Notes on Integral Closure

Dr. Leslie Roberts

Notes by Susan Cooper

Department of Mathematics and Statistics
Queen's University
Kingston, ON
Canada K7L 3N6

During the Winter and Spring of 1999 I gave nine talks in the Curves Seminar. Thanks to Susan Cooper for the excellent note taking, and to Tony Geramita for suggesting that they be included in this volume of the Curves Seminar proceedings. The notes are aimed at beginning graduate students, and anyone else who might be interested. They form part of the background for my current research with Les Reid. As usual I did not get as far as I had hoped. The next topics in the main part of the notes would have been seminormal monoids and their monoid rings, and closure operations on ideals (see [8] and [9] for some idea of where things are going). The final two lectures contain little in the way of proofs, and are only intended to indicate further development of the subject.

Lecture 1

Let A and B be commutative rings which share the same identity 1. We will consider the ring extension $A \subseteq B$. In these notes all rings will be commutative with identity unless otherwise stated. Also all ring extensions will be such that the rings involved share the same identity.

Definition 1.1 An element x in B is said to be *integral over* A if there exist $a_0, \ldots, a_{n-1} \in A$ such that

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0.$$

We refer to this as an *equation of integral dependence of* x *over* A.

Theorem 1.2 *The following are equivalent;*

(1) *x is integral over A,*

(2) *The subring $A[x]$ of B is a finitely generated A-module,*

(3) *There is a faithful $A[x]$-module which is a finitely generated A-module.*

Before proving this theorem we recall that M is a *faithful* A-module if $aM = 0$ implies that $a = 0$, where $a \in A$. For example, any commutative ring A with identity 1 is a faithful A-module since if $a \in A$, where a is nonzero, then $a1$ is nonzero.

Proof (1) $\Longrightarrow$ (2): $A[x]$ is generated as an A-module by $1, x, x^2, \ldots$ The equation $x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$ can be used to replace the generator $x^i, i \geq n$ by an A-linear combination of $1, x, \ldots, x^{n-1}$.
(2) $\Longrightarrow$ (3): $A[x]$ is a faithful $A[x]$-module which is a finitely generated A-module.
(3) $\Longrightarrow$ (1): Let M be a faithful $A[x]$-module generated as an A-module by $m_1, m_2, \ldots, m_n$. Then $xm_i = a_{i1}m_1 + \cdots + a_{in}m_n$, where $1 \leq i \leq n$ and $a_{ij} \in A$. Thus we get the following system of equations

$$\begin{array}{ccccccccc}
(x-a_{11})m_1 & - & a_{12}m_2 & - & \cdots & - & a_{1n}m_n & = & 0 \\
-a_{21}m_1 & + & (x-a_{22})m_2 & - & \cdots & - & a_{2n}m_n & = & 0 \\
 & & & & & & & \vdots & \\
-a_{n1}m_1 & - & a_{n2}m_2 & - & \cdots & + & (x-a_{nn})m_n & = & 0
\end{array}$$

Now let D be the coefficient matrix of this system. That is,

$$D = \begin{bmatrix} (x-a_{11}) & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & (x-a_{22}) & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & (x-a_{nn}) \end{bmatrix}$$

So now the equations above can be written in the matrix form

$$D \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{1}$$

Let D_{ij} be the $(n-1)\times(n-1)$ matrix obtained by crossing out the i^{th} row and j^{th} column of D. We also recall that the classical adjoint, adj D, of D is the $n \times n$ matrix where the ij entry is given by $(-1)^{i+j} \det D_{ji}$. Moreover,

$$(\text{adj } D)(D) = \begin{bmatrix} \det D & 0 & \cdots & 0 \\ 0 & \det D & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & & & \det D \end{bmatrix}$$

Thus if we multiply (1) on the left by adj D we get

$$\begin{bmatrix} \det D & 0 & \cdots & 0 \\ 0 & \det D & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & & & \det D \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Which implies that $(\det D)m_i = 0, 1 \leq i \leq n$. But now $\det D \in A[x]$ and $(\det D)M = 0$. But M is a faithful $A[x]$-module and so it must be that

$\det D = 0$. Now $\det D$ is of the form $x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $a_i \in A$, and hence x is integral over A as required. $\square$

Definition 1.3 The extension $A \subseteq B$ is *integral* if every element $x \in B$ is integral over A.

Lemma 1.4 *The set* ${}_B\overline{A} = \{x \in B \mid x \textit{ is integral over } A\}$ *is a ring containing* A.

Proof Let $x, y \in {}_B\overline{A}$. Consider $A[x, y]$. This is a finitely generated A-module. To see this let (1) $x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$ and (2) $y^m + b_{m-1}x^{m-1} + \cdots + b_0 = 0$ be equations of integral dependence for x and y respectively. Then $A[x, y]$ is generated as an A-module by $x^i y^j, 0 \leq i < n, 0 \leq j < m$. Now $x + y \in A[x, y]$. Thus, $A[x, y]$ is a faithful $A[x + y]$-module containing $x + y$, so by Theorem 1.2 (3) $x + y$ is integral over A. Similarly for xy. $\square$

EXERCISE: Given the equations of integral dependence (1) and (2) in the above proof, how might one find explicit equations of integral dependence over A for $x + y$ and xy?

Definition 1.5 (0) ${}_B\overline{A}$ is the *integral closure* of A in B.

(1) The extension $A \subseteq B$ is *integral* if ${}_B\overline{A} = B$.

(2) A is *integrally closed in* B if $A = {}_B\overline{A}$.

NOTE: From now on we will drop the phrase "the extension" and assume this is what we are referring to when discussing closure.

Theorem 1.6 *Consider the extensions* $A \subseteq B \subseteq C$. *Then* $A \subseteq C$ *is integral if and only if* $A \subseteq B$ *and* $B \subseteq C$ *are both integral.*

Proof The proof is left as an exercise with the note that the transitivity property (i.e. $A \subseteq B$ and $B \subseteq C$ both integral implies $A \subseteq C$ integral) is especially important. $\square$

CONSEQUENCE OF TRANSITIVITY: ${}_B\overline{A}$ is integrally closed in B.

EXAMPLES

(1) $\mathbb{Z}$ is integrally closed in $\mathbb{Q}$.

(2) If K is a finite algebraic extension of $\mathbb{Q}$, then $A = {}_K\overline{\mathbb{Z}}$ is the algebraic integers of K.

(3) If $K = \mathbb{Q}(\sqrt{d})$, where d is a square free integer, then the integers A in K are given by the following two cases.

(a) If $d \equiv 2$ or $3 \pmod 4$, then $A = \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\}$.

(b) If $d \equiv 1 \pmod 4$, then $A = \{(a + b\sqrt{d})/2 \mid a, b \in \mathbb{Z}$ and a, b have the same parity$\}$.

Let us check that $x = (a + b\sqrt{d})/2$ is an integer (i.e. x is integral over $\mathbb{Z}$), where $d \equiv 1 \pmod 4$ and a, b are both odd.

If x is an integer then we have that $2x - a = b\sqrt{d}$ and hence $4x^2 - 4ax + a^2 = b^2 d$ or $4x^2 - 4ax + a^2 - b^2 d = 0$. Thus, $x^2 - ax + (a^2 - b^2 d)/4 = 0$. We see that if $(a^2 - b^2 d)/4 \in \mathbb{Z}$, then x is integral over $\mathbb{Z}$. Now both a and b are odd and hence $a^2 \equiv b^2 \equiv 1 \pmod 4$. This fact, along with $d \equiv 1 \pmod 4$, gives us that $a^2 - b^2 d \equiv 0 \pmod 4$. Thus, $(a^2 - b^2 d)/4 \in \mathbb{Z}$ and the result that x is integral over $\mathbb{Z}$ is achieved.

REFERENCE: The material of this lecture is in many books. I followed mostly [7], Chapter 9.

Lecture 2

Let A be a reduced noetherian ring and $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_r$ be the minimal primes of A. Then $(0) = \cap_{i=1}^r \mathfrak{p}_i$. Then we have inclusions $A \hookrightarrow \prod_{i=1}^r (A/\mathfrak{p}_i) \hookrightarrow \prod_{i=1}^r K_i = K$, where K_i is the field of fractions of the integral domain $A/\mathfrak{p}_i$.

If we let S be the multiplicative set of non-zero-divisors in A, then $A_S = K$, i.e. K is the total ring of fractions of A. We have ${}_K\overline{A} = \prod_{i=1}^r \overline{(A/\mathfrak{p}_i)}$, where $\overline{A/\mathfrak{p}_i}$ is the integral closure of $A/\mathfrak{p}_i$ in K_i. This can be found in [1].

SPECIAL CASE

Let $X = \{P_1, \ldots, P_r\}$ be the r points in $\mathbb{P}_k^n$, $R = k[x_0, \ldots, x_n]$, where k is a field. If $p_i = (a_0, a_1, \ldots, a_n)$ is homogeneous co-ordinates for P_i $(a_0 \neq 0)$, then $\mathfrak{P}_i = (a_0 x_j - a_j x_0 | 1 \leq j \leq n)$ is the ideal of P_i, and $I = \cap_{i=1}^r \mathfrak{P}_i$ is the ideal of X. The homogeneous co-ordinate ring of X is $A = R/I$, and the minimal primes of A are $\mathfrak{p}_i = \mathfrak{P}_i/I$.

We have $R/\mathfrak{P}_i \cong k[t_i]$ since $\mathfrak{P}_i$ is generated by n homogeneous linear forms which can be taken as variables. Then $A/\mathfrak{p}_i \cong (R/I)/(\mathfrak{P}_i/I) \cong R/\mathfrak{P}_i \cong k[t_i]$. The above becomes $A \hookrightarrow \prod_{i=1}^r (A/\mathfrak{p}_i) = \prod_{i=1}^r k[t_i] \hookrightarrow \prod_{i=1}^r k(t_i) = K$ and ${}_K\overline{A} = \prod_{i=1}^r k[t_i]$.

Suppose now that $p_1 = (1, a_1, \ldots, a_n)$, $p_2 = (1, b_1, \ldots, b_n)$, $\ldots$, $p_r = (1, r_1, \ldots, r_n)$. Now consider the following mappings $x_0 \to (t_1, t_2, \ldots, t_r)$, $x_1 \to (a_1 t_1, b_1 t_2, \ldots, r_1 t_r), \ldots,$ $x_n \to (a_n t_1, b_n t_2, \ldots, r_n t_r)$. We now see that $A \cong k$-subalgebra of $\prod_{i=1}^r k[t_i]$ given by $n+1$ explicit elements of degree one, one of which is $(t_1, t_2, \ldots t_r)$. The beauty of this is that we compute the Hilbert Function of X via linear algebra and thus need not know I explicitly.

