Title
Combinatorial analysis of continuous problems

Permalink
https://escholarship.org/uc/item/0jt7p5sp

Author
Tsukerman, Emmanuel

Publication Date
2017

Peer reviewed|Thesis/dissertation
Combinatorial analysis of continuous problems

by

Emmanuel Tsukerman

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Sturmfels, Bernd, Co-chair
Professor Williams, Lauren K, Co-chair
Professor Guntuboyina, Adityanand
Professor Rezakhanlou, Fraydoun

Spring 2017
Combinatorial analysis of continuous problems

Copyright 2017
by
Emmanuel Tsukerman
Abstract

Combinatorial analysis of continuous problems

by

Emmanuel Tsukerman

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Sturmfels, Bernd, Co-chair

Professor Williams, Lauren K, Co-chair

Many objects in mathematics, at first sight, seem to belong to the domain of continuous mathematics. These objects are continuous, smooth and infinite, far different from the discrete and finite objects that are the classical domain of combinatorics. Objects of the former type are, for instance, determinants of matrices (which can take on every complex value), Grassmannians (which are smooth manifolds), and the eigenvalues of matrices (which take on any tuple of complex values). In the latter class lie objects such as paths in graphs, finite groups and generating functions. Applications of the study of such finite objects to the continuous ones would seem unlikely, or at least, trivial. For example, one may count the number of minors of a matrix, but that’s about it. As we will demonstrate, however, this is not the case. The field of combinatorics has developed into a mature field of study, and it is the author’s view that combinatorics can be used as a toolbox to obtain interesting and deep information on all areas of mathematics, continuous especially.

In this work, we will demonstrate this by studying three different continuous problems using the techniques of combinatorics. The first problem concerns the study of symmetric matrices and their principal- and almost-principal minors. Here the main result is a proof of a conjectural combinatorial formula of Kenyon and Pemantle (2014) for the entries of a square matrix in terms of its connected principal and almost-principal minors. The second problem is the study of Bruhat interval polytopes. These polytopes arise as the moment-map images of Richardson varieties of flag varieties. Their study is motivated in part by the integrable system called the Toda lattice. Information obtained about these polytopes can be readily translated to information about the Richardson varieties. For instance, the dimension of the polytope will be used to determine when the Richardson variety is toric. The third problem will pertain to the study of the spectral theory of tensors via tropical methods. We show that an elegant theory in which there is a unique tropical eigenvalue arises. We describe briefly how the corresponding eigenvalue informs us of the asymptotic behavior of the corresponding classical eigenvalues.
To my parents
## Contents

Contents   
List of Figures   
1 introduction   
2 Symmetric matrices, Catalan paths, correlations   
  2.1 Introduction   
  2.2 Square matrices and tilings of the half Aztec diamond   
  2.3 Square matrices and Schröder paths   
  2.4 Back to symmetric matrices   
  2.5 Parametrizing Correlation Matrices   
3 Bruhat interval polytopes   
  3.1 Introduction   
  3.2 Background   
  3.3 The generalized lifting property for the symmetric group   
  3.4 Results on Bruhat interval polytopes   
  3.5 A generalization of the recurrence for $R$-polynomials   
  3.6 Background on partial flag varieties $G/P$   
  3.7 Gelfand-Serganova strata, total positivity, and Bruhat interval polytopes for $G/P$   
4 Tropical spectral theory of tensors   
  4.1 Background   
  4.2 Main Results   
  4.3 H-eigenvalues of Tensors   
  4.4 H-eigenvalues of Symmetric Tensors   
  4.5 circulation polytope   
  4.6 Future Directions   
5 Conclusion   
Bibliography
List of Figures

2.1 A Catalan path $C$ in the planar graph $G_4$ with weight $a_{13}a_{23}a_{24}a_{34}$. .......... 4
2.2 The half Aztec diamond $HD_4(2,7)$. The white boxes are to be tiled. .......... 5
2.3 The six tilings of the colored half Aztec diamond $HD_4(2,7)$. ..................... 6
2.4 The graph $G'_4$ encodes the Schröder paths of order 4. ..................... 7
2.5 How to construct a Schröder path from a tiling. ................................. 8
2.6 The six Schröder paths in $S_4$ together with their weights. .................... 9
2.7 A flip of a tiling and the corresponding local move on Schröder paths. 9
2.8 This local move multiplies the weight of the Schröder path by $\frac{b_3}{d_5}$. ........ 10
2.9 The Schröder paths (left) are projected to the Catalan paths (right). 11
2.10 The standard D-vine for four random variables. ................................. 15

3.1 The two polytopes are the permutohedron $Q_{e,w_0} = \text{Perm}_4$, and the Bruhat interval
polytope $Q_{u,v}$ with $v = (2,4,3,1)$ and $u = (1,2,4,3)$. 19
3.2 Generalized lifting property ....................................................... 24
3.3 Example of Theorem 3.3.3 ....................................................... 27
3.4 Bruhat interval $[1234, 1432]$. .................................................. 31
3.5 Bruhat interval $[1234, 3412]$. .................................................. 31
3.6 Bruhat interval $[143265, 254163]$. ............................................. 42
3.7 Bruhat interval $[1324, 4312]$. .................................................. 42
Acknowledgments

I would like to express my gratitude to my advisors, Prof. Bernd Sturmfels and Prof. Lauren Williams, whose expertise and understanding added considerably to my graduate experience and led to the completion of this thesis. I would also like to thank Prof. Fraydoun Rezakhanlou and Prof. Aditya Guntuboyina for taking time out of their busy schedules to serve as part of my committee. I would like to thank my family for the support they provided me through my entire life and in particular, I must acknowledge my mother and father for their love and encouragement. I would also like to thank Dan Daniel Erdmann-Pham for assistance in proof-reading the dissertation. Lastly, I recognize that this research would not have been possible without the financial assistance of NSF (Graduate Research Fellowship), SIAM and the Berkeley math department and express my gratitude to those agencies.
Chapter 1

introduction

Many objects in mathematics, at first sight, seem to belong to the domain of continuous mathematics. These objects are continuous, smooth and infinite, far different from the discrete and finite objects that are the classical domain of combinatorics. Objects of the former type are, for instance, determinants of matrices (which can take on every complex value), Grassmannians (which are smooth manifolds), and the eigenvalues of matrices (which take on any tuple of complex values). In the latter class lie objects such as paths in graphs, finite groups and generating functions. Applications of the study of such finite objects to the continuous ones would seem unlikely, or at least, trivial. For example, one may count the number of minors of a matrix, but that’s about it. As we will demonstrate, however, this is not the case. The field of combinatorics has developed into a mature field of study, and it is the author’s view that combinatorics can be used as a toolbox to obtain interesting and deep information on all areas of mathematics, continuous especially.

In this work, we will demonstrate this by studying three different continuous problems using the techniques of combinatorics. The first problem concerns the study of symmetric matrices and their principal- and almost-principal minors. Here the main result is a proof of a conjectural combinatorial formula of Kenyon and Pemantle (2014) for the entries of a square matrix in terms of its connected principal and almost-principal minors. The second problem is the study of Bruhat interval polytopes. These polytopes arise as the moment-map images of Richardson varieties of flag varieties. Their study is motivated in part by the integrable system called the Toda lattice. Information obtained about these polytopes can be readily translated to information about the Richardson varieties. For instance, the dimension of the polytope will be used to determine when the Richardson variety is toric. The third problem will pertain to the study of the spectral theory of tensors via tropical methods. We show that an elegant theory in which there is a unique tropical eigenvalue arises. We describe briefly how the corresponding eigenvalue informs us of the asymptotic behavior of the corresponding classical eigenvalues.
We begin with our first problem, on the relationship between principal- and almost-principal minors. The sets of matrices and symmetric matrices are topologically just $\mathbb{R}^n$ for an appropriate $n$, and hence are continuous objects. Their minors are polynomials and are continuous functions. Yet, it is fruitful to study the relations between the minors of a matrix using the techniques of combinatorics. Kenyon and Pemantle (2014) gave a formula for the entries of a square matrix in terms of connected principal and almost-principal minors. Each entry is an explicit Laurent polynomial whose terms are the weights of domino tilings of a half Aztec diamond. They conjectured an analogue of this parametrization for symmetric matrices, where the Laurent monomials are indexed by Catalan paths. In this chapter based on [STW16] with Sturmfels and Williams, we give a proof of the Kenyon-Pemantle conjecture, and relate this to a statistics problem pioneered by Joe (2006). Correlation matrices are represented by an explicit bijection from the cube to the elliptope.

2.1 Introduction

In this section we present a formula for each entry of a symmetric $n \times n$ matrix $X = (x_{ij})$ as a Laurent polynomial in $\binom{n+1}{2}$ distinguished minors of $X$. Our result verifies a conjecture of Kenyon and Pemantle from [KP14]. Let $I$ and $J$ be subsets of $[n] = \{1, 2, \ldots, n\}$ with $|I| = |J|$. Let $X_I^J$ denote the minor of $X$ with row indices $I$ and column indices $J$. Here the indices in $I$ and $J$ are always taken in increasing order. The following signed minors will be used:

\[ p_I := (-1)^{|I|/2} \cdot X_I^I \]

and

\[ a_{ij|I} := (-1)^{|I|/2} \cdot X_{ij|I}^{j|I} \quad \text{for } i, j \not\in I, \quad i \neq j. \]

Here $j|I := \{j\} \cup I$. We call $p_I$ and $a_{ij|I}$ the principal and almost-principal minors, respectively. The minors $p_I$, $a_{ij|I}$ and $a_{ji|I}$ are called connected if $1 \leq i < j \leq n$ and $I = \{i+1, i+2, \ldots, j-2, j-1\}$. Note that $p_I$ is not connected when $1$ or $n$ is in $I$. The $1 \times 1$-minors $a_{ij} := a_{ij|\emptyset} = x_{ij}$ and $p_k = x_{kk}$ are connected when $|i - j| = 1$ and $1 \leq k \leq n$.

These definitions make sense for every $n \times n$ matrix $X$, even if $X$ is not symmetric. A general $n \times n$ matrix $X$ has $2^n$ principal minors, of which $\binom{n-2}{2} + n$ are connected. It also
has \( n(n - 1)2^{n-2} \) almost-principal minors, of which \( n(n - 1) \) are connected. A symmetric \( n \times n \) matrix has \( \binom{n}{2}2^{n-2} \) distinct almost-principal minors \( a_{ij} \), of which \( \binom{n}{2} \) are connected.

A Catalan path \( C \) is a path in the \( xy \)-plane which starts at \((0, 0)\) and ends on the \( x \)-axis, always stays at or above the \( x \)-axis, and consists of steps northeast \((1, 1)\) and southeast \((1, -1)\). We say that \( C \) has size \( n \) if its endpoints have distance \( 2n - 2 \) from each other. Let \( C_n \) denote the set of Catalan paths of size \( n \). Its cardinality equals the Catalan number

\[
|C_n| = \frac{1}{n} \binom{2n-2}{n-1}, \quad \text{which is } 1, 2, 5, 14, 42, 132, 429, 1430, 4862 \text{ for } n = 2, \ldots, 10.
\]

Let \( G_n \) denote the planar graph whose nodes are the \( \binom{n+1}{2} \) lattice points \((x, y)\) with \( x \geq y \geq 0 \) and \( x + y \leq 2n - 2 \) even, and edges are northeast and southeast steps. Thus \( C_n \) consists of the paths from \((0, 0)\) to \((2n-2, 0)\) in \( G_n \). We label the nodes and regions of \( G_n \) as follows. We assign the label \( j \) to the node \((2j-2, 0)\), the label \( a_{ij} \) to the node \((i + j - 2, j - i)\), and the label \( p_i \) to the region below that node. Here, \( I = \{i + 1, i + 2, \ldots, j - 1\} \). Thus, in the planar graph \( G_n \), the connected principal and almost-principal minors of \( X \) are identified with the regions and nodes that are strictly above the \( x \)-axis.

The weight \( W_C(C) \) of a Catalan path \( C \) is a Laurent monomial, derived from the drawing of \( C \) in the graph \( G_n \). Its numerator is the product of the labels \( a_{ij} \) of the nodes of \( G_n \) that are local maxima or local minima of \( C \), and its denominator is the product of the labels \( p_i \) of the regions which are either immediately below a local maximum or immediately above a local minimum. Thus \( W_C(C) \) is a Laurent monomial of degree \( \leq 1 \). There is no lower bound on the degree due to minima on the \( x \)-axis; for instance, \( \frac{a_{13}a_{34}}{p_2p_3p_4p_5p_6p_7p_8} \) has degree \(-3\) and appears for \( n = 9 \), associated to the path \( UDUDUUDDUUDD \).

The following result was conjectured by Kenyon and Pemantle in [KP14, Conjecture 1].

**Theorem 2.1.1.** The entries of an \( n \times n \) symmetric matrix \( X = (x_{ij}) \) satisfy the identity

\[
x_{ij} = \sum_C W_C(C),
\]

where the sum is over all Catalan paths \( C \) between node \( i \) and node \( j \) in \( G_n \).

For symmetric matrices of size \( n = 4 \), Theorem 2.1.1 states the following formula:

\[
X = \left( \begin{array}{cccc}
p_1 & a_{12} & \frac{a_{13}}{p_2} & \frac{a_{12}a_{23}}{p_2} \\
* & p_2 & a_{23} & \frac{a_{12}a_{23}}{p_2} \\
* & * & p_3 & \frac{a_{23}a_{34}}{p_3} \\
* & * & * & p_4
\end{array} \right)
\]

(2.2)

The entry \( x_{14} = x_{41} \) is the sum of five Laurent monomials, one for each Catalan path from node 1 to node 4. The last term \( \frac{a_{13}a_{23}a_{24}}{p_2p_3p_4} \) equals \( W_C(C) \) for the path \( C \) shown in Figure 2.1.

Note that the formula for entries \( x_{ij} \) in Theorem 2.1.1 is a Laurent polynomial in the principal and almost principal minors. For arbitrary (as opposed to symmetric) \( n \times n \) matrices, Kenyon and Pemantle stated and proved an analogue of Theorem 2.1.1, which is also a Laurent polynomial in the parameters. In the case of arbitrary matrices, the
Laurentness is a manifestation of the cluster algebra structure which underlies the relations among the principal and almost principal minors of a generic $n \times n$ matrix. This cluster algebra structure is explained in [KP14]. When one passes to symmetric matrices, we lose the cluster algebra structure – the principal and almost principal minors used in Theorem 2.1.1 are no longer algebraically independent – but surprisingly, (2.1) is still a Laurent polynomial.

The proof of Theorem 2.1.1 is given in Section 2.4. We start in Section 2.2 by reviewing a theorem of Kenyon and Pemantle [KP14] which expresses the entries of an arbitrary square matrix in terms of almost-principal and principal minors, as a sum of Laurent monomials that are in bijection with domino tilings of a half Aztec diamond. In Section 2.3, we give a bijection between these domino tilings and Schröder paths, and restate their theorem using Schröder paths. We then prove our theorem by constructing a projection from Schröder paths to Catalan paths and applying the relation (2.7) among minors of symmetric matrices.

In Section 2.5 we connect Theorem 2.1.1 to an application in statistics, developed in work of Joe, Kurowicka and Lewandowski [Joe06, LKJ09]. Namely, we focus on symmetric matrices that are positive definite and have all diagonal entries equal to 1. These are the correlation matrices, and they form a convex set that is known in optimization as the elliptope [BPT13, LP96]. Our formula yields an explicit bijection between the elliptope and the open cube $(-1, 1)^n$.

2.2 Square matrices and tilings of the half Aztec diamond

In this section we review the Kenyon-Pemantle formula in [KP14, Theorem 4.4]. The half Aztec diamond $HD_n$ of order $n$ is the union of the unit squares whose nodes are in the set

$$\{(a, b) \in \mathbb{Z}^2 : |a| \leq n, 0 \leq b \leq n, |a| + |b| \leq n + 1\}.$$ 

We label the boxes in the bottom row of $HD_n$ by the numbers 1 through $2n$, from left to right. We label certain lattice points of $HD_n$ by minors as follows. Fix $b \in [n]$. The connected principal minors $p_I$ such that $|I| = b$ are assigned to the lattice points $(a, b)$ with
a + b even. The connected almost-principal minors \( a_{ij} \) with \( i > j \) and \( |I| = b - 1 \) are assigned to the lattice points \((a, b)\) with \(a + b\) odd. In both cases, the assignment is from left to right using the lexicographic order on \( I \). The case \( n = 4 \) is shown in Figure 2.2.

Figure 2.2: The half Aztec diamond \( HD_4(2, 7) \). The white boxes are to be tiled.

Fix integers \( a \) and \( b \) such that \( a \) is even, \( b \) is odd, and \( 1 < a < b < 2n \). We define the colored half Aztec diamond \( HD_n(a, b) \) by coloring the boxes of \( HD_n \) black, grey, or white. First color boxes \( a \) and \( b \) in the bottom row black. Let \( L_a \) be the diagonal line of slope 1 through box \( a - 1 \), and let \( L_b \) be the line of slope \(-1\) through box \( b + 1 \). If a box (or any part of it) lies to the left of \( L_a \) or to the right of \( L_b \), then color it grey. All other boxes are white. A domino tiling (or simply a tiling) of \( HD_n(a, b) \) is a tiling of the white boxes by \( 1 \times 2 \) and \( 2 \times 1 \) rectangles. Let \( \mathcal{A}_n(a, b) \) denote the set of tilings of \( HD_n(a, b) \). Figure 2.3 shows the set \( \mathcal{A}_4(2, 7) \), i.e. the six tilings of \( HD_4(2, 7) \), with lines \( L_2 \) and \( L_7 \) superimposed on the tilings.

Each tiling \( T \) of the colored half Aztec diamond \( HD_n(a, b) \) gets a Laurent monomial weight, which we now define. We regard \( T \) as a simple graph whose nodes are the lattice points of \( HD_n \), and whose edges are induced by the edges of the rectangles in the tiling together with the edges of the unit squares outside the tiling. An interior lattice point of \( HD_n(a, b) \) is a lattice point which lies strictly to the right of \( L_a \), strictly to the left of \( L_b \), and above the line \( x = 0 \). The interior lattice points that will concern us are shown in bold in Figures 2.2 and 2.3. Each of these is labeled by a variable \( v_\ell \) which is a connected principal or almost-principal minor. The weight \( W_\mathcal{A}(T) \) of a tiling \( T \in \mathcal{A}_n(a, b) \) is defined to be the Laurent monomial

\[
W_\mathcal{A}(T) := \prod_\ell v_\ell^{d(\ell) - 3},
\]

where \( \ell \) ranges over the interior lattice points of \( HD_n(a, b) \) and \( d(\ell) \) is the degree of \( \ell \) in \( T \).

**Theorem 2.2.1** (Kenyon-Pemantle [KP14]). The entries of an \( n \times n \) matrix \( X = (x_{ij}) \) satisfy

\[
x_{ij} = \sum_{T \in \mathcal{A}_n(2j, 2i - 1)} W_\mathcal{A}(T) \quad \text{for } i > j.
\]
Theorem 4.4 in [KP14] also gives an analogous formula for \( x_{ij} \) with \( i < j \), but we omit it, as the statement would require giving several technical definitions which are not needed here.

**Example 2.2.2.** Figure 2.3 shows the six tilings of \( HD_4(2, 7) \) with their weights. By Theorem 2.2.1, the upper right matrix entry for \( n = 4 \) is the sum of these six Laurent monomials:

\[
x_{41} = \frac{a_{21}a_{32}a_{43}}{p_2p_3} + \frac{a_{31}a_{24}a_{33}}{p_2p_3} + \frac{a_{21}a_{42}a_{33}}{p_2p_3a_{32}} + \frac{a_{31}a_{24}a_{23}}{p_2p_3a_{32}} + \frac{a_{41}a_{23}}{p_2p_3a_{32}}.
\]

(2.3)

The full \( 4 \times 4 \) matrix is shown on page 8 of [KP14], albeit with different notation.

### 2.3 Square matrices and Schröder paths

In this section we continue our discussion of arbitrary square matrices. A **large Schröder path** \( S \) is a path in the \( xy \)-plane which starts at \((0, 0)\), always stays at or above the \( x \)-axis,
and consists of steps which are either northeast \((1,1)\), southeast \((1,-1)\), or horizontal \((2,0)\); cf. [Sta99]. For brevity, we will henceforth omit the adjective “large”. A Schröder path has order \(n\) if it ends at \((2n-4,0)\). Let \(G'_n\) denote the planar graph whose nodes are the lattice points \((x,y)\) with \(0 \leq y \leq x\) and \(x + y \leq 2n-4\) even, with edges given by northeast, southeast and horizontal steps. The set \(\mathcal{S}_n\) of Schröder paths of order \(n\) is identified with the left-to-right paths in \(G'_n\) from \((0,0)\) to \((2n-4,0)\). The cardinality of \(\mathcal{S}_n\) is the Schröder number, which is given by the generating function seen in [Sta99, Exercise 6.40, page 241]:

\[
\sum_{n=2}^{\infty} |\mathcal{S}_n| z^{n-2} = \frac{1 - z - \sqrt{1 - 6z + z^2}}{2z} = 1 + 2z + 6z^2 + 22z^3 + 90z^4 + 394z^5 + 1806z^6 + \cdots
\]

The graph \(G'_n\) is labeled by connected minors. We assign \(a_{ij}\) to the node \((i+j-3, i-j-1)\) for \(i > j\), and we assign \(p_I\) to the triangle below that node. We refer to \((2i-2,0)\) as node \(i\). Figure 2.4 shows the case \(n = 4\). The six Schröder paths in \(\mathcal{S}_4\) are shown in Figure 2.6.

![Figure 2.4: The graph \(G'_4\) encodes the Schröder paths of order 4.](image)

We now define the weight \(W_{\mathcal{S}}(S)\) of a Schröder path \(S\) on \(G'_n\). We regard \(S\) as a graph with nodes \(V(S)\) and edges \(E(S)\). Given a Schröder path \(S\) on \(G'_n\), we define the sets

\[
\begin{align*}
\alpha(S) &= \{ v \in V(S) : v \text{ is a weak local maximum of } S \}, \\
\beta(S) &= \{ e \in E(S) : e \text{ is immediately below a weak local minimum of } S \}, \\
\gamma(S) &= \{ e \in E(S) : e \text{ is a horizontal edge of } S \}, \\
\delta(S) &= \{ v \in V(S) : v \text{ is immediately below a horizontal edge of } S \}, \\
\epsilon(S) &= \{ e \in E(S) : e \text{ is immediately below a strict local maximum of } S \}, \\
\zeta(S) &= \{ v \in V(S) : v \text{ is a strict local minimum (but not an endpoint) of } S \}.
\end{align*}
\]

Each of these is regarded as a monomial by taking the product of all labels. Then we define

\[
W_{\mathcal{S}}(S) = \frac{\alpha(S)\beta(S)}{\gamma(S)\delta(S)\epsilon(S)\zeta(S)}.
\] (2.4)

Figure 2.6 shows the six Schröder paths for \(n = 4\), together with their weights. The sum of these weights is the Laurent polynomial in (2.3), which evaluates to the matrix entry \(x_{41}\).

The main result of this section is a reformulation of Theorem 2.2.1 in terms of Schröder paths. We write \(\mathcal{S}_n(a,b)\) for the set of all Schröder paths from node \(a\) to node \(b\) in \(G'_n\).
Theorem 2.3.1. The entries of an $n \times n$ matrix $X = (x_{ij})$ satisfy
\[
x_{ij} = \sum_{S \in \mathcal{A}_n(j,i-1)} W_S(S) \quad \text{for } i > j.
\]

We shall present a weight-preserving bijection $\Phi : \mathcal{A}_n(2j,2i-1) \rightarrow \mathcal{S}_n(j,i-1)$ between tilings and Schröder paths. Note that we can superimpose the graph $G'_n$ on the graph $HD_n$ so that the labels (connected minors) match up. When we do this, the node $j$ (respectively, $i-1$) of $G'_n$ gets identified with the top right corner of the square $2j$ (respectively, the top left corner of the square $2i-1$) in $HD_n$. We draw a Schröder path $\Phi(T)$ on top of a tiling $T$, as in Figure 2.5. We may then think of the path as an element of $\mathcal{S}_n(j,i-1)$.

![Figure 2.5: How to construct a Schröder path from a tiling.](image)

More formally, given $T \in \mathcal{A}_n(2j,2i-1)$, the path $\Phi(T) \in \mathcal{S}_n(j,i-1)$ is defined as follows. Its starting point is the top right corner of square $2j$ in $HD_n(2j,2i-1)$. We inductively add steps to $\Phi(T)$ depending on the local behavior of the tiling, as shown in Figure 2.5. Let $x$ denote the endpoint of the path that we have built so far. Then we proceed as follows:

- If there is a vertical tile to the east of $x$, then we add a northeast step to our path.
- If there is a vertical tile to the southeast of $x$, such that $x$ is at its northwest corner, then we add a southeast step to our path.
- If there is a horizontal tile to the southeast of $x$, then add an east step to our path.
- If $x$ is already at the top left corner of square $2i-1$, then we stop.

The map $\Phi$ maps the six tilings in Figure 2.3 to the six Schröder paths in Figure 2.6.

Lemma 2.3.2. The map $\Phi : \mathcal{A}_n(2j,2i-1) \rightarrow \mathcal{S}_n(j,i-1)$ is well-defined and is a bijection.

Proof. This is the solution to Exercise 6.49 in [Sta99], based on an idea of Dana Randall. □

Proposition 2.3.4 states that this bijection is weight-preserving. First, another lemma:

Lemma 2.3.3. The local move shown in the top of Figure 2.7 alters the weight of both the tiling and the corresponding Schröder path by the same factor: when passing from the left to the right, the exponents of $b$ and $h$ increase by 1, while those of $d$ and $f$ decrease by 1.
Figure 2.6: The six Schröder paths in $S_4$ together with their weights.

