that there could be more than one optimal solution to the optimization problem. That is the case in the example. The solution $(1,1,3)$ is also optimal.

We still did not explain how to find the feasible solution $(5,0,2)$ to the problem. The solution is to introduce artificial variables. We make a new problem with $n + d$ variables.

Proposition 6.4.10 *Let $A \in \mathbb{N}^{d\times n}, b \in \mathbb{N}^d$. There exists a feasible $v \in \mathbb{N}^n$ with $Av = b$ if and only if the problem*

$$\begin{aligned}&\text{minimize} \sum_{i=n+1}^{n+d} u_i \\ &\text{subject to } [A|I]u = b \text{ and } u \in \mathbb{N}^n \times \mathbb{N}^d\end{aligned} \tag{5}$$

has an optimal solution u with optimal value 0. Here I denotes the $d \times d$ identity matrix.

Proof. If the new problem has an optimal solution with optimal value 0, then the last d entries of u must be zero. But this means that if we let v be the first n entries of u we have $b = [A|I]u = Av$. Hence v is a feasible solution to the inequality system of Equation 4.

On the other hand, if $Av = b$ has a solution, then indeed the optimal value of the new problem is zero. □

The point of the proposition above is that it is trivial to find a feasible solution to the problem in Equation 5. Namely, we just take the vector $(0, \dots, 0, b_1, \dots, b_d)$.

Example 6.4.11 Continued. To find just one way of paying 29 units, we introduce the artificial coin with value one. We now solve the problem:

$$\begin{aligned}&\text{minimize } (0,0,0,1)^t \cdot u \\ &\text{subject to } [3\ 5\ 7\ 1]u = 29 \text{ and } u \in \mathbb{N}^3 \times \mathbb{N}^1\end{aligned} \tag{6}$$

We compute the Gröbner basis:

$$\{w^3 - x, zw - xy, z^3 - x^7, yw - x^2, yz - x^4,$$
$$y^2 - xz, xw^2 - y, x^2w - z, x^3y - z^2\}$$

with respect to an ordering $\prec_{(0,0,0,1)}$. We now take the feasible solution $(0, 0, 0, 29)$ of the new problem and convert it to the monomial w^{29}. The remainder produced by the division algorithm:

$$w^{29} \to xw^{26} \to \cdots \to x^9w^2 \to x^8y \to x^5z^2$$

is x^5z^2, corresponding the the feasible solution $(5, 0, 2)$ of the original problem.

We present the complete algorithm for solving an integer programming problem using toric ideals:

Algorithm 6.4.12
Input: *$A \in \mathbb{N}^{d\times n}, b \in \mathbb{N}^d, \omega \in \mathbb{R}^n$ such that A has a positive vector in its rowspace.*
Output: *A vector $v \in \mathbb{N}^n$ such that $Av = b$ and $\omega \cdot v$ is smallest possible among such vectors.*

- *Let $B := [A|I]$.*

- *Compute the toric ideal I_B using Algorithm 6.3.2/6.3.7.*

- *Compute the Gröbner basis $\mathcal{G}_{\prec_{(0,\dots,0,1,\dots,1)}}(I_B)$, where $\prec$ is some term ordering.*

- *Compute the remainder x^u of $x^{(0,\dots,0,b_1,\dots,b_d)}$ reduced by $\mathcal{G}_{\prec_{(0,\dots,0,1,\dots,1)}}(I_B)$ using the division algorithm.*

- *If $\exists i \in \{n+1, n+2, \dots, n+d\} : u_i \neq 0$ then the problem $Av = b$, $v \in \mathbb{N}^n$ has no solution and the algorithm terminates.*

- *Compute the toric ideal I_A using Algorithm 6.3.2/6.3.7.*

- *Compute the Gröbner basis $\mathcal{G}_{\prec_\omega}(I_A)$.*

- *Compute the remainder x^v of x^u reduced by $\mathcal{G}_{\prec_\omega}(I_A)$ using the division algorithm.*

- *The vector $v \in \mathbb{N}^n$ is an optimal point of the optimization problem.*

Remark 6.4.13 The toric ideals I_A and I_B are strongly related. In fact in Algorithm 6.3.7 $J = I_B$. Therefore it is not necessary to compute I_A from scratch.

Remark 6.4.14 The optimization problem in Equation 4 above is known as an integer programming problem. Such problems can be very difficult to solve. If we change the requirement $v \in \mathbb{N}^n$ to $v \in \mathbb{R}^n_{\geq 0}$ then the problem becomes a *linear programming* problem. Linear programming problems can be solved with Dantzig's *simplex algorithm* (and also with Algorithm 3.1.3, how?) and even algorithms with polynomial time complexity are known. The most important open problem in theoretical computer science ("P≠NP?") essentially asks if integer programming really is harder than linear programming.

7 Regular triangulations and secondary fans

A triangle in the two-dimensional plane has 3 vertices. The generalisation of a triangle to higher dimensions is a d-simplex.

Definition 7.0.15 A d-dimensional polytope $P \subseteq \mathbb{R}^n$ is called a *d-simplex* if it has exactly $d+1$ vertices. A pointed $d+1$-dimensional polyhedral cone $C \subseteq \mathbb{R}^n$ is called a $(d+1$-$)$*simplicial* cone if it has exactly $d+1$ rays.

We notice that the convex hull of any non-empty subset of the $d+1$ vertices is a face of P and every face is of this form. Similarly, any subset of the $d+1$ rays gives a face of C.

Simplicial cones appear in *triangulations* of cones over a finite set of vectors. In this section we will study such triangulations and see how they are connected to initial ideals of toric ideals.

7.1 Simplicial complexes and Stanley-Reisner ideals

Let $[n]$ denote the set $\{1,2,\dots,n\}$. A polyhedral complex consisting either only of simplices or only of simplicial cones is called simplicial. Forgetting the geometry, simplicial polyhedral complexes can be studied purely combinatorially. For this purpose, any subset of $[n]$ is called an *abstract simplex.*

Definition 7.1.1 A(n abstract) *simplicial complex* Δ on $[n]$ is a set of subsets of $[n]$ such that whenever $s \in \Delta$, then every subset of s is also in Δ. The subsets in Δ are called the *faces* of Δ.

The reader should compare this definition to Definition 3.4.1.

Example 7.1.2 The following is an example of a simplicial complex on [6].

$$\Delta = \{\{1,2,5\},\{2,3,6\},\{3,1,4\},\{1,5,4\},\{2,6,5\},\{3,4,6\},\{4,5,6\},$$

$$\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{2,3\},\{2,5\},\{2,6\},\{3,4\},\{3,6\},\{4,5\},\{4,6\},\{5,6\},$$
$$\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},\emptyset\}.$$

Simplicial complexes are connected to polynomial ideals as follows.

Definition 7.1.3 Let Δ be a simplicial complex on $[n]$. The Stanley-Reisner ideal of Δ is defined as

$$\langle \prod_{i\in S} x_i : S \subseteq [n], S \notin \Delta \rangle \subseteq k[x_1,\dots,x_n].$$

Notice, the Stanley-Reisner ideal is "generated by the minimal non-faces of Δ".

Example 7.1.4 The Stanley-Reisner ideal of Δ from Example 7.1.2 is

$$\langle x_1x_6, x_2x_4, x_3x_5, x_1x_2x_3 \rangle \subseteq k[x_1,\dots,x_6].$$

This takes some time to see.

7.2 Radical ideals

Definition 7.2.1 Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal. We define the *radical* of I to be the set

$$\sqrt{I} = \{f : \exists m \in \mathbb{N} : f^m \in I\}.$$

Notice that $I \subseteq \sqrt{I}$. An ideal I for which $I = \sqrt{I}$ is called a *radical* ideal.

Example 7.2.2 The radical of $\langle a^2 + b^2 + 2ab \rangle \subseteq k[a, b]$ is $\langle a + b \rangle \subseteq k[a, b]$.

