Symmetry in Toric Geometry

Dissertation
zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von M.Sc., Lamprini Ananiadi

geb. am 26. November 1992 in Athen, Griechenland

genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Thomas Kahle
(Otto-von-Guericke-Universität Magdeburg)

Prof. Dr. Tim Römer
(Universität Osnabrück, Insitut für Mathematik)

eingereicht am: 28.04.2020

Verteidigung am: 31.08.2020

Zusammenfassung

Familien algebraischer Varietäten, die durch Monome parametrisiert sind, tauchen in verschiedenen Bereichen der Mathematik, wie zum Beispiel der Statistik, der kommutativen Algebra oder der Kombinatorik auf. Solche Varietäten nennt man torische Varietäten und ihre Untersuchung bildet das Feld der torischen Geometrie. Die Ideale, die solche Varietäten definieren, sind sogenannte Binomialideale. Die torische Geometrie ist häufig durch ein enges Zusammenspiel zwischen algebraischer und polyedrischer Geometrie geprägt, da sich strukturelle Aussagen über eine torische Varietät meist kombinatorisch interpretieren lassen. Ziel dieser Arbeit ist es, Fortschritte in zwei verschiedenen Bereichen an diesem Schnittpunkt zweier Felder zu machen.

Im ersten Teil dieser Arbeit, in Kapitel 2 und Kapitel 3, untersuchen wir kombinatorische Objekte im Bereich der torischen Geometrie modulo Symmetrie. Genauer gesagt betrachten wir die korrespondierenden Kegel von Familien von Binomialidealen, deren Anzahl an Variablen unbeschränkt ist, die sich aber modulo Symmetrie stabilisieren. Mit Stabilisierung modulo Symmetrie ist gemeint, dass diese Ideale von den Orbits einer Wirkung der unendlichen symmetrischen Gruppe auf endlich vielen Polynomen erzeugt sind. In Kapitel 2 berechnen wir explizit die Gleichungen und Ungleichungen, die die Facetten der genannten Kegel definieren. Anhand dieser Berechnungen zeigen wir die sogenannte kombinatorische Stabilisierung bestimmter Familien von Kegeln. In Kapitel 3 formulieren wir Kriterien für die Stabilisierung modulo Symmetrie für allgemeine Familien polyedrischer Kegel. Hier ist das zentrale Resultat die Aussage, dass die Stabilisierung einer Familie polyedrischer Kegel modulo Symmetrie die Stabilisierung der korrespondierenden Familie von Monoiden impliziert.

Der zweite Teil dieser Arbeit, Kapitel 4, beschäftigt sich mit der Ermittlung der Erzeuger von Idealen, die eine wichtige Rolle bei der Betrachtung spezieller statistischer Modelle, sogenannter Staged Trees, spielen. Unser Hauptresultat bezieht sich auf den Fall, in dem diese Ideale von Binomen erzeugt sind und besagt, dass in diesem Falle die Erzeuger eine quadratische Gröbnerbasis bilden und die Initialideale quadratfrei sind. Dies impliziert, dass für das Polytop der korrespondierenden torischen Varietät eine unimodulare Triangulierung existiert.

Abstract

Families of algebraic varieties that are parametrized by monomials appear in various areas of mathematics, such as statistics, commutative algebra and combinatorics. Such varieties are referred to as toric varieties and they are the structural objects of toric geometry. The ideals defining toric varieties are prime binomial ideals. Toric geometry is a field of rich interaction between algebraic and polyhedral geometry. We can take statements from algebraic geometry and look for their combinatorial interpretation and vice versa. The main contributions of this thesis are divided in two parts and aim at making progress in two special topics in this intersection.

The first part of the thesis, Chapter 2 and Chapter 3, investigates the combinatorial objects arising in toric geometry up to symmetry. Such objects are cones emerging from families of binomial ideals in an increasing number of variables that stabilize up to symmetry, that is, they are generated by the orbit under the action of the infinite symmetric group on finite sets of polynomials. In Chapter 2, we explicitly compute the facets of these cones by providing the description of their defining inequalities and equations. Based on this, we are able to deduce the combinatorial stabilization of the families of cones of our interest. Afterwards, in Chapter 3, we formulate criteria for stabilization up to symmetry for any family of polyhedral cones by looking at families of monoids. The main outcome here is that when a family of cones stabilizes up to symmetry then also the underlying family of monoids stabilizes.

The second part of this thesis, Chapter 4, addresses the problem of implicitly computing generating sets of the ideals defining combinatorial objects from statistics called staged trees. The main result states that when a staged tree is defined by a toric ideal, then this ideal is generated by a quadratic Gröbner basis and has squarefree initial ideal. As a consequence, the polytope corresponding to this toric variety has a unimodular triangulation.

Contents

Introduction

In algebra and particularly in computational mathematics, one often studies systems of polynomial equations in several variables. The set of solutions to a system of polynomial equations is an algebraic variety which is the building block of algebraic geometry. Although the algorithmic study of algebraic varieties is a hard and complicated task, many recent achievements in symbolic algebra provide the framework for experimental research and conjectures.

An inspiring result that established the development of algebraic geometry is the Hilbert basis theorem [Eis95, Corollary 1.5]. This result was stated and proven in 1888 by David Hilbert. It roughly states that every algebraic variety defined by an ideal in finitely many variables is carved out by finitely many polynomial equations. Here the assumption that the ideal is defined by finitely many variables is necessary, for if we require infinitely many variables then the ideal generated by these variables is not finitely generated. Among the different proofs of the basis theorem, a proof by Paul Gordan [Gor99] in 1899 is of particular interest. In this proof the idea of Gröbner bases made its first appearance. A Gröbner basis of an ideal is a special generating set that allows for a unique representation of the ideal and very often it is convenient for computations. The method for computing Gröbner bases was introduced in 1965 in the dissertation of Bruno Buchberger [Buc65] and it is named after his advisor Wolfgang Gröbner. Nowadays, Gröbner bases are practical algorithmic tools for solving systems of polynomial equations and consequently for efficiently computing with algebraic varieties.

In contrast to the above beautiful results by Hilbert and Buchberger, the situation changes dramatically when the number of variables in systems of polynomial equations increases unboundedly. In such cases ideals generated by these variables are known to not be finitely generated in general. As a result, the systems of interest cannot be reduced to simpler ones and the study of infinite dimensional algebraic varieties becomes very difficult. One approach to this problem is the subject of the newly established discipline called asymptotic algebra. Here, the main idea for computing with families of infinite dimensional algebraic varieties is to pass to the limit object and to examine if this admits a representation as a finite set of polynomials up to the action of some group or monoid. In this case, the finite generation of the limit object is connected with the stabilization of the family, which, in turn, makes the computation with such families easier.

The first known result of finite generation in asymptotic algebra was given by Cohen in 1967 during his study on problems related with group theory and metabelian varieties [Coh67]. He combined the Hilbert basis theorem with the study of well quasi ordered sets to show that ideals of polynomial rings with infinitely many variables are finitely generated up to the action of some monoid. Based on this result, in 1987, Cohen [Coh87] and his student P. Emmott [Emm87] developed an algorithmic theory which generalizes the reduction algorithm of Buchberger to systems of polynomials in infinitely many variables. Their method is known as equivariant Gröbner bases. At present, there is an implementation of equivariant Gröbner bases in the computer algebra system *Macaulay2*. This is due to Hillar, Krone and Leykin [HKL] and it is based on the work of Draisma and Brouwer [BD09].

The concept of finite generation in infinite dimensional polynomial rings was rediscovered several decades later by Aschenbrenner and Hillar [AH07, AH08] to approach problems in algebraic statistics [HS12]. Since in their computations the use of some natural symmetries among the variables is highlighted, the field of study was named commutative algebra up to symmetry, which is a subfield of the more general branch of asymptotic algebra. In the next few paragraphs we review some results in asymptotic algebra that are of particular interest.

The work of Draisma [Dra10] explores finiteness properties of algebraic varieties that appear in algebraic statistics and chemistry. The statistical model under consideration is the factor analysis model [DSS07] while the chemistry problem is related to chirality measurements [RSU67]. In both situations, a family of varieties in infinitely many variables is studied and the problem of finding an implicit finite description for all the varieties in the family is addressed. The outcome of this article is that there exists a finite description up to the action of a monoid. This description is different for any of the problems.

In their recent work, Nagel and Römer [RN17] follow a methodology similar to the one in [AH07, HS12] to study bivariate Hilbert series of ideals in infinite dimensional polynomial rings that are invariant under the action of the group of all permutations of the natural numbers or under the action of related monoids. Any such ideal is described as the union of ideals in lower dimensional polynomial rings which form an invariant chain. As a main result, the authors show the rationality of the Hilbert series which allows them to estimate the Krull dimension and the multiplicity of ideals in an invariant chain. The same result is obtained in [KLS16] in terms of formal languages. In [GN18] the authors specialize the computation of Hilbert series for the case of monomial ideals that are invariant under the action of the monoid of strictly increasing functions and they provide an implicit formula (Theorem 2.4 and Theorem 3.3 of [GN18]). A similar implicitization result is obtained in [MN19] for ideals defining hierarchical models [Sul18]. The subsequent articles [LNNR19, LNNR18, Mur19] study the asymptotic behavior of other invariants of chains of invariant ideals.

The use of commutative algebra up to symmetry is a fundamental tool in measuring the complexity of a homogeneous ideal in a polynomial ring. This is a problem related to the computation of the minimal free resolution of the ideal. The projective dimension of the ideal is an important invariant of such a resolution that counts the number of steps required to compute a minimal resolution. The Hilbert's Syzygy Theorem [Eis95, Corollary 19.7] states that every graded ideal over a polynomial ring with n variables has projective dimension at most n. Stillman's Conjecture [PS09, Problem 3.14] improves this bound. It asserts that the projective dimension of an ideal in a polynomial ring with n variables that is generated by finitely many homogeneous polynomials can be bounded by a number that is independent of n. This conjecture was first proven by Ananyan and Hochster [AH20] and was later proven by Erman, Sam and Snowden in [ESS19] and Draisma, Lasoń and Leykin in [DLL19]. The last two proofs are based on topological Noetherianity techniques introduced by Derksen, Eggermont and Snowden in [DES17] and Draisma in [Dra19].

Since its introduction, the use of symmetry in commutative algebra has played a crucial role in the study of algebro-geometric objects that have a rich combinatorial structure. A special class of these objects is referred to as toric varieties which are the structural objects of the subfield of algebraic geometry called toric geometry. Very often we use monomial maps or embeddings of semigroups in a lattice to define toric varieties. In the first case, a toric variety is determined by the zero set of the kernel of a monomial map. Such kernels form the defining ideals of toric varieties which are called toric ideals in the literature. They are defined as those prime ideals that are generated by monomial differences, called binomials. In the second case, the study of toric varieties is related with to the study of polyhedral cones, the structural objects of polyhedral geometry. This correspondence between polyhedral and toric geometry is fundamental to us because many properties of the semigroups defining the toric varieties are determined by properties of the convex cones. For instance a result due to Gordan states that a finitely generated polyhedral cone gives rise to a finitely generated semigroup.

Toric varieties arise naturally when studying problems originating from algebraic statistics. Frequently, we want to understand stabilization properties of families of toric varieties when some of the defining parameters grow to infinity. As mentioned earlier in this introduction, stabilization is achieved when the limit object of such a family is described by the action of the infinite symmetric group on a finite set of binomials. It is known that not every family of toric varieties stabilizes up to symmetry. For instance, the no hope theorem of De Loera and Onn identifies large families of objects with no hope of finite generation in the limit [DLO06]. On the positive side, the independent set theorem of Hillar and Sullivant [HS12, Theorem 4.7] and the more general result of Draisma, Eggermont, Krone and Leykin [DEKL13] describe large classes of toric varieties that stabilize up to symmetry. Related to this is the work of Kahle, Krone and Leykin [KKL14] that considers the problem of implicitly characterizing the generating sets of large families of toric varieties that stabilize up to symmetry.

The main objective of this thesis is first the study of the combinatorial objects arising in toric geometry up to symmetry and then the characterization of the generating sets of the ideals defining staged trees. The organisation of the thesis is the following.

In Chapter 1 we define well quasi orders, a fundamental notion when studying finite generation results. We also define direct limits, which are basic constructions from category theory, in order to introduce important limit objects that are needed in the development of this thesis. Later we introduce invariant ideals, that is, ideals which are closed under the action of some monoid (with emphasis on the case where the monoid is defined by the infinite symmetric group) and we review Noetherianity results of chains of invariant ideals in infinite dimensional polynomial rings.

In Chapter 2 we summarize the results in [DEKL13] and [KKL14] and we describe the polyhedral cones that correspond to toric varieties that stabilize up to symmetry. To be more precise, we provide an explicit computation of the facets of the cones by providing a description of their defining inequalities and equations (Proposition 2.2.4, Theorem 2.3.12, Theorem 2.3.14). Based on this we are able to deduce the combinatorial stabilization of the families of cones of our interest.

In Chapter 3 we introduce monoids that are closed under the action of the infinite symmetric group or related monoids and we examine when their underlying algebras are finitely generated up to symmetry. We define equivariant families of cones and we show (Theorem 3.4.6) that the property of stabilization up to symmetry is transferred from families of cones to families of monoids that arise by intersecting each of the cones in the family with the ambient space.

Finally, in Chapter 4 we study combinatorial objects, originating from statistics, called staged trees and we determine the generating sets of their defining ideals. In case these defining ideals are toric we show (Theorem 4.4.12) that their generators form a quadratic Gröbner basis whose initial terms are squarefree monomials. As a consequence, the polytope corresponding to this toric variety has a unimodular triangulation.

1 Commutative Algebra up to Symmetry

In this chapter of the thesis we review the basic results related to finite generation up to symmetry of large families of invariant ideals in infinite dimensional polynomial rings. We introduce well quasi orders, which are considered as the starting point for studying finite generation results. Furthermore, we define direct limits, a fundamental tool required later in this thesis.

1.1 Well Quasi Orders

The lemma of Higman is an important result in infinite combinatorics with various applications in Logic and Computer Science. It has been proven several times using different formulations and methods. The structural objects of this result are *well quasi orders* which are important tools when one wants to show the finite termination of algorithms. The notion of a well quasi order is based on the one of a quasi order which we define in the following.

Definition 1.1.1. A **quasi order** is a binary relation $\leq$ over a non-empty set X which is both reflexive, i.e. $x \leq x$ for any $x \in X$, and transitive, i.e. whenever $x_1 \leq x_2$ and $x_2 \leq x_3$ then $x_1 \leq x_3$ for any $x_1, x_2, x_3 \in X$.

We use a pair $(X, \leq)$ to denote a set X that is quasi-ordered by the relation $\leq$.

Remark 1.1.2. Let $(X, \leq)$ be a quasi order. If the relation $\leq$ is also antisymmetric, that is $x_1 \leq x_2$ and $x_2 \leq x_1$ imply $x_1 = x_2$ for any $x_1, x_2 \in X$, then it is a *partial order*. Hence, a partial order is a quasi order. The converse is not always true as the following example demonstrates.

Example 1.1.3. Consider the divisibility relation $|$ over the set $\mathbb{Z}$ of integer numbers given by $a \mid b$ if and only if there exists some $k \in \mathbb{Z}$ such that $b = k \cdot a$ for any $a, b \in \mathbb{Z}$. The pair $(\mathbb{Z}, |)$ is then a quasi order. The divisibility relation is reflexive (any integer number divides itself) and transitive : if $a \mid b$ then there is some $k_1 \in \mathbb{Z}$ with $b = k_1 \cdot a$ and if further $b \mid c$ then there is $k_2 \in \mathbb{Z}$ with $c = k_2 \cdot b$, it then follows that $c = k_2 \cdot b = k_2 \cdot (k_1 \cdot a) = (k_2 \cdot k_1) \cdot a$ which yields that $a \mid c$. However the pair $(\mathbb{Z}, |)$ is not a partial order because the divisibility relation is not antisymmetric. For instance if $a = -2$ and $b = 2$, then $a \mid b$ as $2 = (-1) \cdot a$, similarly $b \mid a$ as $-2 = 1 \cdot b$, but $a \neq b$.

Remark 1.1.4. We can turn a quasi order $\leq$ on the set X into a partial order on the set of equivalence classes of X. For this we define an equivalence relation on the elements of X as follows. For any $x, y \in X$ we write $x \sim y$ if and only if $x \leq y$ and $y \leq x$. Then, the equivalence class of $x \in X$ is the set $[x] = \{y : x \sim y\}$. This partitions X into a set of disjoint equivalence classes $X/\sim = \{[x] : x \in X\}$. We construct a relation $\leq_\sim$ on $X/\sim$ by defining $[x] \leq_\sim [y]$ if $x \leq y$. The relation $\leq_\sim$ is a partial order on $X/\sim$. Reflexivity and transitivity arise from $(X, \leq)$ being a quasi order, while antisymmetry follows from the observation that if $[x] \leq_\sim [y]$ and $[y] \leq_\sim [x]$, then $x \leq y$ and $y \leq x$ which means that $x \sim y$, hence $[x] = [y]$.

Example 1.1.5 (Example 1.1.3 continued)**.** The divisibility order over the set of integer numbers is a quasi order. According to Remark 1.1.4, this quasi order gives rise to a partial order on the set of equivalences classes $\mathbb{Z}/\sim = \{[a] : a \in \mathbb{Z}\}$, where $[a] = -a$ for any $a \in \mathbb{Z}$, because the only way that two distinct integers divide each other is when they are opposite.

Definition 1.1.6. A quasi order $(X, \leq)$ is called a **well quasi order** if

1. it is *well founded*, that is, every strictly decreasing sequence of elements in X is finite, and

2. X does not have *infinite antichains*, that is, any subset $A \subseteq X$ of pairwise incomparable elements is finite.

Example 1.1.7. The pair $(\mathbb{N}, \leq)$, the natural numbers under the standard ordering is a well quasi order.

Example 1.1.8.

- Consider the pair $(\mathbb{Z}, \leq)$ of integral numbers with the standard ordering. This is not a well quasi order. For any $z \in \mathbb{Z}$, consider the sequence

$$z > z - 1 > z - 2 > z - 3 > \dots,$$

and notice that this is an infinite decreasing sequence. Therefore the set of integer numbers is not well founded.

- The pair $(\mathbb{N}, |)$ of natural numbers with the divisibility order is also not a well quasi order. To see this, observe first that the set of prime numbers is infinite. Since no prime number divides another, any two primes are incomparable. Therefore, the set of prime numbers forms an infinite antichain.

Definition 1.1.9. Let $(X, \leq)$ be a quasi order. An infinite sequence $(x_i : i \in \mathbb{N})$ of elements in X is **good**, if $x_i \leq x_j$ for some indices $i < j$. Otherwise it is called **bad**.

Let $(X, \leq)$ be a quasi ordered set. A *final segment* is a subset F of X that is closed upwards, that is, for any $x_1, x_2 \in X$, if $x_1 \in F$ and $x_1 \leq x_2$, then $x_2 \in F$. Given an arbitary subset M of X, we denote the final segment generated by M as follows

$$\mathrm{fin}(M) = \{x_2 \in X : \exists x_1 \in M \text{ such that } x_1 \leq x_2\}.$$

The definition of a well quasi order provided so far is in terms of well founded sequences and antichains. There exist several other conditions which characterize the concept of a well-quasi-order and that can be considered as equivalent definitions. We refer to Kruskal's article [Kru72] for the general theory and to [FT, Theorem 3.2] for a beautiful proof of the next result.

Proposition 1.1.10. The following statements are equivalent for a quasi order $(X, \leq)$.

1. $(X, \leq)$ is a well quasi order.
2. Any final segment of X is finitely generated.
3. If $F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n \subseteq \ldots$ is an ascending chain of final segments of X, then this chain is eventually stable, that is, there exists some natural number $n \in \mathbb{N}$ such that $F_N = F_n$, for any $N \geq n$.
4. Any infinite sequence of elements in X is good.
5. Any infinite sequence of elements in X admits an infinite ascending subsequence.

Remark 1.1.11. The property described in Proposition 1.1.10(3) is known as the ascending chain condition required for stabilization (Noetherianity) results in commutative algebra, while the property in Proposition 1.1.10(2) is characterized by Higman in [Hig52] as the finite basis property.

Given n quasi-ordered sets $(X_1, \leq_1), (X_2, \leq_2), \ldots, (X_n, \leq_n)$, we can form the product quasi-order $(X, \leq^n)$ on the set $X = X_1 \times X_2 \times \cdots \times X_n$ in the following way. Let $x_1, y_1 \in X_1, x_2, y_2 \in X_2, \ldots, x_n, y_n \in X_n$. Then

$$(x_1, \ldots, x_n) \leq^n (y_1, \ldots, y_n) \iff x_1 \leq_1 y_1, \ldots, x_n \leq_n y_n.$$

This order is reflexive and transitive and therefore $(X, \leq^n)$ is a quasi order.

Proposition 1.1.12. Consider the well quasi orders $(X_1,\leq_1),(X_2,\leq_2),\dots,(X_n,\leq_n)$. Then $(X,\leq^n)$ is also a well-quasi order.

Proof. We prove that $(X,\leq^n)$ is a well quasi order showing that a sequence of tuples of X admits an ascending subsequence. Let

$$\mathcal{X}_0 = (x_1,y_1,\dots,z_1),(x_2,y_2,\dots,z_2),\dots,(x_n,y_n,\dots,z_n),\dots$$

be an infinite sequence of tuples of X. Since $(X_1,\leq_1)$ is a well quasi order, the sequence $\mathcal{X}_0$ admits an infinite subsequence that increases in the first component:

$$\mathcal{X}_1 = (x_{i_1},y_1,\dots,z_1) \leq_1 (x_{i_2},y_2,\dots,z_2) \leq_1 \cdots \leq_1 (x_{i_n},y_n,\dots,z_n) \leq_1 \dots,$$

where $1\leq i_1 < i_2 < \cdots < i_n$. Since $(X_2,\leq_2)$ is a well quasi order, the sequence $\mathcal{X}_1$ admits an infinite subsequence that increases both in the first and in the second component:

$$\mathcal{X}_2 = (x_{j_1},y_{j_1},\dots,z_1) \leq_2 (x_{j_2},y_{j_2},\dots,z_2) \leq_2 \cdots \leq_2 (x_{j_n},y_{j_n},\dots,z_n) \leq_2 \dots,$$

where $j_l \in \{i_1,i_2,\dots,i_n\}, l = 1,2,\dots,n$ and $j_1 < j_2 < \cdots < j_n$.
Using the same argument, that $(X_i,\leq_i)$ is a well quasi order, we construct a subsequence $\mathcal{X}_i$ of $\mathcal{X}$ which increases in the first i components. Thus, $\mathcal{X}_n$ is the subsequence of $\mathcal{X}$ in which all the components are increasing. We conclude that every sequence of elements of X admits an increasing subsequence and therefore $(X,\leq^n)$ is a well quasi order. □

If we specialize Proposition 1.1.12 for the set of n−tuples of natural numbers with the component-wise ordering we obtain the following result.

Corollary 1.1.13 (Dickson's Lemma, [Dic13])**.** The pair $(\mathbb{N}^n,\leq^n)$, of the set of all n−tuples of natural numbers with the component-wise order introduced above, is a well quasi order.

Remark 1.1.14. Dickson's Lemma is very important in computational algebraic geometry as it ensures that any monomial ideal of the polynomial ring $R = \mathbb{K}[x_1,\dots,x_n]$ is finitely generated. In order to see this we associate to an n−tuple of natural numbers $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_n) \in \mathbb{N}^n$ a monomial $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$ in R and we consider the monomial ideal $I = \langle \boldsymbol{x}^\alpha : \alpha \in \mathbb{N}^n\rangle$ of R. Since $(\mathbb{N}^n,\leq^n)$ is a well quasi order, it follows from Proposition 1.1.10 that any final segment of $\mathbb{N}^n$ is finitely generated, hence there is a finite subset $\{\boldsymbol{\beta_1},\dots,\boldsymbol{\beta_r} : \boldsymbol{\beta_i} \in \mathbb{N}^n\}$ of $\mathbb{N}^n$ such that the monomials in the set $\{\boldsymbol{x}^{\boldsymbol{\beta_1}},\dots,\boldsymbol{x}^{\boldsymbol{\beta_r}}\}$ generate I. Therefore, Dickson's lemma is a special case of Hilbert basis theorem ([Eis95, Corollary 1.5]) which states that any ideal in R is finitely generated.

For the rest of this section consider an arbitary set X, and denote by X^* the set of all finite words over X. A quasi order $\leq$ on X yields a quasi order $\leq_{\mathrm{H}}$ on X^* as follows

$$(x_1, \ldots, x_p) \leq_{\mathrm{H}} (x'_1, \ldots, x'_q)$$

if and only if there exists a strictly increasing map $\varphi : [p] \to [q]$ such that $x_i \leq_{\mathrm{H}} x_{\varphi(i)}$ for any $i \in [p]$. The content of the following Lemma is that for X^* to be well quasi ordered, it is sufficient that X is well quasi ordered.

Lemma 1.1.15 (Higman's Lemma, [Hig52])**.** If X is a well quasi ordered set, then X^* is also a well quasi ordered set.

For a nice proof of Lemma 1.1.15 we refer to the work [NW63] of Nash-Williams and to the lecture notes [Dra14, Chapter 1] of Jan Draisma. Well quasi orders were extensively used in [AH07, HS12] and [RN19] to show that Gröbner bases of polynomial rings with infinitely many variables are finitely generated up to symmetry.

1.2 Direct Limits

In this section we briefly introduce direct limits. In category theory, the direct limit is a fundamental way to construct a large object by putting together many smaller objects. These objects may be monoids, groups, rings or vector spaces. The way that these objects are put together is specified via a set of homomorphisms between the smaller objects. Before defining direct limits, we review basic notions from category theory. The main reference here is the book of Steve Awodey [Awo10].

Definition 1.2.1. A **category** $\mathcal{C}$ consists of a class $\mathcal{O}b(\mathcal{C})$ of *objects* and a class of maps between these objects such that the following conditions hold.

$C1$ For all $A, B \in \mathcal{O}b(\mathcal{C})$ there is a (possibly empty) set $\mathrm{Mor}_{\mathcal{C}}(A, B)$, called the *set of morphisms* $f : A \to B$ from A to B, such that

$$\mathrm{Mor}_{\mathcal{C}}(A, B) \cap \mathrm{Mor}_{\mathcal{C}}(A', B') = \varnothing, \quad \text{if } (A, B) \neq (A', B')$$

$C2$ If $A, B, C \in \mathcal{O}b$, then there is a *rule of composition*

$$\mathrm{Mor}_{\mathcal{C}}(A, B) \times \mathrm{Mor}_{\mathcal{C}}(B, C) \to \mathrm{Mor}_{\mathcal{C}}(A, C),$$

defined by $(f, g) \mapsto g \circ f$, so that the following rules hold

(a) Associativity: If $f : A \to B, g : B \to C$ and $h : C \to D$ are morphisms in $\mathcal{C}$, then $(h \circ g) \circ f = h \circ (g \circ f)$.

(b) Existence of Identity: For any $A \in \mathcal{O}b$, there is an identity morphism $\mathrm{Id}_A : A \to A$ such that $f \circ \mathrm{Id}_A = f$ and $\mathrm{Id}_A \circ g = g$, for any morphisms $f : A \to B$ and $g : C \to A$ of $\mathcal{C}$.

Example 1.2.2.

- The category **Sets** of sets has as its objects arbitary sets and as morphisms the maps between sets. The rule of composition is the usual composition of maps.
- The category **Rng** of rings has objects all rings not necessarily having an identity element, and the set of morphisms consists of all ring homomorphisms. The rule of composition in this case is the composition of ring homomorphisms.
- Similarly to the categories **Sets**, **Rng**, we form the categories **Grp** of groups and **Mon** of monoids.
- Let $(X, \leq)$ be a quasi order. We can interpret $(X, \leq)$ as a category with objects the elements of X. For any $x, y \in X$, the set $\mathrm{Mor}(x, y)$ of morphisms in this category, is defined by a unique map $f : x \to y$ whenever x, y are comparable, that is whenever $x \leq y$, otherwise it is the empty set. The rule of composition is given by the property of the quasi order to be transitive, that is, if $x, y, z \in X$, then $x \leq y$ and $y \leq z$ implies $x \leq z$. In terms of the setup above, this means that if there exist maps $f : x \to y$ and $g : y \to z$, then $g \circ f : x \to z$. The associativity and existence of identity is then implied by the usual composition of maps.

The starting point for the direct limit construction is a family of objects $\{X_i : i \in I\}$ in some category $\mathcal{C}$, indexed by a non-empty quasi-ordered set I that has an additional property which is described in the following definition.

Definition 1.2.3. Let $(I, \leq)$ be a quasi order. We say that I is a **directed set**, if for any $i_1, i_2 \in I$ there exists $i_3 \in I$ such that $i_1 \leq i_3$ and $i_2 \leq i_3$.

Example 1.2.4. The pair $(\mathbb{N}, \leq)$, of natural numbers together with the standard ordering $\leq$ defines a directed set.

Definition 1.2.5. Let I be a directed set. An **inductive system** in a category $\mathcal{C}$ over I, is a pair $(\boldsymbol{X}, \boldsymbol{\varphi})$, consisting of a family of objects $\boldsymbol{X} = \{X_i : i \in I\}$, together with a collection of maps

$$\boldsymbol{\varphi} = \{\varphi_{i,j} : X_i \to X_j, \text{ for any } i, j \in \mathbb{N} \text{ with } i < j\}$$

that satisfy the following properties

- $\varphi_{i,i} = \mathrm{Id}_{X_i}$, for any $i \in I$, where $\mathrm{Id}_{X_i} : X_i \to X_i$ is the identity map, and,

► $\varphi_{j,k} \circ \varphi_{i,j} = \varphi_{i,k}$, for any $i, j, k \in I$ with $i < j < k$.

Example 1.2.6. In the following examples the families of objects are indexed over the directed set $(\mathbb{N}, \leq)$.

1. Let X be a non-empty set, and let $\boldsymbol{X} = \{X_i : i \in \mathbb{N}\}$ be a family of subsets of X. Assume that these subsets form a chain

$$\varnothing =: X_0 \subseteq X_1 \subseteq X_2 \subseteq \cdots \subseteq X_{i-1} \subseteq X_i \subseteq X_{i+1} \subseteq \dots .$$

   Thus, there are natural inclusion maps $\varphi_{i,j} : X_i \to X_j$ for any pair i, j of natural numbers with $i \leq j$. These maps are such that $\varphi_{i,i} = \mathrm{Id}_{X_i}$ for any $i \in \mathbb{N}$ and $\varphi_{j,k} \circ \varphi_{i,j} = \varphi_{i,k}$ for any $i, j, k \in \mathbb{N}$ with $i \leq j \leq k$. Hence, the pair $(\boldsymbol{X}, \boldsymbol{\varphi})$, where $\boldsymbol{\varphi} = \{\varphi_{i,j} : i, j \in \mathbb{N}, i \leq j\}$, is an inductive system over $\mathbb{N}$.

2. Given some fixed natural number $k \in \mathbb{N}$ and any $n \in \mathbb{N}$, consider the finite set of indeterminates $X_n = \{x_{i,j} : i \in [k], j \in [n]\}$. For any $n \in \mathbb{N}$ denote by $R_n = \mathbb{K}[X_n]$ the commutative ring with variables in X_n and coefficients in some field $\mathbb{K}$. Let $\boldsymbol{R} = \{R_n : n \in \mathbb{N}\}$ be a family of such rings, and assume that there is a chain

$$R_1 \subseteq R_2 \subseteq \cdots \subseteq R_n \subseteq R_{n+1} \subseteq \dots .$$

   Define $\varphi_{m,n} : R_m \to R_n$ to be the natural inclusion map of R_m into R_n for any $m, n \in \mathbb{N}$ with $m \leq n$. Those maps are such that $\varphi_{m,m} = \mathrm{Id}_{R_m}$ for any $m \in \mathbb{N}$, and $\varphi_{n,r} \circ \varphi_{m,n} = \varphi_{m,r}$ for any $m, n, r \in \mathbb{N}$ with $m \leq n \leq r$. Hence the pair $(\boldsymbol{R}, \boldsymbol{\varphi})$, where $\boldsymbol{\varphi} = \{\varphi_{m,n} : m, n \in \mathbb{N}, m \leq n\}$, is an inductive system over $\mathbb{N}$.

3. Let $\boldsymbol{S} = \{\mathrm{Sym}(n) : n \in \mathbb{N}\}$ be a collection of finite symmetric groups. For any $m, n \in \mathbb{N}$ with $m \leq n$ there are natural inclusion maps

$$\varphi_{m,n} : \mathrm{Sym}(m) \longrightarrow \mathrm{Sym}(n),$$
$$\sigma \mapsto \varphi_{m,n}(\sigma)(i) := \begin{cases} \sigma(i), & \text{if } i \in [m], \\ i, & \text{else.} \end{cases}$$

   These maps are such that $\varphi_{m,m} = \mathrm{Id}_{\mathrm{Sym}(m)}$ for any $m \in \mathbb{N}$, and $\varphi_{n,r} \circ \varphi_{m,n} = \varphi_{m,r}$ for any $m, n, r \in \mathbb{N}$ with $m \leq n \leq r$. Hence the pair $(\boldsymbol{S}, \boldsymbol{\varphi})$, where $\boldsymbol{\varphi} = \{\varphi_{m,n} : m, n \in \mathbb{N}, m \leq n\}$, is an inductive system over $\mathbb{N}$.

Definition 1.2.7. The **direct limit** of the inductive system $(\boldsymbol{X}, \boldsymbol{\varphi})$ in the category $\mathcal{C}$ over the directed set I, is an object in $\mathcal{C}$ that is denoted by $\varinjlim X_i$, together with a collection of morphisms

$$\{\varphi_{i,\infty} : X_i \longrightarrow \varinjlim X_i : i \in I\}$$

such that the following properties hold

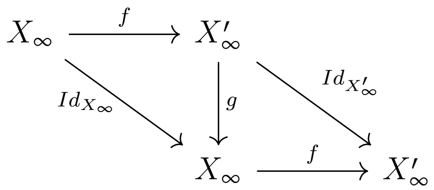
- $\varphi_{i,\infty} = \varphi_{j,\infty} \circ \varphi_{i,j}$ for all indices $i, j \in I$ with $i \leq j$.
- For any object Y in $\mathcal{C}$ and morphisms $f_i : X_i \to Y$ that satisfy the relation $f_j \circ \varphi_{i,j} = f_i$ for any $i \leq j$, there exists a unique morphism $\Phi : \varinjlim X_i \to Y$ such that $f_i = \Phi \circ \varphi_{i,\infty}$ for any $i \in I$. (*universal property for direct limits*)

For simplicity the direct limit of an inductive system $(\boldsymbol{X}, \boldsymbol{\varphi})$ is, from now on, denoted by $X_\infty := \varinjlim X_i$.

Remark 1.2.8. The direct limit of an inductive system does not always exists as Example 1.2.13 shows. Nevertheless if it exists, it is unique up to isomorphism as indicated in the following proposition. The dual of the direct limit, when it exists, is the projective limit.

Proposition 1.2.9. The direct limit (when it exists) is unique up to isomorphism.

Proof. Assume that X_∞, X'_∞ are two different direct limits of an inductive system $(\boldsymbol{X}, \boldsymbol{\varphi})$ in a category $\mathcal{C}$. Let $f : X_\infty \to X'_\infty$ be a morphism from X_∞ to X'_∞ and $g : X'_\infty \to X_\infty$ be a morphism from X'_∞ to X_∞. Then the following diagram



commutes. Note that $f \circ g$ is a morphism from X_∞ to itself and $g \circ f$ is a morphism from X'_∞ to itself. In other words, $f \circ g, g \circ f$ are the identity maps

$$f \circ g = \mathrm{Id}_{X_\infty}, g \circ f = \mathrm{Id}_{X'_\infty},$$

and therefore, the maps f and g are mutual inverses. Hence the maps f and g are isomorphisms in $\mathcal{C}$ and the uniqueness of X_∞ follows. □

Proposition 1.2.10. Direct limits exist in the category **Sets**.

Proof. Let $(\{X_i : i \in I\}, \{\varphi_{i,j} : X_i \to X_j : i, j \in I, i \leq j\})$ be an inductive system in the category **Sets** indexed by a directed set I. Let $\mathcal{X}' = \bigcup_{i \in I} X_i$ and define an equivalence relation on the elements of $\mathcal{X}'$ as follows. If $x_i \in X_i, x_j \in X_j$ for any $i, j \in I$ with $i \leq j$, we say that x_i is equivalent to x_j and we write $x_i \sim x_j$ whenever $\varphi_{i,k}(x_i) = \varphi_{j,k}(x_j)$ for some index $k \in I$ with $i \leq k, j \leq k$. If $x_i \in X_i$ then we denote the equivalence class of x_i by $[x_i]$. Let $\mathcal{X} = \mathcal{X}'/\sim$ be the set of all equivalence classes and consider the natural embedding $\varphi_{i,\infty} : X_i \to \mathcal{X}$ defined by $x_i \mapsto \varphi_{i,\infty}(x_i) = [x_i]$ for any $i \in I$. Notice that

$\varphi_{j,\infty}(\varphi_{i,j}(x_i)) = \varphi_{i,\infty}(x_j) = [x_j] \overset{x_i \sim x_j}{=} [x_i] = \varphi_{i,\infty}(x_i)$ and therefore $\varphi_{i,\infty} = \varphi_{j,\infty} \circ \varphi_{i,j}$ for any $i, j \in I, i \le j$. We claim that $\varinjlim X_i \cong \mathcal{X}$. In order to prove the claim we need to show that $\mathcal{X}$ satisfies the universal property of direct limits. To this end, let Y be some set and consider the maps $f_i : X_i \to Y$ with $f_i = f_j \circ \varphi_{i,j}$ for any $i, j \in I, i \le j$. If $[x_i] \in \mathcal{X} = \bigcup_{i \in I} X_i / \sim$, then there is an index $i \in I$ such that $[x_i] \in X_i / \sim$ or $x_i \in X_i$. Hence it makes sense to define the map $\Phi : \mathcal{X} \to Y$ as $\Phi([x_i]) = f_i(x_i)$ for any $i \in I$. We can check that Φ is well-defined. If $[x_i] = [x_j]$ then there is $k \in I$ with $i \le k, j \le k$ such that $\varphi_{i,k}(x_i) = \varphi_{j,k}(x_j)$, moreover we have $f_i(x_i) = f_k(\varphi_{i,k}(x_i)) = f_k(\varphi_{j,k}(x_j)) = f_j(x_j)$. Furthermore, the map Φ satisfies $f_i = \Phi \circ \varphi_{i,\infty}$. To see this we note that $\Phi([x_i]) = f_i(x_i) = (f_j \circ \varphi_{i,j})(x_i)$ holds for any indices i, j with $i \le j$. This if we choose the index $i \in I$ to be large enough, then the latest holds for any index $j \in I$ that is large enough and approaches the limit. Hence $\Phi([x_i]) = (f_\infty \circ \varphi_{i,\infty})(x_i)$ where $f_\infty = \Phi$. Thus $f_i(x_i) = (\Phi \circ \varphi_{i,\infty})(x_i)$ for any $i \in I$. It remains to show the uniqueness of Φ. On the contrary assume that there exists $\Phi' : \mathcal{X} \to Y$ such that $f_i = \Phi' \circ \varphi_{i,\infty}$ for any $i \in I$. Then, $\Phi([x_i]) = f_i(x_i) = \Phi'(\varphi_{i,\infty}(x_i)) = \Phi'([x_i])$, and so $\Phi = \Phi'$. This proves our claim that $\mathcal{X} \cong \varinjlim X_i$. □

Remark 1.2.11. The direct limits of inductive systems in *concrete categories* exist and admit a characterization similar to the one of the direct limits in **Sets** ([Awo10, Proposition 5.31]). Here by a concrete category ([Awo10, Paragraph 1.4, item 4]) we mean a category whose objects are sets with some additional structure. Morphisms in such categories are functions preserving this structure. For instance, the category **Rng** of rings is concrete, because the objects are rings, i.e. sets together with two binary operations (usually addition and multiplication), and morphisms are just ring homomorphisms (that is functions preserving the ring structure). Other examples of set-based categories are the categories **Grp** of groups and **Mon** of monoids.