Figure 2.7: A flip of a tiling and the corresponding local move on Schröder paths.

Proof. The statement is clear by inspection for the tilings. Checking the assertion for Schröder paths is more complicated. We need to examine the various cases of what the path looks like on the left and right of the square being modified. In other words, we need to specify whether the path increases, stays flat or decreases as it enters node $d$, and ditto for when it leaves node $f$. One such case is seen in Figure 2.8. When we perform that local move, the exponents of $b$ and $h$ in (2.4) increase by 1, while the exponents of both $d$ and $f$ decrease by 1. All other cases are similar. □
Proposition 2.3.4. If $T$ is a tiling in $\mathcal{A}_n(2j, 2i-1)$, where $i > j$, then $W_\mathcal{A}(\Phi(T)) = W_\mathcal{A}(T)$.

Proof. It is well known [ST95] that two domino tilings of a simply connected region can always be connected by a sequence of flips, where a flip is the local move that switches two horizontal tiles for two vertical tiles or vice-versa, as seen in Figure 2.7.

Let $T_0$ be the tiling consisting only of horizontal tiles. The corresponding Schröder path $\Phi(T_0)$ is a horizontal path. Here, the two objects have the same weight:

$$W_\mathcal{A}(\Phi(T_0)) = W_\mathcal{A}(T_0) = \frac{a_{j+1,j}a_{j+2,j+1}a_{j+3,j+2}\ldots a_{i,i-1}}{p_{j+1}p_{j+2}p_{j+3}\ldots p_{i-1}}. \quad (2.5)$$

By Lemma 2.3.3, if $W_\mathcal{A}(\Phi(T)) = W_\mathcal{A}(T)$ and $T'$ is obtained by a flip, then $W_\mathcal{A}(\Phi(T')) = W_\mathcal{A}(T')$. Since the tilings in $\mathcal{A}_n(2j, 2i-1)$ are connected by flips, the assertion follows. \qed

Proof. [Proof of Theorem 2.3.1] This follows from Theorem 2.2.1, Lemma 2.3.2 and Proposition 2.3.4. \qed

2.4 Back to symmetric matrices

The strategy for proving Theorem 2.1.1 is to combine Theorem 2.3.1 with a projection $\pi$ from Schröder paths to Catalan paths. Let $S$ be any Schröder path in $G'_n$. The associated Catalan path $\pi(S)$ in $G_n$ is defined by

- replacing each horizontal step in $S$ with a strict local minimum, i.e. a southeast step followed by a northeast step;
- adding a northeast step at the beginning of $S$ and a southeast step at the end of $S$.

If $S$ starts at $i$ and ends at $j-1$ in $G'_n$ then $\pi(S)$ starts at $i$ and ends at $j$ in $G_n$. Figure 2.9 shows how four of the six Schröder paths in $\mathcal{A}_4'(1,3)$ map to four of the five Catalan paths in $\mathcal{C}_4(1,4)$. The two other Schröder paths in Figure 2.6 map to the Catalan path in Figure 2.1.

Theorem 2.2.1 is an immediate consequence of Theorem 2.3.1 and the following proposition.

Proposition 2.4.1. With labels of the paths coming from a symmetric matrix, the weight of a Catalan path is the sum of the weights of the Schröder paths in its preimage under the

![Figure 2.8: This local move multiplies the weight of the Schröder path by $\frac{bh}{df}$.](image)
The proof will rely on equation (2.7) and Lemma 2.4.2. Using Muir’s law of extensible minors, we obtain the following identity that expresses connected almost-principal minors of $\pi$, i.e.

$$\sum_{S \in \pi^{-1}(C)} W_{\mathcal{S}}(S) = W_{\mathcal{C}}(C).$$

(2.6)

The proof will rely on equation (2.7) and Lemma 2.4.2. Using Muir’s law of extensible minors, we obtain the following identity that expresses connected almost-principal minors of
a symmetric $n \times n$ matrix in terms of connected principal minors:

$$a_{ij}^2 - pr_{i \cup \{i,j\}} - pr_{\{i\}}pr_{\{j\}} = 0, \quad 2 \leq i < j \leq n-1, \quad I = \{i+1, \ldots, j-1\}. \quad (2.7)$$

We now use this identity to prove the following claim.

**Lemma 2.4.2.** Let $S'$ and $S$ be two Schröder paths in $\mathcal{S}_n$ that are related as shown in the bottom row of Figure 2.7 (with $S'$ on the left and $S$ on the right). If the labels come from a symmetric $n \times n$ matrix, then the resulting weights of these paths satisfy

$$W_{\mathcal{S}}(S) + W_{\mathcal{S}}(S') = \frac{e^2}{bh} W_{\mathcal{S}}(S). \quad (2.8)$$

**Proof.** The label $e$ of the local minimum in $S$ is an almost-principal minor, while $b, h, d, f$ are principal minors. By (2.7), it satisfies $e^2 = bh + df$, and hence $\frac{e^2}{bh} = 1 + \frac{df}{bh}$. By Lemma 2.3.3, we have $W_{\mathcal{S}}(S') = \frac{df}{bh} W_{\mathcal{S}}(S)$. This implies $\frac{e^2}{bh} W_{\mathcal{S}}(S) = W_{\mathcal{S}}(S) + W_{\mathcal{S}}(S')$. □

**Example 2.4.3.** Let $S'$ and $S$ be the fourth and fifth Schröder paths in Figure 2.6, with labels given by a symmetric $4 \times 4$ matrix. Using the identity $a_{23} = p_{23} + p_{2p3}$, as in (2.7), we find

$$W_{\mathcal{S}}(S) + W_{\mathcal{S}}(S') = \frac{a_{132a_{243}}}{p_{2p3a_{23}}} + \frac{a_{132a_{243}}}{p_{23a_{23}}} = \frac{a_{132a_{243}a_{23}}}{p_{2p23p3}}.$$

This explains how the six terms in (2.3) become the five terms of $x_{14}$ shown in (2.2). Namely, the weight of the Catalan path in Figure 2.1 is the sum of the fourth and fifth terms in (2.3).

**Proof.** [Proof of Proposition 2.4.1] Let $C$ be a Catalan path with $m$ local minima. It has $m + 1$ local maxima. Let $A_1, \ldots, A_m$ and $A_1', \ldots, A_{m+1}'$ denote the variables at the local minima and maxima, respectively. Let $P_1, \ldots, P_m$ and $P_1', \ldots, P_{m+1}'$ denote the region variables located directly above the minima and directly below the maxima, respectively. Then

$$W_C(C) = \frac{A_1 \cdots A_m A_1' \cdots A_{m+1}'}{P_1 \cdots P_m P_1' \cdots P_{m+1}'} \quad (2.9)$$

We also denote the region variables located directly below the local minima by $P_1', \ldots, P_m'$.

There are $2^m$ Schröder paths that project to $C$ via $\pi$. These correspond to the $2^m$ choices of either preserving a local minimum, or replacing it by a horizontal edge. We denote the Schröder paths in $\pi^{-1}(C)$ by $S_{d_1 d_2 \ldots d_m}$, where $d_i = 0$ if the local minimum at $A_i$ was preserved and $d_i = 1$ if it was replaced by a horizontal edge. By Lemma 2.4.2, we have

$$W_{\mathcal{S}}(S_{0d_2 \ldots d_m}) + W_{\mathcal{S}}(S_{1d_2 \ldots d_m}) = \frac{A_1^2}{P_1 P_1} W_{\mathcal{S}}(S_{0d_2 \ldots d_m}),$$

$$W_{\mathcal{S}}(S_{d_10d_3 \ldots d_m}) + W_{\mathcal{S}}(S_{d_11d_3 \ldots d_m}) = \frac{A_2^2}{P_2 P_2} W_{\mathcal{S}}(S_{d_10d_3 \ldots d_m}),$$

$$
\vdots
$$

By aggregating these identities, we obtain

$$\sum_{S \in \pi^{-1}(C)} W_{\mathcal{S}}(S) = \sum_{d_1, \ldots, d_m \in \{0,1\}} W_{\mathcal{S}}(S_{d_1 d_2 \ldots d_m}) = \frac{A_1^2 A_2^2 \cdots A_m^2}{P_1 P_1 P_2 P_2 \cdots P_m P_m} W_{\mathcal{S}}(S_{00 \ldots 0}).$$
But, now it follows from (2.4) and (2.9) that
\[
W_{\gamma}(S_{00...0}) = \frac{A'_1 \cdots A'_{m+1} P''_m \cdots P''_1}{P'_1 \cdots P'_{m+1} A_1 \cdots A_m} = \frac{P_1 P''_1 P_2 P''_2 \cdots P_m P''_m}{A_1^2 A_2^2 \cdots A_m^2} W_C(C).
\]

Therefore the sum of the weights of the Schröder paths in \(\pi^{-1}(C)\) is equal to \(W_C(C)\). □

**Remark 2.4.4.** The expression in Theorem 2.1.1 is not the only way to express the entries of a symmetric matrix in terms of the \(\binom{n}{2} + \binom{n-2}{2} + n\) connected almost-principal and principal minors. The prime ideal of polynomial relations among these minors is generated by the \(\binom{n+1}{2}\) quadrics in (2.7). To show this, we argue as follows. First, in Theorem 2.1.1 we have expressed the \(\binom{n+1}{2}\) algebraically independent matrix entries \(x_{ij}\) in terms of these minors. This ensures that the algebra generated by these minors has Krull dimension \(\binom{n+1}{2}\). Hence their relation ideal has codimension \(\binom{n-2}{2} = \binom{n}{2} + \binom{n-2}{2} + n - \binom{n+1}{2}\). The \(\binom{n-2}{2}\) relations (2.7) lie in that ideal and they generate a complete intersection. Our final claim is that this complete intersection is a prime ideal. We deduce this from the fact that none of the \(a_{ij} a_{jk}\) has a square root in the subalgebra generated by the principal minors. For a concrete example consider \(n = 4\). Here, our ideal of relations is the principal ideal \((a_{23}^2 - p_2 p_3 - p_{23})\).

### 2.5 Parametrizing Correlation Matrices

We now specialize to real symmetric \(n \times n\) matrices that are positive definite and have all diagonal entries equal to 1. Such matrices are known as correlation matrices. They play an important role in statistics, notably in the study of multivariate normal distributions. The set \(\mathcal{E}_n\) of all \(n \times n\) correlation matrices is an open convex set of dimension \(\binom{n}{2}\). Its closure is a convex body, known in optimization theory [BPT13, LP96] under the name elliptope.

In certain statistical applications it is desirable to generate random correlation matrices. Specifically, one wishes to sample from the uniform distribution on the elliptope \(\mathcal{E}_n\). A solution to this problem was given by Joe [Joe06] and further refined by Lewandowski et al. [LKJ09]. The underlying geometric idea is to construct a parametrization from the standard cube:

\[
\Psi : (-1,1)^n \to \mathcal{E}_n.
\]

The papers [Joe06, LKJ09] describe such maps \(\Psi\) that are algebraic and bijective, so they identify the open cube with the open elliptope. However, the construction is recursive. In what follows we revisit the formula in [Joe06] and we make it completely explicit. Remarkably, it is precisely the restriction of our Laurent polynomial parametrization in Theorem 2.1.1 to the region where all connected principal minors \(p_I\) are positive and \(p_1 = \cdots = p_n = 1\). It would be interesting to explore practical implications of our formulation for sampling from \(\mathcal{E}_n\).

Let \(X = (x_{ij})\) be a real symmetric \(n \times n\) matrix. We assume that \(X\) is positive definite, i.e. all principal minors \(p_I\) are strictly positive. In statistics, such an \(X\) serves as the covariance matrix of a normal distribution on \(\mathbb{R}^n\), whose partial correlations are given by

\[
\rho_{ij|I} = \frac{(-1)^{\lfloor |I|/2 \rfloor} a_{ij|I}}{\sqrt{p_{iI} \cdot p_{jI}}} \quad \text{where } i, j \notin I \text{ and } i < j.
\]
For $I = \emptyset$, we obtain the $\binom{n}{2}$ entries of the correlation matrix $Y = (y_{ij})$, namely
\[
y_{ij} = \rho_{ij} = \frac{a_{ij}}{\sqrt{p_i p_j}} = \frac{x_{ij}}{\sqrt{x_{ii} x_{jj}}} \quad \text{for } 1 \leq i < j \leq n.
\]

The partial correlation $\rho_{ij|I}$ in (2.10) is called connected if $I = \{i+1, i+2, \ldots, j-2, j-1\}$.

**Theorem 2.5.1.** The $\binom{n}{2}$ entries $y_{ij}$ of a correlation matrix can be written uniquely in terms of the $\binom{n}{2}$ connected partial correlations $\rho_{ij|I}$. Explicit formulas are derived from those in Theorem 2.1.1 by first replacing each occurrence of a parameter $a_{ij|I}$ by $(-1)^{|I|/2} \rho_{ij|I} \sqrt{p_i p_j}$ and thereafter replacing each occurrence of a parameter $p_{r,r+1,\ldots,s}$ by the product of the $\binom{s-r+1}{2}$ expressions $(-1)^{|I|/2}(1 - \rho^2_{ij|I})$ where $r \leq i < j \leq s$ and $I = \{i+1, i+2, \ldots, j-1\}$. The resulting map $\Psi : (\rho_{ij|I}) \mapsto (y_{ij})$ is a bijection between $(-1, 1)^{\binom{n}{2}}$ and $E_n$.

**Proof.** The replacement formula for $a_{ij|I}$ is seen in (2.10). The formula for the signed principal minors $p_{r,r+1,\ldots,s}$ in terms of connected partial correlations is due to Joe [Joe06, Theorem 1]. It can be derived by recursively applying the following version of (2.7) in concert with (2.10):
\[
p_{ij|I} = \frac{a_{ij|I}^2 - p_{iI} p_{jI}}{p_I}, \quad I = \{i+1, \ldots, j-1\}. \quad (2.11)
\]

Our formulas give an algebraic map $\Psi : (\rho_{ij|I}) \mapsto (y_{ij})$ between affine spaces of dimension $\binom{n}{2}$. This map is invertible on $E_n$ because each partial correlation $\rho_{ij|I}$ can be written via (2.10) in terms of the entries $y_{ij}$ of the correlation matrix. All partial correlations are real numbers strictly between $-1$ and $1$. The connected partial correlations $\rho_{ij|I}$ can vary freely, as explained in [Joe06, page 2179]. From this, we get the desired bijection. \(\square\)

We now illustrate our parametrization of correlation matrices in the two smallest cases.

**Example 2.5.2 ($n = 3$).** We consider the open 3-dimensional cube defined by the inequalities
\[-1 < \rho_{12}, \rho_{23}, \rho_{13|2} < 1.\]

Our bijection $\Psi$ identifies each point in this cube with a $3 \times 3$ correlation matrix:
\[
\begin{bmatrix}
y_{12} & y_{13} \\
y_{12} & 1 & y_{23} \\
y_{13} & y_{23} & 1
\end{bmatrix}
= \begin{bmatrix}
1 & \rho_{12} & \rho_{12} \rho_{23} - \rho_{13|2} (1 - \rho_{12}^2)^{\frac{1}{2}} (1 - \rho_{23}^2)^{\frac{1}{2}} \\
\rho_{12} & 1 & \rho_{12} \rho_{23} - \rho_{13|2} (1 - \rho_{12}^2)^{\frac{1}{2}} (1 - \rho_{23}^2)^{\frac{1}{2}} \\
\rho_{12} \rho_{23} - \rho_{13|2} (1 - \rho_{12}^2)^{\frac{1}{2}} (1 - \rho_{23}^2)^{\frac{1}{2}} & \rho_{12} \rho_{23} - \rho_{13|2} (1 - \rho_{12}^2)^{\frac{1}{2}} (1 - \rho_{23}^2)^{\frac{1}{2}} & 1
\end{bmatrix}.
\]

One checks that this matrix is positive definite, and, as in [Joe06, Theorem 1], its determinant
\[
\det(Y) = (1 - \rho_{12}^2)(1 - \rho_{23}^2)(1 - \rho_{13|2}^2)
\]
defines the facets of the cube. It is instructive to draw how the boundary of the cube maps onto the boundary of the elliptope $E_3$. The latter is depicted in [BPT13, Figure 5.8, page 232].
The combinatorics of our planar graph $G_n$ and its Catalan paths can be seen in a different guise in [Joe06, LKJ09]. These correspond to the structures called D-vines in these papers. Figure 2.10 shows the standard D-vine for $n = 4$. Its edges are naturally labeled with the six coordinates of the cube, namely $\rho_{12}, \rho_{23}, \rho_{34}, \rho_{13|2}, \rho_{24|3}, \rho_{14|23}$. These correspond to the six almost-principal minors $a_{ij|I}$ in the labeled graph $G_4$ in Figure 2.1.

![Figure 2.10: The standard D-vine for four random variables.](image)

**Example 2.5.3** ($n = 4$). The $4 \times 4$ correlation matrix $Y$ is obtained from the matrix $X$ in (2.2) by performing the replacements that are described in Theorem 2.5.1. We first substitute

\[
\begin{align*}
a_{12} &= \rho_{12}\sqrt{p_1p_2}, \\
a_{23} &= \rho_{23}\sqrt{p_2p_3}, \\
a_{34} &= \rho_{34}\sqrt{p_3p_4}, \\
a_{13|2} &= -\rho_{13|2}\sqrt{p_{12}p_{23}}, \\
a_{24|3} &= -\rho_{24|3}\sqrt{p_{23}p_{34}}, \\
a_{14|23} &= -\rho_{14|23}\sqrt{p_{123}p_{234}}.
\end{align*}
\]

and then we eliminate the connected principal minors as follows:

\[
\begin{align*}
p_1 &= 1, \\
p_2 &= 1, \\
p_3 &= 1, \\
p_4 &= 1, \\
p_{12} &= -(1 - \rho_{12}^2), \\
p_{23} &= -(1 - \rho_{23}^2), \\
p_{34} &= -(1 - \rho_{34}^2), \\
p_{123} &= -(1 - \rho_{12}^2)(1 - \rho_{23}^2)(1 - \rho_{13|2}^2) \\
\text{and} \\
p_{234} &= -(1 - \rho_{23}^2)(1 - \rho_{34}^2)(1 - \rho_{24|3}^2).
\end{align*}
\]

This results in the formulas for the six entries of $Y$ in terms of $\rho_{12}, \rho_{23}, \rho_{34}, \rho_{13|2}, \rho_{24|3}, \rho_{14|23}$. These give the bijection $\Psi$ between the cube and the elliptope, both of dimension six. It is instructive to verify that $\det(Y)$ is the product of the facet-defining expressions $(1 - \rho_{\bullet})^2$. 

Chapter 3

Bruhat interval polytopes

We consider now our second problem, of studying Bruhat interval polytopes. The work here is based on joint work with L. Williams in [TW15]. Bruhat interval polytopes arise as the moment-map images of Richardson varieties of flag varieties. By studying the polytopes using the combinatorial theory of Coxeter groups, we will be able to obtain information about Richardson varieties, “continuous objects”. For example, whether they are toric varieties.

Let $u$ and $v$ be permutations on $n$ letters, with $u \leq v$ in Bruhat order. A Bruhat interval polytope $Q_{u,v}$ is the convex hull of all permutation vectors $z = (z(1), z(2), \ldots, z(n))$ with $u \leq z \leq v$. Note that when $u = e$ and $v = w_0$ are the shortest and longest elements of the symmetric group, $Q_{e,w_0}$ is the classical permutohedron. Bruhat interval polytopes were studied recently in [KW15] by Kodama and Williams, in the context of the Toda lattice and the moment map on the flag variety.

We give an inequality description and a dimension formula for Bruhat interval polytopes, and prove that every face of a Bruhat interval polytope is a Bruhat interval polytope. A key tool in the proof of the latter statement is a generalization of the well-known lifting property for Coxeter groups. Motivated by the relationship between the lifting property and $R$-polynomials, we also give a generalization of the standard recurrence for $R$-polynomials. Finally, we define a more general class of polytopes called Bruhat interval polytopes for $G/P$, which are moment map images of (closures of) totally positive cells in $(G/P)_{\geq 0}$, and are a special class of Coxeter matroid polytopes. Using tools from total positivity and the Gelfand-Serganova stratification, we show that the face of any Bruhat interval polytope for $G/P$ is again a Bruhat interval polytope for $G/P$.

3.1 Introduction

The classical permutohedron is the convex hull of all permutation vectors $(z(1), z(2), \ldots, z(n)) \in \mathbb{R}^n$ where $z$ is an element of the symmetric group $S_n$. It has many beautiful properties: its edges are in bijection with cover relations in the weak Bruhat order; its faces can be described explicitly; it is the Minkowski sum of matroid polytopes; it is the moment map image of the complete flag variety.

The main subject of this paper is a natural generalization of the permutohedron called a Bruhat interval polytope. Let $u$ and $v$ be permutations in $S_n$, with $u \leq v$ in (strong)
CHAPTER 3. BRUHAT INTERVAL POLYTOPES

Bruhat order. The *Bruhat interval polytope* (or *pairmutohedron*) $Q_{u,v}$ is the convex hull of all permutation vectors $z = (z(1), z(2), \ldots, z(n))$ with $u \leq z \leq v$. Note that when $u = e$ and $v = w_0$ are the shortest and longest elements of the symmetric group, $Q_{e,w_0}$ is the classical permutohedron. Bruhat interval polytopes were recently studied in [KW15] by Kodama and the second author, in the context of the Toda lattice and the moment map on the flag variety $\mathcal{Fl}_n$. A basic fact is that $Q_{u,v}$ is the moment map image of the Richardson variety $\mathcal{R}_{u,v} \subset \mathcal{Fl}_n$. Moreover, $Q_{u,v}$ is a Minkowski sum of matroid polytopes (in fact of *positroid polytopes* [ARW16]) [KW15], which implies that $Q_{u,v}$ is a *generalized permutohedron* (in the sense of Postnikov [Pos09]).

The goal of this paper is to study combinatorial aspects of Bruhat interval polytopes. We give a dimension formula for Bruhat interval polytopes, an inequality description of Bruhat interval polytopes, and prove that every face of a Bruhat interval polytope is again a Bruhat interval polytope. In particular, each edge corresponds to some edge in the (strong) Bruhat order. The proof of our result on faces uses the classical result (due to Edelman [Ede81] in the case of the symmetric group, and subsequently generalized by Proctor [Pro82] and then Björner-Wachs [BW82]) that the order complex of an interval in Bruhat order is homeomorphic to a sphere. Our proof also uses a generalization of the lifting property, which appears to be new and may be of interest in its own right. This Generalized lifting property says that if $u < v$ in $S_n$, then there exists an *inversion-minimal* transposition $(ik)$ (see Definition 3.3.2) such that $u \leq v(ik) \leq v$ and $u < u(ik) \leq v$. One may compare this with the usual lifting property, which says that if $u < v$ and the simple reflection $s_i \in D_r(v) \setminus D_r(u)$ is a right-descent of $v$ but not a right-descent of $u$, then $u \leq vs_i < v$ and $u < us_i \leq v$. Note that in general such a simple reflection $s_i$ need not exist.

The usual lifting property is closely related to the $R$-polynomials $R_{u,v}(q)$. Recall that the $R$-polynomials are used to define Kazhdan-Lusztig polynomials [KL79], and also have an interesting geometric interpretation: the Richardson variety $\mathcal{R}_{u,v}$ may be defined over a finite field $\mathbb{F}_q$, and the number of points it contains is given by the $R$-polynomial $R_{u,v}(q) = \#\mathcal{R}_{u,v}(\mathbb{F}_q)$. A basic result about the $R$-polynomials is that if $s_i \in D_r(v) \setminus D_r(u)$, then $R_{u,v}(q) = qR_{us,vs}(q) + (q-1)R_{u,vs}(q)$. We generalize this result, showing that if $t = (ik)$ is inversion-minimal, then $R_{u,v}(q) = qR_{ut,vt}(q) + (q-1)R_{u,vt}(q)$.

Finally we give a generalization of Bruhat interval polytopes in the setting of partial flag varieties $G/P$. More specifically, let $G$ be a semisimple simply connected linear algebraic group with torus $T$, and let $P = P_J$ be a parabolic subgroup of $G$. Let $W$ be the Weyl group, $W_J$ the corresponding parabolic subgroup of $W$, and let $t$ denote the Lie algebra of $T$. Let $\rho_J$ be the sum of fundamental weights corresponding to $J$, so that $G/P$ embeds into $\mathbb{P}(V_{\rho_J})$. Then given $u \leq v$ in $W$, where $v \in W_J$ is a minimal-length coset representative in $W/W_J$, we define the corresponding *Bruhat interval polytope for $G/P$* to be

$$Q^J_{u,v} := \text{Conv}\{z \cdot \rho_J \mid u \leq z \leq v\} \subset t^*_R.$$  

In the $\mathcal{Fl}_n$ case – i.e. the case that $G = \text{SL}_n$ and $P$ is the Borel subgroup of upper-triangular matrices – the polytope $Q^J_{u,v}$ is a Bruhat interval polytope as defined earlier. In the Grass-
CHAPTER 3. BRUHAT INTERVAL POLYTOPES

mammian case – i.e. the case that $G = SL_n$ and $P$ is a maximal parabolic subgroup – the Bruhat interval polytopes for $G/P$ are precisely the positroid polytopes, which were studied recently in [ARW16]. As in the $Fl_n$ case, Bruhat interval polytopes for $G/P$ have an interpretation in terms of the moment map: we show that $Q_{u,v}$ is the moment-map image of the closure of a cell in Rietsch's cell decomposition of $(G/P)_{\geq 0}$. It is also the moment-map image of the projection to $G/P$ of a Richardson variety. We also show that the face of a Bruhat interval polytope for $G/P$ is a Bruhat interval polytope. Along the way, we build on work of Marsh-Rietsch [MR05] to give an interpretation of Rietsch’s cell decomposition of $(G/P)_{\geq 0}$ in terms of the Gelfand-Serganova stratification of $G/P$. In particular, each cell of $(G/P)_{\geq 0}$ is contained in a Gelfand-Serganova stratum. In nice cases (for example $G/B$ and the Grassmannian $Gr_{k,n}$) it follows that Rietsch’s cell decomposition is the restriction of the Gelfand-Serganova stratification to $(G/P)_{\geq 0}$.