Lemma 7.2.3 *The radical of an ideal is an ideal.*

Proof. Left to the reader. □

Lemma 7.2.4 *The radical of a monomial ideal $I \subseteq k[x_1, \ldots, x_n]$ is a monomial ideal generated by monomials of the form $\prod_{i \in S} x_i$ where $S \subseteq [n]$. Furthermore, any ideal generated by such monomials is a radical ideal.*

Proof. Suppose $f \in \sqrt{I}$. Then for some $m \in \mathbb{N}$ we have $f^m \in I$. Consider a term of f and let $S \subseteq [n]$ index the variables x_i appearing in this term. There is also a term in f^m involving only these variables ($f_{|x_i=0 \text{ for } i \notin S} \neq 0$ implies $(f^m)_{|x_i=0 \text{ for } i \notin S} = (f_{|x_i=0 \text{ for } i \notin S})^m \neq 0$ because $k[x_1, \ldots, x_n]$ is an integral domain). Since I is a monomial ideal the term is in I and therefore $\prod_{i \in S} x_i$ is in $\sqrt{I}$. Hence, for every $f \in \sqrt{I}$ we have shown that its terms are in $\sqrt{I}$ which proves that $\sqrt{I}$ is a monomial ideal. Moreover, $\sqrt{I}$ is generated by terms of form $\prod_{i \in S} x_i$ with $S \subseteq [n]$.

To prove the last claim, suppose I is generated by such monomials. We must show that $\sqrt{I} \subseteq I$. Let $f \in \sqrt{I}$ be a monomial with $f^m \in I$. Then there exists $S \subseteq [n]$ such that $\prod_{i \in S} x_i \in I$ and $\prod_{i \in S} x_i | f^m$, implying $\prod_{i \in S} x_i | f$. This prove that $f \in I$ as desired. □

Example 7.2.5 We have $\sqrt{\langle a^3 b, c^2, cb^2 \rangle} = \langle ab, c, cb \rangle = \langle ab, c \rangle \subseteq k[a, b, c]$.

7.3 Regular triangulations of vector configurations

In what follows we let $A \in \mathbb{Z}^{d \times n}$ and think of the columns $a_1, \ldots, a_n$ of A as being vectors of a vector configuration. For simplicity we assume that $(1, \ldots, 1)$ is in the rowspace of A. Identifying $[n]$ with the columns of A we wish to define a simplicial complex $\Delta_\omega(A)$ on $[n]$ for $\omega \in \mathbb{R}^n$. A set $S \subseteq [n]$ is in $\Delta_\omega(A)$ if and only if there exists $v \in \mathbb{R}^d$ such that $v \cdot a_i - \omega_i = 0$ for $i \in S$ and $v \cdot a_i - \omega_i < 0$ for $i \notin S$. In other words, we take the vectors $a_1, \ldots, a_n$ and lift them into $\mathbb{R}^{d+1}$ by appending the additional coordinates $\omega_1, \ldots, \omega_n$. We now form the cone $P_\omega := \text{cone}((a_1, \omega_1), \ldots, (a_n, \omega_n))$. This cone has "lower faces", namely faces of the form $\text{face}_{(v,-1)}(P_\omega)$ with $v \in \mathbb{R}^d$. For each such face we take the indices of the lifted vectors in that face and let them form a simplex in $\Delta_\omega(A)$. As Example 7.3.2 shows this definition does not work for all ω, so we need to be very careful.

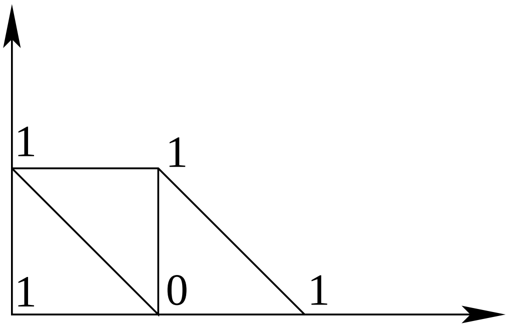

Figure 15: The projection of the lower faces of the 3-dimensional polytope constructed in Example 7.3.1 gotten by intersecting P_ω with $\{0\} \times \mathbb{R}^3$. The numbers indicate the lifted heights of the vectors.

Example 7.3.1 Consider

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

We choose $\omega = (1,1,0,1,1)$. We now construct the lifted cone P_ω and look at its lower faces. The cone is 4-dimensional, but since the vector $(1,1,1,1,1)$ is in the row space of A all the column vectors are contained in an affine hyperplane $H \subseteq \mathbb{R}^3$ not passing through zero. Therefore, finding lower faces of P_ω is the same as finding lower faces of $P_\omega \cap (H \times \mathbb{R})$. Thus finding the simplices of $\Delta_\omega(A)$ amounts to finding the lower faces of a 3-dimensional polytope. From Figure 15 we read off

$$\Delta_\omega(A) = \{\{1,2,3\},\{2,3,4\},\{3,4,5\},\{1,2\},\{1,3\},\{2,3\},\{2,4\},$$
$$\{3,4\},\{3,5\},\{4,5\},\{1\},\{2\},\{3\},\{4\},\{5\},\emptyset\}$$

Example 7.3.2 Choose $\omega = 0$ and consider

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

The square itself $\{1,2,3,4\}$ becomes a subset in $\Delta_\omega(A)$. But $\{1,2,3\} \notin \Delta_\omega(A)$, so $\Delta_\omega(A)$ is not a simplicial complex.

We will say that ω is *generic* if $\Delta_\omega(A)$ is a simplicial complex. A sufficient condition for ω to be generic is that the lifted vectors are linearly independent. In particular, any set of $d+1$ of the lifted vectors is not contained in a d-dimensional subspace. Example 7.3.2 does not satisfy this requirement.

By a *regular triangulation* of the column vectors of A we mean a simplicial complex of form $\Delta_\omega(A)$ or sometimes the polyhedral complex obtained by projecting the lower faces of P_ω to $\mathbb{R}^d$. In general, a triangulation of the columns

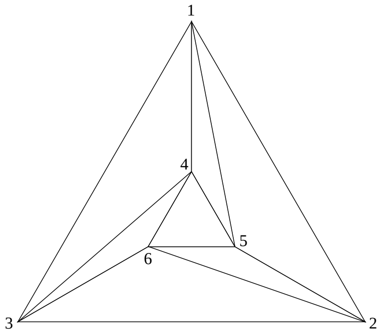

Figure 16: This picture can be interpreted as the simplicial complex of Example 7.1.2 or as the triangulation of Example 7.3.3. In the latter case, what we see is actually how the positive orthant $\mathbb{R}^3_{\geq 0}$ is subdivided into simplicial cones.

of A is a polyhedral complex consisting of simplicial cones which cover cone(A) and whose rays are generated by columns of A. The following example shows that not every triangulation is regular. That is, not every triangulation comes from a lift ω.

Example 7.3.3 [13, Example 8.2] Consider the matrix with columns $a_1, \dots, a_6$

$$A = \begin{bmatrix} 4 & 0 & 0 & 2 & 1 & 1 \\ 0 & 4 & 0 & 1 & 2 & 1 \\ 0 & 0 & 4 & 1 & 1 & 2 \end{bmatrix}.$$

The cone spanned by the columns of A is the positive orthant $\mathbb{R}^3_{\geq 0}$. We may cover the positive orthant with the simplicial cones shown in Figure 16. However, there is no $\omega \in \mathbb{R}^6$ inducing this as a regular triangulation, because the appearance of $\{1, 5\}$ as a lower face implies

$$\frac{4}{5}\omega_4 + \frac{1}{5}\omega_2 > \frac{4}{5}\omega_5 + \frac{1}{5}\omega_1$$

since $\frac{4}{5}a_4 + \frac{1}{5}a_2 = \frac{4}{5}a_5 + \frac{1}{5}a_1$. Similarly

$$\frac{4}{5}\omega_5 + \frac{1}{5}\omega_3 > \frac{4}{5}\omega_6 + \frac{1}{5}\omega_2$$

$$\frac{4}{5}\omega_6 + \frac{1}{5}\omega_1 > \frac{4}{5}\omega_4 + \frac{1}{5}\omega_3$$

Adding up the three inequalities we get the contradiction

$$\frac{4}{5}(\omega_4 + \omega_5 + \omega_6) + \frac{1}{5}(\omega_1 + \omega_2 + \omega_3) > \frac{4}{5}(\omega_4 + \omega_5 + \omega_6) + \frac{1}{5}(\omega_1 + \omega_2 + \omega_3).$$

7.4 Sturmfels' Theorem

We now presents the theorem which ties initial ideal, radicals, triangulations and Stanley-Reisner ideals together. For the proof we need to be a bit careful about what *generic* means in this case.