Definition 2.1 A ring A is *integrally closed* if it is integrally closed in its total ring of fractions A_S, where S is the multiplicative set of non-zero-divisors.

FACT: If A is a unique factorization domain (UFD), then A is integrally closed. Examples include $\mathbb{Z} \subset \mathbb{Q}$, or $k[x_1, x_2, \ldots, x_n] \subset k(x_1, x_2, \ldots, x_n)$ where k is a field.

RELATIONS WITH PRIME IDEALS

The following relations are discussed in greater detail in [5]. Let $R \subseteq S$ be integral.

(1) If P and Q, where $P \subset Q$, are distinct prime ideals of S then $P \cap R \subset Q \cap R$ are distinct primes of R.

(2) If $P_1 \subset P_2 \subset \cdots \subset P_n$ is a chain of distinct prime ideals of R, then there exists a chain of distinct prime ideals $P_1' \subset P_2' \subset \cdots \subset P_n'$ of S such that $P_i' \cap R = P_i$

(3) $\text{Krulldim} R = \text{Krulldim} S$.

GOAL: Our main goal now is to show that if $A \subseteq B$ is an extension of graded rings, then ${}_B\overline{A}$ is a graded subalgebra of B. The argument that we use follows [1].

Lemma 2.2 *Let A be a commutative ring with identity 1, and $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ be a monic polynomial in $A[x]$, where x is an indeterminate over A and $n > 0$. Then there is an extension ring B of A such that, in $B[x]$, f has a root.*

Proof Let $B = A[x]/(f(x))$. Then the canonical image $\bar{x}$ of x in B is a root of f. To complete the proof we should check that the canonical map $A \to B$ is an inclusion. Let $g(x) = b_mx^m + \cdots + b_1x + b_0 \in A[x]$, with $b_m \neq 0$. Then $g(x)f(x) = b_mx^{m+n} + \cdots + b_0a_0 \notin A$, since b_mx^{m+n} is nonzero. Thus $(f(x)A[x]) \cap A = (0)$ so $A \hookrightarrow B$. $\square$

Corollary 2.3 *Let A be a commutative ring with 1 and $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ be a monic polynomial in $A[x]$. Then there is an extension ring B of A and elements $x_i \in B$ such that $f(x) = (x-x_1)(x-x_2)\cdots(x-x_n)$.*

Lemma 2.4 *Let $A \subset B$ be an extension of commutative rings with 1 and $f(x) \in A[x]$ be a monic polynomial. If $f(x) = g(x)h(x)$, where g, h are monic polynomials in $B[x]$, then $g, h \in ({}_B\overline{A})[x]$. That is, the coefficients of g and h are integral over A.*

Proof In some extension ring C of B, $g(x)$ and $h(x)$ factor as the product of monic linear factors. That is, $g(x) = (x-a_1)(x-a_2)\cdots(x-a_n)$ and $h(x) = (x-b_1)(x-b_2)\cdots(x-b_m)$, where $a_i, b_i \in C$. We then have $f(a_i) = f(b_i) = 0 \forall i$. Thus, $a_i, b_i \in {}_C\overline{A}$. Now the coefficients of g are elementary symmetric functions in the a_i, and hence $g \in ({}_C\overline{A})[x]$. But a_i are also in B since $g \in B[x]$. Therefore, $g(x) \in ({}_B\overline{A})[x]$ as claimed. Similarly, $h(x) \in ({}_B\overline{A})[x]$. $\square$

EXAMPLE

$A = k[t^2, t^3] \subset B = k[t]$. $f = (x^2 - t^2) = (x - t)(x + t)$ in $B[x]$. By Lemma 2.4 t is integral over A (which is obvious anyway since $t^2 \in A$).

Theorem 2.5 *Let $A \subset B$ and suppose $f \in B[x]$ is integral over $A[x]$. Then $f \in ({}_B\overline{A})[x]$.*

Corollary 2.6 ${}_{B[x]}\overline{A[x]} = ({}_B\overline{A})[x]$.

PROOF OF THEOREM 2.5

Proof Suppose f is a root of $F(y) = y^n + a_{n-1}y^{n-1} + \cdots + a_1 y + a_0$, where $a_i \in A[x]$. Take $N > \deg f$ so that $-g = -f + x^N \in B[x]$ is monic in x. That is, $g = f - x^N$. Then g is a root of $G(y) = F(y + x^N) \in A[x][y]$. Now $G(y) = y^n + b_{n-1}y^{n-1} + \cdots + b_1 y + b_0$, where $b_i \in A[x]$. Thus, $g^n + b_{n-1}g^{n-1} + \cdots + b_1 g + b_0 = 0$, where everything lives in $B[x]$. Equivalently, $-g(g^{n-1} + b_{n-1}g^{n-2} + \cdots + b_1) = b_0$. Now $b_0 \in A[x]$ and $b_0 = G(0) = F(x^N)$ which is monic if $N > \deg a_i$. Therefore, $g^{n-1} + b_{n-1}g^{n-2} + \cdots + b_1$ is monic. Thus, by Lemma 2.4, $g \in ({}_B\overline{A})[x]$ and hence $f = g + x^N \in ({}_B\overline{A})[x]$ as required. $\square$

Gregor Kemper raised the following question:

Question: Let $L_1, \ldots, L_r$ be linear subspaces of k^{n+1}, with ideals $I_1, \ldots, I_r$ respectively ($I_i \subset k[X_0, \ldots, X_n]$). Let $I = I_1 \cap \cdots \cap I_r$. Find a bound on the degrees of the generators of I.

If the conjectured bound r holds then Gregor informs us that great things will happen in invariant theory ([3]). No counterexample is known.

The conjectured bound is easily seen to hold if all the L_i are one dimensional (corresponding to r points in $\mathbb{P}^n_k$, WLOG none on the hyperplane $X_0 = 0$). For let $R = k[X_0, \ldots, X_n]$, $S = k[X_1, \ldots, X_n]$, and $J = (I, X_0)/(X_0)$, J an ideal of S. Then by the graded Nakayama lemma $f_1, \ldots, f_d$ generate I if and only if their canonical images $\overline{f}_1, \ldots, \overline{f}_d$ generate J. The Hilbert function of S/J is the difference sequence of the Hilbert function of R/I, hence becomes 0 one degree after the Hilbert function of the points becomes constant. If there are r points the Hilbert function of R/I becomes r in at most degree $r - 1$, so the largest degree generator of J, and hence also of I, is in degree r. This maximum occurs if and only if the points lie in a line.

Lecture 3

Definition 3.1 A ring A is said to be $\mathbb{Z}$-*graded* if A can be expressed as $A = \oplus_{i\in\mathbb{Z}} A_i$ as abelian groups, with $A_iA_j \subseteq A_{i+j}$.

For any $a \in A$ we can write $a = \sum_{i\in\mathbb{Z}} a_i$, with $a_i \in A_i, a_i \neq 0$ for only a finite number of i.

If we define the map $j_A : A \to A[x, x^{-1}]$ by taking $a \in A$ to the finite sum $\sum a_i x^i$, then we get the following diagram of ring homomorphisms.

$$\begin{tikzcd} A \arrow[r] \arrow[d, "j_A"] & B \arrow[d, "j_B"] \\ A[x, x^{-1}] \arrow[r] & B[x, x^{-1}] \end{tikzcd}$$

Theorem 3.2 *Let $A \subset B$ be an extension of $\mathbb{Z}$-graded rings. Then ${}_B\overline{A}$ is $\mathbb{Z}$-graded (compatibility with gradings on A and B). Moreover, if B is reduced and $A_i = 0$ for $i < 0$, then $({}_B\overline{A})_i = 0$ for $i < 0$.*

Proof Take $f \in {}_B\overline{A}$. Then $j_B(f)$ is integral over $A[x, x^{-1}]$. Let $(j_B(f))^n + a_{n-1}(j_B(f))^{n-1} + \cdots + a_1(j_B(f)) + a_0 = 0$, where $a_i \in A[x, x^{-1}]$. There exists $N \geq 0$ such that $x^N j_B(f) \in B[x]$ and $x^N a_i \in A[x] \forall i$. Then we have $(x^N j_B(f))^n + x^N a_{n-1}(x^N j_B(f))^{n-1} + \cdots + x^{N(n-1)} a_1 (x^N j_B(f)) + x^{Nn} a_0 = 0$ is an equation of integral dependence of $x^N j_B(f) \in B[x]$ over $A[x]$. By Theorem 2.5 the coefficients of $x^N j_B(f)$ are in ${}_B\overline{A}$. But these are just the components of f. Therefore, ${}_B\overline{A}$ is a graded subring of B. The last assertion is easy to prove and is left as an exercise for the reader. $\square$

REFERENCE: Theorem 3.2 is Proposition 20 in Section 1 of [1].

Theorem 3.3 *If A is a graded integral domain $A = \oplus_{i\in\mathbb{Z}} A_i, K =$ quotient field of A, then ${}_K\overline{A}$ is graded (compatibility with the grading on A).*

Proof Let A be an integral domain such that A is $\mathbb{Z}$-graded, K be the quotient field of A, and d = gcd of all degrees of non - zero homogeneous elements of A. By dividing all degrees by d we can assume $d = 1$. Define S to be the set of non - zero homogeneous elements of A. If $f = s_1/s_2$, where $s_2 \neq 0$ is homogeneous in A, we can define $\deg f = \deg s_1 - \deg s_2$. Let $K_0 = \{s_1/s_2 | \deg s_1 = \deg s_2, s_1, s_2 \in S\}$ and note that K_0 is a subfield of K. Furthermore, A_S contains an element f of degree 1 and $A_S \cong K_0[f, f^{-1}]$, where f is an indeterminate over k_0. Thus ${}_{A_S}\overline{A}$ is graded by Theorem 3.1, and equals ${}_K\overline{A}$ since $K_0[f, f^{-1}]$ is integrally closed. □

Theorem 3.4 *If A is a reduced noetherian graded ring, K = total ring of fractions of A, then ${}_K\overline{A}$ is graded (compatibility with the grading on A).*

Proof If we assume that A is reduced noetherian then $A \hookrightarrow \prod_{i=1}^r (A/\mathfrak{p}_i)$, where $\mathfrak{p}_i$ are minimal primes. If A is also $\mathbb{Z}$-graded, then the $\mathfrak{p}_i$ are homogeneous (see [10, pages 153-154]). So the $A/\mathfrak{p}_i$ are graded integral domains with quotient field K_i, and $K = \prod_{i=1}^r K_i$ is the ring of fractions of A. So we now have ${}_K\overline{A} = \prod_{i=1}^r (\overline{A/\mathfrak{p}_i})$, where $\overline{A/\mathfrak{p}_i} = {}_{K_i}(\overline{A/\mathfrak{p}_i})$ is graded by Theorem 3.3. Thus, ${}_K\overline{A}$ is graded. □

Suppose now that A is not reduced with K being the total ring of fractions. If $x \in K$, x nilpotent, then x is integral over A, so $x \in {}_K\overline{A}$.