The structure of this paper is as follows. In Section 3.2 we provide background and terminology for posets, Coxeter groups, permutohedra, matroid polytopes, and Bruhat interval polytopes. In Section 3.3 we state and prove the Generalized lifting property for the symmetric group. We then use this result in Section 3.4 to prove that the face of a Bruhat interval polytope is a Bruhat interval polytope. Section 3.4 also provides a dimension formula for Bruhat interval polytopes, and an inequality description for Bruhat interval polytopes. In Section 3.5 we give a generalization of the usual recurrence for $R$-polynomials, using the notion of an inversion-minimal transposition on the interval $(u,v)$. The goal of the remainder of the paper is to discuss Bruhat interval polytopes for $G/P$. In Section 3.6 we provide background on generalized partial flag varieties $G/P$, including generalized Plücker coordinates, the Gelfand-Serganova stratification of $G/P$, total positivity, and the moment map. Finally in Section 3.7, we show that each cell in Rietsch’s cell decomposition of $(G/P)_{\geq 0}$ lies in a Gelfand-Serganova stratum, and we use this result to prove that the face of a Bruhat interval polytope for $G/P$ is again a Bruhat interval polytope for $G/P$.

3.2 Background

In this section we will quickly review some notation and background for posets and Coxeter groups. We will also review some basic facts about permutohedra, matroid polytopes, and Bruhat interval polytopes. We will assume knowledge of the basic definitions of Coxeter systems and Bruhat order; we refer the reader to [BB05] for details. Note that throughout this paper, Bruhat order will refer to the strong Bruhat order.

Let $P$ be a poset with order relation $\prec$. We will use the symbol $\preceq$ to denote a covering relation in the poset: $u \preceq v$ means that $u < v$ and there is no $z$ such that $u < z < v$. Additionally, if $u < v$ then $[u,v]$ denotes the (closed) interval from $u$ to $v$; that is, $[u,v] = \{ z \in P \mid u \leq z \leq v \}$. Similarly, $(u,v)$ denotes the (open) interval, that is, $(u,v) = \{ z \in P \mid u < z < v \}$.

The natural geometric object that one associates to a poset $P$ is the geometric realization of its order complex (or nerve). The order complex $\Delta(P)$ is defined to be the simplicial complex whose vertices are the elements of $P$ and whose simplices are the chains $x_0 < x_1 < \cdots < x_k$ in $P$. Abusing notation, we will also use the notation $\Delta(P)$ to denote the geometric realization of the order complex.
Let \((W, S)\) be a Coxeter group generated by a set of simple reflections \(S = \{s_i \mid i \in I\}\). We denote the set of all reflections by \(T = \{wsw^{-1} \mid w \in W\}\). Recall that a reduced word for an element \(w \in W\) is a minimal length expression for \(w\) as a product of elements of \(S\), and the length \(\ell(w)\) of \(w\) is the length of a reduced word. For \(w \in W\), we let \(D_R(w) = \{s \in S \mid ws < w\}\) be the right descent set of \(w\) and \(D_L(w) = \{s \in S \mid sw < w\}\) the left descent set of \(w\). We also let \(T_R(w) = \{t \in T \mid \ell(wt) < \ell(w)\}\) and \(T_L(w) = \{t \in T \mid \ell(tw) < \ell(w)\}\) be the right associated reflections and left associated reflections of \(w\), respectively.

The (strong) Bruhat order on \(W\) is defined by \(u \leq v\) if some substring of some (equivalently, every) reduced word for \(v\) is a reduced word for \(u\). The Bruhat order on a Coxeter group is a graded poset, with rank function given by length.

When \(W\) is the symmetric group \(S_n\), the reflections are the transpositions \(T = \{(ij) \mid 1 \leq i < j \leq n\}\), the set of permutations which act on \(\{1, \ldots, n\}\) by swapping \(i\) and \(j\). The simple reflections are the reflections of the form \((ij)\) where \(j = i + 1\). We also denote this simple reflection by \(s_i\). An inversion of a permutation \(z = (z(1), \ldots, z(n)) \in S_n\) is a pair \((ij)\) with \(1 \leq i < j \leq n\) such that \(z(i) > z(j)\). It is well-known that \(\ell(z)\) is equal to the number of inversions of the permutation \(z\).

Note that we will often use the notation \((z_1, \ldots, z_n)\) instead of \((z(1), \ldots, z(n))\).

We now review some facts about permutohedra, matroid polytopes, and Bruhat interval polytopes.

**Definition 3.2.1.** The usual permutohedron \(\text{Perm}_n\) in \(\mathbb{R}^n\) is the convex hull of the \(n!\) points obtained by permuting the coordinates of the vector \((1, 2, \ldots, n)\).

Bruhat interval polytopes, as defined below, were introduced and studied by Kodama and Williams in [KW15], in connection with the full Kostant-Toda lattice on the flag variety.

**Definition 3.2.2.** Let \(u, v \in S_n\) such that \(u \leq v\) in (strong) Bruhat order. We identify each permutation \(z \in S_n\) with the corresponding vector \((z(1), \ldots, z(n))\) \(\in \mathbb{R}^n\). Then the Bruhat interval polytope \(Q_{u,v}\) is defined as the convex hull of all vectors \((z(1), \ldots, z(n))\) for \(z\) such that \(u \leq z \leq v\).

See Figure 3.1 for some examples of Bruhat interval polytopes.

![Figure 3.1: The two polytopes are the permutohedron \(Q_{e,w_0} = \text{Perm}_4\), and the Bruhat interval polytope \(Q_{u,v}\) with \(v = (2, 4, 3, 1)\) and \(u = (1, 2, 4, 3)\).](image-url)
We next explain how Bruhat interval polytopes are related to matroid polytopes, generalized permutohedra, and flag matroid polytopes.

**Definition 3.2.3.** Let \( \mathcal{M} \) be a nonempty collection of \( k \)-element subsets of \([n]\) such that: if \( I \) and \( J \) are distinct members of \( \mathcal{M} \) and \( i \in I \setminus J \), then there exists an element \( j \in J \setminus I \) such that \((I \setminus \{i\}) \cup \{j\} \in \mathcal{M}\). Then \( \mathcal{M} \) is called the set of bases of a matroid of rank \( k \) on the ground set \([n]\); or simply a matroid.

**Definition 3.2.4.** Given the set of bases \( \mathcal{M} \subset \binom{[n]}{k} \) of a matroid, the matroid polytope \( \Gamma_{\mathcal{M}} \) of \( \mathcal{M} \) is the convex hull of the indicator vectors of the bases of \( \mathcal{M} \):

\[
\Gamma_{\mathcal{M}} := \text{Conv}\{e_I \mid I \in \mathcal{M}\} \subset \mathbb{R}^n,
\]

where \( e_I := \sum_{i \in I} e_i \), and \( \{e_1, \ldots, e_n\} \) is the standard basis of \( \mathbb{R}^n \).

Note that “a matroid polytope” refers to the polytope of a specific matroid in its specific position in \( \mathbb{R}^n \).

**Definition 3.2.5.** The flag variety \( \text{Fl}_n \) is the variety of all flags

\[
\text{Fl}_n = \{V \bullet = V_1 \subset V_2 \subset \cdots \subset V_n = \mathbb{R}^n \mid \dim V_i = i\}
\]

of vector subspaces of \( \mathbb{R}^n \).

**Definition 3.2.6.** The Grassmannian \( Gr_{k,n} \) is the variety of \( k \)-dimensional subspaces of \( \mathbb{R}^n \)

\[
Gr_{k,n} = \{V \subset \mathbb{R}^n \mid \dim V = k\}.
\]

Note that there is a natural projection \( \pi_k : \text{Fl}_n \to Gr_{k,n} \) taking \( V \bullet = V_1 \subset \cdots \subset V_n \) to \( V_k \).

Note also that any element \( V \in Gr_{k,n} \) gives rise to a matroid \( \mathcal{M}(V) \) of rank \( k \) on the ground set \([n]\). First represent \( V \) as the row-span of a full rank \( k \times n \) matrix \( A \). Given a \( k \)-element subset \( I \) of \( \{1, 2, \ldots, n\} \), let \( \Delta_I(A) \) denote the determinant of the \( k \times k \) submatrix of \( A \) located in columns \( I \). This is called a Plücker coordinate. Then \( V \) gives rise to a matroid \( \mathcal{M}(V) \) whose bases are precisely the \( k \)-element subsets \( I \) such that \( \Delta_I(A) \neq 0 \).

One result of [KW15, Section 6] (see also [KW15, Appendix]) is the following. See Section 3.6 for the definition of \( \mathcal{R}_{u,v};>0 \).

**Proposition 3.2.7.** Choose \( u \leq v \in S_n \). Let \( V \bullet = V_1 \subset \cdots \subset V_n \) be any element in the positive part of the Richardson variety \( \mathcal{R}_{u,v};>0 \). Then the Bruhat interval polytope \( Q_{u,v} \) is the Minkowski sum of \( n-1 \) matroid polytopes:

\[
Q_{u,v} = \sum_{k=1}^{n-1} \Gamma_{\mathcal{M}(V_k)}.
\]

In fact each of the polytopes \( \Gamma_{\mathcal{M}(V_k)} \) is a positroid polytope, in the sense of [ARW16], and \( Q_{u,v} \) is a generalized permutohedron, in the sense of Postnikov [Pos09].
We can compute the bases $\mathcal{M}(V_k)$ from the permutations $u$ and $v$ as follows.

\[ \mathcal{M}(V_k) = \{ I \in \binom{[n]}{k} \mid \text{there exists } z \in [u,v] \text{ such that } I = \{ z^{-1}(n), z^{-1}(n-1), \ldots, z^{-1}(n-k+1) \} \}. \]

Therefore we have the following.

**Proposition 3.2.8.** For any $u \leq v \in S_n$, the Bruhat interval polytope $Q_{u,v}$ is the Minkowski sum of $n-1$ matroid polytopes

\[ Q_{u,v} = \sum_{k=1}^{n-1} \Gamma_{\mathcal{M}_k}, \]

where

\[ \mathcal{M}_k = \{ I \in \binom{[n]}{k} \mid \text{there exists } z \in [u,v] \text{ such that } I = \{ z^{-1}(n), z^{-1}(n-1), \ldots, z^{-1}(n-k+1) \} \}. \]

Positroid polytopes are a particularly nice class of matroid polytopes coming from positively oriented matroids. A generalized permutohedron is a polytope which is obtained by moving the vertices of the usual permutohedron in such a way that directions of edges are preserved, but some edges (and higher dimensional faces) may degenerate. See [ARW16] and [Pos09] for more details on positroid polytopes and generalized permutohedra.

There is a generalization of matroid called flag matroid, due to Gelfand and Serganova [GS87], [BGW03, Section 1.7], and a corresponding notion of flag matroid polytope. A convex polytope $\Delta$ in the real vector space $\mathbb{R}^n$ is called a (type $A_{n-1}$) flag matroid polytope if the edges of $\Delta$ are parallel to the roots of type $A_{n-1}$ and there exists a point equidistant from all of its vertices.

The following result follows easily from Proposition 3.2.7.

**Proposition 3.2.9.** Choose $u \leq v \in S_n$. Then the Bruhat interval polytope $Q_{u,v}$ is a flag matroid polytope.

**Proof.** Let $V_\bullet = V_1 \subset \cdots \subset V_n$ be any element in the positive part of the Richardson variety $\mathcal{R}_{u,v,>0}$. By Proposition 3.2.7, $Q_{u,v} = \sum_{k=1}^{n-1} \Gamma_{\mathcal{M}_k}$. Then [BGW03, Theorem 1.7.3] implies that the collection of matroids $\mathcal{M}_\bullet = \{ \mathcal{M}(V_1), \ldots, \mathcal{M}(V_{n-1}) \}$ forms a flag matroid. By [BGW03, Theorem 1.13.5], it follows that the flag matroid polytope associated to $\mathcal{M}_\bullet$ is the Minkowski sum of the matroid polytopes $\Gamma_{\mathcal{M}(V_1)}, \ldots, \Gamma_{\mathcal{M}(V_{n-1})}$. Therefore $Q_{u,v}$ is a flag matroid polytope. $\square$

We can use Proposition 3.2.9 to prove the following useful result.

**Proposition 3.2.10.** Let $Q_{u,v}$ be a Bruhat interval polytope. Consider a face $F$ of $Q_{u,v}$. Let $\mathcal{N}$ be the set of permutations which label vertices of $F$. Then $\mathcal{N}$ contains an element $x$ and an element $y$ such that

\[ x \leq z \leq y \quad \forall z \in \mathcal{N}. \]

**Proof.** By Proposition 3.2.9, $Q_{u,v}$ is a flag matroid polytope. It follows from the definition that every face of a flag matroid polytope is again a flag matroid polytope, and therefore the face $F$ is a flag matroid polytope. By [BGW03, Section 6.1.3], every flag matroid is a
Coxeter matroid, and hence the permutations \( \mathcal{N} \) labeling the vertices of \( F \) are the elements of a Coxeter matroid (for \( S_n \), with parabolic subgroup the trivial group). But now by the Maximality Property for Coxeter matroids [BGW03, Section 6.1.1], \( \mathcal{N} \) must contain a minimal element \( x \) such that \( x \leq z \) for all \( z \in \mathcal{N} \), and \( \mathcal{N} \) must contain a maximal element \( y \) such that \( y \geq z \) for all \( z \in \mathcal{N} \). \( \Box \)

### 3.3 The generalized lifting property for the symmetric group

The main result of this section is Theorem 3.3.3, which is a generalization (for the symmetric group) of the classical lifting property for Coxeter groups. This result will be a main tool for proving that every face of a Bruhat interval polytope is a Bruhat interval polytope.

We start by recalling the usual lifting property.

**Proposition 3.3.1 (Lifting property).** Suppose \( u < v \) and \( s \in D_R(v) \setminus D_R(u) \). Then \( u \leq vs < v \) and \( u < us \leq v \).

**Definition 3.3.2.** Let \( u, v \in S_n \). A transposition \((ik)\) is inversion-minimal on \((u, v)\) if the interval \([i, k]\) is the minimal interval (with respect to inclusion) which has the property

\[
v_i > v_k, \quad u_i < u_k.
\]

**Theorem 3.3.3 (Generalized lifting property).** Suppose \( u < v \) in \( S_n \). Choose a transposition \((ij)\) which is inversion-minimal on \((u, v)\). Then \( u \leq v(ij) < v \) and \( u < u(ij) \leq v \).

We note that there are pairs \( u < v \) where \( D_R(v) \setminus D_R(u) \) is empty, and hence one cannot apply the Lifting property. In contrast, Lemma 3.3.4 below shows that for any pair \( u < v \) in \( S_n \), there exists an inversion-minimal transposition \((ij)\). Hence it is always possible to apply the Generalized lifting property.

**Lemma 3.3.4.** Let \((W,S)\) be a Coxeter group. Take \( u, v \in W \) distinct. If \( \ell(v) \geq \ell(u) \) then there exists a reflection \( t \in T \) such that

\[
v > vt, \quad u < ut.
\]

**Proof.** Recall that \( T_R(w) = \{ t \in T \mid wt < w \} \). The lemma will follow if we show that \( T_R(v) \not\subseteq T_R(u) \). Assume by contradiction that \( T_R(v) \subseteq T_R(u) \). By [BB05, Corollary 1.4.5], for any \( x \in W \), \( |T_R(x)| = \ell(x) \). Since \( \ell(v) \leq \ell(u) \), we must have \( T_R(v) = T_R(u) \). By [BB05, Chapter 1 Exercise 11], this contradicts \( v \neq u \). \( \Box \)

Lemma 3.3.4 directly implies the following corollary.

**Corollary 3.3.5.** Let \( v, u \in S_n \) be two distinct permutations. If \( \ell(v) \geq \ell(u) \) then there exists an inversion-minimal transposition on \((u, v)\).

In preparation for the proof of Theorem 3.3.3, it will be convenient to make the following definition.
Definition 3.3.6. A pattern of length \( n \) is an equivalence class of sequences \( x_1x_2\ldots x_n \) of distinct integers. Two such sequences \( x_1x_2\ldots x_n, y_1y_2\ldots y_n \) are in the same equivalence class ("have the same pattern") if

\[
x_i > x_j \iff y_i > y_j \quad \text{for all } i, j \text{ such that } 1 \leq i, j \leq n.
\]

Denote by \( \text{Patt}_n \) the set of patterns of length \( n \).

There is a canonical representative for each pattern \( x \in \text{Patt}_n \) obtained by replacing each \( x_i \) with

\[
\bar{x}_i := \# \{ j \in [n] : x_j \leq x_i \}.
\]

For example, the canonical representative of 523 is 312.

Definition 3.3.7. Let \( x, y \in \text{Patt}_n \) for some \( n \). Call \((x, y)\) an Inversion-Inversion pair if the following condition holds:

\[
\text{for all } i < j, \quad x_i > x_j \implies y_i > y_j.
\]

Notice that this statement is independent of the choice of representatives.

It is easy to see that if \((x, y)\) is an Inversion-Inversion pair, then so is \((x_1\ldots \hat{x}_k \ldots x_n, y_1\ldots \hat{y}_k \ldots y_n)\) for any \( k \).

In preparation for the proof of Theorem 3.3.3, we first state and prove Lemmas 3.3.8, 3.3.10, and 3.3.11.

Lemma 3.3.8. Let \( u, v \in S_n \). The following are equivalent:

(i). The transposition \((ik)\) is inversion-minimal on \((u, v)\)

(ii). The patterns \( x = x_i \ldots x_k := v_i \ldots v_k \) and \( y = y_i \ldots y_k := u_ku_{i+1}u_i \ldots u_{k-2}u_{k-1}u_i \) form an Inversion-Inversion pair \((x, y)\) with \( \bar{x}_k = \bar{x}_i + 1 \) and \( \bar{y}_k = \bar{y}_i + 1 \).

Proof. First note that (ii) obviously implies (i). We now prove that (i) implies (ii). Assume transposition \((ik)\) is inversion-minimal. We show that the following two cases cannot hold:

Case 1: there is some \( i < j < k \) with \( v_j \in [v_k, v_i] \).

Looking at intervals \([i, j]\) and \([j, k]\), we have \( v_i > v_j \) and \( v_j > v_k \). By minimality of \([i, k]\), this implies that \( u_i > u_j \) and \( u_j > u_k \), contradicting \( u_i < u_k \).

Case 2: there is some \( i < j < k \) with \( u_j \in [u_i, u_k] \).

Looking at intervals \([i, j]\) and \([j, k]\) again, we have \( u_i < u_j \) and \( u_j < u_k \). By minimality of \([i, k]\), this implies that \( v_i < v_j \) and \( v_j < v_k \), contradicting \( v_i > v_k \).

(ii) \implies (i). Since \( \bar{x}_k > \bar{x}_1 \) and \( \bar{y}_k > \bar{y}_1 \), we see that \( v_i > v_k \) and \( u_i < u_k \). Assume by contradiction that \([p, q]\) is a strict subset of \([i, k]\) such that \( v_p > v_q \) and \( u_p < u_q \). Since \( \bar{x}_k = \bar{x}_i + 1 \) and \( \bar{y}_k = \bar{y}_i + 1 \), for any \( j \in (i, k) \),

\[
x_j > x_k \iff x_j > x_i
\]

\[
y_j > y_k \iff y_j > y_i.
\]
Equivalently, for any \( j \in (i, k) \),
\[
 v_j > v_k \iff v_j > v_i \\
 u_j > u_k \iff u_j > u_i.
\]
If \( \{p, q\} \cap \{i, k\} = \emptyset \), then we clearly obtain a contradiction. If \( p = i \), then
\[
 v_i > v_q, u_i < u_q \implies x_i > x_q, y_k < y_q \implies x_i > x_q, y_i < y_q,
\]
which is a contradiction. A similar argument shows that \( q = k \) leads to a contradiction. □

Lemma 3.3.8 implies the following result.

**Corollary 3.3.9.** Let \( u, v \in S_n \) and let \((ik)\) be inversion-minimal on \((u, v)\). Then
\[
 v(ik) \preceq v \quad \text{and} \quad u \preceq u(ik).
\]

![Generalized lifting property](image)

**Figure 3.2:** Generalized lifting property

**Lemma 3.3.10.** Let \( x, y \in \text{Patt}_n \) with \( \overline{x}_n = \overline{x}_1 + 1 \) and \( \overline{y}_n = \overline{y}_1 + 1 \). If \((x, y)\) is an Inversion-Inversion pair, then \( \overline{x}_1 = \overline{y}_1 \).

**Proof.** Define
\[
 I_{i,j}(x) := \begin{cases} 1 & \text{if } x_i > x_j, \\ 0 & \text{if } x_i < x_j. \end{cases}
\]
This function is well-defined on patterns. Let
\[
 f(x, y) := \sum_{1 \leq i < j \leq n} I_{i,j}(x)(1 - I_{i,j}(y)).
\]
With this notation, \((x, y)\) is an Inversion-Inversion pair if and only if \( f(x, y) = 0 \).

Note that
\[
 f(x, y) = \ell(x) - \sum_{1 \leq i < j \leq n} I_{i,j}(x)I_{i,j}(y).
\]

The pairs
\[
 (a) : (x_1 \cdots x_{n-1}, y_1 \cdots y_{n-1})
\]
and
\[
 (b) : (x_2 \cdots x_n, y_2 \cdots y_n)
\]
are Inversion-Inversion pairs. The conditions on \(x_1, x_n\) imply that
\[
I_{1,j}(x) = 1 - I_{j,n}(x), \quad \forall 1 < j < n \tag{3.2}
\]
and similarly for \(y\). Since \(f(x_1 \cdots x_{n-1}, y_1 \cdots y_{n-1}) = 0\), and using \(I_{1,n}(x) = 0\),
\[
(a): \ell(x) - \sum_{1<j<n} I_{1,j}(x) = \sum_{1<i<j<n} I_{i,j}(x)I_{i,j}(y) \tag{3.3}
\]
Applying condition (3.2) to (3.3), and simplifying, we get
\[
(a): \ell(x) - (n - 2) + \sum_{1<j<n} I_{1,j}(x) = \sum_{1<i<j<n} I_{i,j}(x)I_{i,j}(y) \tag{3.4}
\]
Similarly, since \(f(x_2 \cdots x_n, y_2 \cdots y_n) = 0\) and \(I_{1,n}(x) = 0\),
\[
(b): \ell(x) - \sum_{1<j<n} I_{1,j}(x) = \sum_{1<i<j\leq n} I_{i,j}(x)I_{i,j}(y) \tag{3.5}
\]
Using condition (3.2) with \(x\) replaced with \(y\), equation (3.5) reduces to
\[
\ell(x) - \sum_{1<j<n} I_{1,j}(x) = \sum_{1<i<j<n} I_{i,j}(x)I_{i,j}(y) + \sum_{1<i<j<n} I_{i,j}(x)I_{i,j}(y)
\]
\[
= \sum_{1<j<n} (1 - I_{1,j}(x))(1 - I_{1,j}(y)) + \sum_{1<i<j<n} I_{i,j}(x)I_{i,j}(y)
\]
\[
= (n - 2) - \sum_{1<j<n} (I_{1,j}(x) + I_{1,j}(y)) + \sum_{1<j<n} I_{1,j}(x)I_{1,j}(y) + \sum_{1<i<j<n} I_{i,j}(x)I_{i,j}(y). \tag{3.6}
\]
Comparing (3.4) and (3.6) we see that
\[
\sum_{1<j<n} I_{1,j}(x) = \sum_{1<j<n} I_{1,j}(y)
\]
which can only happen if \(x_1 = y_1\). \(\Box\)

**Lemma 3.3.11.** Suppose that \((ik)\) is inversion-minimal on \((u, v)\). Then for every \(i < j < k\), we have
\[
u_j > u_i \iff u_j > u_k \iff v_j > v_k \iff v_j > v_i.
\]

**Proof.** By Lemma 3.3.8, the patterns \(x = v_i \cdots v_k\) and \(y = u_ku_{i+1} \cdots u_{k-1}u_i\) form an
Inversion-Inversion pair \((x, y)\) with \(\bar{x}_k = \bar{x}_1 + 1\) and \(\bar{y}_k = \bar{y}_1 + 1\). By Lemma 3.3.10, \(\bar{x}_1 = \bar{y}_1\).
It follows that
\[
\#\{j : i < j < k, v_j > v_k\} = \#\{j : i < j < k, u_j > u_i\}.
\]
We also see that \(\#\{j : i < j < k, v_j > v_k\} = \#\{j : i < j < k, v_j > v_i\}\) and \(\#\{j : i < j < k, u_j > u_i\} = \#\{j : i < j < k, u_j > u_k\}\). By minimality of \([i, k]\), for every \(i < j < k\),
\[
v_j > v_k \implies u_j > u_k.
\]
Consequently, for every $i < j < k$,

$$u_j > u_i \iff u_j > u_k \iff v_j > v_k \iff v_j > v_i. \quad (3.7)$$

Finally we are ready to prove Theorem 3.3.3.

