2014

Fast Algorithms for Analyzing Partially Ranked Data

Matthew McDermott
Harvey Mudd College

Fast Algorithms for Analyzing Partially Ranked Data

Matthew McDermott

Michael Orrison, Advisor

Mohamed Omar, Reader

HARVEY MUDD COLLEGE

Department of Mathematics

May, 2014

Abstract

Imagine your local creamery administers a survey asking their patrons to choose their five favorite ice cream flavors. Any data collected by this survey would be an example of *partially ranked data*, as the set of all possible flavors is only ranked into subsets of the chosen flavors and the non-chosen flavors. If the creamery asks you to help analyze this data, what approaches could you take? One approach is to use the natural symmetries of the underlying data space to decompose any data set into smaller parts that can be more easily understood. In this work, I describe how to use permutation representations of the symmetric group to create and study efficient algorithms that yield such decompositions.

Contents

Acknowledgments

First and foremost, I would like to thank my advisor, Professor Orrison. Without his guidance and mentorship I could not have completed this project. Additionally, the input and feedback of my second reader, Professor Omar, was invaluable. The support I received from my academic advisor, Professor Karp, was also invaluable throughout my growth as a mathematician. Finally, neither I nor this project would be where they are without the tremendous support of the entire Harvey Mudd College Mathematics Deptartment.

Chapter 1

Introduction

1.1 A Motivating Example

Imagine your local creamery administers a survey to its patrons, asking them to choose their five favorite ice cream flavors. Once the creamery tabulates the survey results, how can they best analyze this data? Simply finding the most often chosen set of flavors is informative, but does not really indicate much about the overall preferences of their consumers. For example, merely looking at the tabulated survey results, it would be very difficult to ascertain whether or not patrons were making flavor preferences that truly depended on a *group* of flavors, or merely an *individual* favorite flavor. In other words, imagine most patrons really just had one favorite flavor and, when asked to choose their top five, they just picked their favorite and selected four more at random. On the other hand, perhaps most patrons remember only their first favorite, second favorite, and third favorite, and choose the last two at random. These are both very different consumer mentalities, but when looking at only the tabulated survey results, it would be difficult to determine which mindset best described the consumer behavior.

Surveys like this, that ask participants to rank subsets of choices, as opposed to ranking all choices, reveal what is known as *partially ranked data*, as opposed to *fully ranked data*. In particular, the creamery's survey reveals partially ranked data because any data gathered from this survey can only rank the subset of chosen flavors above the not-chosen flavors. Analysis of partially ranked data is particularly challenging because the lack of a full ranking can mask various relationships between the data. One powerful way to analyze systems of partially ranked data defined on a finite set of options, X, is to examine how the data behaves under the action of a group G that acts on X. Group actions are particularly useful in data analysis because they capture the abstract notion of *symmetry* of the set X. For example, if X is our set of all possible choices of our five ice cream flavors, then the action of the symmetric group S_n on this set will capture all of the symmetries that exist in X by capturing every possible re-labeling of the flavors of X. As there are no relations on the set X, every re-labeling is a valid symmetry. On the other hand, if the set of options are vertices in a graph, (perhaps X represents a computer network, and a company wishes to analyze user traffic) then not all relabeling of vertices would respect the structure of the graph (namely, the edges). For any finite collection of options X, we can analyze how a data set over X behaves under the action of a group G that is a subgroup of the group of automorphisms of X. This group action will then capture some of the symmetries of X, and analyzing any data set in the context of this action will inform how it respects or violates these symmetries. In particular, if we express our data set over X as an element in a meaningful vector space (which will be our data space), then we can extend the action of G on X to a representation of G over the data space.

With this, we can use well established representation theoretic techniques to express any data set as a collection of well understood parts in the context of this group action. For the creamery's survey, this means writing our survey results in a manner that helps isolate first order effects, such as when patrons only care about their one favorite flavor, from higher order effects, such as when consumers truly care about groups of flavors. In fact, these ideas are useful far beyond merely analyzing ice cream preferences. Machine learning and artificial intelligence applications routinely employ these types of analyses to condense highly interconnected data into simpler, lower dimensional information. Generalized algebraic voting theory and algebraic statistics also often use these methods. For some examples of applications beyond what I look at here, or examples following different methodologies than I do, consider Clausen and Baum (1993), Huang (2008), Huang et al. (2008), Kondor et al. (2007), Daugherty et al. (2009), or Kondor and Barbosa (2010).

Though the method of analysis described above is very useful, it has a significant problem which hampers its applicability when analyzing rich, high-dimensional data sets. Though there are many tools to find appropriate data decompositions, the algorithmic complexity of transforming the data from the natural language of the tabulated surveys into this new interpretation can sometimes be quite slow. For real world applications, this complexity is of prime importance, as ideally this creamery would be able to run statistics on its daily sales of multi-flavor orders, or at least be able to analyze its static survey results with minimal computational effort. In this work, I focus on novel transformation algorithms that are much faster than the naïve, standard, mechanisms and I analyze the complexity of these transformations from a theoretical standpoint. In particular, I examine partially ordered data on the set of numbers from 1 through n, $X = \{1, \ldots, n\}$ through the action of the symmetric group on n elements, S_n.

1.2 An Overview of the Mathematics

In order to understand the transformations and data analysis methods we will devise, we first need to have a comprehensive understanding of the acting group. In this case, we focus explicitly on $G = S_n$, which is a well characterized, finite, non-abelian group. We will assume that the reader is reasonably familiar with its definition and basic properties.

Beyond understanding the mechanics of our group S_n, we need to understand the space in which our data lives. In order to capture the notion of arbitrary, partially ranked data on $X = \{1, \ldots, n\}$, we will introduce the mathematical notion of a tabloid, which is a collection of rows that are filled with numbers. These tabloids will enable us to define our set of voting options, and from there, our data space, which, we shall see, can be realized as a complex vector space. The action of S_n on X will then induce a representation of S_n over this space, and this will drive our transformation algorithms, which will be realized as straightforward change of basis algorithms. We will define these constructions explicitly and explore the basic mechanics of our data space in Section 2.1.

To fully understand the representation of S_n over our data space, we will examine its *irreducible decomposition*. To find the irreducible decomposition of our data space, we must must first understand the irreducible representations of S_n in general. These representations can be described explicitly through the use of *standard Young tableaux*. Using Young tableaux, we can specify exactly the irreducible decomposition of any representation of S_n over a tabloid space. This will be essential when establishing bounds on the complexity of our algorithms. This subject is detailed in Section 2.2.

With this background information established, we will move to Chapter 3, in which we will analyze the concept of a *symmetry adapted basis*, which will be our mechanism for decomposing data sets into well understood parts under the action of S_n. In this chapter, we will also detail the algorithms used in this work to transform from a basis natural for data collection to a basis natural for data analysis. Finally, we will analyze the results of this study in Chapter 4, and discuss the remaining problems and open questions in Chapter 5.

Chapter 2

Background Information

2.1 The Tabloid Space $\mathbb{C}X^\lambda$

In order to analyze partially ranked data, like the ice cream data from our creamery's survey in Chapter 1, we will first investigate partially ranked data in general over a set of options $\{1, \ldots, n\}$. How can we capture these partial rankings? We need a mathematical structure that separates $\{1, \ldots, n\}$ into subsets (or choices) of a particular size, such that there is some consistency in the ranking between these groups. We will capture this information via *tableaux* and *tabloids*.

2.1.1 Tableaux and Tabloids

In order to properly define tableaux and tabloids, we must first recall the notion of a *composition* of n.

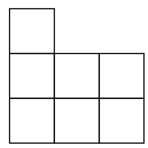
Definition 2.1. Let $n \in \mathbb{Z}_{\geq 0}$. Then, $\eta = (\eta_1, \ldots, \eta_r)$ is a *composition of* n, denoted $\eta \vDash n$, if $\sum_{i=1}^{r} \eta_i = n$ and $\eta_i \in \mathbb{Z}_{\geq 0}$.