Theorem 7.4.1 (Sturmfels, 1991) *Let $A \in \mathbb{Z}^{d\times n}$ with a positive vector in its rowspace, then for generic $\omega \in \mathbb{R}^n$ we have $\sqrt{\mathrm{in}_\omega(I_A)}$ is the Stanley-Reisner ideal of $\Delta_\omega(A)$.*

Proof. We must prove

$$\sqrt{\mathrm{in}_\omega(I_A)} = \langle \prod_{i\in S} x_i : s \subseteq [n] \wedge s \notin \Delta_\omega(A)\rangle.$$

For "$\supseteq$" let $S \subseteq [n]$ be a non-face of $\Delta_\omega(A)$. This means that s does not define a lower face of P_ω. In other words, there does not exist $v \in \mathbb{R}^d$ such that

$$\forall i \in [n] \setminus S : v \cdot a_i - \omega_i < 0$$

$$\text{and } \forall i \in S : v \cdot a_i - \omega_i = 0$$

Equivalently, there does not exist $v \in \mathbb{R}^{d+1}$ with $v_{n+1} \leq -1$ such that

$$\forall i \in [n] \setminus S : [a_i^t, \omega_i]v \leq -1$$

$$\text{and } \forall i \in S : [a_i^t, \omega_i]v \leq 0 \text{ and } -[a_i^t, \omega_i]v \leq 0$$

The condition of the non-existence of v above is equivalent to the polyhedral cone $P_{A'b'} = \{v \in \mathbb{R}^{d+1} : A'v \leq b'\}$ being empty, where

$$A' = \begin{bmatrix} MA^T & M\omega \\ \mathbf{0} & 1 \end{bmatrix}$$

where M has a row e_i for each $i \notin S$ and rows e_i and $-e_i$ for each $i \in S$. The vector b' is -1 for each of the first set of rows of M and on the last entry. By Farkas' Lemma 3.3.6 we can find $y \in \mathbb{R}_{\geq 0}^{n+|S|+1}$ such that $y^tA' = 0$ and $y^tb' = (-1)$. Since the set of possible y vectors is described by inequalities with rational coordinates, we may assume that $y \in \mathbb{Q}_{\geq 0}^{n+|S|+1}$. (The y could be found using Fourier-Motzkin Algorithm 3.1.3 which produces rational output on rational input.) Let y' be the subvector of the first $n + |S|$ coordinates of y. Then $y'^tMA^t = 0$. In particular $M^ty' \in \ker(A)$. Suppose $M^ty' = 0$ then y has to be zero on coordinates indexed by $[n] \setminus S$. But then $y^tb' = (-1)$ gives that y has last coordinate 1 which together with $y^tA' = 0$ implies $y'^tM\omega = -1$. That contradicts $y^tA' = 0$.

We scale M^ty' positively to get a vector $u \in \mathbb{Z}^n \setminus \{0\}$ with the property $Au = 0$. Therefore $x^{u^+} - x^{u^-} \in I_A$. By construction of M, x^{u^-} involves only variables indexed by S. We just need to argue that $\mathrm{in}_\omega(x^{u^+} - x^{u^-}) = -x^{u^-}$ which would imply $x^{u^-} \in \mathrm{in}_\omega(I_A)$ and therefore $\prod_{i\in S} x_i \in \sqrt{\mathrm{in}_\omega(I_A)}$. We know already from $y^tA' = 0$ (since last coordinate of y is non-negative) that $u\cdot\omega \leq 0$. Since ω is generic we get $u \cdot \omega < 0$ as wanted.

For "$\subseteq$" suppose that for some subset $S \subseteq [n]$ we have $\prod_{i \in S} x_i \in \sqrt{\text{in}_\omega(I_A)}$. Then there must exist a vector $u \in \mathbb{Z}^n$ such that $Au = 0$, $u_i \geq 0$ for $i \notin S$ and $u \cdot \omega < 0$. We can express this vector as a non-negative combination of the rows of M, say $u = y'^t M$. With the definition of A' as above, $y'^t M\omega < 0$ implies that we can find a last positive coordinate for $y = (y', \cdot)$ such that $y^t A' = 0$. Moreover, by the choice of b and since $u_i \geq 0$ for $i \notin S$ we have $y^t b' < 0$ which means that from the inequality description of $P_{A'b'}$ we make non-negative combinations to reach an inequality $0 = y^t A' \leq y^t b' < 0$ which cannot be satisfied. Hence S is a non-face. $\square$

Remark 7.4.2 In the proof $u \neq 0$ is in the lattice kernel of A. If $\omega \in \mathbb{R}^n$ has the property that $\omega \cdot u \neq 0$ for all $u \neq 0$ in the lattice kernel of A, then ω is sufficiently generic for the theorem to hold. Such generic vectors exist in any open ε-ball (for $n > 0$). For $n = 2$ take for example $\omega \in \mathbb{Q} \times (\mathbb{R} \setminus \mathbb{Q})$.

Example 7.4.3 Let

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$

We use Algorithm 6.3.2 to compute the toric ideal $I_A \subseteq k[a, b, c, d]$. We already found a basis for the lattice kernel in Exercise 4, Sheet 3. For example this one:

$$C = \{(1, -2, 1, 0)^t, (2, -3, 0, 1)^t\}$$

The algorithm tells us first to consider $J_C = \langle ac - b^2, a^2d - b^3\rangle$. We must now compute $I_A = ((((J_c : a^\infty) : b^\infty) : c^\infty) : d^\infty)$. The first step is to compute a reduced Gröbner basis of J_C with respect to a graded reverse lexicographic $\prec$ with $d \prec c \prec b \prec a$. We get

$$\{\underline{b^2} - ac, \underline{abc} - a^2d, \underline{a^2c^2} - a^2bd\}.$$

We can divide out by a (repeatedly) in the second and third generator and get the following Gröbner basis for $(J_C : a^\infty)$:

$$\{\underline{b^2} - ac, \underline{bc} - ad, \underline{c^2} - bd\}.$$

We change the term ordering and compute the reduced Gröbner basis

$$G = \{\underline{c^2} - bd, \underline{bc} - ad, \underline{b^2} - ac\}.$$

Here b does not divide any polynomial, so this is a Gröbner basis for $((J_C : a^\infty) : b^\infty)$. We repeat this process for c and d, but in both iterations, we cannot divide by the variable. Hence the Gröbner basis above is already a Gröbner basis for I_A.

Let's read off the Gröbner cone for the reduced Gröbner basis G above. We will use Corollary 4.2.4. The vector $\omega \in \mathbb{R}^4$ is in $C_\prec(I_A)$ if and only if

$$\text{in}_\prec(\text{in}_\omega(c^2 - bd)) = c^2 \text{ and } \text{in}_\prec(\text{in}_\omega(bc - ad)) = bc \text{ and } \text{in}_\prec(\text{in}_\omega(b^2 - ac)) = b^2.$$

With matrix representation we may write this as

$$\begin{bmatrix} 0 & 1 & -2 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 1 & 0 \end{bmatrix} \omega \leq 0.$$

The second inequality is a consequence of the first and the third. Therefore $C_{\prec}(I_A) \subseteq \mathbb{R}^4$ has just two facets. Its lineality space is two-dimensional.