Proof. [Proof of Theorem 3.3.3]. Choose $u < v$ in $S_n$, and a transposition $(ij)$ which is inversion-minimal on $(u, v)$. By Corollary 3.3.9, to prove Theorem 3.3.3, it suffices to show that $u \leq v(ij)$ and $u(ij) \leq v$.

We use induction on $k = j - i$. The base case $k = 1$ holds by the lifting property (Proposition 3.3.1).

Now consider $k > 1$. Since $(ij)$ is inversion-minimal on $(u, v)$, we have $v_i > v_j$ and $u_i < u_j$.

**Case 1:** Suppose that (a) $v_i > v_{i+1}$ and $u_i > u_{i+1}$, or (b) $v_i < v_{i+1}$ and $u_i < u_{i+1}$.

We have (a) $u > us_i$ and $v > vs_i$ or (b) $u < us_i$ and $v < vs_i$. Clearly $((i+1)j)$ is inversion-minimal on $(us_i, us_i)$, and since $u < v$, we have $us_i < vs_i$. By induction,

$$us_i \leq vs_i((i+1)j) \quad \text{and} \quad us_i((i+1)j) \leq vs_i.$$ 

Notice that $s_i((i+1)j)s_i = (ij) = t$.

In case (a), we claim that $s_i \notin D_R(us_i) \cup D_R(vs_i((i+1)j))$. To see this, note first that $v_{i+1} < v_j$; otherwise we’d have $v_{i+1} > v_j$ and also $u_{i+1} < u_j$, which would contradict our assumption that the interval $[i, j]$ is inversion-minimal on $(u, v)$. Therefore $s_i \notin D_R(vs_i((i+1)j))$, and the claim follows. But now the claim together with $us_i \leq vs_i((i+1)j)$ implies that $us_i^2 \leq vs_i((i+1)j)s_i$ and hence $u \leq vt$.

In case (b), we claim that $s_i \in D_R(us_i((i+1)j)) \cap D_R(vs_i)$. To see this, note first that $u_{i+1} > u_j$; otherwise we’d have $u_{i+1} < u_j$ and also $v_{i+1} > v_j$, which would contradict our assumption that transposition $(ij)$ is inversion-minimal on $(u, v)$. Therefore $s_i \in D_R(us_i((i+1)j))$, and the claim follows. But now the claim together with $us_i((i+1)j) \leq vs_i$ implies that $us_i((i+1)j)s_i \leq vs_i^2$, and hence $ut \leq v$.

**Case 2:** Suppose that $v_{j-1} > v_j$ and $u_{j-1} > u_j$, or $v_{j-1} < v_j$ and $u_{j-1} < u_j$.

This case is analogous to Case 1.

**Case 3:** Suppose that neither of the above two cases holds.

Since $(ij)$ is inversion-minimal on $(u, v)$, we must have $v_i < v_{i+1}$ and $v_{j-1} < v_j$. Since $v_i > v_j$, there exists some $m_1 \in (i, j-1)$ such that $v_{m_1} > v_{m_1+1}$. By minimality, $u_{m_1} > u_{m_1+1}$.

By Lemma 3.5.2, $(ij)$ is inversion-minimal on $(us_{m_1}, us_{m_1})$. If $us_{m_1}$ and $vs_{m_1}$ do not satisfy the conditions of Cases 1 or 2, then we may find $m_2 \in (i, j-1)$ and then $(ij)$ is inversion-minimal on $(us_{m_2}, us_{m_2}, vs_{m_2}, us_{m_2})$. Such a sequence $m_1, m_2, \ldots$ clearly terminates. Assume that it terminates at $k$, so that $(ij)$ is inversion-minimal on $(us_{m_1}s_{m_2} \cdots s_{m_k}, us_{m_1}s_{m_2} \cdots s_{m_k})$ and the hypotheses of Case 1 or 2 are satisfied for $us_{m_1}s_{m_2} \cdots s_{m_k}$ and $us_{m_1}s_{m_2} \cdots s_{m_k}$. Set $\Pi_k := s_{m_1}s_{m_2} \cdots s_{m_k}$. We then have

$$u\Pi_k \leq u\Pi kt, \quad v\Pi kt \leq v\Pi k, \quad u\Pi kt \leq v\Pi k, \quad u\Pi k \leq v\Pi kt.$$ 

We show now that for $1 \leq p \leq k$, if

$$u\Pi p \leq u\Pi pt, \quad v\Pi pt \leq v\Pi p, \quad u\Pi pt \leq v\Pi p, \quad u\Pi p \leq v\Pi pt$$ 

then

$$u\Pi p \leq u\Pi pt, \quad v\Pi pt \leq v\Pi p, \quad u\Pi pt \leq v\Pi p, \quad u\Pi p \leq v\Pi pt$$ 

for all $1 \leq p \leq k$.
then

\[ u_{p-1} < u_{p-1}t, \quad v_{p-1}t < v_{p-1}, \quad u_{p-1}t \leq v_{p-1}, \quad u_{p-1} \leq v_{p-1}t. \]

Note that for any \( m \), \( ts_m = s_m t \). Therefore \( u_{p}t = u_{p-1}ts_m \) and \( v_{p}t = v_{p-1}ts_m \).

This implies that \( u_{p}t = u_{p-1}ts_m \geq u_{p-1}t \) and \( v_{p}t = v_{p-1}ts_m \leq v_{p-1}t \). \( \square \)

\[ v = 3241 \]
\[ t = (24) \]
\[ 3142 \]
\[ 2341 \]
\[ (14) \]
\[ t = (24) \]
\[ u = 2143 \]

Figure 3.3: Example of Theorem 3.3.3

**Example 3.3.12.** The following example shows that the converse to Theorem 3.3.3 does not hold: it is not necessarily the case that if the Bruhat relations

\[ v(ik) \preceq v \quad u \prec u(ik) \quad u \preceq v(ik) \quad u(ik) \preceq v \]

hold, then \( (ik) \) is inversion-minimal on \((u,v)\). Take \( v = 4312 \), \( u = 1243 \) and \( (ik) = (24) \).

Then

\[ v(ik) \preceq v \quad u \prec u(ik) \quad u \preceq v(ik) \quad u(ik) \preceq v \]

but also \( v_2 > v_3 \) and \( u_2 < u_3 \).

As a corollary of Generalized lifting, we have the following result, which says that in an interval of the symmetric group we may find a maximal chain such that each transposition connecting two consecutive elements of the chain is a transposition that comes from the atoms, and similarly, for the coatoms.

**Corollary 3.3.13.** Let \([u,v] = \subseteq S_n\) and let \( T(v) := \{ t \in T : v \succ vt \geq u \}\) and \( T(u) := \{ t \in T : u \prec ut \leq v \}\). There exist maximal chains \( C_v : u = x(0) \prec x(1) \prec x(2) \prec \ldots \prec x(l) = v \) and \( C_u : u = y(0) \prec y(1) \prec y(2) \prec \ldots \prec y(l) = v \) in \( I \) such that \( x(i)_{(i)}^{-1}x(i+1) \in T(v) \) and \( y(i)_{(i)}^{-1}y(i+1) \in T(u) \) for each \( i \).

**Proof.** By the Generalized lifting property, there exists a transposition \( t = (ij) \) such that \( u \preceq vt \preceq v \) and \( u \preceq ut \leq v \). But now since \( u \preceq ut \leq v \), we can apply the Generalized lifting property to the pair \( ut \leq v \), and inductively construct the maximal chain \( C_u \). The construction of \( C_u \) is analogous. \( \square \)

We plan to study the Generalized lifting property for other Coxeter groups in a separate paper.
3.4 Results on Bruhat interval polytopes

In this section we give some results on Bruhat interval polytopes. We show that the face of a Bruhat interval polytope is a Bruhat interval polytope; we give a dimension formula; we give an inequality description; and we give a criterion for when one Bruhat interval polytope is a face of another.

Faces of Bruhat interval polytopes are Bruhat interval polytopes

The main result of this section is the following.

Theorem 3.4.1. Every face of a Bruhat interval polytope is itself a Bruhat interval polytope.

Our proof of this result uses the following theorem. It was first proved for the symmetric group by Edelman [Ede81], then generalized to classical types by Proctor [Pro82], and then proved for arbitrary Coxeter groups by Björner and Wachs [BW82].

Theorem 3.4.2. [BW82] Let \((W, S)\) be a Coxeter group. Then for any \(u \leq v\) in \(W\), the order complex \(\Delta(u, v)\) of the interval \((u, v)\) is PL-homeomorphic to a sphere \(S^{\ell(u,v)-2}\). In particular, the Bruhat order is thin, that is, every rank 2 interval is a diamond. In other words, whenever \(u \leq v\) with \(\ell(v) - \ell(u) = 2\), there are precisely two elements \(z(1), z(2)\) such that \(u < z(i) < v\).

We will identify a linear functional \(\omega\) with a vector \((\omega_1, \ldots, \omega_n) \in \mathbb{R}^n\), where \(\omega : \mathbb{R}^n \to \mathbb{R}\) is defined by \(\omega(e_i) = \omega_i\) (and extended linearly).

Proposition 3.4.3. Choose \(u \leq v\) in \(S_n\), and let \(\omega : \mathbb{R}^n \to \mathbb{R}\) be a linear functional which is constant on a maximal chain \(C\) from \(u\) to \(v\). Then \(\omega\) is constant on all permutations \(z\) where \(u \leq z \leq v\).

Proof. We will use the topology of \(\Delta(u, v)\) to prove that \(\omega\) is constant on any maximal chain from \(u\) to \(v\). If \(\ell(v) - \ell(u) = 1\), there is nothing to prove. If \(\ell(v) - \ell(u) = 2\), then the interval \([u, v]\) is a diamond. By the Generalized lifting property (Theorem 3.3.3), there exists a transposition \(t = (ij)\) such that \(u < vt < v\) and \(u < ut < v\). Without loss of generality, \(C\) is the chain \(u < vt < v\). But then since \(\omega(vt) = \omega(v)\), we must have \(\omega_i = \omega_j\). It follows that \(\omega(ut) = \omega(u)\), and hence \(\omega\) is constant on both maximal chains from \(u\) to \(v\).

If \(\ell(v) - \ell(u) \geq 3\), then the order complex \(\Delta(u, v)\) is a PL sphere of dimension at least 1, and hence it is connected in codimension one. Therefore we can find a path of maximal chains \(C = C_0, C_1, \ldots, C_N\) in \((u, v)\) starting with \(C\), which contains all maximal chains of \((u, v)\) (possibly some occur more than once), and which has the property that for each adjacent pair \(C_i\) and \(C_{i+1}\), the two chains differ in precisely one element. Since the Bruhat order is thin, \(C_i\) must contain three consecutive elements \(a < z(1) < b\), and \(C_{i+1}\) is obtained from \(C_i\) by replacing \(z(1)\) by \(z(2)\), the unique element other than \(z(1)\) in the interval \((a, b)\). Suppose by induction that \(\omega\) is constant on \(C_0, C_1, \ldots, C_i\). Since \(\omega(a) = \omega(z(1)) = \omega(b)\) and \(\ell(b) - \ell(a) = 2\), we have observed in the previous paragraph that \(\omega\) must be constant on \([a, b]\). Therefore \(\omega\) attains the same value on \(z(2)\) and hence on all of \(C_{i+1}\). \(\Box\)
Corollary 3.4.4. If a linear functional $\omega : \mathbb{R}^n \to \mathbb{R}$, when restricted to $[u,v]$, attains its maximum value on $u$ and $v$, then it is constant on $[u,v]$.

Proof. By Proposition 3.4.3, it suffices to show that there is a maximal chain $C_0 = \{u = z(0) \leq z(1) \leq \cdots \leq z(t) = v\}$ on which $\omega$ is constant. By the Generalized lifting property, there exists a transposition $t = (ij)$ such that $u \leq vt \leq v$ and $u \leq ut \leq v$. Since $u \leq ut$ and $vt \leq v$, we have $u_i < u_j$ and $v_i > v_j$. Since $\omega(ut) \leq \omega(u)$, it follows that $\omega_i \leq \omega_j$. Similarly, $\omega(vt) \leq \omega(v)$ implies that $\omega_j \geq \omega_i$. Therefore $\omega_i = \omega_j$, and hence $\omega(ut) = w(u) = w(v)$. But now since $u \leq ut \leq v$, with $\omega(ut) = \omega(v)$, we can apply the Generalized lifting property to the pair $ut \leq v$, and inductively construct the desired maximal chain. \[\square\]

The dimension of Bruhat interval polytopes

In this section we will give a dimension formula for Bruhat interval polytopes. We will then use it to determine which Richardson varieties in $\mathbb{F}_n$ are toric varieties, with respect to the usual torus action on $\mathbb{F}_n$. Recall that a Richardson variety $R_{u,v}$ is the intersection of opposite Schubert (sometimes called Bruhat) cells; see Section 3.6 for background on Richardson varieties.

Definition 3.4.5. Let $u \leq v$ be permutations in $S_n$, and let $C : u = x_{(0)} \leq x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(t)} = v$ be any maximal chain from $u$ to $v$. Define a labeled graph $G_C$ on $[n]$ having an edge between vertices $a$ and $b$ if and only if the transposition $(ab)$ equals $x_i^{-1}x_{(i+1)}$ for some $0 \leq i \leq l - 1$. Define $B_C = \{B^1, B^2, \ldots, B^r\}$ to be the partition of $[n] = \{1, 2, \ldots, n\}$ whose blocks $B^j$ are the connected components of $G_C$. Let $\#B_C$ denote $r$, the number of blocks in the partition.

We will show in Corollary 3.4.8 that the partition $B_C$ is independent of $C$; and so we will denote this partition by $B_{u,v}$.

Theorem 3.4.6. The dimension $\dim Q_{u,v}$ of the Bruhat interval polytope $Q_{u,v}$ is

$$\dim Q_{u,v} = n - \#B_{u,v}.$$ 

The equations defining the affine span of $Q_{u,v}$ are

$$\sum_{i \in B^j} x_i = \sum_{i \in B^j} u_i (= \sum_{i \in B^j} v_i), \quad j = 1, 2, \ldots, \#B_{u,v}. \quad (3.8)$$
Before proving Theorem 3.4.6, we need to show that $B_{u,v}$ is well-defined. Given a subset $A \subseteq [n]$, let $e_A$ denote the $0 - 1$ vector in $\mathbb{R}^n$ with a 1 in position $a$ if and only if $a \in A$.

**Lemma 3.4.7.** Let $C$ be a maximal chain in $[u, v] \subset S_n$. Let $B_C = \{B^1, \ldots, B^r\}$ be the associated partition of $[n]$. Then a linear functional $\omega : \mathbb{R}^n \to \mathbb{R}$ is constant on the interval $[u, v]$ if and only if

$$\omega = \sum_{j=1}^r c_j e_{B^j}$$

for some coefficients $c_j$.

**Proof.** Using the definition of the partition $B_C$, it is immediate that $\omega$ is constant on the chain $C$ if and only if it has the form $\sum_{j=1}^r c_j e_{B^j}$. The lemma now follows from Proposition 3.4.3. $\square$

**Corollary 3.4.8.** The partition $B_C$ is independent of the choice of $C$.

**Proof.** Let $B_C = \{B^1_C, B^2_C, \ldots, B^r_C\}$. Take $\omega = e_{B^j_C}$. By Lemma 3.4.7, $\omega$ is constant on $[u, v]$ and therefore on any other chain $C'$. Consequently, there exist some elements $B^j_{C'}, \ldots, B^{j_k}_{C'}$ of $B_{C'}$ such that

$$B^j_C = B^j_{C'} \sqcup \ldots \sqcup B^{j_k}_{C'}$$

It follows that $B'_{C'}$ is a refinement of $B_C$. Similarly, $B_C$ is a refinement of $B'_{C'}$. $\square$

**Definition 3.4.9.** Let $u \leq v$ be permutations in $S_n$, and let $T(u) := \{ t \in T : u < ut \leq v \}$ and $T(v) := \{ t \in T : v > vt \geq u \}$ be the transpositions labeling the cover relations corresponding to the atoms and coatoms in the interval. Define a labeled graph $G^{\text{at}}$ (resp. $G^{\text{coat}}$) on $[n]$ such that $G^{\text{at}}$ (resp. $G^{\text{coat}}$) has an edge between $a$ and $b$ if and only if the transposition $(ab) \in T(u)$ (resp. $(ab) \in T(v)$). Let $B^{\text{at}}_{u,v}$ be the partition of $[n]$ whose blocks are the connected components of $G^{\text{at}}$. Similarly, define partition $B^{\text{coat}}_{u,v}$ whose blocks are the connected components of $G^{\text{coat}}$.

**Proposition 3.4.10.** Let $[u, v] \subset S_n$. The partitions $B^{\text{at}}_{u,v}$ and $B^{\text{coat}}_{u,v}$ are equal to $B_{u,v}$. Consequently, the labeled graphs $G^C, G^{\text{at}}$ and $G^{\text{coat}}$ all have the same connected components.

**Proof.** The result follows from Corollary 3.3.13 and Corollary 3.4.8. $\square$

We now prove Theorem 3.4.6.

**Proof.** [Proof of Theorem 3.4.6] We begin by showing that any point $(x_1, x_2, \ldots, x_n) \in Q_{u,v}$ satisfies the independent equations (3.8). By Lemma 3.4.7, the linear functional $\omega = e_{B^j}$ is constant on $[u, v]$. Since $e_{B^j}(x_1, x_2, \ldots, x_n) = \sum_{i \in B^j} x_i$, (3.8) holds.

Now suppose that there exists another affine space

$$\sum_{i=1}^n a_i x_i = c$$

(3.9)

to which $Q_{u,v}$ belongs. By assumption, the linear functional $a = (a_1, \ldots, a_n)$ is constant on $Q_{u,v}$, so by Lemma 3.4.7, $a = \sum_j c_j e_{B^j}$.
for some coefficients $c_j$. Therefore equation (3.9) is a linear combination of equations (3.8).

\[ \square \]

**Example 3.4.11.** Consider the intervals $[1234, 1432]$ and $[1234, 3412]$ in Figures 3.4 and 3.5. We see that $B_{1234,1432} = |1|234|$ and $B_{1234,3412} = |1234|$, so that the dimensions are 2 and 3, respectively.

![Figure 3.4: Bruhat interval [1234, 1432].](image)

![Figure 3.5: Bruhat interval [1234, 3412].](image)

We now turn to the question of when the Richardson variety $R_{u,v}$ is a toric variety. Our proof uses Proposition 3.7.12, which will be proved later, using properties of the moment map.

**Proposition 3.4.12.** The Richardson variety $R_{u,v}$ in $F_l_n$ is a toric variety if and only if the number of blocks $\#B_{u,v}$ of the partition $B_{u,v}$ satisfies $\#B_{u,v} = n - \ell(v) + \ell(u)$. Equivalently, $R_{u,v}$ is a toric variety if and only if the labeled graph $G^C$ is a forest (with no multiple edges).
Proof. By Proposition 3.7.12, $R_{u,v}$ is a toric variety if and only if $\dim Q_{u,v} = \ell(v) - \ell(u)$. The first statement of the proposition now follows from Theorem 3.4.6.

We will prove the second statement from the first. Note that $C$ is a chain with $\ell(v) - \ell(u)$ edges. Let us consider the process of building the graph $G$ by adding one edge at a time while reading the edge-labels of $C$, say from top to bottom. We start out with a totally disconnected graph on the vertices $[n]$. Adding a new edge will either preserve the number of connected components of the graph, or will decrease it by 1. In order to arrive at a partition with every new edge added. But this will happen if and only if the graph $G$ we construct is a forest (with no multiple edges). \endproof

Given a labeled graph $G$, we will say that a cycle $(v_0, v_1, \ldots, v_k)$ with $v_k = v_0$ is increasing if $v_0 < v_1 < \ldots < v_{k-1}$. We shall call a labeled graph with no increasing cycles an increasing-cycle-free labeled graph.

**Lemma 3.4.13.** The labeled graphs $G^{at}$ and $G^{coat}$ are increasing-cycle-free. In particular, they are simple and triangle-free.

**Proof.** From the definition, it is clear that the graphs are simple. Assume by contradiction that $C = (v_0, v_1, \ldots, v_k)$ is an increasing cycle in $G^{at}$. By properties of Bruhat order on the symmetric group, the existence of an edge $\{a, b\}$ with $a < b$ implies that $u(a) < u(b)$ and for any $a < c < b$, $u(c) \not\in [u(a), u(b)]$. Looking at edges $\{v_i, v_{i+1}\}, i = 0, 1, \ldots, k-2$, of cycle $C$, we see that

$$u(v_0) < u(v_1) < \ldots < u(v_{k-1}).$$

However, edge $\{v_0, v_{k-1}\}$ implies that $u(v_i) \not\in [u(v_0), u(v_{k-1})]$ for any $1 \leq i \leq k-2$, which is a contradiction. The proof for $G^{coat}$ is analogous. \endproof

Following Björner and Brenti [BB05], we call the face poset of a $k$-gon a $k$-crown. Any length 3 interval in a Coxeter group is a $k$-crown [BB05, Corollary 2.7.8]. It is also known that in $S_n$, the values of $k$ can only be 2, 3 or 4.

**Remark 3.4.14.** Using Proposition 3.4.10 and Lemma 3.4.13, it is easy to show that any $k$-crown must have $k \leq 4$. Indeed, the graph $G^C$ has 3 edges, and therefore at least $n - 3$ connected components. By Proposition 3.4.10, the graph $G^{at}$ has the same connected components as $G^C$ and $k$ edges. By Lemma 3.4.13 it is simple and triangle-free. Consequently, if $k > 4$ then $G^{at}$ must have at most $n - 4$ components.

**Lemma 3.4.15.** Let $[u, v]$ be a 4-crown and let $C : u = x(0) \leq x(1) \leq x(2) \leq x(3) = v$ be any maximal chain. The graph $G^C$ is a forest. In particular, if we set $t_i := x_{(i)}^{1} x_{(i)}^{+1}$ for $0 \leq i \leq 2$, then $t_0 \neq t_2$ since there are no multiple edges.

**Proof.** The graph $G^{at}$ has 4 edges. By the discussion above, the smallest cycle $G^{at}$ can have is of length 4. Therefore $G^{at}$ has at most $n - 3$ connected components.

Assume by contradiction that $G^C$ is not a forest. Then the graph $G^C$, which has 3 edges, has at least $n - 2$ connected components. But the number of connected components of $G^C$ and $G^{at}$ must be equal, so we obtain a contradiction. \endproof

**Corollary 3.4.16.** A Richardson variety $R_{u,v}$ in $\Fl_n$ with $\ell(v) - \ell(u) = 3$ is a toric variety if and only if $[u, v]$ is a 3-crown or a 4-crown.
Proof. The interval \([u, v]\) is a \(k\)-crown for \(k = 2, 3\) or \(4\). If \(k = 2\) or \(3\), then, by Lemma 3.4.13 the graph \(G^a\) must have \(n - k\) connected components. Consequently, \(k\) cannot equal to \(2\). For \(k = 3\), we see that the Richardson variety is toric. For \(k = 4\), the result follows from Lemma 3.4.15. □

**Faces of Bruhat interval polytopes**

Using the results of prior sections, we will give a combinatorial criterion for when one Bruhat interval polytope is a face of another (see Theorem 3.4.19). First we need a few lemmas.

Let \(T(x, X) := \{t \in T : \exists z \triangleright x, z \in X\}\) and \(\overline{T}(x, X) := \{t \in T : \exists z \triangleleft x, z \in X\}\) be the transpositions labeling cover relations of an element \(x\) in a set \(X\). In the following we use the convention that \(i < j\) in \((i, j)\).

**Lemma 3.4.17.** Let \(\omega : \mathbb{R}^n \to \mathbb{R}\) be a linear functional satisfying
\[
\omega(b) = \omega(a) \geq \omega(d)
\]
for \(a, b, c, d \in S_n\) the elements of a Bruhat interval of length 2:

\[
\begin{diagram}
  & d & \\
  c & \downarrow & b \\
  & a
\end{diagram}
\]

Then either \(\omega(c) \leq \omega(a)\) or \(\omega(c) \leq \omega(d)\).

Proof. By Theorem 3.3.3, there exists \((i, j)\) such that either
\[
a^{(i, j)} \triangleleft b\) and \(c^{(i, j)} \triangleleft d
\]
or
\[
a^{(i, j)} \triangleleft c\) and \(b^{(i, j)} \triangleleft d.
\]
In the former case, \(\omega(a) = \omega(b) \iff \omega(c) = \omega(d)\). In the latter case, \(\omega(a) \geq \omega(c) \iff \omega(b) \geq \omega(d)\). □

**Lemma 3.4.18.** Let \(x, y, u, v \in S_n\) such that \(u \leq x \leq y \leq v\). Assume that \(\omega : \mathbb{R}^n \to \mathbb{R}\) is a linear functional satisfying
\[
w := \omega(z) \text{ for all } z \in [x, y], \quad (3.10)
\]
\[
\omega(y) \geq \omega(b) \text{ for all } b \triangleright y, b \in [u, v], \quad (3.11)
\]
and
\[
\omega(x) \geq \omega(a) \text{ for all } a \triangleleft x, a \in [u, v]. \quad (3.12)
\]
Then for any \(z \in [x, v] \cup [u, y]\), \(\omega(z) \leq w\).
Proof. By Corollary 3.3.13, the result holds for any $z \in [y, v] \cup [u, x]$. Indeed, given a $z \in [y, v]$, we can construct a chain from $y$ to $z$ whose transpositions are in $\overline{T}(y, [y, z]) \subset T(y, [y, v])$. Analogously, for $z \in [u, x]$, we can construct a chain from $z$ to $x$ with transpositions in $\overline{T}(x, [z, x]) \subset T(x, [u, x])$. Applying (3.11) and (3.12), respectively, yields the result.