As an aside, note that compositions are occasionally defined as infinite lists of non-negative integers, as opposed to finite lists, with the stipulation that there exists some $N \in \mathbb{N}$ such that after this index, every element of the composition is zero. In other words, if $\eta = (\eta_1, \eta_2, \ldots)$, with $\eta_i \in \mathbb{Z}_{\geq 0}$, then η is a composition if there exists $N \in \mathbb{N}$ such that for all $k \geq N$, $\eta_k = 0$. Our definition is, in fact, equivalent to this definition for our purposes; simply by setting $\eta = (\eta_1, \ldots, \eta_N)$ gives a finite list with the requisite properties, and the same sum as η.

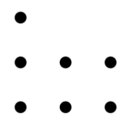
With compositions of n, we can now define the primary structure that we will use to capture partially ranked data.

Definition 2.2. Let $n \in \mathbb{Z}_{\geq 0}, \eta = (\eta_1, \ldots, \eta_r) \vDash n$. A *Young diagram* of shape η is a collection of r left justified rows of boxes, such that the ith row has length η_i. These diagrams are also called *Ferrers diagrams*, though then they are typically illustrated with rows of dots instead of boxes.

Example 2.1. Let $n = 7$, and $\eta = (1, 3, 3)$. Then, the Young diagram *(left)* or Ferrers diagram *(right)* of shape η is given by



and



How does this object help us categorize partially ranked data? Imagine a restaurant has 7 entrés and asks its patrons to choose first their favorite dish, then their next three top dishes, and finally their three least favorite dishes (which are just those that remain). We see that any response can be encoded by simply filling in the boxes of the Young diagram of shape η with the numbers $1, \ldots, 7$ (corresponding to entré one through seven). This filled diagram could then be read to catalog the favorite as the first ranked element in the first row, the next top three as the second ranked subset in the second row, and the remaining subset (also of size three) in the final row.

To capitalize on the insight gained in the preceding example, we see that any Young diagram can be seen to encode a particular survey for partially ranked data. Filling the boxes in a Young diagram of shape $\eta \vDash n$ will uniquely specify a ranking of a subset of size η_1 as first, and a subset of size η_2 as second, and so on and so forth, and thus correspond to a response to the given survey. To formalize this notion of a filled Young diagram, we make the following definition.

Definition 2.3. Let $n \in \mathbb{Z}_{\geq 0}$, with $\eta = (\eta_1, \ldots, \eta_r) \vDash n$. Then, a *Young tableaux* of shape η is any filling of the Young diagram of shape η with the numbers $\{1, \ldots, n\}$. If t is a Young tableaux of shape η, we let t_i be the set of numbers in the ith row of t.

We see by the preceding argument that any Young tableaux t of shape η corresponds to a ranking of subset t_1 above t_2, above t_3, and so on. Note that $|t_i| = \eta_i$, as the ith row contains exactly η_i boxes, and each Young tableaux is a filled with n distinct elements.

Example 2.2. If $\eta = (1,3,3)$, as in Example 2.1, then a possible Young tableaux of shape η is

$$t = \begin{array}{|c|c|c|} \cline{1-1} 4 & \multicolumn{2}{c}{} \\ \hline 3 & 7 & 6 \\ \hline 5 & 2 & 1 \\ \hline \end{array}.$$

We then see

$$t_1 = \{4\}$$
$$t_2 = \{3,7,6\}$$
$$t_3 = \{5,2,1\}$$

If this records a response to the survey described in Example 2.1, then the respondent has ranked the subsets $\{4\}$, $\{3,7,6\}$, and $\{5,2,1\}$, in that order.

Example 2.3. If our creamery has distributed a survey asking participants to indicate their top five flavors, out of 12 total flavors, then this corresponds to Young diagram (or Ferrers diagram) of shape

$$\begin{array}{|c|c|c|c|c|c|c|} \cline{1-5} & & & & & \multicolumn{2}{c}{} \\ \hline & & & & & & \\ \hline \end{array}$$

A given response could be

$$t = \begin{array}{|c|c|c|c|c|c|c|} \cline{1-5} 10 & 1 & 5 & 9 & 7 & \multicolumn{2}{c}{} \\ \hline 8 & 11 & 3 & 2 & 4 & 6 & 12 \\ \hline \end{array}$$

with

$$t_1 = \{10,1,5,9,7\}$$
$$t_2 = \{8,11,3,2,4,6,12\}.$$

Note that here, our Young diagram has $n = 12$ total boxes, even though the respondents are only asked to choose their favorite set of 5 flavors. This is because the Young diagram, and any associated tableaux, captures all ranked subsets. In this case, participants are really ranking two subsets: the five chosen flavors, and everything else (in this case, the seven remaining flavors).

To reflect the fact that we are interested in the rankings of subsets, not ordered lists, we introduce the notion of *row equivalence* of Young tableaux.

Definition 2.4. Two Young tableaux s, t are *row equivalent*, denoted $\equiv$, if they are of the same shape and $s_i = t_i$ for all applicable i.

Example 2.4. Consider the tableaux below:

$$t = \begin{array}{|c|c|c|c|c|cc} \hline 10 & 1 & 5 & 9 & 7 \\ \hline 8 & 11 & 3 & 2 & 4 & \multicolumn{1}{c|}{6} & \multicolumn{1}{c|}{12} \\ \hline \end{array}$$

$$s = \begin{array}{|c|c|c|c|c|cc} \hline 1 & 10 & 7 & 5 & 9 \\ \hline 2 & 11 & 3 & 8 & 6 & \multicolumn{1}{c|}{4} & \multicolumn{1}{c|}{12} \\ \hline \end{array}$$

$$p = \begin{array}{|c|c|c|c|c|cc} \hline 2 & 3 & 7 & 9 & 11 \\ \hline 1 & 4 & 5 & 6 & 8 & \multicolumn{1}{c|}{10} & \multicolumn{1}{c|}{12} \\ \hline \end{array}.$$

Then, $s \equiv t$ but $p \not\equiv s, t$.

It is clear that $\equiv$ is indeed an equivalence relationship; as such, it is natural to think of constructing equivalence classes of all tableaux under $\equiv$. With this, we make the following definition.

Definition 2.5. An equivalence class of tableaux under $\equiv$ is called a *tabloid*. Tabloids are typically denoted by a representative of the associated equivalence class, such that the entries in each row progress in increasing order from left to right. They are typically drawn in the same way as tableaux, but without the vertical lines between the columns.

Example 2.5. The associated tabloids to the tableaux in Example 2.4 are

$$t = s = \begin{array}{ccccccc} \hline 1 & 5 & 7 & 9 & 10 \\ \hline 2 & 3 & 4 & 6 & 8 & 11 & 12 \\ \hline \end{array}, \quad p = \begin{array}{ccccccc} \hline 2 & 3 & 7 & 9 & 11 \\ \hline 1 & 4 & 5 & 6 & 8 & 10 & 12 \\ \hline \end{array}.$$

Throughout the rest of this work, we will consistently work with tabloids rather than tableaux. Thus, equivalence under $\equiv$ will be assumed, and we will simply refer to tabloids by their associated representative tableaux, as in Example 2.5.

With the notation built up thus far, we can now describe the results of any survey in terms of tabloids shaped by compositions of n. These tabloids enable us to see the rankings of various subsets: the first row of the tabloid corresponds to the highest ranked subset, the second row to the second highest ranked subset, and so on. In any real survey, the ordering of the rows of our tabloids, and thus the ordering of the ranked subsets, is of singular importance (it is crucial to know if everyone loved these three flavors or if everyone hated them). Mathematically speaking, the analysis we will use will examine either case in exactly the same manner, so long as the data presents the responses in a consistently ordered way. In this way, the actual order of the rows in a given Young diagram does not matter for the sake of the data analysis, despite the fact that it matters greatly for the data collection. As such, we will make a simplifying assumption for much of the remainder of the work that will greatly aid in the presentation of several key theorems. To begin, recall the following definition.

Definition 2.6. Let $n \in \mathbb{Z}_{\geq 0}$. Then $\lambda = (\lambda_1, \ldots, \lambda_r)$ is a *partition* of n, denoted $\lambda \vdash n$, if λ is a composition of n and $\lambda_i \geq \lambda_{i+1}$, for all applicable i.

