

STURMFELS EXPANDED

Niels Lauritzen

The author's address:

NIELS LAURITZEN

DEPARTMENT OF MATHEMATICAL SCIENCES

UNIVERSITY OF AARHUS, DENMARK

EMAIL: niels@imf.au.dk

URL: <http://www.imf.au.dk/~niels/>

Contents

- Preface** **v**

- 1 Gröbner Basics** **1**
 - 1.1 The universal Gröbner basis 2
 - 1.2 Homogeneous ideals 3
 - 1.3 Term orders and weight vectors 3
 - 1.3.1 Computing $\text{in}_\omega(I)$ 4
 - 1.4 Separation 5
 - 1.5 Initial ideals using weight vectors 5
 - 1.6 The Gröbner region of an ideal 6

- 2 The Gröbner walk** **9**
 - 2.1 Prelude on term orders 9
 - 2.2 Recollections 10
 - 2.3 A somewhat different Gröbner fan 10
 - 2.4 How to compute the Gröbner fan in small cases 11
 - 2.5 Wall crossing 11

- 3 Hilbert functions and best term order** **13**
 - 3.1 Hilbert functions 13
 - 3.2 Is something already a Gröbner basis? 14

- 4 LLL** **15**
 - 4.1 Lattices 15
 - 4.2 Naive search for a shortest vector 15
 - 4.3 The algorithm of Gauss in the plane 16
 - 4.4 Gram-Schmidt 17
 - 4.5 Idea of the LLL-algorithm 17

- 5 Toric ideals** **21**
 - 5.1 Computing kernels 21

5.2	Computing images	23
5.2.1	An application	24
5.3	Toric homomorphisms	24
5.4	Kernels of toric homomorphisms	25
5.5	Computing the kernel of a toric homomorphism	26
5.5.1	The Di Biase - Urbanke algorithm	28
A	Review of convexity	29
A.1	Separation	29
A.2	Polyhedra	30
A.3	Minkowski sums	30

Preface

These notes contain explanatory material for Sturmfels' book "Gröbner bases and convex polytopes". They cover the pages 1–7, 22 (Subroutine 3.7), 25, 31, 32, 113, 114, 115 with background material on the LLL algorithm and lattice reduction. There is an incomplete appendix on convexity. A future project is to write this out containing a survey of details from the Springer GTM book of Arne Brøndsted with a proof of separation coming from dimension 2 and using induction.

Chapter 1

Gröbner Basics

Let $R = k[X_1, \dots, X_n] = \{\sum_{v \in \mathbb{N}^n} a_v X^v \mid a_v \neq 0 \text{ for only finitely many } v \in \mathbb{N}^n\}$, where k is a field. Let \leq be a term order. Given an ideal I we let $\text{in}_{\leq}(I)$ denote the ideal

$$\langle \text{in}_{\leq}(f) \mid f \in I \rangle.$$

Recall that a finite set $G = (g_1, \dots, g_m) \subseteq I$ is called a Gröbner basis for I if

$$\text{in}_{\leq}(I) = \langle \text{in}_{\leq}(g_1), \dots, \text{in}_{\leq}(g_m) \rangle.$$

In this case it is easy to prove that $I = \langle g_1, \dots, g_m \rangle$. If G is a Gröbner basis, then $f \in I \iff f^G = 0$, where f^G denotes the remainder of f divided by G using the division algorithm. We have the following very useful result.

Proposition 1.0.1 The set $S = \{[X^v] \mid v \in \mathbb{N}^n, X^v \notin \text{in}_{\leq}(I)\}$ is a k -vector space basis of

$$A = k[X_1, \dots, X_n]/I.$$

Proof. Let $\lambda_1, \dots, \lambda_r \in k$. Suppose that $\lambda_1[X^{v_1}] + \dots + \lambda_r[X^{v_r}] = 0$ for suitable $v_1, \dots, v_r \in \mathbb{N}^n$, such that $X^{v_i} \notin \text{in}_{\leq}(I)$. Then $f = \lambda_1 X^{v_1} + \dots + \lambda_r X^{v_r} \in I$. If $f \neq 0$, then $\text{in}_{\leq}(f) \in \text{in}_{\leq}(I)$. This is a contradiction. Thus $\lambda_1 = \dots = \lambda_r = 0$. Let $G = (g_1, \dots, g_r)$ be a Gröbner basis of I with respect to \leq . Then $\text{in}_{\leq}(I) = \langle \text{in}_{\leq}(g_1), \dots, \text{in}_{\leq}(g_r) \rangle$. Given $[f] \in A$, we see that $[f] = [f^G]$ and f^G is a sum of terms not in $\text{in}_{\leq}(I)$. This shows that S generates A . \square

Exercise 1.0.2 Suppose that \leq_1 and \leq_2 are two term orders and I an ideal of R . Show that $\text{in}_{\leq_1}(I) = \text{in}_{\leq_2}(I)$ if $\text{in}_{\leq_1}(I) \subseteq \text{in}_{\leq_2}(I)$.

1.1 The universal Gröbner basis

Proposition 1.1.1 Let $n \geq 2$. There are uncountably many term orders on \mathbb{N}^n .

Proof. Fix $n = 2$. Let \leq_r denote the term order on \mathbb{N}^2 given by $u \leq_r v$ if and only if

$$(1, r) \cdot u < (1, r) \cdot v \text{ or } ((1, r) \cdot u = (1, r) \cdot v \text{ and } u \leq_{\text{lex}} v),$$

where $r \in \mathbb{R}$ and $r \geq 0$ (why is $r < 0$ NOT allowed?). If $r \neq s$ we may find $v \in \mathbb{Z}^2$ such that $(1, r) \cdot v > 0$ and $(1, s) \cdot v < 0$. Now write $v = v_1 - v_2$, where $v_1, v_2 \in \mathbb{N}^2$. Then $v_1 \geq_r v_2$ but $v_1 \leq_s v_2$. This shows that $\leq_r \neq \leq_s$ for $r \neq s$. \square

Theorem 1.1.2 Let I be an ideal of R . Then

$$T = \{ \text{in}_{\leq}(I) \mid \leq \text{ term order} \}$$

is a finite set.

Proof. Suppose that T is infinite. Let $f \in I \setminus \{0\}$. Then f contains a term m_1 belonging to an infinite subset T_1 of initial ideals in T . We can find an initial ideal $\text{in}_{\leq_1}(I)$ in T_1 , such that $\langle m_1 \rangle \subsetneq \text{in}_{\leq_1}(I)$. This means that there exists a term $m \in \text{in}_{\leq_1}(I) \setminus \langle m_1 \rangle$, such that

$$[m] = \lambda_1[n_1] + \cdots + \lambda_r[n_r],$$

where n_1, \dots, n_r are terms $\notin \langle m_1 \rangle$ by Proposition 1.0.1. Thus we may find $f_1 \in I \setminus \{0\}$ with no terms in $\langle m_1 \rangle$. Let m_2 denote a term belonging to an infinite subset of initial ideals T_2 in T_1 . As before we may find $f_2 \in I \setminus \{0\}$ with no terms in $\langle m_1, m_2 \rangle$. Continuing like this we get a infinite ascending chain $\langle m_1 \rangle \subsetneq \langle m_1, m_2 \rangle \subsetneq \dots$. This is a contradickson. \square

Definition 1.1.3 Let I be an ideal of R . A universal Gröbner basis for I is a (finite) subset $G \subseteq I$, which is a Gröbner basis for every term order \leq .

Corollary 1.1.4 Let I be an ideal of R . Then I has a universal Gröbner basis.

Proof. Suppose that $\text{in}_{\leq_1}(I), \dots, \text{in}_{\leq_m}(I)$ are the finitely many initial ideals. Then the union of Gröbner bases wrt. \leq_1, \dots, \leq_m is a universal Gröbner basis. \square

Example 1.1.5 The polynomials coming from taking the 2×2 -minors of a $2 \times n$ -matrix

$$\begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \end{pmatrix}$$

form a universal Gröbner basis for the ideal they generate. Take as an example $n = 3$. Then

$$G = \{X_{11}X_{22} - X_{21}X_{12}, X_{11}X_{23} - X_{21}X_{13}, X_{12}X_{23} - X_{22}X_{13}\}$$

is a universal Gröbner basis for $\langle G \rangle$. This is seen by analyzing the 8 possible initial terms in the three polynomials and using Buchbergers S -criterion.