Example 1.2.12.

- The direct limit of the inductive system $(\boldsymbol{R}, \boldsymbol{\varphi})$ introduced in Example 1.2.6(2) is defined in terms of the direct limit of the family of indeterminates $\{X_n : n \in \mathbb{N}\}$. We have that $X_\infty = \{x_{i,j} : i \in [k], j \in \mathbb{N}\}$, and therefore $R_\infty = \mathbb{K}[X_\infty]$ is the polynomial ring in infinitely many variables over a field $\mathbb{K}$.

- Suppose that we want to define the direct limit of the inductive system $(\boldsymbol{S}, \boldsymbol{\varphi})$ of Example 1.2.6(3). If $\mathfrak{S}_\infty = \varinjlim \mathrm{Sym}(n)$, then we can implicitly characterize this direct limit making the following observation. Since $\mathfrak{S}_\infty$ exists, there are maps

$$\varphi_{n,\infty} : \mathrm{Sym}(n) \to \mathfrak{S}_\infty$$
$$\sigma \mapsto \varphi_{n,\infty}(\sigma) = \{\sigma \in \mathrm{Sym}(n) : \sigma(i) = i \ \forall i \in \mathbb{N} \backslash [n]\}, \quad \forall n \in \mathbb{N}.$$

Hence we may define $\mathfrak{S}_\infty$ as the set of all permutations of the natural numbers that fix all but finitely many numbers. Notice that $\mathfrak{S}_\infty$ together with composition of permutations operation is a group which is often referred in the literature as the *infinite symmetric group*. This group has a crucial role in the following sections and chapters of the thesis.

Example 1.2.13. Consider the category **FinSets** of finite sets and morphisms the maps between them. Although there exists an inductive system in this category, the direct limit does not exist, for if it existed it would had been an infinite set.

1.3 Equivariant Noetherianity

In the last section of this chapter, we review Noetherianity results for infinite dimensional polynomial rings. We study ideals in infinitely many variables that are invariant under the action of the infinite symmetric group and we provide instances for finite generation up to symmetry. Finally, we introduce a Gröbner basis theory for the ideals of interest. The main references for Subsection 1.3.1 are the articles [Coh67, AH07, HS12]. In Subsection 1.3.2 we summarize the results in [Coh87, AH08, HS12, HKL18]. The article [Dra10] and the lecture notes [Dra14] are excellent sources for an introduction to equivariant Noetherianity.

1.3.1 Invariant Ideals

Throughout this paragraph, let $\mathbb{K}$ be a field, R a $\mathbb{K}$–algebra and Π a monoid. We assume that Π acts on R in terms of $\mathbb{K}$–algebra homomorphisms, that is through a map

$$\begin{aligned}\rho : \Pi \times R &\longrightarrow R,\\ (\pi, f) &\mapsto \rho((\pi, f)) =: \pi(f), \quad \forall \pi \in \Pi, \forall f \in R.\end{aligned}$$

Associated to R and Π is the skew monoid ring $S = R \star \Pi$ whose elements are finite sums $\sum_{\pi\in\Pi} f_\pi \pi$, where $f_\pi \in R$ for each $\pi \in \Pi$ and $f_\pi = 0$ for all but finitely many $\pi \in \Pi$. The binary operation of addition is given coefficient-wise by

$$\left(\sum_{\pi\in\Pi} f_\pi \pi\right) + \left(\sum_{\pi\in\Pi} g_\pi \pi\right) = \sum_{\pi\in\Pi} (f_\pi + g_\pi)\pi,$$

while multiplication is defined distributively via the formula

$$\left(\sum_{\pi\in\Pi} f_\pi \pi\right) \cdot \left(\sum_{\tau\in\Pi} g_\tau \tau\right) = \sum_{\pi,\tau\in\Pi} f_\pi \pi(g_\tau)\pi\tau = \sum_{\nu\in\Pi} \left(\sum_{\substack{\pi,\tau\in\Pi\\ \pi\tau=\nu}} f_\pi \pi(g_\tau)\right)\nu,$$

where $\pi(g_\tau)$ denotes the element of R obtained by π acting on g_τ. The left action of S on R, gives R the structure of an S−module.

Definition 1.3.1. An ideal $I \subseteq R$ is called **Π−invariant** if it is closed under the action of Π, that is,

$$\Pi I := \{\pi(f) : \pi \in \Pi, f \in I\} \subseteq I, \quad \forall \pi \in \Pi.$$

Let P be a $\mathbb{K}$−algebra. The homomorphism $\phi : R \to P$ is called **Π−equivariant** if $\pi(\phi(f)) = \phi(\pi(f))$ for any $\pi \in \Pi$ and any $f \in R$.

Following the discussion before Definition 1.3.1 we conclude that Π−inariant ideals are exactly the S−submodules of R. Therefore, the study of finiteness properties of Π−invariant ideals is connected with the study of Noetherianity properties of the skew monoid ring S. Such results are known in the literature of non-commutative algebra, like for example [TYL01, JM01].

Definition 1.3.2. A Π−invariant ideal I of R is Π−**finitely generated** if there exists a finite set F such that I is generated by the Π−orbits of the elements in F. In this case, we write $I = \langle F\rangle_\Pi$.

In this thesis we are mainly interested in the case where Π is the infinite symmetric group $\mathfrak{S}_\infty$ of Example 1.2.12, or certain related monoids. We also consider the case where the $\mathbb{K}$−algebra R has the structure of a polynomial ring, $R = \mathbb{K}[X]$ with X being a (not necessarily finite) set of indeterminates. The following examples are instances of $\mathfrak{S}_\infty$−invariant ideals, some of them being $\mathfrak{S}_\infty$−finitely generated, but the last one which is not.

Example 1.3.3. Let $X = \{x_1, x_2, \dots\}, R = \mathbb{K}[X]$, and let $\Pi = (\mathfrak{S}_\infty, \circ)$. Then Π acts on R by permuting the variables on X, so that $\sigma(x_i) := x_{\sigma(i)}$ for any $\sigma \in \mathfrak{S}_\infty, i \in \mathbb{N}$. The ideal $I = \langle x_1, x_2, \dots\rangle$ of R is $\mathfrak{S}_\infty$−invariant but is not finitely generated over R. Nevertheless, we observe that the $\mathfrak{S}_\infty$−orbit of $\{x_1\}$ generates I.

Example 1.3.4. For some $n \in \mathbb{N}$ consider the matrix $X = (x_{i,j} : i, j \in [n]) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. Let $R_n = \mathbb{K}[x_{i,j} : i, j \in [n]]$ be the polynomial ring with variables the entries of the matrix X over a field $\mathbb{K}$, and consider the ideal I_n of R_n that is generated by the 2×2 minors of X, i.e.

$$I_n = \langle x_{i_1,j_1}x_{i_2,j_2} - x_{i_1,j_2}x_{i_2,j_1} : i_1, i_2, j_1, j_2 \in [n]\rangle, \quad \forall n \geq 2.$$

Assume that the monoid $(\mathrm{Sym}(n), \circ)$ acts on X by simultaneously permuting its rows and columns. In terms of the variables of R_n, this means that $\sigma(x_{i,j}) := x_{\sigma(i),\sigma(j)}$ for any $\sigma \in \mathrm{Sym}(n)$ and any $i, j \in [n]$. The ideals I_n are $\mathrm{Sym}(n)$−invariant for any $n \in \mathbb{N}, n \geq 2$.

For increasing n the number of generators of the ideal I_n increases as well. However, for $n \geq 4$, the generators of I_n lie in the $\mathrm{Sym}(n)$–orbit of the set

$$G = \{x_{1,1}x_{2,2} - x_{1,2}x_{2,1},\ x_{1,1}x_{2,3} - x_{1,3}x_{2,1},\ x_{1,2}x_{3,4} - x_{1,4}x_{3,2}\}.$$

This is based on the observation that for any $n \geq 4$ the $\mathrm{Sym}(n)$–orbit of the minor $x_{1,1}x_{2,2}-x_{1,2}x_{2,1}$ are those minors of X involving elements in the diagonal of the matrix, the $\mathrm{Sym}(n)$–orbit of the minor $x_{1,1}x_{2,3}-x_{1,3}x_{2,1}$ are minors of X involving an element in the diagonal and another non-diagonal element, finally the $\mathrm{Sym}(n)$–orbit of the minor $x_{1,2}x_{3,4} - x_{1,4}x_{3,2}$ produce all other minors of X. For any $m, n \in \mathbb{N}$ there are natural inclusion relations $I_m \subseteq I_n$ whenever $m \leq n$. Hence, we can define the ideal $I = \bigcup_{n\in\mathbb{N}} I_n$ of the ring $R = \mathbb{K}[x_{i,j} : i, j \in \mathbb{N}]$. From its definition, the ideal I is $\mathfrak{S}_\infty$–invariant and is generated by the $\mathfrak{S}_\infty$–orbits of G, that is

$$I = \langle x_{1,1}x_{2,2} - x_{1,2}x_{2,1},\ x_{1,1}x_{2,3} - x_{1,3}x_{2,1},\ x_{1,2}x_{3,4} - x_{1,4}x_{3,2}\rangle_{\mathfrak{S}_\infty}.$$

Example 1.3.5 ([HS12], Example 3.8)**.** Let $X = \{x_{i,j} : i, j \in \mathbb{N}\}$ and $R = \mathbb{K}[X]$. Assume that $\mathfrak{S}_\infty$ acts on the variables of X through $\sigma(x_{i,j}) = x_{\sigma(i),\sigma(j)}$ for any $\sigma \in \mathfrak{S}_\infty$ and any $i, j \in \mathbb{N}$. The ideal

$$I = \langle x_{1,1}, x_{1,2}x_{2,1}, x_{1,2}x_{2,3}x_{3,1}, \dots \rangle \subseteq R$$

is $\mathfrak{S}_\infty$–invariant but is not $\mathfrak{S}_\infty$–finitely generated.

Whenever a Π–invariant ideal is Π–finitely generated, it makes sense to look for Noetherianity type results, where by Noetherianity here we mean the following.

Definition 1.3.6. A polynomial ring in infinitely many variables is **Π–Noetherian** if every Π–invariant ideal is Π–finitely generated.

We conclude this section with the following definition which implies $\mathfrak{S}_\infty$–finite generation of families of $\mathfrak{S}_\infty$–invariant ideals. This is according to [KKL14, HKL18].

Definition 1.3.7. Let $R = \mathbb{K}[X]$ and consider the element $f \in R$. The **width** of f is the minimal $n \in \mathbb{N}$ such that all the permutations $\sigma \in \mathfrak{S}_\infty$ that fix the set $[n]$ also fix f. In case such n does not exist we say that f has infinite width.

In the next definition we use the width of an element f in the ring $R = \mathbb{K}[X]$ to define the width of an ideal.

Definition 1.3.8. Let I be an $\mathfrak{S}_\infty$–invariant ideal of R. The **n–th truncation** of I is the set

$$I_n := \{f \in I : \text{the width of } f \text{ is at most equal to } n\}.$$

The set I_n is naturally an $\mathfrak{S}_\infty$–invariant ideal of R. We define the **width** of I as the minimal $n \in \mathbb{N}$ such that the n–th truncation determines it up to the action of $\mathfrak{S}_\infty$, that is $I = \langle I_n \rangle_{\mathfrak{S}_\infty}$. In case such natural number exists, we say that I has **finite width**, otherwise it has **infinite width**.

Example 1.3.9. The $\mathfrak{S}_\infty$–invariant ideal I of Example 1.3.4, defined as the limit of a family of ideals that correspond to the minors of a square matrix, has width equal to four.

An important consequence of Definition 1.3.8 is that in order for I to be $\mathfrak{S}_\infty$–finitely generated there must exist sufficiently large $n \in \mathbb{N}$ such that the ideal I_n is generated by the $\mathrm{Sym}(n)$–orbit of a finite set. By Definition 1.3.7, this is the case when I_n is generated by elements in R that have finite width. If we consider the chain of ideals

$$I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq I_{n+1} \subseteq \cdots \subseteq I = \bigcup_{n \in \mathbb{N}} I_n, \tag{1.1}$$

where each of the ideals I_n is a $\mathrm{Sym}(n)$–invariant ideal of a finite dimensional polynomial ring, then the $\mathfrak{S}_\infty$–finite generation of I implies the stabilization of the chain (1.1) at width n. This stabilization implies the Noetherianity of the infinite dimensional polynomial ring R up to the action of the infinite symmetric group. In the next paragraph we connect the study of Noetherianity of infinite dimensional polynomial rings with the theory of Gröbner bases.

1.3.2 Equivariant Gröbner bases

Assume that $R = \mathbb{K}[X]$ is a polynomial ring in infinitely many variables over a field $\mathbb{K}$ and that Π is a monoid acting on R in terms of monoid homomorphisms. If X^* is the free commutative monoid generated by X, that is the monoid of all finite sequences of elements from X with sequence concatenation as operation, then the monomials of R are exactly the elements of X^*.

Definition 1.3.10. A **monomial order** on the set X^* is a well order $\leq$ with the additional property that $u \leq v$ implies $uw \leq vw$ for any $u, v, w \in X^*$.

As mentioned in Remark 1.1.14, Dickson's lemma implies the finite generation of monomial ideals in polynomial rings with finitely many variables. This can be extended to any ideal I in a finite dimensional polynomial ring as follows. Consider a monomial order $\leq$ on the monomials of the ring and let $M = \{m_1, \ldots, m_r\}$ be the set of all monomials in $\{\mathrm{in}_\leq(f) : f \in I\}$ that are minimal with respect to $\leq$. If $f_1, \ldots, f_r$ are polynomials in I with $\mathrm{in}_\leq(f_i) = m_i$ for any $i \in \{1, \ldots, r\}$, then the ideal $\mathrm{in}_\leq(I) = \langle \mathrm{in}_\leq(f_1), \ldots, \mathrm{in}_\leq(f_r)\rangle$ is a finitely generated monomial ideal and the set $\{f_1, \ldots, f_r\}$ is a finite Gröbner basis for I. Moreover, the ideal I is finitely generated. This relation between well orders and finite generation of ideals is significant for Noetherianity results of high dimensional polynomials rings. In order to establish a Gröbner basis theory for ideals in polynomial rings with infinitely many variables we need to make sure that the action of the monoid Π on the elements of R does not violate the property of the monomial order being a well-order. We require the following definition.

Definition 1.3.11. A **Π–compatible order** on the set X^* is a monomial order, denoted $\leq_\Pi$, with the following property

$$u \leq_\Pi v \Rightarrow \pi(u) \leq \pi(v), \quad \forall u, v \in X^*, \forall \pi \in \Pi.$$

It is noted in [BD09, Remark 2.1] that the infinite symmetric group $\mathfrak{S}_\infty$ does not respect monomial orders. Additionally, [LNNR19, Example 2.2] shows that initial ideals of $\mathfrak{S}_\infty$–invariant ideals are not $\mathfrak{S}_\infty$–invariant in general. To overcome this difficulty one introduces the monoid of strictly increasing functions

$$\mathrm{Inc}(\mathbb{N}) = \{\pi : \mathbb{N} \to \mathbb{N} : \pi(i) < \pi(i+1), \ \forall i \in \mathbb{N}\}. \tag{1.2}$$

Although the elements of $\mathrm{Inc}(\mathbb{N})$ are not permutations, the $\mathrm{Inc}(\mathbb{N})$–orbit of a polynomial are naturally contained in the $\mathfrak{S}_\infty$–orbits. Hence an $\mathfrak{S}_\infty$–invariant ideal $I \subseteq R$ is also $\mathrm{Inc}(\mathbb{N})$–invariant. Furthermore, the $\mathfrak{S}_\infty$–orbit of a polynomial is expressed as the union of finitely many $\mathrm{Inc}(\mathbb{N})$–orbits. Precisely, if $f \in R$ has width equal to k, then

$$\mathfrak{S}_\infty \cdot f = \bigcup_{\sigma \in \mathrm{Sym}(k)} \mathrm{Inc}(\mathbb{N})(\sigma \cdot f).$$

This is because the $\mathrm{Sym}(k)$–orbit of f produces polynomials in R that arise after permuting the indices of the variables in f. Between these elements, there is a permutation which gives the minimum possible values to the indices of f which together with the action of $\mathrm{Inc}(\mathbb{N})$ produces all the elements in the $\mathfrak{S}_\infty$–orbit of f. Based on these observations the following holds.

Lemma 1.3.12 ([HKL18])**.** An $\mathfrak{S}_\infty$–invariant ideal is $\mathfrak{S}_\infty$–finitely generated if and only if it is $\mathrm{Inc}(\mathbb{N})$–finitely generated. If the ring R is $\mathrm{Inc}(\mathbb{N})$–Noetheriant, then it is also $\mathfrak{S}_\infty$–Noetherian.

The following definition generalizes Gröbner bases in the case of monoid actions.

Definition 1.3.13. Let I be a Π–invariant ideal and $\leq_\Pi$ be a Π–compatible monomial order. A **Π-Gröbner basis** (or **equivariant Gröbner basis**) of I with respect to $\leq_\Pi$, is a set $G \subseteq I$, such that the Π–orbits of the initial terms in G generates the initial ideal of I, i.e.

$$\mathrm{in}_{\leq_\Pi}(I) = \langle \mathrm{in}_{\leq_\Pi}(\pi(g)) : \pi \in \Pi, g \in G\rangle.$$

Remark 1.3.14. The requirement in Definition 1.3.13 that $\leq_\Pi$ is a Π–compatible monomial order is necessary to make sure that the ideal $\mathrm{in}_{\leq_\Pi}(I)$ is Π–invariant. Then, we have that $\mathrm{in}_{\leq_\Pi}(\pi(g)) = \pi(\mathrm{in}_{\leq_\Pi}(g))$ for any $\pi \in \Pi$ and any polynomial g in R. Moreover, we also have that

$$\begin{aligned}\mathrm{in}_{\leq_\Pi}(I) &= \langle \mathrm{in}_{\leq_\Pi}(\pi(g)) : \pi \in \Pi, g \in G\rangle = \langle \pi(\mathrm{in}_{\leq_\Pi}(g)) : \pi \in \Pi, g \in G\rangle \\ &= \langle \mathrm{in}_{\leq_\Pi}(g) : g \in G\rangle_\Pi = \langle \mathrm{in}_{\leq_\Pi} G\rangle_\Pi,\end{aligned}$$

and therefore $\mathrm{in}_{\leq_\Pi}(I)$ is Π–finitely generated by the initial terms of polynomials in G.

A Π–invariant ideal does not always admit a finitely generated Gröbner basis. Consider for instance the ideal of Example 1.3.5. It is neither $\mathfrak{S}_\infty$–finitely generated nor $\mathrm{Inc}(\mathbb{N})$–finitely generated. Moreover, there does not exist a finite subset of its generating set that suffices to generate it up to symmetry. Therefore the ideal in this example does not posses a finitely generated equivariant Gröbner basis.

The authors of [HS12] formulated a criterion assuring the existence of a finitely generated equivariant Gröbner basis. They did that using a Π–divisibility order that we now define.

Definition 1.3.15. The Π–divisibility order on X^* is the relation $|_\Pi$ such that

$$u \,|_\Pi\, v \iff \exists \pi \in \Pi : \pi(u) \mid v \quad \forall u, v \in X^*.$$

Here $\pi(u) \mid v$ for some $\pi \in \Pi$ means that there exists $w \in X^*$ such that $v = \pi(u)w$.

The Π–divisibility order on X^* is *reflexive*, we have $u \,|_\Pi\, u$ if we set π to be the identity element e_Π of the monoid Π, and *transitive*, for any u, v, w in X^* the relations $u \,|_\Pi\, v$ and $v \,|_\Pi\, w$ imply $u \,|_\Pi\, w$. Indeed, the relation $u \,|_\Pi\, v$ implies that $\pi(u) \mid v$ for some $\pi \in \Pi$, which means that there exists $v' \in X^*$ such that $\pi(u)v' = v$. Similarly, the relation $v \,|_\Pi\, w$, implies that $\sigma(v) \mid w$ for some $\sigma \in \Pi$, which means that there is some $w' \in X^*$ such that $\sigma(v)w' = w$. Combining the above information yields $\sigma(\pi(u)) \mid w$ and therefore $u \,|_\Pi\, w$. Hence, Π–divisibility is a quasi order. The following proposition shows when Π–divisibility is a well quasi order.

Proposition 1.3.16 (Theorem 2.12, [HS12])**.** Every Π–invariant ideal I of R has a finite Π–Gröbner Basis if and only if $(X^*, |_\Pi)$ is a well quasi order.

Remark 1.3.17. The concept of finite equivariant Gröbner bases is used by the authors of [HS12] and [RN17] to show that chains of Π–invariant ideals are finally stable. This in turn shows the Noetherianity of polynomial rings in infinitely many variables.

We conclude this subsection with the following result, which is the main theorem in [HS12]. In order to formulate it, let $X = \{x_{i,j} : i \in [c], j \in \mathbb{N}\}$ and consider the polynomial ring $R = \mathbb{K}[X]$. We assume that the monoid $\mathrm{Inc}(\mathbb{N})$ acts on the variables of R through $\pi(x_{i,j}) = x_{i,\pi(j)}$ for any $\pi \in \mathrm{Inc}(\mathbb{N})$ and any $i \in [c], j \in \mathbb{N}$.

Proposition 1.3.18 (Theorem 3.1, [HS12])**.** The ring R with the $\mathrm{Inc}(\mathbb{N})$–action on its variables introduced above is $\mathrm{Inc}(\mathbb{N})$–Noetherian.

One of the motivations for studying Noetherianity results of high dimensional polynomial rings comes from algebraic statistics and particularly from statistical models that are parametrized by monomial maps [DS95]. The kernels of such maps are toric ideals. In many cases toric ideals that are invariant under the action of the infinite symmetric group have been studied and are proven to be $\mathfrak{S}_\infty$–finitely generated. Examples of $\mathfrak{S}_\infty$–finite generation are the independent set theorem of Hillar and Sullivant [HS12, Theorem 4.7] as well as the following results.

Proposition 1.3.19 (Theorem 2.1, [DLST95])**.**
Consider the following $\mathfrak{S}_\infty$–equivariant monomial map

$$\begin{aligned} \mathbb{K}[y_{\{i,j\}} : i \neq j \in \mathbb{N}] &\to \mathbb{K}[t_i : i \in \mathbb{N}] \\ y_{\{i,j\}} &\mapsto t_i t_j, \end{aligned}$$

where $\mathfrak{S}_\infty$ acts on the variables of $\mathbb{K}[y_{\{i,j\}} : i \neq j \in \mathbb{N}]$ permuting both indices i, j. Then the ideal $\ker(\phi) = \langle y_{\{1,2\}} y_{\{3,4\}} - y_{\{1,4\}} y_{\{2,3\}} \rangle_{\mathfrak{S}_\infty}$.

Proposition 1.3.20 (Proposition 4.1, [KKL14])**.** Consider the following $\mathfrak{S}_\infty$–equivariant monomial map

$$\begin{aligned} \phi : \mathbb{K}[y_{(\alpha_1,\dots,\alpha_k)} : \alpha_1, \dots, \alpha_k \in \mathbb{N} \text{ distinct}] &\to \mathbb{K}[z_{i,j} : i \in [k], j \in \mathbb{N}] \\ y_{(\alpha_1,\dots,\alpha_k)} &\mapsto z_{1,\alpha_1} \cdots z_{k,\alpha_k}, \end{aligned}$$

where $\mathfrak{S}_\infty$ acts on the variables $y_{(\alpha_1,\dots,\alpha_k)}$ by permuting the indices α_i for any $i \in [k]$. Then $\ker(\phi) = \langle y_{1,2} y_{2,3} y_{3,1} - y_{2,1} y_{3,2} y_{1,3}, y_{1,2} y_{3,4} - y_{1,4} y_{3,2} \rangle_{\mathfrak{S}_\infty}$.

A negative result formulated in [DLO06] demonstrates monomial maps whose kernels are not $\mathfrak{S}_\infty$–finitely generated. Nevertheless, the main result in [DEKL13] shows the $\mathfrak{S}_\infty$–finite generation of kernels of monomial maps where the target polynomial ring has variables with at most one index increasing to infinity. The subsequent article [KKL14] provide explicit formulas for the generators of these kernels. These results motivate the study of the polyhedral objects in the next two chapters.

2 Cones up to Symmetry

Noetherianity results of high dimensional polynomial rings are influenced from the study of families of toric varieties many of which emerge from algebraic statistics. The articles [DEKL13, KKL14, HdC16] have established fundamental results for understanding the behavior of these families in the limit. In this chapter we study the polyhedral objects that correspond to infinite dimensional toric varieties and we deduce characterizations of them in the limit.

2.1 Equivariant Toric Varieties

Throughout this section let $\mathbb{K}$ be a field, or any Noetherian ring, and let Y be a set of indeterminates indexed by the set of natural numbers. Consider the polynomial ring $\mathbb{K}[Y]$ with variables the elements in Y and coefficients in $\mathbb{K}$ and assume that $\mathfrak{S}_\infty$ acts on this ring in terms of ring automorphisms, by permuting the variables in Y. For some natural number k consider a second set of indeterminates $Z = \{z_{i,j} : i \in [k], j \in \mathbb{N}\}$ and let $\mathbb{K}[Z]$ be the polynomial ring with variables the elements in Z and coefficients in $\mathbb{K}$. Suppose that an $\mathfrak{S}_\infty$ action on $\mathbb{K}[Z]$ is given by $\sigma \circ z_{i,j} = z_{i,\sigma(j)}$ for any $i \in [k], j \in \mathbb{N}$ and any $\sigma \in \mathfrak{S}_\infty$.

According to the Noetherianity result of Hillar and Sullivant (Proposition 1.3.18) the polynomial ring $\mathbb{K}[Z]$ is always $\mathfrak{S}_\infty$–Noetherian. On the contrary the polynomial ring $\mathbb{K}[Y]$ need not be $\mathfrak{S}_\infty$–Noetherian. Consider for instance the case where $Y = \{y_{i,j} : i \neq j \in \mathbb{N}\}$ and assume an $\mathfrak{S}_\infty$ action on the elements of Y via $\sigma \circ y_{i,j} = y_{\sigma(i),\sigma(j)}$ for any $i \neq j \in \mathbb{N}$ and any $\sigma \in \mathfrak{S}_\infty$. Then Example 1.3.5 and [AH07, Proposition 5.2] verify that $\mathbb{K}[Y]$ is not $\mathfrak{S}_\infty$–Noetherian. The next result shows that infinite dimensional toric ideals are finitely generated up to symmetry.

Proposition 2.1.1. [DEKL13, Theorem 1.1] Suppose that $\mathfrak{S}_\infty$ has finitely many orbits on Y. Let $\phi : \mathbb{K}[Y] \to \mathbb{K}[Z]$ be an $\mathfrak{S}_\infty$–equivariant homomorphism that maps any variable in Y to a monomial in $\mathbb{K}[Z]$. Then the ideal $\ker(\phi)$ is $\mathfrak{S}_\infty$–finitely generated and $\operatorname{im}(\phi) \cong \mathbb{K}[Y]/\ker(\phi)$ is $\mathfrak{S}_\infty$–Noetherian.

In the following let $X = \{x_i : i \in \mathbb{N}\}$ be a set of indeterminates and assume that $\mathfrak{S}_\infty$ acts on X by permuting its variables. Denote by $\mathbb{K}[X]$ the polynomial ring with variables the elements in X and coefficients in $\mathbb{K}$. As a consequence of Proposition 2.1.1 we have the following result.

Proposition 2.1.2. For any natural number k, the sequence of kernels of monomial maps

$$\pi_k : \mathbb{K}[Y] \to \mathbb{K}[X], \quad \pi_k(y_{i_1 \cdots i_k}) = x_{i_1}^{a_1} \cdots x_{i_k}^{a_k} \tag{2.1}$$

where $i_1, \ldots i_k \in \mathbb{N}$ are distinct indices and $a_1, \ldots, a_k \in \mathbb{N}$ stabilizes up to symmetry.

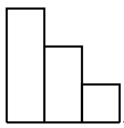
Remark 2.1.3. Proposition 2.1.2 was first stated and proven in [AH07] for the square-free case, that is, when $a_1 = \cdots = a_k = 1$, while the more general statement was formulated as a conjecture.

A key result in algebraic statistics due to Diaconis and Sturmfels states that Markov bases are the exponent vectors of the generators of toric ideals (see [DS95, Theorem 3.1] for more details). Since the kernels of monomial maps are toric ideals, the last two propositions can be reformulated to that Markov bases corresponding to infinite dimensional $\mathfrak{S}_\infty$–equivariant monomial maps have a finite number of generators up to symmetry. The authors of [KKL14] showed that finite equivariant Markov bases and finite equivariant lattice generating sets exist for monomial maps of the form

$$\pi_2 : \mathbb{K}[Y] \to \mathbb{K}[X], \quad y_{i_1 i_2} \mapsto x_{i_1}^{a_1} x_{i_2}^{a_2}$$

for any distinct $i_1, i_2 \in \mathbb{N}$ and any $a_1, a_2 \in \mathbb{N}$ with $a_1 > a_2$ and $\gcd(a_1, a_2) = 1$. A useful tool for their computations was the representation of the monomials in the right hand-side of the map π_k in (2.1) in terms of *box piles*. Every monomial in $\mathbb{K}[X]$ can be specified by the exponents of the variables it contains. For instance the monomial x_i^a is represented by a column of height a in position i of some diagram.

Example 2.1.4. The box pile representation of the monomial $x_1^3 x_2^2 x_3$ is the following



It consists of three boxes, one of height three, one of height two and one of height one.

When a finite or infinite symmetric group acts on the monomials in $\mathbb{K}[X]$ by permuting the variables then the order of the columns in the box pile representation is irrelevant. Therefore we may assume that the exponents on the right hand-side of (2.1) are ordered as $a_1 > a_2 > \cdots > a_k$ for some $k \in \mathbb{N}$.

Definition 2.1.5. For $k, n \in \mathbb{N}$ with $n \geq k$ let $a = (a_1, \dots, a_k, 0, \dots, 0) \in \mathbb{Z}^n_{\geq 0}$ be the exponent vector of a monomial $x^a \in \mathbb{K}[X]$. Assume that the non-zero entries of a are coprime and are ordered as $a_1 > \cdots > a_k$. A **(k,n)–box pile generator** is any vector arising from a by permuting its coordinates.

In the following section we study cones generated by the set of all (k,n)–box pile generators arising from a vector a. We are interested in understanding the behavior of these cones in the limit.

2.2 Box Pile Cones

For $n, k \in \mathbb{N}, n \leq k$, let $a = (a_1, \dots, a_k, 0, \dots, 0) \in \mathbb{Z}^n_{\geq 0}$ and denote by a^σ the (k,n)–box pile generators obtained by applying the permutation $\sigma \in \mathrm{Sym}(n)$ on a.

Definition 2.2.1. The **(k,n)–box pile cone** is the conic hull of the set of all (k,n)–box pile generators. If we denote by $C_{(k,n)}$ the (k,n)–box pile cone, then

$$\begin{aligned} C_{(k,n)} &:= \mathrm{cone}(a^\sigma : \sigma \in \mathrm{Sym}(n)) \\ &= \left\{ \sum_{\sigma \in \mathrm{Sym}(n)} \sum_{i=1}^{s} \lambda_i a^\sigma : \lambda_i \in \mathbb{R}_{\geq 0}, i = 1, \dots, s \right\} \subseteq \mathbb{R}^n \end{aligned}$$

where s is the multinomial coefficient $\frac{n!}{(n-k)!}$.

Example 2.2.2. For $k = 2$ and $n = 3$, the $(2,3)$–box pile generator consists of the vectors $(a_1, a_2, 0), (a_1, 0, a_2), (a_2, a_1, 0), (a_2, 0, a_1), (0, a_1, a_2), (0, a_2, a_1)$ and the $(2,3)$–box pile cone $C_{(2,3)} \subset \mathbb{R}^3$ is the convex cone that is generated by these six vectors.

Remark 2.2.3. Suppose that we start with a pile of k boxes any pair of them having distinct heights. Then the box pile cone $C_{(k,k+1)} \subseteq \mathbb{R}^{k+1}$ is a cone over a k–dimensional permutahedron. In the more general case where some of the box heights coincide or when $n > k+1$ then the box pile cone is affinely isomorphic to a cone over the convex hull of the $\mathrm{Sym}(n)$–orbits of a vector in $\mathbb{R}^n$ with some repeated entries.

For $k = 1$ and any natural number n, the $(1,n)$–box pile cone is the cone generated by the standard unit vectors in $\mathbb{Z}^n$. Such a cone is trivially determined by inequalities of the form $x_i \geq 0$ for any $i \in [n]$. The first non-trivial case of box pile cones appears when $k = 2$ and n is any natural number. In this case the inequality description of the $(2,n)$–box pile cone is due to the following result.

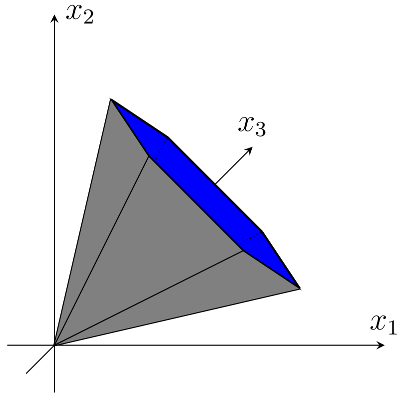


Figure 2.1: The cone $C_{(2,3)} \subseteq \mathbb{R}^3$ generated by the set of $(2,3)$−box pile generators for the vector $(2,1,0)$.

Proposition 2.2.4. The following hyperplanes cut out the $(2,n)$−box pile cone $C_{(2,n)}$

$$\begin{aligned} x_i &\geq 0, \quad \text{for any } i \in [n], \\ a_1(x_1 + \dots + \hat{x}_i + \dots + x_n) &\geq a_2 x_i, \quad \text{for any } i \in [n], \end{aligned} \tag{2.2}$$

where $\hat{x}_i$ means that x_i is omitted from the summation.

Proof. Let $\widetilde{C}$ be the cone generated by the hyperplanes of the statement. Equivalently, $\widetilde{C} = \operatorname{cone}(\operatorname{Sym}(n) \circ \{w_1, w_2\})$ where $w_1 = (1, 0, \dots, 0)$ and $w_2 = (a_1, \dots, a_1, -a_2)$. We will show that $\widetilde{C} = C^*_{(2,n)}$, where $C^*_{(2,n)}$ is the dual of $C_{(2,n)}$. We first observe that $\widetilde{C} \subset C^*_{(2,n)}$ since for any $x \in \widetilde{C}$ the inequalities in (2.2) are valid on $C_{(2,n)}$. For the other direction, assume $x \in C^*_{(2,n)}$. Then, by the definition of the dual cone we have that $\langle x, y \rangle \geq 0$ for any $y \in C_{(2,n)}$. If y is an extreme ray of $C_{(2,n)}$, then since the extreme rays of $C_{(2,n)}$ are permutations of the vector $(a_1, a_2, 0, \dots, 0)$ the last yields

$$a_1 x_i + a_2 x_j \geq 0, \quad \text{for any } i, j \in [n], i \neq j.$$

Since a_1, a_2 are strictly positive numbers the last inequality is equivalent to

$$\frac{a_1}{a_2} x_i + x_j \geq 0, \quad \text{for any } i, j \in [n], i \neq j. \tag{2.3}$$

Without loss of generality we may assume that the coordinates of x can be ordered as follows

$$x_1 \geq x_2 \geq \dots \geq x_n.$$

Equation (2.3) is valid when $x \geq 0$. If all the entries of the vector x are positive, i.e. if $x_i \geq 0$ for any $i \in [n]$, then $x \in \operatorname{cone}(\operatorname{Sym}(n) \circ w_1)$. Therefore we may assume that

$x < 0$. In that case equation (2.3) is valid if the vector x has at most one negative entry. Suppose that $x_n < 0$. To prove that $x \in \widetilde{C}$ we need to show that x can be written as a linear combination of w_1, w_2 and their permutations with non-negative scalars. We have that

$$x = \left(x_1 + \frac{a_1}{a_2}x_n\right)w_1^{(1)} + \cdots + \left(x_{n-1} + \frac{a_1}{a_2}x_n\right)w_1^{(n-1)} + \left(-\frac{1}{a_2}x_n\right)w_2, \tag{2.4}$$

where $w_1^{(i)}$ denotes the vector permutation of w_1 that has a 1 in the i–th coordinate and zeros elsewhere. According to equation (2.3) the coefficients $x_i + \frac{a_1}{a_2}x_n$ are non-negative for any $i \in [n-1]$. Moreover from the assumption that $x_n < 0$ we have that $-\frac{a_1}{a_2}x_n > 0$. The proposition follows by expanding the right hand-side of equation (2.4) and verifying that this linear combination gives the vector x. □

Proposition 2.2.4 shows that when we embed the cone $C_{(2,n)} \subseteq \mathbb{R}^n$ into higher dimensions then, while new inequalities appear, the combinatorial types (2.2) remain fixed. In this thesis, we refer to this phenomenon as the *combinatorial stabilization* of the family $\{C_{(2,n)} : n \geq 2\}$ of $(2,n)$–box pile cones. The same phenomenon is observed for the family of (k,n)–box pile cones when $k > 2$. From computations using the mathematical software *Sage* we conjecture the following.