Since $(1,1,1,1)$ is in the row-space of A, the toric ideal I_A is homogeneous in the total grading. By Proposition 5.5.3 the Gröbner fan covers all of $\mathbb{R}^4$.

We would like to find another Gröbner cone. We can use Algorithm 5.9.4 of the Gröbner walk to find one of the two neighbouring cones. Continuing in this way, we get all cones of the form $C_{\prec'}(I_A)$ where $\prec'$ is a term ordering. Furthermore this procedure also gives us all of the initial ideals $\text{in}_{\prec'}(I_A)$:

$\mathcal{G}_{\prec'}(I_A)$	$\text{in}_{\prec'}(I_A)$	$\sqrt{\text{in}_{\prec'}(I_A)}$
$\{bd - c^2, ad - bc, ac - b^2\}$	$\langle bd, ad, ac\rangle$	$\langle bd, ad, ac\rangle$
$\{bd - c^2, b^2 - ac, ad - bc\}$	$\langle bd, b^2, ad\rangle$	$\langle b, ad\rangle$
$\{bd - c^2, bc - ad, b^2 - ac, ad^2 - c^3\}$	$\langle bd, bc, b^2, ad^2\rangle$	$\langle b, ad\rangle$
$\{c^3 - ad^2, bd - c^2, bc - ad, b^2 - ac\}$	$\langle c^3, bd, bc, b^2\rangle$	$\langle c, b\rangle$
$\{c^2 - bd, ad - bc, ac - b^2\}$	$\langle c^2, ad, ac\rangle$	$\langle c, ad\rangle$
$\{c^2 - bd, bc - ad, ac - b^2, a^2d - b^3\}$	$\langle c^2, bc, ac, a^2d\rangle$	$\langle c, ad\rangle$
$\{c^2 - bd, bc - ad, b^3 - a^2d, ac - b^2\}$	$\langle c^2, bc, b^3, ac\rangle$	$\langle c, b\rangle$
$\{c^2 - bd, bc - ad, b^2 - ac\}$	$\langle c^2, bc, b^2\rangle$	$\langle c, b\rangle$

The Gröbner fan is shown in Figure 17. We see that there are 4 different radical ideals $\sqrt{\text{in}_{\prec'}(I_A)}$. According to Theorem 7.4.1 there are 4 different regular triangulations on the columns of A. The lifts inducing these are shown in Figure 18. The simplicial complexes for these triangulations are:

- $\{\{1,2\},\{2,4\},\{1\},\{2\},\{4\},\emptyset\}$
- $\{\{1,3\},\{3,4\},\{1\},\{3\},\{4\},\emptyset\}$
- $\{\{1,2\},\{2,3\},\{3,4\},\{1\},\{2\},\{3\},\{4\},\emptyset\}$
- $\{\{1,4\},\{1\},\{4\},\emptyset\}$.

We check that each of these have a Stanley-Reisner ideal equal to one of the radicals listed in the right column of the table above.

The example leads us to the following definition.

Definition 7.4.4 Let $A \in \mathbb{Z}^{d\times n}$ be a matrix with a positive vector in its row space. Let M be a monomial ideal which is the radical of an initial ideal of I_A. A *secondary cone* of A is the union of all Gröbner cones $C_{\prec}(I_A)$ with $\sqrt{\text{in}_{\prec}(I_A)} = M$. The collection all secondary cones (and their faces) is the *secondary fan* of A.

Usually people define the secondary fan is terms of triangulations and not in terms of radicals and initial ideals. We end with presenting the following theorem without a proof.

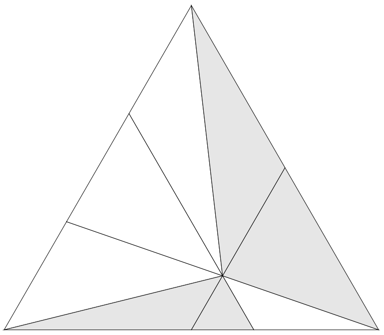

Figure 17: The Gröbner fan in Example 7.4.3 is 4-dimensional so we cannot draw it. Instead we draw its intersection with a triangle. The triangle intersects all 8 full-dimensional Gröbner cones. The triangles whose initial ideals have the same radical are next to each other. The colors indicate how the cones are grouped according to radical.

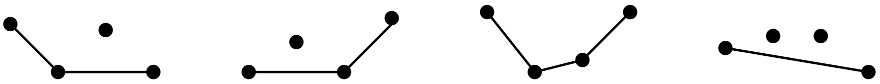

Figure 18: The lifts which induce the four triangulations of the vector configuration in Example 7.4.3.

Theorem 7.4.5 *The secondary fan of a matrix $A \in \mathbb{Z}^{d \times n}$ is a polyhedral fan.*

8 A brief introduction to tropical geometry

The tropical semi-ring $R := (\mathbb{R}, \oplus, \odot)$ consists of the real numbers with two operations: tropical plus $\oplus$ and tropical times $\odot$ where:

$$x \oplus y := \max(x, y) \text{ and } x \odot y := x + y$$

This is almost a ring in the sense that for all $x, y, z \in \mathbb{R}$:

$$x \odot (y \oplus z) = x \odot y \oplus x \odot z$$

Moreover, 0 is the neutral element for $\odot$ and we could include $-\infty$ in $\mathbb{R}$ to get a neutral element for $\oplus$. However, there can be no (tropical) additive inverses since for example $x \oplus 5 = -\infty$ has no solution.

Tropical polynomial functions are piecewise linear. Hence their graphs become polyhedral complexes as the following example shows.

Example 8.0.6 Let $p = 2 \oplus 1 \odot x \oplus (-1) \odot x \odot x$. Figure 19 shows the graph of $p(x) = \max(2, \mathrm{x} + 1, 2\mathrm{x} - 1)$.

We wish to define the "zero set" or "roots" of a tropical polynomial. To make a quadratic polynomial (with a constant term) have two roots (with multiplicity), the right thing to do, is to define the zero set to be the set of point where the maximum is attained at least twice. In our example the roots are 1 and 2.

The amazing fact is that a lot of properties are preserved when studying the tropical semi-ring rather than a field such as $(\mathbb{R}, +, \cdot)$ or $(\mathbb{C}, +, \cdot)$. We will see a few such properties in the following and see how tropical geometry is closely related to the topics of this course.

8.1 Tropical hypersurfaces

Let's now consider a tropical polynomial f in n variables $x_1, \ldots, x_n$. We define its *tropical hypersurface* $T(f)$ to be the set of $x \in \mathbb{R}^n$ such that the maximum in the expression $f(x)$ is attained at least twice. The tropical hypersurface can also be thought of as a polyhedral complex.

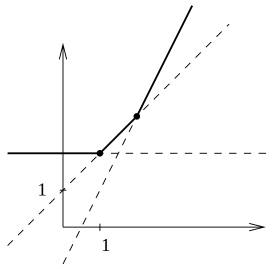

Figure 19: The graph of $p(x) = \max(2, \mathrm{x} + 1, 2\mathrm{x} - 1)$ in Example 8.0.6

Example 8.1.1 Consider the tropical polynomial

$$f(x,y) = (-1) \oplus (0) \odot x \oplus (-1) \odot x \odot x \oplus (-1) \odot y \oplus (-1) \odot x \odot y.$$

Its tropical hypersurface is shown in Figure 20. In particular the maximum is attained three times at each of the three points $(-1,0), (0,1)$ and $(1,1)$.

It is interesting to compare Figure 20 to Figure 15. For fixed support $\mathrm{supp}(f)$, the combinatorial types of tropical hypersurfaces that are defined as the coefficients vary are exactly indexed by the cones of a secondary fan.