There is no reason to expect ${}_K\overline{A}$ to be gradeable and indeed there are counterexamples in [6].

COMPLETE INTEGRAL CLOSURE

Recall one of our characterizations of integral closure form Theorem 1.2 (2): $b \in B$ is integral over A if and only if the A-submodule $A[b]$ of B is a finitely generated A-module. We now weaken this to the condition that $A[b]$ is contained in a finitely generated A-submodule of B. We call such an element b *almost integral* over A (this notation is used in [5]).

Definition 3.5 (1) The *complete integral closure* of A in B is $A_0 = \{x \in B | x$ is almost integral over $A\}$.

(2) A is *completely integrally closed* in B if $A = A_0$.

(3) The *complete integral closure of* A is the complete integral closure of A in its total ring of fractions.

EXAMPLE

Let D be an integrally closed domain with quotient field K. For example we can use $D = \mathbb{Z}, K = \mathbb{Q}$. Let $A = D \oplus xK[x] = \{a_0 + a_1x + \cdots + a_nx^n | a_0 \in D, a_i \in K, i \geq 1\}$. Thus we see that the quotient field of A is $K(X)$. The reader is left with the exercise of showing that A is integrally closed. We claim that the complete integral closure of A is $K[x]$. A partial proof of this claim is as follows. If $a/b \in K - D$, then $(a/b)^i, i \geq 0$, all lie in the finitely generated A-module $AX^{-1} \subset K[X]$.

We have the following pullback diagram which is one natural way in which A can arise:

$$\begin{tikzcd} A \arrow[r] \arrow[d] & K[x] \arrow[d, "x \mapsto 0"] \\ D \arrow[r] & K \end{tikzcd}$$

Theorem 3.6 *Let $A \subset B$.*

(1) *The complete integral closure A_0 of A in B is a subring of B containing A.*

(2) ${}_B\overline{A} \subset A_0$.

(3) *If A is noetherian, then ${}_B\overline{A} = A_0$.*

(4) *A_0 need not be completely integrally closed in B.*

Proof We first prove (1). If $x, y \in A_0$, then $A[x] \subseteq \mathcal{M}_1, A[y] \subseteq \mathcal{M}_2$, where $\mathcal{M}_1, \mathcal{M}_2$ are finitely generated A-submodules of B containing 1. Then $x + y, xy \in \mathcal{M}_1\mathcal{M}_2$, which is finitely generated over A by products of generators. So $x + y, xy \in A_0$ and A_0 is a ring.
Proofs of (2) and (3) are easy and the details are left to the reader. Proof of (4) is done by a counterexample which can be found in [5, page 133]. $\square$

OTHER FACTS

Let $A \subset B$ and A_0 be the complete integral closure of A in B.

(1) A_0 is integrally closed in B.

(2) If B is the total ring of fractions of A, then $x \in A_0 \Leftrightarrow \exists b \in A$, where b is a non-zero divisor such that $bx^i \in A \forall i \geq 0$.

NOTE: For (2) above we can take b to be a common denominator of the generators of the finitely generated A-submodule of B that contains $A[x]$.

Lecture 4

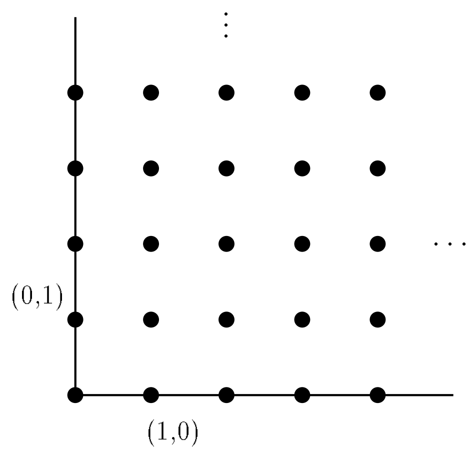
EXAMPLES

Let k be a field.

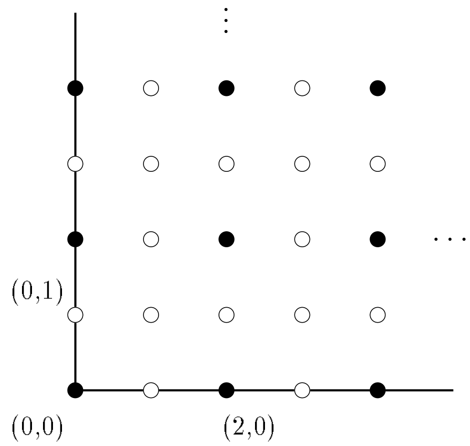
(1) $A = k[x^2, y^2] \subset C = k[x, y]$. C is integral over A, i.e. ${}_C\overline{A} = C$. However, A is integrally closed (in its own field of fractions) since A is just a polynomial ring in the indeterminates x^2 and y^2.

(2) Now consider $A = k[x^2, y^2] \subset B = k[x^2, xy, y^2] \subset C = k[x, y]$. C is integral over B. We claim that B is integrally closed (in its own field of fractions). We will see this later, but in the meantime the reader may wish to find their own proof. Observe that the fields of fractions of $A \subset B \subset C$ are given by $k(x^2/y^2, y^2) \subset k(x/y, y^2) \subset k(x, y)$.

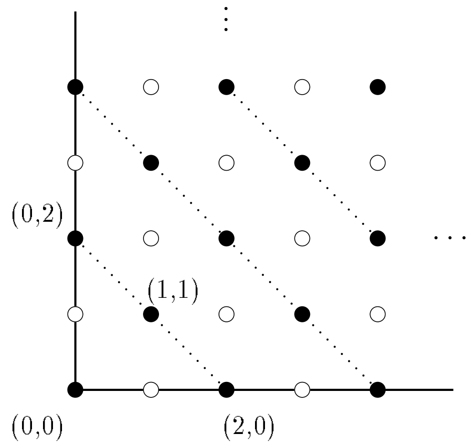
(3) $C = k[x, y]$ has k-basis $x^i y^j$, for $i, j \in \mathbb{N}$, where $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$. We can then identify $x^i y^j$ with $(i, j) \in \mathbb{R}^2$. Thus the above basis of C is identified with the points $\mathbb{N}^2$ in the first quadrant with integer co-ordinates. These correspond to the solid dots below.



Similarly, $A = k[x^2, y^2]$ has k-basis $\{x^i y^j \mid i, j \text{ are both even } \}$ which correspond to the solid dots below.



Finally, a k-basis for $B = k[x^2, xy, y^2]$ is $\{x^i y^j \mid i + j \text{ is even }\}$ which corresponds to the solid dots below.



NOTES

(1) Multiplication of monomials corresponds to addition in $\mathbb{N}^2$. Namely $(x^a y^b)(x^c y^d) = x^{a+c} y^{b+d} \leftrightarrow (a, b) + (c, d) = (a + c, b + d)$.

(2) $\mathbb{N}^2$ has a commutative, associative operation $+$ with unit element $(0, 0)$.

These observations motivate the following definitions.

Definition 4.1 (1) A (commutative) *monoid* is a set S with one commutative, associative operation $+$ with unit element 0 (i.e. $0+s=s+0=s\,\forall s \in S$).

(2) The *monoid ring* $k[S]$ as a k-module is free with basis S. We will write the basis element corresponding to $s \in S$ as X^s. Then multiplication is defined by $(aX^s)(bX^t) = abX^{s+t}$.

IN OUR EXAMPLES

(1) In $C = k[\mathbb{N}^2]$ we have $x^{(a,b)} \leftrightarrow x^a y^b$.

(2) $B = k[x^2, xy, y^2] = k[S]$, where S is the submonoid of $\mathbb{N}^2$ defined by $S = \{(i,j) \in \mathbb{N}^2 \mid i+j \text{ is even }\}$.

(3) $A = k[x^2, y^2] = k[R]$, where R is the submonoid of $\mathbb{N}^2$ defined by $R = \{(a,b) \in \mathbb{N}^2 \mid a, b \text{ are both even }\}$.

We will mostly work with submonoids of $\mathbb{Z}^n$, but we will give a few general definitions and results first.

Definition 4.2 A (commutative) monoid S is called *cancellative* if $s+t = s+u \implies t = u\,\forall s,t,u \in S$.

Given any (commutative) monoid S we can define the quotient group G of S as follows. G consists of pairs (a,b), with $a, b \in S$, modulo the equivalence relation $(a,b) \equiv (c,d) \Leftrightarrow a+d = b+c$ and operation defined by $(a,b)+(c,d) = (a+c, b+d)$. Note that (a,b) is intended to represent $a-b$. We also have that the inverse of (a,b) is (b,a). In addition, the 0 element is the class of $(0,0) \equiv (a,a)\,\forall a$. We should also mention that there is a canonical homomorphism of monoids $S \to G$ given by $s \mapsto (s,0)$.

G is characterized by its universal mapping property. That is, if H is an abelian group and $S \to H$ is a homomorphism of monoids, then there is a unique homomorphism of abelian groups $G \to H$ such that the following diagram commutes

$$\begin{tikzcd} S \arrow[d] \arrow[dr] & \\ G \arrow[r] & H \end{tikzcd}$$

NOTE: From this point on we will drop the phrase "(commutative) monoid" and assume that the monoids referred to are commutative.

CLAIM: S is cancellative $\Leftrightarrow$ the canonical map $S \to G$ is an inclusion.

CLAIM: If S is a cancellative monoid and $R \subset S$ is a submonoid, then the induced homomorphism on the quotient groups is an inclusion.