We can see that every composition $\eta \vDash n$, there corresponds a partition $\overline{\eta} \vdash n$ simply by rearranging the components of η to progress in non-increasing order. With this in mind, we will presume throughout the rest of this work that all data is presented via surveys that are characterized by *partitions* of n, not *compositions*. As we mentioned before, this assumption will greatly simplify the presentation of several results and, as we can simply rearrange the rows from any survey characterized by a composition η into the partition $\overline{\eta}$, this assumption can be made without loss of generality.

With this tabloid framework built up, let us clarify some notation we'll use throughout this work.

Notation. Given $\eta = (\eta_1, \ldots, \eta_r) \vDash n$, we will use $p(\eta) = \{i | \eta_i > 0\}$ to describe the indices corresponding to nonzero parts of η, as a composition, or, equivalently, the indices of rows of a Young diagram of shape η with nonzero length. Further, let

$$\binom{m}{\lambda} = \binom{m}{\lambda_1, \ldots, \lambda_k} = \binom{m}{\lambda_1}\binom{m-\lambda_1}{\lambda_2}\cdots\binom{m-\lambda_1-\cdots-\lambda_{r-1}}{\lambda_k}.$$

To describe all tabloids of a given shape, we make the following definition.

Definition 2.7. Given $\lambda \vdash n$, define

$$X^\lambda = \{\text{Tabloids } t | t \text{ is of shape } \lambda\}.$$

We can see that X^λ is a finite set; in particular, there are only as many tabloids of shape λ as there are ways to fill a tabloid of shape λ with the numbers 1 through n. This is inherently a combinatorial question, and we can answer it with the following theorem.

Proposition 2.1. *Let $\lambda \vdash n$. Then,*

$$|X^\lambda| = \binom{n}{\lambda} = \frac{n!}{\lambda!}$$

Proof. Let $\lambda = (\lambda_1, \ldots, \lambda_r)$. We see that in order to fill a tabloid of shape λ, you must fill every row, and the order in filling that row does not matter. Thus,

$$|X^\lambda| = \binom{n}{\lambda_1, \lambda_2, \ldots, \lambda_r} = \binom{n}{\lambda}$$

But, we can expand this algebraically and see

$$\begin{aligned}
|X^\lambda| &= \binom{n}{\lambda} \\
&= \binom{n}{\lambda_1}\binom{n-\lambda_1}{\lambda_2}\cdots\binom{n-\lambda_1-\cdots-\lambda_{r-1}}{\lambda_r} \\
&= \frac{n!}{\lambda_1!(n-\lambda_1)!} \cdot \frac{(n-\lambda_1)!}{\lambda_2!(n-\lambda_1-\lambda_2)!} \cdots \frac{(n-\lambda_1-\cdots-\lambda_{r-1})!}{\lambda_r!} \\
&= \frac{n!}{\lambda_1!\lambda_2!\cdots\lambda_r!} \\
&= \frac{n!}{\lambda!},
\end{aligned}$$

as desired. □

Note that this combinatorial argument can also be derived from the idea that we could fill this tabloid each way that we could fill its cells, but then we would have overcounted by exactly a factor of $\lambda_i!$ for each row i, as the order of numbers in each row does not matter.

We now define the primary space of interest in this work.

Definition 2.8. Let $\lambda \vdash n$, and define

$$\mathbb{C}X^\lambda = \{f : X^\lambda \to \mathbb{C}\}$$

to be the vector space of complex-valued functions over X^λ. It is clear, given that $\mathbb{C}X^\lambda$ is a space of functions, that under the standard pointwise addition and multiplication of scalars $\mathbb{C}X^\lambda$ is a complex vector space.

This space will be our *data space* in the remainder of our investigations. Why is this a sensible definition? Recall that our creamery is collecting data via tabloids of shape $\eta = (5,7)$, which has associated partition $\overline{\eta} = \lambda = (7,5)$, as in Example 2.3. Then, we see that any survey response for the creamery merely specifies a choice from the set of options X^λ. So, any dataset, which is merely a tabulation of survey results, can be seen as a function $f : X^\lambda \to \mathbb{Z}_{\geq 0}$, simply by specifying that for all tabloids t, $f(t)$ yields the number of respondents who selected option t. In other words, if

$$t = \begin{array}{c} \hline 2 \quad 3 \quad 4 \quad 6 \quad 8 \quad 11 \quad 12 \\ \hline 1 \quad 5 \quad 7 \quad 9 \quad 10 \\ \hline \end{array},$$

then $f(t)$ yields the number of respondents who chose flavors $1,5,7,9$, and 10 to be their list of five favorites, and who did not choose any of the remaining flavors. As $\mathbb{Z}_{\geq 0} \subset \mathbb{C}$, we see that thus $f \in \mathbb{C}X^\lambda$. We choose to extend the data space to all functions defined over $\mathbb{C}$, as opposed to just $\mathbb{Z}_{\geq 0}$ as $\mathbb{C}$ is a field (and, moreover, an algebraically closed field). In fact, defining our data space over $\mathbb{C}$ instead of $\mathbb{Z}_{\geq 0}$ gives it the structure of a complex vector space. This will allow us to realize our desired transformation of our dataset (or, data function in $\mathbb{C}\mathrm{X}^\lambda$) as a literal *change of basis* transformation.

Readers interested in exploring these topics further should consult Daugherty et al. (2009), which greatly informed the material of this section. Additionally, the theory of tableaux and tabloids is far richer than the sliver presented here; readers interested in seeing other facets of this theory should consult Fulton (1996).

2.1.2 Properties of $\mathbb{C}X^\lambda$

Proposition 2.2.

$$\dim(\mathbb{C}X^\lambda) = |X^\lambda|.$$

Proof. To prove this, we will construct a basis for $\mathbb{C}X^\lambda$ of size $|X^\lambda|$.

For each $t \in X^\lambda$, define $\delta_t : X^\lambda \to \mathbb{C}$ such that

$$\delta_t : s \mapsto \begin{cases} 1 & \text{if } s = t \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathcal{C} = \{\delta_t | t \in X^\lambda\}$. Now, consider any function $f \in \mathbb{C}X^\lambda$. Using the standard, pointwise definitions of function addition and scalar multiplication, we see

$$f = \sum_{t \in X^\lambda} f(t)\delta_t$$

Thus, $\langle \mathcal{C} \rangle = \mathbb{C}X^\lambda$.

Next, let $\{\alpha_t \in \mathbb{C} | t \in X^\lambda\}$ be such that

$$\sum_{t \in X^\lambda} \alpha_t \delta_t = 0$$

Then, we see that, for any $s \in X^\lambda$,

$$\begin{aligned} \alpha_s &= \sum_{t \in X^\lambda} \alpha_t \delta_t(s) \\ &= \left(\sum_{t \in X^\lambda} \alpha_t \delta_t \right)(s) \\ &= 0 \end{aligned}$$

Thus, $\mathcal{C}$ is linearly independent, and therefore a basis of $\mathbb{C}X^\lambda$. But, it is clear that $|\mathcal{C}| = |X^\lambda|$. Therefore, $\dim\left(\mathbb{C}X^\lambda\right) = |X^\lambda|$, as desired. $\square$

We will refer to $\mathcal{C}$ as the *delta basis*. Note that there is a canonical bijection between $\mathcal{C}$ and X^λ via $\delta_t \leftrightarrow t$. As such, we can also view the delta basis as consisting of literal tabloids in X^λ; then, linear combinations of these basis vectors correspond to formal linear combinations of tabloids in X^λ themselves. Under this restructuring, the space $\mathbb{C}X^\lambda$ is realized as the vector space given by all formal, complex linear combinations of tabloids in X^λ. This definition leads naturally into the action of S_n on $\mathbb{C}X^\lambda$.

2.1.3 The Action of S_n

To define this action, let us first define the action of S_n over X^λ.

Definition 2.9. Let $t \in X^\lambda$ and $\sigma \in S_n$. Then, we define $\sigma t = s$, where s is the tabloid resulting from permuting the positions of the entries of t according to σ.

Example 2.6. For example, consider $\sigma = (13)(25)$. Then

$$t = \begin{array}{ccc} \hline 1 & 2 & 3 \\ \hline 4 & 5 & \\ \cline{1-2} \end{array} \quad \text{and} \quad \sigma t = \begin{array}{ccc} \hline 3 & 5 & 1 \\ \hline 4 & 2 & \\ \cline{1-2} \end{array} = \begin{array}{ccc} \hline 1 & 3 & 5 \\ \hline 2 & 4 & \\ \cline{1-2} \end{array}.$$

Note that the action of S_n on X^λ has no conception of the order of the rows of λ. Thus, our presumption that data will be presented via tabloids defined by partitions of n is indeed justified. To describe this action further, recall the following definition.