1.2 Homogeneous ideals

Let $(S, +)$ be a commutative semigroup with neutral element 0. A ring R is called S -graded if

$$R = \bigoplus_{s \in S} R_s,$$

where R_s are subgroups of $(R, +)$, $1 \in R_0$ and $R_s R_t \subseteq R_{s+t}$. Notice that R_0 is a subring of R .

An element $r \in R$ is called homogeneous if $r \in R_s$ for some $s \in S$. An ideal $I \subseteq R$ is called homogeneous if $I = \bigoplus_{s \in S} I \cap R_s$. This can be translated into the fact that if $f = f_{s_1} + f_{s_2} + \cdots + f_{s_r} \in I$, where $f_{s_i} \in R_{s_i}$, then $f_{s_i} \in I$.

Proposition 1.2.1 An ideal $I \subseteq R$ is homogeneous if and only if it can be generated by homogeneous elements.

Proof. Exercise. \square

Example 1.2.2 Let $\omega = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$. Then $S = \mathbb{N}\omega_1 + \cdots + \mathbb{N}\omega_n \subseteq \mathbb{R}$ is a semigroup. There is a natural S -grading on the polynomial ring $R = k[X_1, \dots, X_n]$ given by

$$R_s = \text{Span}_k \{X^v \mid v \in \mathbb{N}^n, v \cdot \omega = s\},$$

where $s \in S$.

1.3 Term orders and weight vectors

Let $\omega \in \mathbb{R}^n$ and $f = \sum_{v \in \mathbb{N}^n} a_v X^v \in R$. Then we let

$$\text{in}_\omega(f) = \sum_{v' \in \mathbb{N}^n} a_{v'} X^{v'},$$

where we sum over the vectors v' maximizing $\omega \cdot v'$ in $\{v \mid a_v \neq 0\}$. Given an ideal $I \subseteq R$, we let

$$\text{in}_\omega(I) = \{\text{in}_\omega(f) \mid f \in I\}.$$

Clearly $\text{in}_\omega(I)$ is a homogeneous ideal in the grading given by ω (as in Example 1.2.2). Notice that this is not necessarily a monomial ideal.

Example 1.3.1 Let $f = X_1^5 X_2 + X_1^4 X_2^4 + X_1^4 + X_1^2 X_2^5 + X_1 X_2^2 + X_2^6 + X_2$. Examples with $\omega = (1, 1), (1, 2)$. Determine initial ideals. Some terms can NEVER be initial terms!!

Definition 1.3.2 In the same way as in Proposition 1.1.1 we construct a term order \leq_ω for $\omega \geq 0$.

1.3.1 Computing $\text{in}_\omega(I)$

For a total order (not necessarily a term order) on monomials, we let $\text{in}_\leq(I) = \langle \text{in}_\leq(f) \mid f \in I \rangle$. For arbitrary $\omega \in \mathbb{R}^n$ (the condition $\omega \geq 0$ is dropped) we have a total order \leq_ω .

Proposition 1.3.3 Let \leq be a total order on monomials, $\omega \in \mathbb{R}^n$ and I an ideal in the polynomial ring $k[X_1, \dots, X_n]$. Then

$$\text{in}_\leq(\text{in}_\omega(I)) = \text{in}_{\leq_\omega}(I).$$

Proof. For every $f \in I$ we have the formula

$$\text{in}_\leq(\text{in}_\omega(f)) = \text{in}_{\leq_\omega}(f).$$

This shows that $\text{in}_{\leq_\omega}(I) \subseteq \text{in}_\leq(\text{in}_\omega(I))$. For the other inclusion, note that $\text{in}_\leq(\text{in}_\omega(I))$ is generated by $\text{in}_\leq(g)$, where $g \in \text{in}_\omega(I)$. We have to use that $\text{in}_\omega(I)$ is a homogeneous ideal. Write

$$g = g_{s_1} + \dots + g_{s_t}.$$

Then $\text{in}_\leq(g) = \text{in}_\leq(g_{s_i})$ for some $s_i \in \{s_1, \dots, s_t\}$. But g_{s_i} is a homogeneous element and as such it must be $= \text{in}_\omega(f)$, where $f \in I$. Thus

$$\text{in}_\leq(g) = \text{in}_\leq(g_{s_i}) = \text{in}_\leq(\text{in}_\omega(f)).$$

□

We can use Gröbner basis theory to compute the ideal $\text{in}_\omega(I)$. Let \leq be any term order and $G = \{g_1, \dots, g_r\}$ a Gröbner basis of I wrt. \leq_ω , where $\omega \geq 0$. Then

$$\{\text{in}_\omega(g_1), \dots, \text{in}_\omega(g_r)\}$$

is a Gröbner basis of $\text{in}_\omega(I)$ wrt. \leq . This is a consequence of Proposition 1.3.3.

1.4 Separation

Recall that a non-empty subset $C \subseteq \mathbb{R}^n$ is called convex if

$$\lambda x + (1 - \lambda)y \in C$$

for every $x, y \in C$ and $\lambda \in [0, 1]$. An affine hyperplane H in \mathbb{R}^n is given by

$$H = \{(v_1, \dots, v_n) \mid \lambda_1 v_1 + \dots + \lambda_n v_n = \alpha\},$$

where $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (not all zero) and $\alpha \in \mathbb{R}$. We associate the closed half spaces

$$H_- = \{(v_1, \dots, v_n) \mid \lambda_1 v_1 + \dots + \lambda_n v_n \leq \alpha\}$$

and

$$H_+ = \{(v_1, \dots, v_n) \mid \lambda_1 v_1 + \dots + \lambda_n v_n \geq \alpha\}$$

with H . Now we can state a fundamental separation theorem.

Theorem 1.4.1 Let C_1, C_2 be convex subsets of \mathbb{R}^n , such that $C_1 \cap C_2 = \emptyset$. Then there exists an affine hyperplane H such that $C_1 \subseteq H_-$ and $C_2 \subseteq H_+$.

Proof. This is a consequence of ([1], Theorem 11.3). \square

1.5 Initial ideals using weight vectors

Theorem 1.5.1 Let \leq be a term order and I an ideal of R . Then

$$\text{in}_\omega(I) = \text{in}_{\leq}(I)$$

for some vector $\omega \in \mathbb{N}^n$.

Proof. Let $G = \{g_1, \dots, g_r\}$ be a Gröbner basis for I wrt. \leq . Suppose we have found $\omega \in \mathbb{N}^n$, such that $\text{in}_{\leq}(g_1) = \text{in}_\omega(g_1), \dots, \text{in}_{\leq}(g_r) = \text{in}_\omega(g_r)$. Then

$$\text{in}_{\leq}(I) = \langle \text{in}_{\leq}(g_1), \dots, \text{in}_{\leq}(g_r) \rangle = \langle \text{in}_\omega(g_1), \dots, \text{in}_\omega(g_r) \rangle.$$

This proves that $\text{in}_{\leq}(I) \subseteq \text{in}_\omega(I)$. Therefore $\text{in}_{\leq}(I) \subseteq \text{in}_{\leq_\omega}(I)$ and Exercise 1.0.2 shows that $\text{in}_{\leq}(I) = \text{in}_{\leq_\omega}(I)$. If $\text{in}_{\leq}(I) \subsetneq \text{in}_\omega(I)$, then $\text{in}_{\leq}(I) \subsetneq \text{in}_{\leq_\omega}(I)$, so $\text{in}_{\leq}(I) = \text{in}_\omega(I)$.