Definition 4.3 A monoid S is *finitely generated* if there exist a finite number of elements $s_1, s_2, \dots, s_m \in S$ such that for any $s \in S$, $s = a_1 s_1 + a_2 s_2 + \cdots + a_m s_m$, where $a_i \in \mathbb{N}$.

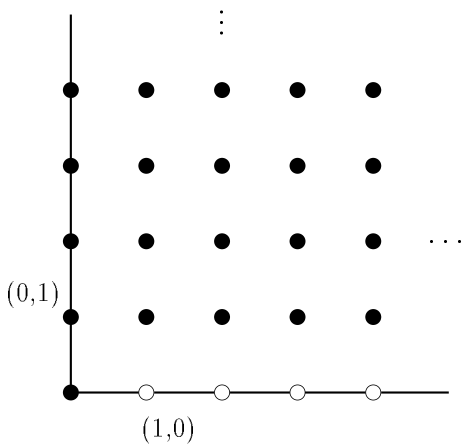
IN OUR EXAMPLES

(1) $\mathbb{N}^2$ is generated by $\{(1,0),(0,1)\}$.

(2) $R = \{(a,b) \mid a, b \text{ are both even }\}$ is generated by $\{(2,0),(0,2)\}$.

(3) $S = \{(a,b) \mid a+b \text{ is even }\}$ is generated by $\{(2,0),(1,1),(0,2)\}$.

(4) The submonoid $\{(a,j) \in \mathbb{N}^2 \mid j > 0\} \cup \{(0,0)\}$ of $\mathbb{N}^2$, which correspond with the solid dots below, is not finitely generated since it requires all the generators $(i,1)$, for $i \geq 0$.



Definition 4.4 (1) A non-zero element s in a monoid S is called *irreducible* if s cannot be written as the sum of two other non-zero elements of S.

(2) An element s in a monoid S is called *invertible* if $-s \in S$.

NOTE: A submonoid of $\mathbb{Z}^n$ might or might not contain non-zero invertible elements.

EXAMPLES

(1) In $S = \{(a, b) \in \mathbb{N}^2 \,|\, a + b \text{ is even }\}$, the element $(1, 1)$ is irreducible.

(2) $\mathbb{N}^2$ has no non-zero invertible elements.

(3) The upper-half plane given by $\{(i, j) \in \mathbb{Z}^2 \,|\, j \geq 0\}$ does have non-zero invertible elements, namely the points $(i, 0)$, where $i \neq 0$, are non-zero invertibles.

REFERENCE FOR MONOIDS: For more material on monoids see [4].

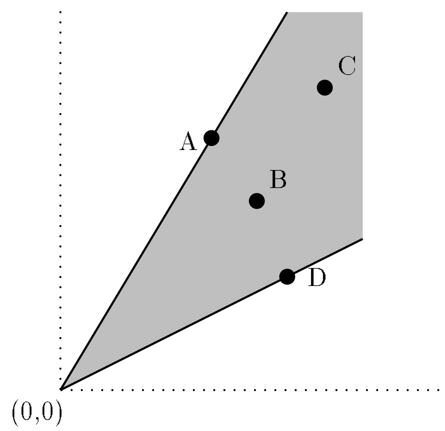
Lecture 5

Definition 5.1 A non-empty subset D of $\mathbb{R}^n$ is called a *cone* if it is closed under positive real linear combinations. That is, for $\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_n \in D$ and $\lambda_i \in \mathbb{R}_+ = \{\lambda \in \mathbb{R} | \lambda \geq 0\}$ we have $\lambda_1 \underline{s}_1 + \lambda_2 \underline{s}_2 + \cdots + \lambda_n \underline{s}_n \in D$.

Definition 5.2 If $S \subset \mathbb{R}^n$ then $\mathbb{R}_+ S = \{\lambda_1 \underline{s}_1 + \cdots + \lambda_m \underline{s}_m | \lambda_i \in \mathbb{R}_+\}$ is the smallest cone containing S and is called the *cone generated by* S.

EXAMPLES

(1) Consider the following cone



Here $\mathbb{R}_+\{A, B, C, D\} = \mathbb{R}_+\{A, D\}$ is the shaded area.

(2) Let a typical element $\underline{x} \in \mathbb{R}^2$ be denoted $\underline{x} = (x_1, x_2)$. Then $\{\underline{x} \in \mathbb{R}^2 | x_2 \geq 0\}$ needs three generators. Three such generators are $(1, 0)$, $(-1, 0), (0, 1)$.

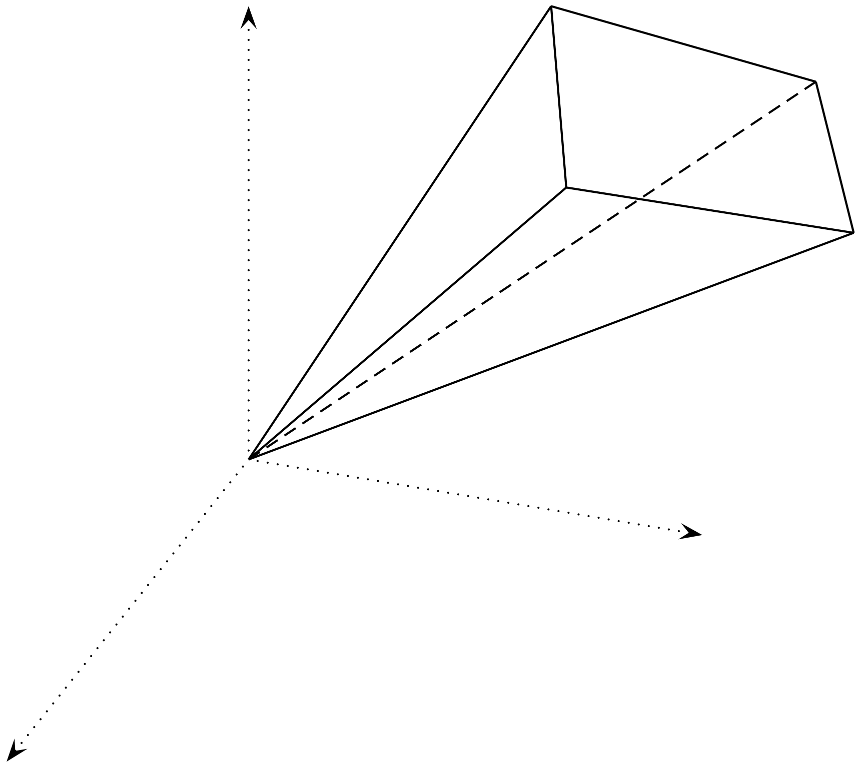
(3) $\mathbb{R}^2$ itself is a cone which needs three generators such as $(1, 1), (1, -1)$, $(-1, 0)$.

Definition 5.3 A cone is *finitely generated* if it is generated by a finite set.

NOTES

(1) Observe that if a cone is finitely generated then it is closed.

(2) For $n \geq 3$ there can be an arbitrarily large number of generators. For example, the cross section by the plane $x_1+x_2+x_3=1$ can be a convex polygon with arbitrarily many sides.



Definition 5.4 The set $\{\underline{x} \in \mathbb{R}^n | \underline{a} \cdot \underline{x} \geq 0\}$ for some fixed $\underline{a} \in \mathbb{R}^n, \underline{a} \neq 0$ is called a *closed linear half space* in $\mathbb{R}^n$.

EXAMPLE

If $a = (a_1, a_2, \dots, a_n)$, then $a_1x_1 + a_2x_2 + \cdots a_nx_n = 0$ is a hyperplane through the origin. We take this hyperplane plus one side to get a closed linear half space in $\mathbb{R}^n$.

FACTS

(A) The finitely generated cones in $\mathbb{R}^n$ are exactly the subsets of $\mathbb{R}^n$ that are the intersection of a finite number of closed linear half spaces.

Fact (A) is intuitively "obvious" but a proof does take a bit of work. The proof is omitted here but can be found as Theorem 1.3 in [11, page 30].

Before we state our next fact we require the following definition.

Definition 5.5 A finitely generated cone C in $\mathbb{R}^n$ is *rational* if C can be generated by a finite set $\{\underline{v}_1, \dots, \underline{v}_m\}$, where $\underline{v}_i \in \mathbb{Q}^n$.

(B) C is rational if and only if the equations of the hyperplanes in (A) can be taken with rational coefficients.

We omit the proof of (B) but remark that the proof is easier than the proof of (A).

NOTE: If S is a submonoid of $\mathbb{Z}^n$, then the cone $\mathbb{R}_+S$ is rational.

Theorem 5.6 *Let S be a finitely generated submonoid of $\mathbb{Z}^n$. Then S has no non-zero invertible elements if and only if S lies "strictly" on one side of a hyperplane. That is, if and only if there exists a hyperplane H such that $H \cap \mathbb{R}_+S = \{\underline{0}\}$.*

NOTE: $\mathbb{R}^n - H$ has two connected components. $\mathbb{R}_+S - \{\underline{0}\}$ being connected, must then lie in one of the components (i.e. lies on one side of H).