Now let $q \in [x, y]$. We show that

$$\omega(q) \geq \omega(z) \forall z \geq q, z \in [u, v].$$

Proceed by induction on $m := \ell(y) - \ell(q) \geq 0$. The base case holds by assumption. Consider now such a $q$ with $m \geq 1$. Suppose that $z \in [u, v]$, where $z \geq q$, and take $q' \geq q$ with $q' \in [x, y]$. We have the following diagram

for some $z' \geq z, q'$, with $z' \in [u, v]$. The existence of such a $z'$ comes from the fact that Bruhat order is a directed poset [BB05, Proposition 2.2.9] and from the structure of its length 2 intervals [BB05, Lemma 2.7.3]. By induction, $\omega(q') \geq \omega(z')$. We also know that $\omega(q) = \omega(q')$. Applying Lemma 3.4.17 completes the induction.

We have shown in particular that $\omega(x) \geq \omega(z)$ for all $x \leq z \in [u, v]$. Applying Corollary 3.3.13 shows that $\omega(x) \geq \omega(z) \forall z \in [x, v]$. By symmetry, $\omega(y) \geq \omega(z) \forall z \in [u, y]$. $\square$

**Theorem 3.4.19.** Let $[x, y] \subset [u, v]$. We define the graph $G_{x,y}^{u,v}$ as follows:

1. The nodes of $G_{x,y}^{u,v}$ are $\{1, 2, \ldots, n\}$, with nodes $i$ and $j$ identified if they are in the same part of $B_{x,y}$.

2. There is a directed edge from $i$ to $j$ for every $(i, j) \in \overline{T}(y, [u, v])$.

3. There is a directed edge from $j$ to $i$ for every $(i, j) \in \overline{T}(x, [u, v])$.

Then the Bruhat interval polytope $Q_{x,y}$ is a face of the Bruhat interval polytope $Q_{u,v}$ if and only if the graph $G_{x,y}^{u,v}$ is a directed acyclic graph.

Proof. Assume first that $\omega : \mathbb{R}^n \to \mathbb{R}$ is a linear functional with $\omega|_{Q_{u,v}}$ maximized exactly on $Q_{x,y}$. Since $\omega$ is constant on $[x, y]$, $\omega$ is compatible with the partition $B_{x,y}$, i.e. $\omega_i = \omega_j$ whenever $i$ and $j$ are in the same part of $B_{x,y}$. From the definition of $\omega$,

$$\omega(y) > \omega(b) \text{ for all } b \geq y, b \in [u, v], \text{ and}$$

$$\omega(x) > \omega(a) \text{ for all } a \leq x, a \in [u, v].$$
Equivalently,
\[ \omega_i < \omega_j \text{ for all } (i, j) \in T(y, [u, v]), \text{ and} \]
\[ \omega_i > \omega_j \text{ for all } (i, j) \in T(x, [u, v]). \]

Label each vertex \( k \in G_{x,y}^{u,v} \) with the number \( \omega_k \). If \( G_{x,y}^{u,v} \) has a directed edge from \( i \) to \( j \), then \( \omega_i < \omega_j \). It follows that \( G_{x,y}^{u,v} \) is acyclic.

Conversely, we assume that \( G_{x,y}^{u,v} \) is acyclic. Consequently, there exists a linear ordering \( L \) of the vertices of \( G_{x,y}^{u,v} \) such that for every directed edge \( i \to j \) from vertex \( i \) to vertex \( j \), \( i \) comes before \( j \) in the ordering (i.e., \( i \to j \implies L(i) < L(j) \)). Define \( \omega : \mathbb{R}^n \to \mathbb{R} \) via
\[ \omega_i := L(i). \]

Since \( \omega \) is constant on each block of \( B_{x,y} \), \( \omega \) is constant on \([x, y]\). Also, \( \omega_i < \omega_j \) for all \( (i, j) \in T(y, [u, v]) \), and (3.13) \( \omega_i > \omega_j \) for all \( (i, j) \in T(x, [u, v]) \), (3.14)

so that
\[ \omega(y) > \omega(b) \text{ for all } b > y, b \in [u, v], \text{ and} \]
\[ \omega(x) > \omega(a) \text{ for all } a < x, a \in [u, v]. \]

We show now that these conditions imply that \( \omega \) defines \( Q_{x,y} \) as a face of \( Q_{u,v} \). Indeed, note that \( y \) is a vertex of \( Q_{u,v} \), and by Theorem 3.4.1, any edge of \( Q_{x,y} \) emanating from \( y \) corresponds to a cover relation of \( y \). Thus, if \( f \) is an edge vector emanating from \( y \), then \( y + f = z \) for some \( z \geq y \) or \( z \leq y \) in \([u, v]\). Consequently, by Lemma 3.4.18,
\[ \omega(z) \leq \omega(y) \implies \omega(f) \leq 0. \]

This argument shows that \( \omega(f) \leq 0 \) for any edge vector. By convexity, \( Q_{u,v} \) is contained in the polyhedral cone spanned by the edges emanating from \( y \). Therefore \( \omega(y) \geq \omega(z) \forall z \in [u, v] \).

Similarly, \( \omega(x) \geq \omega(z) \forall z \in [u, v] \). It follows that \([x, y]\) is a subset of the face defined by \( \omega \). By Theorem 3.4.1, this face corresponds to an interval, which we showed contains \([x, y]\). Inequalities (3.13), (3.14) imply that this interval is no larger than \([x, y]\).

From the proof we obtain

**Corollary 3.4.20.** The normal cone of \( Q_{x,y} \) in \( Q_{u,v} \) is the set of linear functionals \( \omega = (\omega_i) \) compatible with \( G_{x,y}^{u,v} \):

1. \( \omega_i = \omega_j \) if \( i, j \) are identified nodes of \( G_{x,y}^{u,v} \);

2. \( \omega_i < \omega_j \) if there is a directed edge from \( i \) to \( j \) in \( G_{x,y}^{u,v} \).

**Example 3.4.21.** Set \([u, v] = [1243, 4132]\). Let us verify using Theorem 3.4.19 that the BIP \( Q_{2143,4132} \) corresponding to \([x, y] = [2143, 4132] \) is a face of \( Q_{1243,4132} \).
The interval \([x, y]\) along with its neighbors in \([u, v]\) are
\[
\begin{array}{c}
4132 \\
4123 \\
3142 \\
2143 \\
1243
\end{array}
\]

Therefore the graph \(G_{x,y}^{u,v}\) is
\[
\begin{array}{c}
1, 3, 4 \\
2
\end{array}
\]
which is clearly acyclic.

**Diameter of Bruhat interval polytopes**

In this section, we show that the diameter of \(Q_{u,v}\) is equal to \(\ell(v) - \ell(u)\) (see Theorem 3.4.24). Let
\[
\overline{E}(x, [u, v]) := \{z - x \in \mathbb{R}^n : z \geq x, z \in [u, v]\}, \quad \underline{E}(x, [u, v]) := \{z - x \in \mathbb{R}^n : z \leq x, z \in [u, v]\}.
\]
We note that \(\overline{E}(x, [u, v]), \underline{E}(x, [u, v]) \subseteq [u, v] - x\).

For a set \(\{v_1, \ldots, v_m\} \subseteq \mathbb{R}^n\), define
\[
\text{Cone}(\{v_1, \ldots, v_m\}) := \mathbb{R}_{\geq 0}v_1 + \ldots + \mathbb{R}_{\geq 0}v_m.
\]
Lemma 3.4.22. If \( e \in \overline{E}(x, [u, v]) \), then \( e \not\in \text{Cone}(\overline{E}(x, [u, v])) \). Similarly, if \( e \in \overline{E}(x, [u, v]) \), then \( e \not\in \text{Cone}(\overline{E}(x, [u, v])) \).

Proof. Recall that a set of simple roots for type A is given by

\[ e_i - e_{i+1}, \quad i = 1, 2, \ldots, n - 1, \]

and that any other root can be expressed uniquely as a linear combination of simple roots with integral coefficients of the same sign.

Let \( e \in \overline{E}(x, [u, v]) \). Then \( e \) is of the form

\[ e = c(e_j - e_i), \quad i < j, c > 0, \]

which is in the cone of negative roots. On the other hand, \( \text{Cone}(\overline{E}(x, [u, v])) \) is a subset of the cone of positive roots. It follows that \( e \not\in \text{Cone}(\overline{E}(x, [u, v])) \). The argument for \( e \in \overline{E}(x, [u, v]) \) is analogous. \( \square \)

Lemma 3.4.23. Let \( x \in (u, v) \subset S_n \). The sets \( \overline{E}(x, [u, v]) \) and \( \overline{E}(x, [u, v]) \) each contain an edge of \( Q_{u,v} \) incident to \( x \).

Proof. As a consequence of Theorem 3.4.1, the edges of \( Q_{u,v} \) incident to \( x \) are a subset of \( \overline{E}(x, [u, v]) \). Assume by contradiction that all edges of \( Q_{u,v} \) incident to \( x \) are in \( \overline{E}(x, [u, v]) \). Then by convexity of \( Q_{u,v} \),

\[ Q_{u,v} \subset x + \text{Cone}(\overline{E}(x, [u, v])) \Rightarrow \text{Cone}(Q_{u,v} - x) \subset \text{Cone}(\overline{E}(x, [u, v])). \]

Since \( \text{Cone}(\overline{E}(x, [u, v])) \subset \text{Cone}(Q_{u,v} - x) \), we have \( \text{Cone}(\overline{E}(x, [u, v])) \subset \text{Cone}(\overline{E}(x, [u, v])) \). By assumption, \( x \not\in \{u, v\} \), so that \( \overline{E}(x, [u, v]) \neq \emptyset \), contradicting Lemma 3.4.22. The argument for \( \overline{E}(x, [u, v]) \) is analogous. \( \square \)

Theorem 3.4.24. The diameter of \( Q_{u,v} \) is equal to \( \ell(v) - \ell(u) \).

Proof. Since any edge of a Bruhat Interval polytope corresponds to a cover relation in Bruhat order, the distance from \( u \) to \( v \) is at least \( \ell(v) - \ell(u) \). By Lemma 3.4.23, there is a path from \( u \) to \( v \) which takes steps up in the Hasse diagram. Therefore the distance from \( u \) to \( v \) is \( \ell(v) - \ell(u) \).

Next, take \( x \neq y \) in \( [u, v] \). By Lemma 3.4.23, a path from \( x \) to \( y \) can be formed by either taking a path of length \( \ell(v) - \ell(x) \) from \( x \) to \( v \) and then of length \( \ell(v) - \ell(y) \) from \( v \) to \( y \), or of length \( \ell(x) - \ell(u) \) from \( x \) to \( u \) and then of length \( \ell(y) - \ell(u) \) from \( u \) to \( y \). One of these paths must be of length less than or equal to \( \ell(v) - \ell(u) \), since the sum of the lengths is \( 2(\ell(v) - \ell(u)) \). \( \square \)

**An inequality description of Bruhat interval polytopes**

Using Proposition 3.2.8, which says that Bruhat interval polytopes are Minkowski sums of matroid polytopes, we will provide an inequality description of Bruhat interval polytopes.
We first need to recall the notion of the rank function $r_M$ of a matroid $M$. Suppose that $M$ is a matroid of rank $k$ on the ground set $[n]$. Then the rank function $r_M : 2^{[n]} \to \mathbb{Z}_{\geq 0}$ is the function defined by

$$r_M(A) = \max_{I \in M} |A \cap I| \text{ for all } A \in 2^{[n]}.$$ 

There is an inequality description of matroid polytopes, using the rank function.

**Proposition 3.4.25** ([Wel76]). Let $M$ be any matroid of rank $k$ on the ground set $[n]$, and let $r_M : 2^{[n]} \to \mathbb{Z}_{\geq 0}$ be its rank function. Then the matroid polytope $\Gamma_M$ can be described as

$$\Gamma_M = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i \in A} x_i = k, \sum_{i \in A} x_i \leq r_M(A) \text{ for all } A \subset [n] \right\}.$$ 

Using Proposition 3.4.25 we obtain the following result.

**Proposition 3.4.26.** Choose $u \leq v \in S_n$, and for each $1 \leq k \leq n - 1$, define the matroid

$$M_k = \{ I \in \binom{[n]}{k} \mid \text{there exists } z \in [u, v] \text{ such that } I = \{z(1), \ldots, z(k)\} \}.$$ 

Then

$$Q_{u,v} = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i \in [n]} x_i = \frac{n + 1}{2}, \sum_{i \in A} x_i \leq \sum_{j=1}^{n-1} r_{M_j}(A) \text{ for all } A \subset [n] \right\}.$$ 

**Proof.** We know from Proposition 3.2.8 that $Q_{u,v}$ is the Minkowski sum

$$Q_{u,v} = \sum_{k=1}^{n-1} \Gamma_{M_k},$$

where $M_k$ is defined as above. But now Proposition 3.4.26 follows from Proposition 3.4.25 and the observation (made in the proof of [ABD10, Lemma 2.1]) that, if a linear functional $\omega$ takes maximum values $a$ and $b$ on (faces $A$ and $B$ of) polytopes $P$ and $Q$, respectively, then it takes maximum value $a + b$ on (the face $A + B$ of) their Minkowski sum. \(\Box\)

**Example 3.4.27.** Consider $u = 1324$ and $v = 2431$ in $S_4$. We will compute the inequality description of $Q_{u,v}$. First note that $[u, v] = \{1324, 1342, 1423, 1432, 2314, 2341, 2413, 2431\}$. We then compute:

- $M_1 = \{\{1\}, \{2\}\}$, a matroid of rank 1 on [4].
- $M_2 = \{\{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$, a matroid of rank 2 on [4].
- $M_3 = \{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}\}$, a matroid of rank 3 on [4].

Now using Proposition 3.4.26, we get

$$Q_{u,v} = \{ \mathbf{x} \in \mathbb{R}^4 \mid \sum_{i \in [4]} x_i = 10, x_1 + x_2 + x_3 \leq 6, x_1 + x_2 + x_4 \leq 6, x_1 + x_3 + x_4 \leq 6, x_2 + x_3 + x_4 \leq 6, x_1 + x_2 \leq 4, x_1 + x_3 \leq 5, x_1 + x_4 \leq 5, x_2 + x_3 \leq 5, x_2 + x_4 \leq 5, x_3 + x_4 \leq 3, x_1 \leq 3, x_2 \leq 3, x_3 \leq 2, x_4 \leq 2. \}$$
3.5 A generalization of the recurrence for $R$-polynomials

The well-known $R$-polynomials were introduced by Kazhdan and Lusztig as a useful tool for computing Kazhdan-Lusztig polynomials [KL79]. $R$-polynomials also have a geometric interpretation in terms of Richardson varieties. More specifically, the Richardson variety $R_{u,v}$ (see Section 3.6 for the definition) may be defined over a finite field $\mathbb{F}_q$, and the number of points it contains is given by the $R$-polynomial $R_{u,v}(q) = \# R_{u,v}(\mathbb{F}_q)$.

The $R$-polynomials may be defined by the following recurrence.

**Theorem 3.5.1.** [BB05, Theorem 5.1.1] There exists a unique family of polynomials $\{R_{u,v}(q)\}_{u,v \in W} \subset \mathbb{Z}[q]$ satisfying the following conditions:

1. $R_{u,v}(q) = 0$, if $u \not\leq v$.
2. $R_{u,v}(q) = 1$, if $u = v$.
3. If $s \in D_R(v)$, then
   
   $$R_{u,v}(q) = \begin{cases} 
   R_{us,vs}(q) & \text{if } s \in D_R(u), \\
   qR_{us,vs}(q) + (q-1)R_{u,vs}(q) & \text{if } s \not\in D_R(u). 
   \end{cases}$$

It is natural to wonder whether one can replace $s$ with a transposition $t$ whenever the Generalized lifting property holds. More precisely, suppose that $t$ is a transposition such that

$$vt \leq v \quad u \leq ut \quad u \leq vt \quad ut \leq v. \quad (3.15)$$

Is it true that

$$R_{u,v}(q) = qR_{ut,vt}(q) + (q-1)R_{u,vt}(q)? \quad (3.16)$$

In general, the answer is no. For example, one can check that $u = 1324$, $v = 4231$ and $t = (24)$ give a counterexample. However, when $t$ is an inversion-minimal transposition on $(u, v), (3.16)$ does hold. We’ll use the next lemma to prove this.

**Lemma 3.5.2.** Let $u, v \in S_n$ and suppose that $(ik)$ is inversion-minimal on $(u, v)$. Assume further that $v_j > v_{j+1}$ and $u_j > u_{j+1}$ for some $j$ such that $i < j < k - 1$. Then $(ik)$ is inversion-minimal on $(vs_j, us_j)$.

**Proof.** The result follows directly from the definition. \[\square\]

**Proposition 3.5.3.** Let $u, v \in S_n$ with $v \geq u$. Let $t = (ij)$ be inversion-minimal on $(u, v)$. Then

$$R_{u,v}(q) = qR_{ut,vt}(q) + (q-1)R_{u,vt}(q).$$
Similarly, by Lemma 3.3.11, we have $u_{i+1} > v_i$. Assume the inductive hypothesis and consider $\ell > 1$. Since $(ij)$ is inversion-minimal on $(u,v)$, we have $v_i > v_j$ and $u_i < u_j$.

**Case 1:** Suppose that $v_i > v_{i+1}$ and $u_i > u_{i+1}$ or $v_i < v_{i+1}$ and $u_i < u_{i+1}$.

We have $R_{u,v}(q) = R_{us_i,vs_i}(q)$. Let $t'$ be the transposition ($(i+1)j$). Clearly $t'$ is inversion-minimal on $(vs_i,us_i)$. By induction,

$$R_{us_i,vs_i}(q) = qR_{us_i,t',vs_i}(q) + (q-1)R_{us_i,vs_i,t'}(q).$$

By Lemma 3.3.11, we have $v_{i+1} > v_j \iff u_{i+1} > u_j$. Using this and the fact that $s_i t' s_i = (ij) = t$, we see that

$$R_{us_i,t',vs_i}(q) = R_{uts_i,uts_i}(q) = R_{ut,vt}(q).$$

Similarly, by Lemma 3.3.11, we have $u_{i+1} > u_i \iff v_{i+1} > v_i \iff v_i > v_{i+1}$. This implies that

$$R_{us_i,vs_i,t'}(q) = R_{u,v}(q).$$

Putting everything together, we have the desired equality

$$R_{u,v}(q) = R_{us_i,vs_i}(q) = R_{ut,vt}(q) + R_{u,v}(q).$$

**Case 2:** Suppose that $v_{j-1} > v_j$ and $u_{j-1} > u_j$ or $v_{j-1} < v_j$ and $u_{j-1} < u_j$.

This case is analogous to Case 1.

**Case 3:** Suppose that neither of the above two cases holds.

Since $(ij)$ is inversion-minimal on $(u,v)$, we must have $v_i < v_{i+1}$ and $v_{j-1} < v_j$. Since $v_i > v_j$, there exists some $m_1 \in (i,j-1)$ such that $v_{m_1} > v_{m_1+1}$. Using the fact that $(ij)$ is inversion-minimal on $(u,v)$, we must have $u_{m_1} > u_{m_1+1}$. By Lemma 3.5.2, $(ij)$ is inversion-minimal on $(vs_{m_1},us_{m_1})$. If $us_{m_1}$ and $vs_{m_1}$ do not satisfy the conditions of Cases 1 or 2, then we may find $m_2 \in (i,j-1)$ and then $(ij)$ is inversion-minimal on $(vs_{m_1}s_{m_2},us_{m_1}s_{m_2})$. Such a sequence $m_1,m_2,...$ clearly terminates. Assume that it terminates at $k$, so that $(ij)$ is inversion-minimal on $(vs_{m_1}s_{m_2}...s_{m_k},us_{m_1}s_{m_2}...s_{m_k})$ and the hypotheses of Case 1 or 2 are satisfied for $vs_{m_1}s_{m_2}...s_{m_k}$ and $us_{m_1}s_{m_2}...s_{m_k}$. Set $\Pi_k := s_{m_1}s_{m_2}...s_{m_k}$. We then have

$$R_{u\Pi_k,v\Pi_k}(q) = R_{u\Pi_k,t,v\Pi_k}(q) + qR_{u\Pi_k,v\Pi_k,t}(q).$$

To prove Proposition 3.5.3, it suffices to show that for $1 \leq p \leq k$, if

$$R_{u\Pi_p,v\Pi_p}(q) = qR_{u\Pi_p,t,v\Pi_p}(q) + (q-1)R_{u\Pi_p,v\Pi_p,t}(q) \quad (3.17)$$

then

$$R_{u\Pi_{p-1},v\Pi_{p-1}}(q) = qR_{u\Pi_{p-1},t,v\Pi_{p-1}}(q) + (q-1)R_{u\Pi_{p-1},v\Pi_{p-1},t}(q). \quad (3.18)$$

By Proposition 3.5.1, we have that

$$R_{u\Pi_p,v\Pi_p}(q) = R_{u\Pi_{p-1},v\Pi_{p-1}}(q).$$
Note that for any $m$, $ts_m = s_m t$. Therefore $u\Pi_p t = u\Pi_{p-1} ts_m$ and $v\Pi_p t = v\Pi_{p-1} ts_m$. This implies that

$$R_{u\Pi_p t, v\Pi_p t}(q) = R_{u\Pi_{p-1} ts_m, v\Pi_{p-1} ts_m}(q).$$

Similarly, we have that

$$R_{u\Pi_p, v\Pi_p t}(q) = R_{u\Pi_{p-1}, v\Pi_{p-1} t}(q).$$

This shows that (3.17) implies (3.18).

**Remark 3.5.4.** The above statement and proof hold mutatis mutandis for the $\tilde{R}$-polynomials, which are a renormalization of the $R$-polynomials.

**Example 3.5.5.** Take $u = 21345$, $v = 53421$ and $t = (13)$. We have

$$R_{u, v}(q) = q^8 - 4q^7 + 7q^6 - 8q^5 + 8q^4 - 8q^3 + 7q^2 - 4q + 1$$

and

$$R_{u, vt}(q) = q^6 - 4q^5 + 7q^4 - 8q^3 + 7q^2 - 4q + 1$$

and

$$R_{u, vt}(q) = q^7 - 4q^6 + 7q^5 - 8q^4 + 8q^3 - 7q^2 + 4q - 1.$$

**Definition 3.5.6.** A matching of a graph $G = (V, E)$ is an involution $M : V \rightarrow V$ such that \{v, M(v)\} $\in E$ for all $v \in V$.

**Definition 3.5.7.** Let $P$ be a graded poset. A matching $M$ of the Hasse diagram of $P$ is a special matching if for all $x, y \in P$ such that $x \leq y$, we have $M(x) = y$ or $M(x) \leq M(y)$.

It is known that special matchings can be used to compute $R$-polynomials:

**Theorem 3.5.8.** [BCM06, Theorem 7.8] Let $(W, S)$ be a Coxeter system, let $w \in W$, and let $M$ be a special matching of the Hasse diagram of the interval $[e, w]$ in Bruhat order. Then

$$R_{u, w}(q) = q^c R_{M(u), M(w)}(q) + (q^c - 1) R_{u, M(w)}(q)$$

for all $u \leq w$, where $c = 1$ if $M(u) \gg u$ and $c = 0$ otherwise.

One might guess that the Generalized lifting property is compatible with the notion of special matching. More precisely, one might speculate that if $[u, v] \subset S_n$ and $t$ is inversion-minimal on $(u, v)$ then there is a special matching $M$ of $[u, v]$ such that $M(u) = ut$ and $M(v) = vt$. The following gives an example of this.

**Example 3.5.9.** Take $u = 143265$ and $v = 254163$. Then $t = (36)$ is inversion-minimal on $(u, v)$. Suppose that a special matching $M$ of $[u, v]$ (see Figure 3.6) satisfies $M(v) = vt$ and $M(u) = ut$. Then we must have $M(154263) = 153264$ and $M(243165) = 245163$. Observe that the result is a multiplication matching. Similarly, if we take $t = (14)$, another inversion-minimal transposition on $(u, v)$, we again obtain a multiplication matching.
Figure 3.6: Bruhat interval $[143265, 254163]$.

The following example shows that it is not the case that an inversion-minimal transposition must be compatible with a special matching. This makes Proposition 3.5.3 all the more surprising, and shows that it cannot be deduced using special matchings.

**Example 3.5.10.** Take $u = 1324$ and $v = 4312$. Then $t = (24)$ is inversion-minimal on $(u, v)$. Suppose that a special matching $M$ of $[u, v]$ (Figure 3.7) satisfies $M(v) = vt$, i.e., sends $4312$ to $4213$. Then

$$
M(4132) = 4123, \ M(1432) = 1423, \ M(1342) = 1324, \ M(3142) = 3124, \ M(3412) = 3214, \ M(2413) = 2314.
$$

But $M(2314) = 2413 \ngeq 1342 = M(1324)$, which is a contradiction.

Figure 3.7: Bruhat interval $[1324, 4312]$. 
3.6 Background on partial flag varieties \( G/P \)

Preliminaries

over \( \mathbb{R} \), with split torus \( T \). We identify \( G \) (and related spaces) with their real points and consider them with their real topology. Let \( t \) denote the Lie algebra of \( T \), \( t_{\mathbb{R}} \) denote its real part, and let \( t_{\mathbb{R}}^* \) denote the dual of the torus. Let \( \Phi \subset t_{\mathbb{R}}^* \) denote the set of roots, and choose a system of positive roots \( \Phi^+ \). We denote by \( B = B^+ \) the Borel subgroup corresponding to \( \Phi^+ \) and by \( U^+ \) its unipotent radical. We also have the opposite Borel subgroup \( B^- \) such that \( B^+ \cap B^- = T \), and its unipotent radical \( U^- \). For background on algebraic groups, see e.g. [Hum75].