Conjecture 2.2.5. The following hyperplanes cut out the (k,n)–box pile cone $C_{(k,n)}$

$$\begin{aligned} x_i \geq 0, &\quad \forall i \in [n], \\ \left(\sum_{i=1}^{l} a_i\right)\left(\sum_{j=1}^{n} x_j \backslash \sum_{m=1}^{l} x_m\right) \geq \left(\sum_{i=l+1}^{k} a_i\right)\left(\sum_{m=1}^{l} x_m\right), &\quad \forall l \in [k-1] \end{aligned} \tag{2.5}$$

where $\sum_{j=1}^{n} x_j \backslash \sum_{m=1}^{l} x_m$ means that exactly l summands are omitted from $\sum_{j=1}^{n} x_j$, for any $l \in [k-1]$.
In the special case $k = 1$, the hyperplane $x_i \geq 0$ for any $i \in [n]$ cuts out the cone $C_{(k,n)}$, while if $k = n$, then the hyperplanes $\left(\sum_{i=1}^{l} a_i\right)\left(\sum_{j=1}^{n} x_j \backslash \sum_{m=1}^{l} x_m\right) \geq \left(\sum_{i=l+1}^{n} a_i\right)\left(\sum_{m=1}^{l} x_m\right)$ for any $l \in [n-1]$ cut out $C_{(n,n)}$.

We close this section with the following remark.

Remark 2.2.6. The stabilization results described in Section 1.3 require that the elements of an infinite dimensional polynomial ring have finite width. This is very different from the combinatorial stabilization described before, to that the width of vectors required for the inequality description is infinite, however stabilization follows from the combinatorics of the cones.

2.3 Symmetrized matrix cones

So far we have studied cones generated by box pile generators, that is cones generated by the $\mathfrak{S}_\infty$–orbit of a vector that has a finite number of non-zero entries and infinitely many zero entries. Suppose now that we replace the non-zero integer entries in the box pile generators with finite dimensional integer column vectors and similarly with the zero entries. That way we obtain matrices with a finite number of non-zero columns and infinitely many zero columns. In this section we study cones generated by the vectorization of those matrices and we present implicit characterizations of them.

For $k, n \in \mathbb{N}, k \leq n$, consider the matrix

$$\mathcal{A}_{k,n} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,k} & 0 & \cdots & 0 \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,k} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,k} & 0 & \cdots & 0 \end{bmatrix} \in \mathrm{Mat}_{m,n}(\mathbb{R}), \quad m \in \mathbb{N}, m \leq k.$$

Assume that the symmetric group $\mathrm{Sym}(n)$ acts on $\mathcal{A}_{k,n}$ by permuting its columns, i.e.

$$\mathcal{A}^{\sigma}_{k,n} = (\alpha_{i,\sigma(j)})_{i\in[m],j\in[n]} \in \mathrm{Mat}_{m,n}(\mathbb{R}), \quad \forall \sigma \in \mathrm{Sym}(n).$$

As a result of this action we obtain $\frac{n!}{(n-k)!}$ matrices. We would like to study the cone generated by the *vectorization* of those matrices. Recall from linear algebra that the vectorization of a matrix is a linear transformation that converts the matrix $X \in \mathrm{Mat}_{m,n}(\mathbb{R})$ into a vector $\mathrm{vect}(X) \in \mathbb{R}^{mn}$.

Example 2.3.1. For $m = k = 2, n = 3$ we consider the matrix

$$\mathcal{A}_{2,3} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & 0 \\ \alpha_{2,1} & \alpha_{2,2} & 0 \end{bmatrix} \in \mathrm{Mat}_{2,3}(\mathbb{R}).$$

The group $\mathrm{Sym}(3)$ acts on $\mathcal{A}_{2,3}$ by permuting its columns and producing the following six matrices

$$\begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & 0 \\ \alpha_{2,1} & \alpha_{2,2} & 0 \end{bmatrix}, \begin{bmatrix} \alpha_{1,1} & 0 & \alpha_{1,2} \\ \alpha_{2,1} & 0 & \alpha_{2,2} \end{bmatrix}, \begin{bmatrix} \alpha_{1,2} & \alpha_{1,1} & 0 \\ \alpha_{2,2} & \alpha_{2,1} & 0 \end{bmatrix}, \begin{bmatrix} \alpha_{1,2} & 0 & \alpha_{1,1} \\ \alpha_{2,2} & 0 & \alpha_{2,1} \end{bmatrix}, \begin{bmatrix} 0 & \alpha_{1,1} & \alpha_{1,2} \\ 0 & \alpha_{2,1} & \alpha_{2,2} \end{bmatrix}, \begin{bmatrix} 0 & \alpha_{1,2} & \alpha_{1,1} \\ 0 & \alpha_{2,2} & \alpha_{2,1} \end{bmatrix}.$$

The vectorization of the above matrices consists of the following six dimensional vectors

$$\begin{aligned} &(\alpha_{1,1}, \alpha_{1,2}, 0, \alpha_{2,1}, \alpha_{2,2}, 0), (\alpha_{1,1}, 0, \alpha_{1,2}, \alpha_{2,1}, 0, \alpha_{2,2}), (\alpha_{1,2}, \alpha_{1,1}, 0, \alpha_{2,2}, \alpha_{2,1}, 0) \\ &(\alpha_{1,2}, 0, \alpha_{1,1}, \alpha_{2,2}, 0, \alpha_{2,1}), (0, \alpha_{1,1}, \alpha_{1,2}, 0, \alpha_{2,1}, \alpha_{2,2}), (0, \alpha_{1,2}, \alpha_{1,1}, 0, \alpha_{2,2}, \alpha_{2,1}). \end{aligned} \tag{2.6}$$

Definition 2.3.2. The **(k,n)–symmetrized matrix cone** of $\mathcal{A}_{k,n} \in \mathrm{Mat}_{m,n}(\mathbb{R})$, denoted $C_{\mathcal{A}_{k,n}}$, is the convex cone generated by the vectorization of the $\mathrm{Sym}(n)$–orbit of $\mathcal{A}_{k,n}$, that is,

$$\begin{aligned} C_{\mathcal{A}_{k,n}} &= \mathrm{cone}(\mathrm{vect}(\mathcal{A}_{k,n}^{\sigma}) : \sigma \in \mathrm{Sym}(n)) \\ &= \left\{ \sum_{\sigma \in \mathrm{Sym}(n)} \sum_{i=1}^{s} \lambda_i \mathrm{vect}(\mathcal{A}_{k,n}^{\sigma}) : \lambda_1, \dots, \lambda_s \in \mathbb{R}_{\geq 0},\ s = \frac{n!}{(n-k)!} \right\} \subseteq \mathbb{R}^{mn}. \end{aligned}$$

Example 2.3.3. The $(2,3)$–symmetrized matrix cone $C_{\mathcal{A}_{2,3}}$ of the matrix $\mathcal{A}_{2,3}$ in Example 2.3.1 is the convex cone generated by the six vectors in (2.6).

Remark 2.3.4. In case $m = 1$, then $\mathcal{A}_{k,n}$ is a matrix with just one row and is therefore a vector. In this case, the cone $C_{\mathcal{A}_{k,n}}$ coincides with the (k,n)–box pile cone $C_{(k,n)}$ and it therefore admits the characterization in Proposition 2.2.4 and Conjecture 2.2.5.

In the following set $m = k$. In order to characterize the cone $C_{\mathcal{A}_{k,n}}$ we distinguish between different cases regarding the rank of the matrix $\mathcal{A}_{k,n}$. Assume first that $\mathrm{rank}(\mathcal{A}_{k,n}) = 1$. Then the rows of $\mathcal{A}_{k,n}$ are linearly dependent and we can write $\mathcal{A}_{k,n}$ in the form

$$\mathcal{A}_{k,n} = \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,k} & 0 & \cdots & 0 \\ \lambda_1\alpha_{1,1} & \cdots & \lambda_1\alpha_{1,k} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{k-1}\alpha_{1,1} & \cdots & \lambda_{k-1}\alpha_{1,k} & 0 & \cdots & 0 \end{bmatrix} \in \mathrm{Mat}_{k,n}(\mathbb{R}). \tag{2.7}$$

where $\lambda_1, \dots, \lambda_{k-1} \in \mathbb{N}$. We have the following result.

Lemma 2.3.5. Let $C = \mathrm{cone}(v_1, \dots, v_s) \subseteq \mathbb{R}^n$ be a convex cone defined by the inequalities $a_1, \dots, a_k \in (\mathbb{R}^n)^*$ for some $n \in \mathbb{N}$. If

$$\widetilde{C} = \mathrm{cone}((v_1, \lambda_1 v_1, \dots, \lambda_{k-1} v_1), \dots, (v_s, \lambda_1 v_s, \dots, \lambda_{k-1} v_s) : \lambda_i \in \mathbb{N}\ \forall i \in [k-1]) \subseteq \mathbb{R}^{kn},$$

then $\widetilde{C}$ is defined by lifted inequalities $a_1, \dots, a_k \in (\mathbb{R}^{kn})^*$ and equations $\lambda_{k-1} x_i = x_{(k-1)n+i}$, $\lambda_{k-1} x_{ln+i} = \lambda_{l+1} x_{(k-1)n+i}$ for any $i \in [n]$ and any $l \in [k-2]$.

By lifted inequalities in Lemma 2.3.5 we mean that the variables x_i in the inequalities defining the cone $C \subseteq \mathbb{R}^n$ are substituted by variables $x_{(k-1)n+i}$ for any $i \in [n]$.

Proof. For $k = 2$, we show that $\widetilde{C} = \mathrm{cone}((v_1, \lambda v_1), \dots, (v_s, \lambda v_s) : \lambda \in \mathbb{N}) \subseteq \mathbb{R}^{2n}$ is defined by lifted inequalities $a_1, \dots, a_k \in (\mathbb{R}^{2n})^*$ and equations $\lambda x_i = x_{n+i}$ for any $i \in [n]$. For any $j \in [s]$ and some $\lambda \in \mathbb{N}$, the vector $(v_j, \lambda v_j) \in \mathbb{R}^{2n}$ is obtained by vectorizing the product $\begin{pmatrix} 1 \\ \lambda \end{pmatrix} v_j = \begin{pmatrix} v_j \\ \lambda v_j \end{pmatrix}$. In order to describe the space of linear conditions valid on the cone $\widetilde{C}$ it is enough to compute the kernel of the matrix $\begin{pmatrix}1\\ \lambda\end{pmatrix}$ which consists of the vector $z = (\lambda, -1)$. It follows that for $x \in (\mathbb{R}^{2n})^*$ we have that

$$(\lambda, -1) \cdot \begin{pmatrix} x_1 & \cdots & x_n \\ x_{n+1} & \cdots & x_{2n} \end{pmatrix} = 0 \iff \lambda x_i = x_{n+i} \quad \forall i \in [n].$$

Regarding the inequality description of $\widetilde{C}$, this follows from the description of the dual cone $(\widetilde{C})^*$, that is from the set

$$(\widetilde{C})^* = \{x \in (\mathbb{R}^{2n})^* : \langle x, y\rangle \geq 0, \ \forall y \in \widetilde{C}\}.$$

Any $y \in \widetilde{C}$ is written as a sum $y = \sum_{i=1}^{s} k_i(v_i, \lambda v_i)$, where $k_i \in \mathbb{R}_+$ for any $i \in [s]$. If we set $v_i = (\alpha_{i1}, \ldots, \alpha_{in}) \in \mathbb{R}^n$, then $(v_i, \lambda v_i) = (\alpha_{i1}, \ldots, \alpha_{in}, \lambda\alpha_{i1}, \ldots, \lambda\alpha_{in}) \in \mathbb{R}^{2n}$ for any $i \in [s]$. It follows that for $x \in (\mathbb{R}^{2n})^*$ we have

$$\begin{aligned} \langle x, y\rangle \geq 0 &\iff \sum_{i=1}^{s} k_i(x_1\alpha_{i,1} + \cdots + x_n\alpha_{i,n} + \lambda\alpha_{i,1}x_{n+1} + \cdots + \lambda\alpha_{i,n}x_{2n}) \geq 0 \\ &\iff \sum_{i=1}^{s} k_i\left(\frac{\alpha_{i,1}}{\lambda}x_{n+1} + \cdots + \frac{\alpha_{i,n}}{\lambda}x_{2n} + \lambda\alpha_{i,1}x_{n+1} + \lambda\alpha_{i,n}x_{2n}\right) \geq 0 \\ &\iff \sum_{i=1}^{s} k_i\left(\alpha_{i,1}\left(\frac{1+\lambda^2}{\lambda}\right)x_{n+1} + \cdots + \alpha_{i,n}\left(\frac{1+\lambda^2}{\lambda}\right)x_{2n}\right) \geq 0 \\ &\iff \sum_{i=1}^{s} k_i(\alpha_{i,1}x_{n+1} + \cdots + \alpha_{i,n}x_{2n}) \geq 0 \end{aligned}$$

for any $i \in [n]$. Notice here that the set of points $x \in (\mathbb{R}^{2n})^*$ satisfying the inequality $\sum_{i=1}^{s} k_i(\alpha_{i,1}x_{n+1} + \cdots + \alpha_{i,n}x_{2n}) \geq 0$, is exactly the set of points in the dual of the cone C. We conclude that the cone $\widetilde{C}$ is defined by the lifted inequalities defining C and the equations $\lambda x_i = x_{n+i}$ for any $i \in [n]$.
Following exactly the same lines in the proof for the case $k = 2$, one can show the more general case. For the space of linear conditions on the cone

$$\widetilde{C} = \text{cone}((v_1, \lambda_1 v_1, \ldots, \lambda_{k-1}v_1), \ldots, (v_s, \lambda_1 v_s, \ldots, \lambda_{k-1}v_s) : \lambda_i \in \mathbb{N} \ \forall i \in [k-1]) \subseteq \mathbb{R}^{kn}$$

one needs to solve the system $z \cdot X = 0$, where z is any vector in the kernel of $(1, \lambda_1, \ldots, \lambda_{k-1})^{\mathrm{T}}$ and $X \in \mathrm{Mat}_{k,n}(\mathbb{R})$. We note that an element in the i–th row and j–th column of X is of the form $x_{(i-1)n+j}$ for any $i \in [k], k \geq 2$, and any $j \in [n]$. The inequality description for $\widetilde{C}$ follows from the description of the dual cone $(\widetilde{C})^*$ and by appropriately substituting the values $x_{(i-1)n+j}$ for any $i \in [k-1], j \in [n]$ in the inner product $\langle x, y\rangle \geq 0$, where $x \in (\mathbb{R}^{kn})^*$ and y is any point in $\widetilde{C}$, using the equations defining $\widetilde{C}$. □

Substituting $k = 2$ in (2.7) yields

$$\mathcal{A}_{2,n} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & 0 & \dots & 0 \\ \lambda\alpha_{1,1} & \lambda\alpha_{1,2} & 0 & \dots & 0 \end{bmatrix}, \tag{2.8}$$

where $\lambda \in \mathbb{N}$, and we have the following description for the cone $C_{\mathcal{A}_{2,n}}$.

Proposition 2.3.6. The following hyperplanes and halfspaces define the cone $C_{\mathcal{A}_{2,n}}$ where $\mathcal{A}_{2,n}$ is the rank one matrix in (2.8)

$$\begin{aligned} x_{n+j} &\geq 0, && \text{for any } j \in [n], \\ \alpha_{1,1}(x_{n+1} + \cdots + \hat{x}_{n+j} + \cdots + x_{2n}) &\geq \alpha_{1,2}x_{n+j}, && \text{for any } j \in [n], \\ \lambda x_j - x_{n+j} &= 0, && \text{for any } j \in [n], \end{aligned} \tag{2.9}$$

where $\hat{x}_{n+j}$ means that x_{n+j} is omitted from the summation.

Proof. The proposition follows if we specialize $C = C_{(2,n)} \subseteq \mathbb{R}^n$ and $\widetilde{C} = C_{\mathcal{A}_{2,n}} \subseteq \mathbb{R}^{2n}$, in Lemma 2.3.5, where C is the $(2,n)$−box pile cone of Proposition 2.2.4. □

Now we use Lemma 2.3.5 with the setup that C is the (k,n)−box pile cone $C_{(k,n)}$ whose description was conjectured in Conjecture 2.2.5, and $\widetilde{C}$ is the (k,n)−symmetrized matrix cone $C_{\mathcal{A}_{k,n}}$ for the rank one matrix (2.7) to obtain the following characterization.

Conjecture 2.3.7. The following hyperplanes and halfspaces define the cone $C_{\mathcal{A}_{k,n}}$, where $\mathcal{A}_{k,n}$ is the rank one matrix in (2.7)

$$\begin{aligned} x_{n+j} &\geq 0, && \text{for all } j \in [n], \\ \left(\sum_{i=1}^{s} \alpha_{1,i}\right)\left(\sum_{j=1}^{n} x_{(k-1)n+j} \setminus \sum_{m=1}^{s} x_{(k-1)n+m}\right) &\geq \left(\sum_{i=s+1}^{k} \alpha_{1,i}\right)\left(\sum_{m=1}^{s} x_{(k-1)n+m}\right), && \text{for all } s \in [k-1], \end{aligned}$$

where $\sum_{j=1}^{n} x_{(k-1)n+j} \setminus \sum_{m=1}^{s} x_{(k-1)n+m}$ means that exactly s summands are omitted from $\sum_{j=1}^{n} x_{(k-1)n+j}$ for any $s \in [k-1]$, and $\lambda_{k-1}x_j = x_{(k-1)n+j}, \lambda_{k-1}x_{ln+j} = \lambda_{l+1}x_{(k-1)n+j}$ for any $j \in [n]$ and any $l \in [k-2]$.

Remark 2.3.8. The combinatorial stabilization of box pile cones observed in Section 2.2 and the descriptions of the (k,n)−symmetrized matrix cones of Proposition 2.3.6 and Conjecture 2.3.7 imply the combinatorial stabilization of the cones $C_{\mathcal{A}_{k,n}}$, where $\operatorname{rank}(\mathcal{A}_{k,n}) = 1$ and $n \to \infty$.

We now provide a way of obtaining Proposition 2.3.6 and Conjecture 2.3.7 as tensor products of convex cones. When the matrix $\mathcal{A}_{2,n}$ has rank one, then it can be expressed as the tensor product of two vectors, one having length two and another one having length n. Suppose that $u = (u_1, u_2) \in \mathbb{R}^2$ and $w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$, are such that

$$\mathcal{A}_{2,n} = u \otimes w.$$

We can compute the values of the entries of u and w from $\mathcal{A}_{2,n}$. To be more precise, for any $i \in [2]$ and any $j \in [n]$, the values of u_i and w_j are the solutions of the following system of linear equations

$$u_1 w_1 = \alpha_{1,1},\ u_1 w_2 = \alpha_{1,2},\ u_1 w_3 = 0,\ \ldots,\ u_1 w_n = 0$$
$$u_2 w_1 = \lambda\alpha_{1,1},\ u_2 w_2 = \lambda\alpha_{1,2},\ u_2 w_3 = 0,\ \ldots,\ u_2 w_n = 0.$$

Solving this system, we find that

$$u = (r, \lambda r),\ w = \left(\frac{\alpha_{1,1}}{r}, \frac{\alpha_{1,2}}{r}, 0, \ldots, 0\right) \quad \forall r \in \mathbb{R}, r \neq 0.$$

Without loss of generality set $r = 1$ so that $u = (1, \lambda)$, and $w = (\alpha_{1,1}, \alpha_{1,2}, 0, \ldots, 0)$. We show in Theorem 2.3.12 that the cone $C_{\mathcal{A}_{2,n}}$ is the *tensor product* of the cone generated by the vector u and the box pile cone for the vector w. The following definition is according to [Sch74, MN91, Mul97].

Definition 2.3.9. Let $C_1 \subseteq \mathbb{R}^n, C_2 \subseteq \mathbb{R}^m$ be two convex cones. The **projective tensor product cone** $C_{\text{p}}(C_1, C_2)$ of C_1 and C_2 is

$$C_{\text{p}}(C_1, C_2) = \text{cone}(e \otimes f : e \in C_1, f \in C_2) \subseteq \mathbb{R}^n \otimes \mathbb{R}^m \cong \mathbb{R}^{nm}.$$

The **injective tensor product cone** $C_{\text{i}}(C_1, C_2)$ of C_1 and C_2 is

$$C_{\text{i}}(C_1, C_2) = \{X \in \text{Mat}_{n\times m}(\mathbb{R}) : (\lambda \otimes \mu)X \geq 0, \lambda \in C_1^*, \mu \in C_2^*\} \subseteq (\mathbb{R}^{nm})^*.$$

It follows from the definition of the injective and projective tensor product cone that $C_{\text{i}}(C_1, C_2)$ is the dual of $C_{\text{p}}(C_1, C_2)$.

Remark 2.3.10. Tensor products of convex cones were introduced in [Mul97] during the study of problems related to shape preserving interpolation and approximation.

Lemma 2.3.11. Let C_u be the polyhedral cone generated by the vector $u = (u_1, u_2) \in \mathbb{R}^2_+$. Then, the hyperplane $x_2 \geq 0$ and the halfspace $u_2 x_1 = u_1 x_2$ cut out the cone C_u.

Proof. The kernel of $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ consists of the vector $z = (u_2, -u_1)$. Therefore for any vector $x = (x_1, x_2) \in \mathbb{R}^2$ we have that

$$\langle x, z\rangle = 0 \iff u_2 x_1 - u_1 x_2 = 0$$

which proves the validity of the linear conditions on the cone C_u. Denote by C_u^* the dual of the cone C_u. Then, any point $x \in C_u^*$ satisfies the inequality constraint $\langle y, x\rangle \geq 0$ for any $y \in C_u$, hence

$$u_1 x_1 + u_2 x_2 \geq 0. \tag{2.10}$$

Since the equation $u_2 x_1 - u_1 x_2 = 0$ is valid, then $x_1 = \frac{u_1}{u_2} x_2$. Substituting this to the equation (2.10) yields

$$u_1 x_1 + u_2 x_2 \geq 0 \Rightarrow u_1 \left(\frac{u_1}{u_2} x_2\right) + u_2 x_2 \geq 0 \Rightarrow \frac{u_1^2}{u_2} x_2 + u_2 x_2 \geq 0$$
$$\overset{u_2>0}{\Rightarrow} u_1^2 x_2 + u_2^2 x_2 \geq 0 \Rightarrow (u_1^2 + u_2^2) x_2 \geq 0 \overset{u_1^2+u_2^2>0}{\Rightarrow} x_2 \geq 0$$

which is the desired hyperplane description. □

Theorem 2.3.12. Let $\mathcal{A}_{2,n} \in \mathrm{Mat}_{2,n}(\mathbb{R})$ with $\mathrm{rank}(\mathcal{A}_{2,n}) = 1$. Then

$$C_{\mathcal{A}_{2,n}} = C_{\mathrm{p}}(C_u, C_{(2,n)}),$$

where C_u is the cone of Lemma 2.3.11 for $u = (1, \lambda)$, and $C_{(2,n)}$ is the $(2,n)$−box pile cone of Proposition 2.2.4 for the vector $w = (\alpha_{1,1}, \alpha_{1,2}, 0, \ldots, 0)$. Moreover the description of the equations and inequalities that define the cone $C_{\mathcal{A}_{2,n}}$ arises by tensoring the respective descriptions of the cones C_u and $C_{(2,n)}$.

Proof. For the first part of the theorem observe that the generators of the cone $C_{\mathcal{A}_{2,n}}$ are exactly those vectors obtained by vectorizing the matrices that arise as the tensor product of the vector u that generates the cone C_u with any vector obtained by the action of $\mathrm{Sym}(n)$ on the vector $w = (\alpha_{1,1}, \alpha_{1,2}, 0, \ldots, 0)$. The latest are the generators of the $(2,n)$−box pile cone $C_{(2,n)}$ (see Definition 2.1.5 and Definition 2.2.1). For the second part we need to take a closer look at the equations and inequalities defining the cones C_u and $C_{(2,n)}$. For the inequalities defining $C_{\mathcal{A}_{2,n}}$ we have

$$((0,1) \otimes (0, \ldots, 0, 1, 0, \ldots, 0)) \cdot X \geq 0 \Rightarrow \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} \cdot X \geq 0 \Rightarrow x_{n+j} \geq 0 \ \forall j \in [n],$$

and

$$((0,1) \otimes (\alpha_{1,1}, \ldots, -\alpha_{1,2}, \ldots, \alpha_{1,1})) \cdot X \geq 0 \Rightarrow \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \alpha_{1,1} & \cdots & -\alpha_{1,2} & \cdots & \alpha_{1,1} \end{bmatrix} \cdot X \geq 0$$
$$\Rightarrow \alpha_{1,1}(x_{n+1} + \cdots + \hat{x}_{n+j} + \cdots + x_{2n}) - \alpha_{1,2} x_{2,j} \geq 0$$
$$\Rightarrow \alpha_{1,1}(x_{n+1} + \cdots + \hat{x}_{n+j} + \cdots + x_{2n}) \geq \alpha_{1,2} x_{2,j},$$

for any $j \in [n]$. For the equations defining $C_{\mathcal{A}_{2,n}}$ notice the following. From the equation $\lambda x_1 = x_2$ in the description of the cone C_u, we get two inequalities, namely $\lambda x_1 \geq x_2$

and $\lambda x_1 \le x_2$. The first inequality $\lambda x_1 \ge x_2$ implies that $(\lambda, -1) \cdot x \ge 0$, while the second inequality $\lambda x_1 \le x_2$, implies that $-(\lambda, -1) \cdot x \ge 0$. Hence

$$\begin{aligned}((\lambda, -1) \otimes (0, \dots, 1, \dots, 0)) \cdot X \ge 0 &\Rightarrow \begin{bmatrix} 0 & \cdots & \lambda & \cdots & 0 \\ 0 & \cdots & -1 & \cdots & 0 \end{bmatrix} \cdot X \ge 0 \\ &\Rightarrow \lambda x_j - x_{n+j} \ge 0 \\ &\Rightarrow \lambda x_j \ge x_{n+j} \ \forall j \in [n]\end{aligned}$$

and

$$\begin{aligned}(-(\lambda, -1) \otimes (0, \dots, 1, \dots, 0)) \cdot X \ge 0 &\Rightarrow -((\lambda, -1) \otimes (0, \dots, 1, \dots, 0)) \cdot X \ge 0 \\ &\Rightarrow -\begin{bmatrix} 0 & \cdots & \lambda & \cdots & 0 \\ 0 & \cdots & -1 & \cdots & 0 \end{bmatrix} \cdot X \ge 0 \\ &\Rightarrow \lambda x_j - x_{n+j} \le 0 \\ &\Rightarrow \lambda x_j \le x_{n+j} \ \forall j \in [n].\end{aligned}$$

It therefore follows that $\lambda x_j = x_{n+j}$ for any $j \in [n]$. □

Theorem 2.3.12 can be generalized for any (k, n)−symmetrized matrix cone of a rank one matrix.

Conjecture 2.3.13. Let $\mathcal{A}_{k,n}$ be the rank one matrix in (2.7). The (k, n)−symmetrized matrix cone $C_{\mathcal{A}_{k,n}}$ is obtained as the projective tensor product cone of the (k, n)−box pile cone $C_{(k,n)}$ and the cone generated by the vector $u = (1, \lambda_1, \dots, \lambda_{k-1}) \in \mathbb{R}^k$. Moreover the combinatorial data defining $C_{\mathcal{A}_{k,n}}$ arise by tensoring the respective data of these two cones.

For the rest of the section we step away from the assumption that $\operatorname{rank}(\mathcal{A}_{2,n}) = 1$ and we aim at implicitly characterizing $(2, n)$−symmetrized matrix cones for matrices

$$\mathcal{A}_{2,n} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & 0 & \cdots & 0 \\ \alpha_{2,1} & \alpha_{2,2} & 0 & \cdots & 0 \end{bmatrix} \in \operatorname{Mat}_{2,n}(\mathbb{R})$$

with $\operatorname{rank}(\mathcal{A}_{2,n}) = 2$. We assume that the submatrix

$$\mathcal{A}_2 = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} \in \operatorname{Mat}_{2,2}(\mathbb{R})$$

of $\mathcal{A}_{2,n}$ is invertible. In this case we can rewrite $\mathcal{A}_{2,n}$ as the product

$$\mathcal{A}_{2,n} = \mathcal{A}_2 \cdot \mathcal{I}_{2,n}, \tag{2.11}$$

where $\mathcal{I}_{2,n} \in \mathrm{Mat}_{2,n}(\mathbb{Z})$ is the zero extended identity matrix

$$\mathcal{I}_{2,n} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \in \mathrm{Mat}_{2,n}.$$

When the symmetric group $\mathrm{Sym}(n)$ acts on $\mathcal{A}_{2,n}$ by permuting its columns, the resulting matrices are the same with those obtained by the product of $\mathcal{A}_2$ with any of the matrices in the $\mathrm{Sym}(n)$–orbit on $\mathcal{I}_{2,n}$.

Let $\mathrm{vect}(\mathcal{A}_2) \cdot \mathrm{vect}(\mathcal{I}_{2,n})$ be the vector in $\mathbb{R}^{2n}$ whose first n coordinates arise from the sum

a_1(value of ith coordinate of $\mathrm{vect}(\mathcal{I}_{2,n})$)+$a_2$(value of $(n+i)$th coordinate of $\mathrm{vect}(\mathcal{I}_{2,n})$)

and the rest n coordinates arise from the sum

a_3(value of ith coordinate of $\mathrm{vect}(\mathcal{I}_{2,n})$)+$a_4$(value of $(n+i)$th coordinate of $\mathrm{vect}(\mathcal{I}_{2,n})$),

for any $i \in \{1,\dots,n\}$. The operation defined above is well-defined as it follows from the matrix multiplication. That way, $\mathrm{vect}(\mathcal{A}_{2,n}) = \mathrm{vect}(\mathcal{A}_2) \cdot \mathrm{vect}(\mathcal{I}_{2,n})$, and for any $\sigma \in \mathrm{Sym}(n)$ we have

$$\mathrm{vect}(\mathcal{A}_{2,n}^{\sigma}) = \mathrm{vect}(\mathcal{A}_2) \cdot \mathrm{vect}(\mathcal{I}_{2,n}^{\sigma}).$$

Taking into account this observation we have for the cone $C_{\mathcal{A}_{2,n}}$ the following

$$\begin{aligned} C_{\mathcal{A}_{2,n}} &= \mathrm{cone}(\mathrm{vect}(\mathcal{A}_2) \cdot \mathrm{vect}(\mathcal{I}_{2,n}^{\sigma}) : \sigma \in \mathrm{Sym}(n)) \\ &= \mathrm{vect}(\mathcal{A}_2) \cdot \mathrm{cone}(\mathrm{vect}(\mathcal{I}_{2,n}^{\sigma}) : \sigma \in \mathrm{Sym}(n)) \\ &= \mathrm{vect}(\mathcal{A}_2) \cdot C_{\mathcal{I}_{2,n}}. \end{aligned}$$

For the second equation notice that any $x \in C_{\mathcal{A}_{2,n}}$ is expressed as

$$x = \sum_{\sigma \in \mathrm{Sym}(n)} \lambda_\sigma (\mathrm{vect}(\mathcal{A}_2) \cdot y^\sigma) = \mathrm{vect}(\mathcal{A}_2) \cdot \sum_{\sigma \in \mathrm{Sym}(n)} \lambda_\sigma y^\sigma, \quad \forall y \in \mathrm{vect}(\mathcal{I}_{2,n}),$$

hence $x \in \mathrm{vect}(\mathcal{A}_2) \cdot C_{\mathcal{I}_{2,n}}$, and vice versa. We claim that in order to explicitly describe the cone $C_{\mathcal{A}_{2,n}}$ it is enough to know the explicit description of the cone $C_{\mathcal{I}_{2,n}}$. In particular we claim that the following holds.

Theorem 2.3.14. Let $C^*_{\mathcal{A}_{2,n}}, C^*_{\mathcal{I}_{2,n}}$, be the dual cones of the $(2,n)$–symmetrized matrix cones $C_{\mathcal{A}_{2,n}}$, and $C_{\mathcal{I}_{2,n}}$ respectively. Then,

$$C^*_{\mathcal{A}_{2,n}} = (\mathrm{vect}((\mathcal{A}_2^{\mathrm{T}})^{-1})) \cdot C^*_{\mathcal{I}_{2,n}}.$$

The proof of Theorem 2.3.14 follows from the more general Lemma 2.3.16 while the implicit inequalities and equations for the cone $C_{\mathcal{A}_{2,n}}$ are given by combining Lemma 2.3.16 with the following proposition.

Proposition 2.3.15. The following inequalities and equations define the cone $C_{\mathcal{I}_{2,n}}$.

$$\begin{aligned} x_{1,j} &\geq 0, \quad \forall j \in [n], \\ x_{2,j} &\geq 0, \quad \forall j \in [n], \\ (x_{2,1} + \cdots + \hat{x}_{2,j} + \cdots + x_{2,n}) &\geq x_{1,j}, \quad \forall j \in [n], \\ \text{and,} \quad \sum_{j=1}^{n}(x_{1,j} - x_{2,j}) &= 0. \end{aligned} \tag{2.12}$$

Proof. Let $Y \in \mathrm{Mat}_{2,n}(\mathbb{R})$ and for any generator $\alpha \in \mathbb{R}^{2n}$ of the cone $C_{\mathcal{I}_{2,n}}$ consider the equations $\langle \alpha, \mathrm{vect}(Y) \rangle = 0$. Solving the system of these equations yields

$$y_{1,1} = y_{1,2} = \cdots = y_{1,n}, \; y_{2,1} = y_{2,2} = \cdots = y_{2,n}, \text{ and } y_{1,i} = -y_{2,i} \; \forall i \in \{1, \dots, n\}.$$

and therefore

$$Y = \begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,n} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,n} \end{bmatrix} = y_{1,1} \cdot \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \end{bmatrix},$$

which shows validity of linear conditions on the cone. Let $\widetilde{C}$ be the cone generated by the inequalities and equations of the statement. Equivalently,

$$\widetilde{C} = \mathrm{cone}(\mathrm{vect}(K^\sigma, L^\sigma, M^\sigma, \pm N^\sigma) : \sigma \in \mathrm{Sym}(n)),$$

where $K = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, L = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}, M = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 \end{bmatrix}, N = \begin{bmatrix} 1 & \cdots & 1 \\ -1 & \cdots & -1 \end{bmatrix}$. We will show that $\widetilde{C} = C_{\mathcal{I}_{2,n}}$. The inclusion $\widetilde{C} \subseteq C^*_{\mathcal{I}_{2,n}}$ follows from the observation that each of the generators of $\widetilde{C}$ satisfies the inequality constraint imposed by the dual cone $C^*_{\mathcal{I}_{2,n}}$, and therefore, any point in $\widetilde{C}$, which is by definition a linear combination of the generators with non-negative scalars, also satisfies this inequality constraint. We must therefore show that $C^*_{\mathcal{I}_{2,n}} \subseteq \widetilde{C}$. If $X \in \mathrm{Mat}_{2,n}(\mathbb{R})$ then $\mathrm{vect}(X) \in C^*_{\mathcal{I}_{2,n}}$ if and only if $\langle \alpha, \mathrm{vect}(X) \rangle \geq 0$ holds for any $\alpha \in C_{\mathcal{I}_{2,n}}$. This implies

$$x_{1,j_1} + x_{2,j_2} \geq 0, \qquad \forall j_1 \neq j_2 \in [n]. \tag{2.13}$$

If all the entries of X are non-negative, then

$$\mathrm{vect}(X) = \sum_{j=1}^{n}(x_{1,j}\mathrm{vect}(K^{(j)}) + x_{2,j}\mathrm{vect}(L^{(j)}))$$

hence $\mathrm{vect}(X) \in \mathrm{cone}(\mathrm{vect}(K^\sigma, L^\sigma) : \sigma \in \mathrm{Sym}(n))$. Assume therefore that at most one of the entries of X is negative. If all the entries in the first row of X are positive, then at most one of the $x_{2,j}$'s must be negative. Without loss of generality assume the following ordering on the entries of X

$$x_{2,1} \geq x_{2,2} \geq \cdots \geq x_{2,n} \tag{2.14}$$

and let $x_{2,n} < 0$. Because of the linear equation, we can set

$$X' = X - X_{\min} N,$$

where $X_{\min} := \min_{1 \leq j \leq n} x_{1,j}$ is the minimum value of the entries in the first row of X and is therefore equal to the value of the entry $x_{1,n}$. Then,

$$\begin{aligned} X' = X - X_{\min} N &= \begin{bmatrix} x_{1,1} & \cdots & x_{1,n-1} & x_{1,n} \\ x_{2,1} & \cdots & x_{2,n-1} & x_{2,n} \end{bmatrix} - x_{1,n} \begin{bmatrix} 1 & \cdots & 1 & 1 \\ -1 & \cdots & -1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} x_{1,1} - x_{1,n} & \cdots & x_{1,n-1} - x_{1,n} & 0 \\ x_{2,1} + x_{1,n} & \cdots & x_{2,n-1} + x_{1,n} & x_{2,n} + x_{1,n} \end{bmatrix}. \end{aligned}$$

Since $x_{1,n}$ is the element in the first row of X with minimum value, it follows that the entries $x_{1,1}, x_{1,2}, \ldots, x_{1,n-1}$ have values greater or equal than k, and therefore, $x'_{1,j} = x_{1,j} - k \geq 0$ for any $1 \leq j \leq n-1$. Suppose that the entry $x'_{2,n} = x_{2,n} + x_{1,n}$ is negative. Then we have the following expression for the vectorization $\mathrm{vect}(X')$,

$$\begin{aligned} \mathrm{vect}(X') &= (x'_{2,1} - x'_{2,n-1})\mathrm{vect}(L^{(1)}) + \cdots + (x'_{2,n-2} - x'_{2,n-1})\mathrm{vect}(L^{(n-2)}) \\ &\quad + (x'_{1,1} + x'_{2,n})\mathrm{vect}(K^{(1)}) + \cdots + (x'_{1,n-1} + x'_{2,n})\mathrm{vect}(K^{(n-1)}) \\ &\quad + (x'_{1,n} + x'_{2,n-1})\mathrm{vect}(K^{(n)}) + (x'_{2,n-1} - x'_{2,n})\mathrm{vect}(M^{(n)}) \\ &\quad + (-x'_{2,n})\mathrm{vect}(N), \end{aligned} \tag{2.15}$$

and therefore

$$\begin{aligned} \mathrm{vect}(X) &= (x_{2,1} - x_{2,n-1})\mathrm{vect}(L^{(1)}) + \cdots + (x_{2,n-2} - x_{2,n-1})\mathrm{vect}(L^{(n-2)}) \\ &\quad + (x_{1,1} + x_{2,n})\mathrm{vect}(K^{(1)}) + \cdots + (x_{1,n-1} + x_{2,n})\mathrm{vect}(K^{(n-1)}) \\ &\quad + (x_{1,n} + x_{2,n-1})\mathrm{vect}(K^{(n)}) + (x_{2,n-1} - x_{2,n})\mathrm{vect}(M^{(n)}) \\ &\quad + (x_{1,n} - x_{2,n})\mathrm{vect}(N). \end{aligned}$$

From inequality (2.13), it follows that $x_{1,j} + x_{2,n} \geq 0$ for any $j \in [n-1]$, and also $x_{1,n} + x_{2,n-1} \geq 0$. The ordering (2.14) implies the non-negativity of the coefficients $x_{2,j} - x_{2,n-1}$ for any $j \in [n-2]$, and $x_{2,n-1} - x_{2,n}$. Finally, since $x_{2,n} < 0$, it follows that $x_{1,n} - x_{2,n} > 0$. Hence $\mathrm{vect}(X) \in \mathrm{cone}(K, L, M, N)$.
The proof is very similar in the case where the entry $x'_{2,n}$ is positive. The only difference is in the sign of the last summand in equation (2.15). □

Lemma 2.3.16. Let $C \subseteq \mathbb{R}^n$ be a convex cone with dual C^*. If $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ is any invertible matrix, then

$$(AC)^* = (A^{\mathrm{T}})^{-1}C^*,$$

where $(AC)^*$ is the dual cone of the image of C under A.

