Spring March 13, 2019

Dissertation_Davis.pdf

brian davis, *University of Kentucky*

ABSTRACT OF DISSERTATION

Brian Davis

The Graduate School
University of Kentucky
2019

Lattice Simplices:
Sufficiently Complicated

ABSTRACT OF DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Arts and Sciences
at the University of Kentucky

By
Brian Davis
Lexington, Kentucky

Director: Dr. Benjamin Braun, Associate Professor of Mathematics
Lexington, Kentucky 2019

ABSTRACT OF DISSERTATION

Lattice Simplices:
Sufficiently Complicated

Simplices are the "simplest" examples of polytopes, and yet they exhibit much of the rich and subtle combinatorics and commutative algebra of their more general cousins. In this way they are sufficiently complicated — insights gained from their study can inform broader research in Ehrhart theory and associated fields.

In this dissertation we consider two previously unstudied properties of lattice simplices; one algebraic and one combinatorial. The first is the Poincaré series of the associated semigroup algebra, which is substantially more complicated than the Hilbert series of that same algebra. The second is the partial ordering of the elements of the fundamental parallelepiped associated to the simplex.

We conclude with a proof-of-concept for using machine learning techniques in algebraic combinatorics. Specifically, we attempt to model the integer decomposition property of a family of lattice simplices using a neural network.

KEYWORDS: Polytopes, Simplices, Poincaré series, Fundamental parallelepiped, Machine Learning

Author's signature: Brian Davis

Date: March 5, 2019

Lattice Simplices:
Sufficiently Complicated

By
Brian Davis

Director of Dissertation: Benjamin Braun

Director of Graduate Studies: Peter Hislop

Date: March 5, 2019

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Brian Davis

The Graduate School
University of Kentucky
2019

Lattice Simplices:
Sufficiently Complicated

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Brian Davis
Lexington, Kentucky

Director: Dr. Benjamin Braun, Associate Professor of Mathematics
Lexington, Kentucky 2019

ACKNOWLEDGMENTS

With thanks to Matt, who opened the door, and Ben, who held it open.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Preliminaries

1.1 Constructions

Definition 1.1.1. *For a collection $\mathcal{A} = \{a_0, \ldots, a_m\}$ of points in $\mathbb{R}^d$, we define their* ***convex hull*** *to be the set*

$$\operatorname{conv}(\mathcal{A}) := \left\{ \sum_{i=0}^{m} \gamma_i a_i \ \textit{ such that } 0 \le \gamma_i \textit{ and } \sum_{i=0}^{d} \gamma_i = 1 \right\} \subset \mathbb{R}^d,$$

so–named because it is the smallest convex set containing $\mathcal{A}$.

Definition 1.1.2. *In the case that $m = d$ and the set $\mathcal{A}^\circ := \{(a_1 - a_0), \ldots, (a_d - a_0)\}$ is a vector space basis of $\mathbb{R}^d$, then we call $\Delta = \operatorname{conv}(\mathcal{A})$ a d-simplex. We call the a_i's the* ***vertices*** *of Δ, and if each a_i is an integer point, i.e., lies in $\mathbb{Z}^d$, we call Δ a* ***lattice simplex****.*

Definition 1.1.3. *For a collection $\mathcal{A} = \{a_0, \ldots, a_m\}$ of points in $\mathbb{R}^d$, we define their* ***conical hull*** *to be the set*

$$\left\{ \sum_{i=0}^{m} \gamma_i a_i \ \textit{ such that } 0 \le \gamma_i \right\} \subset \mathbb{R}^d.$$

Notice that the conical hull is unbounded, as in particular it contains the rays $\mathbb{R}_{\geq 0} a_i$ for $0 \le i \le m$.

We are particularly interested in conical hulls of the following kind:

Definition 1.1.4. *Let $\mathcal{A} = \{a_0, \ldots, a_d\}$ and $\Delta = \operatorname{conv}(\mathcal{A})$ be a lattice d-simplex. Then the* ***cone over*** *Δ is the conical hull of the points $\{(1, a_0), \ldots, (1, a_d)\} \subset \mathbb{R}^{d+1}$, and is denoted $\operatorname{cone}(\Delta)$.*

Definition 1.1.5. *A* ***semigroup*** *is a set Λ with an operation $+$ satisfying the following axioms for all α, β, and σ in Λ:*

$$\alpha + \beta \in \Lambda \quad (\textit{Closure}),$$

$$\alpha + (\beta + \sigma) = (\alpha + \beta) + \sigma \quad (\textit{Associativity}).$$

It is a generalization of an ordinary algebraic group; note that we do not require the existence of either an identity element or inverses.

Definition 1.1.6. *The semigroup $(\Lambda, +)$ associated to a d-simplex Δ is the intersection*

$$\Lambda := \operatorname{cone}(\Delta) \cap \mathbb{Z}^{d+1},$$

with $+$ given by the usual coordinate-wise addition on $\mathbb{Z}^{d+1}$.

Ehrhart Theory

Our interest in the semi-group Λ has a historical basis. In the 1960's, Eugène Ehrhart (a High School math teacher) proved the following result in the paper [10]:

Theorem 1.1.7. *Let $\mathcal{A} = \{a_0, \ldots, a_m\}$ be a collection of integer points (each a_i has all integer coordinates), and $n\mathcal{A} = \{n\cdot a_0, \ldots, n\cdot a_m\}$. Define* $\mathrm{ehr}(n)$ *to be the function which maps a non-negative integer n to the cardinality of the set*

$$\mathrm{conv}(n\mathcal{A}) \cap \mathbb{Z}^d.$$

Then $\mathrm{ehr}(n)$ *is given by a polynomial in n of degree d.*

We call $\mathrm{ehr}(n)$ the Ehrhart polynomial of $\mathrm{conv}(A)$, and define a formal power series called the Ehrhart series by

$$\mathrm{Ehr}(t) := \sum_{n\geq 0} \mathrm{ehr}(n)\, t^n.$$

Stanley showed in [22, Cor. 1.3] that because $ehr(n)$ is a polynomial in n of degree d, we have that

$$\mathrm{Ehr}(t) = \frac{f(t)}{(1-t)^{d+1}},$$

where $f(t)$ is a polynomial in z of degree at most d.

Readers with a background in commutative algebra might find a strong resemblance between this result and the Hilbert series of a graded algebra – a connection which we will now make precise.

Definition 1.1.8. *For K a field, a K-algebra R is called* ***graded*** *with respect to $\mathbb{Z}^n$ if it can be written as a direct sum*

$$R = \bigoplus_{\alpha\in\mathbb{Z}^n} R_\alpha,$$

where for $x \in R_\alpha$ and $y \in R_\beta$, we have that $x \cdot y \in R_{\alpha+\beta}$.

Definition 1.1.9. *The* ***Hilbert function*** *$h(\alpha)$ of a $\mathbb{Z}^n$-graded K-algebra R is the function whose input is an element α of $\mathbb{Z}^n$ and whose output is the K vector space dimension of R_α. Its ordinary generating function*

$$H_R(\boldsymbol{t}) := \sum_{\alpha\in\mathbb{Z}^n} h(\alpha)\boldsymbol{t}^\alpha,$$

where $\boldsymbol{t}^\alpha = t_1^{\alpha_1}\cdots t_n^{\alpha_n}$, is known as the Hilbert series of R.

Definition 1.1.10. *The* ***semigroup algebra*** *$K[\Lambda]$ associated to a semi-group $\Lambda \subset \mathbb{Z}^{d+1}$ is the K vector space with basis $\{e_\alpha\}_{\alpha\in\Lambda}$ equipped with the product*

$$e_\alpha \cdot e_\beta = e_{\alpha+\beta}.$$

It is immediate that $K[\Lambda]$ is a $\mathbb{Z}^{d+1}$–graded K–algebra, and that the Hilbert function is given by

$$h(\alpha) = \begin{cases} 1 \text{ if } \alpha \in \Lambda, \text{ and} \\ 0 \text{ otherwise.} \end{cases}$$

It is common to "coarsen" the grading of $K[\Lambda]$ by considering it to be $\mathbb{Z}$-graded algebra with grading given by the zeroth coordinate of its $\mathbb{Z}^{d+1}$–grading.

The seemingly arbitrary definition of the cone over a simplex Δ is shown to be natural and helpful by the following observation.

Definition 1.1.11. *For a point $x = (x_0, x_1, \ldots, x_d)$ in $\mathbb{R}^{d+1}$, we define the **height** of x to be*

$$\text{height}(x) = x_0.$$

Letting X_n denote the collection of points $x \in \mathbb{R}^{d+1}$ with height equal to n, we have the set equality

$$X_n \cap \text{cone}(\Delta) \;=\; \{(n, n \cdot x) \in \mathbb{R}^{d+1} \text{ such that } x \in \Delta\}.$$

Observe that the set $\mathbb{Z}^{d+1} \cap X_n \cap \text{cone}(\Delta)$ is in bijection with the set of lattice points of $n\Delta$ (by dropping the zeroth coordinate). Consequently, for Λ the semigroup associated to Δ, the Hilbert function $h(n)$ of the $\mathbb{Z}$-graded algebra $K[\Lambda]$ is precisely $\text{ehr}(n)$ of the simplex Δ! From this we conclude that $h(n)$ is in fact a polynomial, and that its generating function, the Hilbert series, is of the form

$$H_{K[\Lambda]}(t) \;:=\; \sum_{n \in \mathbb{Z}} h(n) t^n \;=\; \frac{f(t)}{(1-t)^{d+1}}$$

for some polynomial $f(t)$ of degree at most d.

Note that this is the Hilbert series of the $\mathbb{Z}$-graded ring Λ. In general, we consider the $\mathbb{Z}^{d+1}$-graded case, for which the denominator has a similar but slightly more complicated form. In particular the Hilbert series of the $\mathbb{Z}^{d+1}$-graded ring $K[\Lambda]$ has $d+1$ variables.

This connection motivates the study of simplices Δ through algebraic properties of the associated object $K[\Lambda]$. In this spirit we introduce another algebraic invariant of $K[\Lambda]$, its Poincaré series.

1.2 The Poincaré series

Definition 1.2.1. *Given a collection of vector spaces $\{F_i\}_{i \in \mathbb{Z}_{\geq 0}}$, together with linear maps ∂_i from F_i to F_{i-1}, we call the sequence*

$$F : \quad F_0 \xleftarrow{\partial_1} F_1 \xleftarrow{\partial_2} \cdots \xleftarrow{\partial_i} F_i \xleftarrow{\partial_{i+i}} F_{i+1} \xleftarrow{\partial_{i+2}} \cdots$$

*a **complex** of vector spaces if the image of ∂_{i+1} is contained in the kernel of ∂_i for all $i \geq 1$.*

Definition 1.2.2. *The i'th homology of the complex F is the quotient vector space*

$$H_i(F) := \ker \partial_i / \operatorname{im} \partial_{i+1}.$$

It has vector space dimension

$$\begin{aligned} \dim_K H_i(F) &= \dim_K \ker \partial_i - \dim_K \operatorname{im} \partial_{i+1} && (1.1) \\ &= \dim_K F_i - (\dim_K \operatorname{im} \partial_i + \dim_K \operatorname{im} \partial_{i+1}). && (1.2) \end{aligned}$$

Our motivation for consider complexes of vector spaces is that they arise naturally when studying free resolutions of K-algebras, which were originally defined in order to study Hilbert series.

Definition 1.2.3. *Let M be a finitely generated graded module over R, F_i be a free R-module and ∂_i be a graded R-module homomorphism such that the image of ∂_{i+1} is equal to the kernel of ∂_i for all $i \geq 1$. Then the complex F is a **free resolution** of M over R if $M \cong F_0 / \operatorname{im} \partial_1$.*

Because it is graded, we may split the free resolution F into a direct sum of K vector space complexes by writing each F_i as a direct sum $\bigoplus_{\alpha \in \mathbb{Z}^n} F_{i,\alpha}$.

Recall that the tensor product $M \otimes N$ of two R-modules M and N may be written as

$$M \otimes N = \left\{ \sum_i x_i \otimes y_i \text{ such that } x_i \in M,\, y_i \in N \right\},$$

and satisfies $(rx) \otimes y = x \otimes (ry)$ for all $r \in R$, $x \in M$, and $y \in N$, and that for (F, ∂) a complex of free R-modules, we can define a tensor complex $(M \otimes F, \operatorname{Id} \otimes \partial)$.

Definition 1.2.4. *The Betti number $\beta^R_{i,\alpha}(M)$ of a graded R-module M is the vector space dimension of the i'th homology of the graded component of $M \otimes F$ of degree α.*

Definition 1.2.5. *The Poincaré series $P^M_R(z; \boldsymbol{t})$ is the ordinary generating function for the Betti numbers of the R-module M, i.e.,*

$$P^M_R(z; \boldsymbol{t}) = \sum_{\alpha \in \mathbb{Z}^n} \sum_{i \geq 0} \beta^R_{i,\alpha}(M) z^i \boldsymbol{t}^\alpha.$$

In the case that R is a polynomial ring in n variables, the Hilbert Syzygy Theorem says that the Poincaré series $P^M_R(z; \boldsymbol{t})$ is a polynomial of z-degree at most n for any finitely generated R-module M. However, when R is not a polynomial ring, the growth of the Betti numbers is not so simple — the Poincaré series may not even be rational!

A question of rationality

We call a $\mathbb{Z}^n$-graded algebra R **connected** if $R_0 \cong K$ (as in the case of a semi-group ring $K[\Lambda]$ associated to a lattice simplex Δ). By a slight abuse of notation, we write

$$\mathfrak{m} := \bigoplus_{\alpha \in \Lambda \setminus 0} R_\alpha$$

and

$$K \cong R/\mathfrak{m}$$

as R-modules.

It has been shown [16] that if the Poincaré series for the ground field K as an R-module is rational for all R, then the Poincaré series is rational for any finitely generated module. Hence the question of Serre-Kaplansky:

Question 1.2.1. *Is the Poincaré series of the ground field K over R rational for all K-algebras R?*

This question was answered in the negative by Anick [1], and much subsequent work has focused on determining the properties of R that lead to rationality or irrationality.

In this work we will consider the case when R is a quotient of a polynomial ring by an ideal I, and when we reference the Poincaré series, we always consider the module to be the ground field K.

When I is generated by monomials, as in the case of Stanley-Reisner theory, the Poincaré series is known to be rational. Berglund, Blasiak, and Hersh [3] describe a combinatorial method for computing the rational form. Less is known about quotients by another important class of ideals in combinatorics, toric ideals [19]. An example of a toric ring with transcendental Poincaré series was found by Roos and Sturmfels [21], and it is known by work of Gasharov, Peeva, and Welker [12] that quotients arising from generic toric ideals have rational Poincaré series. Another relevant line of investigation is the rationality of Poincaré series for certain Gorenstein rings. Elias and Valla proved [11] that the Poincaré series of an almost stretched Gorenstein local ring of dimension d and embedding codimension h is given by

$$\frac{(1+z)^d}{1-hz+z^2}\,. \tag{1.3}$$

It is not known whether rationality or irrationality of the Poincaré series is the more "common" property for toric rings.

When R is equal to $K[\mathbb{Z}^{d+1} \cap \operatorname{cone}(\Delta)]$ and is Koszul [17, Cor. 4.3], then we have the following equality:

$$P_R^K(-1,t)H_R(t) \;=\; 1. \tag{1.4}$$

The original result is, in fact, stronger, and implies (via the rationality of $H_R(t)$), that the Poincaré series itself is rational.

Combinatorialists have historically approached the question of rationality from the observation [18] that the Betti numbers record simplicial homology of open intervals in Λ, and use combinatorial tools to show that $K[\Lambda]$ is Koszul. We will focus instead on the non-Koszul case.

Rationality equals recurrence

Recall that for a sequence $(b_0, b_1, \dots)$, the generating function

$$B(z) \;=\; \sum_{i\geq 0} b_i z^i$$

has a rational form if and only if there exists a polynomial

$$g(z) \;=\; 1 + \sum_{j=1}^{d} g_i \, z^i$$

such that the product

$$B(z) \cdot g(z)$$

is equal to a polynomial $f(z)$, i.e.,

$$B(z) = \frac{f(z)}{g(z)}.$$

One consequence of this observation is that $B(z)$ is rational if and only if there exists a natural number d such that for all $D > d$, we have the equality

$$b_D + \sum_{j=1}^{d} g_j b_{D-j} = 0,$$

which can be read as the linear recurrence

$$b_D = (-g_1)b_{D-1} + \cdots + (-g_d)b_{D-d}.$$

1.3 The Fundamental Parallelepiped

We now introduce the fundamental parallelepiped, a distinguished subset of cone(Δ).

Definition 1.3.1. *For a lattice d-simplex Δ with vertices v_0 through v_d, the* ***fundamental parallelepiped*** Π_Δ *is the set*

$$\Pi_\Delta \;:=\; \left\{ \sum_{i=0}^{d+1} \gamma_i (1, v_i) \;\textit{such that}\; 0 \leq \gamma_i < 1 \right\} \subset \operatorname{cone}(\Delta).$$

Our interest in the fundamental parallelepiped Π_Δ arises mainly from the following fact: every element of the semi-group Λ may be written uniquely as a non-negative integer combination of the $(1, v_i)$'s and a lattice point in Π_Δ. To see this, note that because any element z of $\Lambda = \operatorname{cone}(\Delta) \cap \mathbb{Z}^{d+1}$ lies in cone(Δ), it is a non-negative linear combination of the $(1, v_i)$'s, i.e., there exist non-negative real coefficients g_i such that

$$z = \sum_{i=1}^{d+1} g_i (1, v_i) = \left(\sum_{i=1}^{d+1} \lfloor g_i \rfloor (1, v_i) \right) + \left(\sum_{i=1}^{d+1} \{g_i\} (1, v_i) \right)$$

where $\{g_i\}$ means the fractional part of g_i. By setting γ_i equal to $\{g_i\}$, we see that any point z may be written as a non-negative *integral* combination of the $(1, v_i)$'s and an integer point in $\Pi_\Delta \cap \mathbb{Z}^{d+1}$. In particular, the set $\Lambda = \operatorname{cone}(\Delta) \cap \mathbb{Z}^{d+1}$ has a finite minimal additive generating set.

Definition 1.3.2. *The minimal additive generating set of* Λ *is finite, and is called the* ***Hilbert basis*** $\mathcal{H}$ *of* $\operatorname{cone}(\Delta)$. *It consists of the* $(1, v_i)$*'s and elements* h_1 *through* h_m *in* Π_Δ *such that*

$$\Lambda = \left\{ \left(\sum_{i=0}^{d} r_i (1, v_i) \right) + \left(\sum_{j=1}^{m} s_i \, h_i \right) \ \textit{such that } r_i \, , \, s_j \in \mathbb{Z}_{\geq 0} \right\}.$$

The Hilbert basis consists of the cone generators $(1, v_i)$ together with the additively minimal elements h_j of $\Pi_\Delta \cap \mathbb{Z}^{d+1}$.

A presentation of $K[\Lambda]$

Because $K[\Lambda]$ is finitely generated (by its Hilbert basis $\mathcal{H}$) it has a presentation

$$0 \to \ker \varphi \to K[V_0, \ldots, V_d, x_1, \ldots, x_m] \xrightarrow{\varphi} K[\Lambda] \to 0, \tag{1.5}$$

where the map φ is defined by the image of variables: the image of V_i is the vector space basis element $e_{(1,v_i)}$ associated with the Hilbert basis element $(1, v_i)$ in Λ, and the image of x_i is e_{h_i} where the h_i are the remaining elements of the Hilbert basis. This defines a surjective degree map $\deg(\cdot)$ from the set of monomials of $K[V_1, \ldots, V_{d+1}, x_1, \ldots, x_m]$ onto Λ by

$$\deg \left(\prod V_i^{s_i} \cdot \prod x_j^{r_j} \right) = \sum s_i (1, v_i) + \sum r_j h_j \, .$$

Extending $\deg(\cdot)$ K-linearly, we see that $\ker \varphi$ is the toric ideal I generated by all binomials

$$\mathbf{V}^{u_V} \mathbf{x}^{u_x} - \mathbf{V}^{w_V} \mathbf{x}^{w_x}$$

such that $\deg (\mathbf{V}^{u_V} \mathbf{x}^{u_x}) = \deg(\mathbf{V}^{w_V} \mathbf{x}^{w_x})$.

There is a second algebra associated to Λ.

Definition 1.3.3. *The* ***Fundamental Parallelepiped Algebra*** $\operatorname{FPA}(\Delta)$ *associated with the simplex* Δ *may be constructed in two ways; firstly as the quotient*

$$K[V_0, \ldots, V_d, x_1, \ldots, x_m] \, / \, \ker \varphi + \big(V_0, \ldots, V_d \big),$$

and secondly as the algebra with K *vector space basis*

$$\big\{ e_\sigma \ \textit{such that } \sigma \in \mathbb{Z}^{d+1} \cap \Pi_\Delta \big\}$$

and with multiplication given by

$$e_\sigma \cdot e_\mu = \begin{cases} e_{\sigma+\mu} & \textit{if } \sigma + \mu \in \mathbb{Z}^{d+1} \cap \Pi_\Delta, \ \textit{and} \\ 0 & \textit{otherwise.} \end{cases}$$

Example 1.3.4. *Computation in Macaulay2 [14] gives that for the 2-simplex Δ with vertices $(1,0)$, $(0,1)$, and $(-2,-3)$,* $\operatorname{cone}(\Delta)$ *has Hilbert basis (and associated variables) given by the columns below:*

$$\begin{array}{ccccccc} V_1 & V_2 & V_3 & x_1 & x_2 & x_3 & x_4 \end{array}$$
$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & 0 & 0 & -1 & -1 \\ 1 & 0 & -3 & 0 & -1 & -1 & -2 \end{pmatrix}$$

The associated toric ideal I is given by

$$\begin{aligned} I =&(V_1x_2 - x_1^2, V_2V_3 - x_2x_4, V_2x_4 - x_2^2, V_3x_2 - x_4^2, V_1x_4 - x_1x_3, \\ &x_1x_4 - x_2x_3, V_2x_3 - x_1x_2, V_3x_1 - x_3x_4, V_1V_3 - x_3^2). \end{aligned}$$

The algebra $K[\Lambda]$ is isomorphic to $K[V_1, V_2, V_3, x_1, x_2, x_3, x_4]/I$ and

$$\operatorname{FPA}(\Delta) \cong K[x_1, x_2, x_3, x_4]/(x_1^2, x_2x_4, x_2^2, x_3^2, x_4^2, x_1x_3, x_1x_4 - x_2x_3, x_1x_2, x_3x_4)\,.$$

One inspiration for defining this algebra is the fact that, due to an argument presented earlier, every element of Λ may be written uniquely as a non-negative sum of points $(1, v_i)$ and a single point in Π_Δ. Consequently,

$$\begin{aligned} H_{K[\Lambda]}(\boldsymbol{t}) &= H_{K[x_1,\ldots,x_m]}(\boldsymbol{t}) \cdot H_{\operatorname{FPA}(\Delta)}(\boldsymbol{t}) \\ &= \frac{H_{\operatorname{FPA}(\Delta)}(\boldsymbol{t})}{\prod_{i=1}^m (1 - \boldsymbol{t}^{h_i})}. \end{aligned}$$

Because by construction Π_Δ is bounded, it contains a finite number of lattice points, and hence $H_{\operatorname{FPA}(\Delta)}(\boldsymbol{t})$ is a polynomial in t_0 through t_d.