Definition 2.10. Let X be a finite set, G a finite group such that G acts on X. The action of G on X is said to be *transitive* if, for all $x, y \in X$, there exists a $g \in G$ such that $gx = y$. The *stabilizer* of an element $x_0 \in X$ is the set

$$\operatorname{stab}(x_0) = \{g \in G | gx_0 = x_0\}.$$

Recall that $\operatorname{stab}(x)$ is a subgroup of G for all $x \in X$. The *orbit* of an element $x_0 \in X$ is the set

$$Gx_0 = \{x \in X | \exists g \in G \text{ such that } gx_0 = x\}.$$

The set of all orbits of G is denoted X/G.

We see that for any $\lambda \vdash n$, S_n acts transitively on the set X^λ—as S_n contains all possible permutations of $1, \dots, n$, it can certainly permute these equivalence classes of arrangements of boxes filled with $1, \dots, n$. With this in mind, consider the following theorem regarding transitive group actions.

Theorem 2.1. *Let G be a finite group, X a finite set on which G acts transitively. Then, their exists a bijection of sets between X and $G/\operatorname{stab}(x_0)$ that respects the action of G, for all $x_0 \in X$ (with the standard action of G on $G/\operatorname{stab}(x_0)$).*

Proof. As G acts transitively on X, for each $y \in X$, there exists a $g_y \in G$ such that $g_y x_0 = y$. Consider the map of sets $\varphi : X \to G/\operatorname{stab}(x_0)$, defined by

$$\varphi(y) = g_y \operatorname{stab}(x_0).$$

We will prove this map is a bijection that respects the action of the group.

Let us first show that it is injective. Imagine $\varphi(y) = \varphi(x)$ for $x, y \in X$. This implies that the cosets $g_y \operatorname{stab}(x_0)$ and $g_x \operatorname{stab}(x_0)$ are the same. But, then $g_x^{-1} g_y \in \operatorname{stab}(x_0)$, so $(g_x^{-1} g_y)x_0 = x_0$. Therefore,

$$\begin{aligned}(g_x^{-1} g_y)x_0 &= g_x^{-1}(g_y x_0) \\ &= g_x^{-1} y \\ &= x_0.\end{aligned}$$

Thus, $y = g_x x_0 = x$, which implies that φ is injective.

We now show that φ is surjective. Let $h\,\mathrm{stab}(x_0)$ be a coset of $\mathrm{stab}(x_0)$ in G. Then, $hx_0 = x$, for some $x \in X$. Consider the element $h^{-1}g_x \in G$. We see that $h^{-1}g_x(x_0) = h^{-1}(x) = x_0$, by the definition of x. Thus, $h^{-1}g_x \in \mathrm{stab}(x_0)$, so h and g_x are in the same coset of $\mathrm{stab}(x_0)$. Therefore, $\varphi(x) = g_x\,\mathrm{stab}(x_0) = h\,\mathrm{stab}(x_0)$, which implies that φ is surjective, and thus a bijection of sets.

To see that φ respects the action of the group, let $h \in G$. Then,

$$\begin{aligned}\varphi(hx) &= g_{hx}\,\mathrm{stab}(x_0) \\ &= hg_x\,\mathrm{stab}(x_0) && \text{as } hg_x(x_0) = hx = g_{hx}(x_0) \\ &= h\varphi(x),\end{aligned}$$

as desired. Thus, φ is a bijection of sets between X and $G/\,\mathrm{stab}(x_0)$ respecting the action of the group. □

This theorem is useful because it enables one to view the action of a finite group G on a set as the action of G on cosets of G. This can greatly simplify the representation theory of the resulting permutation module. Though we won't explore this connection extensively in this work, it primarily relates to the the symmetric group through the following example.

Example 2.7. Let $\lambda = (\lambda_1, \dots, \lambda_r) \vdash n$ and t a tabloid of shape λ, with, as defined previously, associated row sets $t_1, \dots, t_r$. Then, we can see that the stabilizer of t under the action of S_n is precisely the subgroup containing all permutations that only permute the row sets of t. Why? Under our row equivalence relationship, any permutation of this form will only adjust our equivalence class representative and not meaningfully change the tabloid t. On the other hand, we see that any permutation that permutes anything between the row sets of t will *not* fix t, as changing the row sets will produce an inequivalent tabloid. Thus,

$$\mathrm{stab}(t) \cong S_{t_1} \times S_{t_2} \times \cdots \times S_{t_r},$$

where we use notation S_{t_i} to denote the group of permutations of the set t_i. But, S_{t_i} is isomorphic to the symmetric group $S_{|t_i|}$, so we can realize this as

$$\mathrm{stab}(t) \cong S_{\lambda_1} \times \cdots \times S_{\lambda_r},$$

Thus, by the preceding theorem

$$X^\lambda \cong S_n/(S_{\lambda_1} \times \cdots \times S_{\lambda_r}),$$

where the isomorphism implied above respects the action of S_n.

We can extend the action of S_n on tabloids shaped by partitions of n to an action defined on tabloids shaped by *compositions* of n as well. Recall that for any composition $\eta \vDash n$, there exists a partition $\bar{\eta} \vdash n$, simply by reordering η in non-increasing order. Furthermore, we see that the action of S_n over a tabloid $t \in X^\lambda$ for some partition λ doesn't ever rely on the rows of t having any particular ordering. Thus, for some composition η, we can define the action on a tabloid t of shape η simply by permuting t to have shape $\bar{\eta}$ in the canonical way and using our action defined over $X^{\bar{\eta}}$, then permuting back to η. To that end, we will routinely use notation $\mathbb{C}X^\eta$ to mean, more technically, $\mathbb{C}X^{\bar{\eta}}$.

Definition 2.11. To extend the action of S_n on X^λ to one on $\mathbb{C}X^\lambda$, let $v = \sum \alpha_i t_i \in \mathbb{C}X^\lambda$. Then, define

$$\sigma v = \sum \alpha_i \left(\sigma t_i\right).$$

In order to realize the power of this action, recall the following definitions.

Definition 2.12. Let G be a finite group. Then,

$$\mathbb{F}G = \{f : G \to \mathbb{F}\} = \{\sum_{g \in G} \alpha_g G | \alpha_g \in \mathbb{F}\}$$

is the *group algebra* over $\mathbb{F}$. Recall that the group algebra is a ring, under the operation *convolution*, such that

$$(f * g)(x) = \sum_{y \in G} f(y) g(y^{-1} x).$$

We typically care about the *complex group algebra*, $\mathbb{C}G$.

Definition 2.13. Let G be a finite group. A pair (ρ, V) is a *representation* of G if V is a vector space over field $\mathbb{F}$ and $\rho : G \to \mathrm{GL}(V)$ is a group homomorphism. By encoding V with respect to a given basis, ρ, also encoded with respect to the basis, becomes a map from G to a group of invertible matrices.

The representation (ρ, V) can be extended to an *algebra representation*, which is a linear transformation from $\mathbb{F}G$ to $\mathcal{L}(V)$, formed by extending ρ linearly, using the elements of G as a basis for $\mathbb{F}G$ (much like the delta basis for $\mathbb{C}X^\lambda$).

Definition 2.14. Let R be a ring with 1. An abelian group M is a *left R-module* if R has an action on M that, for all $m, n \in M$ and $r, s \in R$, satisfies the properties

1. $r(m+n) = rm + rn$
2. $(r+s)m = rm + sm$
3. $(rs)m = r(sm)$
4. $1m = m$.

If M is a left R-module, it is often denoted ${}_RM$.

Note that $\mathbb{C}G$ contains an isomorphic "copy" of the complex numbers by simply scaling the identity in G. Thus, any $\mathbb{C}G$ module M can be thought of as a complex vector space, as it is an abelian group (M) with a defined notion of action by complex scalars (via the action of $\mathbb{C}G$).