Proof. Let $C \subseteq \mathbb{R}^n$ be a cone with dual

$$C^* = \{x \in (\mathbb{R}^n)^* : \langle x, y\rangle \geq 0 \ \forall y \in C\}.$$

For any invertible matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, we would like to characterize the dual cone

$$(AC)^* = \{z \in (\mathbb{R}^n)^* : \langle z, w\rangle \geq 0 \ \forall w \in AC\}.$$

The condition $\forall w \in AC$ can be equivalently written as $w = Ay$ for any $y \in C$, thus,

$$(AC)^* = \{z \in (\mathbb{R}^n)^* : \langle z, Ay\rangle \geq 0 \ \forall y \in C\}.$$

Notice that

$$\langle z, Ay\rangle = z^{\mathrm{T}}Ay = z^{\mathrm{T}}(A^{\mathrm{T}})^{\mathrm{T}}y = (A^{\mathrm{T}}z)^{\mathrm{T}}y = \langle A^{\mathrm{T}}z, y\rangle$$

hence

$$\begin{aligned}(AC)^* &= \{z \in (\mathbb{R}^n)^* : \langle A^{\mathrm{T}}z, y\rangle \geq 0 \ \forall y \in C\} \\ &= \{z \in (\mathbb{R}^n)^* : A^{\mathrm{T}}z \in C^*\}.\end{aligned}$$

Setting $A^{\mathrm{T}}z = x$, it follows that $z = (A^{\mathrm{T}})^{-1}x$ and therefore

$$(AC)^* = \{(A^{\mathrm{T}})^{-1}x \in (\mathbb{R}^n)^* : x \in C^*\} = (A^{\mathrm{T}})^{-1}C^*.$$

□

3 Equivariant Monoids

In this chapter we introduce equivariant monoids, that is commutative monoids modulo a symmetric group action. Our main goal is to study finiteness properties of their underlying algebras and to investigate the behavior of families of equivariant cones in the limit. We start our study by briefly recalling affine monoids and the lemma of Gordan which establishes the connection between affine monoids and finitely generated cones. The main reference here is the book of Bruns and Gubeladze [BG09]. After that we introduce the objects of interest and we extend Gordan's lemma to our setup.

3.1 Affine Monoids

A monoid is a non-empty set M together with a binary operation $+ : M \times M \to M$ which is associative and has an identity element. As mentioned in the introductory part of this chapter, we are mainly interested in monoids that are commutative and are in some sense finitely generated.

Definition 3.1.1. An **affine monoid** M is a finitely generated submonoid of a lattice $\mathbb{Z}^n$ for some $n \in \mathbb{N}$, that is,

- $M \subset \mathbb{Z}^n,\ M + M \subset M,\ 0 \in M$, and,
- there exist $m_1, \ldots, m_r \in M$ such that

$$M = \{\alpha_1 m_1 + \cdots + \alpha_r m_r : \alpha_i \in \mathbb{Z}_{\geq 0},\ \text{for all } i = 1, \ldots, r\}.$$

From now on we use the notation $M = \langle m_1, \ldots, m_r \rangle$ for the commutative monoid $M \subseteq \mathbb{Z}^n$ that is generated by the elements $m_1, \ldots, m_r$.

Example 3.1.2. The monoid $M = \mathbb{Z}_+^2$ is an affine monoid. It is generated by the two standard unit vectors $e_1 = (1, 0), e_2 = (0, 1)$ in $\mathbb{Z}^2$.

Another not so trivial example of an affine monoid is the following.

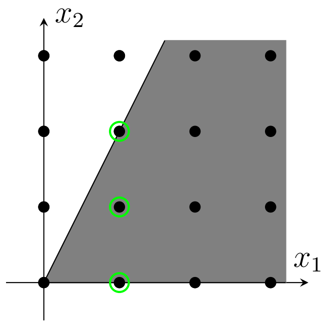


Figure 3.1: The monoid $M = \langle \boldsymbol{x} \in \mathbb{Z}^2 : x_2 \geq 0, 2x_1 \geq x_2 \rangle$. The minimal generators are highlighted with green color.

Example 3.1.3. The monoid $M = \langle \boldsymbol{x} \in \mathbb{Z}^2 : x_2 \geq 0, 2x_1 \geq x_2 \rangle$ is an affine monoid. It is generated by the points $(1,0), (1,1), (2,1)$. This monoid is illustrated in Figure 3.1.

In contrast to the above examples, there are submonoids of $\mathbb{Z}^n$ that are not finitely generated and consequently they are not affine.

Example 3.1.4. Consider the monoid $M = \langle e_1 + ke_2 : k \in \mathbb{N} \rangle$. Then M is a monoid in $\mathbb{Z}^2$. The generating set of M consists of infinitely many elements and it does not have any proper finite subset that generates it. Therefore M is not finitely generated and hence it is not an affine monoid.

Let M be a monoid in $\mathbb{Z}^n$ for some $n \in \mathbb{N}$. Assume that M is generated by $m_1, \ldots, m_n \in \mathbb{Z}^n$. Then the cone associated with M is the cone

$$C(M) = \mathbb{R}_{\geq 0} M = \left\{ \sum_{i=1}^{n} \lambda_i m_i : \lambda_i \in \mathbb{R}_{\geq 0}, \text{ for any } i = 1, \ldots, n \right\}$$

generated by M in $\mathbb{R}^n$. This cone is always a rational polyhedral cone and it is finitely generated whenever the monoid M is so.

Lemma 3.1.5 (Gordan's Lemma)**.** Let $C \subseteq \mathbb{R}^n$ be a finitely generated rational cone. Then $C \cap \mathbb{Z}^n$ is an affine monoid.

We conclude this section with a short reference to monoid algebras and an observation regarding the property of a monoid algebra to be finitely generated. Let $\mathbb{K}$ be a field and $M \subseteq \mathbb{Z}^n$ be a monoid. The *monoid algebra* corresponding to M is denoted $\mathbb{K}[M]$ and is defined as the $\mathbb{K}$−vector space with basis consisting of elements X^m for any $m \in M$. These elements are called the monomials of $\mathbb{K}[M]$. A general element of $\mathbb{K}[M]$ has the form

$$a_1 X^{m_1} + \cdots + a_n X^{m_n}$$

where $a_i \in \mathbb{K}$ and $m_i \in M$ for any $i = 1,\dots,n$. The additive structure of $\mathbb{K}[M]$ is clear while the multiplicative structure arises from the multiplication of monomials $X^m \cdot X^{m'} = X^{m+m'}$ for any $m, m' \in M$.

Example 3.1.6. The monoid algebra corresponding to the monoid $M = \mathbb{Z}^n$ is the Laurent polynomial ring $\mathbb{K}[X_1^{\pm 1},\dots,X_n^{\pm 1}]$, while the monoid algebra for the monoid $M = \mathbb{N}^n$ in the usual polynomial ring in n variables $\mathbb{K}[X_1,\dots,X_n]$.

Let $M = \langle m_1,\dots,m_r\rangle$ be an affine monoid in $\mathbb{Z}^n$. Then the monoid algebra corresponding to M is the set

$$\mathbb{K}[M] = \left\{ \sum_{i=1}^{s} a_i X^{b_i} : s \in \mathbb{N}, a_i \in \mathbb{K}, b_i \in M \text{ for any } i = 1,\dots,s \right\}.$$

We claim that $\mathbb{K}[M]$ is a finitely generated monoid algebra, and particularly that $\mathbb{K}[M] = \mathbb{K}[X^{m_1},\dots,X^{m_r}]$. The inclusion $\mathbb{K}[X^{m_1},\dots,X^{m_r}] \subseteq \mathbb{K}[M]$ is clear because any $f \in \mathbb{K}[X^{m_1},\dots,X^{m_r}]$ is represented as $f = \sum_{i=1}^{r} a_i X^{m_i}$ for any $a_i \in \mathbb{N}$ and any $m_i \in M$ where $i = 1,\dots,r$ and therefore $f \in \mathbb{K}[M]$. For the converse, let $m \in M$ and consider the monomial $f = aX^m \in \mathbb{K}[M]$ for any $a \in \mathbb{K}$. Since M is minimally generated by the $m_1,\dots,m_r$, we can express m as a non-negative linear combination of those minimal generators, i.e. $m = c_1m_1 + \dots + c_rm_r$ for any $c_1,\dots,c_r \in \mathbb{N}$. That way, $f = aX^m = a(X^{m_1})^{c_1}\cdots(X^{m_r})^{c_r}$, hence $f \in \mathbb{K}[X^{m_1},\dots,X^{m_r}]$. This shows that $\mathbb{K}[M] = \mathbb{K}[X^{m_1},\dots,X^{m_r}]$ hence that the monoid algebra $\mathbb{K}[M]$ is finitely generated.

Suppose now that $\mathbb{K}[M]$ is finitely generated as a $\mathbb{K}$–algebra and let $f_1,\dots,f_r$ be a system of generators of $\mathbb{K}[M]$. Then there exists a finite subset $G = \{g_1,\dots,g_r\}$ of M such that $f_i = \sum_{g\in G} r_g X^g$, where $r_g \in \mathbb{K}$. We claim that $M = \mathbb{N}G$, hence M is finitely generated. The inclusion $\mathbb{N}G \subseteq M$ is clear, therefore we need to show that $M \subseteq \mathbb{N}G$. Let $m \in M$, then there exists a polynomial $f \in \mathbb{K}[X_1,\dots,X_n]$ such that $X^m = f(X^{g_1},\dots,X^{g_r})$. Otherwise stated, X^m is a $\mathbb{K}$–linear combination of monomials $(X^{g_1})^{n_1},\dots,(X^{g_r})^{n_r}$ for $n_1,\dots,n_r \in \mathbb{N}$. Since any $\mathbb{K}$–linear combination of the monomials $(X^{g_i})^{n_i}$ can be written as a $\mathbb{K}$–linear combination of monomials X^a with $a \in \mathbb{N}G$, it follows that $X^m \in \mathbb{K}[\mathbb{N}G]$. Then from the definition of $\mathbb{K}[\mathbb{N}G]$ we have that $m \in \mathbb{N}G$, hence $M \subseteq \mathbb{N}G$. This proves the following result.

Proposition 3.1.7. ([BG09, Proposition 2.4]) A monoid $M \subseteq \mathbb{Z}^n$ is finitely generated if and only if the underlying monoid algebra $\mathbb{K}[M]$ is finitely g enerated as a $\mathbb{K}$–algebra.

3.2 Equivariant families of monoids

Let $\{M_n : n \in \mathbb{N}\}$ be a family of monoids, where for each $n \in \mathbb{N}$, M_n is a commutative (not necessarily affine) monoid in $\mathbb{Z}^n$. For any $m, n \in \mathbb{N}, m \leq n$, consider the natural inclusion maps

$$\iota_{m,n} : \mathbb{Z}^m \longrightarrow \mathbb{Z}^n, \quad \boldsymbol{g} \mapsto (\boldsymbol{g}, \boldsymbol{0}).$$

Since M_m is a monoid in $\mathbb{Z}^m$, we can use the map $\iota_{m,n}$ to embed the generators of M_m into $\mathbb{Z}^n$ and we write $\iota_{m,n}(M_m) \subseteq \mathbb{Z}^n$ for the monoid obtained that way. If M_m is an affine monoid in $\mathbb{Z}^m$, then the image $\iota_{m,n}(M_m)$ is an affine monoid in $\mathbb{Z}^n$.

The symmetric group $\mathrm{Sym}(n)$ acts on $\mathbb{Z}^n$ by permuting its coordinates. Hence, if $X \subseteq \mathbb{Z}^m$ is a set, then we denote by

$$\mathrm{Mon}_{m,n}(X) := \langle \mathrm{Sym}(n) \circ \iota_{m,n}(X) \rangle, \quad \forall m \leq n, \tag{3.1}$$

the monoid in $\mathbb{Z}^n$ that is generated by the $\mathrm{Sym}(n)$–orbit of the image $\iota_{m,n}(X)$.

Remark 3.2.1. If M_m is a commutative monoid in $\mathbb{Z}^m$, then $\iota_{m,n}(M_m)$ is a commutative monoid in $\mathbb{Z}^n$ and we define the monoid $\mathrm{Mon}_{m,n}(M_m)$ as in (3.1). Note that the result of acting on $\iota_{m,n}(M_m)$ with the symmetric group $\mathrm{Sym}(n)$ is almost never a monoid. We should therefore consider the monoid closure of the set $\mathrm{Sym}(n) \circ \iota_{m,n}(X)$ when defining the monoid $\mathrm{Mon}_{m,n}(M_m)$.

The same construction can be carried out using the monoid $\mathrm{Inc}(\mathbb{N})$ of strictly increasing functions (see equation (1.2)) and its submonoid

$$\mathrm{Inc}(\mathbb{N})_{m,n} := \{\pi \in \mathrm{Inc}(\mathbb{N}) : \pi(m) \leq n\}, \quad \forall m, n \in \mathbb{N}, m \leq n.$$

Every element $\pi \in \mathrm{Inc}(\mathbb{N})_{m,n}$ can be viewed as a strictly increasing map $\pi : [m] \to [n]$ such that $\pi(i) = j$ for any $i \in [m], j \in [n]$, where $i \leq j$. Now we consider the embedding

$$\iota_\pi : \mathbb{Z}^m \to \mathbb{Z}^n, \quad \boldsymbol{x} \mapsto \boldsymbol{x}'$$

where the j–th coordinate of $\boldsymbol{x}'$ takes value $x_{\pi^{-1}(j)}$ whenever $\pi^{-1}(j) \neq \emptyset$, otherwise it is zero.

Example 3.2.2. If $\boldsymbol{x} = (x_1, x_2) \in \mathbb{Z}^2$, then the coordinates of the vector $\boldsymbol{x}' = \iota_\pi(\boldsymbol{x}) \in \mathbb{Z}^3$ are specified by the values of the elements $\pi \in \mathrm{Inc}(\mathbb{N})_{2,3}$. If $\pi(1) = 1$ then $\pi(2) = 2$ or $\pi(2) = 3$ and if $\pi(1) = 2$ then $\pi(2) = 3$. To each of these elements corresponds a vector $(x_1, x_2, 0), (x_1, 0, x_2)$ and $(0, x_1, x_2)$.

If $X \subset \mathbb{Z}^m$ is a set, then we write

$$\mathrm{Mon}^{\mathrm{Inc}}_{m,n}(X) = \langle \iota_\pi(X) : \pi \in \mathrm{Inc}(\mathbb{N})_{m,n} \rangle, \quad \forall m \leq n, \tag{3.2}$$

for the monoid in $\mathbb{Z}^n$ generated by the set of all the embeddings imposed from the action of $\mathrm{Inc}(\mathbb{N})_{m,n}$ on the elements of X.

Definition 3.2.3. An **$\mathfrak{S}_\infty$–equivariant family of monoids**, is a family $\{M_n : n \in \mathbb{N}\}$ of monoids $M_n \subseteq \mathbb{Z}^n$ such that

$$\mathrm{Mon}_{m,n}(M_m) \subseteq M_n, \quad \text{for any } m,n \in \mathbb{N},\ m \leq n. \tag{3.3}$$

An $\mathfrak{S}_\infty$–equivariant family of monoids $\{M_n : n \in \mathbb{N}\}$ **stabilizes up to symmetry**, if there exists $m \in \mathbb{N}$ such that for any $n \geq m$ the inclusion relation in (3.3) becomes an equality.

The smallest natural number m for which stabilization of an equivariant family of monoids occurs, is called the *stability index* of the family. In this case, we say that the family $\{M_n : n \in \mathbb{N}\}$ *stabilizes at* m.

Remark 3.2.4. As for $\mathfrak{S}_\infty$–equivariant families of monoids, we define analogously $\mathrm{Inc}(\mathbb{N})$–equivariant families of monoids by considering the monoids $\mathrm{Mon}^{\mathrm{Inc}}_{m,n}(M_m)$ instead of the monoids $\mathrm{Mon}_{m,n}(M_m)$ for any $m,n \in \mathbb{N}, m \leq n$ in Definition 3.2.3.

Example 3.2.5. Any finite set $X \subset \mathbb{Z}^n$ defines a $\mathfrak{S}_\infty$–invariant family via $M_m = 0$, whenever $m \leq n$ and $M_r = \mathrm{Mon}_{n,r}(X)$, whenever $r \geq n$. The family $\{M_n : n \in \mathbb{N}\}$ defined that way, stabilizes immediately.

The $\mathrm{Inc}(\mathbb{N})$–orbits are naturally contained in the $\mathfrak{S}_\infty$–orbits and this is true for the $\mathrm{Inc}(\mathbb{N})_{m,n}$–orbits and $\mathrm{Sym}(n)$–orbits as well (see [RN17, Lemma 7.5]). Hence the following holds.

Proposition 3.2.6. A $\mathfrak{S}_\infty$–equivariant family of monoids is also $\mathrm{Inc}(\mathbb{N})$–equivariant.

The main goal of this section is to understand stabilization criteria of equivariant families of monoids. We note here that stabilization is equivalent to finite generation in the limit. In order to write this limit, we use the standard embeddings $\iota_{m,n}$ considered at the beginning of the section to construct the ascending chain of monoids

$$M_1 \subseteq M_2 \subseteq \cdots \subseteq M_n \subseteq M_{n+1} \subseteq \ldots,$$

where for each $n \in \mathbb{N}$, M_n is a commutative monoid in $\mathbb{Z}^n$. Consider the set $M_\infty = \bigcup_{n\in\mathbb{N}} M_n$. Then the set M_∞ equipped with addition operation is a monoid. Indeed, we have that $0 \in M_\infty$ because $0 \in M_1$. Moreover, for any $x,y \in M_\infty$, we have $x+y \in M_\infty$. To see the latest, notice that since $x,y \in M_\infty$, there exist $n_1,n_2 \in \mathbb{N}$ such that $x \in M_{n_1}$ and $y \in M_{n_2}$. Since M_{n_1}, M_{n_2} are monoids in the ascending chain above, then either $M_{n_1} \subseteq M_{n_2}$ or $M_{n_2} \subseteq M_{n_1}$ and there exist an $n \in \mathbb{N}, n = \max\{n_1,n_2\}$ such that $x,y \in M_n$. As M_n is a monoid, it follows that $x+y \in M_n$, hence $x+y \in M_\infty$. Associativity follows from the associativity of the operation of addition.

Definition 3.2.7. Let $\{M_n : n \in \mathbb{N}\}$ be a family of commutative monoids $M_n \subseteq \mathbb{Z}^n$ such that $\iota_{m,n}(M_m) \subseteq M_n$ for any $m, n \in \mathbb{N}, m \le n$. The **limit monoid** of this family is the monoid $M_\infty = \bigcup_{n\in\mathbb{N}} M_n$. This is a monoid in $\bigcup_{n\in\mathbb{N}} \mathbb{Z}^n$.

For any $n \in \mathbb{N}$, the symmetric group $\mathrm{Sym}(n)$ acts on $\mathbb{Z}^n$ by permuting its coordinates and the infinite symmetric group $\mathfrak{S}_\infty$ acts on the set $\bigcup_{n\in\mathbb{N}} \mathbb{Z}^n$ by fixing all but a finite number of permutations. Hence, we assume an action of $\mathfrak{S}_\infty$ on the generators of M_∞ according to the latter.

Definition 3.2.8. The limit monoid M_∞ of an equivariant family of monoids is **$\mathfrak{S}_\infty$–finitely generated**, respectively **Inc($\mathbb{N}$)–finitely generated**, if it is generated by the $\mathfrak{S}_\infty$ orbits, respectively $\mathrm{Inc}(\mathbb{N})$ orbits, on a finite set of generators. In both cases we say that M_∞ is **finitely generated up to symmetry**.

A naive hope would be that every equivariant family of monoids stabilizes up to symmetry. If this is the case the limit monoid M_∞ is finitely generated up to symmetry. One can show the following lemma which states that under the hypothesis that a $\mathfrak{S}_\infty$–equivariant family of monoids is affine, then stabilization is possible.

Lemma 3.2.9. An $\mathfrak{S}_\infty$–equivariant family of affine monoids $\{M_n : n \in \mathbb{N}\}$ stabilizes, if and only if the limit monoid M_∞ is $\mathfrak{S}_\infty$–finitely generated.

Now assume that $\{M_n : n \in \mathbb{N}\}$ is a family of commutative monoids that are not necessarily affine. The following examples show that without the affineness hypothesis, it is not possible to have finite generation in the limit. However, stabilization might follow from the way the family is defined.

Example 3.2.10. Let $M_1 = 0, M_2 = \langle e_1 + ke_2 : k \in \mathbb{N}\rangle$, and define $M_n := \mathrm{Mon}_{2,n}(M_2)$ whenever $n \ge 3$. We can write equivalently $M_n = \langle e_i + ke_j : i \ne j \in [n], k \in \mathbb{N}\rangle$, for any $n \ge 3$. For any $n \ge 3$, the family $\{M_n : n \ge 3\}$ is a $\mathfrak{S}_\infty$–equivariant family of monoids. We have seen in Example 3.1.4 that the monoid M_2 is not affine, hence for each $n \ge 3$, none of the monoids M_n is affine. Therefore, the limit monoid M_∞ cannot be finitely generated up to symmetry. Nevertheless, the family $\{M_n : n \ge 3\}$ stabilizes from the very beginning.

Example 3.2.11. Let $M_1 = 0, M_2 = \langle e_1 + 2e_2\rangle$, and define

$$M_n := \langle \bigcup_{k<n} \mathrm{Mon}^{\mathrm{Inc}}_{k,n}(M_k), e_1 + ne_n\rangle, \quad \forall n > 2.$$

This defines an $\mathrm{Inc}(\mathbb{N})$–equivariant family of monoids $\{M_n : n \in \mathbb{N}\}$, where for any $n \in \mathbb{N}$ each of the monoids M_n is finitely generated. However this family does not stabilize because the elements $e_1 + ne_n$ are new, irredundant generators. As a consequence M_∞ is not finitely generated up to symmetry. For instance for $n = 3$ we have

$$M_3 = \langle \mathrm{Mon}^{\mathrm{Inc}}_{2,3}(M_2), e_1 + 3e_3 \rangle = \langle \underline{e_1 + 2e_2}, \underline{e_1 + 2e_3}, \underline{e_2 + 2e_3}, e_1 + 3e_3 \rangle.$$

Here, the underlined generators are obtained from $\mathrm{Inc}(\mathbb{N})_{2,3}$ acting on M_2 according to (3.2), while the last generator is new. Similarly for $n = 4$, we have the monoid

$$\begin{aligned} M_4 &= \langle \mathrm{Mon}^{\mathrm{Inc}}_{2,4}(M_2), \mathrm{Mon}^{\mathrm{Inc}}_{3,4}(M_3), e_1 + 4e_4 \rangle \\ &= \langle \underline{e_1 + 2e_2}, \underline{e_1 + 2e_3}, \underline{e_1 + 2e_4}, \underline{e_2 + 2e_3}, \underline{e_2 + 2e_4}, \underline{e_3 + 2e_4}, \underline{e_1 + 3e_3}, \underline{e_1 + 3e_4}, \\ &\qquad \underline{e_2 + 3e_4}, e_1 + 4e_4 \rangle. \end{aligned}$$

Like before, the underlined generators in M_4 are elements obtained from the action of $\mathrm{Inc}(\mathbb{N})_{2,4}$ on M_2 and the action of $\mathrm{Inc}(\mathbb{N})_{3,4}$ on M_3, while the generator $e_1 + 4e_4$ is new. In both cases we have containment relations $\mathrm{Mon}^{\mathrm{Inc}}_{2,3}(M_2) \subseteq M_3$ and $\mathrm{Mon}^{\mathrm{Inc}}_{3,4}(M_3) \subseteq M_4$ that make the family $\{M_n : n \in \mathbb{N}\}$ into an $\mathrm{Inc}(\mathbb{N})$–equivariant family of monoids.

Remark 3.2.12. Even if the limit monoid M_∞ of the $\mathrm{Inc}(\mathbb{N})$–equivariant family of monoids given in Example 3.2.11 is not finitely generated up to symmetry, it could still be $\mathfrak{S}_\infty$–finitely generated. However this is not the case. As a consequence we have that while a polynomial ring in infinitely many variables $\mathbb{K}[x_1, x_2, \dots]$ is finitely generated up to symmetry as shown in [HS12, Theorem 3.1], this argument is not true in general for monoid algebras. To be more precise, the monoid algebra $\mathbb{K}[M_\infty]$ generated by the limit monoid M_∞ is not always finitely generated up to symmetry.

3.3 Normal and Saturated Equivariant Monoids

In this section we try to formulate sufficient conditions for invariant families of monoids to stabilize, or equivalently such that the limit monoid M_∞ is finitely generated up to symmetry. The following definition is a generalization of [BG09, Definition 2.21].

Definition 3.3.1. The **normalization** of $\mathrm{Mon}_{m,n}(M_m)$ inside $\mathbb{Z}^n$ is the set

$$\overline{\mathrm{Mon}}_{m,n}(M_m) := \{x \in \mathbb{Z}^n : kx \in \mathrm{Mon}_{m,n}(M_m), k \in \mathbb{N}\}.$$

We say that a family of monoids $\{M_n : n \in \mathbb{N}\}$ is **normal $\mathfrak{S}_\infty$–equivariant** if

$$\overline{\mathrm{Mon}}_{m,n}(M_m) \subseteq M_n \quad \forall m, n \in \mathbb{N}, m \le n. \tag{3.4}$$

A normal $\mathfrak{S}_\infty$–equivariant family of monoids **stabilizes** if there exists $m \in \mathbb{N}$ such that for any $n \ge m$ the inclusion relation (3.4) becomes an equality.

The normalization of a monoid is also called the integral closure of the monoid.

Example 3.3.2. Let $M_1 = 0$ and set $M_n = \mathrm{Mon}_{2,n}(e_1{+}2e_2) = \langle e_i{+}2e_j : i \neq j \in [n]\rangle$ whenever $n \geq 2$. The family of monoids $\{M_n : n \in \mathbb{N}\}$ defined that way is $\mathfrak{S}_\infty$–equivariant but is not normal. For instance, for $n = 3$ we have

$$M_3 = \mathrm{Mon}_{2,3}(e_1 + 2e_2) = \langle e_1 + 2e_2, 2e_1 + e_2, e_1 + 2e_3, 2e_1 + e_3, e_2 + 2e_3, 2e_2 + e_3\rangle$$

and the normalization $\overline{\mathrm{Mon}}_{2,3}(e_1 + 2e_2)$ is given by

$$\overline{\mathrm{Mon}}_{2,3}(e_1 + 2e_2) = \langle e_1 + e_2 + e_3, e_1 + 2e_2, e_1 + 2e_3, e_2 + 2e_1, e_3 + 2e_1, e_2 + 2e_3, e_3 + 2e_2\rangle.$$

The element $e_1 + e_2 + e_3$ is a generator of the normalization of M_3 since $2(e_1 + e_2 + e_3) = (e_1 + 2e_2) + (e_1 + 2e_3)$. Hence $\overline{\mathrm{Mon}}_{2,3}(e_1 + 2e_2) \notin M_3$.

Example 3.3.3. Consider the normal $\mathfrak{S}_\infty$–equivariant family of monoids defined by setting $M_1 = 0$ and $M_n = \overline{\mathrm{Mon}}_{2,n}(e_1 + ke_2 : k \in [n])$ for any $n \geq 2$. For any $n \in \mathbb{N}$ each of the monoids M_n is finitely generated. By computations, we have

$$\begin{aligned}
M_2 &= \langle e_1 + e_2,\ e_1 + 2e_2,\ 2e_1 + e_2\rangle,\\
M_3 &= \langle \mathrm{Mon}_{2,3}(M_2,\ e_1 + 3e_2),\ e_1 + e_2 + e_3\rangle,\\
M_n &= \mathrm{Mon}_{3,n}(M_3,\ e_1 + ne_2)\\
&= \mathrm{Mon}_{3,n}(\mathrm{Mon}_{2,3}(M_2,\ e_1 + 3e_2),\ e_1 + e_2 + e_3,\ e_1 + ne_2)\\
&= \langle \mathrm{Mon}_{2,n}(M_2,\ e_1 + 3e_2),\ e_1 + e_2 + e_3,\ e_1 + ne_2\rangle, \quad \forall n \geq 4.
\end{aligned}$$

We observe that $\mathrm{Mon}_{2,n}(M_2) \subseteq M_n$ for any $n \geq 2$. Hence the family $\{M_n : n \in \mathbb{N}\}$ is an $\mathfrak{S}_\infty$–equivariant family of monoids. This family does not stabilize up to symmetry because for any $n \in \mathbb{N}, n \geq 2$, the elements $e_1 + ne_2$ are irredundant generators. As a consequence, the limit monoid M_∞ is not finitely generated up to symmetry.

The last example motivates for the following result.

Proposition 3.3.4. Any normal $\mathfrak{S}_\infty$–equivariant family of monoids is $\mathfrak{S}_\infty$–equivariant.

Proof. Let $\{M_n : n \in \mathbb{N}\}$ be a normal $\mathfrak{S}_\infty$–equivariant family of monoids. Then by definition $\overline{\mathrm{Mon}}_{m,n}(M_m) \subseteq M_n$ for any $m, n \in \mathbb{N}$ with $m \leq n$. We need to show that $\mathrm{Mon}_{m,n}(M_m) \subseteq \overline{\mathrm{Mon}}_{m,n}(M_m) \subseteq M_n$ for any $m, n \in \mathbb{N}$ with $m \leq n$. This follows by observing that any element $x \in \mathrm{Mon}_{m,n}(M_m)$ is an element in $\mathbb{Z}^n$ which belongs to $\overline{\mathrm{Mon}}_{m,n}(M_m)$ if we choose $k = 1$ in the definition of the normalization. □

Definition 3.3.5. An $\mathfrak{S}_\infty$–equivariant family of monoids $\{M_n : n \in \mathbb{N}\}$ is **saturated** if $M_n = M_\infty \bigcap \mathbb{Z}^n$ for any $n \in \mathbb{N}$. If this equality holds for all but finitely many n, then we call this family **eventually saturated**.

Before showing how saturation and stabilization of an equivariant family of monoids are related to each other, we make the following observations. For any $n \in \mathbb{N}$ the inclusion $M_n \subseteq M_\infty \bigcap \mathbb{Z}^n$ is true for any $\mathfrak{S}_\infty$–equivariant family of monoids. This is because if $x \in M_n$, then $x \in \mathbb{Z}^n$ since M_n is a commutative monoid in $\mathbb{Z}^n$. From the definition of the limit monoid we also have that $x \in M_\infty$. However, the inverse inclusion $M_\infty \bigcap \mathbb{Z}^n \subseteq M_n$, does not hold in general as the following example demonstrates.

Example 3.3.6. The normal $\mathfrak{S}_\infty$–equivariant family introduced in Example 3.3.3 is not saturated. For instance, for $n = 2$, the monoid $M_\infty \cap \mathbb{Z}^2$ is generated by elements of the form $e_1 + ke_2, ke_1 + e_2$ for any $k \in \mathbb{N}$, while the monoid M_2 is generated by the elements $e_1 + e_2, e_1 + 2e_2, 2e_1 + e_2$.

The following example shows that when an $\mathfrak{S}_\infty$–equivariant family of monoids stabilizes, then the family is eventually saturated.

Example 3.3.7. Consider the family of monoids defined by setting $M_1 = 0, M_2 = \langle e_1 + 2e_2, 2e_1 + e_2\rangle$ and $M_n = \mathrm{Mon}_{2,n}(M_2)$ for any $n \geq 2$. This is an $\mathfrak{S}_\infty$–equivariant family of monoids which stabilizes at index $k = 2$. For any $n \in \mathbb{N}$ the monoid M_n is generated by the $\mathrm{Sym}(n)$–orbit of $e_1 + 2e_2$. Those are elements of the form $e_i + 2e_j$ for any $i \neq j \in \{1, \dots, n\}$. The limit monoid M_∞ is, by definition, the monoid generated by the union of all generators of the monoids M_n, for each $n \in \mathbb{N}$. Thus, the intersection $M_\infty \cap \mathbb{Z}^n$ yields all n–dimensional points among the generators of M_∞ which are exactly the generators of M_n. For instance, if $n = 3$, then

$$M_\infty \cap \mathbb{Z}^3 = \langle e_1 + 2e_2, 2e_1 + e_2, e_1 + 2e_3, 2e_1 + e_3, e_2 + 2e_3, 2e_2 + e_3\rangle = M_3.$$

We have the following result.

Lemma 3.3.8. Any $\mathfrak{S}_\infty$–equivariant family of affine monoids that stabilizes is eventually saturated.

Proof. Let $\{M_n : n \in \mathbb{N}\}$ be a normal $\mathfrak{S}_\infty$–equivariant family of affine monoids. If this family stabilizes, then according to Lemma 3.2.9, the limit monoid M_∞ is finitely generated up to symmetry. In particular, if the stability index of the family is $r \in \mathbb{N}$, then

$$M_\infty = M_1 \cup \dots \cup M_r \cup \dots \cup \mathrm{Mon}_{r,n}(M_r) \cup \dots, \quad \forall r \leq n.$$

Since the family of monoids is normal $\mathfrak{S}_\infty$–equivariant, the only r–dimensional elements among the generators of M_∞ are exactly the generators of M_r. Because of the stabilization of the family, it holds that $\mathrm{Mon}_{r,n}(M_r) = M_n = M_\infty \cap \mathbb{Z}^n$ for any $n \geq r$. Therefore, the family $\{M_n : n \in \mathbb{N}\}$ is eventually saturated. □

Remark 3.3.9. The opposite direction of Lemma 3.3.8 does not hold in general and further assumptions are required.

3.4 Equivariant Gordan's Lemma

Suppose that we are given a family of rational polyhedral cones $\{C_n : n \in \mathbb{N}\}$. If we intersect each of the local cones C_n with $\mathbb{Z}^n$, then we obtain a family of monoids $\{M_n : n \in \mathbb{N}\}$, where M_n is a monoid in $\mathbb{Z}^n$ for any $n \in \mathbb{N}$. In the particular case where the cones in the family are finitely generated, then by Gordan's lemma (Lemma 3.1.5) we know that the family of monoids consists of affine monoids. We would like to study the behavior of families of cones up to symmetry and to further examine whether similar results with the finite case hold.

Let $\{C_n : n \in \mathbb{N}\}$ be a family of convex polyhedral cones, where $C_n \subseteq \mathbb{R}^n$ for any $n \in \mathbb{N}$. If $C_m \subseteq \mathbb{R}^m$ is a cone, then we embed C_m into $\mathbb{R}^n$ by applying the standard embeddings $\iota_{m,n}$ of Section 3.2 to the generators of C_m. We write $\iota_{m,n}(C_m)$ for the embedded cone in $\mathbb{R}^n$. The symmetric group $\mathrm{Sym}(n)$ acts on $\mathbb{R}^n$ by permuting coordinates. If $X \subseteq \mathbb{R}^m$, then we denote by

$$\mathrm{cone}_{m,n}(X) = \mathrm{cone}(\iota_{m,n}(X)^\sigma : \sigma \in \mathrm{Sym}(n)), \quad \forall m \leq n$$

the cone generated by the set of all permutations $\sigma \in \mathrm{Sym}(n)$ of the image $\iota_{m,n}(X)$.

Definition 3.4.1. An **$\mathfrak{S}_\infty$–equivariant family of cones**, is a family $\{C_n : n \in \mathbb{N}\}$ of convex cones $C_n \subseteq \mathbb{R}^n$ such that

$$\mathrm{cone}_{m,n}(C_m) \subseteq C_n, \quad \text{for any } m, n \in \mathbb{N},\ m \leq n. \tag{3.5}$$

An $\mathfrak{S}_\infty$–equivariant family of cones **stabilizes** if there exists $m \in \mathbb{N}$ such that for any $n \geq m$ the inclusion relation in equation (3.5) becomes an equality. In this case the family of cones stabilizes at index m.

Example 3.4.2. Let $C_1 = 0$ and $C_n = \mathrm{cone}_{2,n}((e_1 + ne_2)^\sigma : \sigma \in \mathrm{Sym}(n))$ for any $n \geq 2$ so that $C_2 = \mathrm{cone}((e_1 + 2e_2)^\sigma : \sigma \in \mathrm{Sym}(2))$, $C_3 = \mathrm{cone}((e_1 + 3e_2)^\sigma : \sigma \in \mathrm{Sym}(3))$, etc. The family of cones $\{C_n : n \in \mathbb{N}\}$ defined that way is $\mathfrak{S}_\infty$–equivariant. Indeed, we have $\mathrm{cone}_{n-1,n}(C_{n-1}) \subseteq C_n$. If $x \in \mathrm{cone}_{n-1,n}(C_{n-1})$ then

$$x = \sum_{\sigma \in \mathrm{Sym}(n)} \sum_{i=1}^{s} \lambda_i (e_1 + (n-1)e_2)^\sigma, \quad \forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}_{\geq 0},\ s = \frac{n!}{(n-2)!}.$$

We can express x as a linear combination of the extreme rays of C_n as follows,

$$\begin{aligned} x = &\frac{(n^2-n-1)\lambda_1 + \lambda_2}{n^2-1}(ne_1 + e_2) + \frac{\lambda_1 + (n^2-n-1)\lambda_2}{n^2-1}(e_1 + ne_2) + \cdots + \\ &\frac{(n^2-n-1)\lambda_{s-1} + \lambda_s}{n^2-1}(ne_{n-1} + e_n) + \frac{\lambda_{s-1} + (n^2-n-1)\lambda_s}{n^2-1}(e_{n-1} + ne_n) \end{aligned}$$

where $\frac{(n^2-n-1)\lambda_i+\lambda_j}{n^2-1} = \frac{(n(n-1)-1)\lambda_i+\lambda_j}{n^2-1} \geq \frac{\lambda_i+\lambda_j}{3} > 0$ for any pair of consecutive integers $i,j \in [s]$, hence $x \in C_n$. Since the converse inclusion does not hold, the family of cones does not stabilize.