Because the generators $(1, v_i)$ form a linear system of parameters for $K[\Lambda]$, there is an analogous result in the world of Poincaré series.

Theorem 1.3.5. *[2, Prop. 3.3.5] For the $\mathbb{Z}$-graded algebra $K[\Lambda]$, we have the following equality:*

$$\begin{aligned} P^K_{K[\Lambda]}(z; \boldsymbol{t}) &= (1 + z\boldsymbol{t})^{d+1} \cdot P^K_{\operatorname{FPA}(\Delta)}(z; \boldsymbol{t}) \\ &= P^K_{K[V_0,\ldots,V_d]}(z; \boldsymbol{t}) \cdot P^K_{\operatorname{FPA}(\Delta)}(z; \boldsymbol{t}). \end{aligned}$$

For the remainder of this work we will focus on computing the Poincaré series of the Fundamental Parallelepiped Algebra $\operatorname{FPA}(\Delta)$.

1.4 Bar Resolution

One standard resolution of K as a module over a graded K-algebra is the Bar resolution. In the definition we use the bar symbol $|$ to mean a tensor over K, and reserve the tensor symbol $\otimes$ to mean a tensor over the ring under consideration.

Definition 1.4.1. *The Bar resolution* $\mathbb{B}$ *of the module* K *over the* $\mathbb{Z}^n$*-graded* K *algebra* $\mathrm{FPA}(\Delta)$ *has graded components* $[\mathbb{B}_i]_\alpha$ *with vector space basis given by* $\lambda_0 \,|\, \cdots \,|\, \lambda_i$ *such that* λ_0 *is in* Π_Δ*, each* λ_j *is in* $\Pi_\Delta \backslash 0$ *(for* $j \geq 1$*), and* $\sum_{j=0}^{i} \lambda_j = \alpha$*. The differential map* ∂_i *acts by sending* $\lambda_0 \,|\, \cdots \,|\, \lambda_i$ *to the sum*

$$\sum_{j=0}^{i-1} (-1)^j \lambda_0 \,|\, \cdots \,|\, \lambda_{j-1} \,|\, \lambda_j + \lambda_{j+1} \,|\, \lambda_{j+2} \,|\, \cdots \,|\, \lambda_i$$

in $\mathbb{B}_{i-1}$.

Recall that in order to compute the Betti number $\beta_{i,\alpha}$ we must compute homology in the tensor complex $B := K \otimes \mathbb{B}$. Because we identify K with the vector sub-space R_0 with basis e_0, we see that $[B_i]_\alpha$ is generated as a vector space by the collection $\{e_0 \otimes \lambda_0 \,|\, \lambda_1 \,|\, \cdots \,|\, \lambda_i\}$. Observe that unless λ_0 is equal the point 0 in Λ, the product $e_0 \otimes \lambda_0$ is equal to zero, since for σ not equal to zero, $e_0 \cdot e_\sigma$ is equal to zero in the module K, and hence

$$e_0 \otimes e_\sigma \;=\; e_0 \cdot e_\sigma \otimes e_0 \;=\; 0 \otimes e_0 \;=\; 0.$$

Consequently, for $i \geq 1$, $[B_i]_\alpha$ has a vector space basis in bijection with the collection of $\lambda_1 \,|\, \cdots \,|\, \lambda_i$ such that each λ_j is in $\Pi_\Delta \backslash 0$ and $\sum_{j=1}^{i} \lambda_j = \alpha$. We further have that $[B_0]_\alpha$ is the trivial vector space unless α is zero in Λ, and that $[B_0]_0$ is isomorphic to K.

Unimodular simplices

If the matrix whose columns are given by $(1, v_i)$ has determinant $\pm\boldsymbol{v}$, we say that the simplex Δ has **normalized volume** $\boldsymbol{v}$. Since $\boldsymbol{v}$ is precisely the index of the sub-lattice generated by $(1, v_0)$ through $(1, v_d)$ in the lattice generated by the Hilbert basis, we see that the normalized volume is equal to the number of lattice points in Π_Δ. If the normalized volume of Δ is equal to one, then we call Δ a **unimodular** simplex. In this case, it is clear that that $\mathrm{FPA}(\Delta)$ is one-dimensional as a K vector space, and has basis e_0. Consequently, $[B_i]_\alpha$ has empty basis (and dimension zero) unless α is equal to zero in Λ and $i = 0$. It follows that the complex B is given by

$$0 \leftarrow K \leftarrow 0 \leftarrow 0 \leftarrow \cdots$$

and that

$$\beta_{i,\alpha} = \begin{cases} 1 & \text{if } i = 0 \text{ and } \alpha = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Thus for unimodular simplex Δ, $P^K_{\mathrm{FPA}(\Delta)}(z; \boldsymbol{t}) = 1$. The result is consistent with the fact that $K[\Lambda]$ is a polynomial ring in the case that Δ is unimodular.

Partially ordered sets

We now recall some basic definitions from poset theory. A partially ordered set $(P, \prec)$ satisfies three axioms for all x, y, and z in P:

$$x \prec x \quad \text{(reflexivity)},$$

$$x \prec y \text{ and } y \prec x \text{ imply that } x = y \quad \text{(antisymmetry)},$$

$$x \prec y \text{ and } y \prec z \text{ imply that } x \prec z \quad \text{(reflexivity)}.$$

An ordered list of elements $x_0, \ldots x_\ell$ satisfying $x_i \prec x_{i+1}$ is called a chain. The length of the chain is the integer ℓ. For a collection of elements such that no pair are related in P, we call the collection an antichain.

If $x \prec y$, $x \neq y$, and there exists no z such that x, z, y form a chain, then we say that y covers x. The graph whose vertices are the elements of P and with edges $\{x, y\}$ such that y covers x is called the Hasse diagram of P. If the Hasse diagram of P has no cycles and is connected, then we call P a tree.

The set of lattice points $\mathbb{Z}^{d+1} \cap \Pi_\Delta$ has a very natural partial order structure.

Definition 1.4.2. *The set $\mathbb{Z}^{d+1} \cap \Pi_\Delta$ is partially ordered by letting $\sigma \prec \mu$ if and only if $\mu - \sigma$ is an element of $\mathbb{Z}^{d+1} \cap \Pi_\Delta$. We call this poset the **fundamental parallelepiped poset** $\mathrm{P}(\Delta)$.*

Observe that the zero element of Λ is below every other element of $\mathrm{P}(\Delta)$, and that the minimal elements of $\mathrm{P}(\Delta)\backslash 0$ are precisely the elements $h_1, \ldots, h_m$ of the Hilbert basis of cone(Δ). It is also helpful to notice that each edge $\sigma \prec \mu$ of the Hasse diagram of $\mathrm{P}(\Delta)$ may be labeled with the height of $\mu - \sigma$ (neccesarily the height of a Hilbert basis element), and that the height of any element is equal to the sum of the heights of the edge labels in any maximal chain from 0 to that element. Lastly we report the fact (not difficult to derive from the definitions) that d is an upper bound for the height of any element of $\mathrm{P}(\Delta)$, and $d-1$ is an upper bound for the height of any minimal element.

Antichain simplices

As a small demonstration of leveraging knowledge about the partial order $\mathrm{P}(\Delta)$ in order to compute the Poincaré series, consider a simplex Δ such that $P(\Delta)\backslash 0$ is an antichain. We will call such a simplex **antichain**.

In the case of an antichain simplex, the differential map is uniformly zero, since $e_{\lambda_j} \cdot e_{\lambda_{j+1}}$ equals zero for all j. By Equation (1.2) of Definition 1.2.2, $\beta_{i,\alpha}$ is equal to the dimension of $[B_i]_\alpha$. By considering the recurrence (for large i and α)

$$\dim_K [B_i]_\alpha = \sum_{\substack{\sigma \in \mathrm{P}(\Delta) \\ \sigma \neq 0}} \dim_K [B_i]_{\alpha - \sigma},$$

we see that

Theorem 1.4.3. *For an antichain simplex Δ, the Poincaré series is given by*

$$P^K_{\mathrm{FPA}(\Delta)}(z;\boldsymbol{t}) = \left(1 - \sum_{\substack{\sigma\in \mathrm{P}(\Delta)\\ \sigma\neq 0}} z\boldsymbol{t}^\sigma\right)^{-1}.$$

Note that Theorem 1.4.3 does not require that $K[\Lambda]$ be Koszul, and establishes the rationality of the Poincaré series. In fact, if the minimal elements of $P(\Delta)$ do not have (geometric) height equal to one, $K[\Lambda]$ *is not* Koszul! The simplest examples of antichain simplices are the 1-simplices (intervals $[0,n] \subset \mathbb{R}^1$), for the following reason: elements of $\mathrm{P}(\Delta)$ have height bound by $d = 1$, so that all non-zero elements have height equal to 1. It follows that any cover relation $x \prec y$ must be labeled 0. From this we conclude that $y - x$ has height zero and hence *is equal to* zero, the unique element of $\mathrm{P}(\Delta)$ of height zero. This implies that $x = y$, which contradicts y covering x, and so we conclude that there is no such covering relation and that $\mathrm{P}(\Delta)\backslash 0$ is an antichain.

In the case that Δ is a 2 or 3-simplex, it is sufficient that Δ be an empty simplex. An empty simplex is one whose only lattice points are its vertices. Consequently, no elements of $\mathrm{P}(\Delta)$ have height equal to one, and similarly no cover relation has label one. It follows that for any chain $0 \prec \sigma \prec \mu$, the height of μ is at least four. Thus for d equal to two or three and Δ empty, $\mathrm{P}(\Delta)$ is an antichain.

We will further explore antichain simplices experimentally in chapter 3.

In the next chapter we will use a different resolution in order to compute the Poincaré series: the minimal free resolution.

Definition 1.4.4. *An R-resolution (F,∂) is* ***minimal*** *if, for each $i \geq 0$, the image of ∂_{i+1} restricted to each summand of $F_i \cong \bigoplus_n R$ is contained in $\mathfrak{m}$.*

Because the tensor product distributes across direct sums, the tensor complex $K \otimes F_i$ is isomorphic to a direct sum of copies of K. Hence if (F,∂) is minimal, then $\mathrm{Id} \otimes \partial_i$ is the zero map for all i. It follows that the dimension of the i'th homology in degree α, also known as the Betti number $\beta_{i,\alpha}$, is equal to $\dim_K[F_i]\alpha$.

Chapter 2 The Poincaré series of a family of lattice simplices

For a family of lattice simplices described in Definition 2.1.1 of this chapter and denoted by Δ_2^m, we prove in Theorem 2.3.1 that the Poincaré series for their associated algebra $\mathrm{FPA}(\Delta_2^m)$ using a fine grading is rational, with structure similar to Equation (1.3) of Chapter 1. However, the recurrence given in (1.3) is realized only after specializing our fine grading to a coarse grading and then algebraically canceling. Our method of proof is to produce an explicit resolution of K over $\mathrm{FPA}(\Delta_2^m)$ such that the resolution reflects the recursive structure encoded in the denominator of the finely-graded Poincaré series. We show that $\mathrm{FPA}(\Delta_2^m)$ is not Koszul, and therefore rationality does not follow from Equation (1.4). We believe that the results in this chapter will be of interest to both geometric combinatorialists and commutative algebraists, for the following reasons.

- There has been fruitful investigation of the Hilbert series of $K[\Lambda]$, i.e. the Ehrhart series of Δ, in relation to the geometry and arithmetics of Δ. We believe that a similar investigation should be conducted for Poincaré series. Our work is a contribution in this direction.
- For an arbitrary lattice simplex Δ, the arithmetic properties of the fundamental parallelepiped of P should significantly impact the behavior of the Poincaré series for $\mathrm{FPA}(\Delta)$. This influence should be more subtle than the interpretation of the numerator of the Hilbert series of $K[\Lambda]$. Our results show how this works in a special case.
- Our results demonstrate how interactions between multivariate and univariate rational generating functions that are "typical" in combinatorics can create subtle complications when attempting to use rational Poincaré series to inform the construction of minimal resolutions, adding complexity to Avramov's proposed application of using rational Poincaré series to construct minimal resolutions.

The remainder of this chapter is structured as follows. In Section 2.1 we describe a family of lattice simplices and their associated semigroup algebras. In Section 2.2 we present a tree whose weighted rank generating function is equal to the Poincaré series of the Fundamental Parallelepiped Algebra, and whose structure is related to the rationality of that formal power series. In Section 2.3 we state and prove our main result, Theorem 2.3.1, which gives a rational expression for the fine graded Poincaré series of the Fundamental Parallelepiped Algebra of an infinite family of lattice simplices.

2.1 The Fundamental Parallelepiped Algebra of Δ_2^m

The following family of simplices are the main objects under investigation in this chapter.

Definition 2.1.1. *Let* Δ_2^m *be the* $(m+1)$*–simplex whose vertices are the standard basis vectors in* $\mathbb{R}^{m+1}$ *together with the point* $(-2,\ldots,-2,-2m-1) \in \mathbb{R}^{m+1}$.

The simplices Δ_2^m form a subfamily of lattice simplices recently studied by Braun, Davis, and Solus [6, 4] in the context of reflexive simplices having the integer decomposition property and also having a unimodal Ehrhart h^*-polynomial.

Theorem 2.1.2. *Let* $\mathrm{FPA}(\Delta_2^m)$ *denote the fundamental parallelepiped algebra for* Δ_2^m. *The following isomorphism holds for all* $m \geq 1$*:*

$$\mathrm{FPA}(\Delta_2^m) \cong K[x_1,x_2,x_3,x_4]/(x_1^2,x_2^2,x_3^2,x_4^{m+1},x_1x_2,x_1x_3,x_2x_4^m,x_3x_4^m,x_2x_3-x_1x_4)$$

Further, the quotient algebra has a K*-vector space basis given by the equivalence classes represented by the elements of*

$$\{1,x_1,x_4^{\ell+1},x_1x_4^{\ell+1},x_2x_4^{\ell},x_3x_4^{\ell}\}_{0\leq \ell \leq m-1}\,.$$

Proof. We first describe the fundamental parallelepiped Π for Δ_2^m and identify additive relations among the generators of the lattice points in it. As shown in [4], lattice points in Π are parameterized by integers b in $[0,4m+1]$, with each b corresponding to the lattice point

$$z_b := \begin{bmatrix} b - m\lfloor \frac{b}{2m+1} \rfloor - \lfloor b/2 \rfloor \\ -\lfloor \frac{b}{2m+1} \rfloor \\ \vdots \\ -\lfloor \frac{b}{2m+1} \rfloor \\ -\lfloor b/2 \rfloor \\ \vdots \\ -\lfloor b/2 \rfloor \end{bmatrix} \in \Pi\,.$$

Considering the cases $b < 2m+1$ and $b \geq 2m+1$, and then considering the parity of b, we see that for each choice $1 \leq h \leq m$ of zeroth coordinate, we get exactly four solutions (presented as column vectors below):

$$\begin{array}{cccc} z_{2h-1} & z_{2h} & z_{2(m+h)-1} & z_{2(m+h)} \end{array}$$
$$\begin{pmatrix} h & h & h & h \\ 0 & 0 & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & -1 & -1 \\ -h+1 & -h & -(m+h)+1 & -(m+h) \\ \vdots & \vdots & \vdots & \vdots \\ -h+1 & -h & -(m+h)+1 & -(m+h) \end{pmatrix}.$$

By Theorem 4.1 of [4], the simplex Δ_2^m has the integer decomposition property, implying that $\mathrm{FPA}(\Delta_2^m)$ is generated by elements with zeroth coordinate equal to 1, i.e., e_{z_1}, e_{z_2}, $e_{z_{2m+1}}$, and $e_{z_{2m+2}}$. Thus we may assume without loss of generality that additive identities in the fundamental parallelepiped have the form $z_b + z_{b'} = z_c + z_{c'}$, where z_b and z_c have zeroth coordinate equal to 1 and the zeroth coordinate of $z_{b'}$ and $z_{c'}$ is h. It follows by inspection that every such identity is of the form $z_1 + z_{2(m+h)} = z_2 + z_{2(m+h)-1}$ for some h between 2 and m. Every such identity may be written as

$$(h-1)z_2 + (z_1 + z_{2m+2}) = (h-1)z_2 + (z_2 + z_{2m+1})\,,$$

and is therefore a consequence of the primitive additive identity

$$z_1 + z_{2m+2} = z_2 + z_{2m+1}\,.$$

Now that we have a better understanding of the structure of $\mathrm{FPA}(\Delta_2^m)$, we get close to our desired isomorphism by constructing the map

$$\psi : K[x_1, x_2, x_3, x_4]/(x_1x_4 - x_2x_3) \to \mathrm{FPA}(\Delta_2^m) = K[\Lambda]/\left(e_{(1,v_i)}\right)$$

defined by algebraically extending the map on variables given by

$$x_1 \mapsto e_{z_{2m+1}}, \quad x_2 \mapsto e_{z_{2m+2}}, \quad x_3 \mapsto e_{z_1}, \quad x_4 \mapsto e_{z_2}\,.$$

To verify that ψ is well-defined, consider a pair of monomials $\prod_i x_i^{s_i}$ and $\prod_j x_j^{t_j}$ that are in the same equivalence class. Then $(\prod_i x_i^{s_i}) - (\prod_j x_j^{t_j})$ is in the ideal $(x_1x_4 - x_2x_3)$, so that $(\prod_i x_i^{s_i}) - (\prod_j x_j^{t_j}) = t(x_1x_4 - x_2x_3)$ for some t. It follows that $\psi(\prod_i x_i^{s_i}) - \psi(\prod_j x_j^{t_j}) = \psi(t)(e_{z_{2m+1}+z_2} - e_{z_{2m+2}+z_1})$ is zero, since, as we have seen, $z_1 + z_{2m+2} = z_2 + z_{2m+1}$. It is straightforward to verify that the homomorphism ψ is surjective.

We next determine the kernel of ψ. Observe that since $2z_1$ is not among z_3, z_4, z_{2m+3}, and z_{2m+4}, we can conclude that $2z_1$ is not in Π. We can similarly conclude that $2z_{2m+1}$, $2z_{2m+2}$, $z_1 + z_{2m+1}$, and $z_{2m+1} + z_{2m+2}$ are not in Π. We additionally see that $z_{4m+1} = mz_2 + z_{2m+1}$, so that $mz_2 + z_1$, $mz_2 + z_{2m+2}$, and $(m+1)z_2$ are not in Π, since Π contains a unique element with zeroth coordinate equal to $m+1$. Since $z_1 + z_1$ is an element of Λ but not Π, we conclude that $z_1 + z_1 = v_i + z$ for some z in Λ. Thus $e_{z_1+z_1} = e_{z_1}^2 = 0$ in $\mathrm{FPA}(\Delta_2^m)$. Similarly

$$e_{z_1}^2 = e_{z_1}e_{z_{2m+1}} = e_{z_{2m+1}}^2 = e_{z_{2m+1}}e_{z_{2m+2}} = e_{z_{2m+2}}^2 = e_{z_1}e_{z_2}^m = e_{z_{2m+2}}e_{z_2}^m = e_{z_2}^{m+1} = 0$$

in $\mathrm{FPA}(\Delta_2^m)$. Thus the kernel of ψ contains $(x_3^2, x_1x_3, x_1^2, x_1x_2, x_2^2, x_3x_4^m, x_2x_4^m, x_4^{m+1})$.

Finally, we count equivalence classes of monomials in the ring

$$K[x_1, x_2, x_3, x_4]/(x_1^2, x_2^2, x_3^2, x_4^{m+1}, x_1x_2, x_1x_3, x_2x_4^m, x_3x_4^m, x_2x_3 - x_1x_4)\,.$$

We only need to consider the monomials 1 and variables multiplied by powers of x_4, since x_ix_j is either zero or equal to x_4r_k for some k. It follows that it is a $(4m+2)$–dimensional K-vector space with basis

$$\{1, x_1, x_4^{\ell+1}, x_1x_4^{\ell+1}, x_2x_4^{\ell}, x_3x_4^{\ell}\}_{0\le\ell\le m-1},$$

and with a surjective ring homomorphism $\hat{\psi}$ to the $(4m+2)$–dimensional vector space $\mathrm{FPA}(\Delta_2^m)$, i.e., $\hat{\psi}$ is a ring isomorphism from

$$K[x_1, x_2, x_3, x_4]/(x_1^2, x_2^2, x_3^2, x_4^{m+1}, x_1x_2, x_1x_3, x_2x_4^m, x_3x_4^m, x_2x_3 - x_1x_4)$$

to the Fundamental Parallelepiped Algebra $\mathrm{FPA}(\Delta_2^m)$. □

As Koszul algebras must be defined by quadratically-generated ideals, the following corollary is immediate.

Corollary 2.1.3. *For $m \geq 2$, $\mathrm{FPA}(\Delta_2^m)$ is not Koszul.*

2.2 A Weighted Tree Encoding Betti Numbers

Our goal in this section is to define for Δ_2^m a Λ–weighted tree T whose weighted rank generating function

$$T(z; \boldsymbol{y}) := \sum_{\epsilon \in T} z^{\mathrm{rank_K}(\epsilon)} \boldsymbol{y}^{\deg(\epsilon)}$$

is equal to the $(\mathbb{N} \times \mathbb{Z}^{d+1})$–graded Poincaré series

$$P^K_{\mathrm{FPA}(\Delta_2^m)}(z, \mathbf{y}) := \sum_{\alpha \in \Lambda} \sum_{i \geq 0} \dim_K \beta_{i,\alpha} z^i \boldsymbol{y}^\alpha ,$$

where $\boldsymbol{y}^\alpha$ means the multinomial $y_0^{\alpha_0} \cdots y_n^{\alpha_n}$. To construct our weighted tree, we require the following general construction.