Suppose now that

$$\begin{aligned} g_1 &= c_{10}X^{a_{10}} + c_{11}X^{a_{11}} + \dots + c_{1j_1}X^{a_{1j_1}} \\ g_2 &= c_{20}X^{a_{20}} + c_{21}X^{a_{21}} + \dots + c_{2j_2}X^{a_{2j_2}} \\ &\vdots \end{aligned}$$

where $\text{in}_{\leq}(g_1) = c_{10}X^{a_{10}}$, $\text{in}_{\leq}(g_2) = c_{20}X^{a_{20}}$, \dots . We must prove that

$$C = \{x \in \mathbb{R}_+^n \mid x \cdot (a_{i0} - a_{ij}) > 0, \text{ for } i = 1, \dots, r; j = 1, \dots, j_i\}$$

contains a $\omega \in \mathbb{N}^n$. Suppose that C is empty and let v_1, \dots, v_N denote the vectors $a_{i0} - a_{ij}$. Then the subset

$$M = \{(x \cdot v_1, \dots, x \cdot v_N) \mid x \in \mathbb{R}_+^n\}$$

is convex and disjoint from (the interior of) \mathbb{R}_+^n . Using Theorem A.1.1 we may find $\lambda_1, \dots, \lambda_N, \alpha \in \mathbb{R}$ such that

$$\lambda_1 x \cdot v_1 + \dots + \lambda_N x \cdot v_N \leq \alpha \leq \lambda_1 y_1 + \dots + \lambda_N y_N$$

for every $x \in \mathbb{R}_+^n$ and $(y_1, \dots, y_N) \in \mathbb{R}_+^n$. This shows that $\lambda_1, \dots, \lambda_N \geq 0$ and $\alpha = 0$. Thus

$$x \cdot (\lambda_1 v_1 + \dots + \lambda_N v_N) \leq 0$$

for every $x \in \mathbb{R}_+^n$. Therefore $\lambda_1 v_1 + \dots + \lambda_N v_N \leq 0$. Check that one may assume that $\lambda_1, \dots, \lambda_N \in \mathbb{N}$. Writing out we get the inequality

$$\lambda_1 a_{10} + \dots + \lambda_N a_{r0} \leq \lambda_1 a_{11} + \dots + \lambda_N a_{rj_r}.$$

This shows that

$$X^{\lambda_1 a_{10} + \dots + \lambda_N a_{r0}} \leq X^{\lambda_1 a_{11} + \dots + \lambda_N a_{rj_r}},$$

which is a contradiction as $X^{a_{10}}, \dots, X^{a_{r0}}$ are the initial terms. So we may find $\omega \in \mathbb{N}^n$ such that

$$\langle \text{in}_{\omega}(g_1), \dots, \text{in}_{\omega}(g_r) \rangle = \langle \text{in}_{\leq}(g_1), \dots, \text{in}_{\leq}(g_r) \rangle.$$

□

1.6 The Gröbner region of an ideal

Let I be an ideal of R . Then we let

$$GR(I) = \{\omega \in \mathbb{R}^n \mid \text{in}_{\omega}(I) = \text{in}_{\omega'}(I), \text{ for some } \omega' \in \mathbb{R}_+^n\}.$$

Definition 1.6.1 Let $f = a_1 X^{v_1} + \dots + a_r X^{v_r} \in R$. Then the Newton polytope of f is

$$\text{New}(f) = \text{conv}\{v_1, \dots, v_r\}.$$

Example 1.6.2 Let $I = \langle f \rangle$, where

$$f = X_1^4 + X_1^5 X_2^2 + X_1^4 X_2^4 + X_1^2 X_2^5 + X_2^6 + X_2 + X_1 X_2^2.$$

Then $GR(I)$ is a union of normal cones of certain faces of the Newton polytope of f . Show that there are only three terms of the above seven terms that can occur as initial terms. These correspond to the normal cones in the Gröbner region!

Proposition 1.6.3 Let $I \subseteq R$ be a homogeneous ideal with respect to the the vector $(d_1, \dots, d_n) \in \mathbb{N}^n$, where $d_i > 0$. Then $GR(I) = \mathbb{R}^n$.

Proof. Let $\omega \in R^n$. We wish to prove that $\text{in}_\omega(I) = \text{in}_{\omega'}(I)$ for some $\omega' \geq 0$. Choose λ such that $\omega' = \omega + \lambda(d_1, \dots, d_n) \geq 0$. The claim is that ω' does the job. Let $f = f_0 + f_1 + \dots + f_r$. Then $f_0, \dots, f_r \in I$ (since I is homogeneous) and

$$\text{in}_\omega(f) = \text{in}_\omega(f_{i_1}) + \dots + \text{in}_\omega(f_{i_s})$$

for suitable f_{i_1}, \dots, f_{i_s} . But $\text{in}_\omega(f_{i_j}) = \text{in}_{\omega'}(f_{i_j})$ for $j = 1, \dots, s$. This shows that $\text{in}_\omega(I) \subseteq \text{in}_{\omega'}(I)$. The other inclusion is similar. \square

Proposition 1.6.4 Let $I \subseteq R$ be a homogeneous ideal wrt. to the vector $(d_1, \dots, d_n) \in \mathbb{N}_{>0}^n$ and $\omega, \omega' \in \mathbb{R}^n$. Then

$$\text{in}_{\omega'}(\text{in}_\omega(I)) = \text{in}_{\omega + \epsilon\omega'}(I)$$

for $\epsilon > 0$ sufficiently small.

Proof. Fix a universal Gröbner basis $G = \{g_1, \dots, g_r\}$ for I . Then for $\epsilon > 0$ (sufficiently small) we have

$$\text{in}_{\omega'}(\text{in}_\omega(g)) = \text{in}_{\omega + \epsilon\omega'}(g),$$

where $g \in G$. This implies that

$$\text{in}_{\omega'}(\text{in}_\omega(I)) \supseteq \text{in}_{\omega + \epsilon\omega'}(I). \quad (*)$$

We wish to show that the right and left hand sides are identical. To this end we may assume that $\omega' \geq 0$ and $\omega + \epsilon\omega' \geq 0$ because of Proposition 1.6.3. Using \leq on both sides we get

$$\text{in}_{\leq\omega'}(\text{in}_\omega(I)) \supseteq \text{in}_{\leq\omega + \epsilon\omega'}(I).$$

But the left hand side is

$$\text{in}_{\leq\omega'}(I).$$

So we have inclusion between two initial ideals of I . This means that the inclusion in $(*)$ cannot be strict. \square

Chapter 2

The Gröbner walk

2.1 Prelude on term orders

Let $M = \{w_1, \dots, w_m\} \subseteq \mathbb{R}^n$. We may view M as an $m \times n$ matrix in the natural way. Then we define $u \leq_M v$ if $u \cdot w_i = v \cdot w_i$ for every $i = 1, \dots, m$ or

$$\begin{aligned} w_1 \cdot u &= w_1 \cdot v \\ &\vdots \\ w_l \cdot u &= w_l \cdot v \\ w_{l+1} \cdot u &< w_{l+1} \cdot v \end{aligned}$$

for some $l = 1, \dots, m - 1$. This defines a relation \leq_M on \mathbb{N}^n , which is reflexive and transitive. It is antisymmetric if $\text{Ker}(M) \cap \mathbb{Z}^n = \{0\}$. It is easy to see that

$$u \leq_M v \implies u + w \leq_M v + w,$$

for every $u, v, w \in \mathbb{N}^n$. Clearly $0 \leq_M v$ for every $v \in \mathbb{N}^n$ if $w_1, \dots, w_m \in \mathbb{R}_+^n$.

Project 2.1.1 Prove that every term order \leq on \mathbb{N}^n can be realized via a certain matrix M as above.

Example 2.1.2 We give some examples of term orders on \mathbb{N}^3 , which can be realized using suitable matrices.

i) Let

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then \leq_M is the usual lexicographic order on \mathbb{N}^3 .

ii) Let

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then \leq_M is the so-called degree reverse lexicographic order. It is also given by the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}.$$

2.2 Recollections

Recall the construction of the Gröbner fan of an ideal $I \subseteq k[x_1, \dots, X_n] = k[\mathbb{N}^n]$. We let

$$C[\omega] = \{v \in \mathbb{R}^n \mid \text{in}_v(I) = \text{in}_\omega(I)\}.$$

This is a cone in \mathbb{R}^n . If I is homogeneous with respect to $d = (d_1, \dots, d_n) \in \mathbb{N}^n$, where $d_i > 0$, then

$$C[\omega] \cap \mathbb{R}_+^n \neq \emptyset.$$

Thus $C[\omega]$ has nonempty intersection with the positive orthant. This is the content of Proposition 1.6.3. Ulf proved the following important result.

Theorem 2.2.1 The collection $\{\overline{C[\omega]} \mid \omega \in \mathbb{R}^n\}$ is a fan (called the Gröbner fan of the ideal I).

2.3 A somewhat different Gröbner fan

Suppose that I is an arbitrary ideal in $k[\mathbb{N}^n]$ (not necessarily homogeneous). Let

$$G[\omega] = \{v \in \mathbb{R}_+^n \mid \text{in}_v(I) = \text{in}_\omega(I)\}.$$

Then $\{\overline{G[\omega]} \mid \omega \in \mathbb{R}_+^n\}$ is a fan. If \leq is a term order we know that $\text{in}_{\leq}(I) = \text{in}_\omega(I)$ for $\omega \in \mathbb{N}^n$.