Example 3.4.3. Consider the family $\{C_{(k,n)} : k \leq n\}$ of (k,n)–box pile cones $C_{(k,n)} \subseteq \mathbb{R}^n$ introduced in Section 2.2 of Chapter 2. This family is by definition $\mathfrak{S}_\infty$–equivariant and it stabilizes at index $k \in \mathbb{N}$. To see this, note that

$$\begin{aligned}\operatorname{cone}_{k,n}(C_{(k,k)}) &= \operatorname{cone}_{k,n}((a_1,\dots,a_k)^\sigma : \sigma \in \operatorname{Sym}(k)) \\ &= \operatorname{cone}((a_1,\dots.a_k,0,\dots,0)^\sigma : \sigma \in \operatorname{Sym}(n)) = C_{(k,n)}.\end{aligned}$$

When a family of cones stabilizes it is natural to ask what are the properties of the underlying family of monoids. From Gordan's Lemma (Lemma 3.1.5) we know that the property of being finitely generated is transferred from cones to monoids. Hence a family of finitely generated cones gives rise to a family of affine monoids. The following conjecture states that the stabilization property of a family of cones is adopted by the family of monoids. We refer to this result as the equivariant Gordan's Lemma, as it generalizes Gordan's Lemma for equivariant families of cones and monoids.

Conjecture 3.4.4 (Equivariant Gordan)**.** Let $\{C_n : n \in \mathbb{N}\}$ be an $\mathfrak{S}_\infty$–equivariant family of cones that stabilizes. Let $M_n = C_n \cap \mathbb{Z}^n$ for any $n \in \mathbb{N}$. Then the family $\{M_n : n \in \mathbb{N}\}$ is a normal $\mathfrak{S}_\infty$–equivariant family of monoids and it stabilizes.

Definition 3.4.5. The **equivariant Hilbert basis** of an $\mathfrak{S}_\infty$–equivariant family of cones that stabilizes is the up to symmetry minimal generating set of the underlying normal $\mathfrak{S}_\infty$–equivariant family of monoids described in Conjecture 3.4.4.

As an evidence for Conjecture 3.4.4 we provide the following result.

Theorem 3.4.6. The family $\mathfrak{C} = \{C_{(2,n)} : n \geq 2\}$ of $(2,n)$–box pile cones $C_{(2,n)} \subseteq \mathbb{R}^n$ is $\mathfrak{S}_\infty$–equivariant and stabilizes up to symmetry. Moreover if $M_n = C_{(2,n)} \cap \mathbb{Z}^n$ for any $n \geq 2$, then the family $\mathfrak{M} = \{M_n : n \geq 2\}$ is a normal $\mathfrak{S}_\infty$–equivariant family of monoids and it stabilizes up to symmetry. The stabilization index of the family depends on the values of the non-zero entries of the $(2,n)$–box pile generators.

Remark 3.4.7. As stated in Theorem 3.4.6 the stabilization index of the family $\mathfrak{M}$ depends on the values of the entries in the box pile generators. From computations using the computer algebra software *Macaulay2*, we noticed that when $C_{(2,n)}$ is generated by the $\operatorname{Sym}(n)$–orbit of the vector $(\alpha,1,0,\dots,0)$, then the family of monoids stabilizes at index three, while when $C_{(2,n)}$ is generated by the $\operatorname{Sym}(n)$–orbit of the vector $(a_1,a_2,0,\dots,0)$ with $a_2 > 1$, then the family stabilizes at index a_2+1. These cases are studied in detail in the following proof.

proof of Theorem 3.4.6. The $\mathfrak{S}_\infty$–equivariance and the stabilization of the family $\mathfrak{C}$ follows from the definition of box pile cones (see Definition 3.4.1). We show the stabilization of the family $\mathfrak{M}$ by distinguishing between two cases regarding the values of the entries in the $(2,n)$–box pile generators of the cones $C_{(2,n)}$.

Assume first that the cone $C_{(2,n)}$ is generated by the $\mathrm{Sym}(n)$–orbit of the vector $(\alpha,1,0,\dots,0)$ for some $\alpha\in\mathbb{Z}$ with $\alpha>1$. We claim that the monoid M_n is generated by the vectors

$$(1,1,0,\dots,0)^\sigma,\ (1,1,1,0,\dots,0)^\sigma,\ (\alpha-l,1,0,\dots,0)^\sigma,\quad \forall\sigma\in\mathrm{Sym}(n), \tag{3.6}$$

for any $l=0,\dots,\alpha-1$. If not, then there exists a point $x\in C_{(2,n)}\cap\mathbb{Z}^n$ such that $x-m\notin C_{(2,n)}$, where m is any vector in (3.6). We show by contradiction that this is not the case. We distinguish between different cases regarding the number s of non-zero entries in x.

If $s=2$, then $x=(x_1,x_2,0,\dots,0)$ and without loss of generality assume that $x_1\geq x_2$. Since $x\in C_{(2,n)}\cap\mathbb{Z}^n$, it satisfies the inequality description of $C_{(2,n)}$ given in Proposition 2.2.4, hence $\alpha x_2\geq x_1$. If $x_2=1$ then $\alpha\geq x_1\geq 1$ and if we subtract $(\alpha-l,1,0,\dots,0)$ from x for any $l\in\{0,\dots,\alpha-1\}$ we obtain the zero vector which belongs in $C_{(2,n)}$. If $x_2\neq 1$ then $x-(1,1,0,\dots,0)\in C_{(2,n)}$ because

$$\alpha(x_1-1)=\alpha x_1-\alpha\overset{x_1\geq x_2}{\geq}\alpha x_2-\alpha=\alpha(x_2-1)>x_2-1.$$

If $s\geq 3$, then $x=(x_1,\dots,x_s,0,\dots,0)$ and without loss of generality assume that $x_1\geq\dots\geq x_s$. We show that if we subtract a permutation of the vector $(1,1,1,0,\dots,0)$ from x, then the resulting vector is in $C_{(2,n)}$. We note that

$$\begin{aligned}\alpha(x_1+\dots+x_{s-2}-1+x_{s-1}-1)&=\alpha(x_1+\dots+x_{s-1})-2\alpha\\&\geq\alpha(s-1)x_s-2\alpha,\quad s\geq 3,\ \text{hence } s-1\geq 2,\\&\geq 2\alpha(x_s-1)\\&>x_s-1.\end{aligned}$$

This shows that the point $x'=(x_1,\dots,x_{s-2}-1,x_{s-1}-1,x_s-1,0,\dots,0)$ is an element in $C_{(2,n)}$ as it satisfies the inequality description given in Proposition 2.2.4. The above proof contradicts the hypothesis that an element in $C_{(2,n)}\cap\mathbb{Z}^n$ cannot be written as a linear combination of the vectors in (3.6) and the initial claim is true. The stabilization of the family $\mathfrak{M}$ follows from the observation that for any $n\geq 3$,

$$\begin{aligned}M_n&=\langle(1,1,0,\dots,0)^\sigma,(\alpha-l,1,0,\dots,0)^\sigma,(1,1,1,0,\dots,0)^\sigma:\sigma\in\mathrm{Sym}(n),l=0,\dots,\alpha-1\rangle\\&=\langle\iota_{3,n}((1,1,0),(\alpha-l,1,0),(1,1,1))^\sigma:\sigma\in\mathrm{Sym}(n),l=0,\dots,\alpha-1\rangle\\&=\mathrm{Mon}_{3,n}(M_3),\end{aligned}$$

where $M_3 = \langle (1,1,0)^\sigma, (\alpha - l, 1, 0)^\sigma, (1,1,1)^\sigma : \sigma \in \mathrm{Sym}(3), l = 0, \ldots, \alpha - 1\rangle \subseteq \mathbb{Z}^3$. Hence the family stabilizes at index 3.

For the rest of the proof we consider the more general case, where $C_{(2,n)}$ is generated by the $\mathrm{Sym}(n)$–orbits on the vector $(\alpha, \alpha-1, 0, \ldots, 0)$ for any $\alpha \geq 3$. We claim that the monoid M_n is generated by the vectors

$$(\alpha, \alpha-1, 0, \ldots, 0)^\sigma,\ (1,1,0,\ldots,0)^\sigma,\ (1,1,1,0,\ldots,0)^\sigma,\ (\alpha, \alpha - l - 1, l, 0, \ldots, 0)^\sigma, \tag{3.7}$$

for any $l = 1, \ldots, \lfloor \frac{\alpha}{2} \rfloor$ and

$$(\alpha, \alpha - k - r, k-1, 1, \ldots, 1, 0, \ldots, 0)^\sigma, \quad \forall \sigma \in \mathrm{Sym}(n), \tag{3.8}$$

for any $k = 2, \ldots, \lfloor \frac{\alpha}{2} \rfloor$, where $r = q - 3$ indicates the number of entries that are equal to one, for any $q = 4, \ldots, \alpha$. Here $\lfloor \frac{\alpha}{2} \rfloor = \max\{\omega \in \mathbb{Z} : \omega \leq \frac{\alpha}{2}\}$ is the floor of $\frac{\alpha}{2}$.

Suppose that the claim is false. Then there exists a vector $x \in M_n$ such that $x - m \notin C_{(2,n)}$, where m is any vector in the above claim. We prove the claim by contradiction following the same strategy as before. Let $x = (x_1, \ldots, x_s, 0, \ldots, 0) \in M_n$ and without loss of generality assume that $x_1 \geq \cdots \geq x_s$.

If $s = 2$, then by Proposition 2.2.4 the inequality $\alpha x_2 \geq (\alpha - 1)x_1$ is valid in $C_{(2,n)}$. If $x_2 = 1$, then $x_1 = 1$ and if we subtract the vector $(1,1,0,\ldots,0)$ from x then we obtain the zero vector which is an element of $C_{(2,n)}$. If $x_2 = \alpha - 1$ then $\alpha \geq x_1 \geq \alpha - 1$. If $x_1 = \alpha$ then we subtract the vector $(\alpha, \alpha - 1, 0, \ldots, 0)$ from x to obtain the zero vector. Otherwise $x_1 = \alpha - 1$ and we subtract the vector $(\alpha - 1)(1,1,0,\ldots)$ from x to obtain again the zero vector. In case $x_2 \neq 1, x_2 \neq \alpha - 1$, then $x - (1,1,0,\ldots,0) \in C_{(2,n)}$ because

$$\alpha(x_1 - 1) \geq \alpha(x_2 - 1) > (\alpha - 1)(x_2 - 1).$$

If $s = 3$ then $x = (x_1, x_2, x_3, 0, \ldots, 0) \in M_n$. We distinguish between the following cases regarding the entries of the vector x. If $x_1 = x_2 = x_3$, then $x - (1,1,1,0,\ldots,0) \in C_{(2,n)}$ since

$$\alpha(x_1 - 1 + x_2 - 1) = 2\alpha(x_3 - 1) > (\alpha - 1)(x_3 - 1).$$

Otherwise, $x - (\alpha - l - 1, l, \alpha, 0, \ldots, 0) \in C_{(2,n)}$ for any $l = 1, \ldots, \lfloor \frac{\alpha}{2} \rfloor$ because

$$\alpha(x_1 - \alpha + l + 1 + x_2 - l) = \alpha(x_1 + x_2) - \alpha(\alpha - 1) \geq (\alpha - 1)(x_3 - \alpha),$$

where the last inequality follows from the fact that $x \in C_{(2,n)}$ and therefore its coordinates satisfy the inequality $\alpha(x_1 + x_2) \geq (\alpha - 1)x_3$.

For $s \geq 4$ the point $x - (\alpha - k - r, k - 1, 1, \ldots, 1, \alpha, 0, \ldots, 0) \in C_{(2,n)}$, for any $k = 2, \ldots, \lfloor \frac{\alpha}{2} \rfloor$, where $r = q - 3$ for $q = 4, \ldots, \alpha$, because

$$\begin{aligned}\alpha(x_1 - \alpha + k + r + x_2 - k + 1 + x_3 - 1 + \cdots + x_{s-1} - 1) &= \alpha(x_1 + x_2 + \cdots + x_{s-1}) - \alpha(\alpha - 1)\\ &\geq (\alpha - 1)x_s - \alpha(\alpha - 1)\\ &= (\alpha - 1)(x_s - \alpha).\end{aligned}$$

The above contradicts the hypothesis that an element in M_n cannot be written as a linear combination of the vectors (3.7), (3.8). Hence, the initial claim is true. The stabilization index of the family $\mathfrak{M}$ in this case is equal to the maximum number of non-zero entries among the generators of M_n, hence the number of non-zero entries of the vector (3.8). This number is equal to $3 + r = 3 + (q - 3) = 3 + (\alpha - 3) = \alpha$, where the second equality is obtained by giving q the maximum possible value it can take.

One can show that the result obtained in the second part of the proof holds for any family of $(2, n)$−box pile cones $C_{(2,n)} = \text{cone}_{2,n}((a, b, 0, \dots)^\sigma : \sigma \in \text{Sym}(n))$ for any entries a, b with $a > b > 1$ and for any $n \geq 2$. Exhibiting the explicit description of the minimal generators of the monoid M_n in this case is complicated, however following exactly the same arguments as above one can prove that the stabilization index of the family $\mathfrak{M}$ of underlying monoids equals $b + 1$.

□

By Definition 3.4.5 and using the proof of Theorem 3.4.6 we conclude that the equivariant Hilbert basis for the family of $(2, n)$−box pile cones is equal to the $\mathfrak{S}_\infty$−orbits on the vectors in (3.6) whenever $C_{(2,n)} = \text{cone}((\alpha, 1, 0, \dots, 0)^\sigma : \sigma \in \text{Sym}(n), \alpha > 1)$. In case $C_{2,n} = \text{cone}((\alpha, \alpha - 1, 0, \dots, 0)^\sigma : \sigma \in \text{Sym}(n), \alpha > 2)$, then the equivariant Hilbert basis is equal to the $\mathfrak{S}_\infty$−orbits on the vectors in (3.7) and (3.8).

4 Gröbner Bases for Staged Trees

In this chapter of the thesis we are concerned with the problem of determining generators of the toric ideal associated with a combinatorial object called a Staged Tree. We show that in the case of a balanced and stratified staged tree, the generating set of the underlying toric ideal forms a quadratic Gröbner basis with squarefree initial terms. The proofs of the main results presented here are based on a toric fiber product construction due to Sullivant [Sul07]. This chapter consists of results obtained in a joint paper with Eliana Duarte [DA19].

4.1 Basic definitions for Staged Trees

Let $\mathcal{T} = (V, E)$ be a directed rooted tree graph with vertex set V and set E of directed edges. We only consider trees $\mathcal{T} = (V, E)$ where no two directed edges point to the same vertex, and all elements in E are oriented away from the root. For any $v, w \in V$ the directed edge from v to w in E is denoted by (v, w). The set of *children* of v is $\mathrm{ch}(v) := \{u \in V : (v, u) \in E\}$ and the set of outgoing edges from the vertex v is $E(v) := \{(v, u) : u \in \mathrm{ch}(v)\}$. If $E(v) = \varnothing$, then we refer to the vertex v as a *leaf* of $\mathcal{T}$. We denote by $v \to w$ the directed path with head the vertex v and tail the vertex w and by $E(v \to w)$ the set of all edges in this path. If $\mathcal{L}$ is a set of labels, then consider the map $\theta : E \to \mathcal{L}$ which assigns to each $e \in E$ a unique label from $\mathcal{L}$. Given a vertex $v \in V$, we write $\theta_v := \{\theta(e) : e \in E(v)\}$ for the set of edge labels attached to v. The following definition introduces the main objects of this chapter.

Definition 4.1.1. Let $\mathcal{L}$ be a set of labels. A tree $\mathcal{T} = (V, E)$ together with a labeling $\theta : E \to \mathcal{L}$ is a **staged tree**, if

- for any $v \in V$, $|\theta_v| = |E(v)|$, and
- for any $v, w \in V$, the sets θ_v, θ_w are either equal or disjoint.

Using Definition 4.1.1 we define an equivalence relation on the set of vertices of $\mathcal{T}$. Namely, two vertices v, w are equivalent if and only if $\theta_v = \theta_w$. We refer to the partition

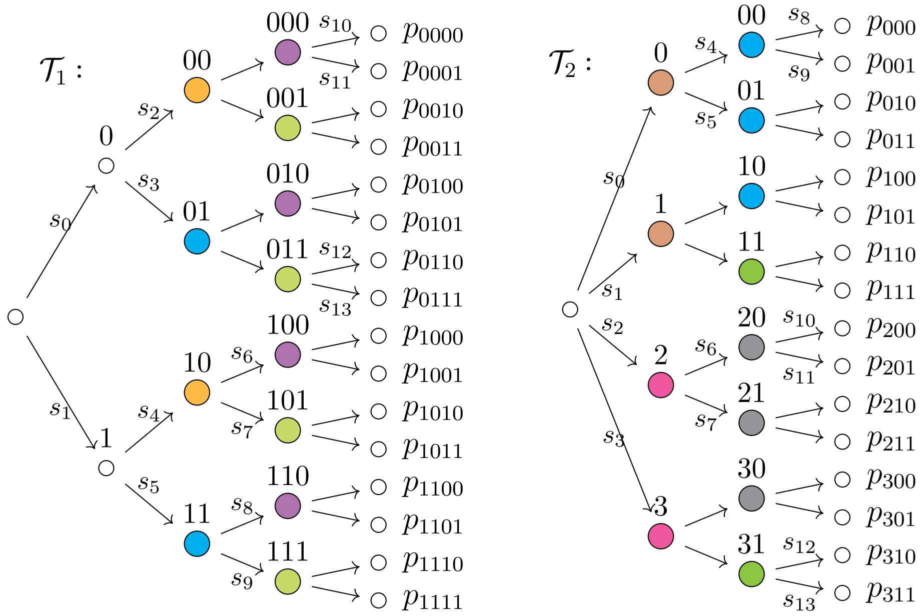


Figure 4.1: Examples of a staged trees.

induced by this equivalence relation on the set V as the *set of stages* of V and to a single element in this partition as a *stage*. We use the pair $(\mathcal{T}, \theta)$ to denote a staged tree $\mathcal{T}$ together with a labeling rule θ. For simplicity we often drop the use of θ and we write $\mathcal{T}$ for a staged tree.

Let $(\mathcal{T}, \theta)$ be a staged tree. In order to define the toric ideal associated to this staged tree, we define two polynomial rings. The first ring is $\mathbb{R}[p]_\mathcal{T} := \mathbb{R}[p_\lambda : \lambda \in \Lambda]$, where Λ is the set of all root-to-leaf paths in $\mathcal{T}$. The second ring is $\mathbb{R}[\Theta]_\mathcal{T} := \mathbb{R}[z, \mathcal{L}]$, with variables the labels in $\mathcal{L}$ together with a homogenizing variable z. Consider the ring homomorphism

$$\begin{aligned} \varphi_\mathcal{T} : \mathbb{R}[p]_\mathcal{T} &\to \mathbb{R}[\Theta]_\mathcal{T} \\ p_\lambda &\mapsto z \cdot \prod_{e \in E(\lambda)} \theta(e) \end{aligned} \tag{4.1}$$

Definition 4.1.2. The **toric staged tree ideal** associated to $(\mathcal{T}, \theta)$ is the kernel $\ker(\varphi_\mathcal{T})$ of the homomorphism $\varphi_\mathcal{T}$.

The toric staged tree ideal defines the toric variety specified as the closure of the image of the monomial parameterization $\Phi_\mathcal{T} : (\mathbb{C}^*)^{|\mathcal{L}|} \to \mathbb{P}^{|\Lambda|-1}$ given by $(\theta(e) \in \mathcal{L}) \mapsto z \cdot (\prod_{e \in E(\lambda)} \theta(e))_{\lambda \in \Lambda}$. We use the homogenizing variable z in the map (4.1) to consider the projective toric variety in $\mathbb{P}^{|\Lambda|-1}$.

Example 4.1.3. The staged tree $\mathcal{T}_1$ in Figure 4.1 has label set $\mathcal{L} = \{s_0, \dots, s_{13}\}$. Each vertex in $\mathcal{T}_1$ is identified by a sequence of 0's and 1's and each edge has a label associated to it. The root-to-leaf paths in $\mathcal{T}_1$ are denoted by p_{ijkl} for any $i,j,k,l \in \{0,1\}$. A vertex in $\mathcal{T}_1$ represented by a blank cycle indicates a stage consisting of a single vertex. The vertices in $\mathcal{T}_1$ that have the same color correspond to vertices having the same stage. For instance, the purple vertices, that is the vertices in the set $\{000, 010, 100, 110\}$, are in the same stage and therefore they have the same set $\{s_{10}, s_{11}\}$ of attached edge labels. The map $\Phi_{\mathcal{T}_1}$ maps the vector $(s_0, \dots, s_{13})$ to the vector

$$\begin{aligned}(&s_0s_2s_6s_{10}, s_0s_2s_6s_{11}, s_0s_2s_7s_{12}, s_0s_2s_7s_{13}, s_0s_3s_8s_{10}, s_0s_3s_8s_{11}, s_0s_3s_9s_{12}, s_0s_3s_9s_{13},\\ &s_1s_4s_6s_{10}, s_1s_4s_6s_{11}, s_1s_4s_7s_{12}, s_1s_4s_7s_{13}, s_1s_5s_8s_{10}, s_1s_5s_8s_{11}, s_1s_5s_9s_{12}, s_1s_5s_9s_{13}).\end{aligned}$$

The toric ideal $\ker(\varphi_{\mathcal{T}_1})$ is generated by a quadratic Gröbner basis with squarefree initial ideal.

We are interested in relating the combinatorial properties of the staged tree $(\mathcal{T}, \theta)$ with the properties of the toric ideal $\ker(\varphi_{\mathcal{T}})$. The two definitions that are relevant for the statement of the main theorem, Theorem 4.4.12, are the definition of balanced staged tree and of stratified staged tree. In the following we look into the definitions and consequences of these two notions.

Definition 4.1.4. Let $\mathcal{T}$ be a tree. For $v \in V$, the **level of v**, denoted $l(v)$, is the number of edges in the unique path from the root of $\mathcal{T}$ to v. If all the leaves in $\mathcal{T}$ have the same level, then the **level of $\mathcal{T}$** is equal to the level of any of its leaves. A staged tree $(\mathcal{T}, \theta)$ is **stratified** if all its leaves have the same level and if every two vertices in the same stage have the same level.

The staged trees $\mathcal{T}_1, \mathcal{T}_2$ in Figure 4.1 are stratified. Particularly, all the leaves of $\mathcal{T}_1$ have level equal to 4, similarly all the leaves of $\mathcal{T}_2$ have level equal to 3, and every two vertices with the same color appear in the same level. Notice that the combinatorial condition of $(\mathcal{T}, \theta)$ being stratified implies the algebraic condition that the map $\varphi_{\mathcal{T}}$ is squarefree.

In the rest of this section we focus on defining balanced staged trees. This definition is formulated in terms of polynomials associated to each vertex. Those are called interpolating polynomials in the following and their properties are very useful for the proof of important statements in the following sections.

Definition 4.1.5. Let $(\mathcal{T}, \theta)$ be a staged tree, $v \in V$ and $\mathcal{T}_v$ be the subtree of $\mathcal{T}$ rooted at v. The tree $\mathcal{T}_v$ is a staged tree with the induced labeling from $\mathcal{T}$. Let Λ_v be the set of all v–to-leaf paths in $\mathcal{T}$ and define

$$t(v) := \sum_{\lambda \in \Lambda_v} \prod_{e \in E(\lambda)} \theta(e).$$

The polynomial $t(v)$ is called the **interpolating polynomial** of $\mathcal{T}_v$. Two staged trees are **polynomially equivalent** if they have the same set of edge labels and their interpolating polynomials coincide.

Remark 4.1.6. When v is the root of $\mathcal{T}$, the polynomial $t(v)$ is the interpolating polynomial of $\mathcal{T}$.

The interpolating polynomial of a staged tree is an important tool in the study of the statistical properties of staged tree models. It was first introduced in [GS18] to enumerate all possible staged trees that define the same staged tree model and was further studied in [GBRS18]. Although these two articles are written for a statistical audience, their symbolic algebra approach to the study of statistical models proves to be very important for the use of these models in practice. A useful property of the interpolating polynomials is stated in the following lemma.

Lemma 4.1.7 (Theorem 1, [GBRS18])**.** Let $(\mathcal{T}, \theta)$ be a staged tree, $v \in V$ and assume that $\mathrm{ch}(v) = \{v_0, v_1, \dots, v_k\}$. Then the polynomial $t(v)$ admits the recursive representation $t(v) = \sum_{i=0}^{k} \theta(v, v_i) t(v_i)$.

Definition 4.1.8. Let $(\mathcal{T}, \theta)$ be a staged tree and v, w be two vertices in the same stage with $\mathrm{ch}(v) = \{v_0, \dots, v_k\}, \mathrm{ch}(w) = \{w_0, \dots, w_k\}$. After a possible permutation of the elements in $\mathrm{ch}(w)$ assume that $\theta(v, v_i) = \theta(w, w_i)$ for all $i \in \{0, \dots, k\}$. Suppose that the vertices v, w satisfy the **balanced condition**

$$t(v_i)t(w_j) = t(w_i)t(v_j), \quad \forall i \neq j \in \{0, \dots, k\} \tag{$\star$}$$

in $\mathbb{R}[\Theta]_{\mathcal{T}}$. We call the staged tree $(\mathcal{T}, \theta)$ **balanced** if every pair of vertices in the same stage satisfy the balanced condition ($\star$).

Example 4.1.9. The staged tree $\mathcal{T}_2$ in Figure 4.1 is stratified but is not balanced since the two orange vertices, that is the vertices 0 and 1, do not satisfy the balanced condition ($\star$). Precisely we have that

$$t(00)t(11) = (s_8 + s_9)(s_{12} + s_{13}) \neq t(10)t(01) = (s_8 + s_9)^2.$$

Although the balanced condition in Definition 4.1.8 seems to be algebraic and hard to check, in many cases it is very combinatorial. To formulate a precise instance where this is true we need the following definition.

Definition 4.1.10. Let $(\mathcal{T}, \theta)$ be a staged tree. The vertices v, w are in the same **position** if they are in the same stage and $t(v) = t(w)$.

Remark 4.1.11. The notion of position was formulated in [SA08]. Intuitively it means that if we regard the subtrees $\mathcal{T}_v$ and $\mathcal{T}_w$ as representing the unfolding of a sequence of events, then the future of v and w is essentially the same.

Lemma 4.1.12. Let $(\mathcal{T}, \theta)$ be a stratified staged tree. Suppose that two vertices in $\mathcal{T}$ that are in the same stage are also in the same position. Then $(\mathcal{T}, \theta)$ is balanced.

Proof. In order to show that $(\mathcal{T}, \theta)$ is a balanced staged tree we need to prove that every pair of vertices in the same stage in $\mathcal{T}$ satisfies equation $(\star)$ of Definition 4.1.8.

Let v, w be two vertices in $\mathcal{T}$ with $\mathrm{ch}(v) = \{v_0, \dots, v_k\}, \mathrm{ch}(w) = \{w_0, \dots, w_k\}$. If v, w are in the same stage, then by the discussion following Definition 4.1.1 we have that $\theta_v = \theta_w$ or equivalently that $\theta(v, v_i) = \theta(w, w_i)$ for any $i \in \{0, \dots, k\}$ and after possibly permuting the vertices in $\mathrm{ch}(w)$. Since v, w are in the same position by Definition 4.1.10 we have $t(v) = t(w)$. Using Lemma 4.1.7 we rewrite this relation as follows

$$t(v) = t(w) \Rightarrow \sum_{i=0}^{k} \theta(v, v_i) t(v_i) = \sum_{i=0}^{k} \theta(w, w_i) t(w_i) \Rightarrow \sum_{i=0}^{k} \theta(v, v_i)(t(v_i) - t(w_i)) = 0.$$

Since $(\mathcal{T}, \theta)$ is a stratified staged tree, the variables in the polynomials $t(v_i), t(w_i)$ are disjoint from the variables in $\{\theta(v, v_0), \dots, \theta(v, v_k)\}$. Hence, $t(v_i) = t(w_i)$ for any $i \in \{0, \dots, k\}$. It follows that $t(v_i)t(w_j) = t(w_i)t(v_j)$ for any $i \neq j \in \{0, \dots, k\}$. Therefore $(\mathcal{T}, \theta)$ is balanced. □

Example 4.1.13. The staged tree $\mathcal{T}_1$ in Figure 4.1 is balanced. This can be seen by noting that the orange vertices are in the same position and similarly for the blue vertices. In contast to this, the staged tree $\mathcal{T}_2$ is not balanced because the orange vertices, as well as the pink vertices, are not in the same position.

4.2 Equations for Staged Trees

In this section we follow the treament in [DG18] to define equations of a given staged tree. These equations are given as generators of two ideals corresponding to paths and maximal path extensions of $\mathcal{T}$. In order to describe these ideals we first need to introduce some special notation.

Let $(\mathcal{T}, \theta)$ be a staged tree. Given a vertex $v \in V$ denote by $[v]$ the set of all root-to-leaf-paths in $\mathcal{T}$ passing through v. Set

$$p_{[v]} := \sum_{j \in [v]} p_j. \tag{4.2}$$

Let $v, w \in V$ be two vertices in the same stage in $\mathcal{T}$. Then by the discussion following Definition 4.1.1 we have $\theta_v = \theta_w$. Assuming that $\mathrm{ch}(v) = \{v_0, \dots, v_k\}, \mathrm{ch}(w) =$

$\{w_0, \dots, w_k\}$, the last condition implies that $\theta(v, v_i) = \theta(w, w_i)$ for any $i \in \{0, \dots, k\}$, and after a possible permutation of the vertices in $\mathrm{ch}(w)$.

Example 4.2.1. Consider the staged tree $\mathcal{T}_1$ in Figure 4.1. There are three root-to-leaf path in $\mathcal{T}_1$ passing through the vertex 10, namely $p_{1000}, p_{1001}, p_{1010}, p_{1011}$. Hence

$$p_{[10]} = p_{1000} + p_{1001} + p_{1010} + p_{1011}.$$

Definition 4.2.2. Let $(\mathcal{T}, \theta)$ be a staged tree. For any pair of vertices v, w in the same stage in $\mathcal{T}$ the **ideal of paths** is an ideal in $\mathbb{R}[p]_{\mathcal{T}}$ defined by

$$I_{\mathrm{paths}} := \sum_{v \sim w \text{ in } \mathcal{T}} I_{v \sim w},$$

where

$$I_{v\sim w} = \langle p_{[v_i]} p_{[w_j]} - p_{[w_i]} p_{[v_j]} : i \neq j \in \{0, \dots k\} \rangle.$$

The ideal I_{paths} in Definition 4.2.2 is called the *ideal of paths* because each generator $p_{[vi]} p_{[wj]} - p_{[wi]} p_{[vj]}$ of $I_{v\sim w}$ corresponds to a pair of undirected paths $(v_i \to w_j, w_i \to v_j)$ in $\mathcal{T}$. To be more precise, to the unique path starting at vertex v_i and ending at vertex w_j, denoted $v_i \to w_j$, corresponds the product $p_{[v_i]} p_{[w_j]}$. Similarly, to the unique path $w_i \to v_j$ corresponds the product $p_{[w_i]} p_{[v_j]}$. From these two products, we form the difference $p_{[v_i]} p_{[w_j]} - p_{[w_i]} p_{[v_j]}$ and we refer to it as the path difference associated to $(v_i \to w_j, w_i \to v_j)$. It follows from [DG18, Lemma 9] that $I_{\mathrm{paths}} \subseteq \ker(\varphi_{\mathcal{T}})$. However in many cases this inclusion is strict and $\ker(\varphi_{\mathcal{T}})$ contains more elements. One way to find more generators of $\ker(\varphi_{\mathcal{T}})$ is to consider extensions of path differences.

Example 4.2.3. The path differences defining the tree $\mathcal{T}_1$ in Figure 4.1 are the following. The binomials

$$p_{0000}p_{0101} - p_{0100}p_{0001}, p_{0000}p_{1001} - p_{0001}p_{1000}, p_{0000}p_{1101} - p_{0001}p_{1100},$$
$$p_{0100}p_{1001} - p_{0101}p_{1000}, p_{0100}p_{1101} - p_{0101}p_{1100}, p_{1000}p_{1101} - p_{1001}p_{1100}$$

define the stage that corresponds to the purple vertices. The binomials

$$p_{0010}p_{0111} - p_{0011}p_{0110}, p_{0010}p_{1011} - p_{0011}p_{1010}, p_{0010}p_{1111} - p_{0011}p_{1110},$$
$$p_{0110}p_{1011} - p_{0111}p_{1010}, p_{0110}p_{1111} - p_{0111}p_{1110}, p_{1010}p_{1111} - p_{1011}p_{1110}$$

define the stage that corresponds to the green vertices. Finally, the differences

$$(p_{0000} + p_{0001})(p_{1010} + p_{1011}) - (p_{1000} + p_{1001})(p_{0010} + p_{0011}),$$
$$(p_{0100} + p_{0101})(p_{1110} + p_{1111}) - (p_{1100} + p_{1101})(p_{0110} + p_{0111})$$

define the stages that are associated with the orange and blue vertices.

Definition 4.2.4. A pair of paths $(u_1 \to u_2, u_3 \to u_4)$ is an **extension** of the path $(v_1 \to v_2, v_3 \to v_4)$ if the following hold:

- The path $u_1 \to u_2$ is obtained by adding l edges at the head or tail of $v_1 \to v_2$.
- The path $u_3 \to u_4$ is obtained by adding l edges at the head or tail of $v_3 \to v_4$.
- The edges $e_i \in E(u_1 \to u_2) \backslash E(v_1 \to v_2)$ and $e_i' \in E(u_3 \to u_4) \backslash E(v_3 \to v_4)$ satisfy

$$\prod_{i=1}^{l} \theta(e_i) = \prod_{i=1}^{l} \theta(e_i'). \tag{4.3}$$

A path extension $(u_1 \to u_2, u_3 \to u_4)$ is a **maximal extension**, if it is not possible to add edges to the pair $(u_1 \to u_2, u_3 \to u_4)$ in such a way that equation (4.3) is satisfied.

Remark 4.2.5. The last condition in Definition 4.2.4 implies that if $p_{[v_1]}p_{[v_2]} - p_{[v_3]}p_{[v_4]} \in \ker(\varphi_{\mathcal{T}})$, then the difference $p_{[u_1]}p_{[u_2]} - p_{[u_3]}p_{[u_4]}$ is also an element in $\ker(\varphi_{\mathcal{T}})$.

Definition 4.2.6. The **ideal of maximal path extensions**, is an ideal in $\mathbb{R}[p]_{\mathcal{T}}$ defined by

$$I_{\text{mpaths}} := \sum_{v \sim w \text{ in } \mathcal{T}} I_{\max(v \sim w)},$$

where $I_{\max(v \sim w)}$ is the ideal generated by path differences associated to all maximal extensions of $(v_i \to w_j, w_i \to v_j)$ for all $i \neq j \in \{0, \dots, k\}$.

Example 4.2.7. Let us consider the staged tree $\mathcal{T}_1$ in Figure 4.1. The pair of paths $(000 \to 101, 100 \to 001)$ has the path difference

$$p_{[000]}p_{[101]} - p_{[100]}p_{[001]} = (p_{0000} + p_{0001})(p_{1010} + p_{1011}) - (p_{1000} + p_{1001})(p_{0010} + p_{0011})$$

associated to it. A possible extension of the path $(000 \to 101, 100 \to 001)$ by one edge is given by $(0000 \to 101, 1000 \to 001)$ because $\theta(0000, 000) = \theta(1000, 001) = s_{10}$. The path difference associated to this extension is

$$p_{0000}p_{[101]} - p_{1000}p_{[001]} = p_{0000}(p_{1010} + p_{1011}) - p_{1000}(p_{0010} + p_{0011}).$$

We can further extend this path using an edge with label s_{12} or s_{13}. If we consider the extension $(0000 \to 1011, 1000 \to 0011)$, then this yields the binomial path difference

$$p_{0000}p_{1011} - p_{1000}p_{0011}.$$

Example 4.2.8. Here we give an example of a path equation that cannot be extended to a binomial equation. To this end we consider the staged tree $\mathcal{T}_2$ in Figure 4.1 and the pair of paths $(20 \to 31, 30 \to 21)$ with associated path difference

$$p_{[20]}p_{[31]} - p_{[30]}p_{[21]} = (p_{200} + p_{201})(p_{310} + p_{311}) - (p_{300} + p_{301})(p_{210} + p_{211}).$$

We can extend this path using an edge with label s_{10} or s_{11}. In each of these cases we get the polynomials

$$p_{200}(p_{310} + p_{311}) - p_{300}(p_{210} + p_{211}), p_{201}(p_{310} + p_{311}) - p_{301}(p_{210} + p_{211}).$$

However, we cannot extend the paths $(200 \to 31, 300 \to 21)$ and $(201 \to 31, 301 \to 21)$ further to get binomial differences because $\theta(21s, 21) \neq \theta(31s, 31)$ for any $s = 0, 1$.

4.3 Toric Fiber Products for Staged Trees

The toric fiber product of two homogeneous ideals is an operation that produces a new homogeneous ideal. This construction was first introduced in [Sul07] and has since then been studied in [KR14, EKS14]. Toric fiber products have been used in algebraic statistics to obtain implicit equations of statistical models. For instance in [DBM10] the authors use the toric fiber product to construct model invariants for phylogenetic trees. In this section we establish the framework to obtain toric staged trees via a toric fiber product construction. The main idea behind this construction is that following a set of gluing rules we can build a staged tree that is balanced and stratified from simpler pieces.

4.3.1 Toric Fiber Products Basics

In this subsection we recall basic facts of toric fiber products. The main reference here is the paper of Seth Sullivant, [Sul07].

In the following, let r be a positive integer and let s, t be two vectors of positive integers in $\mathbb{Z}_+^r$. Given a positive integer m, we denote by $[m]$ the set of the first m nonzero integers, i.e. $[m] := \{1, 2, \dots, m\}$. Consider the homogeneous, multigraded polynomial rings

$$\mathbb{K}[x] := \mathbb{K}[x_j^i \mid i \in [r], j \in [s_i]] \quad \text{and} \quad \mathbb{K}[y] := \mathbb{K}[y_k^i \mid i \in [r], k \in [t_i]]$$

having the same multigrading

$$\deg(x_j^i) = \deg(y_k^i) = \boldsymbol{a}^i \in \mathbb{Z}^d.$$

Denote by $\mathcal{A} = \{\boldsymbol{a}^1, \dots, \boldsymbol{a}^r\}$ the set of all multidegrees of these variables and assume that there exists a vector $w \in \mathbb{Q}^d$ such that $\langle w, \boldsymbol{a}^i \rangle = 1$ for any $\boldsymbol{a}^i \in \mathcal{A}$. If $I \subseteq \mathbb{K}[x]$ and $J \subseteq \mathbb{K}[y]$ are homogeneous ideals, then the quotient rings $R = \mathbb{K}[x]/I$ and $S = \mathbb{K}[y]/J$ are also multigraded rings. Let

$$\mathbb{K}[z] := \mathbb{K}[z^i_{jk} \mid i \in [r], j \in [s_i], k \in [t_i]]$$

and consider the ring homomorphism

$$\begin{aligned} \phi_{I,J} : \mathbb{K}[z] &\to R \otimes_{\mathbb{K}} S \\ z^i_{jk} &\mapsto \overline{x^i_j} \otimes \overline{y^i_k}, \end{aligned}$$

where $\overline{x^i_j}$ and $\overline{y^i_k}$ are the equivalence classes of x^i_j and y^i_k respectively.