Definition 2.2.1. *Let Δ be a lattice simplex with fundamental parallelepiped algebra $R = \mathrm{FPA}(\Delta)$ described as a quotient of a polynomial ring with a monomial term order $\prec_R$. Assume that we have a* distinguished monomial basis *for R consisting of all monomials outside the $\prec_R$-leading term ideal for the defining ideal of R. Let d be a map between free finitely generated Λ-graded R-modules M and N, where there is an ordering $\prec$ on the generators of N. Consider a generator ϵ of M, and let δ be the $\prec$-minimal support of $d(\epsilon)$ and s the $\prec_R$-maximal monomial of $d(\epsilon)$ supported on δ. If δs is distinct for each ϵ, then we say that M* can be ordered with respect to *d. If M can be ordered with respect to d, we define an ordering of the generators of M as follows: $\epsilon \prec \epsilon'$ if $\delta \prec \delta'$ or if $\delta = \delta'$ and $s' \prec_R s$. In this case, we define the* leading term *map $\mathrm{LT}(\cdot)$ on the graded components of M which projects each element onto the summand generated by its $\prec$-minimal support. For notational convenience, we define the* leading coefficient $\mathrm{LC}(\cdot)$ *of an element to be the $\prec_R$-maximal monomial of its leading term.*

For a given complex (F, d) we denote by $F_{\leq n}$ the truncated complex

$$F_{\leq n} \ : \quad F_0 \xleftarrow{d_1} F_1 \xleftarrow{d_2} \cdots \xleftarrow{d_n} F_n \, .$$

Observe that for a Λ-graded complex F of free finitely generated R-modules, if $\mathrm{LT}(\cdot)$ is defined for the truncated complex $F_{\leq n}$ and the leading terms of $d_{n+1}(\epsilon)$ for generators ϵ of F_{n+1} are all distinct, then F_{n+1} can be ordered with respect to d_{n+1}. In this case, we may define $\mathrm{LT}(\cdot)$ on F_{n+1}.

Definition 2.2.2. *For R corresponding to Δ_2^m as given in Theorem 2.1.2, specify the ordering $\prec_R$ of the monomial K-basis by using the lexicographic order induced by the ordering*

$$1 \prec_R x_1 \prec_R x_2 \prec_R x_3 \prec_R x_4 \, ,$$

i.e. on our basis elements we have $x_i x_4^j \prec_R x_k x_4^\ell$ if $j < \ell$ or $j = \ell$ and $x_i \prec_R x_k$.

Example 2.2.3. *Let our simplex be Δ_2^m with R as given in Theorem 2.1.2. Consider the complex $F_{\leq 2}$ below:*

$$F_{\leq 2} \ : \quad R \xleftarrow[d_1]{} R^4 \xleftarrow[d_2]{} R^{15}$$

where the map $d_1 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix}$ sends each $\delta_i \to x_i$ and d_2 is given by the matrix

$$\begin{array}{c} \\ \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{array} \begin{array}{c} \begin{array}{ccccccccccccccc} \epsilon_1 & \epsilon_2 & \epsilon_3 & \epsilon_4 & \epsilon_5 & \epsilon_6 & \epsilon_7 & \epsilon_8 & \epsilon_9 & \epsilon_{10} & \epsilon_{11} & \epsilon_{12} & \epsilon_{13} & \epsilon_{14} & \epsilon_{15} \end{array} \\ \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x_1 & x_2 & x_3 & x_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_1 & x_2 & x_3 & x_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -x_1 & 0 & 0 & -x_1 & -x_2 & 0 & -x_1 & 0 & -x_3 & x_2x_4^{m-1} & x_3x_4^{m-1} & x_4^m \end{pmatrix} \end{array}.$$

We see that F_1 can be ordered with respect to d_1, with the result that $\delta_1 \prec \delta_2 \prec \delta_3 \prec \delta_4$. Further, it is straightforward to verify that F_2 can be ordered with respect to d_2, and hence the leading term of the element $d_2(\epsilon_4)$ of F_1 is $\mathrm{LT}(d_2(\epsilon_4)) = x_4\delta_1$ and the leading coefficient is $\mathrm{LC}(d_2(\epsilon_4)) = x_4$.

Construction 2.2.4. *As in Definition 2.2.1, assume Δ is a lattice simplex with fundamental parallelepiped algebra $R = \mathrm{FPA}(\Delta_2^m)$ described as a quotient of a polynomial ring with a monomial term order $\prec_R$, together with a distinguished monomial basis. Assume that F is a resolution of a module M over R such that F_n can be ordered with respect to d_n and the order on F_n is defined in this manner, with associated maps* LT *and* LC. *Construct a Λ-weighted tree T whose elements are the generators of the summands of F, and whose cover relations are given by $\epsilon \gtrdot \delta$ if $\mathrm{LT}(d(\epsilon)) = s\delta$. This also defines a labeling η of the cover relations of T where $\eta(\epsilon, \delta) := \mathrm{LC}(d(\epsilon)) = s \in R$ (by construction a monomial).*

Note that if F_0 is cyclic, then T is ranked, with the rank of an element equal to the graph distance between an element and the root of the tree in the Hasse diagram. For each element ϵ in T, there is a unique path $\hat{0} = t_0 \lessdot t_1 \lessdot \cdots \lessdot t_\ell = \epsilon$, where $\hat{0}$ is the generator of F_0. We define the degree of ϵ in T to be

$$\deg(\epsilon) = \sum_{i=0}^{\ell-1} \deg(\eta(t_i, t_{i+1})) \in \Lambda \,. \tag{2.1}$$

This definition agrees with the internal degree of the summand generated by ϵ, and the length ℓ of the chain from $\hat{0}$ to ϵ is precisely the homological degree where the summand sits. Thus there is a degree preserving bijection between summands of the complex F and elements of T, so that $T(z;\boldsymbol{t}) = \sum_{\epsilon \in T} z^{\mathrm{rank}_{\mathrm{K}}(\epsilon)} \boldsymbol{y}^{\deg(\epsilon)}$ is equal to

$$F(z;\boldsymbol{t}) = \sum_{k \geq 0} \sum_{\alpha \in \Lambda} \mathrm{rank}_{\mathrm{K}} \left[F_k\right]_\alpha z^k \boldsymbol{y}^\alpha \,.$$

Example 2.2.5. *For the complex $F_{\leq 2}$ of Example 2.2.3, the tree $T_{\leq 2}$ is depicted in Figure 2.1. Example 2.2.3 implies that the cover label $\eta(\delta_1, \gamma)$ is equal to x_1*

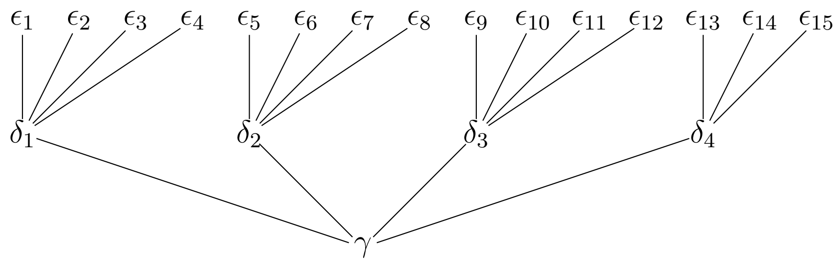


Figure 2.1: The tree $T_{\leq 2}$ for Example 2.2.3

and the cover label $\eta(\epsilon_2, \delta_1)$ *is equal to* x_2*, thus* $\deg(\epsilon_2)$ *is equal to* $\deg(\eta(\delta_1, \gamma)) + \deg(\eta(\epsilon_2, \delta_1)) = \deg(x_1) + \deg(x_2)$*. After making a similar argument for each basis element in* $F_{\leq 2}$*, it follows that the generating function* $T_{\leq 2}(z; \boldsymbol{t})$ *is given by*

$$T_{\leq 2}(z; \boldsymbol{y}) = 1 + z(\boldsymbol{y}^{\deg(x_1)} + \boldsymbol{y}^{\deg(x_2)} + \boldsymbol{y}^{\deg(x_3)} + \boldsymbol{y}^{\deg(x_4)}) + z^2(\boldsymbol{y}^{\deg(x_1)+\deg(x_1)} + \cdots + \boldsymbol{y}^{\deg(x_4)+\deg(x_4)})\,.$$

The following lemma gives a sufficient condition for $F(z; \boldsymbol{y})$ to be a rational function.

Lemma 2.2.6. *Assume the setting of Construction 2.2.4 and let* $\{\lambda_i\}_{i \in [n]}$ *denote the subset of elements of the distinguished monomial basis of* R *that appear as labels in* T*. Let the associated* η*-labeled tree* T *have the property that the multiset* $\{\eta(\epsilon, \delta) : \epsilon \gtrdot \delta\}$ *depends only on* $\mathrm{LC}(d(\delta))$*, i.e. for* δ *with* $\mathrm{LC}(d(\delta)) = \lambda_j$*, there exists exactly* $a_{i,j}$ *elements* ϵ *in* T *with* $\eta(\epsilon, \delta) = \lambda_i$ *(note that by hypothesis* $a_{i,j}$ *is either zero or one). Let* A *be the* $n \times n$ *matrix with entries* $A_{i,j} = a_{i,j} z \boldsymbol{y}^{\deg(\lambda_i)}$*. Then the generating function*

$$F(z; \boldsymbol{y}) = \sum_{k \geq 0} \sum_{\alpha \in \Lambda} \mathrm{rank}_{\mathrm{K}} \left[F_k\right]_\alpha z^k \boldsymbol{y}^\alpha$$

has a rational representation of the form

$$F(z; \boldsymbol{y}) = \frac{f(z; \boldsymbol{y})}{\chi(z; \boldsymbol{y}, 1)}\,,$$

where $\chi(z; \boldsymbol{y}, t)$ *is the characteristic polynomial of the matrix* A *and the* z*-degree of* $f(z; \boldsymbol{y})$ *is at most that of* $\chi(z; \boldsymbol{y}, 1)$*.*

Proof. Let $b^i_{k,\alpha}$ be the number of rank k elements ϵ of T having degree α and with $\mathrm{LC}(d(\epsilon)) = \lambda_i$, so that

$$\sum_{i=1}^{n} b^i_{k,\alpha} = \mathrm{rank}_{\mathrm{K}}[F_k]_\alpha.$$

Define B to be the $n \times 1$ matrix whose i-th entry is $b^i_{1,\deg(\lambda_i)} z \boldsymbol{y}^{\deg(\lambda_i)}$. Note that $b^i_{1,\deg(\lambda_i)}$ is equal to 1 if λ_i is equal to a single variable in R, and is equal to 0 otherwise.

We prove by induction the claim that the matrix A^kB is given by

$$\left(A^kB\right)_i = \sum_{\alpha\in\Lambda} b^i_{k+1,\alpha} z^{k+1}\boldsymbol{y}^\alpha.$$

The base case $k=0$ is trivial. Assume the induction hypothesis and write $\left(A^kB\right)_i$ as

$$\begin{aligned}\left(A^kB\right)_i &= \sum_{j=1}^n A_{i,j}\left(A^{k-1}B\right)_j \\ &= \sum_{j=1}^n A_{i,j}\left(\sum_{\alpha\in\Lambda} b^j_{k,\alpha} z^k\boldsymbol{y}^\alpha\right) \\ &= \sum_{\alpha\in\Lambda}\sum_{j=1}^n A_{i,j} b^j_{k,\alpha} z^k\boldsymbol{y}^\alpha \\ &= \sum_{\alpha\in\Lambda}\sum_{j=1}^n a_{i,j} b^j_{k,\alpha} z^{k+1}\boldsymbol{y}^{\alpha+\deg(\lambda_i)}.\end{aligned}$$

Observe that the coefficient of $z^{k+1}\boldsymbol{y}^\mu$ in the last line above is equal to $\sum_{j=1}^n a_{i,j} b^j_{k,\mu-\deg(\lambda_i)}$, and that, by equation (2.1), the product $a_{i,j} b^j_{k,\mu-\deg(\lambda_i)}$ is precisely the number of elements ϵ of T of degree μ and rank $k+1$ such that $\deg(\mathrm{LC}(d(\epsilon))) = \deg(\lambda_i)$. Thus

$$\left(A^kB\right)_i = \sum_{\mu\in\Lambda} b^i_{k+1,\mu} z^{k+1}\boldsymbol{y}^\mu,$$

completing the proof of the claim.

Defining $\mathbf{1}_n$ to be the $1\times n$ matrix of 1's, note that

$$\mathbf{1}_n\cdot A^\ell\cdot B = \sum_{\alpha\in\Lambda}\mathrm{rank}_{\mathrm{K}}[F_{\ell+1}]_\alpha\ z^{\ell+1}\boldsymbol{y}^\alpha.$$

Let $\chi\in K[z;\boldsymbol{y},t]$ be the characteristic polynomial of A, so that $\chi(z;\boldsymbol{y},A)=0$, and let

$$\chi = t^d + \sum_{i=0}^{d-1}\chi_i t^i,$$

with $\chi_i = \sum_j q_{i,j} z^{r_{i,j}}\boldsymbol{y}^{s_{i,j}}$, where $q_{i,j}\in K$. Then $A^d + \sum_{i=0}^{d-1}\sum_j q_{i,j} z^{r_{i,j}}\boldsymbol{y}^{s_{i,j}} A^i$ is the zero matrix. Left multiplying by $\mathbf{1}_n A^{k-1}$ and right multiplying by B yields

$$\mathbf{1}_n A^{d+k-1}B + \sum_{i=0}^{d-1}\sum_j q_{i,j} z^{r_{i,j}}\boldsymbol{y}^{s_{i,j}}\mathbf{1}_n\cdot A^{i+k-1}\cdot B = 0.$$

Thus for all $k\geq 1$,

$$\sum_{\alpha\in\Lambda}\mathrm{rank}_{\mathrm{K}}[F_{d+k}]_\alpha z^{d+k}\boldsymbol{y}^\alpha - \sum_{\mu\in\Lambda}\sum_{i=0}^{d-1}\sum_j q_{i,j}\,\mathrm{rank}_{\mathrm{K}}[F_{i+k}]_\mu z^{i+k+r_{i,j}}\boldsymbol{y}^{\mu+s_{i,j}} = 0.$$

In particular, the coefficient of $z^{d+k}\boldsymbol{y}^{\alpha}$ on the left hand side is zero, so that

$$\operatorname{rank}_{\mathrm{K}}[F_{d+k}]_{\alpha}+\sum_{i=0}^{d-1}\sum_{j}q_{i,j}\operatorname{rank}_{\mathrm{K}}[F_{d+k-r_{i,j}}]_{\alpha-s_{i,j}}=0.$$

Since this is also the coefficient of $z^{d+k}\boldsymbol{y}^{\alpha}$ in the product $\chi(z;\boldsymbol{y},1)\cdot F(z;\boldsymbol{y})$, we see that this product is a power series in $K[[z;\boldsymbol{y}]]$ which vanishes in z-degree greater than d. Since F_ℓ is a finite direct sum for each ℓ, the product is in fact a polynomial in $K[z;\boldsymbol{y}]$ and the result follows. $\square$

We have described a sufficient condition for the rank generating function of a complex to have a rational generating function. What remains is to connect the ranks of the graded components of the complex to the Betti numbers. Recall from the previous chapter that, if the complex is a minimal free resolution, then the Betti numbers are given by the ranks of the graded components of the resolution.

2.3 Rationality and Δ_2^m

In this section we prove the following theorem, our main result in this chapter.

Theorem 2.3.1. *For the simplex Δ_2^m with $R=\mathrm{FPA}(\Delta_2^m)$ as given by Theorem 2.1.2, the R–module K has a minimal free resolution F satisfying the hypotheses of Lemma 2.2.6. The matrix A resulting from Lemma 2.2.6 in this case is given by*

$$\begin{array}{c} \\ x_1\\ x_2\\ x_3\\ x_4\\ x_2x_4^{m-1}\\ x_3x_4^{m-1}\\ x_4^m\end{array}\begin{array}{c}\begin{array}{ccccccc} x_1 & x_2 & x_3 & x_4 & x_2x_4^{m-1} & x_3x_4^{m-1} & x_4^m\end{array}\\ \begin{pmatrix}
z\boldsymbol{y}^{\deg(x_1)} & z\boldsymbol{y}^{\deg(x_1)} & z\boldsymbol{y}^{\deg(x_1)} & 0 & z\boldsymbol{y}^{\deg(x_1)} & z\boldsymbol{y}^{\deg(x_1)} & 0\\
z\boldsymbol{y}^{\deg(x_2)} & z\boldsymbol{y}^{\deg(x_2)} & z\boldsymbol{y}^{\deg(x_2)} & 0 & z\boldsymbol{y}^{\deg(x_2)} & z\boldsymbol{y}^{\deg(x_2)} & z\boldsymbol{y}^{\deg(x_2)}\\
z\boldsymbol{y}^{\deg(x_3)} & z\boldsymbol{y}^{\deg(x_3)} & z\boldsymbol{y}^{\deg(x_3)} & 0 & z\boldsymbol{y}^{\deg(x_3)} & z\boldsymbol{y}^{\deg(x_3)} & z\boldsymbol{y}^{\deg(x_3)}\\
z\boldsymbol{y}^{\deg(x_4)} & z\boldsymbol{y}^{\deg(x_4)} & z\boldsymbol{y}^{\deg(x_4)} & 0 & z\boldsymbol{y}^{\deg(x_4)} & z\boldsymbol{y}^{\deg(x_4)} & z\boldsymbol{y}^{\deg(x_4)}\\
0 & 0 & 0 & z\boldsymbol{y}^{\deg(x_2x_4^{m-1})} & 0 & 0 & 0\\
0 & 0 & 0 & z\boldsymbol{y}^{\deg(x_3x_4^{m-1})} & 0 & 0 & 0\\
0 & 0 & 0 & z\boldsymbol{y}^{\deg(x_4^{m})} & 0 & 0 & 0
\end{pmatrix}\end{array} \tag{2.2}$$

and thus the $(\Lambda\times\mathbb{N})$-graded Poincaré series $P^K_{\mathrm{FPA}(\Delta_2^m)}(z;\boldsymbol{y})$ is given by

$$\frac{1+z\boldsymbol{y}^{\deg(x_4)}}{1-z\big(\boldsymbol{y}^{\deg(x_1)}+\boldsymbol{y}^{\deg(x_2)}+\boldsymbol{y}^{\deg(x_3)}\big)-z^2\big(\boldsymbol{y}^{\deg(x_2x_4^2)}+\boldsymbol{y}^{\deg(x_3x_4^2)}+\boldsymbol{y}^{\deg(x_4^3)}\big)+z^3\boldsymbol{y}^{\deg(x_1x_4^3)}}.$$

Corollary 2.3.2. *Using Equation (1.4) and the specialization $\boldsymbol{y}\mapsto(1,\ldots,1)$, the Poincaré series of the Ehrhart ring $K[\Delta]$ of the lattice simplex Δ_2^m is given by*

$$P^K_{K[\Lambda]}(z)=\frac{(1+z)^{m+2}}{1-4z+z^2}.$$

Remark 2.3.3. *Note that the structure of the Poincaré series in a single variable in this case does not fully represent the structure of the minimal resolution we construct. Rather, there is cancellation after specialization. This indicates that while rational single-variable Poincaré series can be useful for asymptotic approximation of Betti numbers, to inspire explicit construction of minimal resolutions, sometimes a more complex multivariate rational function is required.*

Proof of Theorem 2.3.1. Our proof will proceed as follows: first, we establish the hypotheses required for inductively constructing our resolution; second, for the inductive step, we identify kernel elements; third, we prove those kernel elements generate the kernel; fourth, we prove that the resulting resolution is minimal; fifth, we show that this resolution results in a rational Poincaré series.

Step 1: Establish inductive hypotheses for constructing the resolution. We will begin with the initial complex given in Example 2.2.3. Using this as a base case, we will inductively construct a minimal free resolution F of the type we desire. To verify that the complex $F_{\leq 2}$ in Example 2.2.3 is exact at F_1, assume that f is an element in the kernel of d_1 with leading term supported on some δ_i. If i is equal to 1, 2, or 3, we may reduce f by subtracting a monomial multiple of one of the elements $d_2(\epsilon_1), \dots, d_2(\epsilon_{12})$ in a way that strictly reduces the leading term of f; this is possible since no element of the kernel of d_1 can have a unit as a leading coefficient. By iterated reductions of this type, we produce an element in the kernel supported on only δ_4. By the definition of R and d_1, such an element must be a linear combination of $d_2(\epsilon_{13})$, $d_2(\epsilon_{14})$, and $d_2(\epsilon_{15})$, and thus our complex is exact.

It is straightforward to verify that our base case given by $F_{\leq 2}$ in Example 2.2.3 satisfies the following four hypotheses. To state the hypotheses, suppose for the sake of induction that we have produced a complex $F_{\leq n}$ that is exact except at F_0 and F_n.

Hypothesis (Ordering): Assume that for each i, F_i is ordered with respect to d_i, and no element of the kernel of d_n has leading coefficient equal to a unit.

Hypothesis (Generator Poset): For each generator ϵ of F_i, where $1 \leq i \leq n$, each coefficient, up to K-scalar, of $d_i(\epsilon) - \mathrm{LT}(d_i(\epsilon))$ is either zero or lies strictly below $\mathrm{LC}(d_i(\epsilon))$ in the poset in Figure 2.2.

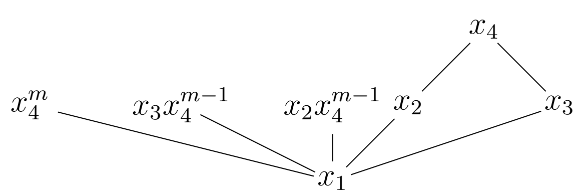


Figure 2.2: A useful partial ordering

Hypothesis (Cover Condition): For each generator δ of a summand of F_i, where i is at most $n-1$ and n is at least two, let the leading coefficient $\mathrm{LC}(d_i(\delta))$ be the monomial s. Then the following holds:

- If $s \in \{x_1, x_2, x_3, x_2x_4^{m-1}, x_3x_4^{m-1}\}$, then there are exactly four elements covering δ in $T_{\leq n}$, and their leading coefficients are x_1, x_2, x_3, and x_4.
- If $s = x_4$, then there are exactly three elements covering δ in $T_{\leq n}$, and their leading coefficients are $x_2x_4^{m-1}$, $x_3x_4^{m-1}$, and x_4^m.
- If $s = x_4^m$, then there are exactly three elements covering δ in $T_{\leq n}$, and their leading coefficients are x_2, x_3, and x_4.

These are the only values that s takes.