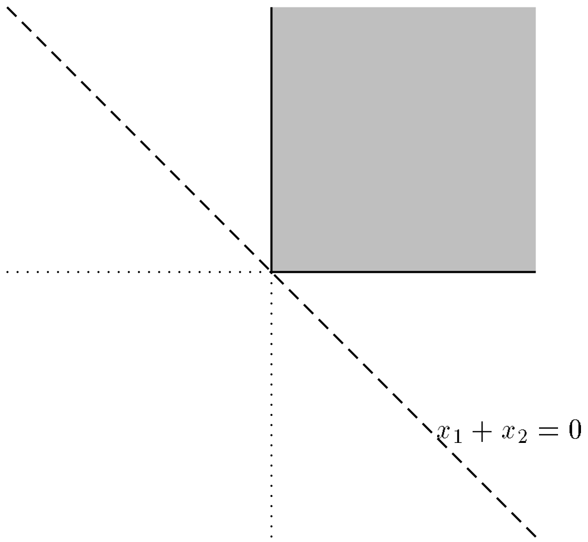
Proof The direction $\Longleftarrow$ is obvious.
Let us now show the direction $\Longrightarrow$. Suppose $S \cap (-S) = \{\underline{0}\}$. We know by (A) that $\mathbb{R}_+S = \{\underline{x} \in \mathbb{R}^n | \underline{x} \cdot \underline{a}_i \geq 0\}$ for some set $\underline{a}_1, \dots, \underline{a}_m$ of elements of $\mathbb{Q}^n$. We claim that the $\underline{a}_i$ must span $\mathbb{R}^n$. To see this assume otherwise. Then $E = \{\underline{x} | \underline{x} \cdot \underline{a}_i = 0\}$, for $1 \leq i \leq m$, is of dimension greater than zero

and is contained in both $\mathbb{R}_+S$ and $\mathbb{R}_+(-S)$. But now since E is rational there exists $\underline{v} \in E \cap \mathbb{Q}^n, \underline{v} \neq 0$. Then $\underline{v} \in (\mathbb{R}_+S) \cap \mathbb{Q}^n \implies \underline{v} \in \mathbb{Q}_+S$. The details of this last implication are left to the reader. Similarly, $\underline{v} \in \mathbb{Q}_+(-S)$. Clearing denominators we get $n\underline{v} \in S \cap (-S)$ for some positive integer $n \neq 0$. This gives a contradiction and we can therefore conclude that $\underline{a}_i$ span $\mathbb{R}^n$. We now have $\mathbb{R}_+S = \{\underline{x}|\underline{x} \cdot \underline{a}_i \geq 0\}$, where $\underline{a}_i \in \mathbb{Q}^n$ for $1 \leq i \leq m$, and the $\underline{a}_i$ span $\mathbb{R}^n$. We claim that $\mathbb{R}_+S$ lies strictly on one side of the hyperplane $\{H = \underline{x}|(\underline{a}_1 + \cdots + \underline{a}_m) \cdot \underline{x} \geq 0\}$. For suppose that $\underline{x} \in \mathbb{R}_+S \cap H$. Then $\underline{a}_i \cdot \underline{x} \geq 0$ since $\underline{x} \in \mathbb{R}_+S$ and $\underline{a}_1 \cdot \underline{x} + \underline{a}_2 \cdot \underline{x} + \cdots + \underline{a}_m \cdot \underline{x} = 0$ since $\underline{x} \in H$. This implies that $\underline{a}_i \cdot \underline{x} = 0$ for all i. Hence $\underline{x} = \underline{0}$ since the $\underline{a}_i$ span $\mathbb{R}^n$. $\square$

Definition 5.7 A cone is *d-dimensional* if the smallest linear subspace containing it is of dimension d.

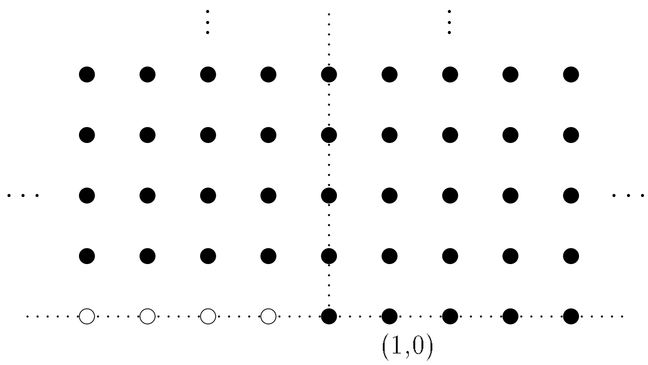
EXAMPLES

(1) For the cone given by $x_1 \geq 0, x_2 \geq 0$ pictured below, we have $\underline{a}_1 = (1,0)$ and $\underline{a}_2 = (0,1)$, so that $\underline{a}_1 + \underline{a}_2 = (1,1)$ and the cone lies strictly on one side of the hyperplane $x_1 + x_2 = 0$.

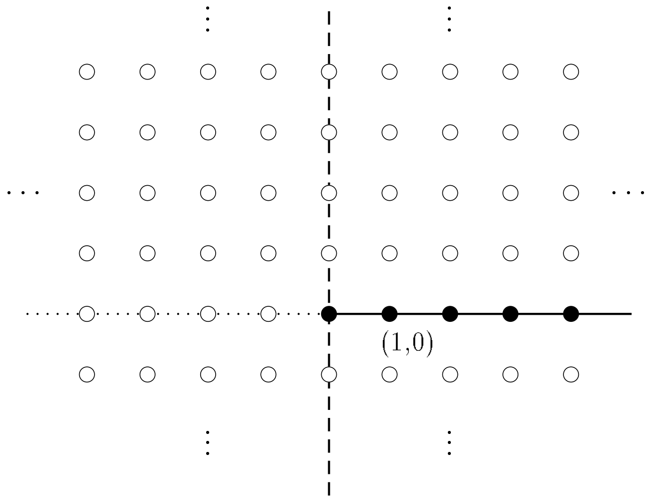


(2) The semigroup S with generators $(1,0)$ and $\{(i,1)|i \leq 0\}$ is not finitely generated and contains no non-zero invertible elements. S does not lie

strictly on one side of a hyperplane, showing that Theorem 5.6 requires that S be finitely generated. The semigroup S is illustrated by the solid dots in the diagram below. Thanks to Les Reid for this example.



(3) The non-negative half of the x_1-axis in $\mathbb{R}^2$ is described by inequalities $x_1 \geq 0, x_2 \geq 0$ and $-x_2 \geq 0$. The $\underline{a}_i$'s are given by $(1,0),(0,1),(0,-1)$. This cone is one dimensional. Here $\underline{a}_1 + \underline{a}_2 + \underline{a}_3 = (1,0)$, and the cone lies strictly on one side of the hyperplane $x_1 = 0$. The integer points in the cone are just $\mathbb{N}$.



Lecture 6

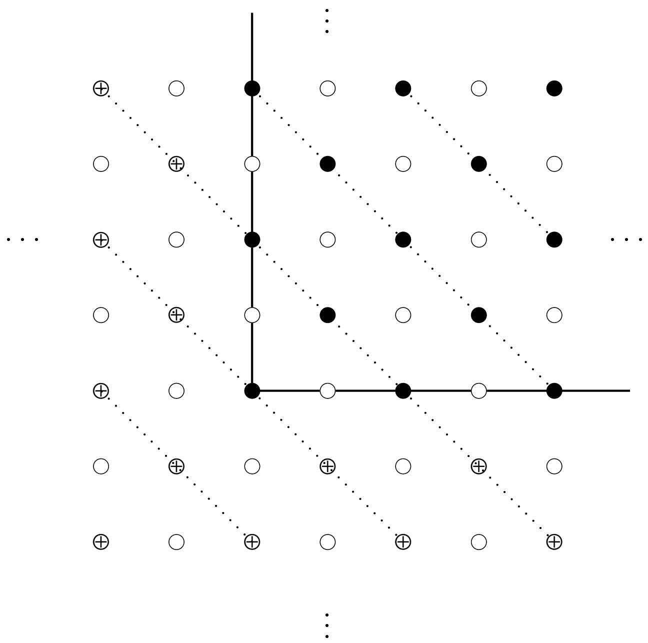
Definition 6.1 A cancellative monoid S (with quotient group G) is called *normal* if $x \in G, nx \in S \implies x \in S$, where $n \in \mathbb{Z}, n > 0$.

RECALL: If S is cancellative then $S \subset G$.

EXAMPLES

(1) $\mathbb{N}^2$ is normal. In this case $G = \mathbb{Z}^2$.

(2) $\{(x, y) \in \mathbb{N}^2 | x + y \text{ is even}\}$ is normal, where $G = \{(x, y) \in \mathbb{Z}^2 | x + y$ is even$\}$. In the figure below S is represented as usual by solid dots and the points of $G \backslash S$ are represented by $\oplus$'s.



Theorem 6.2 *Let S be a submonoid of $\mathbb{Z}^n$ with quotient group G. Then S is normal if and only if $S = G \cap (\mathbb{R}_+ S)$.*

Before proving Theorem 6.2 we need some preliminaries.

Lemma 6.3 *Every subgroup of $\mathbb{Z}^n$ is isomorphic to $\mathbb{Z}^m$, for some $m \leq n$.*

NOTE: Lemma 6.3 applies in particular to the G of Theorem 6.2. We omit the proof of Lemma 6.3. This is a well known property of finitely generated abelian groups.

Theorem 6.4 below is known as *Carathéodory's Theorem*. We give the proof as found in [11].

Theorem 6.4 *Let D be a d-dimensional cone in $\mathbb{R}^m$, generated by $X = \{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n\}$. If $\underline{x} \in D$ then $\underline{x}$ is a positive linear combination of some d elements of X. That is, $\underline{x} = \sum_{i=1}^{d} \lambda_i \underline{x}_{j_i}$, where $\lambda_i \in \mathbb{R}$ and $\lambda_i \geq 0$.*

Proof Proof by contradiction: Suppose that Theorem 6.4 is false. Let the minimal expression for $\underline{x}$ as a positive linear combination of elements of X contain d' elements, where $d' > d$. That is, $\underline{x} = \sum_{i=1}^{d'} \lambda_i \underline{x}_{j_i}$, where $\lambda_i > 0$ and $\{\underline{x}_{j_1}, \underline{x}_{j_2}, \ldots, \underline{x}_{j_{d'}}\} \subseteq X$. Then the $\lambda_i \underline{x}_{j_i}, 1 \leq i \leq d'$ are linearly dependent. There is a relation of the form $\sum_{i=1}^{d'} c_i \lambda_i \underline{x}_{j_i} = 0$, where one of the c_i's is 1 and the rest are ≤ 1. Then $\underline{x} = \sum_{i=1}^{d'} (1 - c_i) \lambda_i \underline{x}_{j_i}$ is an expression for $\underline{x}$ as a positive linear combination of fewer than d' elements. This is a contradiction and thus Carathéodory's Theorem follows. $\square$

Lemma 6.5 *If S and G are as in Theorem 6.2, then $G \cap (\mathbb{R}_+ S) = G \cap (\mathbb{Q}_+ S)$, where $\mathbb{Q}_+ = \{x \in \mathbb{Q} | x \geq 0\}$.*

Proof The inclusion $G \cap (\mathbb{R}_+ S) \supseteq G \cap (\mathbb{Q}_+ S)$ is obvious. So take $\underline{x} \in G \cap (\mathbb{R}_+ S)$. By Carathéodory's Theorem, $\underline{x} = \sum_{i=1}^{m} \lambda_i \underline{s}_i$, where $\lambda_i \in \mathbb{R}_+$ and the $\underline{s}_i$ are linearly independent generators of S. We have $\underline{x}, \underline{s}_i \in \mathbb{Q}^n \subset \mathbb{R}^n$. Now $\{\underline{s}_i\}$ can be extended to a $\mathbb{Q}$-basis of $\mathbb{Q}^n$ (which is also a $\mathbb{R}$-basis of $\mathbb{R}^n$.) From this it follows that $\lambda_i \in \mathbb{Q}_+$ and the lemma follows. $\square$

PROOF OF THEOREM 6.2

Proof ($\Longleftarrow$). Suppose $S = G \cap (\mathbb{R}_+ S)$ and $\underline{x} \in G, n\underline{x} \in S$, with $n \in \mathbb{Z}, n > 0$. Since $n > 0$ and $\mathbb{R}_+ S$ is a cone we have $n\underline{x} \in \mathbb{R}_+ S \implies \underline{x} \in \mathbb{R}_+ S$. Now $x \in G \cap (\mathbb{R}_+ S) = S$. That is, S is normal. ($\Longrightarrow$). Suppose S is normal.