8.2 Enumerative geometry

Since the tropical hypersurface of Example 8.1.1 has dimension 1 it is also called a *tropical curve*. Other examples of tropical curves are the tropical lines which in the plane are tropical hypersurfaces defined by polynomials of the form

$$a \odot x \oplus b \odot y \oplus c$$

where $a, b, c \in \mathbb{R}$. A tropical line consists of three halflines meeting in a point and going off in directions north-east, west and south. As the coefficients a, b, c vary, the tropical line is translated around in the plane.

Two *generic* tropical curves intersect in exactly one point as shown in Figure 21. Moreover, given two *generic* points in the plane, exactly one tropical line passes through them. These statements do not hold for all lines/points.

Another surprising fact is that Bezout's Theorem holds tropically. If a tropical curve of degree m and a tropical curve of degree n in the plane intersect in more than mn points then they must intersect in infinitely many points. Indeed, the tropical curve of degree 2 in Figure 20 intersects a tropical line in either $1, 2$ or infinitely many points.

Finding tropical counts such as mn above and more complicated counts is the topic of *enumerative tropical geometry.* Many of these counts translate to the classical setting where $(\mathbb{R}, +, \cdot)$ or $(\mathbb{C}, +, \cdot)$ are considered.

8.3 Tropical varieties

The situation is slightly simpler if all coefficients are 0, which is OK tropically! Then the tropical hypersurface of a polynomial f is the union of the $n-1$

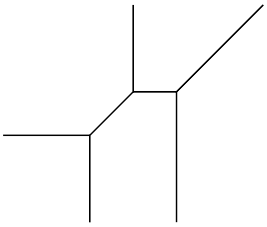

Figure 20: The tropical hypersurface of Example 8.1.1.

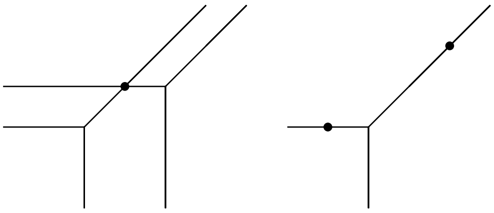

Figure 21: Two tropical lines intersecting in a point. A tropical line passing through two given points.

dimensional cones in the normal fan of the Newton polytope of f. Another way to say this is that $\omega \in \mathbb{R}^n$ is in $T(f)$ if and only if $\text{in}_\omega(f)$ is not a single term.

Since we are working with tropical polynomials we did not in fact define their initial forms. However, in tropical geometry, it is still fruitful to study polynomial ideals of usual polynomial rings. We make the following definition.

Definition 8.3.1 Let $I \subseteq k[x_1, \dots, x_n]$ be a polynomial ideal. We define the *tropical variety* of I to be

$$T(I) := \{\omega \in \mathbb{R}^n : \text{in}_\omega(I) \text{ contains no monomial}\}.$$

This set turns out to be closed. If I is homogeneous the cones of the Gröbner fan of I index all initial ideals of I. Therefore $T(I)$ is the support of a subfan of the Gröbner fan. This subfan we will also call the tropical variety of I.

Essentially we know how to compute Gröbner fans using the Gröbner walk. Therefore, we could in theory find all Gröbner cones of I. For each Gröbner cone we could check if $\text{in}_\omega(I)$ contains a monomial. That would give us a way to compute $T(I)$ as a polyhedral fan. One way to check if $\text{in}_\omega(I)$ contains a monomial is by computing the saturation $(\text{in}_\omega(I) : x_1 \cdots x_n^\infty)$ and checking if it is $\langle 1 \rangle$.

Example 8.3.2 Let $A \in \mathbb{C}^{2\times 5}$ be a matrix with *generic* (or random) entries. Let $p_{12}, \dots, p_{45}$ denote the ten 2×2 subdeterminants. Notice that $p_{12}p_{34} - p_{13}p_{24} + p_{14}p_{23} = 0$. Let $I_{2,5} \subseteq \mathbb{C}[p_{12}, \dots, p_{45}]$ denote the ideal generated by all such relations. The *tropical Grassmannian2,5* is the tropical variety $T(I_{2,5})$. This tropical variety is a polyhedral fan in $\mathbb{R}^{10}$. Using the method described above it is possible to compute that $T(I_{2,5})$ has a 5-dimensional lineality, ten 6-dimensional cones and fifteen 7-dimensional cones. (Analogously to classical Grassmannians, the points of $T(I_{2,5})$ define *tropical linear subspaces* of $\mathbb{R}^5$.)

We started this subsection by setting all coefficients equal to 0 to make tropical geometry match the Gröbner basis theory. Another approach is to introduce term orderings which do not only compare exponent vectors but also coefficients. That leads to a generalized notion of Gröbner bases which is better suited for tropical geometry. That and much much more is explained in the upcoming book by Diane Maclagan and Bernd Sturmfels:"Introduction to Tropical Geometry" `http://homepages.warwick.ac.uk/staff/D.Maclagan/papers/TropicalBook.html`

A Exercises

A.1 First sheet

Choose 4 of the following exercises and hand in solutions by February 5th (Wednesday).

1. Let $F \subseteq k[x_1, \ldots, x_n]$. Recall that we defined $\langle F\rangle := \{\sum_{i=1}^{m} g_i f_i : m \in \mathbb{N} \wedge g_i \in k[x_1, \ldots, x_n] \wedge f_i \in F\}$. Prove that $\langle F\rangle$ is an ideal. Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal. Prove that $F \subseteq I$ implies $\langle F\rangle \subseteq I$.
   We conclude that $\langle F\rangle$ is the (unique) *smallest ideal containing* F.

2. Prove that $I := \langle 2x^3y - 3x^5y^2, xy^2, 5x^6y + x^9, x^4y^2\rangle \subseteq \mathbb{Q}[x, y]$ is a monomial ideal. Draw its staircase diagram and find the unique minimal monomial generating set for I.

3. Let $f_1 := x^2 + y^2 - 1$ and $f_2 = x + y - 1$. Prove that $\{f_1, f_2\}$ is a minimal generating set for the ideal $I := \langle f_1, f_2\rangle \subseteq \mathbb{C}[x, y]$. Find a different minimal generating set for I.
   (Hint: To show that $\langle f_1, f_2\rangle \neq \langle f_2\rangle$ you can for example find a point in $V(\langle f_2\rangle)$ which is not in $V(\langle f_1, f_2\rangle)$.)

4. Let $I \subseteq k[x_1, \ldots, x_n]$ be a monomial ideal and let $f \in I$. Show that every term of f is in I.

5. Let $f = x^3 + y^3 + z^3 + xyz \in k[x, y, z]$. Draw the Newton polytope of f. Does there exist a vector $\omega \in \mathbb{R}^3$ such that $\text{in}_\omega(f) = xyz$? Does there exist a term order $\prec$ such that $\text{in}_\prec(f) = xyz$?

6. Let $n = 1$. Prove that there is only one term ordering on $k[x_1]$.

7. Let $\omega \in \mathbb{R}^n_{\geq 0}$. Let $\preceq$ be some term order on $k[x_1, \ldots, x_n]$. We define the relation $\preceq_\omega$ as follows: $x^u \preceq_\omega x^v \Leftrightarrow \omega \cdot u < \omega \cdot v \vee (\omega \cdot u = \omega \cdot v \wedge x^u \preceq x^v)$. Prove that $\preceq_\omega$ is a term order.

8. Let $n = 2$. Use the exercise above to construct infinitely (or even uncountably) many *different* term orderings on $k[x_1, \ldots, x_n]$.