Remark 2.3.1 If $\text{in}_\omega(I)$ is a monomial ideal, then $G[\omega]$ is an open subset of \mathbb{R}_+^n . We have already seen this using a Gröbner basis argument. Let $G = \{g_1, \dots, g_t\}$ be a Gröbner basis with respect to a term order \leq . Write $g_i = x^{\alpha_i} + \sum_{\beta} c_{i,\beta} x^\beta$. Then we have seen that $\text{in}_\omega(I) = \text{in}_{\leq}(I)$ if and only if $\omega \cdot \alpha_i > \omega \cdot \beta$ (when $c_{i,\beta} \neq 0$). This condition defines an open subset of \mathbb{R}^n .

Definition 2.3.2 If $\text{in}_\omega(I)$ is a monomial ideal, ω is called *generic*.

2.4 How to compute the Gröbner fan in small cases

The strategy for computing the modified Gröbner fan in small cases is quite obvious (once you have spent a few days with it;-)). Compute Gröbner bases for some well known term orders \leq . For each term order we can compute the set of $\omega \in \mathbb{R}_+^n$ such that $\text{in}_{\leq}(I) = \text{in}_{\omega}(I)$. They form an open cone $G[\omega]$. The closures $\overline{G[\omega]}$ are the maximal cones in the Gröbner fan. Ulf explained how the whole fan is given using the maximal cones in it. It is time for an example.

Example 2.4.1 Let $I = \langle X^2 + Y, X^2Y + 1 \rangle$ (have you seen this example before?). Consider two term orders $\leq_1 =$ lexicographic order with $X \geq Y$ and $\leq_2 =$ lexicographic order with $Y \geq X$. The reduced Gröbner basis with respect to \leq_1 is $R_1 = (X^2 + Y, Y^2 - 1)$ and $\text{in}_{\leq_1}(I) = (X^2, Y^2)$. The reduced Gröbner basis with respect to \leq_2 is $R_2 = (X^2 + Y, X^4 - 1)$ and $\text{in}_{\leq_2}(I) = (Y, X^4)$.

Now face the facts! It is easy to realize that $\text{in}_{\omega}(I) = (X^2, Y^2)$ if and only if $2x > y$, where $\omega = (x, y)$. Similarly it follows that $\text{in}_{\omega}(I) = (Y, X^4)$ if and only if $2x < y$. We have our maximal cones and we are happy. Thus

$$\mathbb{R}_+^2 = \overline{G_1} \cup \overline{G_2}$$

where $G_1 = \{(x, y) \in \mathbb{R}_+^2 \mid 2x > y\}$ and $G_2 = \{(x, y) \in \mathbb{R}_+^2 \mid 2x < y\}$. This should be sketched in a drawing!! At the same time we have proved that

$$(X^2 + Y, Y^2 - 1, X^4 - 1)$$

is a **universal Gröbner basis** for I .

2.5 Wall crossing

Sometimes it is impossible to compute the Gröbner basis with respect to a specific term order \leq . In many cases the big obstacle is the lexicographic order. This is usually a difficult term order for computing Gröbner bases. What if we could start at some other point of the Gröbner fan and move towards the cone containing the lexicographic order? Each time we cross a wall we do some update so when reaching the lexicographic cone we have the Gröbner basis with respect to the lexicographic term order. This sounds like wishful thinking, but it is not!!

First recall that $G[\omega]$ is an open set if and only if $\text{in}_{\omega}(I)$ is a monomial ideal $= \langle \text{in}_{\omega}(f_1), \dots, \text{in}_{\omega}(f_m) \rangle$ for some $f_1, \dots, f_m \in I$. This implies that (f_1, \dots, f_m) is a Gröbner basis for I wrt. \leq_{ω} , where \leq is an arbitrary term order. We say in this case that (f_1, \dots, f_m) is a Gröbner basis wrt. ω .

Remark 2.5.1 Here is a very important remark, which must be reconsidered a few times. Suppose that I is a homogeneous ideal. Then the reduced Gröbner basis with respect to an arbitrary term order consists of homogeneous polynomials! This follows by writing up what Buchbergers algorithm does to a set of homogeneous generators. Every remainder of S -polynomials added to the list during the algorithm must be homogeneous!!

Proposition 2.5.2 Let I be an ideal in $k[\mathbb{N}^n]$. Suppose that $G[\omega_1]$ and $G[\omega_2]$ are open cones and that $\omega \in G[\omega_1] \cap G[\omega_2]$. Let

$$G_1 = \{g_1, \dots, g_t\}$$

be **the reduced** Gröbner basis of I with respect to ω_1 . Then $\text{in}_\omega(G)$ is **the reduced** Gröbner basis of $\text{in}_\omega(I)$ with respect to ω_1 . Now let

$$H = \{h_1, \dots, h_s\}$$

be **the reduced** Gröbner basis of $\text{in}_\omega(I)$ with respect to ω_2 . Write

$$h_i = a_{i1}\text{in}_\omega(g_1) + \dots + a_{it}\text{in}_\omega(g_t)$$

using the division algorithm and let

$$f_i = a_{i1}g_1 + \dots + a_{it}g_t.$$

Then (f_1, \dots, f_s) is a Gröbner basis of I with respect to ω_2 .

Proof. Since

$$\text{in}_{\leq \omega_1}(\text{in}_\omega(I)) = \text{in}_{(\leq \omega_1)_\omega}(I) = \text{in}_{\omega_1}(I) = \langle \text{in}_{\omega_1}(g_1), \dots, \text{in}_{\omega_1}(g_t) \rangle$$

we get that $\text{in}_{\leq \omega_1}(I) \subseteq \text{in}_{(\leq \omega_1)_\omega}(I)$. This implies that

$$\text{in}_{\omega_1}(\text{in}_\omega(I)) = \text{in}_{\omega_1}(I).$$

It follows from the homogeneity of h_i that

$$h_i = \text{in}_\omega(h_i) = \text{in}_\omega(a_{i1}\text{in}_\omega(g_1) + \dots + a_{it}\text{in}_\omega(g_t)) = \text{in}_\omega(a_{i1}g_1 + \dots + a_{it}g_t).$$

This implies that

$$\text{in}_{\omega_2}(f_i) = \text{in}_{\omega_2}(\text{in}_\omega(f_i)) = \text{in}_{\omega_2}(h_i)$$

and therefore that

$$\text{in}_{\omega_2}(I) = \langle \text{in}_{\omega_2}(f_1), \dots, \text{in}_{\omega_2}(f_s) \rangle$$

since $\text{in}_{\omega_2}(I) = \text{in}_{\omega_2}(\text{in}_\omega(I)) = \langle \text{in}_{\omega_2}(h_1), \dots, \text{in}_{\omega_2}(h_s) \rangle$. \square

Chapter 3

Hilbert functions and best term order

3.1 Hilbert functions

Let I be an ideal in $R = k[x_1, \dots, x_n]$, where R is graded using the vector $(1, \dots, 1)$ (the usual grading). The Hilbert function $H_I : \mathbb{N} \rightarrow \mathbb{N}$ is defined as

$$H_I(m) = \dim_k R_m/I_m,$$

where $I_m = I \cap R_m$. Given a term order \leq we know that the set M of monomials $X^v \notin \text{in}_{\leq}(I)$ form a basis of R/I . Letting $M_m = \{X^v \in M \mid |v| = m\}$ we get that

$$|M_m| \leq H_I(m).$$

If I is a homogeneous ideal we have equality. Why is this? Well, suppose that I is homogeneous and let $G = (g_1, \dots, g_t)$ denote the reduced Gröbner basis of I wrt. some term order. Then g_1, \dots, g_t are homogeneous. Therefore we get

$$f \in R_m \implies f^G \in R_m.$$

Thus $f \equiv f^G \pmod{I}$ and therefore $[f] = [f^G]$ is a k -linear combination of monomials in M_m . We have proved for an arbitrary term order \leq and a homogeneous ideal I that

$$H_{\text{in}_{\leq}(I)} = H_I.$$

Remark 3.1.1 Let \leq be a term order and $F = (f_1, \dots, f_r) \subseteq I$, where I is homogeneous. Then F is a Gröbner basis if and only if

$$H_{\langle \text{in}_{\leq}(F) \rangle} = H_I.$$

This observation may come in handy. Sometimes we know the Hilbert function of an ideal (maybe we know a Gröbner basis with respect to another term order).