Hypothesis (Boundary Condition): For each generator ϵ of a summand of F_i, where i is at least one and $\mathrm{LT}(d_i(\epsilon)) = s\delta$:

If $s \in \{x_2x_4^{m-1}, x_3x_4^{m-1}, x_4^m\}$, then one of the following holds:

(i) $d_i(\epsilon) = s\delta$

(ii) $d_i(\epsilon) = s\delta + \sigma x_1\delta'$, where $\sigma \in \{1, -1\}$, $\mathrm{LC}(d_{i-1}(\delta)) = x_4$, and $\mathrm{LC}(d_{i-1}(\delta')) = x_4^m$

If $s \in \{x_2, x_3\}$, then one of the following holds:

(iii) $d_i(\epsilon) = s\delta$ where $\mathrm{LC}(d_{i-1}(\delta)) = t$ and $st = 0$

(iv) $d_i(\epsilon) = s\delta - x_1\delta'$, where $\mathrm{LT}(d_{i-1}(\delta)) = t\gamma$ for $t \in \{x_2x_4^{m-1}, x_3x_4^{m-1}\}$ with $st \neq 0$ and $\mathrm{LT}(d_{i-1}(\delta')) = x_4^m\gamma$

(v) $d_i(\epsilon) = s\delta - x_1\delta'$ where $\mathrm{LT}(d_{i-1}(\delta)) = t\gamma$ for $t \in \{x_1, x_2, x_3\}$ with $st \neq 0$ and $\mathrm{LT}(d_{i-1}(\delta')) = x_4\gamma$

If $s = x_4$, then one of the following holds:

(vi) $d_i(\epsilon) = s\delta$ where $\mathrm{LC}(d_{i-1}(\delta)) \in \{x_2x_4^{m-1}, x_3x_4^{m-1}, x_4^m\}$

(vii) $d_i(\epsilon) = s\delta - \sigma x_1\delta'$ where $\sigma \in \{1, -1\}$, $d_{i-1}(\delta) = t\gamma - x_1\gamma'$ for $t \in \{x_2x_4^{m-1}, x_3x_4^{m-1}\}$ and $\mathrm{LT}(d_{i-1}(\delta')) = x_4\gamma'$

(viii) $d_i(\epsilon) = s\delta - t\delta' + x_1\delta''$ where $d_{i-1}(\delta) = t\gamma - x_1\gamma'$ for $t \in \{x_1, x_2, x_3\}$, $\mathrm{LT}(d_{i-1}(\delta')) = x_4\gamma$, and $\mathrm{LT}(d_{i-1}(\delta'')) = x_4\gamma'$

(ix) $d_i(\epsilon) = s\delta - t\delta'$ for $t \in \{x_1, x_2, x_3\}$ where $\mathrm{LT}(d_{i-1}(\delta)) = t\gamma$ and $\mathrm{LT}(d_{i-1}(\delta')) = x_4\gamma$

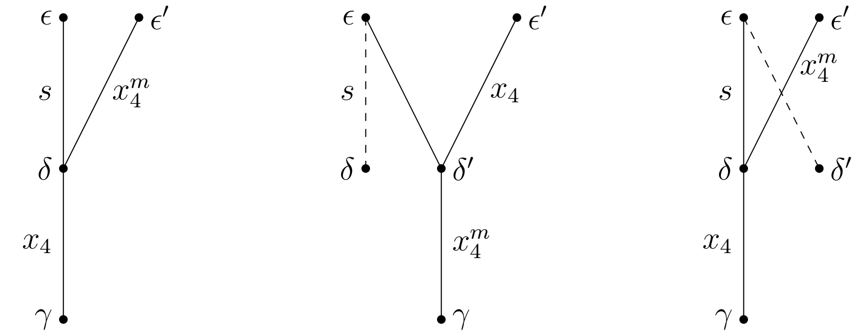
If $s = x_1$, then

(x) $d_i(\epsilon) = s\delta$

Step 2: Inductive construction of kernel elements. Assume that hypotheses (Ordering), (Generator Poset), (Cover Condition), and (Boundary Condition) are satisfied by our complex $F_{\leq n}$, exact except at F_0 and F_n. We will now use hypotheses (Ordering), (Generator Poset), (Cover Condition), and (Boundary Condition) to show that for each generator ϵ of F_n, there exists a set of homogeneous kernel elements whose leading term is supported on ϵ and whose leading coefficients satisfy hypothesis (Cover Condition).

Specifically, assume that ϵ is such that $\mathrm{LC}(d_n(\epsilon)) = s$:

- For each of the cases $s \in \{x_1, x_2, x_3, x_2x_4^{m-1}, x_3x_4^{m-1}\}$ we find an element f_i of $\ker d_n$ with leading term $u\epsilon$ for $u \in \{x_1, x_2, x_3, x_4\}$.
- For $s = x_4$ we find a kernel element f_i with leading term $u\epsilon$ for each $u \in \{x_2x_4^{m-1}, x_3x_4^{m-1}, x_4^m\}$.



$f_i = u\epsilon - x_1\epsilon'$ $\qquad$ $f_i = u\epsilon - \sigma x_1\epsilon'$ $\qquad$ $f_i = u\epsilon - x_1\epsilon'$

Figure 2.3: The case $s \in \{x_2x_4^{m-1}, x_3x_4^{m-1}\}$.

- For $s = x_4^m$ we find a kernel element f_i with leading term $u\epsilon$ for each $u \in \{x_2, x_3, x_4\}$.

Observe that this is precisely what is needed to extend our resolution while satisfying (Cover Condition). We will denote this collection of kernel elements by $\{f_i\}$ and let them be ordered by $\prec$ on the minimal supports (with tie breaking by $\prec_R$ on the leading coefficients).

Let ϵ be a generator of F_n with $\mathrm{LT}(d_n(\epsilon)) = s\delta$ and $\mathrm{LT}(d_{n-1}(\delta)) = t\gamma$. We construct the elements set $\{f_i\}$ in a case-by-case manner as follows.

Case: $s \in \{x_2x_4^{m-1}, x_3x_4^{m-1}\}$**.** We refer to Figure 2.3 throughout this argument. By (Boundary Condition), $d_n(\epsilon)$ is either equal to $s\delta$ or $s\delta + \sigma x_1\delta'$, where $\mathrm{LC}(d_{n-1}(\delta')) = x_4^m$. We suppose $u \in \{x_1, x_2, x_3, x_4\}$ and consider three subcases.

- If $d_n(u\epsilon) = us\delta$ is zero, we set $f_i := u\epsilon$. This can only happen in two situations, either when $s = x_3x_4^{m-1}$ and u is equal to x_1, x_3, or x_4, or else when $s = x_2x_4^{m-1}$ and u is equal to x_1, x_2, or x_4. Note that this assignment of f_i satisfies the four inductive hypotheses.

- If $d_n(\epsilon) = s\delta$ and $d_n(u\epsilon) \neq 0$, then $d_n(u\epsilon)$ must be equal to $x_1x_4^m\delta$. This can only happen when either $s = x_2x_4^{m-1}$ and $u = x_3$ or when $s = x_3x_4^{m-1}$ and $u = x_2$, and both situations require consideration of the binomial relation $x_2x_3 = x_1x_4$. See the left-hand skematic in Figure 2.3 to illustrate the following argument. Since $\delta \in F_{n-1}$ is covered in $T_{\leq n}$ by an element $\epsilon \in F_n$ having $\mathrm{LC}(d_n(\epsilon))$ equal to one of $x_2x_4^{m-1}$ or $x_3x_4^{m-1}$, by (Cover Condition) we have that $\mathrm{LC}(d_{n-1}(\delta)) = x_4$. Thus, again by (Cover Condition), there exists $\epsilon' \in F_n$ with $\mathrm{LT}(d_n(\epsilon')) = x_4^m\delta$. It follows that $d_n(x_1\epsilon') = x_1x_4^m\delta$ since by (Generator Poset), each coefficient of $d_n(\epsilon') - \mathrm{LT}(d_n(\epsilon'))$ is either zero or x_1, and $x_1^2 = 0$. Thus, $d_n(u\epsilon - x_1\epsilon') = 0$ and we can set $f_i := u\epsilon - x_1\epsilon'$. Observe that this assignment of f_i satisfies the four inductive hypotheses.

- If $d_n(\epsilon) = s\delta + \sigma x_1\delta'$, observe that (Boundary Condition) implies that since $d_n(\epsilon) = s\delta + \sigma x_1\delta'$, we have $\mathrm{LC}(d_{n-1}(\delta')) = x_4^m$. We have three subsubcases that arise in this subcase, and note the assignment of f_i given in each of them satisfies the four inductive hypotheses.
  - If $d_n(u\epsilon)$ is equal to zero, we set $f_i := u\epsilon$. This will happen when $u = x_1$ and for certain pairs of u and s when $u = x_2$ or $u = x_3$, with the remaining pairs handled in the subsubcase $su = x_1x_4^m$ below.
  - If $u = x_4$, then $us = 0$ for both possible values of s. Thus, $d_n(u\epsilon) = \sigma x_1x_4\delta'$. See the center skematic in Figure 2.3 illustrating the following argument. Since $\mathrm{LC}(d_{n-1}(\delta')) = x_4^m$, by (Cover Condition) there exists ϵ' such that $\mathrm{LC}(d_n(\epsilon')) = x_4\delta'$. Since by (Boundary Condition) any coefficient of $d_n(\epsilon') - \mathrm{LT}(d_n(\epsilon'))$ is either zero, x_1, x_2, or x_3, and multiplying any of these variables by x_1 results in a zero, we have that $d_n(u\epsilon - \sigma x_1\epsilon') = 0$. Thus, we can set $f_i := u\epsilon - \sigma x_1\epsilon'$.
  - As before, $su = x_1x_4^m$ can only happen in two situations, when either $s = x_2x_4^{m-1}$ and $u = x_3$ or when $s = x_3x_4^{m-1}$ and $u = x_2$. See the right-hand skematic in Figure 2.3 illustrating the following argument. In either event, we have that $d_n(u\epsilon) = x_1x_4^m\delta$, since multiplying x_2 or x_3 by the x_1 in the coefficient of δ' will zero out that term. As in a previous case, by (Cover Condition), we can find an ϵ' such that $\mathrm{LT}(d_n(\epsilon')) = x_4^m\delta$ and $d_n(x_1\epsilon') = x_1x_4^m\delta$. In this case, $d_n(u\epsilon - x_1\epsilon') = 0$, hence we set $f_i := u\epsilon - x_1\epsilon'$.

Case $s = x_4^m$: By (Boundary Condition), $d_n(\epsilon)$ is equal to $s\delta$ or $s\delta + \sigma x_1\delta'$, where $\mathrm{LC}(d_{n-1}(\delta')) = x_4^m$.

- If u equals x_2 or x_3, then since us and ux_1 are both equal to zero, $d_n(u\epsilon) = 0$ and we set $f_i := u\epsilon$.

- If instead u is equal to x_4, then $d_n(u\epsilon)$ is either zero or is equal to $\sigma x_1x_4\delta'$ for some δ' with $\mathrm{LC}(d_{n-1}\delta') = x_4^m$. By (Cover Condition), since $\mathrm{LC}(d_{n-1}\delta') = x_4^m$, there exists a generator ϵ' of F_n with $\mathrm{LT}(d_n(\epsilon')) = x_4\delta'$. By (Boundary Condition) applied to ϵ', since $\mathrm{LC}(d_{n-1}(\delta')) = x_4^m$, we have the stronger condition that $d_n(\epsilon') = x_4\delta'$, so that $d_n(u\epsilon - \sigma x_1\epsilon') = 0$. Thus we set $f_i = u\epsilon - \sigma x_1\epsilon'$.

- If $u = x_1$, then $\mathrm{LT}(d_n(u\epsilon)) = x_1x_4^m\delta$, which is nonzero. By construction, any generator ϵ' which is after ϵ in the ordering $\prec$ of generators of F_n has the leading term of its image under d_n supported on generators that are after or equal to δ in the ordering of generators of d_{n-1}. By (Cover Condition), ϵ is the $\prec$-maximal generator whose image under d_n is supported on δ. Consequently, the equation

$$d_n(u\epsilon) = -d_n\left(\sum_{\epsilon \prec \epsilon_k} r_k\epsilon_k\right)$$

has no solutions for r_k in R, and so no homogeneous kernel element of d_n has leading term $x_1\epsilon$. Since any monomial of R of degree greater than one is divisible by x_4, no minimal generator of $\ker d_n$ has leading term $u\epsilon$ where u is different from x_2, x_3, or x_4.

Case $s \in \{x_1, x_2, x_3\}$: By (Boundary Condition), $d_n(\epsilon) = s\delta$ or $d_n(\epsilon) = s\delta - x_1\delta'$.

- Let $u \in \{x_1, x_2, x_3\}$. Observe that us is either zero or equal to x_2x_3, and that ux_1 is always zero. Thus $d_n(u\epsilon)$ is equal to $us\delta$, and is either zero or equal to $x_1x_4\delta$ (under the equivalence $x_2x_3 = x_1x_4$). In the first case we set $f_i = u\epsilon$. In the second case, we note that since by hypothesis $s\delta := \mathrm{LT}(d_n(\epsilon))$ is among $x_1\delta$, $x_2\delta$, and $x_3\delta$, by (Cover Condition) there exists ϵ' such that $\mathrm{LT}(d_n(\epsilon')) = x_4\delta$. By (Generator Poset), $d_n(x_1\epsilon') = x_1x_4\delta$ since $x_1^2 = x_1x_2 = x_1x_3 = 0$, so that $d_n(u\epsilon - x_1\epsilon') = 0$. We set $f_i = u\epsilon - x_1\epsilon'$.
- Let instead $u = x_4$, and consider the two cases: $d_n(\epsilon) = s\delta$ and $d_n(\epsilon) = s\delta - x_1\delta'$. Again note that by (Cover Condition) there exists ϵ' with $\mathrm{LT}(d_n(\epsilon')) = x_4\delta$ and recall that by (Poset Condition), $d_n(x_1\epsilon') = x_1x_4\delta$.
  - In the case that $d_n(\epsilon) = s\delta = x_1\delta$, we have that $d_n(u\epsilon - x_1\epsilon') = 0$, and we set $f_i = u\epsilon - x_1\epsilon'$. Let instead $d_n(\epsilon) = s\delta$ where s equals x_2 or x_3, and note that by (Boundary Condition), $\mathrm{LC}(d_{n-1}(\delta)) = t$ where $st = 0$. Then by (Boundary Condition) there are two possibilities for $d_n(s\epsilon')$. Because $sx_1 = 0$, either $d_n(s\epsilon')$ is equal to $x_4s\delta$, or else it is equal to $x_4s\delta - st\delta'$ for some generator δ' of F_{n-1}. As we have established, $st = 0$, and so $d_n(s\epsilon') = x_4s\delta$, so that $d_n(u\epsilon - s\epsilon') = 0$. We set $f_i = u\epsilon - s\epsilon'$.
  - If $d_n(\epsilon) = s\delta - x_1\delta'$, then by (Boundary Condition), s is equal to x_2 or x_3. There are two possibilities, both having $\mathrm{LT}(d_{n-1}(\delta)) = t\gamma$ where $st \neq 0$. **The first possibility** is that t is among $\{x_2x_4^{m-1}, x_3x_4^{m-1}\}$ and $\mathrm{LT}(d_{n-1}(\delta')) = x_4^m\gamma$. In this case, by (Cover Condition) there exists ϵ'' such that $\mathrm{LT}(d_n(\epsilon'') = x_4\delta'$. By (Boundary Condition) applied to ϵ', we see that since t is among $\{x_2x_4^{m-1}, x_3x_4^{m-1}\}$, $d_n(\epsilon')$ is either $x_4\delta$ or $x_4\delta - \sigma x_1\delta'$. In either case, $d_n(s\epsilon')$ is equal to $x_4s\delta$. Also by (Boundary Condition), since $\mathrm{LC}(d_{n-1}(\delta'))$ is x_4^{m-1}, we see that $d_n(\epsilon'')$ is equal to $x_4\delta'$. Thus we see that $d_n(u\epsilon - s\epsilon' + x_1\epsilon'') = 0$ and we set $f_i = u\epsilon - s\epsilon' + x_1\epsilon''$.

    The second possibility is that t is among $\{x_1, x_2, x_3\}$ and $\mathrm{LT}(d_{n-1}(\delta')) = x_4\gamma$. In this case we see that by (Boundary Condition), since $sx_1 = 0$, we have that $d_n(s\epsilon')$ is equal to $x_4s\delta - st\delta'$. Observe that for s equal to x_2 or x_3 and t among $\{x_1, x_2, x_3\}$, the fact that st is nonzero implies that st equals $x_2x_3 = x_1x_4$. It follows that $d_n(u\epsilon - s\epsilon') = 0$ and we set $f_i = u\epsilon - s\epsilon'$.

Case $s = x_4$:

- For u any monomial of $\mathbb{N}$–degree less than or equal to $m-1$, or for $u = x_1x_4^{m-1}$, $d_n(u\epsilon)$ is nonzero, and as argued earlier, since the leading coefficient of ϵ is x_4, (Cover Condition) implies that the ordering $\prec$ of the generators of F_n precludes any homogeneous kernel element of d_n having leading term $u\epsilon$.

The monomials having $\mathbb{N}$–degree equal to m are $x_2x_4^{m-1}$, $x_3x_4^{m-1}$, and x_4^m. We show that each of these monomials is the leading coefficient of homogeneous kernel element of d_n having leading support ϵ.

If $u \in \{x_2x_4^{m-1}, x_3x_4^{m-1}\}$, then by (Boundary Condition), either $d(u\epsilon)) = 0$ or $d(u\epsilon) = -x_1x_4^m\delta'$, where $\mathrm{LC}(d(\delta')) = x_4$. In the first case we set $f_i = u\epsilon$. In the latter case, by (Cover Condition) there exists ϵ' such that $\mathrm{LT}(d(\epsilon')) = x_4^m\delta'$. By (Generator Poset), $d_n(x_1\epsilon') = x_1x_4^m\delta'$, so that $d_n(u\epsilon + x_1\epsilon') = 0$ and we set $f_i = u\epsilon + x_1\epsilon'$.

If $u = x_4^m$, then by (Boundary Condition) we have that either $d_n(\epsilon) = x_4\delta$, so that $d_n(u\epsilon) = 0$, or $d_n(u\epsilon) = \sigma x_1x_4^m\delta'$, where $\mathrm{LC}(d_{n-1}(\delta')) = x_4$. In the latter case, we have by (Cover Condition) that there exists ϵ' such that $\mathrm{LT}(d_n(\epsilon')) = x_4^m\delta'$. It follows that $d_n(u\epsilon - \sigma x_1\epsilon') = 0$ and we set $f_i = u\epsilon - \sigma x_1\epsilon'$.

Thus, we have constructed a set of kernel elements $\{f_i\}$ satisfying the properties stated at the beginning of this step.

Step 3: Proof that the elements $\{f_i\}$ generate the kernel. Note that by our inductive construction, no element of the kernel has a leading coefficient equal to a unit in R. Given a homogeneous element f of $\ker d_n$ with $\mathrm{LT}(f) = v\epsilon$, we have by hypothesis (Cover Condition) that $\mathrm{LC}(d_n(\epsilon)) = s$ is among the cases above. If $s \in \{x_1, x_2, x_3, x_2x_4^{m-1}, x_3x_4^{m-1}\}$, then since by hypothesis the leading coefficient of f is divisible by some variable, f may be reduced by one of the f_i's with the same minimal support to a kernel element with strictly larger leading term.

If $s = x_4$, then as established above, v is divisible by one of $x_2x_4^{m-1}$, $x_3x_4^{m-1}$, or x_4^m, and f may be reduced by one of the f_i's with the same minimal support to a kernel element with strictly larger leading term.

As there is a finite collection of possible leading terms, this process will reduce f to zero, showing that the f_i are a generating set for $\ker d_n$.

Step 4: Proof that this is a minimal resolution.

Now that we have proved that the set of f_i's generate the kernel of d_n, we proceed by augmenting $F_{\leq n}$ with a free R-module with basis in bijection with the f_i's. Define a map d_{n+1} sending each new basis element to its associated f_i, and the result is a complex $F_{\leq n+1}$, which is exact except at F_0 and F_{n+1}. We need to show that our choice of kernel generators f_i is a minimal set of generators.

Since none of the f_i whose leading coefficient is a variable can be written as an R-linear combination of the others, we only need consider the minimality of f_i's whose leading coefficient is $x_kx_4^{m-1}$, where k is 2, 3, or 4. Let $\mathrm{LT}(f_i) = x_kx_4^{m-1}\epsilon$, where $\mathrm{LT}(d_n(\epsilon)) = x_4\gamma$ and $f_j \neq f_i$ be such that $[\epsilon](s_jf_j)$, the coefficient of s_jf_j on the summand generated by ϵ, is also $x_kx_4^{m-1}$. Since, by construction, the only coefficients appearing in f_j are in $\{x_1, x_2, x_3, x_4, x_2x_4^{m-1}, x_3x_4^{m-1}, x_4^m\}$, we see by divisibility that $[\epsilon](f_j)$ is x_2, x_3, or x_4. Also by construction, $[\epsilon](f_j) = x_4$ implies that $\mathrm{LT}(f_j) = x_4\epsilon$, which is a contradiction with $\mathrm{LT}(d_n(\epsilon)) = x_4\gamma$, thus $[\epsilon](f_j)$ is x_2 or x_3, and $s_j = x_4^{m-1}$. Again by construction, $\mathrm{LT}(f_j) = x_4\epsilon'$, where $\mathrm{LT}(d_n(\epsilon')) = x_k\gamma$, hence

$\mathrm{LC}(s_j f_j) = x_4^m \neq 0$, so that if $\mathrm{LT}\left(\sum_{\ell \neq j} s_\ell f_\ell\right) = x_k x_4^{m-1}\epsilon$, then there exists ℓ such that $[\epsilon'] s_\ell f_\ell = x_4^m$. But no f_ℓ with $\ell \neq j$ can have $[\epsilon'] f_\ell$ divisible by x_4. Thus f_i is minimal, and the collection of f_i's is in fact a minimal generating set.

Step 5: Verification that the inductive result satisfies the four hypotheses and produces the matrix A, computation of the rational Poincaré series. It is immediate from the construction above that the hypotheses (Ordering), (Generator Poset), (Cover Condition), and (Boundary Condition) are satisfied by the augmented complex $F_{\leq n+1}$. Further, it is immediate from these four hypotheses that the matrix A given in the statement of the theorem is correct. We therefore have established an inductive construction of the minimal free resolution of K over R.