Let \( \Pi = \{ \alpha_i \mid i \in I \} \subset \Phi^+ \) denote the simple roots, and let \( \{ \omega_i \mid i \in I \} \) denote the fundamental weights. For each \( \alpha_i \in \Pi \) there is an associated homomorphism \( \phi_i : SL_2 \to G \). Consider the 1-parameter subgroups in \( G \) (landing in \( U^+, U^- \), and \( T \), respectively) defined by

\[
x_i(m) = \phi_i \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}, \quad y_i(m) = \phi_i \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}, \quad \alpha_i^\vee(\ell) = \phi_i \begin{pmatrix} \ell & 0 \\ 0 & \ell^{-1} \end{pmatrix},
\]

where \( m \in \mathbb{R}, \ell \in \mathbb{R}^*, i \in I \). The datum \((T, B^+, B^-, x_i, y_i; i \in I)\) for \( G \) is called a pinning. The standard pinning for \( SL_n \) consists of the diagonal, upper-triangular, and lower-triangular matrices, along with the simple root subgroups \( x_i(m) = I_n + mE_{i,i+1} \) and \( y_i(m) = I_n + mE_{i+1,i} \) where \( I_n \) is the identity matrix and \( E_{i,j} \) has a 1 in position \((i,j)\) and zeroes elsewhere.

The Weyl group \( W = N_G(T)/T \) acts on \( t_{\mathbb{R}}^* \), permuting the roots \( \Phi \). We set \( s_i := \dot{s}_i T \) where \( \dot{s}_i := \phi_i \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \). Then any \( w \in W \) can be expressed as a product \( w = s_{i_1} s_{i_2} \ldots s_{i_m} \) with \( \ell(w) \) factors. This gives \( W \) the structure of a Coxeter group; we will assume some basic knowledge of Coxeter systems and Bruhat order as in [BB05]. We set \( \dot{w} = \dot{s}_{i_1} \dot{s}_{i_2} \ldots \dot{s}_{i_m} \). It is known that \( \dot{w} \) is independent of the reduced expression chosen.

The \emph{(complete) flag variety} is the homogeneous space \( G/B^+ = G/B \). Note that we will frequently use \( B \) to denote \( B^+ \).

We have two opposite Bruhat decompositions of \( G/B \):

\[
G/B = \bigsqcup_{v \in W} B v B/B = \bigsqcup_{u \in W} B^- u B/B.
\]

We define the intersection of opposite Bruhat cells

\[
\mathcal{R}_{u,v} := (B v B/B) \cap (B^- u B/B),
\]

which is nonempty precisely when \( u \preceq v \) in Bruhat order, and in that case is irreducible of dimension \( \ell(v) - \ell(u) \), see [KL79]. The strata \( \mathcal{R}_{u,v} \) are often called \emph{Richardson varieties}.

Let \( J \subset I \). The parabolic subgroup \( W_J \subset W \) corresponds to a parabolic subgroup \( P_J \) in \( G \) containing \( B \). Namely, \( P_J = \bigsqcup_{w \in W_J} B w B \). There is a corresponding \emph{generalized partial flag variety}, which is the homogeneous space \( G/P_J \).

There is a natural projection from the complete flag variety to a partial flag variety which takes the form \( \pi = \pi_J : G/B \to G/P_J \), where \( \pi(gB) = gP_J \).
Generalized Plücker coordinates and the Gelfand-Serganova stratification of $G/P$

Let $P = P_J$ be a parabolic subgroup of $G$. In [GS87], Gelfand and Serganova defined a new stratification of $G/P$. In the case that $G = SL_n$ and $P$ is a maximal parabolic subgroup, their stratification recovers the well-known matroid stratification of the Grassmannian.

Let $C$ be a Borel subgroup of $G$ containing the maximal torus $T$. The Schubert cells on $G/P$ associated with $C$ are the orbits of $C$ in $G/P$. The Schubert cells are in bijection with $W^J$, and can be written as $C\dot{w}P$ where $w \in W^J$.

**Definition 3.6.1.** The Gelfand-Serganova stratification (or thin cell stratification) of $G/P$ is the simultaneous refinement of all the Schubert cell decompositions described above. The (nonempty) strata in this decomposition are called Gelfand-Serganova strata or thin cells.

In other words, we choose for each Borel subgroup $C$ a Schubert cell associated with $C$. The intersection of all chosen cells, if it is nonempty, is called a Gelfand-Serganova stratum or a thin cell.

There is another way to think about the Gelfand-Serganova stratification, using generalized Plücker coordinates.

Let $J \subset I$ index the simple roots corresponding to the parabolic subgroup $P = P_J$, and let $\rho_J = \sum_{j \in J} \omega_j$. Let $V_{\rho_J}$ be the representation of $G$ with highest weight $\rho_J$, and choose a highest weight vector $\eta_J$. Recall that we have an embedding of the flag variety $G/P \hookrightarrow \mathbb{P}(V_{\rho_J})$ given by

$$gP \mapsto g \cdot \eta_J.$$

Let $\mathcal{A}$ be the set of weights of $V_{\rho_J}$ taken with multiplicity. We choose a weight basis $\{e_\alpha | \alpha \in \mathcal{A}\}$ in $V_{\rho_J}$. Then any point $X \in G/P$ determines, uniquely up to scalar $d$, a collection of numbers $p^\alpha(X)$, where

$$X = d \sum_{\alpha \in \mathcal{A}} p^\alpha(X) e_\alpha. \quad (3.19)$$

Let $W(\rho_J) \subset \mathfrak{t}_R^*$ be the orbit of $\rho_J$ under $W$. Then $W(\rho_J)$ are the extremal weight vectors, that is, they lie at the vertices of some convex polytope $\Delta_P$, and the other elements of $\mathcal{A}$ lie inside of $\Delta_P$, see [Ati82, GS87]. The extremal weight vectors can be identified with the set $W/W_J$ of cosets via the map $w \cdot \rho_J \mapsto wW_J$.

**Definition 3.6.2.** Let $X \in G/P$. The numbers $\{p^\alpha(X) | \alpha \in W(\rho_J)\}$, defined up to scalar, are called the generalized Plücker coordinates of $X$. And the list of $X$ is the subset

$$L_X \subset W(\rho_J) = \{\alpha \in W(\rho_J) | p^\alpha(X) \neq 0\}.$$ 

**Example 3.6.3.** Let $G = SL_n$ and $P = SL_k \times SL_{n-k}$; note that $G/P \cong Gr_{k,n}$. Let $V$ denote the $n$-dimensional vector space with standard basis $e_1, \ldots, e_n$. The vector $e_1 \wedge e_2 \wedge \cdots \wedge e_k$ is a highest weight vector for the representation $\bigwedge^k V$ of $G$, and we have an embedding

$$G/P \hookrightarrow \mathbb{P}(\bigwedge^k V)$$
CHAPTER 3. BRUHAT INTERVAL POLYTOPES

given by

\[ gP \mapsto g(e_1 \wedge e_2 \wedge \cdots \wedge e_k). \]

Expanding the right-hand side in the natural basis, we get

\[ g(e_1 \wedge e_2 \wedge \cdots \wedge e_k) = \sum_{I \in \binom{[n]}{k}} \Delta_I(A)e_{i_1} \wedge \cdots \wedge e_{i_k}, \]

where \( I = \{ i_1 < \cdots < i_k \} \), and \( A = \pi_k(g) \in Gr_{k,n} \) is the span of the leftmost \( k \) columns of \( A \). This shows that the generalized Plücker coordinates agree with the Plücker coordinates from Section 3.2 in the case of the Grassmannian.

Theorem 3.6.4. [GS87, Theorem 1] Two points \( X,Y \in G/P \) lie in the same Gelfand-Serganova stratum if and only if they have the same list.

Total positivity

We start by reviewing the totally nonnegative part \((G/P)_{\geq 0}\) of \( G/P \), and Rietsch’s cell decomposition of it. We then relate this cell decomposition to the Gelfand-Serganova stratification.

Definition 3.6.5. [Lus94] The totally non-negative part \( U^-_{\geq 0} \) of \( U^- \) is defined to be the semigroup in \( U^- \) generated by the \( y_i(t) \) for \( t \in \mathbb{R}_{\geq 0} \).

The totally non-negative part \((G/B)_{\geq 0}\) of \( G/B \) is defined by

\[ (G/B)_{\geq 0} := \{ yB \mid y \in U^-_{\geq 0} \}, \]

where the closure is taken inside \( G/B \) in its real topology.

The totally non-negative part \((G/P_J)_{\geq 0}\) of \( G/P_J \) is defined to be \( \pi_J((G/B)_{\geq 0}) \).

Lusztig [Lus94, Lus98] introduced natural decompositions of \((G/B)_{\geq 0}\) and \((G/P)_{\geq 0}\).

Definition 3.6.6. [Lus94] For \( u,v \in W \) with \( u \leq v \), let

\[ R_{u,v;>0} := R_{u,v} \cap (G/B)_{\geq 0}. \]

We write \( W^J \) (respectively \( W^J_{\max} \)) for the set of minimal (respectively maximal) length coset representatives of \( W/W_J \).

Definition 3.6.7. [Lus98] For \( u \in W \) and \( v \in W^J \) with \( u \leq v \), define \( P^J_{u,v;>0} := \pi^J(R_{u,v;>0}). \)

Here the projection \( \pi^J \) is an isomorphism from \( R_{u,v;>0} \) to \( P^J_{u,v;>0} \).

Lusztig conjectured and Rietsch proved [Rie99, Rie98] that \( R^J_{u,v} \) (and hence \( P^J_{u,v;>0} \)) is a semi-algebraic cell of dimension \( \ell(v) - \ell(u) \). Subsequently Marsh-Rietsch [MR04] provided an explicit parameterization of each cell. To state their result, we first review the notion of positive distinguished subexpression, as in [Deo85] and [MR04].

Let \( v := s_{i_1} \cdots s_{i_m} \) be a reduced expression for \( v \in W \). A subexpression \( u \) of \( v \) is a word obtained from the reduced expression \( v \) by replacing some of the factors with 1. For example, consider a reduced expression in the symmetric group \( S_4 \), say \( s_3s_2s_1s_3s_2s_3 \). Then \( 1s_211s_2s_3 \) is a subexpression of \( s_3s_2s_1s_3s_2s_3 \). Given a subexpression \( u \), we set \( u(k) \) to be the product of the leftmost \( k \) factors of \( u \), if \( k \geq 1 \), and \( u(0) = 1 \).
Definition 3.6.8. [Deo85, MR04] Given a subexpression $u$ of $v = s_{i_1}s_{i_2} \ldots s_{i_m}$, we define

$$J_u^+ := \{k \in \{1, \ldots, m\} \mid u_{(k-1)} < u_k\},$$
$$J_u^- := \{k \in \{1, \ldots, m\} \mid u_{(k-1)} = u_k\},$$
$$J_u^0 := \{k \in \{1, \ldots, m\} \mid u_{(k-1)} > u_k\}.$$

The subexpression $u$ is called non-decreasing if $u_{(j-1)} \leq u_j$ for all $j = 1, \ldots, m$, e.g. if $J_u^0 = \emptyset$. It is called distinguished if we have $u_j \leq u_{(j-1)}s_i$ for all $j \in \{1, \ldots, m\}$. In other words, if right multiplication by $s_i$ decreases the length of $u_{(j-1)}$, then in a distinguished subexpression we must have $u_{(j)} = u_{(j-1)}s_i$. Finally, $u$ is called a positive distinguished subexpression (or a PDS for short) if $u_{(j-1)} < u_{(j-1)}s_i$ for all $j \in \{1, \ldots, m\}$. In other words, it is distinguished and non-decreasing.

Lemma 3.6.9. [MR04] Given $u \leq v$ and a reduced expression $v$ for $v$, there is a unique PDS $u_+$ for $u$ contained in $v$.

Theorem 3.6.10. [MR04, Proposition 5.2, Theorem 11.3] Choose a reduced expression $v = s_{i_1} \ldots s_{i_m}$ for $v$ with $\ell(v) = m$. To $u \leq v$ we associate the unique PDS $u_+$ for $u$ in $v$. Then $J_{u^+}^0 = \emptyset$. We define

$$G_{u_+,v}^{>0} := \left\{ g = g_1g_2 \ldots g_m \mid \begin{array}{ll}
g_i = g_i(p_{\ell}) & \text{if } \ell \in J_{u^+}^+ \\
g_i = s_i & \text{if } \ell \in J_{u^+}^0
g_i = s_i & \text{if } \ell \in J_{u^+}^0
g_i = s_i & \text{if } \ell \in J_{u^+}^0,
\end{array} \right\},$$

(3.20)

where each $p_{\ell}$ ranges over $\mathbb{R}_{>0}$. Then $G_{u_+,v}^{>0} \cong \mathbb{R}_{>0}^{\ell(v)-\ell(u)}$, and the map $g \mapsto gP_J$ defines an isomorphism

$$G_{u_+,v}^{>0} \xrightarrow{\sim} \mathbb{R}_{>0}^{>0}.$$

Remark 3.6.11. Use the notation of Theorem 3.6.10, and now assume additionally that $v \in W^J$. Then Theorem 3.6.10 and Definition 3.6.7 imply that the map $g \mapsto gP_J$ defines an isomorphism

$$G_{u_+,v}^{>0} \xrightarrow{\sim} P_{u,v;}^{J>0}.$$

Definition 3.6.12. Let $T_{>0}$ denote the positive part of the torus, i.e. the subset of $T$ generated by all elements of the form $\alpha_i^{>0}(\ell) = \phi_i \left( \begin{array}{cc} \ell & 0 \\
0 & \ell^{-1} \end{array} \right)$, where $\ell \in \mathbb{R}_{>0}$.

Lemma 3.6.13. Let $t \in T_{>0}$ and $gP_J \in P_{u,v;}^{J>0}$. Then $tgB \in P_{u,v;}^{J>0}$.

Proof. We claim that for any $t \in T_{>0}$ and $a \in \mathbb{R}_{>0}$, we have $ts_i = \hat{s}_it'$ for some $t' \in T_{>0}$, and also $ty_i(a) = y_i(a't)$ for some $a' \in \mathbb{R}_{>0}$. If we can demonstrate this claim, then the lemma follows from the parameterization of cells given in Theorem 3.6.10 and Remark 3.6.11: using the claim, we can simply factor the $t$ all the way to the right where it will get absorbed into the group $P_J$.

To prove the first part of the claim, note that since $\hat{s}_i$ lies in the normalizer of the torus $N_G(T)$, for any $t \in T$ we have that $\hat{s}_it\hat{s}_i^{-1} = t'$ for some $t' \in T$. Moreover, if $t \in T_{>0}$ then
also \( t' \in T_{>0} \): one way to see this is to use the fact that \( T_{>0} \) is the connected component of \( T \) containing 1 [Lus94, 5.10]. Then since \( s_i T_{>0} s_i^{-1} \) is also connected and contains 1, its elements must all lie in \( T_{>0} \).

To prove the second part of the claim, note that by [Lus94, 1.3 (b)], we have \( ty_i(a) = y_i(\chi_i(t)^{-1}a)t \) for any \( i \in I, t \in T \), and \( a \in \mathbb{R} \), where \( \chi_i : T \to \mathbb{R}^* \) is the simple root corresponding to \( i \). When \( t \in T_{>0} \) and \( a \in \mathbb{R}_{>0} \), we have \( \chi_i(t)^{-1}a > 0 \). \( \square \)

Rietsch also showed that the closure of each cell of \( (G/P_I)_{>0} \) is a union of cells, and described when one cell of \( (G/P_I)_{>0} \) lies in the closure of another [Rie06]. Using this description, it is easy to determine the set of 0-cells contained in the closure \( \overline{P_{u,v}^J} \).

**Corollary 3.6.14.** The 0-cells in the closure of the cell \( P_{u,v}^J \) of \( (G/P_I)_{>0} \) are in bijection with the cosets

\[
\{ zW_J \mid u \leq z \leq v \}.
\]

More specifically, those 0-cells are precisely the cells of the form \( P_{\tilde{z}, \tilde{z}}^J \) where \( \tilde{z} \) is the minimal-length coset representative for \( z \) in \( W/W_J \), and \( u \leq z \leq v \).

**Remark 3.6.15.** The 0-cells in \( \overline{P_{u,v}^J} \) are precisely the torus fixed points of \( G/P_I \) that lie in \( \overline{P_{u,v}^J} \).

**Total positivity and canonical bases for simply laced \( G \)**

Assume that \( G \) is simply laced. Let \( U \) be the enveloping algebra of the Lie algebra of \( G \); this can be defined by generators \( e_i, h_i, f_i \) \( (i \in I) \) and the Serre relations. For any dominant weight \( \lambda \in \mathfrak{t}_G^* \), there is a finite-dimensional simple \( U \)-module \( V(\lambda) \) with a non-zero vector \( \eta \) such that \( e_i \cdot \eta = 0 \) and \( h_i \cdot \eta = \lambda(h_i)\eta \) for all \( i \in I \). The pair \( (V(\lambda), \eta) \) is determined up to unique isomorphism.

There is a unique \( G \)-module structure on \( V(\lambda) \) such that for any \( i \in I, a \in \mathbb{R} \) we have

\[
x_i(a) = \exp(ae_i) : V(\lambda) \to V(\lambda), \quad y_i(a) = \exp(af_i) : V(\lambda) \to V(\lambda).
\]

Then \( x_i(a) \cdot \eta = \eta \) for all \( i \in I, a \in \mathbb{R} \), and \( t \cdot \eta = \lambda(t)\eta \) for all \( t \in T \). Let \( B(\lambda) \) be the canonical basis of \( V(\lambda) \) that contains \( \eta \) [Lus90]. We now collect some useful facts about the canonical basis.

**Lemma 3.6.16.** [Lus98, 1.7(a)]. For any \( w \in W \), the vector \( \check{w} \cdot \eta \) is the unique element of \( B(\lambda) \) which lies in the extremal weight space \( V(\lambda)^{\check{w} \cdot \lambda} \). In particular, \( \check{w} \cdot \eta \in B(\lambda) \).

We define \( f_i^{(p)} \) to be \( f_i^p \).

**Lemma 3.6.17.** Let \( s_{i_1} \ldots s_{i_n} \) be a reduced expression for \( w \in W \). Then there exists \( a \in \mathbb{N} \) such that \( f_i^{(a)} \hat{s}_{i_1} \hat{s}_{i_2} \ldots \hat{s}_{i_n} \cdot \eta = \hat{s}_{i_1} \hat{s}_{i_2} \ldots \hat{s}_{i_n} \cdot \eta \). Moreover, \( f_i^{(a+1)} \hat{s}_{i_1} \hat{s}_{i_2} \ldots \hat{s}_{i_n} \cdot \eta = 0 \).

**Proof.** This follows from Lemma 3.6.16 and properties of the canonical basis, see e.g. the proof of [Lus10, Proposition 28.1.4]. \( \square \)
The moment map for $G/P$

In this section we start by defining the moment map for $G/P$ and describing some of its properties. We then give a result of Gelfand-Serganova [GS87] which gives another description of their stratification of $G/P$ in terms of the moment map.

Recall the notation of Section 3.6.

Definition 3.6.18. The moment map on $G/P$ is the map $\mu : G/P \to \mathfrak{t}_\mathbb{R}^*$ defined by

$$\mu(X) = \frac{\sum_{\alpha \in A} |p^\alpha(X)|^2 \alpha}{\sum_{\alpha \in A} |p^\alpha(X)|^2},$$

where

$$X = d \sum_{\alpha \in A} p^\alpha(X)e_\alpha.$$

Given $X \in G/P$, let $TX$ denote the orbit of $X$ under the action of $T$, and $\overline{TX}$ its closure.

Theorem 3.6.19 follows from classical work of Atiyah [Ati82] and Guillemin-Sternberg [GS82]. See also [GS87, Theorem 3.1].

Theorem 3.6.19. [GS87, Theorem 3.1] Let $X \in G/P$. The image $\mu(\overline{TX})$ is a convex polytope, and $\mu$ induces a one-to-one correspondence between the set of orbits of $T$ in $\overline{TX}$ and the set of faces of the polytope $\mu(\overline{TX})$, whereby a $q$-dimensional orbit of $T$ is mapped onto an open $q$-dimensional face of $\mu(\overline{TX})$.

Gelfand and Serganova [GS87] characterized the vertices of $\mu(\overline{TX})$.

Proposition 3.6.20. [GS87, Proposition 5.1] Let $X \in G/P$. Then the vertices of $\mu(\overline{TX})$ are the points $\alpha$ for all $\alpha \in L_X$.

3.7 Gelfand-Serganova strata, total positivity, and Bruhat interval polytopes for $G/P$

In this section we show that each totally positive cell of $(G/P)_{\geq 0}$ lies in a Gelfand-Serganova stratum, and we explicitly determine which one (i.e. we determine the list). We then define a Bruhat interval polytope for $G/P$, and show that each face of a Bruhat interval polytope is a Bruhat interval polytope. Our proof of this result on faces uses tools from total positivity. Allen Knutson has informed us that he has a different proof of this result about faces, using Frobenius splitting [Knu14].

\footnote{In the $G/B$ case, this result was conjectured in Rietsch’s thesis [Rie98]. Moreover Theorem 3.7.1 was partially proved in an unpublished manuscript of Marsh and Rietsch [MR05]. The second author is grateful to Robert Marsh and Konni Rietsch for generously sharing their ideas.}
CHAPTER 3. BRUHAT INTERVAL POLYTOPES

Gelfand-Serganova strata and total positivity

Write \( G/P = G/P_J \). Our goal is to prove the following theorem.

**Theorem 3.7.1.** Let \( u, v \in W \) with \( v \in W^J \) and \( u \leq v \). If \( X \in \mathcal{P}^J_{u,v;>0} \), then the list \( L_X \) of \( X \) is the set \( \{ z \cdot \rho_J \in W(\rho_J) \mid u \leq z \leq v \} \). In particular, \( \mathcal{P}^J_{u,v;>0} \) is entirely contained in one Gelfand-Serganova stratum.

Theorem 3.7.1 immediately implies the following.

**Corollary 3.7.2.** Suppose that whenever \((u, v) \neq (u', v')\) (where \((u, v)\) and \((u', v')\) index cells of \((G/P)_{\geq 0}\)), we have that \( \{ z \cdot \rho_J \in W(\rho_J) \mid u \leq z \leq v \} \neq \{ z \cdot \rho_J \in W(\rho_J) \mid u \leq z \leq v \} \). Then Rietsch’s cell decomposition of \((G/P)_{\geq 0}\) is the restriction of the Gelfand-Serganova stratification to \((G/P)_{\geq 0}\). In particular, her cell decompositions of the totally nonnegative parts of the complete flag variety \((G/B)_{\geq 0}\) and of the Grassmannian \((Gr_{k,n})_{\geq 0}\) are the restrictions of the Gelfand-Serganova stratification to \((G/B)_{\geq 0}\) and \((Gr_{k,n})_{\geq 0}\), respectively.

**Remark 3.7.3.** In general, it is possible for two distinct cells to lie in the same Gelfand-Serganova stratum. For example, let \( W = S_4 = S_{\{1,2,3,4\}} \) and \( W_J = S_{\{1\}} \times S_{\{2,3\}} \times S_{\{4\}} \). Let \((u, v) = (e, 4231)\) and \((u', v') = (1324, 4231)\). Then the minimal-length coset representatives in \( W/W_J \) of both \( \{ z \mid u \leq z \leq v \} \) and \( \{ z \mid u' \leq z \leq v' \} \) coincide and hence \( \{ z \cdot \rho_J \in W(\rho_J) \mid u \leq z \leq v \} = \{ z \cdot \rho_J \in W(\rho_J) \mid u \leq z \leq v \} \). It follows that the cells \( \mathcal{P}^J_{u,v;>0} \) and \( \mathcal{P}^J_{u',v';>0} \) both lie in the same Gelfand-Serganova stratum.

To prove Theorem 3.7.1, we will prove Proposition 3.7.4 and Proposition 3.7.6 below.

**Proposition 3.7.4.** Let \( u, v \in W \) with \( v \in W^J \) and \( u \leq v \). If \( X \in \mathcal{P}^J_{u,v;>0} \), then the list \( L_X \) is contained in \( \{ z \cdot \rho_J \in W(\rho_J) \mid u \leq z \leq v \} \).

Before proving Proposition 3.7.4, we need a lemma about the moment map image of positive torus orbits.

**Lemma 3.7.5.** Let \( X \in G/P \). Recall that \( T_{>0} \) denotes the positive part of the torus. Then \( \mu(TX) = \mu(T_{>0}X) \) and \( \mu(\overline{TX}) = \mu(T_{>0}X) \).

Proof. Recall that the torus \( T \) acts on the highest weight vector \( \eta_J \) of \( V_{\rho_J} \) by \( t \eta_J = \rho_J(t) \eta_J \) for all \( t \in T \). So the action of \( t \in T \) on \( X \in G/P \) will scale the Plücker coordinates of \( X \) by \( \rho_J(t) \).

Since the elements \( \alpha_j^\vee(\ell) \) for \( \ell \in \mathbb{C}^* \) generate \( T \), and we can write any \( \ell \in \mathbb{C}^* \) in the form \( re^{i\theta} \) with \( r \in \mathbb{R}_{>0} \) and \( \theta \in \mathbb{R} \), to prove the lemma, it suffices to show that for any positive \( r \) and real \( \theta \),

\[
\mu(\alpha_j^\vee(re^{i\theta})X) = \mu(\alpha_j^\vee(r)X).
\]  

(3.21)

First suppose that \( e^{i\theta} \) has finite order in the group of complex numbers of norm 1. Then \( \alpha_j^\vee(e^{i\theta}) \) has finite order, and hence \( |\rho_J(\alpha_j^\vee(e^{i\theta}))| = 1 \). But now within the group of unit complex numbers, the elements of finite order are dense. Therefore for any unit complex number \( e^{i\theta} \), we have \( |\rho_J(\alpha_j^\vee(e^{i\theta}))| = 1 \).