3.2 Is something already a Gröbner basis?

Suppose that we have given a set of generators $F = (f_1, \dots, f_n)$ for an ideal I . If F is a Gröbner basis for some term order then we can find $\omega \geq 0$ such that all of $\text{in}_\omega(f_1), \dots, \text{in}_\omega(f_n)$ are monomials. This can be formulated in the language of convex geometry. Let

$$K = \text{New}(f_1) + \dots + \text{New}(f_n).$$

The vertices of K correspond to the case with monomials, since we have the formula

$$\text{face}_\omega(K) = \text{face}_\omega(\text{New}(f_1)) + \dots + \text{face}_\omega(\text{New}(f_n)).$$

Let P denote the set of vertices of K . For each vertex $v \in P$ we compute

$$N_P(v) = \{\omega \in \mathbb{R}^n \mid \omega \cdot v \geq \omega \cdot u, \forall u \in P\}.$$

If $N_P(v)$ intersects \mathbb{R}_+^n we pick ω there and use Buchberger to check if F is a Gröbner basis wrt. ω .

Example 3.2.1 Do this for $F = (X^2 + Y, X^2Y + 1) \subseteq k[X, Y]$. In this case you get K a zonotope (Minkowski sum of line segments).

As described (for once with all the details) in Sturmfels pp. 25–26 one may also use the above search algorithm to find the best term order for example with respect to a Hilbert function criterion.

Chapter 4

LLL

4.1 Lattices

Let b_1, \dots, b_n be linearly independent vectors in \mathbb{R}^n . We let

$$L = \mathbb{Z}b_1 + \dots + \mathbb{Z}b_n$$

denote the lattice spanned by b_1, \dots, b_n . There are many \mathbb{Z} -bases of L . Each of these bases can be gotten from (b_1, \dots, b_n) by an invertible matrix $A \in \text{GL}_n(\mathbb{Z})$. The fun begins when we enter the inner product (\cdot, \cdot) on \mathbb{R}^n . As is well known this gives rise to the norm

$$|x| = \sqrt{(x, x)}$$

on \mathbb{R}^n . We let $m(L) = \min\{|x| \mid x \in L \setminus \{0\}\}$. The problem of finding a shortest vector in L can be formulated as

Problem 4.1.1 Find a vector $v \in L$ such that $|v| = m(L)$.

This is an extremely difficult problem. No polynomial time algorithm is known to solve it. C. Hermite showed in 1846 that

$$m(L) \leq (4/3)^{(n-1)/4} |\det(L)|^{1/n}.$$

4.2 Naive search for a shortest vector

How can we find a shortest vector in L ? Is this a finite problem?? The answer is yes and the reason is quite simple. Let B denote the matrix constructed with b_1, \dots, b_n as columns. If a_1, \dots, a_n denotes the rows of $(B^{-1})^t$, then

$$(a_i, b_j) = \delta_{ij}.$$

So if $u = x_1b_1 + \cdots + x_nb_n$, then $x_k = (u, a_k)$. By Cauchy-Schwarz we have

$$x_k^2 = (u, a_k)^2 \leq (u, u)^2(a_k, a_k)^2.$$

So if x_k denotes the k -th coordinate in a shortest vector v in L , then

$$x_k^2 \leq (b_j, b_j)(a_k, a_k)$$

for every $j = 1, \dots, n$. This shows that a shortest vector is within a finite search (even though it may be very long): Let $m = \min\{(b_j, b_j) \mid j = 1, \dots, n\}$. If $x = (x_1, \dots, x_k)$ is a shortest vector in L then

$$|x_j| \leq m|a_k|.$$

Remark 4.2.1 In fact one may show using the LLL-algorithm that finding a shortest vector can be done in polynomial time when n (the dimension) is fixed.

4.3 The algorithm of Gauss in the plane

Suppose that $L = \mathbb{Z}a + \mathbb{Z}b \subseteq \mathbb{R}^2$, where a and b are linearly independent. Then there is a nice algorithm due to Gauss for finding a shortest vector in L . Suppose that $|a|^2 > |b|^2$. Then we find $x \in \mathbb{R}$ that minimizes

$$|a - xb|^2 = (a - xb, a - xb).$$

The minimum M is assumed for $x = (a, b)/(b, b)$ (bottom point of a parabola). If $M \geq |b|^2$ then b is a shortest vector. In fact suppose that $w \in L$ is a shortest vector. Then $w = ua + vb$. We may assume that $u \neq 0$. Then

$$w = ua + (qu + r)b = u(a + qb) + rb,$$

where $0 \leq r < |u|$. Therefore

$$|w| \geq |u||a + qb| - r|b| \geq (|u| - r)|b|$$

and b is seen to be a shortest vector in L .

If $M = |a - qb|^2 < |b|^2$ we continue the procedure with b and $a - qb$, where q is the integer nearest to x . Why does this algorithm terminate? A straightforward reason is that there are only finitely many elements of L of norm $< R$, where $R > 0$. This does not explain why the algorithm is very effective (as effective as the Euclidean algorithm).

4.4 Gram-Schmidt

Recall the Gram-Schmidt orthogonalization procedure. Suppose that v_1, \dots, v_n is a basis of \mathbb{R}^n . Then we let $v_1^* = v_1$ and

$$v_i^* = v_i - \sum_{j=1}^{i-1} \mu_{ij} v_j^*,$$

for $i > 1$, where

$$\mu_{ij} = \frac{(v_i, v_j^*)}{(v_j^*, v_j^*)}.$$

It is easy to see that $(v_i^*, v_j^*) = 0$ if $i \neq j$ and $(v_i^*, v_i^*) = |v_i^*|^2$. A fundamental invariant of L is

$$d(L) = |\det(b_1, \dots, b_n)|.$$

4.5 Idea of the LLL-algorithm

A \mathbb{Z} -basis b_1, \dots, b_n for a lattice $L \subseteq \mathbb{R}^n$ is called LLL-reduced if

- i) $|\mu_{ij}| \leq 1/2$ for $1 \leq j < i \leq n$.
- ii) $|b_i^* + \mu_{i,i-1} b_{i-1}^*|^2 \geq \frac{3}{4} |b_{i-1}^*|^2$ for $1 < i \leq n$ (Lovasz condition).

Proposition 4.5.1 Let b_1, \dots, b_n be an LLL-reduced basis of L . Then

- i) $d(L) \leq |b_1| \cdots |b_n| \leq 2^{n(n-1)/4} d(L)$.
- ii) If $x \in L \setminus \{0\}$, then $|b_1| \leq 2^{(n-1)/2} |x|$.

Proof. In general we know that $d(L)$ is also the determinant of the matrix with column vectors b_1^*, \dots, b_n^* . Since this is an orthogonal matrix it implies that

$$d(L) = |b_1^*| \cdots |b_n^*|.$$

This observation makes it easy to see that $d(L) \leq |b_1| \cdots |b_n|$. When b_1, \dots, b_n is LLL-reduced it implies that

$$|b_i^*|^2 \geq 1/2 |b_{i-1}^*|^2$$

or that

$$|b_{i-1}^*|^2 \leq 2 |b_i^*|^2.$$

We get thus that $|b_{i-j}^*|^2 \leq 2^j |b_i^*|^2$ for $1 \leq j < i$. This gives

$$\begin{aligned} |b_i|^2 &= |b_i^*|^2 + \mu_{ii-1}^2 |b_{i-1}^*|^2 + \cdots + \mu_{i1} |b_1^*|^2 \\ &\leq |b_i^*|^2 + 1/4 |b_{i-1}^*|^2 + \cdots + 1/4 |b_1^*|^2 \\ &\leq |b_i^*|^2 + 1/2(1 + \cdots + 2^{i-2}) |b_i^*|^2 \\ &= \frac{2^{i-1} + 1}{2} |b_i^*|^2 \\ &\leq 2^{i-1} |b_i^*|^2. \end{aligned}$$

Therefore

$$|b_1|^2 \cdots |b_n|^2 \leq 2^{1+2+\cdots+n-1} d(L)^2$$

giving the desired inequality. Now suppose that $x \in L \setminus \{0\}$. Then

$$x = r_1 b_1^* + \cdots + r_i b_i^* = s_1 b_1 + \cdots + s_i b_i$$

where $r_i = s_i \in \mathbb{Z} \setminus \{0\}$ (transition matrix upper triangular with ones in the diagonal). This implies

$$|x|^2 \geq |b_i^*|^2.$$

Multiply both sides of this inequality with 2^{i-1} to get the desired result. \square

Remark 4.5.2 Here is a nice application of Cramer's rule: If b_1, \dots, b_n is reduced in the sense that

$$|b_1| \cdots |b_n| \leq 2^{n(n-1)/4} d(L),$$

then

$$|x_j| \leq 2^{n(n-1)/4}$$

where $x = (x_1, \dots, x_n)$ is a shortest vector in L . So if we have a polynomial time algorithm for turning a basis of a lattice into an LLL-reduced basis, then we also have a polynomial time algorithm for finding a shortest vector when the dimension n is fixed.