Clearly $S \subset G \cap (\mathbb{R}_+ S)$. Suppose $\underline{x} \in G \cap (\mathbb{R}_+ S)$. Then $\underline{x} = \sum_{i=1}^{m} \lambda_i \underline{c}_i$, where $\lambda_i \in \mathbb{R}, \lambda_i > 0$, and $\underline{c}_i \in S$. By Lemma 6.5 we can take $\lambda_i \in \mathbb{Q}$. Then clearing denominators we find $N \in \mathbb{Z}, N > 0$ such that $N\underline{x} \in S$. But $\underline{x} \in G$ and S is normal so $\underline{x} \in S$ as we wished to show. $\square$

IN GENERAL: For any submonoid S of $\mathbb{Z}^n$ with quotient group G, the proof of Theorem 6.2 shows that $G \cap (\mathbb{R}_+ S) = \{x \in G | nx \in S$ for some integer $n > 0\}$. We call the set $\{x \in G | nx \in S$ for some integer $n > 0\}$ the *normalization of* S.

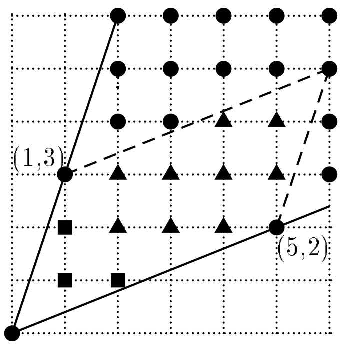
Theorem 6.6 *Let G be a subgroup of $\mathbb{Z}^n$ and let D be a finitely generated rational cone in $\mathbb{R}^n$. Then $S = G \cap D$ is a finitely generated submonoid of $\mathbb{Z}^n$ (which is normal by Theorem 6.2).*

Proof WLOG we can assume that G is of rank n and that D is generated by $\underline{s}_1, \underline{s}_2, \dots, \underline{s}_m$, with $\underline{s}_i \in G$. Let $T = \{\lambda_1 \underline{s}_1 + \lambda_2 \underline{s}_2 + \cdots + \lambda_m \underline{s}_m | 0 \le \lambda_i < 1\}$. Then S is generated by $\{\underline{s}_1, \underline{s}_2, \dots, \underline{s}_m\}$ and the non-zero elements of $G \cap T$. But G is discrete and T is bounded so $G \cap T$ is finite. Therefore, G has a finite number of generators. $\square$

EXAMPLE

Let $D \subset \mathbb{R}^2$ be the cone generated by $(1,3)$ and $(5,2)$. Then T is the parallelogram illustrated below, not including its dashed edges. $T \cap \mathbb{Z}^2$ contains 12 interior points, together with $(0,0)$. The generators of $D \cap \mathbb{Z}^2$ are $(1,3)$, $(5,2)$, and the three square dots. The other interior points of T, indicated by triangles, are by inspection positive integer linear combinations of these points. We note that

$$\left|\det\begin{pmatrix} 1 & 3 \\ 5 & 2 \end{pmatrix}\right| = 13.$$



Lecture 7

In this lecture we relate normal monoids to integrally closed rings. First we prove two preliminary lemmas:

Lemma 7.1 *Let C be a normal, finitely generated submonoid of $\mathbb{Z}^n$ and let C_0 be the subgroup of invertible elements of C. Then $C \cong C_0 \oplus C'$, where C' is a normal monoid with no non-zero invertible elements.*

Proof Let G be the quotient group of C. Then G/C_0 is torsion free since C is normal. This detail is left as an exercise. Hence, $G/C_0 \cong \mathbb{Z}^m$ for some m. We then get the split exact sequence of abelian groups

$$\begin{tikzcd} 0 \arrow[r] & C_0 \arrow[r, "i"] & G \arrow[l, bend left, dotted, "g"] \arrow[r, "\pi"] & \mathbb{Z}^m \arrow[r] & 0 \end{tikzcd}$$

where $g \circ i = 1_{C_0}$. We have a canonical isomorphism $G \cong C_0 \oplus \mathbb{Z}^m$ (given by $g \leftrightarrow (i(g), \pi(g)))$. Define $C' = \pi(C)$, so we get $C = C_0 \oplus C'$. Lastly we check that C' is normal. To do this take $c' \in \mathbb{Z}^m$, such that $dc' \in C'$ for a positive integer d. Then $(0, c') \in G$ and $d(0, c') \in C$. But C is normal and so this implies that $(0, c') \in C$ which implies $c' \in C'$. $\square$

Lemma 7.2 *Let K be a field, where $k \subset K$, and $k \subset D_i$, where the D_i are integrally closed subrings of K (possibly with different fields of fractions). Then $\cap D_i$ is integrally closed.*

Proof Let f be in the field of fractions of $R = \cap D_i$, f is integral over R. Then f is in the field of fractions of each D_i and is integral over D_i. Thus $f \in D_i \forall i \implies f \in R$. $\square$

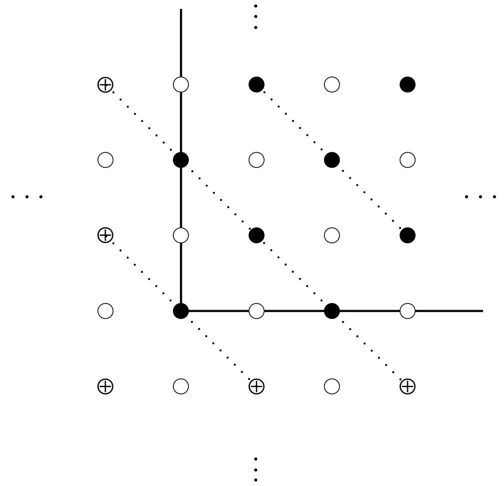
Theorem 7.3 *Let S be a submonoid of $\mathbb{Z}^n$. Then S is normal if and only if $k[S]$ is normal.*

NOTE: Theorem 7.3 is Corollary 12.11 in [4] where the proof is done in the general case. As in [2] we will give the proof where S is a finitely generated submonoid of $\mathbb{Z}^n$. WLOG we may assume the quotient group of S is $\mathbb{Z}^n$.

Proof of Theorem 7.3 First let us prove the direction ($\Leftarrow$). $k[S] \subseteq k[G]$ with the same field of fractions. Suppose $s \in G, ds \in S$ for d a positive integer. Then $(X^s)^d = X^{sd} \in k[S]$. X^s is integral over $k[S]$, and in the field of fractions of $k[S]$. But now $k[S]$ is normal $\implies X^s \in k[S] \implies s \in S$. Now let us show the direction ($\Rightarrow$). So we suppose that S is a normal, finitely generated submonoid of $\mathbb{Z}^n$. Then $\mathbb{R}_+S = \cap_{i=1}^N H_i$, where H_i are rational closed linear half spaces. Since S is normal we then have $S = \mathbb{Z}^n \cap (\mathbb{R}_+S) = \cap_{i=1}^N(\mathbb{Z}^n \cap H_i) \subset k(x_1, x_2, \cdots, x_n)$. Hence, $k[S] = \cap_{i=1}^N k[\mathbb{Z}^n \cap H_i]$. Using Lemma 7.1 it can be shown that $\mathbb{Z}^n \cap H_i \cong \mathbb{Z}^{n-1} \oplus \mathbb{N}$, where $\mathbb{N} = \{0, 1, 2, \cdots\}$. But $k[\mathbb{Z}^{n-1} \oplus \mathbb{N}] = k[x_1, x_1^{-1}, \cdots, x_{n-1}, x_{n-1}^{-1}, x_n]$ which is integrally closed. Thus $k[S]$ is the intersection of integrally closed rings and hence is itself integrally closed by Lemma 7.2. $\square$

EXAMPLES

(1) $k[S] = k[x^2, xy, y^2]$ is integrally closed. Here we recall that $S = \{(i,j) \in \mathbb{N}^2 | i+j \text{ is even}\}, G = \{(i,j) \in \mathbb{Z}^2 | i+j \text{ is even}\}$, and $\mathbb{R}_+S$ is the first quadrant. We see that $(\mathbb{R}_+S) \cap G = S$ and so S is normal. Hence, by Theorem 7.3, $k[S]$ is normal.



We also have in this example that S is the submonoid of $\mathbb{N}^2$ generated by the rows of

$$\begin{array}{cc} c_1 & c_2 \\ \end{array}$$
$$\begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

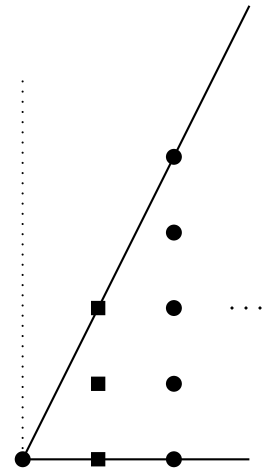
This is isomorphic to the submonoid S' of $\mathbb{N}^2$ generated by the rows of

$$\begin{array}{cc} c_1+c_2 & c_2 \end{array}$$
$$\begin{pmatrix} 2 & 0 \\ 2 & 1 \\ 2 & 2 \end{pmatrix}.$$

The isomorphism $S \cong S'$ is through the mapping $(a-b, b) \leftrightarrow (a, b)$. But this is isomorphic to the submonoid S'' of $\mathbb{N}^2$ generated by the rows of

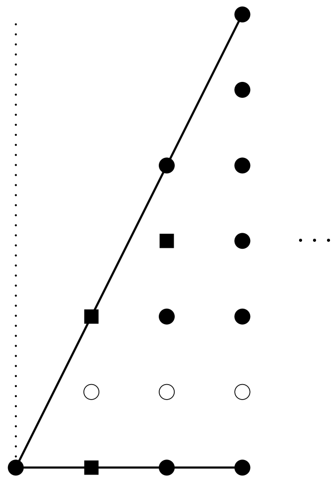
$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

The isomorphisms $S \cong S' \cong S''$ are through the mappings $(2a-b, b) \leftrightarrow (2a, b) \leftrightarrow (a, b)$. We get the following picture for S''.