Now note that since \( \rho_J \) and \( \phi_j \) are homomorphisms, we have

\[
|\rho_J(\alpha_j^\vee(re^{i\theta}))| = |\rho_J(\alpha_j^\vee(r))\rho_J(\alpha_j^\vee(e^{i\theta}))| = |\rho_J(\alpha_j^\vee(r))|.
\]
And since the moment map depends only on the absolute value of the Plücker coordinates, it follows that (3.21) holds. □

We now turn to the proof of Proposition 3.7.4.

Proof. [Proof of Proposition 3.7.4] Consider \( z \cdot \rho_J \in L_X \). Since the extremal weight vectors are in bijection with cosets \( W/W_J \), we may assume that \( z \in W^J \). Proposition 3.6.20 implies that \( z \cdot \rho_J \) is a vertex of \( \mu(TX) \), and Lemma 3.7.5 therefore implies that \( z \cdot \rho_J \in \mu(T_{>0}X) \). Choose \( X' \in \overline{T_{>0}X} \) such that \( \mu(X') = z \cdot \rho_J \). Since \( z \cdot \rho_J \) is a vertex of \( \mu(TX) \), Theorem 3.6.19 implies that \( z \cdot \rho_J \) is the image of a torus fixed point of \( TX \). Therefore \( X' \) is a torus fixed point of \( G/P \) and must necessarily be the point \( \dot{z}P \).

By Lemma 3.6.13, \( T_{>0}X \subset \overline{P_{u,v}^J} \). Therefore \( X' = \dot{z}P \in \overline{P_{u,v}^J} \). It follows that \( \dot{z}P \) is a 0-cell in the closure of \( P_{u,v}^J \), so by Corollary 3.6.14, we must have \( u \leq z \leq v \). □

**Proposition 3.7.6.** Let \( u, v \in W \) with \( v \in W^J \) and \( u \leq v \). If \( X \in P_{u,v}^J \), then the list \( L_X \) contains \( \{ z \cdot \rho_J \in W(\eta_J) \mid u \leq z \leq v \} \).

To prove Proposition 3.7.6, we will need Proposition 3.7.7 below, which follows from [RW08, Lemma 6.1]. In fact the statements in [RW08, Section 6] used the \( \rho \)-representation \( V_\rho \) of \( G \), but the arguments apply unchanged when one uses \( V_{\rho_J} \) in place of \( V_\rho \).

**Proposition 3.7.7.** Consider \( G/P_J \) where \( G \) is simply laced. Let \( \eta_J \) be a highest weight vector of \( V_{\rho_J} \) and let \( B = B(\rho_J) \) be the canonical basis of \( V_{\rho_J} \) [Lus90] which contains \( \eta_J \). Consider the embedding

\[
G/P_J \hookrightarrow \mathbb{P}(V_{\rho_J})
\]

where

\[
gP_J \mapsto \langle g \cdot \eta_J \rangle.
\]

Then if \( gP_J \in (G/P_J)_{>0} \), the line \( \langle g \cdot \eta_J \rangle \), when expanded in \( B \), has non-negative coefficients. Moreover, the set of coefficients which are positive (respectively, zero) depends only on which cell of \( (G/P_J)_{>0} \) the element \( gP_J \) lies in.

We are now ready to prove Proposition 3.7.6. We will first prove it in the simply-laced case, using properties of the canonical basis, following Marsh and Rietsch, and then prove it in the general case, using folding.

Proof. [Proof of Proposition 3.7.6 when \( G \) is simply-laced.] Let \( z \in [u, v] \). We will use induction on \( \ell(v) \) to show that \( z \cdot \rho_J \) is in the list \( L_X \).

By Remark 3.6.11, we can write \( X = gP_J \) for \( g = g_1 \ldots g_n \in G_{u,v}^> \), where \( v = s_{i_1}\ldots s_{i_n} \) is a reduced expression of \( v \). Define \( v' = s_{i_2}\ldots s_{i_n} \), \( g' = g_2\ldots g_n \), and \( X' = g'P_J \). We need to consider two cases: that \( g_1 = y_{i_1}(p_1) \), and \( g_1 = s_{i_1} \).

In the first case, we have \( X' \in P_{u,v'}^J \). So the induction hypothesis implies that

\[
L_{X'} = \{ z \cdot \rho_J \mid u \leq z \leq v' \}.
\]

Here we must have \( L_{X'} \subset L_X \), since \( P_{u,v'}^J \subset \overline{P_{u,v}^J} \). So we are done if \( u \leq z \leq v' \). Otherwise, \( u \leq z \leq v \) but \( z \not\in [u, v'] \). Then any subexpression for \( z \) within \( v \) must use the \( s_{i_1} \), and so \( u \leq s_{i_1}z \leq v' \). And now by induction, we have \( s_{i_1}z \cdot \rho_J \in L_{X'} \).

By Proposition 3.7.7, the line \( \langle X' \cdot \eta_J \rangle = \langle g' \cdot \eta_J \rangle \) is spanned by a vector \( \xi \) which is a non-negative linear combination of canonical basis elements. Since \( s_{i_1}z \cdot \rho_J \in L_{X'} \), we have
that \( \xi = c s_i \cdot z \cdot \eta_J + \text{other terms} \), where \( c \) is positive. By Lemma 3.6.17, when we apply \( y_i \) to \( \xi \) we see that \( \langle X \cdot \eta_J \rangle = \langle c' z \cdot \eta_J + \text{other terms} \rangle \), where \( c' \neq 0 \). Therefore \( z \cdot \rho_J \in L_X \).

In the second case, we have that \( g_1 = s_i, \) so \( u' := s_i u \preceq u \) and \( v' := s_i v \preceq v \). By the induction hypothesis,

\[
L_{X'} = \{ z \cdot \rho_J \mid u' \preceq z \preceq v' \}.
\]

Consider again \( u \preceq z \preceq v \). Since the positive subexpression \( u_+ \) for \( u \) in \( v \) begins with \( s_i \), we must have \( u \not\preceq v' \). But then \( z \not\preceq v' \).

Now \( z \preceq v \) and \( z \not\preceq v' \) implies that any reduced expression for \( z \) in \( v \) must use the \( s_i \). So if we let \( z' := s_i z \), then \( u' \preceq z' \preceq v' \). Therefore by the induction hypothesis, \( z' \cdot \rho_J \in L_{X'}, \) i.e. \( \langle X' \cdot \eta_J \rangle = \langle c' z' \cdot \eta_J + \text{other terms} \rangle \) where \( c' \neq 0 \). But now \( \langle X \cdot \eta_J \rangle = \langle s_i X' \cdot \eta_J \rangle = \langle c s_i z' \cdot \eta_J + \text{other terms} \rangle = \langle c z \cdot \eta_J + \text{other terms} \rangle \). Therefore \( z \cdot \rho_J \in L_X \).

Before proving Proposition 3.7.6 in the general case, we give a brief overview of how one can view each \( \mathcal{G} \) which is not simply laced in terms of a simply laced group \( G \) by “folding.” For a detailed explanation of how folding works, see [Ste08].

If \( \mathcal{G} \) is not simply laced, then one can construct a simply laced group \( G \) and an automorphism \( \tau \) of \( G \) defined over \( \mathbb{R} \), such that there is an isomorphism, also defined over \( \mathbb{R} \), between \( \mathcal{G} \) and the fixed point subset \( G^\tau \) of \( G \). Moreover the groups \( \mathcal{G} \) and \( G \) have compatible pinnings. Explicitly we have the following.

Let \( G \) be simply connected and simply laced. Choose a pinning \((T, B^+, B^-, x_i, y_i, i \in I)\) of \( G \). Here \( I \) may be identified with the vertex set of the Dynkin diagram of \( G \). Let \( \sigma \) be a permutation of \( I \) preserving connected components of the Dynkin diagram, such that, if \( j \) and \( j' \) lie in the same orbit under \( \sigma \) then they are not connected by an edge. Then \( \sigma \) determines an automorphism \( \tau \) of \( G \) such that \( \tau(T) = T; \) and for all \( i \in I \) and \( m \in \mathbb{R} \), we have \( \tau(x_i(m)) = x_{\sigma(i)}(m) \) and \( \tau(y_i(m)) = y_{\sigma(i)}(m) \). In particular \( \tau \) also preserves \( B^+, B^- \). Let \( \overline{\sigma} \) denote the set of \( \sigma \)-orbits in \( I \), and for \( \overline{i} \in \overline{\sigma} \), let

\[
x_\overline{i}(m) := \prod_{i \in \overline{i}} x_i(m), \text{ and } y_\overline{i}(m) := \prod_{i \in \overline{i}} y_i(m).
\]

We also let \( s_\overline{i} := \prod_{i \in \overline{i}} s_i \), and \( \alpha_\overline{i} = \sum_{i \in \overline{i}} \alpha_i \), where \( \{\alpha_i \mid i \in I\} \) is the set of simple roots for \( G \).

Then the fixed point group \( G^\tau \) is a simply connected algebraic group with pinning \((T^\tau, B^{\tau+}, B^{\tau-}, x_\overline{i}, y_\overline{i}, \overline{i} \in \overline{\sigma})\). There exists, and we choose, \( G \) and \( \tau \) such that \( G^\tau \) is isomorphic to our group \( \mathcal{G} \) via an isomorphism compatible with the pinnings. The set \( \{\alpha_\overline{i} \mid \overline{i} \in \overline{\sigma}\} \) is the set of simple roots for \( \mathcal{G} \), and \( \overline{W} := \{s_\overline{i} \mid \overline{i} \in \overline{\sigma}\} \) is the Weyl group for \( \mathcal{G} \). Note that \( \overline{W} \subset W \), where \( W \) is the Weyl group for \( G \). Moreover, any reduced expression \( \overline{v} = (\overline{i}_1, \overline{i}_2, \ldots, \overline{i}_m) \) in \( \overline{W} \) gives rise to a reduced expression \( v \) in \( W \) of length \( \sum_{k=1}^m |\overline{i}_k| \), which is determined uniquely up to commuting elements [Nan05, Prop. 3.3]. To a subexpression \( \overline{u} \) of \( \overline{v} \) we can then associate a unique subexpression \( u \) of \( v \) in the obvious way.

Proof. [Proof of Proposition 3.7.6 in the general case.] Let \( \mathcal{G} \) be a group which is not simply laced and use all the notation above. We have that \( \mathcal{G} \) is isomorphic to \( G^\tau \) via an isomorphism compatible with the pinnings. Let \( \overline{P}_J \) be the parabolic subgroup of \( \mathcal{G} \) determined by the subset \( J \subset \overline{I} \). This gives rise to a subset \( J \subset I \) defined by

\[
J := \bigcup_{\overline{i} \in \overline{J}} \overline{i}.
\]
Now note that
\[
\rho_J = \sum_{i \in J} \omega_i = \sum_{i} \omega_i = \rho_J.
\]
Therefore the highest weight vector \( \eta_J \) of the \( G \)-representation \( V_{\rho_J} \) can also be viewed as a highest weight vector for the \( G \)-representation \( V_{\rho_J} \).

It follows from [RW08, Lemma 6.3] that we have an inclusion \( G^{\geq 0}_{\eta_J} \subset G_{\eta_J}^{\geq 0} \). Therefore the highest weight vector \( \eta_J \) is the sum of fundamental weights \( \sum_{j \in J} \omega_j \). Let \( u, v \in W \) with \( v \in W^J \) and \( u \leq v \). The Bruhat interval polytope \( Q^J_{u,v} \) for \( G/P \) is the convex hull
\[
\text{Conv} \{ z \cdot \rho_J \mid u \leq z \leq v \} \subset t^*_R.
\]

**Lemma 3.7.9.** For any \( X \in \mathcal{P}^J_{u,v} \), \( \mu(TX) = \text{int}(Q^J_{u,v}) \) and \( \mu(TX) = Q^J_{u,v} \), where \( \text{int} \) denotes the interior.

**Proof.** By Theorem 3.7.1, the list of \( X \) is precisely the set \( \{ z \cdot \rho_J \mid u \leq z \leq v \} \). And by Proposition 3.6.20, the vertices of \( \mu(TX) \) are precisely the elements of the list. Therefore \( \mu(TX) = Q^J_{u,v} \). Finally, Theorem 3.6.19 implies that \( \mu(TX) \) maps onto the interior of \( Q^J_{u,v} \).

**Proposition 3.7.10.** We have that \( \mu(\mathcal{P}^J_{u,v}) = \text{int}(Q^J_{u,v}) \) and \( \mu(\mathcal{P}^J_{u,v}) = Q^J_{u,v} \).

**Proof.** By Lemma 3.7.5 and Lemma 3.7.9, for each \( X \in \mathcal{P}^J_{u,v} \), we have \( \mu(T_{>0}X) = \text{int}(Q^J_{u,v}) \). Now using Lemma 3.6.13, we have that \( T_{>0}X \subset \mathcal{P}^J_{u,v} \). It follows that \( \mu(\mathcal{P}^J_{u,v}) = \text{int}(Q^J_{u,v}) \).

To prove the second statement of the proposition, note that since \( \mu \) is continuous, \( \mu(\mathcal{P}^J_{u,v}) \subset \mu(\mathcal{P}^J_{u,v}) \), and hence \( \mu(\mathcal{P}^J_{u,v}) \subset Q^J_{u,v} \). But now by Lemma 3.7.9, for any \( X \in \mathcal{P}^J_{u,v} \), \( \mu(TX) = Q^J_{u,v} \). Since \( TX \subset \mathcal{P}^J_{u,v} \), we obtain \( \mu(\mathcal{P}^J_{u,v}) = Q^J_{u,v} \).

**Remark 3.7.11.** Forgetting about total positivity, one can also consider the moment map images of Richardson varieties \( \mathcal{R}_{u,v} \) and projected Richardson varieties \( \mathcal{P}^J_{u,v} = \pi_J(\mathcal{R}_{u,v}) \), for \( u \leq v \). Using Proposition 3.7.10 and the fact that the torus fixed points of \( \mathcal{R}_{u,v} \) and \( \mathcal{R}_{u,v} \) agree, it follows that \( \mu(\mathcal{R}_{u,v}) = Q_{u,v} \). We similarly have that for \( u \leq v \) with \( v \in W^J \), \( \mu(\mathcal{P}^J_{u,v}) = Q^J_{u,v} \).
Proposition 3.7.12. The Richardson variety \( R_{u,v} \) is a toric variety if and only if \( \dim Q_{u,v} = \ell(v) - \ell(u) \). Similarly, if \( u \leq v \) and \( v \in W^J \), the projected Richardson variety \( P^J_{u,v} \) is a toric variety if and only if \( \dim Q^J_{u,v} = \ell(v) - \ell(u) \).

Proof. The Richardson variety \( R_{u,v} \) is a toric variety if and only if it contains a dense torus. By Theorem 3.6.19, \( R_{u,v} \) contains a dense torus if and only if \( \dim \mu(R_{u,v}) = \dim R_{u,v} \). By Remark 3.7.11, \( \mu(R_{u,v}) = Q_{u,v} \). Also, by [KL79], we have that \( \dim R_{u,v} = \ell(v) - \ell(u) \). Therefore \( R_{u,v} \) is a toric variety if and only if \( \dim Q_{u,v} = \ell(v) - \ell(u) \). The second statement of the proposition follows from the first, using the fact that the projection map \( \pi_J \) is an isomorphism from \( R_{u,v} \) to \( P^J_{u,v} \) whenever \( v \in W^J \). 

We are now ready to prove Theorem 3.7.13.

Theorem 3.7.13. The face of a Bruhat interval polytope for \( G/P \) is a Bruhat interval polytope for \( G/P \).

Proof. Consider a Bruhat interval polytope \( Q^J_{u,v} \) for \( G/P_J = G/P \). By Lemma 3.7.9, we can express \( Q^J_{u,v} = \mu(TX) \) for some \( X \in P^J_{u,v;>0} \).

Now let \( F \) be a face of \( Q^J_{u,v} \). By Theorem 3.6.19, the interior of \( F \) is the image of a \( T \)-orbit of some point \( Y \in TX \). By Lemma 3.7.5, the interior of \( F \) is therefore also the image of a \( T_{>0} \)-orbit of some point \( Y \in TX \). Therefore \( Y \) lies in some cell \( P^J_{a,b;>0} \) with \( a, b \in W, b \in W^J \), \( a \leq b \).

But then \( F = \mu(TY) \) for \( Y \in P^J_{a,b;>0} \) and hence by Lemma 3.7.9, \( F \) is a Bruhat interval polytope.

Corollary 3.7.14. Every edge of a Bruhat interval polytope corresponds to a cover relation in the (strong) Bruhat order.

Proof. Every edge is itself a face of the polytope, so by Theorem 3.7.13, it must come from an interval in Bruhat order. Since the elements of \( W(\rho_J) \) are the vertices of a polytope, none lies in the convex hull of any of the others. So a Bruhat interval polytope with precisely two vertices must come from a cover relation in Bruhat order.

Example 3.7.15. When \( G = \text{SL}_n \) and \( P = B \), a Bruhat interval polytope for \( G/P \) is precisely a Bruhat interval polytope as defined in Definition 3.2.2.

Example 3.7.16. When \( G = \text{SL}_n \) and \( P \) is a maximal parabolic subgroup, \( G/P \) is a Grassmannian, \( (G/P)_{>0} \) is the totally non-negative part of the Grassmannian, and the cells \( P^J_{u,v;>0} \) (for \( u \leq v, u \in S_n, \) and \( v \in W^J \)) are called positroid cells. In this case the moment map images of closures of torus orbits are a special family of matroid polytopes called positroid polytopes. These polytopes were studied in [ARW16]; in particular, it was shown there (by a different method) that a face of a positroid polytope is a positroid polytope.
Chapter 4

Tropical spectral theory of tensors

Our third problem concerns the combinatorial study of the spectral theory of tensors. Eigenpairs of tensors generalize the traditional spectral theory of matrices. They correspond to solutions of systems of multivariate polynomial equations. The tropical limit of these equations defines the tropical eigenpairs. Insights into the tropical limit, obtained by combinatorial means, translate to insights about the classical, continuous, spectral theory. This work is based on [Tsu15].

To be more precise, we will introduce and study tropical eigenpairs of tensors, a generalization of the tropical spectral theory of matrices. We will show the existence and uniqueness of an eigenvalue. We will associate to a tensor a directed hypergraph and define a new type of cycle on such a hypergraph, which we call an H-cycle. The eigenvalue of a tensor will turn out to be equal to the minimal normalized weighted length of H-cycles of the associated hypergraph. We will show that the eigenvalue can be computed efficiently via a linear program.

4.1 Background

A tensor of order $m$ and rank $n$ is an array $A = (a_{i_1...i_m})$ of elements of a field or semifield $K$ (which we shall take to be $\mathbb{R}$ or $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$), where $1 \leq i_1, \ldots, i_m \leq n$. In ordinary arithmetic, given $x \in \mathbb{R}^n$, we define

$$(Ax^{m-1})_i := \sum_{i_2, \ldots, i_m=1}^n a_{ii_2...i_m}x_{i_2} \cdots x_{i_m}.$$  

Here $A$ is being contracted with $x^{\otimes(m-1)}$, the $(m-1)$st tensor power of $x$. An H-eigenpair [Qi05] of a tensor is defined as follows. Define $x^{[m-1]} = (x^{m-1}_i)$. Then an H-eigenpair is a pair $(x, \lambda) \in \mathbb{P}^{n-1} \times \mathbb{R}$ such that

$$Ax^{m-1} = \lambda x^{[m-1]}.$$  

Let $A$ be a $n \times n$ matrix with entries in the tropical semiring $(\mathbb{R}, \oplus, \odot)$. We shall take $\oplus$ to be min throughout. An eigenvalue of $A$ is a number $\lambda$ such that

$$A \odot \mathbf{v} = \lambda \odot \mathbf{v}.$$
The nature of tropical eigenpairs is understood in the setting of matrices ([ST13],[Tra14]) but a survey of the literature shows no prior research on tropical eigenpairs of tensors.

**Definition 4.1.1.** A tropical H-eigenpair for a tensor \((a_{i_1\ldots i_m}) \in \mathbb{R}^{n^m}\) of order \(m\) and rank \(n\) is a pair \((x, \lambda) \in \mathbb{R}^n/\mathbb{R}(1,1,\ldots,1) \times \mathbb{R}\) such that

\[
\bigoplus_{i_2,\ldots,i_m=1}^n a_{i_2\ldots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m} = \lambda \odot x_{i}^{(m-1)}, \quad i = 1, 2, \ldots, n.
\]  

(4.1)

We call \(x\) a tropical H-eigenvector and \(\lambda\) a tropical H-eigenvalue.

In the classical setting, several other definitions of eigenpairs of tensors exist. For instance, an E-eigenpair is defined via the condition

\[Ax^{m-1} = \lambda x.\]

We define the tropicalization here in an analogous manner. In this paper, we focus on H-eigenpairs and only discuss E-eigenpairs for purposes of comparison. Note that the notions of E- and H-eigenpairs coincide with the usual notion of eigenpairs of matrices in the case of matrices.

**Example 4.1.2.** Take \(n = 2\) and \(m = 3\). Then a tropical H-eigenpair \((x, \lambda)\) satisfies

\[
\begin{align*}
\min\{a_{111} + 2x_1, a_{112} + x_1 + x_2, a_{121} + x_1 + x_2, a_{122} + 2x_2\} &= \lambda + 2x_1 \\
\min\{a_{211} + 2x_1, a_{212} + x_1 + x_2, a_{221} + x_1 + x_2, a_{222} + 2x_2\} &= \lambda + 2x_2.
\end{align*}
\]

Define a partial symmetrization map \(\text{PSym}\) on tensors via

\[
(\text{PSym} A)_{i_1\ldots i_n} := \min_{\sigma \in S\{2,\ldots,m\}} A_{i_1,i_{\sigma(2)}\ldots,i_{\sigma(n)}},
\]

where \(S\{2,\ldots,m\}\) is the set of permutations of \(\{2,\ldots,m\}\). Let \(a'_{i_1\ldots i_n} := (\text{PSym} A)_{i_1\ldots i_n}\).

Observe that

\[
\bigoplus_{i_2,\ldots,i_m=1}^n a_{i_2\ldots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m} = \bigoplus_{i_2,\ldots,i_m=1}^n a'_{i_2\ldots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m},
\]

so that \(\text{PSym} A\) has the same H- (and E-) eigenpairs as \(A\). Therefore, without loss of generality, we will assume from now on that all tensors are symmetric in their last \(m-1\) coordinates:

\[
a_{i_1i_2\ldots i_m} = a_{i_1i_{\sigma(2)}\ldots i_{\sigma(m)}}
\]

for all permutations \(\sigma\) of \(\{2,\ldots,m\}\).

**Example 4.1.3.** Take \(n = 3\) and \(m = 3\). Then a tropical H-eigenpair \((x, \lambda)\) satisfies

\[
\begin{align*}
\min\{a_{111} + 2x_1, a_{112} + x_1 + x_2, a_{113} + x_1 + x_3, a_{122} + 2x_2, a_{123} + x_2 + x_3, a_{133} + 2x_3\} &= \lambda + 2x_1 \\
\min\{a_{211} + 2x_1, a_{212} + x_1 + x_2, a_{213} + x_1 + x_3, a_{222} + 2x_2, a_{223} + x_2 + x_3, a_{233} + 2x_3\} &= \lambda + 2x_2 \\
\min\{a_{311} + 2x_1, a_{312} + x_1 + x_2, a_{313} + x_1 + x_3, a_{322} + 2x_2, a_{323} + x_2 + x_3, a_{333} + 2x_3\} &= \lambda + 2x_3.
\end{align*}
\]
4.2 Main Results

We show that

**Theorem 4.2.1.** A tensor $A \in \mathbb{R}^{n \times m}$ has a unique tropical H-eigenvalue $\lambda(A) \in \mathbb{R}$.

The result is all the more striking when compared with the situation of E-eigenpairs. By the Transverse Intersection Lemma of Tropical Geometry [MSon, Theorem 3.4.12], for generic tensors, tropical E-eigenpairs lift to classical eigenpairs. By [CS13], there are \((m-1)^{n-1} \frac{m-2}{m-1} \) classical E-eigenpairs for a tensor of order \(m\) and rank \(n\). Consequently, there can be at most \((m-1)^{n-1} \frac{m-2}{m-1} \) tropical E-eigenpairs.

Experimentally, in 5000 runs of randomly generated symmetric $3 \times 3 \times 3$-tensors, we obtained the following distribution on the number of tropical E-eigenpairs (from 0 eigenpairs to 7): [0, 4007, 7, 950, 0, 6, 1, 29].

Uniqueness and existence for tropical H-eigenpairs can be extended to the case when $A \in \mathbb{R}^{n \times m}$ under suitable technical assumptions:

**Theorem 4.2.2.** Let $A \in \mathbb{R}^{n \times m}$ and assume that for each $i = 1, 2, \ldots, n$, the sets

$$S_i := \{\{i_2^{(i)}, \ldots, i_m^{(i)}\} : a_{i_2^{(i)} \cdots i_m^{(i)}} \neq \infty\}$$

are nonempty and mutually equal. Then $A$ has a unique tropical H-eigenvalue $\lambda(A) \in \mathbb{R}$.

Surprisingly, it turns out that the H-eigenvalue is a solution to a much simpler problem given by a linear program which can be interpreted as merely requiring that $\lambda(A)$ be a subeigenvalue. This is a consequence of a special property of the circulation polytope (to be defined later).