Having constructed our desired minimal free resolution, by Lemma 2.2.6 we have that $P_K^R(z,t)$ is rational of the form

$$\frac{f(z;\boldsymbol{y})}{\chi(z;\boldsymbol{y},1)}$$

where $\chi(z;\boldsymbol{y},t)$ is the characteristic polynomial of the matrix A. Computation in Macaulay2 gives that

$$\begin{aligned}\chi(z;\boldsymbol{y},1) =& 1 - z(\boldsymbol{y}^{\deg(x_1)} + \boldsymbol{y}^{\deg(x_2)} + \boldsymbol{y}^{\deg(x_3)}) \\ &- z^2(\boldsymbol{y}^{\deg(x_2 x_4^m)} + \boldsymbol{y}^{\deg(x_3 x_4^m)} + \boldsymbol{y}^{\deg(x_4^{m+1})}) + z^3 \boldsymbol{y}^{\deg(x_1 x_4^{m+1})}.\end{aligned}$$

Using the minimal resolution construction given above, we compute that $F_{\leq 3}(z;\boldsymbol{y}) \cdot \chi(z;\boldsymbol{y},1)$ is given by

$$1 + z\boldsymbol{y}^{\deg(x_4)} + z^4 E(z;\boldsymbol{y})\,,$$

where $E(z;\boldsymbol{y})$ is a polynomial in $K[z;\boldsymbol{y}]$. By Lemma 2.2.6,

$$f(z;\boldsymbol{y}) = \chi(z;\boldsymbol{y},1) \cdot F(z;\boldsymbol{y})\,,$$

so that

$$f(z;\boldsymbol{y}) - F_{\leq 3}(z;\boldsymbol{y}) \cdot \chi(z;\boldsymbol{y},1) = (F(z;\boldsymbol{y}) - F_{\leq 3}(z;\boldsymbol{y})) \cdot \chi(z;\boldsymbol{y},1)$$

is a polynomial divisible by z^4. Since the z-degree of $f(z;\boldsymbol{y})$ is at most three (by the degree of $\chi(z;\boldsymbol{y},1)$), we see that $f(z;\boldsymbol{y}) = 1 + z\boldsymbol{y}^{\deg(x_4)}$, and the rational form follows. □

Chapter 3 Computing the poset $\mathrm{P}(\Delta)$

Let the matrix V have columns given by $\{(1, v_i)\}_{0\leq i\leq d}$, where the v_i's are the vertices of a lattice d-simplex Δ. Recall that the normalized volume $\boldsymbol{v}$, the number of elements of Π_Δ, may be computed by $\boldsymbol{v} = |\det V|$. Recall also that the set Π_Δ is the image of $[0,1)^{d+1}$ under the linear transformation V, so that the preimage of a lattice point of Π_Δ must be a rational point of $[0,1)^{d+1}$ with denominator $\boldsymbol{v}$. We may therefore compute the set of points in $\Pi_\Delta \cap \mathbb{Z}^{d+1}$ by considering each element of the form

$$\left\{ V \cdot \left(\frac{b_0}{\boldsymbol{v}}, \cdots, \frac{b_d}{\boldsymbol{v}}\right)^T \text{ such that } 0 \leq b_i < \boldsymbol{v} \right\},$$

and throwing out the ones which are not integer points. Unfortunately, this test set grows as $\boldsymbol{v}^{d+1}$, and there is no easy way to describe the lattice points among them.

Normaliz [8] gives a more efficient implementation based on the fact that the matrix V has a representation $V = UH$ where U is a unimodular matrix and H is in Hermite normal form. Bruns et al. [9] show that, for $\{h_{i,i}\}_{0\leq i\leq d}$ given by the diagonal entries of the matrix H, lattice points in

$$[0, h_{0,0}) \times \cdots \times [0, h_{d,d})$$

are representatives of the quotient classes (in $\mathbb{Z}^{d+1}$ modulo the $(1, v_i)$'s), of the elements of $\Pi_\Delta \cap \mathbb{Z}^{d+1}$. It is then sufficient to consider the image under V of the elements $(V^{-1} \cdot x) \mod \mathbb{Z}^{d+1}$ for $x \in [0, h_{0,0}) \times \cdots \times [0, h_{d,d})$. This modular arithmetic is implemented in a computer easily enough, but introduces number theory to any analysis of the poset $P(\Delta)$. Thus, moving forward we make a simplifying assumption.

Definition 3.0.1. *We say that a lattice d-simplex has a **unimodular facet** if there exists a permutation π in $\mathfrak{S}_{d+1}$ such that $\operatorname{conv}(\{v_{\pi_1}, \ldots, v_{\pi_d}\})$ is a unimodular lattice $(d-1)$-simplex.*

If Δ has a unimodular facet, then we may define a lattice preserving transformation taking Δ to $\operatorname{conv}(e_1, \ldots, e_d, z)$ where the e_i are the standard basis vectors of $\mathbb{R}^d$ and z is a lattice point in $\mathbb{Z}^d$. Our goal in this chapter is to find a description of the relations in $P(\Delta)$ in terms of the coordinates of the point z.

3.1 Lattice Simplices with a Unimodular Facet and their Posets

Definition 3.1.1. *Let $\lambda = (\lambda_1, \ldots, \lambda_d)$ be a lattice point in $\mathbb{N}^d$ such that $\sum_{i=1}^d \lambda_i = n$, and $\Delta_\lambda := \operatorname{conv}(e_1, \ldots, e_d, \lambda) \subset \mathbb{R}^d$. We will use the shortened notation $\Pi_\lambda := \Pi_{\Delta_\lambda}$ and $P(\lambda) := P(\Delta_\lambda)$.*

Remark 3.1.2. *The simplices Δ_λ are defined in a similar manner to the simplices $\Delta_{(1,\mathbf{q})}$ studied by Braun, Davis and Solus [7], but are not the same family. For example, it is possible for Δ_λ to contain no interior lattice points, but this is not the case for any $\Delta_{(1,\mathbf{q})}$.*

The following is a straightforward determinant calculation.

Proposition 3.1.3. *The number of lattice points in Π_λ, which is equal to the normalized volume of Δ_λ, is $\sum_{i=1}^d \lambda_i - 1 = n-1$.*

We can describe the integer points in Π_λ using only the entries of λ.

Proposition 3.1.4. *For each integer b with $0 \le b < n-1$, there is a unique lattice point $p(b)$ in Π_λ given by*

$$p(b) = \left(\left(\sum_{i=1}^d \left\lceil \frac{b\lambda_i}{n-1} \right\rceil \right) - b \, , \, \left\lceil \frac{b\lambda_1}{n-1} \right\rceil \, , \, \ldots \, , \, \left\lceil \frac{b\lambda_d}{n-1} \right\rceil \right) . \tag{3.1}$$

Every integer point in Π_λ arises in this manner, and thus we identify the integer b with the lattice point $p(b)$.

Proof. For an element $\sum_{i=1}^d \gamma_i(1, e_i) + \gamma_{d+1}(1, \lambda) \in \Pi_\lambda \cap \mathbb{Z}^{d+1}$, we have

$$\left(\left(\sum_{i=1}^{d+1} \gamma_i \right) \, , \, (\gamma_1 + \gamma_{d+1}\lambda_1) \, , \, \ldots \, , \, (\gamma_d + \gamma_{d+1}\lambda_d) \right) \in \mathbb{Z}^{d+1} .$$

Because of the condition that each γ_i is strictly less than one, for each i we have

$$\gamma_i = \lceil \gamma_{d+1}\lambda_i \rceil - \gamma_{d+1}\lambda_i \, ,$$

thus

$$\begin{aligned} &\left(\gamma_{d+1} + \sum_{i=1}^d \left(\lceil \gamma_{d+1}\lambda_i \rceil - \gamma_{d+1}\lambda_i \right) , \lceil \gamma_{d+1}\lambda_1 \rceil, \ldots, \lceil \gamma_{d+1}\lambda_d \rceil \right) \\ = &\left(\gamma_{d+1} \left(1 - \sum_{i=1}^d \lambda_i \right) + \sum_{i=1}^d \lceil \gamma_{d+1}\lambda_i \rceil, \lceil \gamma_{d+1}\lambda_1 \rceil, \ldots, \lceil \gamma_{d+1}\lambda_d \rceil \right) . \end{aligned}$$

Observe that the first coordinate of this vector is an integer, hence

$$\gamma_{d+1} \left(1 - \sum_{i=1}^d \lambda_i \right) = \gamma_{d+1}(1-n) \in \mathbb{Z} \, .$$

It follows that γ_{d+1} is a rational number of the form $b/(n-1)$, and every lattice point arises in this manner and is of the form

$$\left(\left(\sum_{i=1}^d \left\lceil \frac{b\lambda_i}{n-1} \right\rceil \right) - b \, , \, \left\lceil \frac{b\lambda_1}{n-1} \right\rceil \, , \, \ldots \, , \, \left\lceil \frac{b\lambda_d}{n-1} \right\rceil \right) .$$

Since there are $n-1$ lattice points in Π_λ by Proposition 3.1.3, there must be one unique lattice point for each $0 \le b < n-1$. □

Using the notation from (3.1), for $0 \le b < n-1$ we have that the zeroth coordinate of $p(b)$ is

$$p(b)_0 := \left(\sum_{i=1}^{d} \left\lceil \frac{b\lambda_i}{n-1} \right\rceil \right) - b\,.$$

Recall that we freely identify the integer b with the lattice point $p(b)$. The following lemma provides a connection between the parameterization of the integer points in Π_λ and the order in $P(\lambda)$.

Lemma 3.1.5. *For $i, j \in P(\lambda)$ with $i \neq j$, we have $i \prec j$ if and only if $i < j$ and $p(i) + p(j-i) = p(j)$.*

Proof. For the forward direction, if $i \prec j$, then by Proposition 3.1.4 there exists a point $p(\ell) \in P(\lambda)$ such that $p(i) + p(\ell) = p(j)$. Note that $\ell > 0$ since $p(0) = 0$. It follows that for all $1 \le t \le d$, we have

$$\left\lceil \frac{i\lambda_t}{n-1} \right\rceil + \left\lceil \frac{\ell\lambda_t}{n-1} \right\rceil = \left\lceil \frac{j\lambda_t}{n-1} \right\rceil .$$

Given this, we have that $p(i)_1 + p(\ell)_1 = p(j)_1$ reduces to $i + \ell = j$, forcing $\ell = j - i > 0$, as desired.

For the reverse direction, if $i < j$ and $p(i) + p(j-i) = p(j)$, then we have $i \prec j$ by definition. $\square$

We now give two propositions demonstrating how Lemma 3.1.5 can be used in practice.

Proposition 3.1.6. *If $i \prec j$ in $P(\lambda)$, then also $j - i \prec j$ in $P(\lambda)$.*

Proof. By Lemma 3.1.5, we have $i \prec j$ if and only if $i < j$ and $p(i) + p(j-i) = p(j)$ if and only if $j - i < j$ and $p(i) + p(j-i) = p(j)$ if and only if $j - i \prec j$. $\square$

Proposition 3.1.7. *Let $\lambda = (n-2, 2)$. Then $P(n-2, 2)$ is equal to the following poset on the elements $\{1, 2, \ldots, n-2\}$: The minimal elements of $P(n-2, 2)$ are $\{1, 2, \ldots, \lfloor \frac{n-1}{2} \rfloor\}$ and the maximal elements are $\{\lfloor \frac{n-1}{2} \rfloor + 1, \ldots, n-2\}$. The cover relations are that the maximal element $\lfloor \frac{n-1}{2} \rfloor + j$ covers $\{j, j+1, \ldots, \lfloor \frac{n-1}{2} \rfloor\}$.*

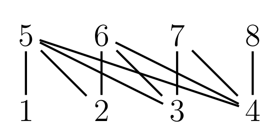


Figure 3.1: The poset $P(8, 2)$.

Proof. By Lemma 3.1.5, we see that $i \prec j$ if and only if $i < j$ and the following hold:

$$\left\lceil \frac{2i}{n-1} \right\rceil + \left\lceil \frac{2(j-i)}{n-1} \right\rceil = \left\lceil \frac{2j}{n-1} \right\rceil \tag{3.2}$$

$$\left\lceil \frac{i(n-2)}{n-1} \right\rceil + \left\lceil \frac{(j-i)(n-2)}{n-1} \right\rceil = \left\lceil \frac{j(n-2)}{n-1} \right\rceil \tag{3.3}$$

It is straightforward to verify that these equations hold for the values claimed in the proposition statement.

To show that no other pairs $i < j$ lead to relations $i \prec j$, suppose that $1 \leq i < j \leq \left\lfloor \frac{n-1}{2} \right\rfloor$. Then in (3.2), we obtain $1 + 1 = 1$, which is false. Similarly, if $\left\lfloor \frac{n-1}{2} \right\rfloor + 1 \leq i < j \leq n-2$, then in (3.2) we obtain $2+2 = 2$, which is again false. □

3.2 Characterizing the Relations in $P(\lambda)$

While Lemma 3.1.5 is a reasonable first tool, as Propositions 3.1.6 and 3.1.7 illustrate, in general it is difficult to compute these relations directly. Thus, we need to create a more sophisticated mechanism through which to study $P(\lambda)$. In this section, we establish in Theorem 3.2.2 a number-theoretic characterization of the relations in $P(\lambda)$. Further, Corollary 3.2.3 provides a particularly simple characterization in the case where each part of λ is relatively prime to $n-1$.

For $0 \leq i < n-1$, define the non-negative integers $r_{t,i}$ and $0 \leq s_{t,i} < n-1$ by

$$i\lambda_t = r_{t,i}(n-1) + s_{t,i}\,. \tag{3.4}$$

Lemma 3.2.1. *We have $i \prec j$ in $P(\lambda)$ if and only if $i < j$ and for every $t \in \{1, \ldots, d\}$ we have*

$$\frac{s_{t,i} + s_{t,j-i} - s_{t,j}}{n-1} = \left\lceil \frac{s_{t,i}}{n-1} \right\rceil + \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil - \left\lceil \frac{s_{t,j}}{n-1} \right\rceil \tag{3.5}$$

Proof. After adding and subtracting (3.4) for the values i, $j-i$, and j, we obtain

$$r_{t,i} + r_{t,j-i} - r_{t,j} = \frac{-s_{t,i} - s_{t,j-i} + s_{t,j}}{n-1}\,. \tag{3.6}$$

By dividing both sides of (3.4) by $n-1$ and taking the ceiling of both sides, we see that

$$\left\lceil \frac{\ell\lambda_t}{n-1} \right\rceil = r_{t,\ell} + \left\lceil \frac{s_{t,\ell}}{n-1} \right\rceil\,. \tag{3.7}$$

Adding (3.7) with itself for ℓ equal to i and $j-i$, then subtracting the equation with $\ell = j$, and further applying (3.6), we obtain

$$\begin{aligned} &\left\lceil \frac{i\lambda_t}{n-1} \right\rceil + \left\lceil \frac{(j-i)\lambda_t}{n-1} \right\rceil - \left\lceil \frac{j\lambda_t}{n-1} \right\rceil \\ &= r_{t,i} + r_{t,j-i} - r_{t,j} + \left\lceil \frac{s_{t,i}}{n-1} \right\rceil + \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil - \left\lceil \frac{s_{t,j}}{n-1} \right\rceil \\ &= \frac{-s_{t,i} - s_{t,j-i} + s_{t,j}}{n-1} + \left\lceil \frac{s_{t,i}}{n-1} \right\rceil + \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil - \left\lceil \frac{s_{t,j}}{n-1} \right\rceil\,. \end{aligned}$$

Recall that $i \prec j$ in $P(\lambda)$ if and only if $p(i) + p(j-i) = p(j)$ if and only if for all t, we have that

$$\left\lceil \frac{i\lambda_t}{n-1} \right\rceil + \left\lceil \frac{(j-i)\lambda_t}{n-1} \right\rceil - \left\lceil \frac{j\lambda_t}{n-1} \right\rceil = 0\,,$$

which by our computation above holds if and only if

$$\frac{s_{t,i} + s_{t,j-i} - s_{t,j}}{n-1} = \left\lceil \frac{s_{t,i}}{n-1} \right\rceil + \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil - \left\lceil \frac{s_{t,j}}{n-1} \right\rceil\,.$$

□

Theorem 3.2.2. *Let λ be an integer point in $\mathbb{Z}^d_{\geq 1}$ with coordinates summing to n. We have $i \prec j$ in $P(\lambda)$ if and only if $i < j$ and for each $t \in \{1, \ldots, d\}$, one of the following holds:*

1. $s_{t,i} > s_{t,j} > 0$,
2. $s_{t,i} = 0$ *and* $s_{t,j} = s_{t,j-i}$, *or*
3. $s_{t,j} = s_{t,i} > 0$ *and* $s_{j-i} = 0$.

Proof. **Forward implication:** Suppose that $i \prec j$ in $P(\lambda)$, and thus by Lemma 3.2.1 the s-values satisfy (3.5). We consider five cases:

- $s_{t,i} = 0$
- $s_{t,i} > s_{t,j} = 0$
- $s_{t,i} = s_{t,j} > 0$
- $s_{t,i} > s_{t,j} > 0$
- $s_{t,j} > s_{t,i} > 0$

Case 1: $s_{t,i} = 0$. If $s_{t,i} = 0$, then by (3.5) we have that

$$\frac{s_{t,j-i} - s_{t,j}}{n-1} = \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil - \left\lceil \frac{s_{t,j}}{n-1} \right\rceil\,.$$

Thus $\dfrac{s_{t,j-i} - s_{t,j}}{n-1}$ is equal to an integer, and the fact that $0 \leq s_{t,\ell} < n-1$ implies that $s_{t,j-i} - s_{t,j} = 0$. Thus, we must have $s_{t,j-i} = s_{t,j}$. This establishes the second condition in the theorem statement.

Case 2: $s_{t,i} > s_{t,j} = 0$. In this case, (3.5) implies

$$\frac{s_{t,i} + s_{t,j-i}}{n-1} = \left\lceil \frac{s_{t,i}}{n-1} \right\rceil + \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil\,.$$

Thus $\dfrac{s_{t,i} + s_{t,j-i}}{n-1}$ is an integer, and again since $0 \leq s_{t,\ell} < n-1$ and $0 < s_{t,i} < n-1$ we have that Thus, it is impossible to have $s_{t,i} > s_{t,j} = 0$.

Case 3: $s_{t,i} = s_{t,j} > 0$. In this case, (3.5) implies

$$\frac{s_{t,j-i}}{n-1} = \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil .$$

This forces $s_{t,j-i} = 0$, resulting in the third condition in the theorem statement.

Case 4: $s_{t,i} > s_{t,j} > 0$. In this case, (3.5) implies $\dfrac{s_{t,i} + s_{t,j-i} - s_{t,j}}{n-1}$ is equal to an integer, and the fact that every $0 \le s_{t,\ell} < n-1$ implies this integer is 0 or 1. Since $n-1 > s_{t,i} - s_{t,j} > 0$, we must have $\dfrac{s_{t,i} + s_{t,j-i} - s_{t,j}}{n-1} = 1$, and also the right-hand side of (3.5) is equal to 1. Thus, the first condition in the theorem statement is possible if $i \prec j$.

Case 5: $s_{t,j} > s_{t,i} > 0$. Following the same logic as in the previous case, we must have $\dfrac{s_{t,i} + s_{t,j-i} - s_{t,j}}{n-1} = 0$ and thus $s_{t,j-i} \neq 0$. But then the right-hand side of (3.5) is equal to 0 while the right-hand side is equal to 1, a contradiction.

Reverse implication: We verify that each of the three conditions listed in the theorem statement imply that (3.5) is valid.

First, by equation (3.6) we have $\dfrac{s_{t,i} + s_{t,j-i} - s_{t,j}}{n-1} \in \mathbb{Z}$. Combining $n-1 > s_{t,i} > s_{t,j} > 0$ and the general bounds $0 \le s_{t,\ell} < n-1$ for all ℓ, it follows that $\dfrac{s_{t,i} + s_{t,j-i} - s_{t,j}}{n-1} = 1$. Thus, $s_{t,i} + s_{t,j-i} - s_{t,j} = n-1$. Since $n-1 > s_{t,i} - s_{t,j} > 0$, we have $s_{t,j-i} = n-1-(s_{t,i} - s_{t,j}) > 0$, and thus

$$\left\lceil \frac{s_{t,i}}{n-1} \right\rceil + \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil - \left\lceil \frac{s_{t,j-i}}{n-1} \right\rceil = 1 .$$

We conclude that equation (3.5) holds.

Second, if $s_{t,i} = 0$ and $s_{t,j} = s_{t,j-i}$, then it is immediate that (3.5) holds.

Finally, if $s_{t,j} = s_{t,i} > 0$ and $s_{j-i} = 0$, then again it is immediate that (3.5) holds. □

The following corollary illustrates a special case of Theorem 3.2.2 that we will focus on in the remainder of this paper.

Corollary 3.2.3. *Let λ be an integer point in $\mathbb{Z}^d_{\ge 1}$ with coordinates summing to n where each coordinate is coprime to $n-1$, i.e. $\gcd(n-1, \lambda_t) = 1$. Then $i \prec j$ in $P(\lambda)$ if and only if $s_{t,i} > s_{t,j} > 0$ for every t.*

Proof. If $\gcd(n-1, \lambda_t) = 1$, then $s_{t,i} \neq 0$ for all i. Thus, the second and third conditions in Theorem 3.2.2 do not apply. □

We can use Corollary 3.2.3 to prove the following structural result regarding $P(\lambda)$ in the case where each part of λ is coprime to $n-1$.

Theorem 3.2.4. *Let λ be an integer point in $\mathbb{Z}^d_{\ge 1}$ with coordinates summing to n such that each λ_t is coprime to $n-1$. Then $P(\lambda)$ is self-dual.*

Proof. We claim that $\varphi : x \to n-1-x$ for $x \in [n-2]$ is an order-reversing poset isomorphism. It is clear that φ is a bijection. To see that φ is order-reversing, observe that by Corollary 3.2.3, we have that $i \prec j$ if and only if

$$s_{t,i} > s_{t,j} \text{ for all } t\,. \tag{3.8}$$

Due to the fact that $\gcd(n-1, \lambda_t) = 1$, we have that $s_{t,i} + s_{t,n-i-i} = n-1$ for all i and t, and thus (3.9) holds if and only if

$$s_{t,n-1-j} > s_{t,n-1-i} \text{ for all } t\,. \tag{3.9}$$

This final condition holds if and only if $n-1-j \prec n-1-i$, as desired. □

3.3 Partitions With One Distinct Part

When $\lambda = (x, x, \ldots, x)$ has v occurrences of x, it is immediate that x is coprime to $n-1 = vx-1$. In this case, $P(\lambda)$ has a direct interpretation as a subposet of $\mathbb{Z}^2$.