Once we have the definition of an LLL-reduced basis, the algorithm for turning a given basis into an LLL-reduced one is quite simple. There are two steps.

- i) If $|\mu_{ij}| > 1/2$ for some $1 \leq j < i \leq n$, then we replace b_i with $b_i - qb_j$, where q is the integer nearest to μ_{ij} . Here the "new" μ becomes $\mu_{ij} - q$. Therefore $|\mu| \leq 1/2$.
- ii) On the other hand if the Lovasz condition fails for a pair $(i-1, i)$ then we simply interchange the two vectors b_{i-1} and b_i .

Why is the Lovasz condition then satisfied for the pair $(i, i - 1)$? Well, let S be the subspace spanned by b_1, \dots, b_{i-2} and let

$$\pi : \mathbb{R}^n = S^\perp \oplus S \rightarrow S^\perp$$

denote the orthogonal projection on S^\perp . Then the Lovasz condition reads

$$|\pi(b_i)|^2 \geq \frac{3}{4} |\pi(b_{i-1})|^2.$$

If

$$|\pi(b_i)|^2 < \frac{3}{4} |\pi(b_{i-1})|^2$$

then clearly

$$|\pi(b_{i-1})|^2 \geq \frac{3}{4} |\pi(b_i)|^2.$$

We keep doing this and stop when we have an LLL-reduced basis. Why does this algorithm terminate? We will use a beautiful classical argument to show this. Let

$$d_i = \det((b_r, b_s)_{1 \leq r, s \leq i}).$$

Notice that this also is the square of the determinant of the lattice spanned by b_1, \dots, b_i . One may prove that

$$d_i = |b_1^*|^2 \cdots |b_i^*|^2.$$

We have the following classical result due to Hermite.

Theorem 4.5.3 There exists a non-zero vector $x \in L \setminus \{0\}$ such that

$$|x|^2 \leq (4/3)^{\frac{n-1}{2}} d(L)^{\frac{1}{n}}.$$

Let

$$D = d_1 \cdots d_{n-1}.$$

The only operation changing the D is the Lovasz swap. Suppose that b_{i-1} and b_i is swapped. This happens only if

$$|b_i^* + \mu_{i,i-1} b_{i-1}^*|^2 < \frac{3}{4} |b_{i-1}^*|^2.$$

This means that d_{i-1} gets multiplied with a factor which is $\leq 3/4$. Notice that d_j , $j \neq i - 1$ are unchanged, so that D gets multiplied with $3/4$. Since each of the d_i are bounded below (this follows from Theorem 4.5.3), D must be bounded below and the algorithm stops after a finite number of steps.

Chapter 5

Toric ideals

We begin this chapter with some basic results in the abstract framework of commutative rings.

Lemma 5.0.4 Let $a_1, \dots, a_n, b_1, \dots, b_n$ be elements of a commutative ring R . Then

$$a_1 \cdots a_n - b_1 \cdots b_n \in \langle a_1 - b_1, \dots, a_n - b_n \rangle.$$

Proof. This is an easy induction. Use that

$$a_1 a_2 \cdots a_n - b_1 b_2 \cdots b_n = a_2 \cdots a_n (a_1 - b_1) + b_1 (a_2 \cdots a_n - b_2 \cdots b_n).$$

□

Proposition 5.0.5 Let $\varphi : R \rightarrow S$ be a ring homomorphism and $s : S \rightarrow R$ another ring homomorphism such that $\varphi s = 1_S$. Then

$$\varphi^{-1}(I) = s(I) + \text{Ker}(\varphi).$$

Proof. If $x \in I$ and $y \in \text{Ker}(\varphi)$, then $\varphi(s(x) + y) = \varphi(s(x)) = x \in I$. Thus $s(I) + \text{Ker}(\varphi) \subseteq \varphi^{-1}(I)$. On the other hand suppose that $x \in \varphi^{-1}(I)$. Then $s(\varphi(x)) - x \in \text{Ker}(\varphi)$. Therefore $x = s(\varphi(x)) - y$, where $y \in \text{Ker}(\varphi)$ so that $\varphi^{-1}(I) \subseteq s(I) + \text{Ker}(\varphi)$. □

5.1 Computing kernels

The purpose of this section is to develop a method to compute the kernel of a ring homomorphism

$$\varphi : k[x_1, \dots, x_m]/J \rightarrow k[y_1, \dots, y_n]/I.$$

NB: Our ring homomorphisms will always fix k . Thus $\varphi(x) = x$ if $x \in k$. We begin in the simplest case, where $I = J = 0$.

Proposition 5.1.1 Let

$$\varphi : k[x_1, \dots, x_m] \rightarrow k[y_1, \dots, y_n]$$

denote a ring homomorphism given uniquely by $\varphi(x_i) = f_i \in k[y_1, \dots, y_n]$. We extend this definition to a ring homomorphism

$$\tilde{\varphi} : k[x_1, \dots, x_m, y_1, \dots, y_n] \rightarrow k[y_1, \dots, y_n]$$

by $\tilde{\varphi}(x_i) = f_i$ and $\tilde{\varphi}(y_j) = y_j$. Then $\text{Ker } \varphi = k[x_1, \dots, x_m] \cap \text{Ker } \tilde{\varphi}$ and

$$\text{Ker } (\tilde{\varphi}) = \langle x_1 - f_1, \dots, x_m - f_m \rangle.$$

Proof. If $P \in \langle x_1 - f_1, \dots, x_m - f_m \rangle$, then

$$\tilde{\varphi}(P) = P(f_1, \dots, f_m, y_1, \dots, y_n) = 0.$$

On the other hand if $P(f_1, \dots, f_m, y_1, \dots, y_n) = 0$. Then

$$\begin{aligned} P(x_1, \dots, x_m, y_1, \dots, y_n) &= P(x_1, \dots, x_m, y_1, \dots, y_n) - P(f_1, \dots, f_m, y_1, \dots, y_n) \\ &\in \langle x_1 - f_1, \dots, x_m - f_m \rangle \end{aligned}$$

by Lemma 5.0.4 (why?). \square

Remark 5.1.2 This shows that Gröbner basis theory can be used to compute $\text{Ker } \varphi$ above. Use lex order with $y_1, \dots, y_n \geq x_1, \dots, x_m$. The reduced Gröbner basis G of the ideal

$$\langle x_1 - f_1, \dots, x_m - f_m \rangle$$

in $k[x_1, \dots, x_m, y_1, \dots, y_n]$ is special in that $G \cap k[x_1, \dots, x_m]$ is the reduced Gröbner basis for $\text{Ker } \varphi$ in $k[x_1, \dots, x_m]$ (cf. Algebra 1).

Proposition 5.1.3 Now suppose that

$$\varphi : k[x_1, \dots, x_m]/I \rightarrow k[y_1, \dots, y_n]/J$$

is a ring homomorphism. Then there exists a ring homomorphism

$$\psi : k[x_1, \dots, x_m] \rightarrow k[y_1, \dots, y_n],$$

such that $\varphi(f + I) = \psi(f) + J$ and

$$(K \cap k[x_1, \dots, x_m]) + I,$$

where K is the ideal in $k[x_1, \dots, x_m, y_1, \dots, y_n]$ generated by J and $\text{Ker } (\tilde{\psi})$, where $\tilde{\psi}$ is defined as in Proposition 5.1.1.