**Theorem 4.2.3.** Assume the hypotheses of Theorem 4.2.2. The solution to the linear program

$$\begin{align*}
\text{maximize} & \quad \lambda \\
\text{subject to} & \quad a_{i_1 i_2 \cdots i_m} + x_{i_2} + \ldots + x_{i_m} \geq \lambda + (m-1)x_{i_1}, \quad \forall (i_1, i_2, \ldots, i_m) \in [n]^m
\end{align*}$$

is equal to the H-eigenvalue of $A$. Dually, the H-eigenvalue of $A$ is given by

$$\begin{align*}
\text{minimize} & \quad \sum_{(i_1, i_2, \ldots, i_m) \in [n]^m} a_{i_1 i_2 \cdots i_m} y_{i_1 i_2 \cdots i_m} \\
\text{subject to} & \quad \sum_{(i_1, i_2, \ldots, i_m) \in [n]^m} y_{i_1 i_2 \cdots i_m} ((m-1)e_{i_1} - e_{i_2} - \ldots - e_{i_m}) = 0 \\
& \quad \sum_{(i_1, i_2, \ldots, i_m) \in [n]^m} y_{i_1 i_2 \cdots i_m} = 1 \\
& \quad y_{i_1 i_2 \cdots i_m} \geq 0.
\end{align*}$$

For matrices ($m = 2$), the dual problem has the interpretation of giving the minimal normalized length of any directed cycle in the weighted directed graph associated to the matrix $A$. Motivated by this, we associate to a tensor $A$ a weighted directed hypergraph.

We summarize the relevant notions concerning directed hypergraphs (for more on directed hypergraphs, see [AFF01, GLPN93]). For us, a directed hypergraph $H$ will be a pair $(V, E)$ where $V$ is a set of nodes and $E$, the set of hyperedges, consists of pairs of multiset of nodes. An F-hyperedge is a hyperedge $e = (T(e), H(E))$ such that $|T(e)| = 1$. Here $T$ stands for tail and $H$ for head. An F-hypergraph is a directed hypergraph whose edges are F-hyperedges.
CHAPTER 4. TROPICAL SPECTRAL THEORY OF TENSORS

57

Definition 4.2.4. Given a tensor $A$ of order $m$ and rank $n$, we associate to it a weighted $F$-hypergraph whose vertices are $V(A) := [n]$, hyperarcs are $E(A) := \{(i_1, \{i_2, i_3, \ldots, i_m\}) : i_j \in [n] \forall j = 1, 2, \ldots, m\}$ and weights are given by $W(A)((i_1, \{i_2, \ldots, i_m\})) = a_{i_1i_2\ldots i_m}$.

Definition 4.2.5. Let $H = (V, E)$ be an $F$-hypergraph. A circulation on $H$ is a function $f : E \to \mathbb{R}$ such that

\[
\sum_{e=(v_1,\{v_2,\ldots,v_m\})\in E} f(e)[e_{i_2} + \ldots + e_{i_m} - (m - 1)e_{i_1}] = 0. \tag{4.4}
\]

We can interpret the assignment $f(e)$ as having $f(e)$ units flowing out of $v_{i_1}$ to each of $v_{i_2}, \ldots, v_{i_m}$. Equation (4.4) then implies that for every vertex the amount of flow going out and the amount of flow coming in are equal.

Recall that a tight cycle in a hypergraph is a sequence of edges of the form

\[
(v_1, \{v_2, \ldots, v_k\}), (v_2, \{v_3, \ldots, v_{k+1}\}), \ldots, (v_{r-k+1}, \{v_{r-k+2}, \ldots, v_r\}), (v_{r-k+2}, \{v_{r-k+3}, \ldots, v_r, v_1\}), \ldots, (v_r, \{v_1, \ldots, v_{k-1}\}),
\]

with the vertices not necessarily distinct. Any tight cycle induces a circulation. Indeed, consider the function assigning 1 to a hyperedge of the tight cycle and zero to all other hyperedges. Viewing indices modulo $r$,

\[
\sum_{j=1}^{r} [e_{v_{j+1}} + \ldots + e_{v_{j+m-1}} - (m-1)e_{v_j}] = (e_{v_1} + \ldots + e_{v_r}) + \ldots + (e_{v_1} + \ldots + e_{v_r}) - (m-1)(e_{v_1} + \ldots + e_{v_r}) = 0.
\]

We introduce and study the circulation polytope, which is the feasible set of the linear program (4.3).

Definition 4.2.6. The circulation polytope $H_{n,m}$ is the polytope in $\mathbb{R}^{n^m}$ defined by the inequalities

\[
\begin{align*}
\sum_{(i_1,i_2,\ldots,i_m)\in [n]^m} y_{i_1i_2\ldots i_m}((m-1)e_{i_1} - e_{i_2} - \ldots - e_{i_m}) &= 0 \\
\sum_{(i_1,i_2,\ldots,i_m)\in [n]^m} y_{i_1i_2\ldots i_m} &= 1 \\
y_{i_1i_2\ldots i_m} &\geq 0.
\end{align*}
\]

The $H$-cycles are the vertices of $H_{n,m}$.

In this language, $\lambda(A)$ is equal to the minimal normalized weighted length of any $H$-cycle of $H(A)$.

The circulation polytope is a generalization of the normalized cycle polytope [Tra13] but turns out to be more complicated for $m > 2$. For instance, the vertices are no longer normalized characteristic functions of subsets of the edges (see Example 4.5.1). Nevertheless, we show that it has the following nice property, which explains Theorem 4.2.3.

Theorem 4.2.7. Let $y \in \mathbb{R}^{n^m}$ be a vertex of the circulation polytope $H_{n,m}$. Then $y$ has at most one nonzero entry of the form $y_{j_1\ldots j_n}$ for each $j = 1, 2, \ldots, n$. There exist vertices with $n$ nonzero entries.
CHAPTER 4. TROPICAL SPECTRAL THEORY OF TENSORS

4.3 H-eigenvalues of Tensors

Our setup is that of Theorem 4.2.2:

Restatement of Theorem 4.2.2. Let $A \in \mathbb{R}^{n \times m}$ and assume that for each $i = 1, 2, \ldots, n$, the sets

$$S_i := \{ \{i_2^{(i)}, \ldots, i_m^{(i)}\} : a_{i_2^{(i)} \cdots i_m^{(i)}} \neq \infty \}$$

are nonempty and mutually equal. Then $A$ has a unique tropical H-eigenvalue $\lambda(A) \in \mathbb{R}$.

We begin by proving uniqueness of the H-eigenvalue. The main tool for the proof is to use Gordan’s Theorem [BL06, Theorem 2.2.1] which states that for a matrix $M$

either $\exists x \in \mathbb{R}^n_+ \setminus \{0\}$ such that $Mx = 0$, or $\exists y \in \mathbb{R}^n$ such that $M^t y > 0$.

The main idea in the proof is to perform Gaussian elimination using only additions of positive scalar multiples of rows.

Proposition 4.3.1. Let $\lambda$ be a tropical H-eigenvalue of the tensor $A = (a_{i_1 \cdots i_m})$. For all $(i_1, i_2, \ldots, i_m) \in [n]^m$, define the column vector

$$v_{i_1 \cdots i_m} := -(m - 1)e_{i_1} + e_{i_2} + \ldots + e_{i_m}$$

and let $M$ be the matrix whose columns are $v_{i_1 \cdots i_m}$. Then

$$\lambda = \begin{bmatrix} \minimize \\
\text{subject to} \\
Mc = 0 \\
1^t c = 1 \\
c \geq 0. \end{bmatrix}$$

(4.5)

Proof. Let $x = (x_1, \ldots, x_n)^t$ and $\lambda$ be an H-eigenpair, so that

$$\lambda = a_{i_2^{(i)} \cdots i_m^{(i)}} + x^t v_{i_2^{(i)} \cdots i_m^{(i)}}, \quad i = 1, 2, \ldots, n,$$

for $a_{i_2^{(i)} \cdots i_m^{(i)}} \in \mathbb{R}$. We show that there exists $c \geq 0, c \neq 0$, such that

$$\sum_{i=1}^n c_i v_{i_2^{(i)} \cdots i_m^{(i)}} = 0. \quad (4.6)$$

Form the matrix $W$ whose columns are $v_{i_2^{(i)} \cdots i_m^{(i)}}$. We would like to show that there exists $c \in \mathbb{R}^n \setminus \{0\}$ such that $Wc = 0$. Applying Gordan’s Theorem, we will show that the alternative $\exists y \in \mathbb{R}^n$ such that $W^t y > 0$ leads to a contradiction. We do so as follows. If $W^t$ has a zero row, then we are done. So assume otherwise. We show that we can bring the matrix $W^t$ to row echelon form using the operation of adding to a row a positive scalar multiple of another row. This preserves the positivity condition. Since $W^t$ is not of full rank (each row sum is zero), this will prove the claim.

The assumption that $W^t$ has no zero rows implies that an entry of $W^t$ is negative if and only if it lies along the diagonal. Let $r_i = (r_{i1}, r_{i2}, \ldots, r_{in}), i = 1, 2, \ldots, n$ denote the rows
of \( W^t \). We may add positive scalar multiples of \( r_1 \) to all the remaining rows so as to zero out \( r_{i1} \) for each \( i \geq 2 \). In doing so, either: the \((2,2)\) entry of \( W^t \) remains negative, or the second row of \( W^t \) is now zero. To see why this is so, note that the sum of entries of \( r_2 \) must be zero, while the entries \((2,3), \ldots, (2,n)\) are nonnegative. We proceed in this manner, adding a positive multiple of \( r_i \) to \( r_j \), \( j \geq i + 1 \), until \( W^t \) has reached row echelon form. The existence of \( c \geq 0, c \neq 0 \) satisfying (4.6) now follows.

We then have

\[
\sum_{i=1}^{n} c_i \lambda = \sum_{i=1}^{n} c_i a_{i_1(\cdots)_{i_m}} + \sum_{i=1}^{n} c_i x^t v_{i_2(\cdots)_{i_m}} = \sum_{i=1}^{n} c_i a_{i_2(\cdots)_{i_m}},
\]

so that

\[
\lambda \geq \left[ \begin{array}{c} \text{minimize} \sum_{i_1 \cdots i_m} a_{i_1 \cdots i_m} c_{i_1 \cdots i_m} \\ \text{subject to} \\ Mc = 0 \\ 1^t c = 1 \\ c \geq 0 \end{array} \right].
\]

Conversely, from (4.1), we have

\[
\lambda \leq a_{i_1 \cdots i_m} + x^t v_{i_1 \cdots i_m}
\]

for every choice of indices. Suppose that \( c \in \ker M, c \geq 0 \) and \( 1^t c = 1 \). Then, in particular,

\[
\sum_{i_1 \cdots i_m} c_{i_1 \cdots i_m} v_{i_1 \cdots i_m} = 0.
\]

Since \( c \geq 0 \),

\[
(\sum_{i_1 \cdots i_m} c_{i_1 \cdots i_m}) \lambda(A) \leq \sum_{i_1 \cdots i_m} c_{i_1 \cdots i_m} a_{i_1 \cdots i_m} + \sum_{i_1 \cdots i_m} c_{i_1 \cdots i_m} x^t v_{i_1 \cdots i_m} = \sum_{i_1 \cdots i_m} c_{i_1 \cdots i_m} a_{i_1 \cdots i_m},
\]

It follows that

\[
\lambda \leq \min_{c \in \ker M, c \geq 0, 1^t c = 1} \sum_{i_1 \cdots i_m} c_{i_1 \cdots i_m} a_{i_1 \cdots i_m}.
\]

Next we show the existence of an H-eigenpair. The idea of the proof is to consider a tropical analogue of the proof of the Perron-Frobenius Theorem. For a reference on the Perron-Frobenius Theory of tensors, see [CPZ08].

For the proof of existence, we will be needing the following lemma.

**Lemma 4.3.2.** Let \( c_i, d_i, A_i \in \mathbb{R} \) for \( 1 \leq i \leq k \). Then

\[
\min_i \{c_i - d_i\} \leq \min_i \{c_i + A_i\} - \min_i \{d_i + A_i\} \leq \max_i \{c_i - d_i\}.
\]

Proof. Suppose that

\[
\min_k \{c_k + A_k\} = c_i + A_i, \quad \min_k \{d_k + A_k\} = d_j + A_j.
\]

\( \square \)
By (4.7), $c_i + A_i \leq c_j + A_j$ and $d_j + A_j \leq d_i + A_i$. Therefore

$$\min_i \{c_i + A_i\} - \min_i \{d_i + A_i\} = c_i - d_j + A_i - A_j \leq c_j - d_j,$$

and similarly $c_i - d_j + A_i - A_j \geq c_i - d_i$. \hfill \square

**Proposition 4.3.3.** For any tensor $A$ satisfying the hypotheses of Theorem 4.2.2, there exists a tropical $H$-eigenpair.

Proof. Let $S := S_i$. Define $F : \mathbb{T}^{n-1} \to \mathbb{T}^{n-1}$ by

$$F(x)_i = \frac{\bigoplus_{\{i_2, \ldots, i_m\} \in S} a_{i_1 i_2 \cdots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m}}{m-1}.$$ 

Using the equivalence relation on $\mathbb{T}^{n-1}$, we can view this map as a map $F : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ by

$$F(x_1, x_2, \ldots, x_{n-1})_i = \frac{\bigoplus_{\{i_2, \ldots, i_m\} \in S} a_{i_1 i_2 \cdots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m} - \bigoplus_{\{i_2, \ldots, i_m\} \in S} a_{ni_1 i_2 \cdots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m}}{m-1}$$

for $1 \leq i \leq n - 1$. By Lemma 4.3.2, each coordinate of $F$ is bounded. Thus $F$ is a continuous mapping of a convex set of $\mathbb{R}^{n-1}$ into a bounded closed subset of $\mathbb{R}^{n-1}$, and consequently has a fixed point. This condition translates to having, for each $1 \leq i \leq n - 1$,

$$\bigoplus_{\{i_2, \ldots, i_m\} \in S} a_{i_1 i_2 \cdots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m} = \bigoplus_{\{i_2, \ldots, i_m\} \in S} a_{ni_1 i_2 \cdots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m} + (m-1)x_i,$$

i.e., the existence of an $H$-eigenpair (here $x_n$ is normalized to zero). \hfill \square

**Remark 4.3.4.** The proof of existence can be easily adopted to tropical $E$-eigenpairs; one simply replaces the function

$$F(x)_i = \frac{\bigoplus_{\{i_2, \ldots, i_m\} \in S} a_{i_1 i_2 \cdots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m}}{m-1}$$

with

$$F(x)_i = \bigoplus_{\{i_2, \ldots, i_m\} \in S} a_{i_1 i_2 \cdots i_m} \odot x_{i_2} \odot \cdots \odot x_{i_m}.$$ 

Next we show that the $H$-eigenvalue of a tensor can be computed via a linear program.

**Restatement of Theorem 4.2.3.** Assume the hypotheses of Theorem 4.2.2. The solution to the linear program

$$\text{maximize } \lambda$$

$$\text{subject to } a_{i_1 i_2 \cdots i_m} + x_{i_2} + \ldots + x_{i_m} \geq \lambda + (m-1)x_{i_1}, \forall (i_1, i_2, \ldots, i_m) \in [n]^m$$

(4.8)
is equal to the H-eigenvalue of \( A \). Dually, the H-eigenvalue of \( A \) is given by

\[
\text{minimize } \sum_{i_1,i_2,\ldots,i_m \in [n]^m} a_{i_1i_2\ldots i_m} y_{i_1i_2\ldots i_m}
\]

subject to

\[
\sum_{i_1,i_2,\ldots,i_m \in [n]^m} \lambda (m - 1) e_{i_1} - e_{i_2} - \ldots - e_{i_m} = 0 \\
\sum_{i_1,i_2,\ldots,i_m \in [n]^m} y_{i_1i_2\ldots i_m} = 1 \\
y_{i_1i_2\ldots i_m} \geq 0.
\]

(4.9)

Proof. The dual problem to (4.8) is (4.9). The primal problem is feasible, since we can take \( x_i = 0 \forall i \in [n] \) and \( \lambda = \min_{i_1,i_2,\ldots,i_m \in [n]^m} a_{i_1i_2\ldots i_m} \). The problems (4.5) and (4.9) are identical, hence feasible, and so the result follows by Proposition 4.3.1. \( \Box \)

### 4.4 H-eigenvalues of Symmetric Tensors

In this section, we specialize the results of section 4.3 to symmetric tensors.

**Proposition 4.4.1.** For a symmetric tensor \( A \in \mathbb{R}^{n \times n} \),

\[
\lambda(A) = \min_{i_1i_2\ldots i_m} a_{i_1i_2\ldots i_m}.
\]

**Proof.**

Consider the system (4.8) for an H-eigenpair. Set \( \Delta_{i,j} = x_i - x_j \). In this notation, we have

\[
a_{i_1i_2\ldots i_m} + \Delta_{i_2,i_1} + \Delta_{i_3,i_1} + \ldots + \Delta_{i_m,i_1} \geq \lambda(A) \quad \forall i_1,\ldots,i_m \in [n].
\]

In particular, for every \( \sigma \in S_m \), we have

\[
\lambda(A) \leq a_{i_{\sigma(1)}i_{\sigma(2)}\ldots i_{\sigma(m)}} + \Delta_{i_{\sigma(2)},i_{\sigma(1)}} + \Delta_{i_{\sigma(3)},i_{\sigma(1)}} + \ldots + \Delta_{i_{\sigma(m)},i_{\sigma(1)}}.
\]

Summing such inequalities over \( S_m \) and using the symmetry of \( A \) shows that \( \lambda(A) \leq \min_{i_1i_2\ldots i_m} a_{i_1i_2\ldots i_m} \). Taking \( x_i = 0 \forall i = 1,2,\ldots,n \) in (4.8) and \( \lambda = \min_{i_1i_2\ldots i_m} a_{i_1i_2\ldots i_m} \) yields the result. \( \Box \)

**Proposition 4.4.2.** Suppose that \( A \) is a symmetric tensor whose minimum entries \( a_{i_1i_2\ldots i_m} \) all share the same set of indices \( I = \{i_1,i_2,\ldots,i_m\} \). If \( x \) is an eigenvector of \( A \), then

1. \( x_i = x_j \) for each \( i,j \in I \).
2. \( x_i \leq x_j \) for each \( i \in I \) and \( j \notin I \).

**Proof.** Let \( x_k = \min_i x_i \). For some choice of indices \( k_2,\ldots,k_m \),

\[
a_{k_2k_3\ldots k_m} + \Delta_{k_2,k} + \Delta_{k_3,k} + \ldots + \Delta_{k_m,k} = \lambda.
\]

By choice of \( k \), \( \Delta_{k,l} \geq 0 \) for \( l = 2,3,\ldots,m \). By Proposition 4.4.1, we must have \( \Delta_{k,l} = 0 \) for each \( l \). Since \( \lambda(A) \) is equal to the minimum entry of \( A \), \( a_{kk_2\ldots k_m} \) is a minimum element. Since a minimum element \( a_{i_1\ldots j_m} \) has indices \( I, \{k,k_2,\ldots,k_m\} = I \). This shows (1). Part (2) follows from \( x_k \) being minimal. \( \Box \)
4.5 circulation polytope

In view of (4.9), we can interpret tensors as linear functions, and their minima on the circulation polytope as the H-eigenvalues. The following example shows that, unlike the vertices of the normalized cycle polytope ($m = 2$), the vertices of the circulation polytope are no longer normalized characteristic functions of subsets of the edges.

**Example 4.5.1.** Consider the tensor whose entries are all zero except for $a_{132} = a_{213} = a_{322} = -1$. Taking

$$y_{132} = \frac{2}{9}, \quad y_{213} = \frac{4}{9}, \quad y_{322} = \frac{3}{9}, \quad y_{ijk} = 0 \text{ for all remaining indices}$$

we see that the dual problem (4.9) attains a value of $-1$. The primal problem (4.8) also has a feasible point attaining the value $-1$: we take $\lambda = -1$ and $x_i = 0$ for each $i$. It is easy to check that (4.10) is a vertex of the circulation polytope.

Next we give a proof of Theorem 4.2.7.

**Proof.** (of Theorem 4.2.7) Let $y$ be a vertex of $H_{n,m}$. Take a linear functional $A = (a_{i_1i_2...i_m})$ whose restriction to $H_{n,m}$ is minimized on, and only on, $y$. This minimum value is equal to $\lambda^*$, the solution to the linear program (4.9). By Theorem 4.2.3, the value $\lambda(A)$ is attained by a tropical H-eigenvector of the tensor $A$. From the proof of Proposition 4.3.1, we see that the tropical H-eigenvector $\lambda$ can be written as

$$\lambda = \sum_{i=1}^{n} c_i a_{i_1(i)...i_m}, \quad c \in H_{n,m}.$$  

It follows that $y_{i_1(i)...i_m} = c_i$ and the remaining entries of $y$ are zero.

Next we show that a vertex with $n$ nonzero entries exists. Consider the (one-sided) infinite sequence

$$s = \{1, 2, \ldots, n, 1, 2, \ldots, n, 1, \ldots\}.$$  

We take $y$ to have an entry of $\frac{1}{n}$ at position $(s_j, s_{j+1}, \ldots, s_{j+m})$ for $j = 1, 2, \ldots, n$, and zero otherwise. We also take the linear functional $A$ to have entries of $-1$ at these positions and zeros otherwise. To see that $y \in H_{n,m}$, write $s_j = j - n [j-1]$. Then

$$\sum_{(i_1, i_2, \ldots, i_m) \in [n]^m} y_{i_1i_2...i_m}((m-1)e_{i_1} - e_{i_2} - \ldots - e_{i_m})$$

$$= \sum_{j=1}^{n} \frac{1}{n} ((m-1)e_{j-n[j-1/n]} - e_{j+1-n[j-1/n]} - \ldots - e_{j+m-n[j+m-1/n]})$$

$$= \frac{1}{n} ((m-1)(e_1 + \ldots + e_n) - (e_1 + \ldots + e_n) - \ldots - (e_1 + \ldots + e_n)) = 0.$$  

The dual problem (4.9) takes on the value $-1$. The primal problem (4.8) has a solution $\lambda = -1$, $x_i = 0 \forall i$. This shows that $y$ lies on a face of $H_{n,m}$. To see that it is a vertex, note that by the rearrangement inequality, any other minimizer must have support contained
in the support of $y$. We show that the matrix $W$ whose columns are
\[(m-1)e_{j-n\lfloor j-1\rfloor} - e_{j+1-n\lfloor j\rfloor} - \ldots - e_{j+m-n\lfloor j+m-1\rfloor}, j = 1, 2, \ldots, n,\]
has rank $n-1$. This implies that the kernel is 1-dimensional, so that $y$ is indeed a vertex. Write $m = an + r, 1 \leq r \leq n$. The matrix $W$ is the circulant matrix whose first row is
\[a - (m-1)e_1 + (a+1)e_2 + \ldots + (a+1)e_r + ae_{r+1} + \ldots + ae_n.\]
The associated polynomial is
\[f(x) = a - (m-1) + (a+1)x + \ldots + (a+1)x^{r-1} + ax^r + \ldots + ax^n = a - (m-1) + x + \ldots + x^{r-1} + a(x + \ldots + x^{n-1}).\]
It is known that the rank of a circulant matrix is equal to $n - d$, where $d$ is the degree of $\gcd(f(x), x^n - 1)$ [Ing56]. We show that the only common root to both polynomials is $x = 1$. We have
\[f(1) = a - (m-1) + r - 1 + a(n-1) = 0.\]
Let $\xi \neq 1$ be an $n$th root of unity. Then
\[f(\xi) = -m + 1 + \xi + \ldots + \xi^{r-1} = \frac{1 - \xi^r}{1 - \xi} - m.\]
We see that $f(\xi) \neq 0$ for $m \geq 2$. □

**Corollary 4.5.2.** Let $s_j = j - n\lfloor \frac{j-1}{n} \rfloor$. The $H$-cycle whose edges are
\[\{(s_1, s_2, \ldots, s_{1+m}), (s_2, s_3, \ldots, s_{2+m}), \ldots, (s_n, s_{n+1}, \ldots, s_{n+m})\}\]
isan vertex of the circulation polytope $H_{n,m}$.

**Proof.** This is a consequence of the proof of Theorem 4.2.7. □

### 4.6 Future Directions

For non-generic tensors, there may be infinitely many nonequivalent tropical $H$-eigenvectors. In contrast, for generic matrices, there exists a unique tropical $H$-eigenpair [ST13].

**Question 4.6.1.** What is the number and structure of nonequivalent tropical $H$-eigenpairs of a tensor?

Another direction of study is properties of the circulation polytope.

**Question 4.6.2.** Describe the facial structure of the circulation polytope.

Finally, it is of great interest to understand the relationship between the classical $H$-eigenpairs of a tensor and the tropical ones. For example, for a positive matrix $X$ [ABG98],
\[\lim_{k \to \infty} \rho(X^{(k)})^\frac{1}{k} = \lambda(X).\]
Here $X^{(k)}$ denotes the $k$th Hadamard power of $X$, $\rho(X^{(k)})$ is the Perron-Frobenius eigenvalue of $X^{(k)}$ and $\lambda(X)$ is the tropical eigenvalue of $X$. Work on classical versions of Perron-Frobenius for tensors (e.g., [CPZ08]) suggests a similar result for positive tensors.
Chapter 5

Conclusion

We have seen three ways in which combinatorial methods can be applied to obtain information about a range of continuous objects. These examples suggest that the field of combinatorics has matured greatly since the days of “balls and urns”, to an extent that it may be applied to obtain deep information about diverse fields of mathematics.
Bibliography


BIBLIOGRAPHY