Theorem 3.3.1. *For $\lambda = (x, x, \ldots, x)$ with v occurrences of x, we have that $P(\lambda)$ is isomorphic to the poset with elements*

$$\{(r,p) : 0 \le r < x, 0 \le p < v\} \setminus \{(0,0), (x-1, v-1)\}$$

and order relation $(r,p) \prec (r',p')$ if both $p > p'$ and $r' > r$.

Proof. For $1 \le i \le vx-2$, write

$$i = r_i v + p_i$$

where $0 \le r_i < x$ and $0 \le p_i < v$, but we do not have simultaneously $r_i = x-1$ and $p_i = v-1$. Then

$$\begin{aligned}
s_i &= ix - \left\lfloor \frac{ix}{xv-1} \right\rfloor (xv-1) \\
&= x(r_i v + p_i) - \left\lfloor \frac{(r_i v + p_i)x}{xv-1} \right\rfloor (xv-1) \\
&= xr_i v + xp_i - \left\lfloor \frac{xr_i v - r_i + r_i + p_i x}{xv-1} \right\rfloor (xv-1) \\
&= xr_i v + xp_i - \left(r_i + \left\lfloor \frac{r_i + p_i x}{xv-1} \right\rfloor \right) (xv-1) \\
&= r_i + xp_i - \left\lfloor \frac{r_i + p_i x}{xv-1} \right\rfloor (xv-1) \\
&= r_i + xp_i
\end{aligned}$$

where the final equality is a result of the bounds on r_i and p_i forcing the floor function to be zero. Thus, if $i = r_i v + p_i$ and $j = r_j v + p_j$, then we have $i \prec j$ in $P(\lambda)$ if and only if $i < j$ and $s_i > s_j$, which happens if and only if the following two conditions simultaneously occur:

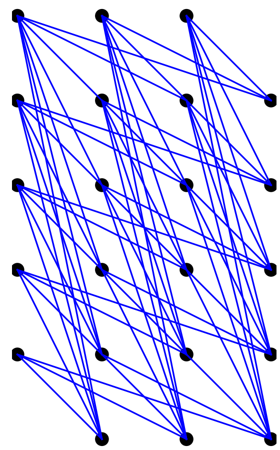


Figure 3.2: $P(4,4,4,4,4,4)$.

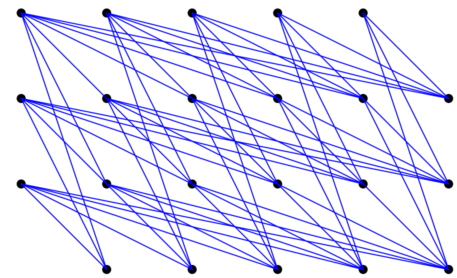


Figure 3.3: $P(6,6,6,6)$.

- $p_i > p_j$ or $p_i = p_j$ with $r_i > r_j$
- $r_j > r_i$ or $r_j = r_i$ with $p_j > p_i$

The only way for both conditions to simultaneously occur is to have $p_i > p_j$ and $r_j > r_i$, and thus our proof is complete. $\square$

The following corollary follows immediately.

Corollary 3.3.2. *The posets for $\lambda = (x, x, \dots, x)$ where x occurs v times and $\lambda' = (v, v, \dots, v)$ where v occurs x times are isomorphic.*

Corollary 3.3.2 is interesting because the two lattice simplices corresponding to λ and λ' are in different dimensions. As an aside, we remark that the order on the lattice points within a rectangular grid given in Theorem 3.3.1 corresponds to the reflexive closure of the direct product of two strict total orders.

Example 3.3.3. *Figures 3.2 and 3.3 show the Hasse diagrams of the posets $P(4,4,4,4,4,4)$ and $P(6,6,6,6)$, respectively, embedded in $\mathbb{Z}^2$ as described in Theorem 3.3.1. This illustrates the isomorphism obtained by switching the roles of x and v.*

3.4 Partitions With Two Distinct Parts

The situation for λ with two distinct parts is significantly more complicated than for one distinct part. Rather than consider arbitrary pairs of distinct parts for λ, we will consider the special case where one of the parts is a multiple of the other. Specifically, we use the following setup.

Setup 3.4.1. *Let $\lambda = (x, \ldots, x, ax, \ldots, ax)$ where the multiplicity of x is $ua + v$ and the multiplicity of ax is $v - (u+1)$, with:*

$$\begin{aligned} 3 &\leq a \leq x \\ 0 &\leq u \leq a - 3 \\ u + 2 \leq v &\leq \min\left(a - 1, \frac{a(x-1)}{x}\right) . \end{aligned}$$

Let $n = |\lambda|$, so that

$$n - 1 = x[(a+1)(v-1) + 1] - 1 = (xa(v-1)) + xv - 1 \,. \tag{3.10}$$

For $0 \leq i \leq n-2$, define as usual

$$s_{1,i} := ix - (n-1)\left\lfloor \frac{ix}{n-1} \right\rfloor \quad \text{and} \quad s_{2,i} := iax - (n-1)\left\lfloor \frac{iax}{n-1} \right\rfloor .$$

As in the proof of Theorem 3.3.1, our analysis will require us to represent i as a quotient with remainder. In this case, we will use a combination of two quotients-with-remainder from applying the division algorithm twice. Observing that $n/x = (a+1)(v-1)+1$, we write $0 \leq i \leq n-2$ as

$$i = \frac{n}{x} r_i + (v-1)p_i + q_i \tag{3.11}$$

subject to the following inequalities:

$$0 \leq r_i < x \tag{3.12}$$

with

$$0 \leq (v-1)p_i + q_i < n/x = (a+1)(v-1) + 1 \,,$$

and

$$0 \leq p_i < a + 2 \tag{3.13}$$

with

$$0 \leq q_i < v - 1 \,, \tag{3.14}$$

where $p_i = a + 1$ implies $q_i = 0$.

Our first goal is to express $s_{1,i}$ and $s_{2,i}$ as explicit functions of r_i, p_i, q_i, x, a, and v.

Lemma 3.4.2. $s_{1,i} = r_i + x\big[(v-1)p_i + q_i\big]$.

Proof. By (3.11), since $ix = nr_i + x\big[(v-1)p_i + q_i\big]$, we have that

$$\begin{aligned} s_{1,i} &= nr_i + x\big[(v-1)p_i + q_i\big] - (n-1)\left\lfloor \frac{nr_i + x\big[(v-1)p_i + q_i\big]}{n-1} \right\rfloor \\ &= r_i + x\big[(v-1)p_i + q_i\big] - (n-1)\left(-r_i + \left\lfloor \frac{nr_i + x\big[(v-1)p_i + q_i\big]}{n-1} \right\rfloor\right) \\ &= r_i + x\big[(v-1)p_i + q_i\big] - (n-1)\left\lfloor \frac{r_i + x\big[(v-1)p_i + q_i\big]}{n-1} \right\rfloor . \end{aligned}$$

Observe that equations (3.10), (3.14), (3.13), and (3.12) imply that

$$0 \le r_i + x\big[(v-1)p_i + q_i\big] \le x - 1 + x(a+1)(v-1) = n-1\,,$$

with equality only if $r_i = x-1$ and $p_i = a+1$ simultaneously. But in this case, we have that

$$ix = (x-1)n + x(a+1)(v-1) = (n-1)x\,,$$

a contradiction with $i \le n-2$. Thus the right hand floor term is zero in our expression for $s_{1,i}$, and the result follows. □

Define the function

$$f(i) :=\ ar_i - (xv-1)p_i + xaq_i\,,$$

and associated set partition $[n-2] = \uplus_k F_k$ given by

$$F_k := \left\{ i : k = -\left\lfloor \frac{f(i)}{n-1} \right\rfloor \right\}.$$

Lemma 3.4.3. *If i is in F_k, then $s_{2,i} = f(i) + k(n-1)$.*

Proof. Observe that

$$\frac{s_{2,i}}{n-1} = \left\{ \frac{iax}{n-1} \right\}.$$

Further, notice that using (3.10) we have

$$\begin{aligned} axi - f(i) &= axi - (ar_i - (xv-1)p_i + xaq_i) \\ &= a(n-1)r_i + p_i(ax(v-1) + (xv-1)) \\ &= (n-1)(ar_i + p_i)\,, \end{aligned}$$

an integer multiple of $n-1$, so that

$$\left\{ \frac{iax}{n-1} \right\} = \left\{ \frac{f(i)}{n-1} \right\}.$$

It follows that

$$s_{2,i} = (n-1)\left(\frac{f(i)}{n-1} - \left\lfloor \frac{f(i)}{n-1} \right\rfloor \right) = f(i) + k(n-1)\,.$$

since $i \in F_k$. □

Theorem 3.4.4. *Suppose that $n-1$ is coprime to both x and a, and let $\ell \in \mathbb{Z}$. If $i \in F_k$ and $j \in F_{k+\ell}$, then $i \prec j$ if and only if $(p_j - p_i, q_j - q_i, r_j - r_i)$ lies in the open polyhedral cone defined by $Cx > (\ell(n-1), 0, 0)^T$, where C is the matrix*

$$C := \begin{bmatrix} xv-1 & -ax & -a \\ 1-v & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Proof. Summarizing our results so far,

1. $i < j$ if and only if $r_i < r_j$ or $r_i = r_j$ and $(v-1)p_i + q_i < (v-1)p_j + q_j$.

2. $s_{1,i} > s_{1,j}$ if and only if $r_i - r_j > x\big[(v-1)(p_j - p_i) + (q_j - q_i)\big]$. If $i < j$, then $r_i - r_j < 0$ or $r_i - r_j = 0$ and $(v-1)(p_j - p_i) + q_j - q_i > 0$. Since $-(x-1) \leq r_i - r_j \leq x - 1$, we see that for $i < j$, we have $s_{1,i} > s_{1,j}$ if and only if $r_i < r_j$ and $\big[(v-1)p_i + q_i\big] > \big[(v-1)p_j + q_j\big]$.

3. If i is in F_k and j is in $F_{k+\ell}$, then $s_{2,i} > s_{2,j}$ if and only if $f(i) > f(j) + \ell(n-1)$, i.e., if and only if $(xv-1)(p_j - p_i) - ax(q_j - q_i) + a(r_j - r_i) > \ell(n-1)$.

Notice that these conditions correspond to affine half-spaces and are simultaneously satisfied exactly exactly when $(p_j - p_i, q_j - q_i, r_j - r_i) \in \mathbb{Z}^3$ lies in the open polyhedral cone $\mathcal{C}^\circ$, defined by the matrix equation $Cx > (\ell(n-1), 0, 0)^T$. $\square$

The following proposition shows that there are a limited number of values for which F_k is non-empty, demonstrating the utility of Theorem 3.4.4.

Proposition 3.4.5. $[n-2] = F_0 \uplus F_1 \uplus F_2$

Proof. We show that $-2(n-1) < f(i) < n-1$ for every $i \in [n-2]$, from which the result follows. To prove $-2(n-1) < f(i)$, we observe the following, using equations (3.12), (3.14), and (3.13) for the first inequality and the fact that (by definition) $v \geq 2$ and $a \geq 1$ for the second inequality:

$$\begin{aligned}
f(i) + 2(n-1) &= ar_i - p_i(xv-1) + xaq_i + 2(xa(v-1) + xv - 1) \\
&= a(r_i + xq_i + 2x(v-1)) - (xv-1)(p_i - 2) \\
&\geq a(2x(v-1)) - (xv-1)(a-1) \\
&= (v-2)ax + xv + a - 1 \\
&\geq 0
\end{aligned}$$

To prove the $f(i) < n-1$, we observe that using the same inequalities as before together with (3.10) we have:

$$\begin{aligned}
n - 1 - f(i) &= n - 1 - ar_i + (xv-1)p_i - xaq_i \\
&\geq n - 1 - a(x-1) - xa(v-2) \\
&= xa(v-1) + xv - 1 - a(x-1) - xa(v-2) \\
&= a - 1 + xv \\
&\geq 0
\end{aligned}$$

$\square$

3.5 The case $v = 2$

Let $v = 2$, so that $\lambda = (x, x, ax) \in \mathbb{Z}^3$. Then $n = x(a+2)$ and $i = (a+2)r_i + p_i$, with $0 \leq p_i < a+2$, and $q_i = 0$ for all i.

Proposition 3.5.1. *For $v = 2$,*

$$r_i + r_{n-1-i} = x - 1 \quad \text{and} \quad p_i + p_{n-1-i} = a + 1.$$

Proof. Notice that

$$\begin{aligned} n &= i + (n - i - 1) + 1 \\ &= (a+2)(r_i + r_{n-1-i}) + p_i + p_{n-1-i} + 1 \\ &= (a+2)x, \end{aligned}$$

so that

$$x = (r_i + r_{n-1-i}) + \frac{p_i + p_{n-1-i} + 1}{a+2}.$$

Since x is an integer, this implies that $(p_i + p_{n-1-i}) \mod (a+2) \equiv a+1$ and $p_i + p_{n-1-i} = a + 1 + k(a+2)$. Since $0 \leq p_j < a+2$, the result follows. □

Proposition 3.5.2. *For $v = 2$,*

$$i \in F_s \quad \text{if and only if} \quad n - 1 - i \in F_{2-s}.$$

Proof. Since by definition $i \in F_s$ if and only if

$$s = -\left\lfloor \frac{ar_i - (2x-1)p_i)}{n-1} \right\rfloor,$$

the proposition is equivalent to the claim that

$$-\left\lfloor \frac{ar_i - (2x-1)p_i}{n-1} \right\rfloor - \left\lfloor \frac{ar_{n-1-i} - (2x-1)p_{n-1-i}}{n-1} \right\rfloor = 2.$$

Using the previous proposition and some tedious but straightforward algebra, we have that

$$\begin{aligned} &-\left\lfloor \frac{ar_i - (2x-1)p_i}{n-1} \right\rfloor - \left\lfloor \frac{ar_{n-1-i} - (2x-1)p_{n-1-i}}{n-1} \right\rfloor \\ &= -\left\lfloor \frac{ar_i - (2x-1)p_i}{n-1} \right\rfloor - \left\lfloor \frac{a(x-1-r_i) - (2x-1)[(a+1) - p_i]}{n-1} \right\rfloor \\ &= -\left\lfloor \frac{ar_i - (2x-1)p_i}{n-1} \right\rfloor - \left\lfloor \frac{-(n-1) - [ar_i - (2x-1p_i]}{n-1} \right\rfloor \\ &= -\left\lfloor \frac{ar_i - (2x-1)p_i}{n-1} \right\rfloor + \left\lceil \frac{(n-1) + [ar_i - (2x-1)p_i]}{n-1} \right\rceil \\ &= 1 + \left(\left\lceil \frac{ar_i - (2x-1)p_i}{n-1} \right\rceil - \left\lfloor \frac{ar_i - (2x-1)p_i}{n-1} \right\rfloor \right), \end{aligned}$$

so that unless $ar_i - (2x-1)p_i$ is a multiple of $(n-1)$, the claim holds.

Let $ar_i - (2x-1)p_i$ equal $k(n-1)$; we will show that k is not an integer. Observe that $p_i = i - (a+2)r_i$ (by definition) and that $n = x(a+2)$. We obtain

$$\begin{aligned} k(n-1) &= ar_i - (2x-1)p_i \\ &= ar_i - (2x-1)(i-(a+2)r_i) \\ &= r_i[a + (2x-1)(a+2)] - (2x-1)i \\ &= r_i[2x(a+2) - 2] - (2x-1)i \\ &= 2r_i(n-1) - (2x-1)i, \end{aligned}$$

so that

$$\begin{aligned} k &= 2r_i - \frac{(2x-1)i}{n-1} \\ &= 2r_i - i\frac{(n-1) - ax}{n-1} \\ &= 2r_i - i + \frac{iax}{n-1}. \end{aligned}$$

Since we assume that both a and x are relatively prime to $n-1$ and i is less than $n-1$, k is not an integer. $\square$

In the case $v = 2$, Theorem 3.4.4 is equivalent to $i \prec j$ if and only if $i \in F_s$, $j \in F_{s+\ell}$, and

$$\begin{bmatrix} 2x-1 & -ax & -a \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_j - p_i \\ 0 \\ r_j - r_i \end{bmatrix} = \begin{bmatrix} (2x-1)(p_j - p_i) - a(r_j - r_i) \\ p_i - p_j \\ r_j - r_i \end{bmatrix} > \begin{bmatrix} \ell(n-1) \\ 0 \\ 0 \end{bmatrix}.$$

We now demonstrate how to use the results of this chapter to construct $P(x, x, ax)$.

Example 3.5.3. *Let $a = x = 3$, so that $n-1$ is equal to 14, and note that this is relatively prime to 3. Since i is in F_0 if and only if $ar_i \geq (2x-1)p_i$, we draw the elements of $P(3,3,3\cdot 3)\backslash 0$ in the plane by*

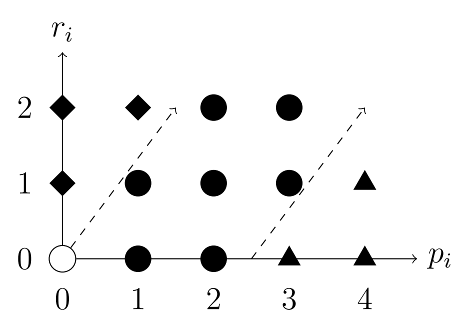


Figure 3.4: Constructing the poset $P(3,3,9)$.

where the diamonds correspond to elements of F_0 *and the triangles correspond to elements of* F_2*. If* $i \in F_s$ *and* $j \in F_{s+1}$*, then* $i \prec j$ *if and only if the point* $(p_j - p_i, r_j - r_i)$ *is among the three points in Figure 3.5. For example, we see that for* $i = 1$*,* $i \prec j$ *if and only if* $(p_j - p_i, r_j - r_i)$ *is among the points* $(-1, 1)$*,* $(-1, 2)$*, and* $(-2, 1)$*. The only suitable values of* j *are 5 and 10. The induced relations in* $P(3, 3, 9)\backslash 0$ *are depicted in Figure 3.6.*

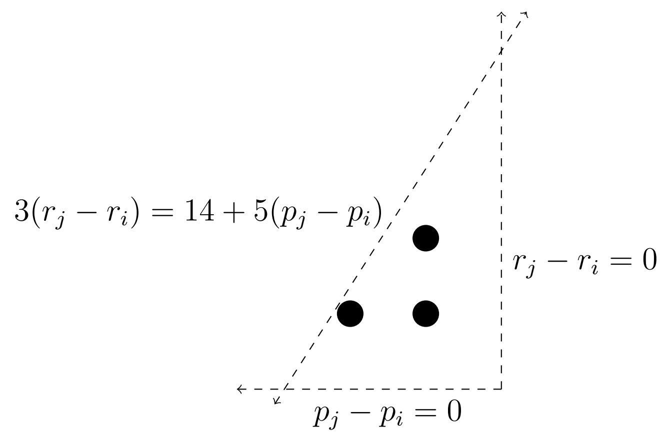


Figure 3.5: Relations in $P(3, 3, 9)$ for $\ell = -1$

The relations in Figure 3.7 contribute the relations in Figure 3.8 (some of which are redundant). This completes our construction of the poset $P(3, 3, 9)$*, depicted in Figure 3.9.*

3.6 Experimental Data

For a fixed dimension d and natural number k, we consider the collection L of all of the integer partitions $\lambda \vdash n$ with exactly d parts, all relatively prime to $n-1$, and having maximum part at most k. We define the ratio $\mathrm{anti}(d, k)$ to be be the proportion of $\lambda \in L$ for which $\Pi_\lambda\backslash 0$ is an antichain (as in Section 1.4.3). We summarize values of $\mathrm{anti}(d, k)$ computed in SAGE [23] in Table 3.6.

While Example 3.5.3 demonstrates that the posets $P(\lambda)$ can be complicated even when the simplex Δ is in low dimension ($d = 3$ in the example), Table 3.6 suggests

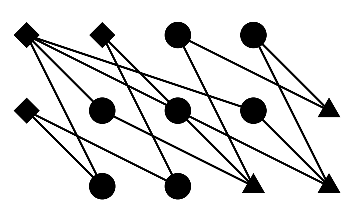


Figure 3.6: Constructing relations.

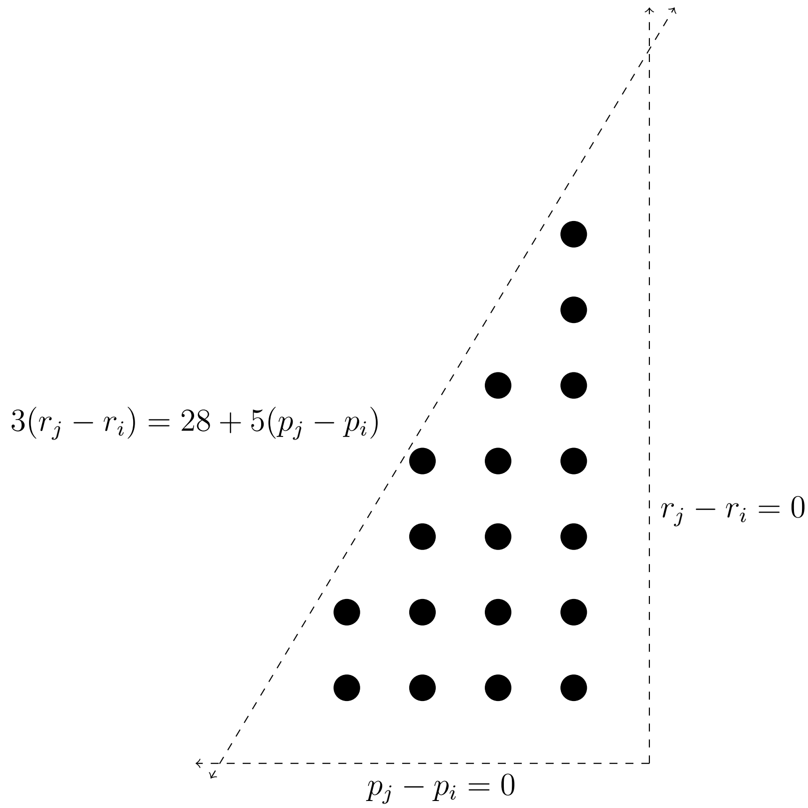


Figure 3.7: Relations in $P(3,3,9)$ for $\ell = -2$

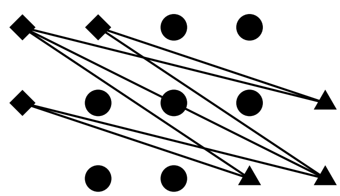


Figure 3.8: Constructing additional relations.