Proof. This is a consequence of Proposition 5.0.5 and the fact that the kernel of

$$k[x_1, \dots, x_m, y_1, \dots, y_n] \xrightarrow{\tilde{\psi}} k[y_1, \dots, y_n] \rightarrow k[y_1, \dots, y_n]/J$$

is $\tilde{\psi}^{-1}(J)$. \square

5.2 Computing images

Suppose that $\varphi : k[x_1, \dots, x_m] \rightarrow k[y_1, \dots, y_n]$ is a ring homomorphism, how do we decide if $g \in k[y_1, \dots, y_n]$ is in $\text{Im } \varphi$? This question can be rephrased: How do we decide if $g \in k[f_1, \dots, f_m]$, where $\varphi(x_i) = f_i$? Again it pays to consider

$$\tilde{\varphi} : k[x_1, \dots, x_m, y_1, \dots, y_n] \rightarrow k[y_1, \dots, y_n].$$

If

$$g = \lambda_1(x_1 - f_1) + \dots + \lambda_m(x_m - f_m) + P, \quad (*)$$

where $\lambda_1, \dots, \lambda_m \in k[x_1, \dots, x_m, y_1, \dots, y_n]$ and $P \in k[x_1, \dots, x_m]$, then clearly

$$g \in k[f_1, \dots, f_m].$$

In fact in this case $g = P(f_1, \dots, f_m)$. On the other hand if $g = P(f_1, \dots, f_m)$ for some polynomial $P \in k[x_1, \dots, x_m]$, then

$$g - P \in \langle x_1 - f_1, \dots, x_m - f_m \rangle$$

by Lemma 5.0.4. This shows that $g \in k[f_1, \dots, f_m]$ if and only if $(*)$ holds. Let G be a Gröbner basis of $\langle x_1 - f_1, \dots, x_m - f_m \rangle$ with respect to lex , where $y_1, \dots, y_n \geq x_1, \dots, x_m$. Then we get

$$g \in k[f_1, \dots, f_m] \iff g^G \in k[x_1, \dots, x_m].$$

The \Leftarrow direction is easy. The \Rightarrow direction calls for the details of the division algorithm. If $g = P(f_1, \dots, f_m)$ for a polynomial $P \in k[x_1, \dots, x_m]$, then

$$g = \lambda_1(x_1 - f_1) + \dots + \lambda_m(x_m - f_m) + P.$$

This does not mean that $P = g^G$!!! It does mean though that we have an expression

$$g = \mu_1 g_1 + \dots + \mu_N g_N + P,$$

where $G = \{g_1, \dots, g_N\}$. Continue the division algorithm from this point! The point is that if $\text{in}_{\leq}(g_j) \mid \text{in}_{\leq}(g)$, then $g_j \in k[x_1, \dots, x_m]$. We constantly move inside $k[x_1, \dots, x_m]$ in the division algorithm and terms transferred to the remainder are all in $k[x_1, \dots, x_m]$. Thus $g^G \in k[x_1, \dots, x_m]$.

5.2.1 An application

Suppose that A is an $n \times m$ matrix with natural numbers as entries and $b \in \mathbb{N}^n$. Can we find $v \in \mathbb{N}^m$ such that

$$Av = b?$$

This problem is NP-complete and has a lot to do with integer linear programming. We will translate it into computing the image of a toric ring homomorphism. Consider

$$\varphi : k[x_1, \dots, x_m] \rightarrow k[t_1, \dots, t_n]$$

given by $\varphi(x^v) = t^{Av}$. Then clearly our question is equivalent to if t^b is in the image of φ and we have developed a Gröbner basis method for answering this.

5.3 Toric homomorphisms

In algebraic geometry a torus T_n is simply the subset $(k^*)^n = k^* \times \dots \times k^*$ of affine n -space $\mathbb{A}^n = k^n$. A torus is a group wrt. multiplication. In fact it is a so-called algebraic group. A toric morphism is an “equivariant” algebraic map $T_n \rightarrow \mathbb{A}^m$ for some $n, m \in \mathbb{N}$. There is an equivalent algebraic description which we will use. A ring homomorphism

$$\varphi : k[x_1, \dots, x_m] \rightarrow k[t_1, t_1^{-1}, \dots, t_n, t_n^{-1}] \subseteq k(t_1, \dots, t_n)$$

is called toric if $\varphi(x_i) = t^{v_i}$, where $v^i \in \mathbb{Z}^n$. If $v = (v_1, \dots, v_n) \in \mathbb{Z}^n$, then the notation t^v means

$$t_1^{v_1} \dots t_n^{v_n}.$$

So a toric homomorphism $\varphi : k[x_1, \dots, x_m] \rightarrow k[t_1, t_1^{-1}, \dots, t_n, t_n^{-1}]$ is given by m vectors $(v_1, \dots, v_m) \subseteq \mathbb{Z}^n$ or in other words an $n \times m$ integral matrix $A_\varphi \in \text{Mat}_{n,m}(\mathbb{Z})$. If $A \in \text{Mat}_{n,m}(\mathbb{Z})$ we let φ_A denote the corresponding toric homomorphism.

Definition 5.3.1 A toric ideal I_A is the kernel of a toric homomorphism φ_A .

Remark 5.3.2 If $v \in \mathbb{Z}^n$ we let $v^+ = \max(v, 0)$ and $v^- = \max(-v, 0)$. Then $v^+, v^- \in \mathbb{N}^n$ and $v = v^+ - v^-$.

Example 5.3.3 An integral matrix $A \in \text{Mat}_{n,m}(\mathbb{Z})$ can be viewed as a \mathbb{Z} -linear map $\mathbb{Z}^m \rightarrow \mathbb{Z}^n$. If $v \in \text{Ker } A$, then $x^{v^+} - x^{v^-} \in \text{Ker } \varphi_A$.

Lemma 5.3.4 Let $A \in \text{Mat}_{n,m}(\mathbb{Z})$ and let $I_A = \text{Ker } \varphi_A$. Then

$$I_A = \langle x^u - x^v \mid u, v \in \mathbb{N}^m, Au = Av \rangle.$$

Proof. The inclusion \supseteq clearly holds. Suppose that $f \in I_A$ and that $f = \sum_{v \in \mathbb{N}^n} a_v x^v$. Then

$$\varphi(f) = \sum_{v \in \mathbb{N}^n} a_v t^{Av}.$$

Fix some term order \leq . If $\text{in}_{\leq}(f) = a_v X^v$, then f must contain a term $a_u X^u$ such that $Av = Au$, but this implies that $g = a_v(X^v - X^u) \in I_A$. Therefore $f - g \in I_A$ and $\text{in}_{\leq}(f - g) < \text{in}_{\leq}(f)$. Now use induction to conclude that $f - g \in \langle x^u - x^v \mid u, v \in \mathbb{N}^n, Au = Av \rangle$. The result follows. \square

Corollary 5.3.5

$$I_A = \langle x^{v^+} - x^{v^-} \mid v \in \text{Ker } A \rangle,$$

A reduced Gröbner basis with respect to an arbitrary term order \leq consists of binomials of the form $x^{v^+} - x^{v^-}$, where $v \in \text{Ker } A$.

Proof. I_A is a prime ideal and $v - u = (v - u)^+ - (v - u)^-$:

$$X^v - X^u = X^{v-(v-u)^+} (X^{(v-u)^+} - X^{(v-u)^-}) \in I_A.$$

Binomials are preserved when taking S -polynomials and remainders with the division algorithm. So starting with a set of binomials we end up with a Gröbner basis consisting of binomials. Suppose that $X^u - X^v$ is a binomial in a reduced Gröbner basis. Then $v - (v - u)^+ = 0$ (why?). \square

5.4 Kernels of toric homomorphisms

Proposition 5.4.1

$$k[t_1, t_1^{-1}, \dots, t_n, t_n^{-1}] \cong k[t_0, t_1, \dots, t_n] / \langle t_0 t_1 \cdots t_n - 1 \rangle.$$

Consider a toric ring homomorphism given by $\varphi(x_i) = t^{v_i}$ with $v_i \in \mathbb{Z}^n$ (not limited to \mathbb{N}^n !!). Then we can compute the kernel of φ by intersecting the ideal

$$\langle t_0 t_1 \cdots t_n - 1, t^{v_1^-} x_1 - t^{v_1^+}, \dots, t^{v_m^-} x_m - t^{v_m^+} \rangle$$

with $k[x_1, \dots, x_m]$. This is a consequence of Proposition 5.1.3.

5.5 Computing the kernel of a toric homomorphism

So given a matrix A with integral entries we have defined a toric ring homomorphism

$$\varphi_A : k[x_1, \dots, x_m] \rightarrow k[t_1^{\pm 1}, \dots, t_n^{\pm 1}].$$

We wish to compute the kernel I_A of φ_A somehow using the related linear algebra kernel of

$$A : \mathbb{Z}^m \rightarrow \mathbb{Z}^n.$$

Given a subset $C \subseteq \text{Ker } A$ we let

$$J_C = \{x^{v^+} - x^{v^-} \mid v \in C\}.$$

We know that $\text{Ker } \varphi = J_C$ for some subset C . Here is a crucial result which is the foundation for finding such a subset. First a definition.