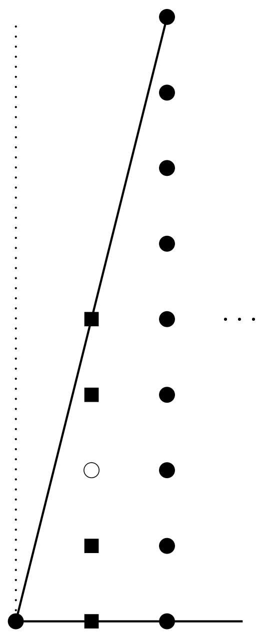


If we then make the following identifications $s \to x^2, t \to x^{-1}y$ we get $k[x^2, xy, y^2] \cong k[s, st, st^2]$. The square dots are the generators of S.

(2) $k[x, xy^2, x^2y^3]$ has integral closure $k[x, xy, xy^2]$. Here we just take all integer points in the cone. Then $\dim_k(k[x, xy, xy^2]/k[x, xy^2, x^2y^3]) = \infty$ with basis being the open circles in the picture.



(3) $k[x, xy, xy^3, xy^4]$ has only xy^2 missing from the cone. The integral closure of $k[x, xy, xy^3, xy^4]$ is $k[x, xy, xy^2, xy^3, xy^4]$ and $\dim_k(k[x, xy, xy^2, xy^3, xy^4]/k[x, xy, xy^3, xy^4]) = 1$.



Moreover, after a change of co-ordinates $k[x, xy, xy^3, xy^4]$ becomes $k[x^4, x^3y, xy^3, y^4]$ which has integral closure $k[x^4, x^3y, x^2y^2, xy^3, y^4]$.

Lecture 8

We conclude this series of lectures by describing (with most proofs omitted) an application of Hilbert functions to a non-trivial combinatorial problem, that of counting points with integer co-ordinates in polytopes.

Let D be a finitely generated rational cone in $\mathbb{R}^n$ and $C = D \cap \mathbb{Z}^n$, a normal submonoid of $\mathbb{Z}^n$. We let $\dim C = \dim k[C] = d$ where $k[C]$ is integrally closed by Theorem 7.3. We also have that $k[C]$ is a finitely generated k-algebra, not necessarily by elements of degree 1, and is graded by $\deg x_i = 1$.

Theorem 8.1 *$A = k[C]$ is Cohen-Macaulay.*

REFERENCE Theorem 8.1 can be found in [2, page 261] and is a famous theorem due to Hochster.

From this point on we will denote Cohen-Macaulay by C.M. We first recall the definition:

Definition 8.2 *A commutative graded ring R is C.M. if R has a regular sequence $f_1, f_2, \ldots, f_d$, where $d = \dim R$. (Furthermore, the f_i can be taken to be homogeneous.)*

RECALL $f_1, f_2, \ldots, f_d$ is a regular sequence if f_i is a non-zero divisor in $R/(f_1, \ldots, f_{i-1}), 1 \leq i \leq d$.

We first note that C.M. has simple consequences for the Hilbert function of A. We see this in the following.

Suppose f_i is of degree e_i. Let $H(A, i) = \dim_k A_i$. Now define $H_A = \sum_{i \geq 0} H(A, i) t^i$. We call H_A the *Hilbert series* of A. We now, for simplicity, take D to be in the first orthant so as to have no negative degrees. Note that f_1 is a non-zero divisor and so $(Af_1)_i = A_{i-e_1} f_1 \cong A_{i-e_1}$ and $H_{Af_1} = t^{e_1} H_A$. We have the exact sequence

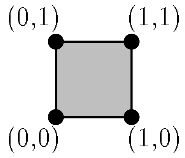
$$0 \to Af_1 \to A \to A/f_1 A \to 0.$$

Thus we see that $\dim_k(A_i) = \dim_k(Af_1)_i + \dim_k(A/f_1A)_i$. This implies that $H_A = H_{Af_1} + H_{A/f_1A}$ and so $H_{A/f_1A} = H_A - H_{Af_1} = (1-t^{e_1})H_A$. Using that we have a regular sequence we can repeat this process obtaining $H_{A/(f_1,\ldots,f_d)A} = (1-t^{e_1})(1-t^{e_2})\cdots(1-t^{e_d})H_A$. And thus we see that $H_A = H_{A/(f_1,\ldots,f_d)A}/(1-t^{e_1})(1-t^{e_2})\cdots(1-t^{e_d})$. Now since $\dim A = d$ we get $A/(f_1,\ldots,f_d)A$ is a finite dimensional k-vector space and so $h(t) = H_{A/(f_1,\ldots,f_d)A}$ is a polynomial. Thus H_A is a rational function. Now $h(t) = 1 + h_1t + h_2t^2 + \cdots + h_et^e$, where $h_e \neq 0$, is the Hilbert function of the 0 dimensional ring $A/(f_1,\ldots,f_d)A$ which implies that $h_i \geq 0, 1 \leq i \leq d$.

The Hilbert function of any finitely generated graded k-algebra (graded in degrees greater than or equal to zero) can be written in the same form $H_A = h(t)/(1-t^{e_1})\cdots(1-t^{e_d})$ where $h(t)$ is a polynomial with $h(1) \neq 0$. However, if A is not C.M. then the h_i need not be greater than or equal to zero and $h(t)$ is not in general the Hilbert function of a k-algebra.

If A is C.M. and generated as a k-algebra by elements of degree 1 (possibly after regrading and changing H_A accordingly) then $1, h_1, h_2, \ldots, h_e$ is an O-sequence implying that $h_i > 0$ for $1 \leq i \leq e$.

We now turn to an application to the *Ehrhart Polynomial*. Consider the square below. This is an example of a *polytope* P.



P

Definition 8.3 (1) *The* convex hull P *of points* $p_1, p_2, \ldots, p_d$ *in* $\mathbb{R}^n$ *is the set of all linear combinations* $\lambda_1p_1 + \lambda_2p_2 + \cdots + \lambda_dp_d$, *where* $\lambda_i \in \mathbb{R}$, $0 \leq \lambda_i \leq 1$, *and* $\lambda_1 + \lambda_2 + \cdots + \lambda_d = 1$.

(2) *A set* $P \subset \mathbb{R}^n$ *is* convex *if given any two points* $A, B \in P$, *the line segment between* A *and* B *lies in* P. *(A convex hull is thus convex.)*

(3) *A* polytope P *is the convex hull of a finite set of points in* $\mathbb{R}^n$ *with integer or rational co-ordinates.*

(4) *Given a polytope* P *we can form the polytope* iP, *where* $i \geq 0$ *is an integer, which consists of all points* $\{ip | p \in P\} = \{p_1 + p_2 + \cdots + p_i | p_j \in P\}$.

In the picture above, P is the convex hull of $(0,0),(1,0),(0,1),(1,1)$.

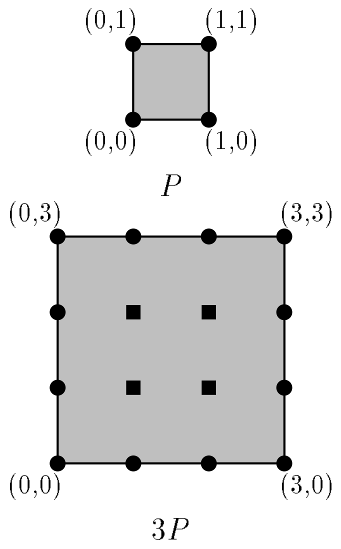
PROBLEM: We now consider the problem of finding the number of points with integer co-ordinates in iP. That is, how many elements are in $(iP) \cap \mathbb{Z}^n$.

NOTATION Given a polytope P we will use the following notation.

(1) $E(P,m)$ denotes the number of points in mP with integer co-ordinates. We call $E(P,m)$ the *Ehrhart function* of P.

(2) $E^+(P,m)$ denotes the number of interior points in mP with integer co-ordinates.

EXAMPLE
We consider the square P mentioned above. In the picture of $3P$ below we indicate interior vertices with square dots.



The square P has vertices $(0,0),(1,0),(0,1),(1,1)$ and mP is the square with vertices $(0,0),(m,0),(0,m),(m,m)$. Observe that $E(P,m) = (m+1)^2$ and $E^+(P,m) = (m-1)^2$.

OBSERVATIONS We make the three following observations on the example above.

(1) $E^+(P,m) = E(P,m-2)$. This appears to be an accident in this case.

(2) $E(P,m)$ and $E^+(P,m)$ are polynomials (which can then be evaluated at negative m). This is not an accident and generalizes with care to other polytopes.

(3) $E^+(P,m) = E(P,-m)$. This also generalizes to other polytopes.

THE EHRHART FUNCTION AS A HILBERT FUNCTION

Given P a polytope in $\mathbb{R}^n$ with rational vertices $\{\underline{v}_1,\ldots,\underline{v}_m\}, \underline{v}_i \in \mathbb{Q}^n \subset \mathbb{R}^n$ we form the cone $D \subseteq \mathbb{R}^{n+1}$ spanned by $\{(\underline{v},1) | \underline{v} \in P\}$. Let $C = D \cap \mathbb{Z}^{n+1}$ and $A = k[C]$. We call A the *Ehrhart Ring* of P and note that A is graded in degrees greater than or equal to 0 by setting the degree $X^{(a,i)} = i$, where $a \in iP, a$ is an integer co-ordinate, and $i \geq 0$ is an integer. Note that the conditions put on a and i are necessary and sufficient for $(a,i) \in C$. We see that $\dim_k A_i =$ the number of integer points in $iP = E(P,i)$ thus interpreting the Ehrhart function as a Hilbert function.

Lecture 9

We begin this final lecture by recalling a claim from Lecture 8. Let A be a finitely generated graded k-algebra, generated by a finite number of homogeneous elements of degree greater than or equal to one. We then have the Hilbert function $H(A,t) = \dim_k A_t$ and Hilbert series $H_A(t) = \sum_{i\geq 0} H(A,i)t^i$. The claim we wish to recall is the following:

$$H_A = h(t)/(1-t^{e_1})(1-t^{e_2})\cdots(1-t^{e_d})$$

where $h(t)$ is a polynomial in which $h(1) \neq 0$.