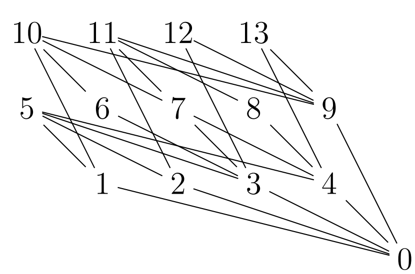


Figure 3.9: The poset $P(3,3,9)$.

Dimension d	anti$(d, \lceil 0.5d \rceil)$	anti(d, d)	anti$(d, \lceil 1.5d \rceil)$	anti$(d, 2d)$	anti$(d, \lceil 2.5d \rceil)$
2	1.0	0.5	0.33	0.2	0.14
3	0.67	0.57	0.5	0.43	0.36
4	0.67	0.56	0.53	0.54	0.49
5	0.64	0.63	0.53	0.45	0.37
6	0.75	0.61	0.52	0.42	0.39

Table 3.1: Density of antichains

that if the maximum coordinates of λ are not too big relative to d, then $P(\lambda)\backslash 0$ is probably an antichain (and thus Theorem 1.4.3 applies).

Chapter 4 Predicting the integer decomposition property via machine learning

Due to the maturity and ubiquity of machine learning techniques and applications, open–source software libraries such as **Tensorflow** have become available to non-specialists. These libraries are typically well–documented, and friendly technical references are freely available online, e.g., [13]. In this environment, it seems natural to ask: *How do we apply machine learning technology to algebraic combinatorics?*

It is not clear how to extract human–understandable meaning from the raw numerical data of, for example, a neural network (for a discussion of comprehensibility, see [24].) We therefore employ these techniques for their prediction and approximation power, rather than for use in theorems and their proofs. There is a long history of using neural networks in order to approximate solutions to combinatorial optimization problems, e.g. the traveling salesman problem [20], and in [15], Gryak, Haralick, and Kahrobaei use machine learning to predict if two elements of a group are conjugate. It seems reasonable, then, to hope that machine learning has some applicability to problems at the intersection of combinatorics and algebra.

We intend for this work to be an introduction to neural networks and a proof of concept for the use of machine learning, and neural networks in particular, in predicting properties relevant to lattice points in polyhedra. As a particular application, we attempt to predict the integer decomposition property (IDP) in a special class of lattice simplices.

In their paper [7], Braun, Davis, and Solus study the infinite family of lattice simplices of the form

$$\Delta_{(1,q)} = \operatorname{conv}\left\{e_1, \ldots, e_d, -\sum_{i=1}^{d} q_i e_i\right\} \subset \mathbb{R}^d,$$

where $q_i \in \mathbb{Z}_{\geq 0}$ for all i, and give necessary and sufficient conditions on the entries q_i (the q**-vector**) for a such a simplex to be IDP in the case that it is reflexive. In the present work we will "train" a neural network to predict if a given example of a $\Delta_{(1,q)}$ simplex is IDP without actually computing the Hilbert basis.

In Section 2 we define the integer decomposition property and interpret it as a composition of functions to be approximated. In Section 3 we develop the general framework for training a neural network using the language of piecewise linear functions[1] and stochastic gradient descent. In Section 4 we discuss a piecewise linear approximation of the integer decomposition property and its accuracy.

[1]This exposition agrees with the more common descriptions of neural networks when restricted to the case that the source of the training data is a well-defined function and we use **ReLU** activation functions.

4.1 Approximating the integer decomposition property as a function

Recall that height(z) is the zeroth coordinate of a point z in Π_Δ. We say that the simplex Δ has the **integer decomposition property** (IDP) if, for each element of the Hilbert basis of cone(Δ), the height is equal to 1. We encode the integer decomposition property as a real-valued function $\mathbb{IDP}$ in Subsection 4.1, and define what it means to approximate the integer decomposition property.

Partitioning Π

We partition Π_Δ into disjoint subsets we call **bins** B_α for α in $\{0, \ldots, d\}^{d+1}$, with $z \in B_\alpha$ if and only if

$$(\lfloor (d+1)\gamma_1 \rfloor, \ldots, \lfloor (d+1)\gamma_{d+1} \rfloor) = \alpha,$$

where the γ_i's are the coefficients of the representation of z in terms of the generators $(1, v_i)$.

Proposition 4.1.1. *Let z be an integer point in B_α. Then*

$$\text{height}(z) = \left\lceil \frac{\sum_{i=1}^{d+1} \alpha_i}{d+1} \right\rceil.$$

Proof. Considering the zeroth coordinate of z, it is clear that

$$\text{height}(z) = \sum_{i=1}^{d+1} \gamma_i.$$

Note that since $\alpha = (\lfloor (d+1)\gamma_1 \rfloor, \ldots, \lfloor (d+1)\gamma_{d+1} \rfloor)$, the inequality

$$(d+1)\gamma_i - 1 < \lfloor (d+1)\gamma_i \rfloor \leq (d+1)\gamma_i$$

implies that

$$\sum_{i=1}^{d+1} \big((d+1)\gamma_i - 1\big) \; < \; \sum_{i=1}^{d+1} \alpha_i \; \leq \sum_{i=1}^{d+1} (d+1)\gamma_i.$$

Thus

$$(d+1)\big(\text{height}(z) - 1\big) < \sum_{i=1}^{d+1} \alpha_i \; \leq (d+1)\,\text{height}(z),$$

and

$$\text{height}(z) - 1 < \frac{\sum_{i=1}^{d+1} \alpha_i}{d+1} \leq \text{height}(z),$$

from which the result follows. □

This leads to the following characterization of the integer decomposition property:

Corollary 4.1.2. *The simplex Δ is IDP if and only if for each z in the Hilbert basis of* cone(Δ), *$z \in B_\alpha$ implies that $\sum_{i=1}^{d+1} \alpha_i$ is at most $d+1$.*

The function $\mathbb{IDP}$

Definition 4.1.3. $\mathbb{IDP}$ *is the* $0/1$ *function which takes as input the vertices of a lattice* d*-simplex* Δ *and returns one if* Δ *is IDP and zero otherwise.*

We were unsuccessful in approximating $\mathbb{IDP}$ directly using techniques presented in this paper. Instead, we find success approximating another function, $\mathbb{HB}$, from which the value of $\mathbb{IDP}$ can be inferred.

For a vector $x \in \mathbb{R}^{(d+1)^{d+1}}$, we consider x to be multi-indexed by the collection of $\alpha \in \{0, \ldots, d\}^{d+1}$.

Definition 4.1.4. $\mathbb{HB}$ *is the function taking as input the vertices of a lattice* d*-simplex and returning an element of* $\{0,1\}^{(d+1)^{d+1}}$, *with coordinate* α *equal to one if and only if there exists a Hilbert basis element in bin* B_α.

Example 4.1.5. *Consider the* $\Delta_{(1,q)}$ *simplex in dimension* $d = 2$ *with* q*-vector* $(2,1)$. *The ray generators are* $v_1 = (1,0,1)$, $v_2 = (0,1,1)$, *and* $v_3 = (-2,-1,1)$. *Computation with* ***Normaliz*** *[8] yields that the set of lattice points in* $\Pi_{\Delta_{(1,g)}}$ *is equal to* $\{(0,0,0),\ (-1,0,1),\ (0,0,1)\}$. *The representation of these points in terms of the ray generators are* $(0,0,0)$, $(0, 1/2, 1/2)$, *and* $(1/2, 1/4, 1/4)$. *The bins* B_α *which contain these points have* α *equal to:*

$$\begin{aligned}
&(\lfloor 3\cdot 0\rfloor, \lfloor 3\cdot 0\rfloor, \lfloor 3\cdot 0\rfloor) = (0,0,0),\\
&(\lfloor 3\cdot 0\rfloor, \lfloor 3\cdot 1/2\rfloor, \lfloor 3\cdot 1/2\rfloor) = (0,1,1), \text{ and}\\
&(\lfloor 3\cdot 1/2\rfloor, \lfloor 3\cdot 1/4\rfloor, \lfloor 3\cdot 1/4\rfloor) = (1,0,0),
\end{aligned}$$

respectively.

When the α *are lexicographically ordered, we may write the image under* $\mathbb{HB}$ *as the vector*

$$(1,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0) \in \mathbb{R}^{3^3}.$$

Let **supp** be the $0/1$ function on $\mathbb{R}^{(d+1)^{d+1}}$ which, for a vector x, returns zero if and only if there exists an index α such that $x_\alpha \neq 0$ and $\sum_{i=1}^{d+1} \alpha_i > d+1$. Then, using Corollary 4.1.2, we may write the functional equality

$$\mathbb{IDP} = \mathbf{supp} \circ \mathbb{HB}.$$

Note that for Example 4.1.5, the non-zero entries are at multi-indices $(0,0,0)$, $(0,1,1)$ and $(1,0,0)$, and that the sum of each individual multi-index is not more than 3. Thus the image of $\mathbb{IDP}$ is equal to 1, indicating that the example is IDP. We can verify this fact by noting that the height of each lement of the Hilbert basis

$$\{v_1, v_2, v_3\} \bigcup \{(-1,0,1), (0,0,1)\}$$

is equal to 1. We remark that it is not true in general that the Hilbert basis elements are the non-zero lattice points of Π_Δ.

We have developed a theoretical framework for approximating the integer decomposition property by approximating the real-valued function $\mathbb{HB}$. One difficulty in the implementation of this scheme is the fact that **supp** is not sensitive to how close to zero a value is. If the entry at some multi-index α in the approximation of $\mathbb{HB}$ is close to but not equal to zero, and $\sum_{i=1}^{d+1} \alpha_i > d+1$, then the image of **supp** will be 0, i.e., our approximation of $\mathbb{IDP}$ will almost always predict that an example is *not* IDP. A standard solution to this issue is to first map our approximation into the open interval $(0,1)$, then choose a value $0 \le \eta \le 1$, then interpret values less than or equal to η as 0 and greater than η as 1. For the the first step, we use the Sigmoid function:

$$\sigma(x) := (1+e^{-x})^{-1},$$

mapping $\mathbb{R}$ one-to-one onto the open interval $(0,1)$. For some fixed $0 \le \eta \le 1$, define

$$\mathbf{cutoff}(x) = \begin{cases} 0 & \text{if } x \le \eta, \text{ and} \\ 1 & \text{otherwise.} \end{cases}$$

The composition of **cutoff** and σ allows us to turn any real-valued function of one variable into a 0/1 function, and by applying it coordinate-wise, we may turn any function $f : \mathbb{R}^u \longrightarrow \mathbb{R}^v$ into $\mathbf{cutoff} \circ \sigma \circ f : \mathbb{R}^u \longrightarrow \{0,1\}^v$. In particular, consider $\mathbb{R}^u$ to be the space parameterizing the vertex sets of lattice d-simplices: $\mathbb{R}^u = \mathbb{R}^d \times \cdots \times \mathbb{R}^d$ with $d+1$ factors $\mathbb{R}^d$ (not all points in the space give rise to full-dimensional simplices.) Further consider $\mathbb{R}^v$ to have basis multi-indexed by $\alpha \in \{0,\ldots,d\}^{d+1}$. Then for any map f from $\mathbb{R}^u$ to $\mathbb{R}^v$, $\mathbf{cutoff} \circ \sigma \circ f$ may be considered as a map from lattice d-simplices to 0/1 vectors indexed by bins B_α.

Continuing Example 4.1.5, consider the function $f : \mathbb{R}^{2\times 3} \longrightarrow \mathbb{R}^{3^3}$ defined by

$$f(\boldsymbol{x}) = \frac{1}{\|\boldsymbol{x}\|}\left(\sum_{i=1}^{27}(-1)^i \cdot e_i\right).$$

Then $f(1,0,0,1,-2,-1) = \frac{1}{\sqrt{7}}\left(\sum_{i=1}^{27}(-1)^i \cdot e_i\right)$. We compute that $\sigma(1/\sqrt{7}) = 0.593$, and that $\sigma(-1/\sqrt{7}) = 0.407$. Thus if:

- $0 \le \eta < 0.407$, then $\mathbf{cutoff} \circ \sigma \circ f(1,0,0,1,-2,-1)$ is the all-ones vector,
- if $0.593 \le \eta \le 1$ then $\mathbf{cutoff} \circ \sigma \circ f(1,0,0,1,-2,-1)$ is the zero vector, and
- if $0.407 \le \eta < 0.593$, then $\mathbf{cutoff} \circ \sigma \circ f(1,0,0,1,-2,-1)$ is the 0/1-vector $\sum_{i=1}^{27} \frac{(1+(-1)^i)}{2} \cdot e_i$.

As this example demonstrates, the quality of the approximation depends heavily on the choice of value for η, as for the fixed function f, $\mathbf{cutoff} \circ \sigma \circ f$ can be correct on 11%, 89%, or 33% of the entries of $\mathbb{HB}$, depending on the choice of η.

Definition 4.1.6. *Let f be any function from $\mathbb{R}^{d(d+1)}$ to $\mathbb{R}^{(d+1)^{d+1}}$. Then we call the (coordinate-wise) composite function*

$$\widehat{\mathbb{IDP}} := \boldsymbol{supp} \circ \boldsymbol{cutoff} \circ \sigma \circ f$$

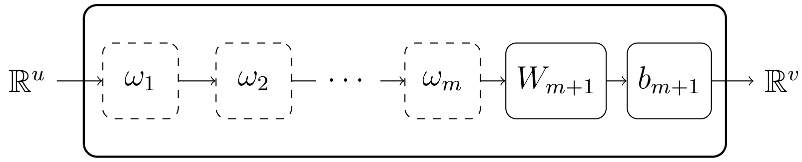


Figure 4.1: The approximation $\widehat{f}$.

an approximation of the integer decomposition property.

Note that when $\sigma \circ f$ closely approximates $\mathbb{HB}$ coordinate-wise, $\widehat{\mathbb{IDP}}$ agrees with $\mathbb{IDP}$. For a given f, we will use the shorthand notation $\widehat{\mathbb{HB}}$ for the composite function **cutoff** $\circ\ \sigma\ \circ\ f$.

4.2 A general approximation method

In this section we describe piecewise linear functions as compositions of affine transformations and a well-behaved piecewise linear function ρ. We next describe the use of a loss function L in quantifying the accuracy of an approximation $\widehat{f}$ of a function f. We then describe an algorithm called gradient descent, which deforms the piecewise linear function $\widehat{f}$ in order to minimize the loss function L with respect to the target function f.

Let f be any set map from $\mathbb{R}^u$ to $\mathbb{R}^v$. We will approximate f by constructing a *random* initial "approximation" $\widehat{f}$, which we will deform until we have a sufficiently accurate approximation.

For a positive integer m, fix m positive integers ℓ_1 through ℓ_m, as well as a small real $\epsilon > 0$. We will call this the collection of **hyper-parameters**. Choose matrices $W_k \in \mathbb{R}^{\ell_{k-1} \times \ell_k}$ for $1 \leq k \leq m+1$, where we set $\ell_0 = u$ and $\ell_{m+1} = v$ (the dimensions of the domain and codomain of f.) Additionally, for each k, choose vectors $b_k \in \mathbb{R}^{\ell_k}$. The entries $(W_k)_{i,j}$ are called **weights**, and the $(b_k)_i$ are called **biases**. Generally, the initial values are randomized by an algorithm we will not discuss here. We will consider each such collection of **parameters** to be a point

$$p = \big(W_1, b_1, \ldots, W_{m+1}, b_{m+1}\big)$$

in the space of parameters $\mathbb{R}^{(\ell_0+1)\times\ell_1} \times \cdots \times \mathbb{R}^{(\ell_m+1)\times\ell_{m+1}}$. We define ρ to be the function which returns the coordinate-wise maximum of 0 and the identity, i.e., $\rho(x_i) = \max(0, x_i)$. The map ρ is an example of an **activation function** and is called **ReLU** (Rectified Linear Unit.) Let ω_k be the affine map $x \mapsto W_k(x) + b_k$ composed with ρ. Then the approximation $\widehat{f}$ is the function

$$\widehat{f}(x;p) = W_{m+1} \circ\ \omega_m \circ \cdots \circ\ \omega_1(x) + b_{m+1}.$$

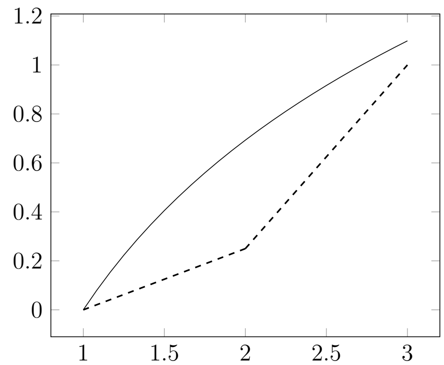


Figure 4.2: The function f and the approximation $\widehat{f}$ (dashed).

Example 4.2.1. *Let* $f(x) = \log(x)$*. We will approximate* f *on the interval* $[1, 3]$*. Let* $m = 1$ *and* $\ell_1 = 2$*. We initially set the parameters* $p = (\omega_1, b_1, \omega_2, b_2) \in \mathbb{R}^{(1+1)\times 2} \times \mathbb{R}^{(2+1)\times 1}$ *by*

$$W_1 = [0.75, -0.5]^T \quad b_1 = [-0.75, 1] \quad W_2 = [1, 1] \quad b_2 = [-0.5].$$

The resulting approximation, which we expect to be poor because it knows nothing about the function it is supposed to approximate, is given by the piecewise linear function (the dotted graph in Figure 4.2)

$$\begin{aligned}
\widehat{f}(x; p) &= [1, 1]\rho\left(\begin{bmatrix} 0.75 \\ -0.5 \end{bmatrix} [x] + \begin{bmatrix} -0.75 \\ 1 \end{bmatrix}\right) + [-0.5] \\
&= 1 \cdot \rho(0.75x - 0.75) + 1 \cdot \rho(-0.5x + 1) - 0.5 \\
&= \begin{cases} 0.25x - 0.25 & 1 \le x \le 2 \\ 0.75x - 1.25 & 2 < x \le 3. \end{cases}
\end{aligned}$$

Loss functions and gradient descent

We measure the quality of the approximation via a **loss function** $L(x; p)$ which we attempt to minimize. By minimizing its value at many "training" points x distributed throughout the domain, we hope that the value of the approximation $\widehat{f}$ will be close to that of f at points outside of training set, i.e., that the magnitude of the loss function will be small at new points as well.

One example of a loss function is the Euclidean distance

$$D(x; p) = \left\| f(x) - \widehat{f}(x, p) \right\|.$$

Continuing Example 4.2.1,

$$D(x; p) = \left\| \log(x) - \big((W_2)_{1,1}\rho\big((W_1)_{1,1}(x) + b_{1,1}\big) + (W_2)_{1,2}\rho\big((W_1)_{2,1}(x) + b_{1,2}\big) + b_{2,1}\big) \right\|,$$

and for our specific parameters p,

$$D(x;p) = \Big\|\log(x) - \big(1\cdot\rho\big(0.75x - 0.75\big) + 1\cdot\rho\big(0.75x + 1\big) - 0.5\big)\Big\|\,.$$

Although for fixed parameters p, the loss $L(x;p)$ is a function of x, the "learning" step of machine learning happens by interpreting it as a function of p, holding x fixed. We can imagine L as a surface above the parameterization space which is fixed by the choice of hyper-parameters and x. In order to improve our approximation $\widehat{f}$ at a particular point x in the domain, we modify its parameters in such a way that that the value of the loss function L is reduced, i.e., "moving downhill" on the surface L.

We compute the gradient ∇L with respect to the parameters p at the point (x,p) and update the parameters by $p \mapsto p - \epsilon\nabla L$. The value of ϵ is chosen small enough that $L(x;p-\epsilon\nabla L) < L(x;p)$. When we repeatedly apply this process for points x sampled uniformly at random, this method is called **stochastic gradient descent** or SGD. In practice, for reasons of computational efficiency and stability, a batch of points are sampled and the mean of the gradients is used for the update. This is known as **mini batch** SGD.

Continuing our example, fix $x = 1.5$ and use the chain rule to compute that

$$\begin{aligned}\nabla D(1.5;p) &= \left\langle \frac{\partial D}{\partial \omega_1}, \frac{\partial D}{\partial b_1}, \frac{\partial D}{\partial \omega_2}, \frac{\partial D}{\partial b_2} \right\rangle_{x=1.5} \\ &= \langle -1.5\,,\,-1.5\,,\,-1\,,\,-1\,,\,-0.375\,,\,-0.25\,,\,-1\rangle\,.\end{aligned}$$

Then for $\epsilon = 0.02$, the update $p' = p - \epsilon\,\nabla D(1.5;p)$ is given by

$$\omega_1 = [0.78\,,\,-0.47]^T \quad b_1 = [-0.73\,,\,1.02] \quad \omega_2 = [1.0075\,,\,1.0075] \quad b_2 = [-0.48].$$

The resulting updated approximation is

$$\widehat{f}(x;p') = \begin{cases} 0.312x - 0.187 & 1 \le x \le 2.17 \\ 0.786x - 1.215 & 2.17 < x \le 3 \end{cases}.$$

Training and Validation

In practice, we perform the update step many thousands of times at x-values distributed throughout the domain. Often we gather a large collection of pairs $(x, f(x))$ called a **training set** to store for later use in the update process, rather than computing the value of f when needed. When generating this collection is costly, as in the case of the function $\mathbb{IDP}$, we use each pair from the collection multiple times over, in some cases as many as 100 times. By analogy with polynomial approximation, where we fit a polynomial to a finite set of points on the graph of a function, one may wonder if, when reusing sampled points in refining our approximation, we are simultaneously losing accuracy at other points in the domain. The short answer is yes.

This phenomenon of **overfitting** is a principal concern in the process of refining our approximation, and there are some standard techniques for mitigating its effect, including:

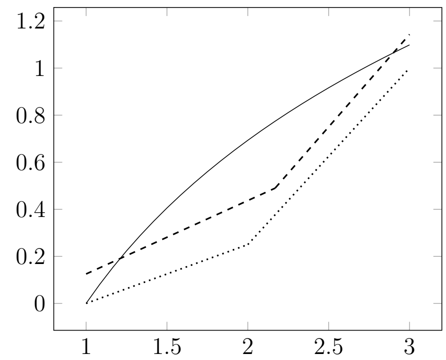


Figure 4.3: The approximations $\widehat{f}(x;p)$ (dotted) and $\widehat{f}(x;p')$ (dashed).