Definition 5.5.1 Let R be a commutative ring, $f \in R$ and I an ideal of R . Then we define the ideal

$$I : f^\infty = \{x \in R \mid x f^m \in I, m \gg 0\}.$$

Exercise 5.5.2 Check that this really is an ideal. Also check the identity

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

Proposition 5.5.3 If C is a generating set for $\text{Ker } A$, then

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

Proof. Let $C = \{v_1, \dots, v_r\}$. If $(x_1 \cdots x_n)^m f \in J_C$ then $\varphi(x_1 \cdots x_n)^m \varphi(f) = 0$ and therefore $f \in I_A$. This shows \subseteq . On the other hand suppose that $x^{v^+} - x^{v^-} \in \text{Ker } \varphi$, where $v = \lambda_1 v_1 + \cdots + \lambda_r v_r \in \text{Ker } A$. Then $v^+ - v^- = (\lambda_1 v_1)^+ - (\lambda_1 v_1)^- + \cdots + (\lambda_r v_r)^+ - (\lambda_r v_r)^-$. This translates into the following identity for rational functions

$$\frac{x^{v^+}}{x^{v^-}} = \frac{x^{(\lambda_1 v_1)^+}}{x^{(\lambda_1 v_1)^-}} \cdots \frac{x^{(\lambda_r v_r)^+}}{x^{(\lambda_r v_r)^-}} = \left(\frac{x^{v_1^+}}{x^{v_1^-}}\right)^{\lambda_1} \cdots \left(\frac{x^{v_r^+}}{x^{v_r^-}}\right)^{\lambda_r}.$$

First subtract 1 from both sides and then multiply with a sufficiently high power of $x_1 \cdots x_n$ to see the inclusion \supseteq and don't forget to use the fact that I_A is a prime ideal. Writing things a bit out we get

$$x^w (x^{v^+} - x^{v^-}) \in J_C = \langle \pm(x^{v_1^+} - x^{v_1^-}), \dots, \pm(x^{v_r^+} - x^{v_r^-}) \rangle.$$

□

Here is a very interesting corollary.

Corollary 5.5.4 If C contains a vector u with strictly positive entries, then

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

Proof. We get that $x^u - 1 \in J_C$ and thereby that $x^{nu} - 1 \in J_C$ for $n \in \mathbb{N}$. Suppose that $x^w f \in J_C$ for $w \gg 0$. Then we may adjust w such that $x^w f \in J_C$ and $x^w - 1 \in J_C$. This gives that $f \in J_C$. \square

The result enabling us to compute $J_C : (x_1 \cdots x_n)^\infty$ is the following.

Proposition 5.5.5 Let I be a homogeneous ideal in $k[x_1, \dots, x_n]$. Let \leq be the graded reverse lex order with x_n smallest and G the **reduced** Gröbner basis of I wrt. \leq . Then

$$G' = \{f \in G \mid x_n \nmid f\} \cup \{f/x_n^m \mid f \in G, x_n \mid f\}$$

is a Gröbner basis of $I : x_n^\infty$, where x_n^m denotes the highest power dividing f .

Proof. We must prove that if $g \in I : x_n^\infty$, then $\text{in}_{\leq}(f) \mid \text{in}_{\leq}(g)$ for some $f \in G'$. Now we know that $x_n^m g \in I$ for some $m \gg 0$, so we may find $g' \in G$ such that $\text{in}_{\leq}(g') \mid \text{in}_{\leq}(x_n^m g) = x_n^m \text{in}_{\leq}(g)$. If $x_n \nmid \text{in}_{\leq}(g')$, then $\text{in}_{\leq}(g') \mid \text{in}_{\leq}(g)$. If $x_n \mid \text{in}_{\leq}(g')$ we use the fundamental observation that if g is homogeneous, then

$$x_n \mid \text{in}_{\leq}(g) \iff x_n \mid g.$$

This is a consequence of the graded reverse lexicographic order with x_n being smallest. \square

Remark 5.5.6 Now we have a straightforward algorithm for computing a generating set for I_A . First compute a \mathbb{Z} -basis B for $\text{Ker } \varphi$. Use the LLL algorithm to reduce the basis B into a basis B' consisting of relatively short vectors. Then start computing

$$J_{B'} : (x_1 \cdots x_n)^\infty$$

using the above results. Reducing the lattice basis first leads to a significant improvement in the running time of the total algorithm.

Remark 5.5.7 This algorithm works only for homogeneous ideals. In order for I_A to be homogeneous we must have that A e. g. has non-negative entries.

5.5.1 The Di Biase - Urbanke algorithm

We will illustrate a beautiful algorithm for computing the kernel of a toric ring homomorphism by means of an example.

Example 5.5.8 It is difficult for us to compute the kernel of φ_A , where $A = (3, 4, 5)$. It is easier if we change A a little bit into the matrix $A_1 = (-3, 4, 5)$. In this case $\text{Ker } A_1$ contains the vector $(3, 1, 1)$ with strictly positive entries. In fact a \mathbb{Z} -basis of $\text{Ker } A_1$ is $(3, 1, 1)$ and $(4, 3, 0)$. So by Corollary 5.5.4 we get

$$I_{A_1} = \langle x^3yz - 1, x^4y^3 - 1 \rangle$$

without going through any horrendous Gröbner basis computations! But we needed I_A and not I_{A_1} . Is there an easy way of changing I_{A_1} into I_A ? The answer is yes. Choose an elimination order \leq eliminating x , for example \leq the lexicographic order with $x \geq y \geq z$. Compute the Gröbner basis

$$(y^5 - z^4, y^3 - xz^3, z - xy^2, y^2 - x^2z^3, x^3yz)$$

of I_{A_1} with respect to \leq . Then flip the x to the other term as follows

$$(y^5 - z^4, xy^3 - z^3, xz - y^2, x^2y^2 - z^3, yz - x^3).$$

These polynomials form a generating set for I_A !!!! You can read the proof of this in Sturmfels Proposition 12.5.

Appendix A

Review of convexity

Recall that a non-empty subset $C \subseteq \mathbb{R}^n$ is called convex if

$$\lambda x + (1 - \lambda)y \in C$$

for every $x, y \in C$ and $\lambda \in [0, 1]$. An affine hyperplane H in \mathbb{R}^n is given by

$$H = \{(v_1, \dots, v_n) \mid \lambda_1 v_1 + \dots + \lambda_n v_n = \alpha\},$$

where $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (not all zero) and $\alpha \in \mathbb{R}$. We associate the closed half spaces

$$H_- = \{(v_1, \dots, v_n) \mid \lambda_1 v_1 + \dots + \lambda_n v_n \leq \alpha\}$$

and

$$H_+ = \{(v_1, \dots, v_n) \mid \lambda_1 v_1 + \dots + \lambda_n v_n \geq \alpha\}$$

with H .

A.1 Separation

Now we can state a fundamental separation theorem.

Theorem A.1.1 Let C_1, C_2 be convex subsets of \mathbb{R}^n , such that $C_1 \cap C_2 = \emptyset$. Then there exists an affine hyperplane H such that $C_1 \subseteq H_-$ and $C_2 \subseteq H_+$.

Proof. This is a consequence of ([1], Theorem 11.3). \square

A.2 Polyhedra

A polyhedron is a subset of \mathbb{R}^n of the form

$$P = \{x \in \mathbb{R}^n \mid Ax \leq b\},$$

where A is an $m \times n$ real matrix and $b \in \mathbb{R}^m$. Geometrically this is an intersection of closed half spaces.

Theorem A.2.1 If $b = 0$, then

$$P = C(u_1, \dots, u_n) = \mathbb{R}_{\geq 0} u_1 + \dots + \mathbb{R}_{\geq 0} u_n.$$

So P is a cone generated by finitely many vectors u_1, \dots, u_n if $b = 0$.

Definition A.2.2 A polytope is a bounded polyhedron.

Theorem A.2.3 A polytope P is the convex hull of a finite set $v_1, \dots, v_N \in \mathbb{R}^n$ of points.

$$P = \text{conv}\{v_1, \dots, v_N\}.$$

Definition A.2.4 Let P be a polyhedron. Then a face of P is a subset of P , where a linear functional is maximized. Thus given $\omega \in \mathbb{R}^n$,

$$\text{face}_\omega(P) = \{u \in P \mid u \cdot \omega \geq v \cdot \omega, \forall v \in P\}.$$

A face of (affine) dimension zero is called a *vertex*. A face of (affine) dimension one is called an *edge*. A face of codimension one in P is called a *facet*.

Definition A.2.5 The normal cone of a polyhedron $P \subseteq \mathbb{R}^n$ at a face $F \subseteq P$ is

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

So the normal cone consists of the vectors of linear functionals maximized on the face.

Exercise A.2.6 A face of a polyhedron is a polyhedron.

A.3 Minkowski sums

The Minkowski sum of two subsets $A, B \subseteq \mathbb{R}^n$ is

$$A + B = \{a + b \mid a \in A, b \in B\}.$$

Bibliography

- [1] Rockafellar, *Convex analysis*, Princeton University Press, 1970.