We know that if A is not C.M., then $h(t)$ may have some coefficients that are zero or negative. In addition, if $h(t) = 1 + h_1 t + h_2 t^2 + \cdots + h_e t^e,\ h_e \neq 0$ then $(1, h_1, \ldots, h_e)$ is called the H-vector of A. However, in general the H-vector won't be the Hilbert function of anything. Our claim above implies that there is a quasi-polynomial $q(m)$ such that $H(A,m) = q(m)$ if $m >> 0$. We now recall the definition of quasi-polynomial:

Definition 9.1 A function $q : \mathbb{Z} \to \mathbb{Z}$ is a *quasi-polynomial* if there exists an integer $g \geq 1$ and polynomials $q_0, q_1, \ldots, q_{g-1}$ such that $q(m) = q_i(m)$ whenever $m \equiv i \mod g,\ 0 \leq i \leq g-1$.

EXAMPLE

Let $A = k[t^2, u^2]$. Then the Hibert function of A is 1 0 2 0 3 0 4 0 5 $\cdots$. We then see that $g = 2$, $q_0(m) = 1 + m/2$, and $q_1(m) = 0$.

The degree of a rational function f/g (f, g polynomials) is defined by $\deg(f/g) = \deg(f) - \deg(g)$. In particular H_A is rational and so $\deg H_A$ is defined. The following fact can be found in [2, page 168].

FACT

Let A be a finitely generated graded k-algebra (in degrees greater than or equal to zero). Let $q(m)$ be the quasi-polynomial such that $H(A,m) = q(m)$ for $m >> 0$. Then degree $H_A = \max\{m | H(A,m) \neq q(m)\}$.

EXAMPLE

Let $A = k[x_0, \ldots, x_n], \deg x_i = 1$. Then $H_A = 1/(1-t)^{n+1}$. Here we have $\deg H_A = 0 - (n+1) = -n-1$, $g = 1$ and $q(i) = \binom{n+i}{n}$. From this we see that $q(i) = 0$ for $i = -1, -2, -3, \ldots, -n$ and $q(i) \neq 0$ for $i = -n-1$. Thus we see that $q(i) = H(A, i)$, $i \geq -n$ and $q(i) \neq H(A, i)$, $i = -n-1$. Hence $\deg H_A = -n-1$ is the largest degree in which $q(i) \neq H(A, i)$ as observed by the fact above.

APPLICATION

We now wish to apply the above to the Ehrhart Series. We begin by recalling definitions and notation from Lecture 8. Let P be a polytope in $\mathbb{R}^n$. We defined $E(P, m)$ to be the number of integer points in mP and $E_P(t) = \sum_{i \geq 0} E(P, i)t^i$. The Ehrhart Ring is then $k[C]$, where $D \subseteq \mathbb{R}^{n+1}$ is the cone spanned by $\{(\underline{v}, 1) | \underline{v} \in P\}$ and $C = D \cap \mathbb{Z}^{n+1}$ is a monoid. Note that D is a finitely generated rational cone which implies that $k[C]$ is a finitely generated k-algebra. In lecture 8 we also saw that $E(P, m) = H(k[C], m)$ where it is understood that $E(P, 0) = 1$. Thus we define the Ehrhart Series to be $E_P(t) = H_{k[C]}(t) = \sum_{i \geq 0} E(P, i)t^i$. With this definition we can apply the above properties of the Hilbert series to the Ehrhart Series.

From the above we can make the following conclusions:

(1) $E_P(t)$ is a rational function.

(2) There is a quasi-polynomial $q(t)$ such that $E(P, m) = q(m)$ for $m >> 0$.

ADDITIONAL FACTS: The proofs of the following facts use that $k[C]$ is C.M. We omit these non-trivial proofs.

(1) $E(P, m) = q(m)$ for all $m \geq 0$. Equivalently $\deg H_A < 0$.

(2) $E^+(P, m) = (-1)^d q(-m)$, where $d = \dim P$ and $E^+(P, m)$ is the number of interior points in mP.

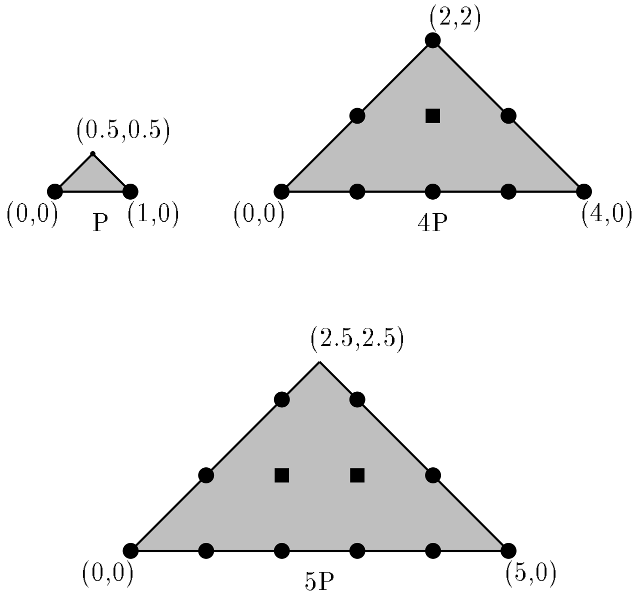
Theorem 9.2 *If the polytope P has vertices with integer co-ordinates (instead of just rational co-ordinates), then $q(t)$ is a polynomial.*

Proof Let $V \subset \mathbb{Z}^n$ be the set of vertices of P, V finite. Define $k[V] = k[X^{(v,1)} | v \in V]$. Thus $k[V] \subset k[C]$. Now if $c \in C$, then there exists $q_v \in \mathbb{Q}_{\geq 0}$ such that $c = \sum_{v \in V} q_v v$. Clearing denominators we get $\lambda c = \sum_{v \in V} q_v \lambda v$ where $\lambda \geq 1$ is an integer which is a common denominator of the q_v's. This implies that $(X^c)^\lambda \in k[V]$. Now $k[C]$ is integral over $k[V]$. We also have that $k[C]$ is a finitely generated k-algebra and hence is also a finitely generated $k[V]$-algebra. This implies that $k[C]$ is a finitely generated $k[V]$-module.

$k[V]$ is "standard". That is, $k[V]$ is generated as a k-algebra by elements of degree 1. Therefore, $k[C]$ has a Hilbert function (see Theorem 4.1.3 of [2]). That is our quasi-polynomial is actually a polynomial as claimed. $\square$

EXAMPLE

Consider the polytope (triangle) P below.



We get the following table.

t	$E(P,t)$		
0	1		
1	2	$=$	$2 = 1 \times 2$
2	4	$=$	$1 + 3$
3	6	$=$	$2 + 4 = 2 \times 3$
4	9	$=$	$1 + 3 + 5$
5	12	$=$	$2 + 4 + 6 = 3 \times 4$
6	16	$=$	$1 + 3 + 5 + 7$

and so on.

We thus see that $E(P,m) = (1 + m/2)^2$ if m is even and $E(P,m) = [(m+1)/2][(m+3)/2]$ if m is odd.

From the even part we get:

$$\begin{array}{rcl}
F &=& 1 + 4t^2 + 9t^4 + 16t^6 + \cdots \\
t^2F &=& t^2 + 4t^4 + 9t^6 + \cdots \\
(1-t^2)F &=& 1 + 3t^2 + 5t^4 + 7t^6 + \cdots \\
t^2(1-t^2)F &=& t^2 + 3t^4 + 5t^6 + \cdots \\
(1-t^2)^2F &=& 1 + 2t^2 + 2t^4 + 2t^6 + \cdots \\
t^2(1-t^2)^2F &=& t^2 + 2t^4 + 2t^6 + \cdots
\end{array}$$

So we see that $(1-t^2)^3F = 1 + t^2$ and hence $F = (1+t^2)/(1-t^2)^3$ for the even part.

Similarly from the odd part we get:

$$\begin{array}{rcl}
G &=& 2t + 6t^3 + 12t^5 + 20t^7 + \cdots \\
G/2 = G' &=& t + 3t^3 + 6t^5 + 10t^7 + \cdots \\
t^2G' &=& t^3 + 3t^5 + 6t^7 + \cdots \\
(1-t^2)G' &=& t + 2t^3 + 3t^5 + 4t^7 + \cdots \\
t^2(1-t^2)G' &=& t^3 + 2t^5 + 3t^7 + \cdots \\
(1-t^2)^2G' &=& t + t^3 + t^5 + t^7 + \cdots
\end{array}$$

Now $(1-t^2)^2G' = t/(1-t^2)$ and so $G' = t/(1-t^2)^3$. That is $(1-t^2)^3G' = t$ and so $G = 2G' = 2t/(1-t^2)^3$.

From the even and odd parts we have that the Ehrhart series for the triangle is $F+G = (1+t^2)/(1-t^2)^3 + 2t/(1-t^2)^3 = (1+2t+t^2)/(1-t^2)^3$. We note that the degree of this Ehrhart series is -4.

References

[1] N. Bourbaki. *Algèbre Commutative*, chapter 5, Entiers. Hermann, 1964.

[2] W. Bruns and J. Herzog. *Cohen-Macaulay Rings*. Cambrige University Press, 1993.

[3] Harm Derksen. Computation of invariants for reductive groups. *Advances in Mathematics*, 141:366–384, 1999.

[4] R. Gilmer. *Commutative Semigroup Rings*. The University of Chicago Press, 1984.

[5] R. Gilmer. *Multiplicative Ideal Theory*, volume 90. Queen's Papers in Pure and Applied Mathematics, 1992.

[6] J. Huckaba and J. Keller. The integral closure of a graded ring. *Communications in Algebra*, 9:1551–1558, 1981.

[7] S. Lang. *Algebra.* Addison-Wesley, second edition edition, 1984.

[8] Les Reid and M.Vitulli. The weak subintegral closure of a monomial ideal. To appear in Comm. in Alg.

[9] Les Reid and Leslie G. Roberts. Monomial subrings and systems of subintegrality. To appear in JPAA.

[10] O. Zariski and P. Samuel. *Commutative Algebra.* D.Van Nostrand, 1960.

[11] G. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics.* Springer-Verlag, 1995.