Definition 4.3.1. The **toric fiber product** of I and J with respect to the multidegrees in $\mathcal{A}$, is $I \times_{\mathcal{A}} J := \ker \phi_{I,J}$.

It is shown in [Sul07, Theorem 12] that under the condition that $\mathcal{A}$ is a linearly independent set, then it is possible to explicitly compute a Gröbner bases of $I \times_{\mathcal{A}} J$ from individual Gröbner bases of I and J. We will use this result to show that the equations defining staged trees can be obtained via a toric fiber product technique. Furthermore, we will show that these equations form a Gröbner basis.

4.3.2 The Tree Gluing construction

Let $(\mathcal{T}, \theta)$ be a staged tree. For the purposes of this paragraph, we recursively define an indexing on the interior vertices of $\mathcal{T}$, i.e. the vertices that are different from the root of $\mathcal{T}$, as follows. The children of the root are indexed by a number in $\{0, 1, \dots, k\}$. If $\boldsymbol{a}$ is the index of the vertex v and $|E(v)| = j + 1$, then we index the children of v by $\boldsymbol{a}0, \boldsymbol{a}1, \dots, \boldsymbol{a}j$. This way every interior vertex in V is indexed by a finite sequence of non-negative integers

$$\boldsymbol{a} = a_1 a_2 \cdots a_l,$$

where l is the level of the vertex indexed by $\boldsymbol{a}$. We denote by $\boldsymbol{i}_{\mathcal{T}}$ the set of indices of the leaves in $\mathcal{T}$. From now on we refer to any interior vertex in V using its index $\boldsymbol{a}$.

Example 4.3.2. The vertices in both trees $\mathcal{T}_1$ and $\mathcal{T}_2$ in Figure 4.1 are labeled according to the rule introduced in the above paragraph.

Definition 4.3.3. If a staged tree only has one level we call it a **one-level tree**. We reserve for it the special notation $(\mathcal{B}, \epsilon)$ where $\mathcal{B} = (V, E)$ is the tree and ϵ its labeling rule. By Definition 4.1.1, the size of the label set of $\mathcal{B}$ is equal to $|E|$. Thus we use ϵk to denote the the image of the k−th element in E under ϵ. We also use the notation $(\mathcal{B}, \{\epsilon 0, \dots, \epsilon m\})$ when we wish to emphasize the label set of the one-level tree.

Definition 4.3.4. Let $(\mathcal{T},\theta)$ be a staged tree and $G = \{G_1,\dots,G_r\}$ be a partition on the set of leaves $\boldsymbol{i}_{\mathcal{T}}$. For any $i \in [r]$, let $\{(\mathcal{B}_i,\epsilon^{(i)}) : i \in [r]\}$ be a collection of one-level trees such that their label sets are pairwise disjoint and disjoint from the label set of $(\mathcal{T},\theta)$. The **gluing component** associated to $\mathcal{T}$ and G is denoted by $\mathcal{T}_G$ and is defined as the disjoint union of the $(\mathcal{B}_i,\epsilon^{(i)})$'s, that is,

$$\mathcal{T}_G := \bigsqcup_{i\in[r]} (\mathcal{B}_i,\epsilon^{(i)}).$$

The gluing component $\mathcal{T}_G$ is a forest of one-level trees, its label set is the union of the label sets of each $(\mathcal{B}_i,\epsilon^{(i)})$. We denote by $[\mathcal{T},\mathcal{T}_G]$ the tree obtained by gluing $\mathcal{B}_i$ to the leaf $\boldsymbol{a}$ for all $\boldsymbol{a} \in G_i$ and all $i \in [r]$.

Remark 4.3.5. The tree $[\mathcal{T},\mathcal{T}_G]$ is a staged tree. Its label set is the union of the labels sets of $(\mathcal{T},\theta)$ and $\mathcal{T}_G$. The labeling rule is inherited from the labelings of $\mathcal{T}$ and $\mathcal{T}_G$ and it satisfies the conditions in Definition 4.1.1. Moreover, $\boldsymbol{i}_{[\mathcal{T},\mathcal{T}_G]} = \{\boldsymbol{a}k : \boldsymbol{a} \in G_i, k \in \boldsymbol{i}_{\mathcal{B}_i}, i \in [r]\}$. The stages in $[\mathcal{T},\mathcal{T}_G]$ are the ones inherited from $\mathcal{T}$ union the new stages determined by G. This means that two vertices $\boldsymbol{a},\boldsymbol{b} \in \boldsymbol{i}_{\mathcal{T}}$ are in the same stage in $[\mathcal{T},\mathcal{T}_G]$ provided $\boldsymbol{a},\boldsymbol{b} \in G_i$.

In the following, let $\mathcal{T}, G, \mathcal{T}_G$ and $[\mathcal{T},\mathcal{T}_G]$ be as in Definition 4.3.4. We explain how $\ker(\varphi_{[\mathcal{T},\mathcal{T}_G]})$ is obtained as the toric fiber product of the ideals $\ker(\varphi_{\mathcal{T}})$ and the zero ideal $\langle 0\rangle$. First, we associate to $\mathcal{T}_G$ the polynomial rings $\mathbb{R}[p]_{\mathcal{T}_G} := \mathbb{R}[p^i_k : k \in \boldsymbol{i}_{\mathcal{B}_i}, i \in [r]]$ and $\mathbb{R}[\Theta]_{\mathcal{T}_G} := \mathbb{R}[\epsilon^{(i)}_k : i \in [r], k \in \boldsymbol{i}_{\mathcal{B}_i}]$ and define

$$\begin{aligned}\varphi_{\mathcal{T}_G} : \mathbb{R}[p]_{\mathcal{T}_G} &\to \mathbb{R}[\Theta]_{\mathcal{T}_G}\\ p^i_k &\mapsto \epsilon^{(i)}_k, \qquad \forall k \in \boldsymbol{i}_{\mathcal{B}_i}, \forall i \in [r].\end{aligned}$$

Since there is a one-to-one correspondence between the variables p^i_k and $\epsilon^{(i)}_k$, the map $\varphi_{\mathcal{T}_G}$ is an isomorphism. In particular, $\ker(\varphi_{\mathcal{T}_G}) = \langle 0\rangle$. Second, using the partition G on $\boldsymbol{i}_{\mathcal{T}}$ we write

$$\mathbb{R}[p]_{\mathcal{T}} = \mathbb{R}[p_{\boldsymbol{j}} : \boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}}] = \mathbb{R}[p^i_{\boldsymbol{j}} : \boldsymbol{j} \in G_i, i \in [r]].$$

Then, $\ker(\varphi_{\mathcal{T}})$ is the kernel of the map

$$\begin{aligned}\varphi_{\mathcal{T}} : \mathbb{R}[p]_{\mathcal{T}} &\to \mathbb{R}[\Theta]_{\mathcal{T}}\\ p^i_{\boldsymbol{j}} &\mapsto z \cdot \prod_{e\in E(\lambda_{\boldsymbol{j}})} \theta(e), \qquad \forall \boldsymbol{j} \in G_i, \forall i \in [r].\end{aligned}$$

We define a multigrading on the variables of $\mathbb{R}[p]_{\mathcal{T}}$ and $\mathbb{R}[p]_{\mathcal{T}_G}$ as follows:

$$\deg(p^i_{\boldsymbol{j}}) = \deg(p^i_k) = \boldsymbol{e}_i, \quad \forall \boldsymbol{j} \in G_i, \forall k \in \boldsymbol{i}_{\mathcal{B}_i}, \forall i \in [r],$$

where $\boldsymbol{e}_i$ is the i–th standard unit vector in $\mathbb{Z}^r$, with a 1 in coordinate i and zeros elsewhere. If $\mathcal{A}$ is the set of all these multidegrees, then $\mathcal{A}$ is linearly independent, as it is the collection of standard unit vectors in $\mathbb{Z}^r$.
Set $R := \mathbb{R}[p]_{\mathcal{T}}/\ker(\varphi_{\mathcal{T}}), S := \mathbb{R}[p]_{\mathcal{T}_G}/\ker(\varphi_{\mathcal{T}_G})$, and let

$$\mathbb{R}[p]_{[\mathcal{T},\mathcal{T}_G]} = \mathbb{R}[p^i_{\boldsymbol{j}k} : \boldsymbol{j} \in G_i, k \in \boldsymbol{i}_{\mathcal{B}_i}, i \in [r]].$$

Consider the ring homomorphism

$$\begin{aligned} \psi : \mathbb{R}[p]_{[\mathcal{T},\mathcal{T}_G]} &\to \mathbb{R}[p]_{\mathcal{T}} \otimes_{\mathbb{R}} \mathbb{R}[p]_{\mathcal{T}_G} \\ p^i_{\boldsymbol{j}k} &\mapsto p^i_{\boldsymbol{j}} \otimes p^i_k, \qquad \forall \boldsymbol{j} \in G_i, \forall k \in \boldsymbol{i}_{\mathcal{B}_i}, \forall i \in [r]. \end{aligned}$$

The ideal $\ker(\psi)$ is the toric fiber product of $\ker(\varphi_{\mathcal{T}})$ and $\langle 0 \rangle$.

Proposition 4.3.6. Let $\mathcal{T}, G, \mathcal{T}_G$ and $[\mathcal{T},\mathcal{T}_G]$ be as in Definition 4.3.4. Suppose that $\ker(\varphi_{\mathcal{T}})$ is homogeneous with respect to the multigrading given by $\mathcal{A}$. Then

$$\ker(\varphi_{[\mathcal{T},\mathcal{T}_G]}) = \ker(\varphi_{\mathcal{T}}) \times_{\mathcal{A}} \langle 0 \rangle.$$

Proof. We need to show that the map $\varphi_{[\mathcal{T},\mathcal{T}_G]}$ factorizes according to ψ. We have that

$$\begin{aligned} \varphi_{[\mathcal{T},\mathcal{T}_G]} : \mathbb{R}[p]_{[\mathcal{T},\mathcal{T}_G]} &\to \mathbb{R}[\Theta]_{[\mathcal{T},\mathcal{T}_G]} \\ p^i_{\boldsymbol{j}k} &\mapsto z \cdot \prod_{e \in E(\lambda_{\boldsymbol{j}k})} \theta(e), \quad \forall \boldsymbol{j} \in G_i, \forall k \in \boldsymbol{i}_{\mathcal{B}_i}, \forall i \in [r]. \end{aligned} \tag{4.4}$$

From the construction of $[\mathcal{T},\mathcal{T}_G]$ we can rewrite equation (4.4) as

$$\begin{aligned} p^i_{\boldsymbol{j}k} \mapsto z \cdot \prod_{e \in E(\lambda_{\boldsymbol{j}k})} \theta(e) &= z \cdot \left(\prod_{e \in E(\lambda_{\boldsymbol{j}})} \theta(e) \right) \theta(\boldsymbol{j}, \boldsymbol{j}k) \\ &= z \cdot \left(\prod_{e \in E(\lambda_{\boldsymbol{j}})} \theta(e) \right) \epsilon^{(i)}_k, \quad \forall \boldsymbol{j} \in G_i, \forall k \in \boldsymbol{i}_{\mathcal{B}_i}, \forall i \in [r]. \end{aligned}$$

Therefore, $\varphi_{[\mathcal{T},\mathcal{T}_G]}$ factors through ψ. Since $\ker(\varphi_{\mathcal{T}})$ is homogeneous with respect to the multigrading given by $\mathcal{A}$, we conclude that $\ker(\varphi_{[\mathcal{T},\mathcal{T}_G]}) = \ker(\varphi_{\mathcal{T}}) \times_{\mathcal{A}} \langle 0 \rangle$, that is $\ker(\varphi_{[\mathcal{T},\mathcal{T}_G]})$ is the toric fiber product of $\ker(\varphi_{\mathcal{T}})$ and $\ker(\varphi_{\mathcal{T}_G}) = \langle 0 \rangle$. □

Definition 4.3.7. Let $(\mathcal{T},\theta)$ be a staged tree of level m. For $1 \le q \le m$ we define $V_{\le q} := \bigcup_{i=1}^q V_i$, where $V_q := \{v \in V : l(v) = q\}$ and $E_{\le q} := \{(v,w) \in E : l(v) \le q, l(w) \le q\}$. The **staged subtree** $(\mathcal{T}^{(q)},\theta)$ is defined as the subtree $\mathcal{T}^{(q)}$ of $\mathcal{T}$ with vertex set $V_{\le q}$ and edge set $E_{\le q}$. The labeling of $\mathcal{T}^{(q)}$ is induced by $\mathcal{T}$.

Example 4.3.8. Consider the staged tree $\mathcal{T}_1$ in Figure 4.1 and set $\mathcal{T} = \mathcal{T}_1^{(3)}$. Then $\mathcal{T}$ is a staged tree with label set $\{s_0, \dots, s_9\}$. The set $\boldsymbol{i}_{\mathcal{T}}$ of indices of the leaves in $\mathcal{T}$ consists of sequences ijk where $i, j, k \in \{0,1\}$. Consider the partition of $\boldsymbol{i}_{\mathcal{T}}$ given by $G = \{G_1 = \{000, 010, 100, 110\}, G_2 = \{001, 011, 101, 111\}\}$ and let $\mathcal{T}_G = (\mathcal{B}_1, \{s_{10}, s_{11}\}) \sqcup (\mathcal{B}_2, \{s_{12}, s_{13}\})$. Using this setup in the discussion following Definition 4.3.4, we see that $\mathcal{T}_1 = [\mathcal{T}, \mathcal{T}_G]$. The two polynomial rings required for the gluing construction are

$$\mathbb{R}[p]_{\mathcal{T}} = \mathbb{R}[p^i_{\boldsymbol{j}} : \boldsymbol{j} \in G_i, i \in [2]],\ \mathbb{R}[\Theta]_{\mathcal{T}_G} = \mathbb{R}[p^i_k : k \in \boldsymbol{i}_{\mathcal{B}_i}, i \in [2]].$$

Using the set G we define a multigrading on the variables of $\mathbb{R}[p]_{\mathcal{T}}, \mathbb{R}[p]_{\mathcal{T}_G}$ as follows

$$\begin{aligned}\deg(p^1_{000}, p^1_{010}, p^1_{100}, p^1_{110}) = \deg(p^1_0, p^1_1) = e_1,\\ \deg(p^2_{001}, p^2_{011}, p^2_{101}, p^2_{111}) = \deg(p^2_0, p^2_1) = e_2,\end{aligned}$$

so that $\mathcal{A} = \{e_1, e_2\} \subseteq \mathbb{Z}^2$ is linearly independent. That way, the ideal

$$\ker(\varphi_{\mathcal{T}}) = \langle p^1_{000}p^2_{101} - p^1_{100}p^2_{001}, p^1_{010}p^2_{111} - p^1_{110}p^2_{011}\rangle$$

is homogeneous with respect to the multigrading in $\mathcal{A}$. Hence Proposition 4.3.6 yields $\ker(\varphi_{\mathcal{T}_1}) = \ker(\varphi_{\mathcal{T}}) \times_{\mathcal{A}} \langle 0\rangle$.

Example 4.3.9. Consider the staged tree $\mathcal{T}_2$ of Figure 4.1 and let $\mathcal{T} = \mathcal{T}_2^{(2)}$ be as in Definition 4.3.7. Then $\mathcal{T}$ is a staged tree with label set $\mathcal{L} = \{s_0, \dots, s_7\}$. The set $\boldsymbol{i}_{\mathcal{T}}$ of indices of the leaves in $\mathcal{T}$ consists of all strings ij with $i = 0, \dots, 3$ and $j = 0, 1$. Consider the partition of $\boldsymbol{i}_{\mathcal{T}}$ given by $G = \{G_1 = \{00, 01, 10\}, G_2 = \{20, 21, 30\}, G_3 = \{11, 31\}\}$ and let $\mathcal{T}_G = (\mathcal{B}_1, \{s_8, s_9\}) \sqcup (\mathcal{B}_2, \{s_{10}, s_{11}\}) \sqcup (\mathcal{B}_3, \{s_{12}, s_{13}\})$. Under this setup, $\mathcal{T}_2 = [\mathcal{T}, \mathcal{T}_G]$. The set G defines a multigrading on $\mathbb{R}[p]_{\mathcal{T}}$ with $\mathcal{A} = \{e_1, e_2.e_3\} \subseteq \mathbb{Z}^3$. The ideal $\ker(\varphi_{\mathcal{T}}) = \langle p^1_{00}p^3_{11} - p^1_{10}p^1_{01}, p^2_{20}p^3_{31} - p^2_{30}p^2_{21}\rangle$ is not homogeneous with respect to the multigrading in $\mathcal{A}$. Thus in this case $\ker(\varphi_{\mathcal{T}_2}) \neq \ker(\varphi_{\mathcal{T}}) \times_{\mathcal{A}} \langle 0\rangle$.

4.3.3 Inductive Tree Gluing

The information provided in Subsection 4.3.2 and particularly in Proposition 4.3.6 provides a way to construct a staged tree from smaller components. However, the new tree obtained that way is not always guaranteed to be balanced. In this paragraph we give a sufficient condition to have a balanced tree gluing. Based on this we formulate an inductive process that allows us to construct a balanced staged tree inductively in a finite number of steps.

Definition 4.3.10. Let $\mathcal{T}$ be a balanced and stratified staged tree. The partition G of the set $\boldsymbol{i}_{\mathcal{T}}$ is called **balanced** if the tree $[\mathcal{T}, \mathcal{T}_G]$ is balanced.

Example 4.3.11. The partition G in Example 4.3.8 is a balanced partition while the partition G given in Example 4.3.9 is not balanced.

Suppose now that we start with a balanced and stratified staged tree. The following result tells us how exactly we should partition the leaves of this tree to continue having a balanced and stratified staged tree.

Proposition 4.3.12. Let $\mathcal{T}, G, \mathcal{T}_G$ be like in Definition 4.3.4 and denote by $\mathcal{S}$ the set of stages in $\mathcal{T}$ that involve vertices that are parents of leaves. If every pair of vertices $\boldsymbol{v}, \boldsymbol{w}$ in $\mathcal{S}$ that are in the same stage satisfy at least one of the conditions

1. $\mathrm{ch}(\boldsymbol{v}) \subset G_i,\ \mathrm{ch}(\boldsymbol{w}) \subset G_j$ for any $i, j \in [r]$ not necessarily distinct,
2. $\{\boldsymbol{v}s, \boldsymbol{w}s\} \subset G_{i_s}$ for any $s \in \{0, \dots, k\}$ and any $i_s \in [r]$,

where $\mathrm{ch}(\boldsymbol{v}) = \{\boldsymbol{v}0, \dots, \boldsymbol{v}k\}, \mathrm{ch}(\boldsymbol{w}) = \{\boldsymbol{w}0, \dots, \boldsymbol{w}k\}$, then G is a balanced partition.

Proof. Suppose first that $\mathrm{ch}(\boldsymbol{v}) \subset G_i, \mathrm{ch}(\boldsymbol{w}) \subset G_j$ for any $i, j \in [r]$ not necessarily distinct. Then $t(\boldsymbol{v}s) = t_{\mathcal{B}_i}, t(\boldsymbol{w}s) = t_{\mathcal{B}_j}$ for any $s \in \{0, \dots, k\}$ and some one-level trees $\mathcal{B}_i, \mathcal{B}_j$ in $\mathcal{T}_G$. It follows that

$$t(\boldsymbol{v}s_1)t(\boldsymbol{w}s_2) = t_{\mathcal{B}_i}t_{\mathcal{B}_j} = t(\boldsymbol{v}s_2)t(\boldsymbol{w}s_1), \quad \forall s_1 \neq s_2 \in \{0, \dots, k\}$$

holds in $\mathbb{R}[p]_{[\mathcal{T}, \mathcal{T}_G]}$. Now assume that $\{\boldsymbol{v}s, \boldsymbol{w}s\} \subset G_{i_s}$ for any $s \in \{0, \dots, k\}$ and any $i_s \in [r]$. Then $t(\boldsymbol{v}s) = t(\boldsymbol{w}s) = t_{\mathcal{B}_{i_s}}$ for any $s \in \{0, \dots, k\}$ and some one-level tree $\mathcal{B}_{i_s}$ in $\mathcal{T}_G$. It follows that

$$t(\boldsymbol{v}s_1)t(\boldsymbol{w}s_2) = t_{\mathcal{B}_{i_{s_1}}}t_{\mathcal{B}_{i_{s_2}}} = t(\boldsymbol{w}s_1)t(\boldsymbol{v}s_2), \quad \forall s_1 \neq s_2 \in \{0, \dots, k\}$$

holds in $\mathbb{R}[p]_{[\mathcal{T}, \mathcal{T}_G]}$. □

For the rest of this paragraph, suppose that $\mathcal{T}_n$ is a balanced an stratified staged tree of level n, for some $n \in \mathbb{N}$. We will explain how $\mathcal{T}_n$ is obtained inductively from other components in a finite number of steps.

We always start this inductive construction with a one-level tree, say $\mathcal{T}_1$. If G^1 is a partition of the set $\boldsymbol{i}_{\mathcal{T}_1}$ of leaves of $\mathcal{T}_1$, and $\mathcal{T}_{G^1}$ is a gluing component, then we can form the balanced and stratified staged tree $\mathcal{T}_2 = [\mathcal{T}_1, \mathcal{T}_{G^1}]$ following the discussion in Subsection 4.3.2. If G^2 is a balanced partition of the set $\boldsymbol{i}_{\mathcal{T}_2}$, then $\ker(\varphi_{\mathcal{T}_2})$ is homogeneous with respect to the set $\mathcal{A}_2$ of multidegrees in $\mathbb{R}[p]_{\mathcal{T}_2}$. Therefore, we can form the balanced and stratified staged tree $\mathcal{T}_3 = [\mathcal{T}_2, \mathcal{T}_{G^2}]$ using Proposition 4.3.6. Continuing that way we form a staged tree $\mathcal{T}_n = [\mathcal{T}_{n-1}, \mathcal{T}_{G^{n-1}}]$ from a balanced and stratified staged tree $\mathcal{T}_{n-1}$ and a gluing component $\mathcal{T}_{G^{n-1}}$, where G^{n-1} is a balanced partition of $\boldsymbol{i}_{\mathcal{T}_{n-1}}$. The set of stages of $\mathcal{T}_n$ is exactly equal to $\bigcup_{j=1}^{n-1} G^j$. Whenever a staged tree $\mathcal{T}$ is constructed in a way such that $\mathcal{T} = \mathcal{T}_n$ for some n, we say that $\mathcal{T}$ is an **inductively constructed** staged tree.

Example 4.3.13. The staged tree $\mathcal{T}_1$ of Figure 4.1 is an inductively constructed staged tree in four steps. Precisely we have $\mathcal{T}_1 = [[[\mathcal{T}_1^{(1)}, \mathcal{T}_{1_{G^1}}], \mathcal{T}_{1_{G^2}}], \mathcal{T}_{1_{G^3}}]$, where $\mathcal{T}_1^{(1)}$ is like in Definition 4.3.7. The sets G^i and the components $\mathcal{T}_{1_{G^i}}$ for $i = 1, 2, 3$ are as follows

$$G^1 = \{\{0\},\{1\}\},\ \mathcal{T}_{1_{G^1}} = (\mathcal{B}_1^1, \{s_2, s_3\}) \sqcup (\mathcal{B}_2^1, \{s_4, s_5\}),$$
$$G^2 = \{\{00,10\},\{01,11\}\},\ \mathcal{T}_{1_{G^2}} = (\mathcal{B}_1^2, \{s_6, s_7\}) \sqcup (\mathcal{B}_2^2, \{s_8, s_9\}),$$
$$G^3 = \{\{000,010,100,110\},\{001,011,101,111\}\},\ \mathcal{T}_{1_{G^3}} = (\mathcal{B}_1^3, \{s_{10}, s_{11}\}) \sqcup (\mathcal{B}_2^3, \{s_{12}, s_{13}\}).$$

4.4 Equations via Tree Gluings

In this section we construct generators of a Gröbner basis for the toric ideal of a balanced and stratified staged tree. The main ingredient for our study is the following result from [Sul07].

Theorem 4.4.1 (Theorem 2.9, [Sul07])**.** Suppose that $\mathcal{A}$ is a linearly independent set. Let $F \subset I$ be a homogeneous Gröbner basis for I with respect to the weight vector ω_1 and let $H \subset J$ be a homogeneous Gröbner basis for J with respect to the weight vector ω_2. Let w be a weight vector such that Quad_B is a Gröbner basis for I_B. Then

$$\mathrm{Lift}(F) \cup \mathrm{Lift}(H) \cup \mathrm{Quad}_B$$

is a Gröbner basis for $I \times_{\mathcal{A}} J$ with respect to the weight order $\phi_B^*(\omega_1, \omega_2) + \epsilon\omega$, for sufficiently small $\epsilon > 0$.

This theorem has two important ingredients. The first is the set of equations denoted by Quad_B, these are equations that emerge from the construction of the toric fiber product. We focus on the description of these equations for the case of balanced and stratified staged trees in Subsection 4.4.1 and we connect them to the ideal I_{paths} from Definition 4.2.2. The second ingredient is the set $\mathrm{Lift}(F) \cup \mathrm{Lift}(H)$, these are the lifts of generators of the ideals I and J respectively. We consider the lifts of equations for inductively constructed staged trees in Subsection 4.4.2 and connect these elements to generators of the ideal I_{mpaths}.

4.4.1 Quadratic Equations

Throughout this section, let $\mathcal{T}_j$ be an inductively constructed balanced and stratified staged tree with $\mathcal{T}_m = [\mathcal{T}_{m-1}, \mathcal{T}_{G^{m-1}}]$ for any $m \in \{2, \dots, j\}$. Set $r_{m-1} := |G^{m-1}|$. We denote by $\mathcal{A}_{m-1}$ the set of multidegrees in $\mathbb{R}[p]_{\mathcal{T}_m}$. The set $\mathcal{A}_{m-1}$ is determined by the

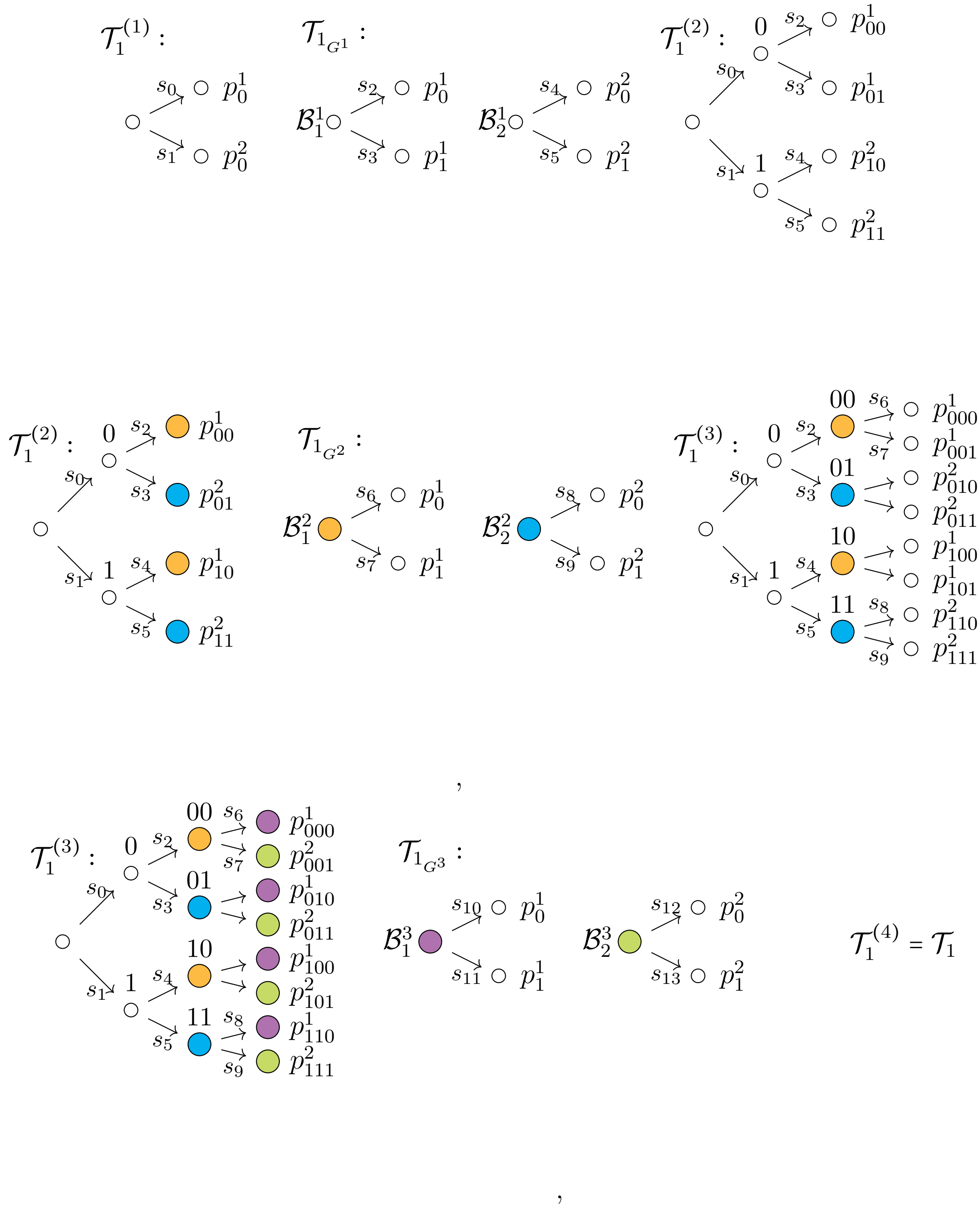


Figure 4.2: Inductive tree gluing to obtain the staged tree $\mathcal{T}_1$ of Figure 4.1.

partition G^{m-1} on the set $\boldsymbol{i}_{\mathcal{T}_{m-1}}$ which justifies the index $m-1$. Consider the monomial parametrization

$$\begin{aligned}\phi_{B_{j-1}} : \mathbb{R}[p]_{\mathcal{T}_j} &\to \mathbb{R}[p^i_{\boldsymbol{v}}, p^i_k : \boldsymbol{v} \in G_i^{j-1}, k \in \boldsymbol{i}_{\mathcal{B}_i^{j-1}}, i \in [r_{j-1}]] \\ p^i_{\boldsymbol{v}k} &\mapsto p^i_{\boldsymbol{v}} p^i_k,\end{aligned}$$

where B_{j-1} denotes the exponent matrix associated to $\phi_{B_{j-1}}$. Set $I_{B_{j-1}} := \ker(\phi_{B_{j-1}})$ and

$$\text{Quad}_{B_{j-1}} := \{\underline{p^i_{\boldsymbol{v}k_1} p^i_{\boldsymbol{w}k_2}} - p^i_{\boldsymbol{w}k_1} p^i_{\boldsymbol{v}k_2} : \boldsymbol{v}, \boldsymbol{w} \in G_i^{j-1}, k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_i^{j-1}}, i \in [r_{j-1}]\}.$$

The elements in $\text{Quad}_{B_{j-1}}$ are homogeneous with respect to the multidegrees in $\mathcal{A}_{j-1}$. Additionally, by [Sul07, Proposition 10], $\text{Quad}_{B_{j-1}}$ is a Gröbner basis of $I_{B_{j-1}}$ with respect to any term order that selects the underlined terms as leading terms. In the following proposition, we show that the elements in $\text{Quad}_{B_{j-1}}$ are exactly the path equations of the stages in the leaves of $\mathcal{T}_{j-1}$.

Proposition 4.4.2. The polynomials in $\text{Quad}_{B_{j-1}}$ are path differences coming from the stages in G^{j-1}. Moreover,

$$I_{B_{j-1}} = \sum_{i\in[r_{j-1}]} \sum_{\boldsymbol{v},\boldsymbol{w}\in G_i^{j-1}} I_{\boldsymbol{v}\sim\boldsymbol{w}}.$$

Proof. The stages on the leaves of $\mathcal{T}_j = [\mathcal{T}_{j-1}, \mathcal{T}_{G^{j-1}}]$ are determined by the partition G^{j-1} on $\boldsymbol{i}_{\mathcal{T}_{j-1}}$. Let $\boldsymbol{v}, \boldsymbol{w}$ be two leaves in $\mathcal{T}_{j-1}$. If $\boldsymbol{v}, \boldsymbol{w} \in G_i^{j-1}$ for some $i \in [r_{j-1}]$, then $\boldsymbol{v}, \boldsymbol{w}$ are in the same stage. The ideal $I_{\boldsymbol{v}\sim\boldsymbol{w}}$ is generated by polynomials $p^i_{[\boldsymbol{v}k_1]} p^i_{[\boldsymbol{w}k_2]} - p^i_{[\boldsymbol{w}k_1]} p^i_{[\boldsymbol{v}k_2]}$ for any $k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_i^{j-1}}$, where $\mathcal{B}_i^{j-1}$ is some one-level tree in $\mathcal{T}_{G^{j-1}}$. Since the vertices $\boldsymbol{v}k_1, \boldsymbol{v}k_2, \boldsymbol{w}k_1, \boldsymbol{w}k_2$ are leaves in $\mathcal{T}_j$, then

$$I_{\boldsymbol{v}\sim\boldsymbol{w}} = \langle p^i_{\boldsymbol{v}k_1} p^i_{\boldsymbol{w}k_2} - p^i_{\boldsymbol{w}k_1} p^i_{\boldsymbol{v}k_2} : k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_i^{j-1}}\rangle.$$

We have that

$$\text{Quad}_{B_{j-1}} = \bigcup_{i\in[r_{j-1}]} \bigcup_{\boldsymbol{v},\boldsymbol{w}\in G_i^{j-1}} \{p^i_{\boldsymbol{v}k_1} p^i_{\boldsymbol{w}k_2} - p^i_{\boldsymbol{w}k_1} p^i_{\boldsymbol{v}k_2} : k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_i^{j-1}}\},$$

and therefore

$$I_{B_{j-1}} = \sum_{i\in[r_{j-1}]} \sum_{\boldsymbol{v},\boldsymbol{w}\in G_i^{j-1}} \langle p^i_{\boldsymbol{v}k_1} p^i_{\boldsymbol{w}k_2} - p^i_{\boldsymbol{w}k_1} p^i_{\boldsymbol{v}k_2} : k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_i^{j-1}}\rangle = \sum_{i\in[r_j]} \sum_{\boldsymbol{v},\boldsymbol{w}\in G_i^{j-1}} I_{\boldsymbol{v}\sim\boldsymbol{w}}.$$

□

Since $\mathcal{T}_j$ is an inductively constructed staged tree, the set of stages in $\mathcal{T}_j$ is exactly the union of all stages produced in each iteration, i.e. $\bigcup_{m=1}^{j-1} G^m$. Set

$$\mathrm{Quad}_{[B_m]} = \bigcup_{i=1}^{r_m} \{p_{[\boldsymbol{v}k_1]}p_{[\boldsymbol{w}k_2]} - p_{[\boldsymbol{w}k_1]}p_{[\boldsymbol{v}k_2]} : \boldsymbol{v}, \boldsymbol{w} \in G_i^m, k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_i^m}\}, \tag{4.5}$$

for any $m \in \{1, \dots, j-1\}$. For $m = j-1$ we specialize $\mathrm{Quad}_{[B_{j-1}]} := \mathrm{Quad}_{B_{j-1}}$.

Corollary 4.4.3. If I_{paths} is the ideal of all path differences in $\mathcal{T}_j$, then

$$I_{\text{paths}} = \sum_{m=1}^{j-1} I_{[B_m]}.$$

Proof. From Proposition 4.4.2 the polynomials in Quad_{B_m} are path differences defining the stages from G^m in $\mathcal{T}_{m+1}$ for any $m = 1, \dots, j-1$. The statement in the corollary follows from the observation that the stages in $\mathcal{T}_j$ is the union of the G^m for any $m = 1, \dots, j-1$ and the fact that any staged tree $\mathcal{T}_m$ is a subtree of $\mathcal{T}_j$. □

Example 4.4.4. Consider the balanced and stratified staged tree $\mathcal{T}_1$ of Figure 4.1 that is iteratively constructed as described in Example 4.3.13. The equations defining $\mathcal{T}_1$ consist of the quadrics

$$\begin{aligned}\mathrm{Quad}_{B_3} = \{&p^i_{\boldsymbol{j}_1k_1}p^i_{\boldsymbol{j}_2k_2} - p^i_{\boldsymbol{j}_1k_2}p^i_{\boldsymbol{j}_2k_1} : i \in [2], \boldsymbol{j}_1, \boldsymbol{j}_2 \in G_i^3, k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_i^3}\}\\
= \{&p^1_{0000}p^1_{0101} - p^1_{0001}p^1_{0100}, p^1_{0000}p^1_{1001} - p^1_{0001}p^1_{1000}, p^1_{0000}p^1_{1101} - p^1_{0001}p^1_{1100},\\
&p^1_{0100}p^1_{1001} - p^1_{0101}p^1_{1000}, p^1_{0100}p^1_{1101} - p^1_{0101}p^1_{1100}, p^1_{1000}p^1_{1101} - p^1_{1001}p^1_{1100},\\
&p^2_{0010}p^2_{0111} - p^2_{0011}p^2_{0110}, p^2_{0010}p^2_{1011} - p^2_{0011}p^2_{1010}, p^2_{0010}p^2_{1111} - p^2_{0011}p^2_{1110},\\
&p^2_{0110}p^2_{1011} - p^2_{0111}p^2_{1010}, p^2_{0110}p^2_{1111} - p^2_{0111}p^2_{1110}, p^2_{1010}p^2_{1111} - p^2_{1011}p^2_{1110}\}\end{aligned}$$

that correspond to the path differences for the stages in the leaves of $\mathcal{T}_4$, and the equations $\mathrm{Quad}_{[B_m]}$ in (4.5) that define the other stages in $\mathcal{T}_4$. These are given as follows. For $m = 1$ the set $\mathrm{Quad}_{[B_1]}$ is the empty set because there are no relations between the vertices and the edge labels in $\mathcal{T}_2$ that define stages, while for $m = 2$,

$$\mathrm{Quad}_{[B_2]} = \{p_{[000]}p_{[101]} - p_{[001]}p_{[100]}, p_{[010]}p_{[111]} - p_{[011]}p_{[110]}\}.$$

The next lemma is useful when proving the main theorem of the chapter.