Recall that for any representation (ρ, V) of a group G, there exists a corresponding $\mathbb{C}G$ module, given by ${}_{\mathbb{C}G}V$, with action defined by, for $x \in \mathbb{C}G$ and $v \in V$, $xv = \rho(x)v$. Similarly, for any $\mathbb{C}G$ module M, we can define a representation (ρ, M), where we realize M as a complex valued vector space, and $\rho(g) : m \mapsto gm$ for all $g \in G$. Thus, these perspectives are interchangeable—every representation of G corresponds exactly to a $\mathbb{C}G$ module, and vice-versa.

With these definitions in mind, we prove the following proposition.

Proposition 2.3. *If X is a finite set, and G a finite group that acts on X, then $\mathbb{C}X = \{f : X \to \mathbb{C}\}$ is a left $\mathbb{C}G$ module. In this case, ${}_{\mathbb{C}G}\mathbb{C}X$ is called a* permutation module, *as the action of any element $g \in G$ on $\mathbb{C}X$ is defined in terms of permuting the basis vectors of $\mathbb{C}X$.*

Proof. This proof is immediate—via the analog to the delta basis $\mathcal{C} = \{\delta_x | x \in X\}$, $\mathbb{C}X$ can be seen as a vector space of formal, complex linear combinations of elements in X. The action of G on X therefore defines a representation of G over $\mathbb{C}X$ defined by permuting the basis elements in $\mathcal{C}$. Therefore, it also defines a left-$\mathbb{C}G$ module, as desired. □

Corollary. *The action of S_n over X^λ gives $\mathbb{C}X^\lambda$ the structure of a $\mathbb{C}S_n$ permutation module, and defines an associated representation of S_n, $(\Theta_\lambda, \mathbb{C}X^\lambda)$, such that for all $\sigma \in S_n$ and $v \in \mathbb{C}X^\lambda$,*

$$(\Theta_\lambda(\sigma))\, v = \sigma v.$$

By restriction, it also thus defines a representation of the subgroups $S_j \leq S_n$ over $\mathbb{C}X^\lambda$. Note that if the context is clear, we will often drop the λ in the notation Θ_λ.

2.1.4 Orbits of $S_{n-1} \leq S_n$

A key facet of our eventual algorithm to analyze partially ranked data will be our ability to act on the space $\mathbb{C}X^\lambda$ with elements in $S_j \leq S_n$, in addition to to S_n itself. As such, something especially relevant to note about this action is the structure of the orbits of $S_j \leq S_n$ over X^λ. Though it is clear that S_n acts transitively on X^λ, S_j permits many orbits over this space. To describe these orbits, we use the following notation.

Definition 2.15. Let $\lambda = (\lambda_1, \ldots, \lambda_r) \vdash n$. Define 1_n^i to be the vector of size n with a 1 at the ith position, and 0s everywhere else. In particular,

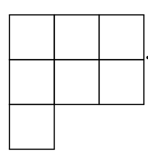
$$1_n^i = (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{n-i}).$$

Often, if the context is clear, we will simply use 1^i. Then, define

$$\lambda^i = \lambda - 1^i,$$

using standard component wise subtraction. Note that λ^i is only a meaningful object in this context if $\lambda_i > 0$; as such, we enforce the stipulation that λ^i is only defined if $i \in p(\lambda)$, which we recall is the set $p(\lambda) = \{i | \lambda_i > 0\}$, corresponding to the set of rows with positive length in a Young diagram of shape λ. As we will prove in a moment, λ^i corresponds to the ith orbit of S_{n-1} over $\mathbb{C}X^\lambda$.

Example 2.8. Let $n = 7$, $\lambda = (3, 3, 1) \vdash n$. This corresponds to the Young diagram

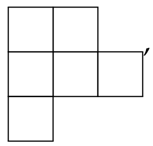
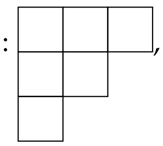
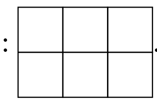


Then,

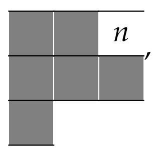
$$\begin{aligned}
\lambda^1 &= \lambda - 1^1 = (2, 3, 1) \\
\lambda^2 &= \lambda - 1^2 = (3, 2, 1) \\
\lambda^3 &= \lambda - 1^3 = (3, 3, 0)
\end{aligned}$$

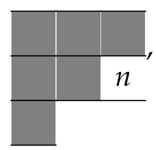
Note that λ^i is a composition of $6 = n - 1$ for all applicable i, but is no longer a partition of $n - 1$ for λ^2, in particular. The λ^i have the associated

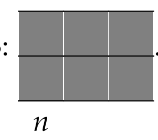
Young diagrams,

λ^1 : , λ^2 : , λ^3 : .

How can we see that these correspond to the orbits of S_{n-1} over X^λ? We see that S_{n-1} can never move the number n, so any two tabloids with n in a different row must be in different orbits. Simultaneously, S_{n-1} can permute the numbers $1, \ldots, n-1$ arbitrarily, so any two tabloids in X^λ with n in the same row must be in the same orbit. Therefore, the orbits of S_{n-1} in X^λ are describable exactly by which row n is in. For λ, we have,

Orbit 1: 

, Orbit 2: 

, Orbit 3: 

.

We can see that, by simply ignoring the fixed box containing n, and only examining the gray boxes remaining, we arrive at orbits characterized by Young diagrams of shape λ_1, λ_2, and λ_3. We will formalize and extend this argument in Theorem 2.2, but this intuition alone illustrates a use for the λ^i. As i specifies a row of λ, we can envision that λ^i catalogs the orbit of S_{n-1} such that n is in row i. Then, the associated Young diagram for λ^i details the shape of the composition that describes the set of tabloids in this particular orbit and the action of S_{n-1} on those tabloids.

In order to describe orbits of all $S_j \leq S_n$, we need a bit more notation.

Definition 2.16. As the orbits of S_j nest in a natural, recursive structure, we expand the definition of λ^i to allow for a multi-index exponent (recall a multi-index is just a list of indices). For some multi-index $l = l_1, l_2, \ldots, l_k$, define

$$\lambda^l = \lambda^{l_1,\ldots,l_{k-1}} - 1^k.$$

Just as in the case of the definition of λ^i, this is only sensible if each subsequent index in l corresponds to a row with positive length. As such, we enforce stipulation that λ^l is defined if and only if $l_1 \in p(\lambda)$, and $l_i \in p\left(\lambda^{l_1,\ldots,l_{i-1}}\right)$ for all $i > 1$. If l satisfies this property, l is said to be a *valid multi-index for shape* λ. For ease of notation, define

$$\lambda^* = \{\lambda^i | i \in p(\lambda)\} = \{\lambda^i | \lambda_i > 0\}$$

Similarly,

$$\lambda^{\ell *} = \{\lambda^{\ell,i} | i \in p(\lambda^{\ell})\}.$$

We also allow nesting, so

$$\lambda^{*^k} = \{\lambda^l | l \text{ is a valid multi-index of length } k \text{ for shape } \lambda\}.$$

Example 2.9. Note that we can realize the definition of λ^{*^k} another way, via

$$\lambda^{*^k} = \{\eta \vDash n-k \mid \eta \text{ is obtainable by decrementing parts of } \lambda\}.$$

This equality holds because decrementing a sequence of k parts of λ (or, equivalently, removing a sequence of boxes from the Young diagram of shape λ) results in a composition of $n-k$ that can be obtained by decrementing parts of λ. This suggests an alternative framework for viewing these λ^{*^k} and, more generally, the orbits of $S_j \leq S_n$ on $\mathbb{C}X^\lambda$. We see that the relation "is obtainable by decrementing parts of" defines a partially ordered set on all compositions. In particular, if η, ν are compositions, then we say that $\eta \leq \nu$ if and only if η can be obtained by decreeing parts of ν. Recall that we can think of compositions as infinite lists that terminate with an infinite sequence of zeros; under this framework, $\eta \leq \nu$ if and only if $\eta_i \leq \nu_i$ for all $i \in \mathbb{N}$. It can be shown that this forms a graded poset, with minimal element $0 = (0, 0, \ldots)$. Then, the set of all compositions obtainable by removing boxes from λ is given by $[0, \lambda]$, and the jth level of this subset of the poset, denoted $[0, \lambda]_j$, is equal to the set $\lambda^{*^{n-j}}$. We will not use this framework in this work, instead relying on the equivalent definition in terms of $\lambda^{l_1, \ldots, l_k}$, but it provides another way to analyze the orbits of $S_j \leq S_n$.