9. In this exercise we prove Lemma 1.4.3. Let $\prec$ be a term ordering, $\omega \in \mathbb{R}^n$ and $f, g \in k[x_1, \ldots, x_n]$.
   - Prove that the ring $k[x_1, \ldots, x_n]$ is an integral domain and if $f \neq 0 \neq g$ then $\text{in}_\prec(fg) = \text{in}_\prec(f)\text{in}_\prec(g)$.
   - Using that $k[x_1, \ldots, x_n]$ is an integral domain, prove that $\text{in}_\omega(fg) = \text{in}_\omega(f)\text{in}_\omega(g)$.

10. Prove that $x^v|x^u$ if and only if $\forall i : v_i \leq u_i$. Prove that if $M \subseteq k[x_1, \ldots, x_n]$ is a set of monomials and $x^u \in \langle M\rangle$ then there exists $x^v \in M$ such that $x^v|x^u$. Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal, $\prec$ a term order and $x^u \in \text{in}_\prec(I)$. Prove that there exists $f \in I$ such that $x^u = \text{in}_\prec(f)$.

A.2 Second sheet

Choose 3 of the following exercises and hand in solutions by February 19th (Wednesday).

1. How is Algorithm 1.5.1 different from [11, Algorithm 1.3.4] in [11]? Do the two algorithms produce the same remainder? Do they produce the same remainder if $\{f_1, \dots, f_s\}$ is a Gröbner basis?

2. Using one of the free computer algebra systems Singular or Macaulay2 (or Risa-Asir, or CoCoA, or SAGE) compute a Gröbner basis for the ideal:

   $$I := \langle x^3 + 5xyz + xy + y^2 + z, y^3 + 4xz - x^2 + y, xz - yx + z^2\rangle \subseteq \mathbb{Q}[x, y, z]$$

   Both Singular and Macaulay2 are installed on the IMF machines (orc05,...) and are invoked by running the commands "Singular" or "M2" in a shell/terminal window.

3. Let $n = 1$ and $f, g \in k[x] \setminus \{0\}$. We already saw on the last sheet that there is a unique term ordering on $k[x]$. Prove that the reduced Gröbner basis of the ideal $\langle f, g\rangle$ has just a single element $h \in k[x]$. Prove that $h|f$ and $h|g$. Prove that if $p \in k[x]$ with $p|f$ and $p|g$ then $p|h$.

   (We call h the *greatest common divisor* of f and g. Usually it is computed by the Euclidean Algorithm for polynomials. This exercise shows that Buchberger's Algorithm is a generalisation of the Euclidean Algorithm.)

4. Using a computer algebra system, write a procedure/function which given a polynomial $f \in \mathbb{Q}[x_1, \dots, x_n]$ and vector $\omega \in \mathbb{Z}^n$ computes the initial form $\text{in}_\omega(f)$.

5. Let $I = \langle x^3 - y^2 + 1, x^4 - xy + 2x + 2\rangle$ be an ideal in $\mathbb{Q}[x, y, z]$. Prove, for example using a computer algebra system, that $[xy^2 + x] \neq [xy + 5]$ holds in the quotient ring $\mathbb{Q}[x, y, z]/I$.

   We let $\mathbb{Z}/7\mathbb{Z}$ be the field with 7 elements. Let $J = \langle x^3 - y^2 + 1, x^4 - xy + 2x + 2\rangle$ be an ideal in the polynomial ring $(\mathbb{Z}/7\mathbb{Z})[x, y, z]$. Prove that $[xy^2 + x] = [xy + 5]$ holds in the quotient ring $(\mathbb{Z}/7\mathbb{Z})[x, y, z]/J$.

6. Using Lemma 1.7.6, make a clever choice of $\prec$ and find a Gröbner basis of the ideal in Exercise 2 with respect to $\prec$ by hand. (Hint: Exercise 7 on the first sheet is useful.)

7. Let $\prec$ be a term ordering and $f \in k[x_1, \dots, x_n] \setminus \{0\}$. Use Lemma 1.4.3 to prove that for $I := \langle f\rangle$ we have $\text{in}_\prec(I) = \langle \text{in}_\prec(f)\rangle$. Use this to prove that $\{f\}$ is a Gröbner basis for I with respect to $\prec$.

8. Prove that for any term ordering $\prec$ and any $f \in k[x_1, \dots, x_n] \setminus \{0\}$ we have $S_\prec(f, f) = 0$. Use Theorem 1.7.2 or Algorithm 1.7.3 to prove that $\{f\}$ is Gröbner basis for $I := \langle f\rangle$ with respect to $\prec$.

A.3 Third sheet

Do three of the exercises below. Hand in solutions by March 5th.

1. Solve the system
$$9z^2 - 17z + 2y^2 = 0$$
$$xy^2 - 9x + 5xz = 0$$
$$x^2 - z^2 + 4x - 2xz = 0$$
over $\mathbb{C}$ using Gröbner bases (and a computer algebra system).

2. Let k be a field. Let $p_1, p_2 \in k^n$ with $p_1 \neq p_2$. Prove that there exists a polynomial $f \in k[x_1, \ldots, x_n]$ such that $f(p_1) \neq f(p_2)$. Prove that there exists $g \in k[x_1, \ldots, x_n]$ such that $g(p_1) = 1$ and $g(p_2) = 0$. Let $q_1, \ldots, q_m \in k^n$ be all different. Prove that there exists $h \in k[x_1, \ldots, x_n]$ such that $h(q_1) = 1$ and $h(q_i) = 0$ for all $i \neq 1$.

3. Let $I = \langle y^5 - y^2 + z, -x^3 + y^6 - y^3 - 1, xy - 1, x^4 + x + y^2 - y^5 \rangle \subseteq \mathbb{C}[x, y, z]$ be an ideal. Choose a term ordering $\prec$. Compute the initial ideal $\text{in}_{\prec}(I)$. Find $\text{std}_{\prec}(I)$. Use the proof of Corollary 1.8.6 to give an upper bound on the number of points in $V(I) \subseteq \mathbb{C}^3$.

4. Compute a lattice basis of the lattice kernel $\ker(A) \cap \mathbb{Z}^4$ for the matrix
$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$
Complete the basis to a lattice basis of $\mathbb{Z}^4$.

5. Compute the reduced row echelon form of the following matrix over $\mathbb{Q}$:
$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 7 & 10 & 13 \end{array}\right].$$
What is the reduced lexicographic Gröbner basis of $\langle x + y + z - 1, x + 2y + 3z - 4, 4x + 7y + 10z - 13 \rangle \subseteq \mathbb{Q}[x, y, z]$?

6. Without using Section 2.1 (other than possibly Theorem 2.1.2) prove that if the groups $\mathbb{Z}^n$ and $\mathbb{Z}^m$ are isomorphic then $n = m$. This shows that the notion of rank is well-defined in Definition 2.1.1. (Hint: linear algebra over $\mathbb{R}$, or observe that the proof of Theorem 2.1.2 does not use the notion of rank, and apply it.)

7. Can every term ordering $\prec$ on monomials in $k[x_1, \ldots, x_n]$ be represented by a matrix $A \in \mathbb{Q}^{d \times n}$?

8. Prove Lemma 2.2.4 without using the trigonometric functions Arg, sin and cos.

B Suggested projects

The most difficult projects have been marked with a *.

Gröbner bases over $\mathbb{Z}$ Buchberger's Algorithm is a generalization of Gauss Elimination. In Section 2.1 we saw that row reduction can be done over the integers. Similarly Buchberger's Algorithm can be improved to $\mathbb{Z}[x_1, \dots, x_n]$. See [2, Chapter 10.1].

Hilbert functions What is the Hilbert function of a homogeneous ideal? Why is it a polynomial for large degree? What is the Hilbert series?

Hilbert driven Buchberger How is it an advantage to know the Hilbert function when computing a Gröbner basis?

Gröbner basis conversion FGLM An alternative to the Gröbner walk for changing a Gröbner basis from one ordering to another is the FGLM procedure. See [6] and [3, page 49-56].

Computing the ideal of a finite set of points How does one compute the ideal of polynomials vanishing on a finite set of points?