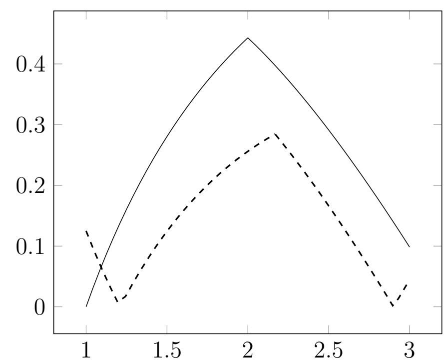


Figure 4.4: The loss function D for $\widehat{f}(x;p)$ and $\widehat{f}(x;p')$ (dashed).

- creating two collections of pairs $(x, f(x))$ — one for training and one for validation. As we train $\widehat{f}$ we simultaneously monitor its accuracy on the **validation set**. If the performance on the validation set worsens while improving on the training set, we stop training.
- introducing a component to the loss function for the magnitudes of the parameters. Experience shows that this method, called **regularization**, reduces overfitting to the training data.
- using the simplest "structure" possible to achieve the desired performance. Complicated models require more training to achieve their optimal performance, and hence increase the number of times training data is reused. We balance the expressive capability of a complicated approximation with the need to minimize overfitting.

m	ℓ_1	ℓ_2	ℓ_3	ℓ_4	ϵ
4	100	400	800	3,000	0.001

Table 4.1: Hyperparameters

4.3 Implementation and Results

Implementation

Our first goal is an approximation of the function $\mathbb{HB}$ restricted to the vertex sets of $\Delta_{(1,q)}$ simplices of dimension $d = 4$ and with q-vector entries bounded by 25. Recall also that, even though the target space of $\mathbb{HB}$ has dimension $(d+1)^{d+1}$, the relevant values are those at indices α whose coordinates sum to more than $d+1$. We restrict to these 2,877 relevant indices. Hence the input to our function is the tuple $(q_1, q_2, q_3, q_4) \in [1, 25]^4$ and the output is in $\mathbb{R}^{2,877}$.

There is no general-purpose best design of hyper-parameters that works for every application of a neural network. In fact, it is possible to approximate with arbitrary accuracy any continuous function on a compact subset of $\mathbb{R}^u$ using only one "hidden layer" ($m = 1$.) The general rule is that higher values of m allow smaller values for the ℓ_i's while maintaining approximation flexibility. Optimizing hyper-parameters is a process that is outside the scope of this work, so we will simply report that, after experimenting with several values of m and ℓ_i's in order to minimize the loss function and computation time, we proceeded using the following choice of hyper-parameters:

Tensorflow produces a neural network with the specified dimensions and initializes the weights and biases automatically. In order to implement mini batch SGD, the user must make more decisions than just specifying the hyper-parameters:

1. Amount of training/validation data:
   We used Normaliz and a script to compute $\mathbb{HB}$ for a sample of size 50,000, 10% of which we reserved for validation.

2. Batch size:
   During training we computed the gradient ∇L for batches of 10 q-vectors at a time and used the mean for the update of the parameters p.

3. Loss function:
   Because the image of $\mathbb{HB}$ is contained by the set $\{0,1\}^{2,877}$, we may consider the approximation to be the composite $\sigma\widehat{f}$ and use the **Binary Cross Entropy** loss function BCE summed entry-wise over

   $$BCE = (\mathbb{HB} - 1) \cdot \log\left(1 - \sigma\widehat{f}\right) - \mathbb{HB} \cdot \log \sigma\widehat{f}.$$

   When the value of $\mathbb{HB}$ is one, the value of BCE is decreased by increasing the value of $\log \sigma\widehat{f}$, i.e., increasing the value of $\widehat{f}$. In this case, minimizing BCE

	PREDICTED 0	PREDICTED 1
ACTUAL 0	2,808	55
ACTUAL 1	0	14

Table 4.2: The confusion table for $\widehat{\mathbb{HB}}(4, 10, 14, 14)$

coincides with minimizing the difference between $\sigma\widehat{f}$ and $\mathbb{HB}$. A similar analysis for the case when $\mathbb{HB}$ equals zero shows that BCE is a measure of the accuracy of $\sigma\widehat{f}$ as an approximation of $\mathbb{HB}$. We used a modification of BCE, which we discuss in Section 4.3

4. Training length:
   We performed roughly 100,000 updates in the process of training the approximation.

The result of this training procedure was a piecewise linear function f. It was the well-defined and deterministic[2] result of the specific choices outlined above.

An approximation $\widehat{\mathbb{HB}}$ requires a choice of cutoff parameter η, and $\widehat{\mathbb{IDP}}$ requires the additional choice of a tolerance parameter τ (introduced in Subsection 4.3). These parameters control the functions **cutoff** and **supp**, respectively. Recall that the resulting approximations are given by

$$\widehat{\mathbb{HB}} := \textbf{cutoff} \circ \sigma \circ f \qquad \text{and} \qquad \widehat{\mathbb{IDP}} := \textbf{supp} \circ \widehat{\mathbb{HB}}$$

We present the results in terms of the values η and τ.

The approximation $\widehat{\mathbb{HB}}$

While it is tempting to present the accuracy of $\widehat{\mathbb{HB}}$ as the percentage of indices on which it agrees with $\mathbb{HB}$, this is problematic due to the scarcity of non-zero entries in any given image of $\mathbb{HB}$. Consider the q-vector $(4, 10, 14, 14)$; there are just 14 non-zero entries among the 2,877 relevant entries in its image under $\mathbb{HB}$. Consequently, an approximation which is uniformly equal to zero would be correct 99.5% of the time, while knowing essentially nothing about the function it is trying to approximate other than that it is typically equal to zero! We therefore present the accuracy in the form of a **confusion table**, which breaks down the indices α along two criteria — firstly depending on whether $\mathbb{HB}_\alpha$ is equal to 1 (positive) or 0 (negative), and secondly whether $\widehat{\mathbb{HB}}_\alpha$ is positive or negative.

Example 4.3.1. *Again using the q-vector* $(4, 10, 14, 14)$*, we set* $\eta = 0.1$ *and present the resulting confusion table below:*

[2]For purposes of analysis and reproducibility, we initialize the computer's randomness generator so that the stochastic processes are, in fact, deterministic, while still having good randomness properties.

η	specificity	sensitivity
0.1	0.981	1.00
0.25	0.986	0.857
0.5	0.993	0.214

Table 4.3: The effect of varying η

	PREDICTED 0	PREDICTED 1
ACTUAL 0	12,726,675	1,573,167
ACTUAL 1	22,569	88,482

Table 4.4: An aggregated confusion table for $\mathcal{S}$ ($\eta = 0.1$)

*Observe that the sum of the table entries is, in fact, 2,877. We call entries appearing in the upper right cell of the table "**false positive**" because the approximation incorrectly predicted that a bin contained a Hilbert basis element. Similarly, entries in the bottom left cell are called "**false negative**".*

We may summarize the table with the pair of ratios

$$\textit{specificity} = \frac{\textit{true negatives}}{\textit{true negatives} + \textit{false positives}}, \quad \textit{and}$$

$$\textit{sensitivity} = \frac{\textit{true positives}}{\textit{true positives} + \textit{false negatives}}.$$

For the present example, they are 98% and 100%, respectively. The specificity and sensitivity vary with the cutoff value η, and are negatively correlated with each other, as demonstrated in Table 4.3.1:

Validation

When we sampled 50,000 examples for training, we reserved 5,000 of them for validation purposes. We now report the performance on this validation set, which we denote $\mathcal{S}$. We aggregate (sum entry-wise) the confusion tables for $\eta = 0.1$ in Table 4.3.

The corresponding aggregated specificity is 89.0%, and sensitivity is 79.7%. One can account for the difference between specificity and sensitivity by recalling the scarcity of non-zero entries of $\mathbb{HB}$, i.e., the low total number of positives. If we use the loss function

$$BCE = (\mathbb{HB} - 1) \cdot \log \left(1 - \sigma \widehat{f}\right) - \mathbb{HB} \cdot \log \sigma \widehat{f}$$

as earlier described for our gradient descent, the resulting approximation will essentially be the constant zero function. In order for the model to learn to identify positives, we must balance the contributions to the loss function associated to positive and negative according to the inverse of their frequency. We accomplish this by

τ \ η	0.5	0.25	0.12	0.05
0	3/7 (42.9%)	3/4 (75.0%)	3/3 (100.0%)	3/3 (100.0%)
10	21/320 (6.6%)	11/38 (29.0%)	8/21 (38.1%)	6/12 (50.0%)
20	46/1026 (4.5%)	21/102 (20.6%)	11/45 (24.4%)	8/27 (29.6%)
30	65/1770 (3.7%)	35/196 (17.9%)	23/103 (22.3%)	16/64 (25.0%)

Table 4.5: The rate of true positives (specificity) for given values of η and τ

introducing a positive term β which we call the **balance** term:

$$L = (\mathbb{HB} - 1) \cdot \log \left(1 - \sigma \widehat{f}\right) - \beta \cdot \mathbb{HB} \cdot \log \sigma \widehat{f}.$$

The results presented in this section correspond to a β value of 10. All other parameters remaining fixed, a higher value, roughly $\beta = 75$, is required in order achieve approximately equal sensitivity and specificity. However, it is not necessarily desirable to match the sensitivity and specificity, as we will discuss.

The approximation $\widehat{\mathbb{IDP}}$

Under the unrealistic assumption that Hilbert basis elements are distributed roughly uniformly among bins, consider an approximation with a specificity of 99.9% applied to the q-vector of an IDP $\Delta_{(1,q)}$ simplex. Because there are 2,877 bins, the probability that all bins will be correctly identified as negative (not containing a Hilbert basis element) can be estimated as $0.999^{2,877} \approx 5.6\%$. Since we expect the incidence of the integer decomposition property to be low, a true positive rate for IDP of 5.6% may result in few or even no examples being correctly predicted as IDP! We have several tools to combat this issue:

1. manipulating the balance term β to produce high specificity (possibly at the expense of sensitivity)
2. manipulating the cutoff value η to produce high specificity (again, at the expense of sensitivity)
3. tolerating some number of positive entries in $\widehat{\mathbb{HB}}$ (under the assumption that many of them are false.)

For this last option we introduce the **tolerance** parameter τ, which sets an upper bound on the number of positive entries before the function $\widehat{\mathbb{IDP}}$ returns that an example is IDP negative. In our original description of $\widehat{\mathbb{IDP}}$, τ was implicitly set to zero.

Table 4.3 records the number of true positives over the total number of positives of $\widehat{\mathbb{IDP}}$ when applied to the sample $\mathcal{S}$ for select values of η and τ.

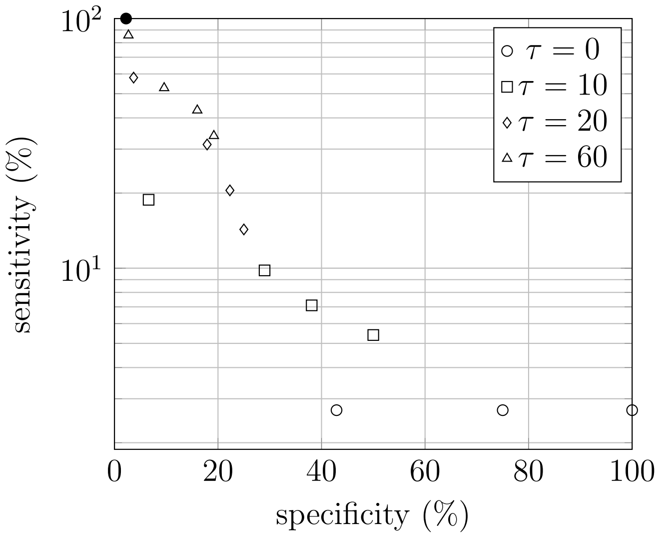


Figure 4.5: The effect of varying η for fixed values of τ.

From Table 4.3, we see that there is not one optimal choice for the values of η and τ, since higher specificity is correlated with few examples being found; the goals of specificity and sensitivity are in tension. Figure 4.5 shows the (log-scale) relationship between specificity and sensitivity induced by varying these values. When we actually checked for IDP using Normaliz, we found 112 positive examples among 5,000. The analogous "specificity" is 2.24%, but the "sensitivity" is 100% — we plot this point $(2.24, 100)$ for reference.

Table 4.3 lists all 112 q-vectors of $\mathcal{S}$ that correspond to IDP $\Delta_{(1,q)}$ simplices according to Normaliz. Recall that the rate of IDP in $\mathcal{S}$ is 2.24%. Table 4.3 lists the subset of $\mathcal{S}$ which are predicted to be IDP when $\eta = 0.5$ and $\tau = 0$, with the correct positive predictions highlighted. Observe that the incidence of IDP among the *predicted* IDP examples is about 43%, much higher than the rate in the sample at large.

We highlight the q-vectors in Table 4.3 that correspond to true IDP positive predictions made by setting $\eta = 0.1$ and $\tau = 65$ (the specificity was 15% and the sensitivity was 58%.)

Discussion

As a demonstration of the utility of the approximation method presented here, we could attempt to advance the previously mentioned work in [4] on $\Delta_{(1,q)}$ simplices by producing a large and diverse collection of IDP examples from which to form conjectures to try to prove. A natural scheme for arriving at such a collection is to first generate a test set, say, all $\Delta_{(1,q)}$ simplices of dimension d with q-vector entries bounded by n, then verify the integer decomposition property with a program like Normaliz, collecting the positive examples. We could augment this scheme with

1,1,1,1	1,1,3,9	1,1,21,24	1,2,14,10	1,2,14,10
1,3,16,3	1,3,24,1	1,4,2,16	1,4,20,20	1,8,1,1
1,10,10,8	1,10,24,24	1,12,4,12	1,15,3,1	1,18,1,6
1,21,1,4	1,24,1,9	1,24,14,2	1,24,17,1	1,24,18,1
1,24,18,4	1,24,24,20	2,2,2,7	2,3,12,18	2,8,8,4
2,10,1,16	2,20,10,5	3,1,1,9	3,6,12,1	3,12,2,24
3,14,21,3	3,19,3,1	3,23,15,3	4,1,1,4	4,8,2,16
4,20,1,14	4,20,10,20	4,23,4,12	4,24,1,16	6,1,2,12
6,2,6,3	6,2,18,9	6,6,6,3	6,14,6,15	6,17,9,18
7,3,21,7	7,7,1,7	7,7,16,16	8,1,8,2	8,2,12,24
8,16,4,2	9,1,1,9	9,6,18,2	9,9,4,4	9,18,4,4
9,18,18,6	9,22,1,11	10,1,5,22	10,5,10,9	10,24,4,1
11,22,5,5	12,1,2,6	12,1,24,19	12,2,3,12	12,2,18,3
12,3,2,6	12,3,11,6	12,6,1,1	12,6,1,3	12,12,4,12
12,16,1,16	12,24,2,24	12,24,6,1	13,2,2,20	14,6,14,7
14,7,2,24	14,7,12,1	15,1,13,15	15,15,1,1	16,1,6,6
16,4,2,16	16,7,16,16	16,8,4,2	16,16,12,3	16,24,1,22
17,1,7,1	17,17,8,4	17,17,17,1	18,1,1,15	18,2,6,6
18,2,22,1	18,10,1,15	19,19,1,16	20,2,1,12	20,8,19,8
20,14,24,1	20,20,1,20	20,20,4,1	20,20,4,20	20,22,1,22
21,21,16,4	22,2,2,22	22,16,4,1	22,16,22,1	22,22,20,1
23,2,2,6	23,18,3,24	23,24,24,12	24,2,1,16	24,4,2,4
24,24,6,24	24,24,23,12			

Table 4.6: The 112 IDP examples in the sample $\mathcal{S}$

1,1,1,1	1,2,10,2	1,3,24,1	2,2,2,7	2,3,4,7
4,3,2,5	11,6,9,6			

Table 4.7: Predicted IDP examples ($\eta = 0.5$, $\tau = 0$)

machine learning by performing an initial sieving step prior to testing with Normaliz. By developing a computationally–cheap approximation to the integer decomposition property, we can reserve the relatively expensive Normaliz computations for those examples that, according to the approximation, are more likely to be IDP.

In the context of this application, the results outlined above point to a tradeoff between the computational efficiency (controlled by the specificity) and the number of examples that are ultimately produced (controlled by the sensitivity). It also seems that the approximation $\widehat{\mathbb{IDP}}$ is biased in favor of repeated entries (see the highlighted examples in Table 4.3,) which brings into question how diverse a set of examples it is capable of producing.

We computed the value of $\widehat{\mathbb{IDP}}$ for all 390,625 $\Delta_{(1,q)}$ simplices with q-vector in $[1,25]^4$ using $\eta = 0.1$ and $\tau = 65$. The computation produced 2,520 predicted positives. We then computed $\mathbb{IDP}$ for these examples and found that 521 were IDP. This corresponds to a specificity of 20.7%. It is impractical to compute $\mathbb{IDP}$ over the entire collection of 390,625 examples in order to compute the sensitivity, so it is not known.

4.4 Concluding Remarks

It is very likely that other choices of hyper-parameters, or even entirely different machine learning techniques, will yield improved performance. However, the results, such as they are, do indicate that functions like $\mathbb{IDP}$ have the potential to be modeled by machine learning techniques. The following remarks point out directions in which this investigation might be continued.

Remark 4.4.1. *Figure 4.5 shows the tradeoff between specificity and sensitivity for an approximation $\widehat{\mathbb{IDP}}$ that is a product of a choice of hyper-parameters, balance β, and training size. It would be useful to see the effect of different values of β in the plot. Does there exist a choice which achieves sensitivity and specificity of 50%?*

Remark 4.4.2. *The intermediate step of computing an approximation of $\mathbb{HB}$ has several potential applications which are not explicitly discussed in this paper. In particular we note that by computing the set of lattice points in each predicted–positive bin, we have an approximation of the Hilbert basis itself.*

If the sensitivity of $\widehat{\mathbb{HB}}$ is high then it is very likely that the Hilbert basis is contained *by the approximated Hilbert basis, and may be recovered by the reduction algorithm used by Normaliz (implemented by a python script, for example.) This could potentially be more efficient than Normaliz, which reduces the entire fundamental parallelepiped, if the specificity is high.*

Remark 4.4.3. *The Ehrhart h^*-vector records the number of lattice points at each height in Π_Δ. In the case that the h^*-vector is the concatenation of two vectors — the first increasing and the second decreasing — we call it* ***unimodal***. *Unimodality is another interesting property to investigate for lattice simplices, see, e.g., [5]. If, rather than recording the presence of a Hilbert basis element, we were to record the number of fundamental parallelepiped points in each bin, we could approximate the h^*-vector using Proposition 4.1.1. Thus we have a framework for predicting both IDP* and *unimodality.*

Bibliography

[1] David Anick. Construction d'espaces de lacets et d'anneaux locaux à séries de Poincaré-Betti non rationnelles. *C. R. Acad. Sci. Paris Sér. A-B*, 290(16):A729–A732, 1980.

[2] Luchezar L. Avramov. Infinite free resolutions. In *Six lectures on commutative algebra*, Mod. Birkhäuser Class., pages 1–118. Birkhäuser Verlag, Basel, 2010.

[3] Alexander Berglund, Jonah Blasiak, and Patricia Hersh. Combinatorics of multigraded Poincaré series for monomial rings. *J. Algebra*, 308(1):73–90, 2007.

[4] B. Braun, R. Davis, and L. Solus. Detecting the Integer Decomposition Property and Ehrhart Unimodality in Reflexive Simplices. *ArXiv e-prints*, August 2016.

[5] Benjamin Braun. Unimodality problems in Ehrhart theory. In *Recent trends in combinatorics*, volume 159 of *IMA Vol. Math. Appl.*, pages 687–711. Springer, [Cham], 2016.

[6] Benjamin Braun and Robert Davis. Ehrhart series, unimodality, and integrally closed reflexive polytopes. *Ann. Comb.*, 20(4):705–717, 2016.

[7] Benjamin Braun, Robert Davis, and Liam Solus. Detecting the integer decomposition property and ehrhart unimodality in reflexive simplices. to appear in *Advances in Applied Mathematics*, preprint at http://arxiv.org/abs/1608.01614.

[8] W. Bruns, B. Ichim, T. R¨omer, R. Sieg, and C. Söger. Normaliz. algorithms for rational cones and affine monoids. Available at https://www.normaliz.uni-osnabrueck.de.

[9] Winfried Bruns, Bogdan Ichim, and Christof Söger. The power of pyramid decomposition in normaliz. *arXiv preprint arXiv:1206.1916*, 2012.

[10] Eugene Ehrhart. *Polynômes arithmétiques et méthode des polyedres en combinatoire*, volume 35. Birkhauser, 1977.

[11] Juan Elias and Giuseppe Valla. A family of local rings with rational Poincaré series. *Proc. Amer. Math. Soc.*, 137(4):1175–1178, 2009.

[12] Vesselin Gasharov, Irena Peeva, and Volkmar Welker. Rationality for generic toric rings. *Math. Z.*, 233(1):93–102, 2000.

[13] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[14] Daniel R. Grayson and Michael E. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at `http://www.math.uiuc.edu/Macaulay2/`.

[15] Jonathan Gryak, Robert Haralick, and Delaram Kahrobaei. Solving the conjugacy decision problem via machine learning. 05 2017.

[16] Tor Holtedahl Gulliksen. Massey operations and the poincaré series of certain local rings. *Preprint series: Pure mathematics http://urn. nb. no/URN: NBN: no-8076*, 1970.

[17] Jürgen Herzog, Vic Reiner, and Volkmar Welker. The Koszul property in affine semigroup rings. *Pacific J. Math.*, 186(1):39–65, 1998.

[18] Olav Arnfinn Laudal and A Sletsjøe. Betti numbers of monoid algebras. applications to 2-dimensional torus embeddings. *Mathematica Scandinavica*, 56(2):145–162, 1985.

[19] Irena Peeva. Infinite free resolutions over toric rings. In *Syzygies and Hilbert functions*, volume 254 of *Lect. Notes Pure Appl. Math.*, pages 233–247. Chapman & Hall/CRC, Boca Raton, FL, 2007.

[20] Jean-Yves Potvin. State-of-the-art survey?the traveling salesman problem: A neural network perspective. *ORSA Journal on Computing*, 5(4):328–348, 1993.

[21] Jan-Erik Roos and Bernd Sturmfels. A toric ring with irrational Poincaré-Betti series. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(2):141–146, 1998.

[22] Richard P Stanley. *Combinatorics and commutative algebra*, volume 41. Springer Science & Business Media, 2007.

[23] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 8.4)*, 2018. http://www.sagemath.org.

[24] Geoffrey G Towell and Jude W Shavlik. Extracting refined rules from knowledge-based neural networks. *Machine learning*, 13(1):71–101, 1993.