Example 2.10. If, again, $n = 7$ and $\lambda = (3, 3, 1)$, then

$$\begin{aligned}
\lambda^{1,2,2} &= \lambda^{1,2} - 1^2 \\
&= \lambda^1 - 1^2 - 1^2 \\
&= \lambda - 1^1 - 1^2 - 1^2 \\
&= (2, 1, 1),
\end{aligned}$$

and

$$\begin{aligned}
\lambda^* &= \{\lambda^1, \lambda^2, \lambda^3\} \\
&= \{(2,3,1), (3,2,1), (3,3,0)\}, \\
\lambda^{*^2} &= \{\lambda^{1,1}, \lambda^{1,2}, \lambda^{1,3}, \lambda^{2,1}, \lambda^{2,2}, \lambda^{2,3}, \lambda^{3,1}, \lambda^{3,2}\} \\
&= \{(1,3,1), (2,2,1), (2,3,0), (2,2,1), (3,1,1), (3,2,0), (2,3,0), (3,2,0)\}.
\end{aligned}$$

With this notation, let us examine the orbits of $S_j \leq S_n$ over X^λ, and how they relate to our data space, $\mathbb{C}X^\lambda$. We begin with a simple lemma.

Lemma 2.1. *Let $\lambda = (\lambda_1, \ldots, \lambda_r) \vdash n$, with $t, t' \in X^\lambda$. Then, t and t' are in the same orbit of S_j if and only if t and t' agree in all boxes containing numbers larger than j.*

Proof. In the forwards direction, we note that if t and t' are in the same orbit of S_j, then there exists some $\sigma \in S_j$ such that $\sigma t = t'$. But, as $\sigma \in S_j$, σ does not permute any $i > j$. Thus, t and t' must agree in the locations of these numbers. In the backwards direction, if t and t' agree in the location of all $i > j$, then there exists some $\sigma \in S_j$ such that $\sigma t = t'$ as S_j contains every possible permutation of the numbers $1, \ldots, j$, and only those numbers can be in different locations in t and t'. Thus, the lemma is proved. □

With this lemma, we can now characterize the orbits of S_j over X^λ, and describe how they inform the permutation module $\mathbb{C}X^\lambda$.

Theorem 2.2. *Let $\lambda = (\lambda_1, \ldots, \lambda_r) \vdash n$. Then, the orbits of S_j over X^λ are precisely described by the set $\lambda^{*^{n-j}}$. Moreover,*

$$ {}_{\mathbb{C}S_j}\mathbb{C}X^\lambda \cong \bigoplus_{\eta \in \lambda^{*^{n-j}}} {}_{\mathbb{C}S_j}\mathbb{C}X^\eta $$

Proof. We recall that

$$ \begin{aligned} \lambda^{*^{n-j}} &= \{\eta \vDash j | \eta \text{ is obtainable by removing boxes from } \lambda.\} \\ &= \{\lambda^l | l \text{ is a valid multi-index of length } n-j \text{ for shape } \lambda\}. \end{aligned} $$

We will first show that there exists a canonical bijection between the orbits of S_j over X^λ and the set $\lambda^{*^{n-j}}$. We see by Lemma 2.1 that there exists a bijection between the orbits of S_j and placements of all numbers $j+1, \ldots, n$ in a tabloid of shape λ. But, what is a placement of numbers of $j \ldots, n$? It is an ordered selection of boxes. But, as row order doesn't matter in a tabloid, this can be seen as an ordered selection of rows of λ. But, this is nothing more than a multi-index l, such that each entry l_i corresponds to a valid row after "removing" a box from each of the rows previously specified by l. How long must this list be? Exactly $n-j$ elements long. The set of all such l is, by definition, bijective with the set $\lambda^{*^{n-j}}$, so there is a bijection between the orbits of S_j in X^λ and elements of $\lambda^{*^{n-j}}$.

We call this bijection *canonical* because of the following. Suppose $t \in X^\lambda$, and $S_j t$ is an orbit of S_j over X^λ. Then, this orbit is associated with $\eta = \lambda^l \in \lambda^{*^{n-j}}$, where $l = l_1, \dots, l_{n-j}$ catalogs the rows of t containing all numbers $i > j$. Moreover, the action of S_j over this orbit is exactly described by the action of S_j on X^η, as S_j cannot permute any element $i > j$.

Now that we have established this bijection, we must show that

$$_{\mathbb{C}S_j}\mathbb{C}X^\lambda \cong \bigoplus_{\eta \in \lambda^{*^{n-j}}} {}_{\mathbb{C}S_j}\mathbb{C}X^\eta$$

To see this, let us note the following. Recall that $\mathbb{C}X^\lambda$ is realizable as the space of all formal, complex linear combinations of elements of X^λ. As the orbits of S_j partition X^λ, this interpretation of $\mathbb{C}X^\lambda$ reveal that each orbit of S_j will form a basis for a $\mathbb{C}S_j$-invariant subspace of $\mathbb{C}X^\lambda$, with subspaces corresponding to distinct orbits disjoint. These invariant subspaces form disjoint submodules of $_{\mathbb{C}S_j}\mathbb{C}X^\lambda$, and the structure of these permutation modules is described exactly by the action of S_j on their bases (which are orbits of S_j). But, we know that if $\eta \in \lambda^{*^{n-j}}$, then the action of S_j on the orbit corresponding to η is described exactly by the action of S_j on X^η. This implies that the submodule given by the invariant subspace spanned by the orbit associated with η is exactly the $\mathbb{C}S_j$ module $\mathbb{C}X^\eta$. Thus, we have that $_{\mathbb{C}S_j}\mathbb{C}X^\lambda$ breaks down into a collection of invariant subspaces, parametrized by $\eta \in \lambda^{*^{n-j}}$, with the action of S_j on the space associated to η described exactly by the module $_{\mathbb{C}S_j}\mathbb{C}X^\eta$. Therefore,

$$_{\mathbb{C}S_j}\mathbb{C}X^\lambda \cong \bigoplus_{\eta \in \lambda^{*^{n-j}}} {}_{\mathbb{C}S_j}\mathbb{C}X^\eta,$$

as desired. □

2.2 Irreducible Representations of The Symmetric Group

In order to understand any dataset that is realized as an element of $\mathbb{C}X^\lambda$, we must learn how to meaningfully decompose the space $\mathbb{C}X^\lambda$ in general. To do so, we will use the notion of an irreducible decomposition of a representation. Recall the following definition.

Definition 2.17. Let G be a finite group, with (ρ, V) a representation of G. A subspace $W \subset V$ is said to be *ρ-invariant* if $\rho(g)W \subset W$ for all $g \in G$. The representation ρ is said to be *irreducible* if the only ρ-invariant subspaces of V are $\{0\}$ and V itself. Note that, in module theoretic terms, $W \subset V$ is ρ-invariant if and only if W is a submodule of ${}_{\mathbb{C}G}V$, and V is irreducible if it permits no nontrivial submodules.

The invariant representations of a group G are supremely meaningful, as they can be used to capture almost all of the algebraic structure of *any* representation of G. To see this, recall the following theorem, whose proof can be found in James and Liebeck (1993), Chapter 8.

Maschke's Theorem. *Let G be a finite group, with (ρ, V) a complex representation of G. If $U \subset V$ is a ρ-invariant subspace, then there exists another ρ-invariant subspace $W \subset V$ such that $V = U \oplus W$.*

Corollary. *If G is a finite group and M is a $\mathbb{C}G$ module, then there exist irreducible submodules $M_1 \ldots M_N$ such that*

$$M = M_1 \oplus M_2 \oplus \cdots \oplus M_{N-1} \oplus M_N.$$

By the definition of $\oplus$, if $m \in M$, there exist $m_i \in M_i$ such that $m = m_1 + \cdots + m_N$. The M_i and m_i are said to be an irreducible decomposition *of M and m, respectively.*

Our mechanism for "meaningfully decomposing" a dataset will be to find its irreducible decomposition, in a particularly meaningful way. To do so, it will be very product to first examine the irreducible representations of S_n in general. We shall see that these can be described exactly by the *standard Young tableaux*.