Tropical geometry What is tropical geometry, and how does in relate to the Gröbner fan?

LLL reduction Sometimes a "small" lattice basis is desirable. Such basis can be computed with the Lenstra Lenstra Lovasz algorithm. One application: I have chosen a polynomial $f \in \mathbb{C}[x, y, z]$ of degree 2 with small coefficients defining a hypersurface $V(f)$. The approximate point $(-1.85395720454315454, -0.95734675483425443 4, 0.74075744188757582084)$ is on the variety. Which polynomial did I choose?

Short rational functions* We have seen in Example **??** that a reduced Gröbner basis for a toric ideal I_A can easily be exponential in size of the bit encoding of A - even when the dimensions of the matrix are fixed. In the paper [5] the authors claim that Gröbner bases for toric ideals can be computed in polynomial time for fixed dimension. What do they mean?

Eigenvalues or Sturm sequences The first step of solving a system of polynomial equations is to compute a Gröbner basis. If $V(I) \subseteq \mathbb{C}^n$ is a finite set then the next step is to compute the eigenvalues of the companion matrix. If real solutions are required then Sturm sequences are a useful tool. Polynomial system solving using eigenvalues is the topic of [3, page 56-69]. Solving polynomial systems over the reals is the topic of [3, page 69-76]. The last exercise on page 76 is to prove Sturm's theorem.

Local orderings* A local ordering is a term ordering where 1 is not necessarily the smallest monomial. Gröbner bases for these orderings are called standard bases. They are generators for ideals in localized polynomial rings. Their construction relies on the more complicated "normal form algorithm" by Mora.

Gröbner bases for modules* An ideal $I \subseteq k[x_1, \dots, x_n]$ is a $k[x_1, \dots, x_n]$-module. Gröbner bases can be defined and computed for submodules of the free module $(k[x_1, \dots, x_n])^m$. See [1, page 140-152].

Primary decomposition of monomial ideals How can one read off a primary decomposition of a staircase diagram?

Generic initial ideals* What is a generic initial ideal? How do we compute it? What is the generic Gröbner fan?

A vector interpretation of Buchberger's algorithm for toric ideals We have already seen that Gröbner bases of toric ideal are generated by binomials. It is is possible to describe Buchberger's algorithm purely using vectors in $\mathbb{Z}^n$.

Gröbner bases with p-adic valuation* Fix a prime p. The p-adic valuation on $\mathbb{Q}$ can be used in the definition of Gröbner bases.

Comparing algorithms for computing toric ideals We have already seen one algorithm for computing for computing toric ideals. Describe the DiBiase-Urbanke Algorithm, implement it (in Singular?), and compare running times.

Gebauer Möller Criteria* What is the Gebauer Möller criteria for eliminating S-polynomials in Buchberger's algorithm, and why does it work?

The integer programming gap* How do we use Gröbner bases to estimate the difference between the optimal value for a in integer programming problem and its LP-relaxation?

Universal Gröbner bases Given a set of polynomials, how do we use Newton polytopes to check if it is a Gröbner basis with respect to any term order? How do we check if there exists a term ordering which it is a Gröbner basis with respect to?

Hilbert's Nullstellensatz We have already used Hilbert's Nullstellensatz when solving polynomial systems. But how does the proof go?

C Notation and conventions

- $\mathbb{N} = \{0, 1, 2, \ldots\}$.
- $x^u = x_1^{u_1} x_2^{u_2} \cdots x_n^{u_n}$ for a vector $u \in \mathbb{Z}^n$.
- $\mathbb{Z}^{d\times n}$ - the set of $d \times n$ matrices with entries in $\mathbb{Z}$.
- $A_{i\cdot}$ - the ith row of a matrix A.
- $A_{\cdot j}$ - the jth column of a matrix A.
- $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$.
- For $U \subseteq \mathbb{R}^n$ the orthogonal complement $U^{\perp} := \{x \in \mathbb{R}^n : \forall y \in U : x \cdot y = 0\}$.
- $\mathcal{G}_{\prec}(I)$ is the reduced Gröbner basis of I w.r.t. $\prec$.
- For $u \in \mathbb{Z}^n$ we define $u^+ \in \mathbb{N}^n$ with $u_i^+ := \max(u_i, 0)$.
- For $u \in \mathbb{Z}^n$ we define $u^- \in \mathbb{N}^n$ with $u_i^- := \max(-u_i, 0)$.
- For $u, v \in \mathbb{Z}^n$ define $u \wedge v$ and $u \vee v$ as follows: $(u \wedge v)_i := \min(u_i, v_i)$ and $(u \vee v)_i := \max(u_i, v_i)$.
- $p_u := x^{u^+} - x^{u^-}$

We use the following conventions. If we apply an associative operation to zero operands we get the neutral element for that operation. For example:

- If we are summing the real numbers in a finite set B, and B happens to be empty, then $\sum_{a\in B} a = 0$, the neutral element for addition in $\mathbb{R}$.
- If we make a union of 0 sets, then we get the empty set $\bigcup_{a\in\emptyset} a = \emptyset$.
- Let B be a set of subsets of $\mathbb{R}^n$, then $\bigcap_{a\in B} a$ is $\mathbb{R}^n$ if $B = \emptyset$.

There is good reason for such a convention. We all remember that a d-dimensional linear subspace $U \subseteq \mathbb{R}^n$ has a basis with d elements. This means that the subspace $\{0\}$ has the empty set as a basis. The set $\{0\}$ would not work as a basis because this set is linearly dependent.

Similarly, in the case of ideals in $k[x_1, \ldots, x_n]$, we have $\langle\emptyset\rangle := \{\sum_{i=1}^m g_i f_i : m \in \mathbb{N} \wedge g_i \in k[x_1, \ldots, x_n] \wedge f_i \in \emptyset\} = \{0\}$ because $m = 0$ is the only choice (otherwise we cannot pick f_i).

D Software introductions

If you are familiar with the text editor Emacs, it is a good idea run the command line software from there. After having started Emacs, hold down the META key (ESC or ALT) and press “x”. Now type “shell” and press ENTER. You now have a working shell with the advantage that you can edit the buffer as any other Emacs buffer and press CTRL-UP to repeat you input. Some systems such as Singular already has this feature built in.

D.1 Singular

Singular is a free Computeralgebra system for computing with polynomials. Singular is specialized in the area of singularity theory and therefore handles local rings and local term orderings which was not covered in this course. The core Singular developer team is located at the Technical University of Kaiserslautern, Germany. Contributors are spread over the world.

We start Singular by typing “Singular” in the shell. To illustrate how the software works we compute the Gröbner basis of Example 1.6.4 by typing

```
ring r=0,(x,y),dp;
ideal I=x2+y2+x2y, x2+xy+x2y;
std(I);
```

and get the result:

```
_[1]=xy-y2
_[2]=y3+x2+y2
_[3]=x3+x2+y2
```

The first line of our input sets up the polynomial ring r. We provide three kinds of information: the characteristic of the ring (we just choose 0 for $\mathbb{Q}$), the variable names, and finally we specify the term order “dp” which means the graded reverse lexicographic ordering. The second line specifies an ideal I by listing a set of generators. In the third line we compute a Gröbner basis of I using the command “std”.

If we want to make sure that Singular computes the reduced Gröbner basis, we need to run the command:

```
option(redSB);
```

before computing a Gröbner basis.

To compute the remainder of a polynomial by division with the Gröbner basis, we first store the Gröbner basis and use the command “reduce”:

```
ideal G=std(I);
reduce(xy3+1,G);
2x2+2y2+1
```

Other term orders can be chosen:

```
ring s=0,(x,y),wp(1,3);
ring t=0,(x,y),lp;
```

Here the first ring uses a term ordering induced by a vector (and tie-broken reverse lexicographically). The second ring uses the lexicographic ordering.