Lemma 4.4.5. Let $\mathcal{T}_j$ be an inductively constructed balanced and stratified staged tree and suppose that $\mathcal{T}_{j+1} = [\mathcal{T}_j, \mathcal{T}_{G^j}]$ is also balanced and stratified. Then the elements in $\mathrm{Quad}_{B_{j-1}}$ are homogeneous with respect to the multigrading in $\mathcal{A}_j$.

Proof. Since $\mathcal{T}_j$ is inductively constructed, there exists a sequence

$$(\mathcal{T}_1, \mathcal{T}_{G^1}), \ldots, (\mathcal{T}_{j-1}, \mathcal{T}_{G^{j-1}})$$

of trees and gluing components from which $\mathcal{T}_j$ is obtained. Set $i \coloneqq j-1$. From Proposition 4.4.2, the quadrics in

$$\mathrm{Quad}_{B_i} = \bigcup_{l=1}^{r_i} \{p_{\boldsymbol{v}k_1} p_{\boldsymbol{w}k_2} - p_{\boldsymbol{w}k_1} p_{\boldsymbol{v}k_2} : \boldsymbol{v}, \boldsymbol{w} \in G_l^i, k_1 \neq k_2 \in \boldsymbol{i}_{B_l^i}\}$$

are the path equations of the stages in $G^i = \{G_1^i, \ldots, G_{r_i}^i\}$ in $\mathcal{T}_{i+1}$ and they are homogeneous with respect to the multigrading in $\mathcal{A}_i$. In order to show that those equations are $\mathcal{A}_{i+1}$–homogeneous we need to jump to the construction of $\mathcal{T}_{i+2} = [\mathcal{T}_{i+1}, \mathcal{T}_{G^{i+1}}]$. This is because from the tree gluing construction, the multidegrees $\mathcal{A}_{i+1}$ are determined by the partition G^{i+1} of the set $\boldsymbol{i}_{\mathcal{T}_{i+1}}$ which allows for the construction of $\mathcal{T}_{i+2}$. An useful argument for the proof is that if $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{i}_{\mathcal{T}_{i+1}}$ are two leaves in $\mathcal{T}_{i+1}$ that belong to the same set G_α^{i+1} of the partition G^{i+1}, then they have the same degree, i.e. $\deg(p_{\boldsymbol{a}}) = \deg(p_{\boldsymbol{b}})$ in $\mathbb{R}[p]_{\mathcal{T}_{i+1}}$.
The staged tree $\mathcal{T}_{i+2}$ is balanced and stratified from the assumption. Since $\mathcal{T}_{i+1}$ is a balanced and stratified staged subtree of $\mathcal{T}_{i+2}$, all stages in G^i satisfy the balanced condition ($\star$) in $\mathbb{R}[\Theta]_{\mathcal{T}_{i+2}}$. In other words, for any $\alpha \in [r_i]$, and any vertices $\boldsymbol{v}, \boldsymbol{w} \in G_\alpha^i$ with $\mathrm{ch}(\boldsymbol{v}) = \{\boldsymbol{v}k : k \in \boldsymbol{i}_{\mathcal{B}_\alpha^i}\}$ and $\mathrm{ch}(\boldsymbol{w}) = \{\boldsymbol{w}k : k \in \boldsymbol{i}_{\mathcal{B}_\alpha^i}\}$ the equation

$$t_{(i+2)}(\boldsymbol{v}k_1) t_{(i+2)}(\boldsymbol{w}k_2) = t_{(i+2)}(\boldsymbol{w}k_1) t_{(i+2)}(\boldsymbol{v}k_2), \quad \forall k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}_\alpha^i}, \tag{4.6}$$

is valid in $\mathbb{R}[\Theta]_{\mathcal{T}_{i+2}}$. Any $\boldsymbol{u} \in \boldsymbol{i}_{\mathcal{T}_{i+1}} = \boldsymbol{i}_{[\mathcal{T}_i, \mathcal{T}_{G^i}]}$ is described as $\boldsymbol{u} \in \{\boldsymbol{v}k, \boldsymbol{w}k : k \in \boldsymbol{i}_{\mathcal{B}_\alpha^i}\}$, hence when constructing the staged tree $\mathcal{T}_{i+2} = [\mathcal{T}_{i+1}, \mathcal{T}_{G^{i+1}}]$, the interpolating polynomial $t_{(i+2)}(\boldsymbol{u})$ in $\mathcal{T}_{i+2}$ is exactly equal to the sum of edge labels of some one level probability tree $\mathcal{B}_\beta^{i+1}$ in $\mathcal{T}_{G^{i+1}}$. From the validity of equation (4.6) we distinguish between two cases for the values of the interpolating polynomials $t_{(i+2)}(\boldsymbol{v}k_1), t_{(i+2)}(\boldsymbol{w}k_2), t_{(i+2)}(\boldsymbol{w}k_1)$ and $t_{(i+2)}(\boldsymbol{v}k_2)$.

If $t_{(i+2)}(\boldsymbol{v}k_1) = t_{(i+2)}(\boldsymbol{w}k_1)$ is equal to the sum of edge labels of some one level probability tree $\mathcal{B}_\gamma^{i+1}$ and $t_{(i+2)}(\boldsymbol{v}k_2) = t_{(i+2)}(\boldsymbol{w}k_2)$ is equal to the sum of edge labels of some one level probability tree $\mathcal{B}_\delta^{i+1}$, where $\mathcal{B}_\gamma^{i+1}, \mathcal{B}_\delta^{i+1}$ are components in $\mathcal{T}_{G^{i+1}}$, then $\{\boldsymbol{v}k_1, \boldsymbol{w}k_1\}$ belong in the same partition G_γ^{i+1} of $\boldsymbol{i}_{\mathcal{T}_{i+1}}$ and $\{\boldsymbol{v}k_2, \boldsymbol{w}k_2\}$ belong in the same partition G_δ^{i+1} of $\boldsymbol{i}_{\mathcal{T}_{i+1}}$. If we associate to the sets $G_\gamma^{i+1}, G_\delta^{i+1}$ multidegrees e_γ and e_δ in $\mathcal{A}_{i+1}$ respectively, then the quadrics $p_{\boldsymbol{v}k_1} p_{\boldsymbol{w}k_2} - p_{\boldsymbol{w}k_1} p_{\boldsymbol{v}k_2} \in \mathrm{Quad}_{\mathcal{B}_i}$ are $\mathcal{A}_{i+1}$–homogeneous polynomials of degree $e_\gamma + e_\delta$. This proves the Lemma.

If $t_{(i+2)}(\boldsymbol{v}k_1) = t_{(i+2)}(\boldsymbol{v}k_2)$ and $t_{(i+2)}(\boldsymbol{w}k_1) = t_{(i+2)}(\boldsymbol{w}k_2)$, then following exactly the same arguments as in the above paragraph we prove the $\mathcal{A}_{i+1}$–homogeneity of the quadrics in $\mathrm{Quad}_{\mathcal{B}_i}$. □

4.4.2 Lifted Equations

In the previous subsection we studied the equations that emerge from the toric fiber product construction. We further noticed that the ideal I_{paths} is described by simply collecting the quadrics in $\text{Quad}_{[B_m]}$ at each step of the inductive toric fiber product. We now look at lifts of equations in $\text{Quad}_{[B_m]}$.

Let $\mathcal{T}, G, \mathcal{T}_G$ and $[\mathcal{T}, \mathcal{T}_G]$ be as in Definition 4.3.4 and denote by $\mathcal{A}$ the multigrading on the rings $\mathbb{R}[p]_{\mathcal{T}}, \mathbb{R}[p]_{[\mathcal{T},\mathcal{T}_G]}$ imposed by G. Recall the definition of the *lifting* of a homogeneous polynomial in $\mathbb{R}[p]_{\mathcal{T}}$ to the polynomial ring $\mathbb{R}[p]_{[\mathcal{T},\mathcal{T}_G]}$ of the toric fiber product. Since we only consider pure quadratic binomials, we restrict the definition of lifts provided in [Sul07] to this particular case. Let

$$f = p^{i_1}_{\boldsymbol{j}_1} p^{i_2}_{\boldsymbol{j}_2} - p^{i_1}_{\boldsymbol{j}_3} p^{i_2}_{\boldsymbol{j}_4} \in \mathbb{R}[p]_{\mathcal{T}}$$

be a homogeneous polynomial with respect to the multigrading given by $\mathcal{A}$, where $\boldsymbol{j}_1, \boldsymbol{j}_3 \in G_{i_1}, \boldsymbol{j}_2, \boldsymbol{j}_4 \in G_{i_2}$ for some $i_1, i_2 \in [r]$. Let $\mathcal{B}_{i_1}, \mathcal{B}_{i_2}$ be one-level probability trees in $\mathcal{T}_G$ and set $k = (k_1, k_2)$ with $k_1 \in \boldsymbol{i}_{\mathcal{B}_{i_1}}, k_2 \in \boldsymbol{i}_{\mathcal{B}_{i_2}}$. Consider the polynomial $f_k \in \mathbb{R}[p]_{[\mathcal{T},\mathcal{T}_G]}$ defined by

$$f_k := p^{i_1}_{\boldsymbol{j}_1 k_1} p^{i_2}_{\boldsymbol{j}_2 k_2} - p^{i_1}_{\boldsymbol{j}_3 k_1} p^{i_2}_{\boldsymbol{j}_4 k_2}.$$

Then from [Sul07] f_k is an element in $\ker(\varphi_{\mathcal{T}}) \times_{\mathcal{A}} \ker(\varphi_{\mathcal{T}_G})$. Hence, $f_k \in \ker(\varphi_{[\mathcal{T},\mathcal{T}_G]})$.

Definition 4.4.6. Let $\mathcal{A}$ be the multigrading of the rings $\mathbb{R}[p]_{\mathcal{T}}, \mathbb{R}[p]_{[\mathcal{T},\mathcal{T}_G]}$ determined by G, and let $F \in \ker(\varphi_{\mathcal{T}})$ be a collection of pure homogeneous binomials with respect to the multigrading in $\mathcal{A}$. We associate to each $f \in F$ the set $T_f = \boldsymbol{i}_{\mathcal{B}_1} \times \boldsymbol{i}_{\mathcal{B}_2}$ of indices and define

$$\text{Lift}(F) := \{f_k : f \in F, k \in T_f\}.$$

The set $\text{Lift}(F)$ is the **lifting** of F to $\ker(\varphi_{\mathcal{T}}) \times_{\mathcal{A}} \ker(\varphi_{\mathcal{T}_G})$.

The information provided in Lemma 4.4.5 allows us to talk about liftings of the pure homogeneous binomials in Quad_{B_m}.

Definition 4.4.7. Let $\mathcal{T}_j$ be an inductively constructed staged tree with $\mathcal{T}_m = [\mathcal{T}_{m-1}, \mathcal{T}_{G^{m-1}}]$ for any $m = 2, \ldots, j$. If q is a non-negative integer with $0 \le m + q \le j - 1$, define by

$$\text{Lift}^q(\text{Quad}_{\text{B}_\text{m}}) := \text{Lift}_{\mathcal{A}_{m+q}}(\ldots(\text{Lift}_{\mathcal{A}_{m+2}}(\text{Lift}_{\mathcal{A}_{m+1}}(\text{Quad}_{B_m})))\ldots),$$

the **degree q lifting of $\text{Quad}_{\text{B}_\text{m}}$.** Here the subscript $\mathcal{A}$ indicates that the argument in $\text{Lift}_{\mathcal{A}}(\cdot)$ must be homogeneous with respect to the multigrading $\mathcal{A}$.

We formulate a lemma that says that level q subtrees of balanced and stratified staged trees are also balanced. This leads us to consider interpolating polynomials of a vertex in two different rings. For a staged tree $\mathcal{T} = (V, E), \theta$ and a vertex $v \in V$ of level q, we write $t_{(j)}(v)$ for the interpolating polynomial of v in the level j subtree $(\mathcal{T}^{(j)}, \theta \mid_{E_{\leq j}})$, where $q \leq j \leq m$ and m is the level of T. Thus $t_{(j)}(v)$ is an element of $\mathbb{R}[\Theta]_{\mathcal{T}^{(j)}}$.

Lemma 4.4.8. Let $(\mathcal{T}, \theta)$ be a staged tree of level m and let q be a positive integer, $1 \leq q \leq m-1$. If $(\mathcal{T}, \theta)$ is balanced and stratified, then the level q subtree $\mathcal{T}^{(q)}$ of $\mathcal{T}$ is also balanced and stratified.

Proof. Let $\boldsymbol{v}, \boldsymbol{w}$ be two vertices in $\mathcal{T}^{(q)}$ that are in the same stage and assume that $\mathrm{ch}(\boldsymbol{v}) = \{\boldsymbol{v}0, \ldots, \boldsymbol{v}k\}, \mathrm{ch}(\boldsymbol{w}) = \{\boldsymbol{w}0, \ldots, \boldsymbol{w}k\}$. Since $(\mathcal{T}, \theta)$ is a balanced staged tree, by Definition 4.1.8, the equation

$$t(\boldsymbol{v}i)t(\boldsymbol{w}j) = t(\boldsymbol{w}i)t(\boldsymbol{v}j) \quad \forall i \neq j \in \{0, \ldots, k\} \tag{4.7}$$

is valid in $\mathbb{R}[\Theta]_{\mathcal{T}}$. In order to show that $\mathcal{T}^{(q)}$ is balanced, we need to prove that latest equation holds in $\mathbb{R}[\Theta]_{\mathcal{T}^{(q)}}$. To this end, if $\boldsymbol{a} \in V_{\leq q}$ we consider the set

$$[\boldsymbol{a}] = \{\boldsymbol{b} \in \boldsymbol{i}_{\mathcal{T}^{(q)}} : \text{the root-to-}\boldsymbol{b} \text{ paths in } \mathcal{T}^{(q)} \text{ pass through } \boldsymbol{a}\}.$$

If $\boldsymbol{c} \in \{\boldsymbol{v}i, \boldsymbol{v}j, \boldsymbol{w}i, \boldsymbol{w}j : i \neq j \in \{0, \ldots, k\}\}$ then

$$t_m(\boldsymbol{c}) = \sum_{\boldsymbol{b} \in [\boldsymbol{c}]} \prod_{e \in E(\lambda_{\boldsymbol{b}})} \theta(e) t_{(m)}(\boldsymbol{b}),$$

where $\lambda_{\boldsymbol{b}}$ is the $\boldsymbol{c}$–to-$\boldsymbol{b}$ path in $\mathcal{T}^{(q)}$, is a polynomial in $\mathbb{R}[\Theta]_{\mathcal{T}}$. Denote by $t_{(m)}(\boldsymbol{c}) \mid_{\mathcal{T}^{(q)}}$ the polynomial obtained from $t_m(\boldsymbol{c})$ when we specialize $t_m(\boldsymbol{b}) = 1$ for any vertex $\boldsymbol{b} \in [\boldsymbol{c}]$. From the assumption that $\mathcal{T}$ is stratified, it follows that $t_{(m)}(\boldsymbol{c}) \mid_{\mathcal{T}^{(q)}}$ is a polynomial in $\mathbb{R}[\Theta]_{\mathcal{T}^{(q)}}$. Precisely, $t_{(m)}(\boldsymbol{c}) \mid_{\mathcal{T}^{(q)}}$ is the interpolating polynomial $t_{(q)}(\boldsymbol{c})$ of $\boldsymbol{c}$ as a vertex in $\mathcal{T}^{(q)}$. Applying this specialization to (4.7) yields the balanced condition for the pair $\boldsymbol{v}, \boldsymbol{w}$ in $\mathbb{R}[\Theta]_{\mathcal{T}^{(q)}}$. □

Lemma 4.4.9. Let $\mathcal{T}_j$ be an inductively constructed balanced and stratified staged tree and suppose $\mathcal{T}_{j+1} = [\mathcal{T}_j, \mathcal{T}_{G^j}]$ is also balanced and stratified. Then, the elements

$$\mathrm{Lift}^{j-2}(\mathrm{Quad}_{B_1}), \mathrm{Lift}^{j-3}(\mathrm{Quad}_{B_2}), \ldots, \mathrm{Lift}(\mathrm{Quad}_{B_{j-2}}), \mathrm{Quad}_{B_{j-1}}$$

are homogeneous with respect to the multigrading in $\mathcal{A}_j$.

Proof. Since $\mathcal{T}_j$ is an inductively constructed balanced and stratified staged tree, there exists a sequence

$$(\mathcal{T}_1, \mathcal{T}_{G^1}), (\mathcal{T}_2, \mathcal{T}_{G^2}), \ldots, (\mathcal{T}_{j-1}, \mathcal{T}_{G^{j-1}})$$

of staged trees and gluing components, from which $\mathcal{T}_j$ is obtained. Moreover by Lemma 4.4.8 each of the trees $\mathcal{T}_1, \dots, \mathcal{T}_{j-1}$ is balanced. Fix $q \in \{0, \dots, j-2\}$ and $i = j-1-q$. We will show that the elements $\mathrm{Lift}^q(\mathrm{Quad}_{B_i})$ are homogeneous with respect to the multigrading in $\mathcal{A}_j$. To this end we show that for $m \in \{0, \dots, q\}$ the elements $\mathrm{Lift}^m(\mathrm{Quad}_{B_i})$ are homogeneous with respect to the multigrading in $\mathcal{A}_{i+m+1}$. We proceed by induction on m.

For $m = 0$, the elements $\mathrm{Lift}^0(\mathrm{Quad}_{B_i}) = \mathrm{Quad}_{B_i}$ are, by Lemma 4.4.5, homogeneous with respect to the multigrading in $\mathcal{A}_{i+1}$. Consequently, all the polynomials in Quad_{B_i} can be lifted to polynomials in $\ker(\varphi_{\mathcal{T}_{i+2}})$.

Suppose that $\mathrm{Lift}^{m-1}(\mathrm{Quad}_{B_i})$ is formed inductively by lifting the equations in Quad_{B_i} and at each step all the equations lift. An element in $\mathrm{Lift}^{m-1}(\mathrm{Quad}_{B_i})$ is a binomial of the form

$$f = p_{\boldsymbol{v}k_1 l} p_{\boldsymbol{w}k_2 u'} - p_{\boldsymbol{w}k_1 u} p_{\boldsymbol{v}k_2 l'},$$

for any $\alpha \in [r_i], \boldsymbol{v}, \boldsymbol{w} \in G^i_\alpha, k_1 \neq k_2 \in \boldsymbol{i}_{\mathcal{B}^i_\alpha}$, where u, u', l, l' are sequences of non-negative integer numbers of length $m-1$. These sequences arise as subindices after $m-1$ times of lifting the binomial f. The equations in $\mathrm{Lift}^{m-1}(\mathrm{Quad}_{B_i})$ define the tree $\mathcal{T}_{i+m}$. We claim that $f \in \mathrm{Lift}^{m-1}(\mathrm{Quad}_{B_i})$ is homogeneous with respect to the multigrading in $\mathcal{A}_{i+m}$.

In order to prove the claim, we follow the same treatment as for $m = 0$. We know that two elements in the same set of the partition G^{i+m} have the same degree with respect to $\mathcal{A}_{i+m}$. This condition can be verified for f checking that the condition

$$t_{(i+m+1)}(\boldsymbol{v}k_1 l) t_{(i+m+1)}(\boldsymbol{w}k_2 u') = t_{(i+m+1)}(\boldsymbol{w}k_1 u) t_{(i+m+1)}(\boldsymbol{v}k_2 l')$$

is valid in $\mathbb{R}[\Theta]_{\mathcal{T}_{i+m+1}}$. If $\boldsymbol{c} \in \{\boldsymbol{v}k_1, \boldsymbol{v}k_2, \boldsymbol{w}k_1, \boldsymbol{w}k_2\}$, then $c \in \boldsymbol{i}_{\mathcal{T}_{i+m}}$. Denote by $[c] := \{\beta \in \boldsymbol{i}_{\mathcal{T}_{i+m}} : \text{the root-to-}\beta \text{ path in } \mathcal{T}_{i+m} \text{ goes through } c\}$. Since $\mathcal{T}_{i+m}$ is a balanced and stratified staged subtree of $\mathcal{T}_{i+m+1}$, the stages in G^i must satisfy condition ($\star$) in $\mathbb{R}[\Theta]_{\mathcal{T}_{i+m+1}}$. In terms of interpolating polynomials, the equation

$$t_{(i+m+1)}(\boldsymbol{v}k_1) t_{(i+m+1)}(\boldsymbol{w}k_2) = t_{(i+m+1)}(\boldsymbol{w}k_1) t_{(i+m+1)}(\boldsymbol{v}k_2) \tag{4.8}$$

must be valid in $\mathbb{R}[\Theta]_{\mathcal{T}_{i+m+1}}$. Equation (4.8) can be rewritten as follows

$$\left(\sum_{\boldsymbol{v}k_1 l \in [\boldsymbol{v}k_1]} \left(\prod_{e \in E(\boldsymbol{v}k_1 \to \boldsymbol{v}k_1 l)} \theta(e)\right) t_{(i+m+1)}(\boldsymbol{v}k_1 l)\right) \cdot \left(\sum_{\boldsymbol{w}k_2 u' \in [\boldsymbol{w}k_2]} \left(\prod_{e \in E(\boldsymbol{w}k_2 \to \boldsymbol{w}k_2 u')} \theta(e)\right) t_{(i+m+1)}(\boldsymbol{w}k_2 u')\right) =$$
$$\left(\sum_{\boldsymbol{w}k_1 u \in [\boldsymbol{w}k_1]} \left(\prod_{e \in E(\boldsymbol{w}k_1 \to \boldsymbol{w}k_1 u)} \theta(e)\right) t_{(i+m+1)}(\boldsymbol{w}k_1 u)\right) \cdot \left(\sum_{\boldsymbol{v}k_2 l' \in [\boldsymbol{v}k_2]} \left(\prod_{e \in E(\boldsymbol{v}k_2 \to \boldsymbol{v}k_2 l')} \theta(e)\right) t_{(i+m+1)}(\boldsymbol{v}k_2 l')\right).$$

If we specialize $t_{(i+m+1)}(\boldsymbol{v}k_1 l) = t_{i+m+1}(\boldsymbol{v}k_2 l') = t_{i+m+1}(\boldsymbol{w}k_1 u) = t_{i+m+1}(\boldsymbol{w}k_2 u') = 1$, then we recover the interpolating polynomials $t_{(i+m)}(\boldsymbol{v}k_1), t_{(i+m)}(\boldsymbol{v}k_2), t_{(i+m)}(\boldsymbol{w}k_1), t_{(i+m)}(\boldsymbol{w}k_2)$

in $\mathbb{R}[\Theta]_{\mathcal{T}_{i+m}}$. In this case we have

$$\left(\sum_{\boldsymbol{v}k_1 l\in[\boldsymbol{v}k_1]}\left(\prod_{e\in E(\boldsymbol{v}k_1\to\boldsymbol{v}k_1 l)}\theta(e)\right)\right)\cdot\left(\sum_{\boldsymbol{w}k_2 u'\in[\boldsymbol{w}k_2]}\left(\prod_{e\in E(\boldsymbol{w}k_2\to\boldsymbol{w}k_2 u')}\theta(e)\right)\right)=$$
$$\left(\sum_{\boldsymbol{w}k_1 u\in[\boldsymbol{w}k_1]}\left(\prod_{e\in E(\boldsymbol{w}k_1\to\boldsymbol{w}k_1 u)}\theta(e)\right)\right)\cdot\left(\sum_{\boldsymbol{v}k_2 l'\in[\boldsymbol{v}k_2]}\left(\prod_{e\in E(\boldsymbol{v}k_2\to\boldsymbol{v}k_2 l')}\theta(e)\right)\right). \tag{4.9}$$

The factors in the above equality are sums of monomials all with coefficients equal to one. Hence, for every pair $\boldsymbol{v}k_1 l\in[\boldsymbol{v}k_1], \boldsymbol{w}k_2 u'\in[\boldsymbol{w}k_2]$ in the product of the left hand-side of (4.9), there is a pair $\boldsymbol{w}k_1 u\in[\boldsymbol{w}k_1], \boldsymbol{v}k_2 l'\in[\boldsymbol{v}k_2]$ in the product of the right hand-side, such that

$$\left(\prod_{e\in E(\boldsymbol{v}k_1\to\boldsymbol{v}k_1 l)}\theta(e)\right)\cdot\left(\prod_{e\in E(\boldsymbol{w}k_2\to\boldsymbol{w}k_2 u')}\theta(e)\right)=\left(\prod_{e\in E(\boldsymbol{w}k_1\to\boldsymbol{w}k_1 u)}\theta(e)\right)\cdot\left(\prod_{e\in E(\boldsymbol{v}k_2\to\boldsymbol{v}k_2 l')}\theta(e)\right).$$

Hence,

$$\sum_{\substack{\boldsymbol{v}k_1 l\in[\boldsymbol{v}k_1],\\ \boldsymbol{w}k_2 u'\in[\boldsymbol{w}k_2]}}\left(\prod_{\substack{e\in E(\boldsymbol{v}k_1\to\boldsymbol{v}k_1 l),\\ e'\in E(\boldsymbol{w}k_2\to\boldsymbol{w}k_2 u'}}\theta(e)\theta(e')\right)(t_{(i+m+1)}(\boldsymbol{v}k_1 l)t_{(i+m+1)}(\boldsymbol{w}k_2 u')-$$
$$-t_{(i+m+1)}(\boldsymbol{w}k_1 u)t_{(i+m+1)}(\boldsymbol{v}k_2 l'))=0.$$

Since $\mathcal{T}_{i+m+1}$ is stratified, the variables involved in the factored monomials above are disjoint from the variables involved in the factors $t_{(i+m+1)}(\boldsymbol{v}k_1 l)t_{(i+m+1)}(\boldsymbol{w}k_2 u')-t_{(i+m+1)}(\boldsymbol{w}k_1 u)t_{(i+m+1)}(\boldsymbol{v}k_2 l')$. Therefore, the last equation holds whenever

$$t_{(i+m+1)}(\boldsymbol{v}k_1 l)t_{(i+m+1)}(\boldsymbol{w}k_2 u')-t_{(i+m+1)}(\boldsymbol{w}k_1 u)t_{(i+m+1)}(\boldsymbol{v}k_2 l')=0$$

for any summand. This proves that $\mathrm{Lift}^{m-1}(\mathrm{Quad}_{B_i})$ is homogeneous with respect to $\mathcal{A}_{i+m}$. □

Proposition 4.4.10. Let $\mathcal{T}_j$ be an inductively constructed balanced and stratified staged tree. If I_{mpaths} is the ideal of maximal path extensions in $\mathcal{T}_j$, then the generating set of I_{mpaths} is exactly the union

$$\mathrm{Lift}^{j-2}(\mathrm{Quad}_{B_1})\cup\mathrm{Lift}^{j-3}(\mathrm{Quad}_{B_2})\cup\cdots\cup\mathrm{Lift}(\mathrm{Quad}_{B_{j-2}})\cup\mathrm{Quad}_{B_{j-1}}.$$

Proof. From Proposition 4.4.2, the equations in $\mathrm{Quad}_{B_{j-1}}$ correspond to path differences for the stages in G^{j-1}. Those path differences are by construction associated to maximal path extensions in $\mathcal{T}_j$. Fix two non-negative integers $q\in\{0,\ldots,j-2\}$ and

$i = j - 1 - q$ and consider the lifts $\mathrm{Lift}^q(\mathrm{Quad}_{B_i})$. From Lemma 4.4.9 the equations in $\mathrm{Lift}^q(\mathrm{Quad}_{B_i})$ are homogeneous with respect to the multigrading in $\mathcal{A}_j$. As a result, the binomials in Quad_{B_i} can be lifted to binomials in $\ker(\varphi_{\mathcal{T}_{j+1}})$. For any $\alpha \in [r_i]$ and any $\boldsymbol{v}, \boldsymbol{w} \in G^i_\alpha$, consider the binomial

$$f = p_{\boldsymbol{v}k_1}p_{\boldsymbol{w}k_2} - p_{\boldsymbol{w}k_1}p_{\boldsymbol{v}k_2} \in \mathrm{Quad}_{B_i},$$

where $k_1 \neq k_2 \in \boldsymbol{i}_{B^i_\alpha}$. The lifting of f to $\ker(\varphi_{\mathcal{T}_{j+1}})$ is a binomial

$$f_{\boldsymbol{\beta}} = p_{\boldsymbol{v}k_1\boldsymbol{\beta_1}}p_{\boldsymbol{w}k_2\boldsymbol{\beta_2}} - p_{\boldsymbol{w}k_1\boldsymbol{\beta_1}}p_{\boldsymbol{v}k_2\boldsymbol{\beta_2}},$$

where $\boldsymbol{\beta_1}, \boldsymbol{\beta_2}$ are sequences of non-negative integers of length m. These sequences arise as subindices after m times of lifting the binomial f. Since $f_{\boldsymbol{\beta}} \in \ker(\varphi_{\mathcal{T}_{j+1}})$, then $\varphi_{\mathcal{T}_{j+1}}(f_{\boldsymbol{\beta}}) = 0$. Using Definition 4.2.4 of extensions of paths, the pair of paths $(\boldsymbol{v}k_1\boldsymbol{\beta_1} \to \boldsymbol{w}k_2\boldsymbol{\beta_2}, \boldsymbol{w}k_1\boldsymbol{\beta_1} \to \boldsymbol{v}k_2\boldsymbol{\beta_2})$ is an extension of the pair $(\boldsymbol{v}k_1 \to \boldsymbol{w}k_2, \boldsymbol{w}k_1 \to \boldsymbol{v}k_2)$. In other words, $f_{\boldsymbol{\beta}}$ is an extension of the path equation f for any choice of the sequences $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$. Since $f_{\boldsymbol{\beta}}$ is a binomial for any choice of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$, we conclude that $\mathrm{Lift}^q(\mathrm{Quad}_{B_i})$ is a set of path differences corresponding to maximal path extensions in $\mathcal{T}_j$. □

Example 4.4.11. The polynomials in $\mathrm{Quad}_{[B_2]}$ of Example 4.4.4 are lifted to the binomials

$$\begin{aligned}\mathrm{Lift}(\mathrm{Quad}_{[B_2]}) = &\{p_{000s}p_{101t} - p_{001t}p_{100s}, p_{010s}p_{111t} - p_{011t}p_{110s} : s \in \boldsymbol{i}_{\mathcal{B}^3_1}, t \in \boldsymbol{i}_{\mathcal{B}^3_2}\}\\ &\{p_{0000}p_{1010} - p_{0010}p_{1000}, p_{0100}p_{1110} - p_{0110}p_{1100}, p_{0000}p_{1011} - p_{0011}p_{1000},\\ &p_{0100}p_{1111} - p_{0111}p_{1100}, p_{0001}p_{1010} - p_{0010}p_{1001}, p_{0101}p_{1110} - p_{0110}p_{1101},\\ &p_{0001}p_{1011} - p_{0011}p_{1001}, p_{0101}p_{1111} - p_{0111}p_{1101}\}.\end{aligned}$$

4.4.3 Gröbner Bases for Staged Trees

Theorem 4.4.12. Let $(\mathcal{T}, \theta)$ be a balanced and stratified staged tree. Then the ideal $\ker(\varphi_{\mathcal{T}})$ is generated by a quadratic Gröbner basis with squarefree initial ideal.

Proof. Let $\mathcal{T}$ be an inductively constructed balanced and stratified staged tree. Then $\mathcal{T} = \mathcal{T}_n$ for some $n \in \mathbb{N}$. By Proposition 4.4.10 $\mathcal{T}$ is defined by elements of the form

$$F_n = \mathrm{Lift}^{n-2}(\mathrm{Quad}_{B_1}) \cup \mathrm{Lift}^{n-3}(\mathrm{Quad}_{B_2}) \cup \cdots \cup \mathrm{Lift}(\mathrm{Quad}_{B_{n-2}}) \cup \mathrm{Quad}_{B_{n-1}}.$$

We claim that F_n is a quadratic Gröbner basis for the ideal $\ker(\varphi_{\mathcal{T}_n})$ with squarefree initial ideal. We prove the claim using induction on n.
Starting with $n = 2$, the set $F_2 = \mathrm{Quad}_{B_1}$ is by [Sul07, Proposition 10] a quadratic Gröbner basis for the ideal $\ker(\varphi_{\mathcal{T}_2}) = \ker(\varphi_{\mathcal{T}_1}) \times_{\mathcal{A}_1} \langle 0 \rangle$ with squarefree initial ideal.

Suppose that the claim is true for $n-1$ so that the elements in F_{n-1} form a quadratic Gröbner basis for the ideal $\ker(\varphi_{\mathcal{T}_{n-1}})$ with a squarefree initial ideal.
We will show that the claim holds for n. Since $\mathcal{T}_n$ is a balanced and stratified staged tree and $\mathcal{T}_{n-1}$ is a staged subtree of $\mathcal{T}_n$, it follows from Lemma 4.4.8 that $\mathcal{T}_{n-1}$ is also balanced and stratified. By Lemma 4.4.9 we know that the equations in F_{n-1} are homogeneous with respect to the multigrading in $\mathcal{A}_n$. As a result they can be lifted to equations in $\ker(\varphi_{\mathcal{T}_n})$. The ideal $\ker(\varphi_{\mathcal{T}_n})$ is generated by the set F_n. The equations in the set F_n consist of the lifts of F_{n-1} to $\ker(\varphi_{\mathcal{T}_n})$ and the equations in $\text{Quad}_{B_{n-1}}$. Hence from the induction hypothesis and [Sul07, Proposition 10] the set F_n is a Gröbner basis for $\ker(\varphi_{\mathcal{T}_n})$. Since the elements in F_n are lifts of the binomials in Quad_{B_j} for any $j = 1, \ldots, n-1$, and since those binomials have squarefree leading terms, we conclude that the initial ideal of $\langle F_n \rangle$ is squarefree and the Theorem follows. □

Theorem 4.4.12 is of great importance in convex geometry as shown in the preceding result.

Corollary 4.4.13. Let $(\mathcal{T}, \theta)$ be a balanced and stratified staged tree. If Δ is the polytope obtained as the convex hull of the lattice points in the exponent matrix of the map $\varphi_{\mathcal{T}}$, then Δ admits a regular unimodular triangulation. Moreover, the toric variety defined by $\ker(\varphi_{\mathcal{T}})$ is Cohen-Macaulay.

Proof. As shown in Theorem 4.4.12, under the assumption that $(\mathcal{T}, \theta)$ is a balanced and stratified staged tree the generating set of the ideal $\ker(\varphi_{\mathcal{T}})$ forms a quadratic Gröbner basis with a squarefree initial ideal. From [Stu96, Corollary 8.9], the latest induces a regular unimodular triangulation of Δ. □

Example 4.4.14. Consider the staged tree $\mathcal{T}_1$ of Figure 4.1. This is an example of a balanced and stratified staged tree. The map $\varphi_{\mathcal{T}_1}$ is given in Example **??** and the exponent matrix of $\varphi_{\mathcal{T}_1}$ is the following

$$
\begin{array}{c}
 \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \\ s_{10} \\ s_{11} \\ s_{12} \\ s_{13}
\end{array}
\begin{array}{cccccccccccccccc}
p_{0000} & p_{0001} & p_{0010} & p_{0011} & p_{0100} & p_{0101} & p_{0110} & p_{0111} & p_{1000} & p_{1001} & p_{1010} & p_{1011} & p_{1100} & p_{1101} & p_{1110} & p_{1111} \\
\end{array}
$$

$$
\begin{array}{c} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \\ s_{10} \\ s_{11} \\ s_{12} \\ s_{13} \end{array}
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1
\end{pmatrix}. \tag{4.10}
$$

The polytope Δ of Corollary 4.4.13 for this example is the convex hull of the vectors $m_1, \dots, m_{16}$ that are columns of the matrix (4.10), i.e. $\Delta = \mathrm{conv}(m_1, \dots, m_{16}) \subseteq \mathbb{R}^{16}$. Since the rank of the matrix (4.10) is equal to eight, Δ is a 0/1 polytope of dimension equal to eight.

4.5 Applications to Algebraic Statistics

Staged tree models are a class of graphical discrete statistical models introduced by Anderson and Smith in [SA08]. While Bayesian networks and decomposable models are defined via conditional independence statements on random variables corresponding to the vertices of a graph, staged tree models encode independence relations on the events of an outcome space represented by a tree. In the statistical literature these models are also referred to as chain event graphs. We refer the reader to the book [SGC17] and to [TSR10] to find out more about their statistical properties, practical implementation and causal interpretation. In this section we give a formal definition of these models and recall results from [DG18] and [Gör17] about their defining equations.

4.5.1 Discrete Statistical Models

The basic setting for a statistical model is the random experiment. The set of all possible outcomes of this random experiment is the *sample space*. We usually denote the sample space by Ω. Depending on the experiment Ω may be finite, countably infinite or uncountably infinite. We refer to the elements of Ω as atoms.

A *random variable* is a random number which is determined by the outcome of a random experiment. It is usually denoted by a capital letter X. Mathematically, X is a real-valued map on the sample space Ω :

$$X : \Omega \to \mathbb{R}, \quad \omega \mapsto X(\omega).$$

For the purposes of this section we consider only *discrete random variables*, meaning that the set $X(\Omega)$ of possible values for X is either finite or countably finite. We denote by $\{X = x\}$ the set of outcomes in Ω for which the value x is assigned to X :

$$\{X = x\} = \{\omega \in \Omega : X(\omega) = x\} = X^{-1}(\{x\}).$$

Let X be a discrete random variable with values in a finite set Ω. If Ω has n elements, then we assume that $\Omega = \{1, \ldots, n\}$ and we identify the probability distribution of X with a vector $p = (p_1, \ldots, p_n) \in \mathbb{R}^n$, where each coordinate p_i is the probability that $\mathbb{P}(X = i)$ for any $i \in \Omega$ as well as $p_i \geq 0$ for any $i \in \Omega$ and $\sum_{i\in\Omega} p_i = 1$. Note that the non-negativity and the sum-to-one condition of the coordinates p_i implies that $0 \leq p_i \leq 1$ for any $i \in \Omega$. The *probability simplex* is then defined as the following set

$$\Delta_{n-1} := \left\{ p \in \mathbb{R}^n : \sum_{i=1}^{n} p_i = 1,\; p_i \geq 0 \;\forall i = 1, \ldots, n \right\}.$$

Following [Sul18, Chapter 5], we define a statistical model as a family of probability distributions over a given space. Hence, when the space is discrete, a discrete statistical model is determined by a set of points lying inside the probability simplex. We often study parametric statistical models, meaning that they are defined as families of distributions over a parameter space together with certain constraints on these parameters.