2.2.1 Irreducible Representations Indexed by Standard Tableaux

Definition 2.18. Given $\lambda \vdash n$, the *standard tableaux* or *standard Young tableaux* of shape λ are the tableaux of shape λ that have the property that every row and column progress in increasing order. The standard tableaux of shape λ are typically ordered, for convenience, with the *last letter ordering*, which orders those tabloids such that those with larger numbers in higher rows come before those with smaller numbers in higher rows. More specifically, let t, s be standard tableaux of shape λ. Then, t precedes s in the last letter ordering if and only if, reading *right to left* by row, starting from the first row, has the property that the first difference encountered between t and s has t with a larger number than s. The collection of standard tableaux in this order are called the *last letter sequence* of shape λ. The length of this sequence is denoted f^{λ}.

Example 2.11. Let $\lambda = (3, 2)$. The last letter sequence associated with λ is

$$\begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array}.$$

This ordering comes because we first have all the standard tableaux with the largest number, 5, (hence the name *last letter ordering*) in the first row, and then all those with 5 in the second row. Inside each of these groups, the same holds true of the next largest letter, 4, and so on and so forth.

If t, s are tableaux of shape λ given by

$$t = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array}, \; s = \begin{array}{|c|c|c|}\hline 1 & 3 & 2 \\ \hline 4 & 5 \\ \cline{1-2}\end{array},$$

then t is a standard tableaux and s is not, as the first row does not progress in increasing order.

Note that the standard tableaux are *tableaux*, not *tabloids*, and thus we do not assume that row equivalence is true equality. However, as the row constraints of a standard tableau uniquely determine the order of its entries, we will never have more than one representative for each equivalence class of row-equivalent tableaux in a list of standard tableaux.

In fact, though its proof is beyond the scope of this work, the irreducible representations of S_n are indexed exactly by the partitions of n, and, for each $\lambda \vdash n$, the associated representation has dimension given by the number of standard tableaux of shape λ. Furthermore, we can discern the matrices associated with these representations directly for all $\sigma \in S_n$. To do so, we will need another definition.

Definition 2.19. Let t be a standard tableau of shape $\lambda \vdash n$. Then, for $1 \leq i, j \leq n$, define the *axial distance* between i and j, $d_t(i, j)$, to be the difference between the number of steps down or to the left and the number of steps up or to the right along any path from i to j in t. This will be well-defined as we can work in a taxicab geometry on the tabloid t.

Example 2.12. Let $t = \begin{array}{|c|c|c|}\hline 1 & 3 & 6 \\ \hline 2 & 7 & 9 \\ \hline 4 & 8 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array}$. Then, $d_t(9, 4) = 3$ and $d_t(8, 3) = -2$

Now, consider the following definition of a very particular representation of S_n, which we adapt from James and Kerber (1982), Section 3.4.

Definition 2.20. Let $\lambda \vdash n$, $\alpha_i = (i\ i+1) \in S_n$ be the ith adjacent transposition in S_n, and $\left(t_1, \ldots, t_{f^\lambda}\right)$ be the last letter sequence of shape λ. Then, define the S_n-representation $\left(\theta_\lambda, \mathbb{C}^{f^\lambda}\right)$, or, in module notation, S^λ, as follows. Let M be the $f^\lambda \times f^\lambda$ zero matrix. For each $1 \leq j \leq f^\lambda$, in order, do the following:

1. If t_j has already been the image of a previous standard tableau, do nothing.
2. If i and $i+1$ are in the same row of t_j, then set $M_{j,j} = 1$.
3. If i and $i+1$ are in the same column of t_j, then set $M_{j,j} = -1$.
4. If i and $i+1$ are not in the same row or column, then $\alpha_i t_j = t_\ell$ for some $j \leq \ell \leq f^\lambda$. Set the sub-matrix $M_{j,\ell} = \begin{pmatrix} M_{j,j} & M_{j,\ell} \\ M_{\ell,j} & M_{\ell,\ell} \end{pmatrix}$ as follows:

$$M_{j,\ell} = \begin{pmatrix} -d^{-1}(i,i+1) & \sqrt{1-d^{-2}(i,i+1)} \\ \sqrt{1-d^{-2}(i,i+1)} & d^{-1}(i,i+1) \end{pmatrix}$$

Then, let $\theta_\lambda(\alpha_i) = M$. As $\{\alpha_i\}$ generates S_n, this defines a representation $\left(\theta_\lambda, \mathbb{C}^{f^\lambda}\right)$ and module S^λ. This construction produces the *Young's orthogonal form* of the module S^λ.

We now state the following theorem, which we will not prove here. Its proof can be found in James and Kerber (1982), Chapter 3, or, for a different treatment, in Fulton (1996), Chapter 7.

Theorem 2.3. *The set S^λ for all $\lambda \vdash n$ is a complete set of irreducible modules of S_n.*

The Young's orthogonal form of any irreducible representation yields a construction for a matrix representation of every group element $\sigma \in S_n$, for every irreducible representation of S_n. As we will see, these matrix forms are essential when forming our final, symmetry adapted basis.

More information surrounding the irreducible representations of S_n can be found in James and Kerber (1982), Fulton (1996), or Sagan (1991). More information on representation theory in general can be found at James and Liebeck (1993), Dummit and Foote (2004), or Fulton and Harris (1991).

2.2.2 Decomposition of a Tabloid Representation

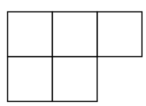
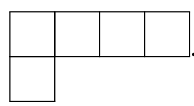
We can form a more explicit characterization of the role these representations play when speaking of tabloid representations, like our data space $\mathbb{C}X^\lambda$. To do so, consider the following definitions.

Definition 2.21. Given two partitions $\lambda = (\lambda_1, \ldots, \lambda_r), \mu = (\mu_1, \ldots, \mu_{r'}) \vdash n$, we say that *$\mu$ dominates λ*, or, *λ precedes μ in the dominance order*, denoted $\mu \trianglerighteq \lambda$ or $\lambda \trianglelefteq \mu$, respectively, if the shape μ can be formed from the shape λ by moving any number of boxes from any row of λ up any number of rows. More formally, $\mu \trianglerighteq \lambda$ if and only if

$$\sum_{i=1}^{j} \lambda_i \leq \sum_{i=1}^{j} \mu_i, \quad \text{for all } 1 \leq j \leq \max(r, r').$$

Note that here we take convention that $\mu_k = 0$ if $k > r'$ (and similarly for λ).

Example 2.13. Let $\lambda = (3,2)$ and $\mu = (4,1)$, illustrated as tabloid shapes via

$\lambda =$  $\mu =$ .

Then, $\mu \trianglerighteq \lambda$.

Example 2.14. Let $\lambda \vdash n$. Then, we see that $(n) \trianglerighteq \lambda$ and $\lambda \trianglerighteq \underbrace{(1,1,\ldots,1)}_{n}$.

Definition 2.22. Let $\lambda = (\lambda_1, \ldots, \lambda_r) \vdash n$. Define the *Kostka set* K^λ to be the multiset given by

$$K^\lambda = \{(i, \lambda_i) | 1 \leq i \leq r\},$$

so K^λ contains the number i exactly as many times as there are boxes in the ith row of λ. Additionally, define the *Kostka constraints* to be filling constraints on a shape $\mu \vdash n$ such that each row of μ progresses in *non-decreasing* order while each column progresses in *strict increasing* order. A *Kostka filling* of μ by λ is a filling of μ from elements of the Kostka set K^λ satisfying the Kostka constraints.

Given these ideas, we can define a Kostka number.

Definition 2.23. Let $\lambda, \mu \vdash n$. Then, define the *Kostka number* $\kappa_{\mu,\lambda}$ to be

$$\kappa_{\mu,\lambda} = \text{ The number of Kostka fillings of } \mu \text{ by } \lambda.$$

Note that if $\mu \ntrianglerighteq \lambda$, then there are zero Kostka fillings of μ by λ, so $\kappa_{\mu,\lambda} = 0$.