From our viewpoint it is unnatural to specify the term ordering at the same time as the polynomial ring. To Singular, however, it is important, because if the ordering is not "global" (meaning 1 is not smallest), then Singular does not complain but computes in a *localization* of the usual polynomial ring.

As the reader might have noticed, Singular uses the C programming language syntax and does indeed contain a complete programming language. More information can be found at the Singular webpage `http://www.singular.uni-kl.de` and the online manual.

D.2 Polymake

Polymake is a free software system for computing with polyhedra. Often Polymake does not do the computations itself, but hands over the data to different kinds of specialized software packages. Polymake provides a uniform, transparent interface to these packages. Polymake is developed in Berlin and Darmstadt in Germany. It does not run on MS-Windows. But one can use the online interface `http://polymake.org/doku.php/boxdoc`

In the following we do a little session with the Polymake type Polytope. The name of the data type "Polytope" is misleading: A Polymake Polytope can be unbounded. We construct a polyhedron in $\mathbb{R}^3$ generated by points and rays.

```
polytope > $p=new Polytope(POINTS=>[[0,2,2,4],[1,1,1,0],[1,1,0,1]]);
```

The first coordinate is special: 1 means a point used for the convex hull in the statement of Theorem 3.2.6, while 0 means a cone generator of the construction.

We make another polyhedron by specifying the defining A matrix and b vector. In our notation INEQUALITIES= $\begin{bmatrix} b & -A \end{bmatrix}$ and we input the rows:

```
polytope > $q=new Polytope(INEQUALITIES=>[[2,-3,1,1],[-2,3,-1,-1],
                               [1,2,4,-4],[1,-4,4,2],[10,1,1,1]]);
```

We can now compute the Minkowski sum.

```
polytope > $r=minkowski_sum($p,$q);
```

We can ask for the rays and vertices:

```
polytope > print $r->VERTICES;
1 13/6 1/3 13/6
1 -1 7/2 -21/2
0 1 1 2
0 0 1 -1
```

And ask for a minimal set of defining inequalities and equations:

```
polytope > print $r->FACETS;
1 1 0 0
1 0 0 0
11/6 -1 1 0
-5/2 1 1 0
polytope > print $r->LINEAR_SPAN;
4 -3 1 1
```

Polymake uses the PERL programming language, allowing users to extend the system with whatever functions they like.

D.3 Gfan

Gfan is a free software system for computing Gröbner fans. On the Gfan webpage `http://home.imf.au.dk/jensen/software/gfan/gfan.html` the software and a manual can be found. The software is also installed on the IMF machines (orc05,...). After logging in, you can type the following

```
/home/jensen/bin/gfan _bases
```

followed by

```
Q[x,y,z]
{x^5 + z^2 + y^3 - 1, y^2 + z + x^2 - 1, z^3 + y^5 + x^6 - 1}
```

to compute all 360 reduced Gröbner bases of the ideal in Example 4.0.22. If you want to compute the Gröbner fan instead, type

```
/home/jensen/bin/gfan _groebnerfan
```

followed by

```
Q[x,y,z]
{x^5 + z^2 + y^3 - 1, y^2 + z + x^2 - 1, z^3 + y^5 + x^6 - 1}
```

. The output explains how the rays of the fan are combined to form all cones.

E Exam topics

At the exam you will be assigned one of the 6 topics below at random. You will have 25 minutes to prepare (alone) after having drawn your topic. After this the 20 minutes exam will start. You should present your topic for about 12-14 minutes. It is good to include (a sketch of) a proof, but definitions and examples are at least as important. You should be prepared to also say something about your project and answer questions about your project and other topics.

Robbiano's characterization of term orderings Suggested things to explain: How a matrix represents a term ordering. Example. What Robbiano's theorem says (Theorem 2.3.5). Proof idea: extend to Laurent monomials, consider convex set X not containing 0, how we get the first row of the matrix A by applying Lemma 2.2.5.

Fourier-Motzkin elimination and its consequences Suggested thing to explain: What a polyhedron is. How we compute the projection of a polyhedron using the Fourier-Motzkin algorithm (Algorithm 3.1.3). Maybe very small example. Explain one of the applications: Theorem 3.1.6, Theorem 3.1.10 or Proposition 3.1.12.

The Gröbner fan of an ideal Suggested things to explain: (What a fan is.) What the notation $C_\prec(I)$ means. Definition 4.3.1. There are many different things one could explain now. But there is little time, so *one needs to choose.* Here is something to choose from:

- The set $C_\prec(I)$ is a polyhedral cone. This follows from Lemma 4.2.1.
- The Gröbner walk. What the idea of Section 5.9 is.
- There are only finitely many initial ideals $\text{in}_\prec(I)$ of an ideal I and hence only finitely many cones in the Gröbner fan (Proposition 4.1.1).

Lattice ideals Suggested things to explain: What a lattice is. What a generating set for a lattice is. What a lattice ideal is. How one computes a generating set for the lattice ideal if a generating set for the lattice is known (Proposition 6.2.8). What the notation $(I : (x_1 \cdots x_n)^\infty)$ means. What Proposition 6.1.3 says. Maybe: Why any reduced Gröbner basis of a lattice ideal consists of binomials. (Notice lattices were defined in Section 2.1 and lattice ideals in Section 6.2.)

Toric ideals and integer programming Suggested things to explain: What the toric ideal of a matrix $A \in \mathbb{N}^{d\times n}$ is. Why it is a lattice ideal. What Fiber(A) is. What Fiber$_\prec(A)$ is. Why it has no cycles and a unique sink Proposition 6.4.7. Why this graph "solves" the minimization problem in (4), page 73 (Lemma 6.4.8). Example.

Regular triangulations and secondary fans Suggested things to explain: How a (generic) vector ω defines a triangulation of a vector configuration A. An example. What the secondary fan is. What a Stanley-Reisner ideal of an abstract simplicial complex is. What Sturmfels' Theorem says. What the idea of the proof is.

References

[1] W.W. Adams and P. Loustaunau. *An introduction to Gröbner bases.* Graduate studies in mathematics. American Mathematical Society, 1994.

[2] Thomas Becker and Volker Weispfenning. *Gröbner bases*, volume 141 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993. A computational approach to commutative algebra, In cooperation with Heinz Kredel.

[3] D.A. Cox, J.B. Little, and D. O'Shea. *Using algebraic geometry.* Graduate texts in mathematics. Springer, 1998.

[4] David Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms.* Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1997. An introduction to computational algebraic geometry and commutative algebra.

[5] Jesus De Loera, David Haws, Raymond Hemmecke, Peter Huggins, Bernd Sturmfels, and Ruriko Yoshida. Short rational functions for toric algebra and applications, 2003.

[6] J. C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *J. Symbolic Comput.*, 16(4):329–344, 1993.

[7] Komei Fukuda, Anders Jensen, and Rekha Thomas. Computing Gröbner fans. *Mathematics of Computation*, 76:2189–2212, 2007, math.AC/0509544.

[8] Anders Nedergaard Jensen. Computing Gröbner fans of toric ideals, 2002. Master thesis, University of Aarhus.

[9] Anders Nedergaard Jensen. Algorithmic aspects of Gröbner fans and tropical varieties, 2007. PhD thesis, University of Aarhus.

[10] Niels Lauritzen. *Concrete abstract algebra.* Cambridge University Press, Cambridge, 2003. From numbers to Gröbner bases.

[11] Diane Maclagan and Rekha R. Thomas. *Computational Algebra and Combinatorics of Toric Ideals.* `http://www.math.rutgers.edu/~maclagan/india/india.html`.

[12] Teo Mora and Lorenzo Robbiano. The Gröbner fan of an ideal. *J. Symbolic Comput.*, 6(2-3):183–208, 1988. Computational aspects of commutative algebra.

[13] Bernd Sturmfels. *Gröbner bases and convex polytopes*, volume 8 of *University Lecture Series.* American Mathematical Society, Providence, RI, 1996.