Definition 4.5.1. Let $\Theta \subseteq \mathbb{R}^d, d \in \mathbb{N}$ be a finite dimensional parameter space. A **discrete and parametric statistical model** on a discrete space Ω is a set of vectors

$$\mathcal{M}_\Psi := \{p_\theta : \theta \in \Theta\} \subset \Delta^{\circ}_{n-1}$$

which lie in the $(n-1)$–dimensional probability simplex, for $n = |\Omega|$. The index Ψ in $\mathcal{M}_\Psi$ is a bijective map

$$\Psi : \Theta \to \mathcal{M}_\Psi, \quad \theta \mapsto p_\theta$$

which identifies a choice of parameters in Θ with a distribution in the model. Note that $\mathcal{M}_\Psi = \Psi(\Theta)$. The map Ψ is the **parametrization** of the model.

Statistical models are often parametrized by polynomial maps. When this is the case, these models form semialgebraic sets.

4.5.2 Conditional Independence

Let $X = (X_1, X_2, \ldots, X_n)$ be a vector of discrete random variables. If the random variable X_i has sample space $[d_i]$ for any $i = 1, \ldots, n$ then the vector X takes its values on a cartesian product space $\Sigma = \prod_{i=1}^n [d_i]$. We write $p_{x_1 \cdots x_n}$ for the probability $\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)$. If A is any subset of $[n]$, then we denote by $X_A = (X_a : a \in A)$ the subvector of A indexed by the elements of A. The sample space of X_A is $\Sigma_A = \prod_{a \in A} [d_a]$.

Definition 4.5.2. Let A, B, C be pairwise disjoint subsets of $[n]$. The random vector X_A is **conditionally independent** of X_B given X_C if and only if

$$\mathbb{P}(X_A = a, X_B = b \mid X_C = c) = \mathbb{P}(X_A = a \mid X_C = c) \cdot \mathbb{P}(X_B = b \mid X_C = c)$$

for any $a \in X_A, b \in X_B$ and $c \in X_C$. The notation $X_A \perp\!\!\!\perp X_B \mid X_C$ is used to denote that the random vector X satisfies the **conditional independence statement** that X_A is conditionally independent on X_B given X_C. When C is the empty set this reduces to the **marginal independence** between X_A and X_B.

Remark 4.5.3. For simplicity, the conditional independence statement $X_A \perp\!\!\!\perp X_B \mid X_C$ is often further abbreviated as $A \perp\!\!\!\perp B \mid C$.

Definition 4.5.4. Let $\mathcal{C}$ be a list of conditional independence statements among the variables in a vector X. The **conditional independence model**, denoted by $\mathcal{M}_{\mathcal{C}}$, is the set of all probability distributions on Σ that satisfy the conditional independence statements in $\mathcal{C}$.

A conditional independence statement $X_A \perp\!\!\!\perp X_B \mid X_C$ translates into the condition that the joint probability distribution of the variables in X satisfies a set of quadratic equations. For elements $a \in X_A, b \in X_B$ and $c \in X_C$ we set $p_{a,b,c,+} = \mathbb{P}(X_A = a, X_B = b, X_C = c)$.

Proposition 4.5.5 (Proposition 4.1.6, [Sul18])**.** If X is a discrete random variable, then the conditional independence statement $X_A \perp\!\!\!\perp X_B \mid X_C$ holds if and only if

$$p_{a_1,b_1,c,+} \cdot p_{a_2,b_2,c,+} - p_{a_1,b_2,c,+} \cdot p_{a_2,b_1,c,+} = 0$$

for any $a_1, a_2 \in \Sigma_A, b_1, b_2 \in \Sigma_B, c \in \Sigma_C$.

The *conditional independence ideal* $I_{A \perp\!\!\!\perp B \mid C}$ is the ideal generated by all the quadratic polynomials of Proposition 4.5.5. If $\mathcal{C}$ is a list of conditional independence statements, then the conditional independence ideal $I_{\mathcal{C}}$ is the sum of all conditional independence ideals associated to statements in $\mathcal{C}$.

4.5.3 The Staged Tree Model

Let $\mathcal{T} = (V, E)$ be a tree graph as in Section 4.1. In Probability theory and Statistics, tree graphs are used to represent all the possible outcomes of experiments in an efficient way. The setup for the statistical study of tree graphs is the following. Any vertex $\boldsymbol{v} \in V$ denotes a different state of the experiment and any edge $e = (\boldsymbol{v}, \boldsymbol{w}) \in E$ denotes the possibility of passing from the state $\boldsymbol{v}$ to the next state $\boldsymbol{w}$. For any $\boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}}$, every root-to-leaf path $\lambda_{\boldsymbol{j}} \in \Lambda$ corresponds to an atom in an induced sample space and depicts the history of a possible single outcome of the experiment.

In Section 4.1 we assigned labels to each of the edges in $\mathcal{T}$ using a set of labels $\mathcal{L}$ and a surjective a map $\theta : E \to \mathcal{L}$, and we denoted by $\theta_{\boldsymbol{v}}$ the set of edge labels attached to a given vertex $\boldsymbol{v}$.

Definition 4.5.6. The staged tree $(\mathcal{T}, \theta)$ is a **staged probability tree** if $\theta_{\boldsymbol{v}} \in \Delta^{\circ}_{|E(\boldsymbol{v})|-1}$ for any $\boldsymbol{v} \in V$.

If we denote by $\boldsymbol{\theta} = (\theta(e) : \theta(e) \in \mathrm{im}(\theta))$ the vector of parameters where each entry is a label in $\mathcal{L}$, then $\boldsymbol{\theta}$ is an element of the parameter space which is a product of probability simplices,

$$\Theta_{\mathcal{T}} = \left\{ \boldsymbol{\theta} : 0 < \theta(e) < 1, \sum_{e \in E(\boldsymbol{v})} \theta(e) = 1 \ \forall \boldsymbol{v} \in V \right\} = \prod_{\boldsymbol{v} \in V} \Delta^{\circ}_{|E(\boldsymbol{v})|-1}.$$

The product of edge labels along a root-to-leaf path $\lambda_{\boldsymbol{j}} \in \Lambda$, for some $\boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}}$,

$$p_{\boldsymbol{\theta}}(\lambda_{\boldsymbol{j}}) = \prod_{e \in E(\lambda_{\boldsymbol{j}})} \theta(e), \quad \boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}}$$

is an *atomic monomial.* It is shown in [Gör17, Proposition 1.6] that under the additional assumption that $\mathcal{T}$ is a probability tree the atomic monomials are atomic probabilities. Consequently the vector $(p_{\boldsymbol{\theta}}(\lambda_{\boldsymbol{j}}) : \boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}})$ defines a probability distribution on $(\mathcal{T}, \theta)$. Therefore, we have the following definition for a staged tree probability model.

Definition 4.5.7. Let $(\mathcal{T}, \theta)$ be a staged tree and let $\boldsymbol{\theta}$ be the vector of parameters associated to the tree. A **staged tree model** $\mathcal{M}_{(\mathcal{T},\theta)}$ is the image of the map

$$\begin{aligned} \Psi_{\mathcal{T}} : \Theta_{\mathcal{T}} &\to \mathcal{M}_{(\mathcal{T},\theta)} \\ \boldsymbol{\theta} &\mapsto p_{\boldsymbol{\theta}} = (p_{\boldsymbol{\theta}}(\lambda_{\boldsymbol{j}}) : \boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}}). \end{aligned}$$

From the discussion preceding Definition 4.5.7 it follows that $\mathcal{M}_{(\mathcal{T},\theta)} \subseteq \Delta^{\circ}_{|\boldsymbol{i}_{\mathcal{T}}|-1}$. Two staged trees $(\mathcal{T}, \theta)$ and $(\mathcal{T}', \theta')$ are *statistically equivalent* if there exists a bijection between the sets $\Lambda_{\mathcal{T}}$ and $\Lambda_{\mathcal{T}'}$ in such a way that the image of $\Psi_{\mathcal{T}}$ is equal to the image of $\Psi_{\mathcal{T}'}$ under this bijection.

Remark 4.5.8. Given an edge $e = (\boldsymbol{v}, \boldsymbol{w})$ in $\mathcal{T}$, the label $\theta(e)$ is known in the statistical literature as the transition probability of passing from the state $\boldsymbol{v}$ to the state $\boldsymbol{w}$ given arrival at $\boldsymbol{v}$.

A staged tree model $\mathcal{M}_{(\mathcal{T},\theta)}$ is a discrete statistical model that can be parametrized by a map

$$\Psi_{\mathcal{T}} : \prod_{\boldsymbol{v}\in V} \Delta^{\circ}_{|E(\boldsymbol{v})|-1} \to \Delta^{\circ}_{|\boldsymbol{i}_{\mathcal{T}}|-1}$$

$$(\theta_{\boldsymbol{v}} : \boldsymbol{v} \in V) \mapsto \left(\prod_{e\in E(\lambda_{\boldsymbol{j}})} \theta(e) : \boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}} \right)$$

so that $\Psi_{\mathcal{T}}(\prod_{\boldsymbol{v}\in V} \Delta^{\circ}_{|E(\boldsymbol{v})|-1}) = \mathcal{M}_{(\mathcal{T},\theta)}$. The domain of $\Psi_{\mathcal{T}}$ is a semialgebraic set given by a product of probability simplices. As a consequence, $\mathcal{M}_{(\mathcal{T},\theta)}$ is also a semialgebraic set. An important property of the staged tree models, as noted in [Gör17], is that the only inequality constraints of the image of $\Psi_{\mathcal{T}}$ are those imposed by the probability simplex, that is $0 \leq p_{\boldsymbol{j}} \leq 1$ for any $\boldsymbol{j} \in \boldsymbol{i}_{\mathcal{T}}$ and $\sum_{\boldsymbol{j}\in\boldsymbol{i}_{\mathcal{T}}} p_{\boldsymbol{j}} = 1$.

Let $(\mathcal{T}, \theta)$ be a staged tree. In Definition 4.1.2 we defined the toric ideal associated to $(\mathcal{T}, \theta)$ as the kernel of a ring homomorphim from the polynomial ring $\mathbb{R}[p]_{\mathcal{T}}$ to the polynomial ring $\mathbb{R}[\Theta]_{\mathcal{T}}$. Now we define the toric ideal associated to the staged tree model $\mathcal{M}_{(\mathcal{T},\theta)}$. To this end let

$$\mathfrak{m} := \langle 1 - \sum_{e\in E(\boldsymbol{v})} \theta(e) : \boldsymbol{v} \in V \rangle$$

be the ideal of $\mathbb{R}[\Theta]_{\mathcal{T}}$ generated by all the sum-to-one conditions in the probability simplex, and consider the quotient ring $\mathbb{R}[\Theta]_{\mathcal{M}_{\mathcal{T}}} := \mathbb{R}[\Theta]_{\mathcal{T}}/\mathfrak{m}$. Denote by π the canonical projection from $\mathbb{R}[\Theta]_{\mathcal{T}}$ to $\mathbb{R}[\Theta]_{\mathcal{M}_{\mathcal{T}}}$.

Definition 4.5.9. Let $\mathcal{M}_{(\mathcal{T},\theta)}$ be a staged tree model and set $\overline{\varphi}_{\mathcal{T}} := \pi \circ \varphi_{\mathcal{T}}$. The ideal $\ker(\overline{\varphi}_{\mathcal{T}})$ is the **staged tree model ideal** associated to $\mathcal{M}_{(\mathcal{T},\theta)}$.

It follows from the above definition that for every staged tree $(\mathcal{T}, \theta)$ the toric staged tree ideal is contained in the staged tree model ideal, i.e. $\ker(\varphi_{\mathcal{T}}) \subseteq \ker(\overline{\varphi}_{\mathcal{T}})$. Note that it is not true in general that these two ideals are equal. However, the result in [DG18, Theorem 10] ensures that if $(\mathcal{T}, \theta)$ is a balanced staged tree, then $\ker(\varphi_{\mathcal{T}}) = \ker(\overline{\varphi}_{\mathcal{T}})$. As a consequence we have the following corollary of Theorem 4.4.12.

Corollary 4.5.10. If $(\mathcal{T}, \theta)$ is a balanced and stratified staged tree, then the ideal $\ker(\overline{\varphi}_{\mathcal{T}})$ admits a quadratic Gröbner basis with squarefree initial ideal.

Corollary 4.5.10 is relevant in statistics because of the importance of Gröbner bases in sampling. For more details on this topic we refer the interested reader to [AHT12]. In the following we provide examples of statistical models that can be represented by staged trees and to which Corollary 4.5.10 can be applied.

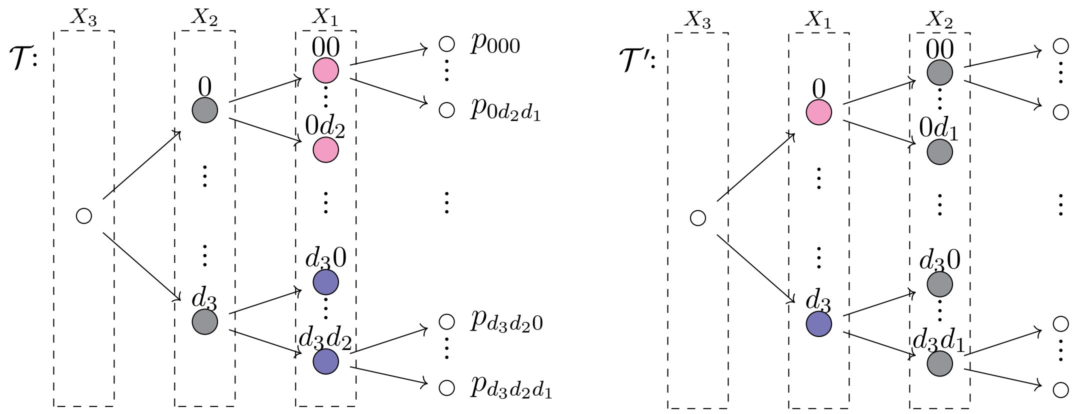


Figure 4.3: The staged trees $\mathcal{T}$ and $\mathcal{T}'$ are statistically equivalent, they represent the contraction axiom for three discrete random variables X_1, X_2 and X_3.

Example 4.5.11. Following the discussion in Subsections 1.2.2, 1.2.3 of [Gör17], the staged tree $\mathcal{T}_1$ in Figure 4.1 is the staged tree representation for the decomposable model associated to the graph $G = [12][23][34]$ on four vertices. Since the staged tree model $\mathcal{M}_{(\mathcal{T}_1,\theta_{\mathcal{T}_1})}$ coincides with the decomposable model given by G, we know from [GMS⁺06] that $\ker(\overline{\varphi}_{\mathcal{T}_1})$ has a quadratic Gröbner basis. The same result is recovered using Corollary 4.5.10.

Example 4.5.12. We consider the contraction axiom for positive distributions using staged tree models. Fix three discrete random variables X_1, X_2, X_3 with state spaces $[d_1+1], [d_2+1], [d_3+1]$ respectively. The contraction axiom states that the set of conditional independence statements $\mathcal{C} = \{X_1 \perp\!\!\!\perp X_2 \,|\, X_3, X_2 \perp\!\!\!\perp X_3\}$ implies the statement $X_2 \perp\!\!\!\perp (X_1, X_3)$. A primary decomposition of the ideal $I_{\mathcal{C}}$ was obtained in [GSS05, Theorem 1]. Here we provide a proof using staged trees, that one of the primary components of $I_{\mathcal{C}}$ is the prime binomial ideal $I_{X_2 \perp\!\!\!\perp (X_1,X_3)}$. As mentioned in [GSS05] this is a well known fact. First we explain how to represent the two statements in $\mathcal{C}$ with a staged tree. Consider the tree $\mathcal{T}$ in Figure 4.3. This tree represents the state space of the vector (X_3, X_2, X_1) as a sequence of events where X_3 takes place first, X_2 second and X_1 third. The vertices of $\mathcal{T}$ are indexed recursively as defined at the beginning of Subsection 4.3.2. The statement $X_2 \perp\!\!\!\perp X_3$ is represented by the stage consisting of the vertices $\{0, \ldots, d_3\}$, these are colored gray in $\mathcal{T}$. The statement $X_1 \perp\!\!\!\perp X_2 \,|\, X_3$ is represented by the stages $S_0, \ldots, S_{d_3}$ where $S_i = \{ij \,|\, j \in \{0, \ldots, d_2\}\}$ and $i \in \{0, \ldots, d_3\}$. These stages mean that for a given outcome of X_3 the unfolding of the event X_2 followed by X_1 behaves as an independence model on two random variables. In Figure 4.3 the stage S_0 is colored in pink and the stage S_{d_3} is colored in purple.

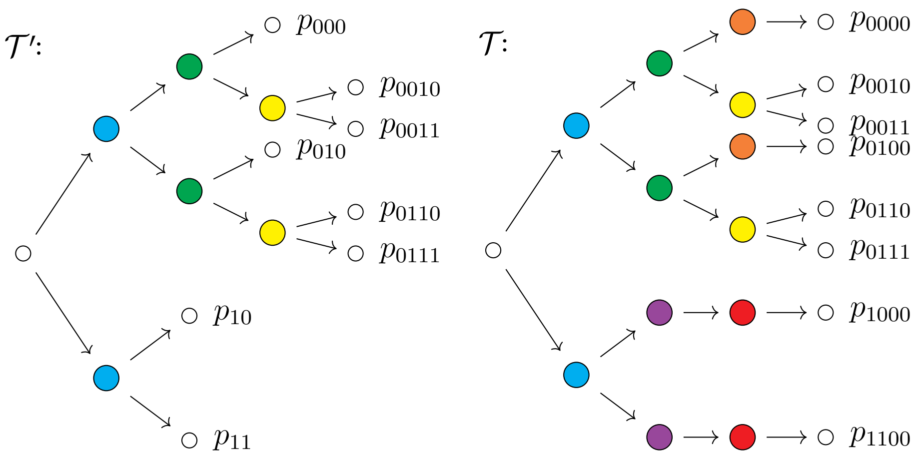


Figure 4.4: The staged trees $\mathcal{T}$ and $\mathcal{T}'$ are statistically equivalent.

Although the gray vertices are not in the same position, we can easily check that $\mathcal{T}$ is balanced and stratified. Therefore $\ker(\varphi_{\mathcal{T}})$ has a quadratic Gröbner basis. Following the proof of Theorem 2.3.14 we can construct this basis explicitly. It consists of a set of quadratic equations given by the elements in Quad_{B_2} coming from the stages in $S_0, \ldots, S_{d_3}$ and the lifts of the equations Quad_{B_1} coming from the stage $\{0, \ldots, d_3\}$. If we swap the order of X_1 and X_2 in $\mathcal{T}$, we obtain the staged tree $\mathcal{T}'$ in Figure 4.3. This tree represents the same statistical model as $\mathcal{T}$ now with the unfolding of events X_3, X_2, X_1. The gray stages in $\mathcal{T}'$ represent the statement $X_2 \perp\!\!\!\perp (X_1, X_3)$. Hence, after establishing the evident bijection between the leaves of $\mathcal{T}$ and $\mathcal{T}'$ we see that $I_{X_2 \perp\!\!\!\perp (X_1,X_3)} = \ker(\varphi_{\mathcal{T}'}) = \ker(\varphi_{\mathcal{T}})$.

The definition of staged tree in [Gör17] requires that each vertex in $\mathcal{T}$ has either no or at least two outgoing edges from v. We stepped away from making this requirement for the staged trees we consider in Section 4.1. In the next lemmas we explain how this mild extension of the definition behaves with respect to condition $(\star)$ and how trees defined according to [Gör17] are recovered from the more general trees we consider. Throughout the next lemmas, we fix a staged tree $(\mathcal{T}, \theta)$ with edge set E and define $E_1 = \{e \in E \mid E(v) = \{e\} \text{ for some } v \in V\}$. For the trees in Figure 4.4, $\mathcal{T}$ has $|E_1| = 6$ while for $\mathcal{T}'$, $|E_1| = 0$.

Lemma 4.5.13. Suppose $(\mathcal{T}, \theta)$ is a staged tree. Let $\mathcal{T}'$ be the staged tree obtained from $\mathcal{T}$ by contracting the edges in E_1. Then $\mathcal{M}_{(\mathcal{T},\theta)} = \mathcal{M}_{(\mathcal{T}',\theta)}$ and $\ker(\overline{\varphi}_{\mathcal{T}}) = \ker(\overline{\varphi}_{\mathcal{T}'})$.

Proof. First, note that the number of root-to-leaf paths in $\mathcal{T}'$ is the same as in $\mathcal{T}$. Moreover, each root-to-leaf path λ' in $\mathcal{T}'$ is obtained from a unique root-to-leaf path

λ in $\mathcal{T}$ by contracting the edges in E_1. Now let λ be a root-to-leaf path in $\mathcal{T}$. The λ-coordinate of the map $\Psi_{\mathcal{T}}$ applied to an element $\boldsymbol{\theta} \in \Theta_{\mathcal{T}}$ is

$$[\Psi_{\mathcal{T}}(\boldsymbol{\theta})]_{\lambda} = \prod_{e \in E(\lambda)} \theta(e) = \prod_{e \in E(\lambda')} \theta(e) = [\Psi_{\mathcal{T}'}(\boldsymbol{\theta}\mid_{\mathcal{T}'})]_{\lambda'}$$

The second equality in the previous equation follows from taking a closer look at $\Theta_{\mathcal{T}}$. Indeed for all $e \in E_1$ we have $\theta(e) = 1$ because of the sum-to-one conditions imposed on $\Theta_{\mathcal{T}}$ in Definition 4.5.7. For the third equality, $\boldsymbol{\theta}\mid_{\mathcal{T}'}$ denotes the restriction of the vector $\boldsymbol{\theta}$ to the edge labels of $\mathcal{T}'$. It follows from the equalities above that the coordinates of $\Psi_{\mathcal{T}}$ and $\Psi_{\mathcal{T}'}$ are equal. Therefore $\mathcal{M}_{(\mathcal{T},\theta)} = \mathcal{M}_{(\mathcal{T}',\theta)}$. A similar argument applied to the maps $\overline{\varphi}_{\mathcal{T}}$ and $\overline{\varphi}_{\mathcal{T}'}$ shows that $\ker(\overline{\varphi}_{\mathcal{T}}) = \ker(\overline{\varphi}_{\mathcal{T}'})$. To carry out this argument we need to reindex the leaves of the trees, this can be done by dropping the index of the elements in E_1. □

We illustrate Lemma 4.5.13 in Figure 4.4 where $\mathcal{T}'$ is obtained from $\mathcal{T}$ by contracting the six edges in E_1. The two staged trees in this figure define the same statistical model.

Lemma 4.5.14. Suppose $(\mathcal{T},\theta)$ is a balanced and stratified staged tree. Let $\mathcal{T}'$ be the tree obtained from $\mathcal{T}$ by contracting the edges in E_1. Then $(\mathcal{T}',\theta)$ is also balanced.

Proof. Suppose $\mathcal{T}$ is balanced and $\boldsymbol{v}, \boldsymbol{w}$ are in the same stage. Following the notation from Definition 4.1.8, we have $t(\boldsymbol{v}i)t(\boldsymbol{w}j) = t(\boldsymbol{w}i)t(\boldsymbol{w}j)$ in $\mathbb{R}[\Theta]_{\mathcal{T}}$, for all $i \neq j \in \{0,1,\ldots,k\}$. If we specialize $\theta(e) = 1$ in this equation for all $e \in E_1$ and since $\mathcal{T}'$ is stratified, then $t(\boldsymbol{v}i)t(\boldsymbol{w}j)\mid_{\theta(e)=1, e\in E_1} = t(\boldsymbol{w}i)t(\boldsymbol{v}j)\mid_{\theta(e)=1,e\in E_1}$ in $\mathbb{R}[\Theta]_{\mathcal{T}'}$. Therefore $\mathcal{T}'$ is also balanced. □

Corollary 4.5.15. Suppose $\mathcal{T}$ is a balanced and stratified staged tree with $|E_1| > 1$. Let $\mathcal{T}'$ be the staged tree obtained from $\mathcal{T}$ by contracting the edges in E_1. Then $\ker(\overline{\varphi}_{\mathcal{T}'})$ is a toric ideal with a quadratic Gröbner basis whose intial ideal is squarefree.

Proof. From Corollary 4.5.10 it follows that $\ker(\overline{\varphi}_{\mathcal{T}})$ is a toric ideal with a quadratic Gröbner basis and squarefree initial ideal. Using Lemma 4.5.13, $\ker(\overline{\varphi}_{\mathcal{T}}) = \ker(\overline{\varphi}_{\mathcal{T}'})$ and the corollary follows. □

We illustrate the result in Corollary 4.5.15 with an example.

Example 4.5.16. Fix $\mathcal{T}$ and $\mathcal{T}'$ to be the staged trees in Figure 4.4. The staged tree $\mathcal{T}'$ is considered in [DG18, Section 6] as an example of the possible unfolding of events in a cell culture. A thorough discussion of this example and its difference with other graphical models is also contained in [DG18, Section 6]. Here we explain how to obtain a Gröbner basis for $\ker(\varphi_{\mathcal{T}'})$ using Corollary 4.5.10. The tree $\mathcal{T}'$ is balanced and statistically equivalent to $\mathcal{T}$. By Corollary 4.5.10, $\mathcal{T}$ has a quadratic Gröbner basis with square free initial ideal. Using the lemmas preceding this example, there is a bijection between the root-to-leaf paths in $\mathcal{T}$ and $\mathcal{T}'$ thus $\mathbb{R}[p]_{\mathcal{T}}$ and $\mathbb{R}[p]_{\mathcal{T}'}$ are isomorphic. Under this isomorphism, the Gröbner basis for $\ker(\overline{\varphi}_{\mathcal{T}})$ is a Gröbner basis for $\ker(\overline{\varphi}_{\mathcal{T}'})$ its generators are

$$p_{0111}p_{10} - p_{0011}p_{110},\ p_{0011}p_{0110} - p_{0010}p_{0111},\ p_{0110}p_{10} - p_{0010}p_{110},$$
$$p_{0010}p_{010} - p_{000}p_{0110},\ p_{0011}p_{010} - p_{000}p_{0111},\ p_{010}p_{10} - p_{000}p_{110}.$$

Bibliography

[AH07] Mathias Aschenbrenner and Christopher J. Hillar. Finite generation of symmetric ideals. *Transactions of the American Mathematical Society*, 359(11):5171–5192, 2007.

[AH08] Matthias Aschenbrenner and Christopher J. Hillar. An algorithm for finding symmetric Gröbner bases in infinite dimensional rings [extended abstract]. In *ISSAC 2008*, pages 117–123. ACM, New York, 2008.

[AH20] Tigran Ananyan and Melvin Hochster. Small subalgebras of polynomial rings and Stillman's Conjecture. *J. Amer. Math. Soc.*, 33(1):291–309, 2020.

[AHT12] Satoshi Aoki, Hisayuki Hara, and Akimichi Takemura. *Markov Bases in Algebraic Statistics*, volume 199. 01 2012.

[Awo10] Steve Awodey. *Category Theory*. Oxford University Press, Inc., USA, 2nd edition, 2010.

[BD09] Andries E. Brouwer and Jan Draisma. Equivariant Gröbner bases and the Gaussian two-factor model, 2009.

[BG09] Winfried Bruns and Joseph Gubeladze. *Polytopes, rings, and K-theory*. Springer Monographs in Mathematics. Springer, Dordrecht, 2009.

[Buc65] B. Buchberger. Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal. Dissertation an dem Math. Inst. der Universität von Innsbruck, 1965.

[Coh67] D. E. Cohen. On the laws of a metabelian variety. *J. Algebra*, 5:267–273, 1967.

[Coh87] Daniel E. Cohen. Closure relations, Buchberger's algorithm and polynomials in infinitely many variables. *Computation theory and Logic*, 270:78–87, 1987.

[DA19] Eliana Duarte and Lamprini Ananiadi. Gröbner bases for staged trees. 2019.

[DBM10] Maria Donten-Bury and Mateusz Michalek. Phylogenetic invariants for group-based models. *arXiv preprint arXiv:1011.3236*, 2010.

[DEKL13] Jan Draisma, Rob Eggermont, Robert Krone, and Anton Leykin. Noetherianity for infinite-dimensional toric varieties. *Algebra and Number Theory*, 9, 06 2013.

[DES17] Harm Derksen, Rob H. Eggermont, and Andrew Snowden. Topological noetherianity for cubic polynomials. *Algebra Number Theory*, 11(9):2197–2212, 2017.

[DG18] Eliana Duarte and Christiane Görgen. Equations defining probability tree models. *Journal of Symbolic Computation (to appear), arXiv:1802.04511*, 2018.

[Dic13] Leonard Eugene Dickson. Finiteness of the Odd Perfect and Primitive Abundant Numbers with n Distinct Prime Factors. *Amer. J. Math.*, 35(4):413–422, 1913.

[DLL19] Jan Draisma, MichałLasoń, and Anton Leykin. Stillman's conjecture via generic initial ideals. *Comm. Algebra*, 47(6):2384–2395, 2019.

[DLO06] Jesús A. De Loera and Shmuel Onn. Markov bases of three-way tables are arbitrarily complicated. *J. Symbolic Comput.*, 41(2):173–181, 2006.

[DLST95] Jesús A. De Loera, Bernd Sturmfels, and Rekha R. Thomas. Gröbner bases and triangulations of the second hypersimplex. *Combinatorica*, 15(3):409–424, 1995.

[Dra10] Jan Draisma. Finiteness for the k-factor model and chirality varieties. *Adv. Math.*, 223(1):243–256, 2010.

[Dra14] Jan Draisma. *Noetherianity up to Symmetry*, pages 33–61. Springer International Publishing, 2014.

[Dra19] Jan Draisma. Topological Noetherianity of polynomial functors. *J. Amer. Math. Soc.*, 32(3):691–707, 2019.

[DS95] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Annals of Statistics*, 26:363–397, 1995.

[DSS07] Mathias Drton, Bernd Sturmfels, and Seth Sullivant. Algebraic factor analysis: tetrads, pentads and beyond. *Probab. Theory Related Fields*, 138(3-4):463–493, 2007.

[Eis95] D. Eisenbud. *Commutative Algebra with a view toward Algebraic Geometry.* Springer, Berlin, 1995.

[EKS14] Alexander Engström, Thomas Kahle, and Seth Sullivant. Multigraded commutative algebra of graph decompositions. *Journal of Algebraic Combinatorics*, 39(2):335–372, 2014.

[Emm87] Phillip Emmott. *Some decision problems in group theory and ring theory.* PhD thesis, Queen Mary, University of London, 1987.

[ESS19] Daniel Erman, Steven V. Sam, and Andrew Snowden. Big polynomial rings and Stillman's conjecture. *Invent. Math.*, 218(2):413–439, 2019.

[FT] Forrest Thurman. "On well-quasi-orderings", (2013). *HIM* 990-2015. 1474. (https://stars.library.ucf.edu/honorstheses1990-2015/1474).

[GBRS18] Christiane Görgen, Anna Bigatti, Eva Riccomagno, and Jim Q. Smith. Discovery of statistical equivalence classes using computer algebra. *Internat. J. Approx. Reason.*, 95:167–184, 2018.

[GMS+06] Dan Geiger, Christopher Meek, Bernd Sturmfels, et al. On the toric algebra of graphical models. *The Annals of Statistics*, 34(3):1463–1492, 2006.

[GN18] Sema Güntürkün and Uwe Nagel. Equivariant Hilbert series of monomial orbits. *Proc. Amer. Math. Soc.*, 146(6):2381–2393, 2018.

[Gor99] P. Gordan. Neuer Beweis des Hilbertschen Satzes über homogene Funktionen. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1899:240–242, 1899.

[Gör17] Christiane Görgen. *An algebraic characterisation of staged trees: their geometry and causal implications.* PhD thesis, University of Warwick, Department of Statistics, 2017.

[GS18] Christiane Görgen and Jim Q. Smith. Equivalence classes of staged trees. *Bernoulli*, 24(4A):2676–2692, 2018.

[GSS05] Luis David Garcia, Michael Stillman, and Bernd Sturmfels. Algebraic geometry of bayesian networks. *Journal of Symbolic Computation*, 39(3-4):331–355, 2005.

[HdC16] Christopher J. Hillar and Abraham Martín del Campo. Corrigendum to "Finiteness theorems and algorithms for permutation invariant chains of Laurent lattice ideals" [J. Symb. Comput. 50 (March 2013) 314–334] [MR2996883]. *J. Symbolic Comput.*, 74:650–652, 2016.

[Hig52] Graham Higman. Ordering by Divisibility in Abstract Algebras. *Proceedings of the London Mathematical Society*, s3-2(1):326–336, 1952.

[HKL] Chris Hillar, Robert Krone, and Anton Leykin. EquivariantGB: Equivariant Groebner bases and related algorithms. Version 0.2. Available at https://github.com/Macaulay2/M2/tree/master/M2/Macaulay2/packages.

[HKL18] Christopher J. Hillar, Robert Krone, and Anton Leykin. Equivariant Gröbner bases. In *The 50th anniversary of Gröbner bases*, volume 77 of *Adv. Stud. Pure Math.*, pages 129–154. Math. Soc. Japan, Tokyo, 2018.

[HS12] Christopher J. Hillar and Seth Sullivant. Finite Gröbner bases in infinite dimensional polynomial rings and applications. *Advances in Mathematics*, 229:1–25, 2012.

[JM01] J.C. Robson J.C. McConnell. *Noncommutative Noetherian Rings*, volume 30. American Mathematical Society, Providence, RI, 2001.

[KKL14] Thomas Kahle, Robert Krone, and Anton Leykin. Equivariant lattice generators and markov bases. *Proceedings of the International Symposium on Symbolic and Algebraic Computation, ISSAC*, 01 2014.

[KLS16] Robert Krone, Anton Leykin, and Andrew Snowden. Hilbert series of symmetric ideals in infinite polynomial rings via formal languages. *Journal of Algebra*, 06 2016.

[KR14] Thomas Kahle and Johannes Rauh. Toric fiber products versus Segre products. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 84, pages 187–201. Springer, 2014.

[Kru72] Joseph B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *J. Combinatorial Theory Ser. A*, 13:297–305, 1972.

[LNNR18] Dinh Le, Uwe Nagel, Hop Nguyen, and Tim Römer. Castelnuovo-Mumford regularity up to symmetry, 2018. arXiv preprint arXiv:1806.00457.

[LNNR19] Dinh Le, Uwe Nagel, Hop Nguyen, and Tim Römer. Codimension and projective dimension up to symmetry. *Mathematische Nachrichten*, 12 2019.

[MN91] Peter Meyer-Nieberg. *Banach lattices*. Universitext. Springer-Verlag, Berlin, 1991.

[MN19] Aida Maraj and Uwe Nagel. Equivariant Hilbert series for hierarchical models. *arXiv preprint arXiv:1909.13026*, 2019.

[Mul97] Bernd Mulansky. Tensor products of convex cones. In Günther Nürnberger, Jochen W. Schmidt, and Guido Walz, editors, *Multivariate Approximation and Splines*, pages 167–176, Basel, 1997. Birkhäuser Basel.

[Mur19] Satoshi Murai. Betti tables of monomial ideals fixed by permutations of the variables. 2019.

[NW63] C. St. J. A. Nash-Williams. *On well-quasi-ordering finite trees*, volume 59, pages 833–835. 1963.

[PS09] Irena Peeva and Mike Stillman. Open problems on syzygies and Hilbert functions. *J. Commut. Algebra*, 1(1):159–195, 2009.

[RN17] Tim Römer and Uwe Nagel. Equivariant hilbert series in non-noetherian polynomial rings. *Journal of Algebra*, 486:204–245, 2017.

[RN19] Tim Römer and Uwe Nagel. FI- and OI- modules with varying coefficients. *Journal of Algebra*, 535:286–322, 2019.

[RSU67] Ernst Ruch, Alfred Schönhofer, and Ivar Ugi. Die Vandermondesche Determinante als Näherungsansatz für eine Chiralitätsbeobachtung, ihre Verwendung in der Stereochemie und zur Berechnung der optischen Aktivität. *Theoretica Chimica Acta*, 7(5):420–432, 1967.

[SA08] Jim Q Smith and Paul E Anderson. Conditional independence and chain event graphs. *Artificial Intelligence*, 172(1):42–68, 2008.

[Sch74] Helmut H. Schaefer. *Banach lattices and positive operators*. Springer-Verlag, New York-Heidelberg, 1974. Die Grundlehren der mathematischen Wissenschaften, Band 215.

[SGC17] Jim Q Smith, Christiane Görgen, and Rodrigo A Collazo. Chain event graphs. *preparation for Chapman & Hall*, 2017.

[Stu96] Bernd Sturmfels. *Gröbner bases and convex polytopes*, volume 8. American Mathematical Soc., 1996.

[Sul07] Seth Sullivant. Toric fiber products. *Journal of Algebra*, 316(2):560–577, 2007.

[Sul18] Seth Sullivant. *Algebraic statistics*, volume 194 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2018.

[TSR10] Peter Thwaites, Jim Q. Smith, and Eva Riccomagno. Causal analysis with chain event graphs. *Artificial Intelligence*, 174(12-13):889–909, 2010.

[TYL01] Tsit-Yuen Lam. *A first course in noncommutative rings*, volume 131. Springer, New York, 2001.

Ehrenerklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; verwendete fremde und eigene Quellen sind als solche kenntlich gemacht.

Ich habe insbesondere nicht wissentlich:

- Ergebnisse erfunden oder widersprüchliche Ergebnisse verschwiegen,
- statistische Verfahren absichtlich missbraucht, um Daten in ungerechtfertigter Weise zu interpretieren,
- fremde Ergebnisse oder Veröffentlichungen plagiiert oder verzerrt wiedergegeben.

Mir ist bekannt, dass Verstöße gegen das Urheberrecht Unterlassungs- und Schadenersatzansprüche des Urhebers sowie eine strafrechtliche Ahndung durch die Strafverfolgungsbehörden begründen kann.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

(Magdeburg, 28. April 2020)

(Unterschrift)

Erklärung zur strafrechtlichen Verurteilung

Ich erkläre hiermit, nicht wegen einer Straftat verurteilt worden zu sein, die Wissenschaftsbezug hat.

(Magdeburg, 28. April 2020)

(Unterschrift)

Declaration of Honor

I hereby declare that I produced this thesis without prohibited assistance and that all sources of information that were used in producing this thesis, including my own publications, have been clearly marked and referenced

In particular I have not wilfully:

- Fabricated data or ignored or removed undesired results.
- Misused statistical methods with the aim of drawing other conclusions than those warranted by the available data.
- Plagiarised data or publications or presented them in a disorted way.

I know that violations of copyright may lead to injunction and damage claims from the author or prosecution by the law enforcement authorities.

This work has not previously been submitted as a doctoral thesis in the same or a similar form in Germany or in any other country. It has not previously been published as a whole.

(Magdeburg, April 28, 2020)

(Signature)

Declaration of criminal convictions

I hereby declare that I have not been found guilty of scientific and/or academic misconduct.

(Magdeburg, April 28, 2020)

(Signature)