Example 2.15. Let $\lambda, \mu \vdash 5$ such that, again,

$$\lambda = \ydiagram{3,2} \qquad \mu = \ydiagram{4,1}.$$

Then, $\kappa_{\mu,\lambda} = 1$. To see this, let us illustrate the multiset K^{λ} directly:

$$K^{\lambda} = \begin{ytableau} 1 & 1 & 1 \\ 2 & 2 \end{ytableau}$$

We see that to satisfy all the constraints, all of the ones must fill the first row of μ. But, this only leaves one other option, so its positions are forced. Thus, we see that the only Kostka filling is

$$\begin{ytableau} 1 & 1 & 1 & 2 \\ 2 \end{ytableau}$$

Example 2.16. A more interesting case is with $\lambda = (3, 1, 1)$ and $\mu = (4, 1)$. Then, we see we wish to fill

$$\ydiagram{4,1}$$

with entries from

$$K^{\lambda} = \begin{ytableau} 1 & 1 & 1 \\ 2 \\ 3 \end{ytableau}$$

Here, we can move either the 3 up to the first row or the 2 up to the first row while still maintaining the ordering constraints. Thus, $\kappa_{\mu,\lambda} = 2$. Both Kostka fillings are illustrated below.

$$\begin{ytableau} 1 & 1 & 1 & 2 \\ 3 \end{ytableau} \qquad \begin{ytableau} 1 & 1 & 1 & 3 \\ 2 \end{ytableau}$$

We will use Kostka numbers in this work primarily through the following theorem, which is proved in Fulton (1996), Chapter 7 and Sagan (1991), Chapter 2.

Theorem 2.4 (Young's Rule)**.** *If $\lambda \vdash n$, then the module ${}_{\mathbb{C}S_n}\mathbb{C}X^\lambda$ decomposes as follows:*

$$
\begin{aligned}
{}_{\mathbb{C}S_n}\mathbb{C}X^\lambda &\cong \bigoplus_{\mu \vdash n} \kappa_{\mu,\lambda} S^\mu \\
&\cong \bigoplus_{\mu \trianglerighteq \lambda} \kappa_{\mu,\lambda} S^\mu
\end{aligned}
$$

Example 2.17. Let $\lambda = (n-k, k)$. Then, by Theorem 2.4, we know that

$$
{}_{\mathbb{C}S_n}\mathbb{C}X^\lambda \cong \bigoplus \cong \bigoplus_{\mu \trianglerighteq \lambda} \kappa_{\mu,\lambda} S^\mu
$$

So, in order to find the irreducible decomposition of $\mathbb{C}X^\lambda$, we need merely find the Kostka numbers associated with shapes that dominate λ. first, we see by inspection that $\mu \trianglerighteq \lambda$ if and only if $\mu = (n-\ell, \ell)$, where $0 \leq \ell \leq k$. However, as tabloids of shape λ have only two rows, we see that K^λ will only ever have, at most, two distinct entries: 1 and 2. Further, we know by the Kostka constraints on the columns that *any* Kostka filling of μ must have all of the 1s in K^λ in the first row of μ. But, then there is only one distinct entry, 2, in K^λ left. Thus, we must continue filling μ by simply filling in the remaining entries with 2s. Therefore, there is only one Kostka filling of μ, so $\kappa_{\mu,\lambda} = 1$, for any $\mu \trianglerighteq \lambda$. Thus,

$$
\mathbb{C}X^\lambda \cong \bigoplus_{\ell \leq k} S^{(n-\ell,\ell)}.
$$

We will refer to this result as the condition on two-rowed tabloids that they are *multiplicity free* because their decompositions require no irreducible representations with multiplicity.

Now that we have detailed the definition of the space ${}_{\mathbb{C}S_n}\mathbb{C}X^\lambda$ and described how it decomposes into irreducible submodules, we are ready to define the type of basis that will be ideal to extract this qualitative information out of any data set $f \in \mathbb{C}X^\lambda$. To do so, we must explore the notion of a symmetry adapted basis.

Chapter 3

Symmetry Adapted Bases

3.1 Introduction

We have claimed that transforming our space $\mathbb{C}X^\lambda$ from the standard, data-collection basis $\mathcal{C} = X^\lambda$ into one that is "symmetry adapted" would reveal qualitatively meaningful information regarding a dataset over X^λ. To understand this claim, recall Theorem 2.4. This theorem tells us that for any $\lambda \vdash n$, the module ${}_{\mathbb{C}S_n}\mathbb{C}X^\lambda$ decomposes according to the following formula:

$$ {}_{\mathbb{C}S_n}\mathbb{C}X^\lambda \cong \bigoplus_{\mu \trianglerighteq \lambda} \kappa_{\mu,\lambda} S^\mu. $$

If we define the space V_μ^λ to be

$$ V_\mu^\lambda = \kappa_{\mu,\lambda} S^\mu = \bigoplus_{i=1}^{\kappa_{\mu,\lambda}} S^\mu, $$

then we see that we also have

$$ {}_{\mathbb{C}S_n}\mathbb{C}X^\lambda \cong \bigoplus_{\mu \trianglerighteq \lambda} V_\mu^\lambda. $$

This type of decomposition of a module is called an *isotypic decomposition*, as the set of subspaces are pairwise inequivalent and invariant, but all multiplicity has been suppressed by combining all spaces of the same "type," or, that correspond to the same irreducible representation. In addition to being a module decomposition, we see that this decomposition is also a vector space decomposition, so any $f \in \mathbb{C}X^\lambda$ will also decompose uniquely into

$$f = \sum_{\mu \trianglerighteq \lambda} f_\mu,$$

where $f_\mu \in V_\mu^\lambda$. It is this decomposition that we need when we want to make sense of partially ranked data. Let us analyze this for a moment. Why is this decomposition better, or more interesting, than any other decomposition of the vector space $\mathbb{C}X^\lambda$? We recall that we use the structure of a group in this question particularly because a group enables us to encapsulate some notion of symmetry on the set of options X^λ. But, if our group is defining our notion of "symmetry," then every way that these symmetries can be informative is contained in the module decomposition of $\mathbb{C}X^\lambda$, as this totally defines the action of G on our data space. Moreover, as each inequivalent irreducible submodule expressed in this decomposition is orthogonal, each component f_μ will reveal a different expression of meaningful symmetry. So, if we apply this analysis to a dataset f, which, we recall, lives in the space $\mathbb{C}X^\lambda$, this will enable much more informative statistics to be drawn about the population than merely examining the expression of f in the standard, delta function basis $\mathcal{C}$.

In order to access this decomposition of any data vector, we will need some way to express this isotypic decomposition of $\mathbb{C}X^\lambda$. This is where the notion of a "symmetry adapted basis" will be so helpful. We make the following definition:

Definition 3.1. Let G be a finite group, (Θ, V) an n-dimensional complex representation of G, and $\theta_\rho, 1 \leq \rho \leq M$ a complete set of irreducible representations of G, whose matrix forms are known. Then, we know that Θ decomposes completely into a direct sum

$$\Theta \cong \bigoplus_{\rho=1}^{M} c_\rho \theta_\rho,$$

where $c_\rho \in \mathbb{Z}_{\geq 0}$. We say that a basis $\mathcal{B}$ for V is *symmetry adapted* if, for any $g \in G$, we have that $[\Theta(g)]_\mathcal{B}$ is block diagonal, with a block for each θ_ρ such that $c_\rho > 0$, such that all c_j blocks corresponding to θ_j are all equal to the known matrix $\theta_j(g)$, as opposed to just merely equivalent.

These types of bases enable us to capture the structure of an isotypic decomposition in a particularly natural way. In particular, we can see that the isotypic decomposition of a module encapsulates symmetries by condensing all components corresponding to a given irreducible representation into one, invariant subspace. In this way, an isotypic decomposition highlights the relevant information, namely, to which irreducible representation a component of the space corresponds, and masks the irrelevant information, such as to which isomorphic copy of this irreducible representation the component corresponds. Similarly, when expressed in a symmetry adapted basis, each irreducible representation is expressed differently, but the isomorphic copies of each irreducible representation behave in *exactly* the same way under the action of any group element.

In order to make use of a symmetry adapted basis, we will need some algorithm to produce one for the symmetric group. In Section 3.2, we'll explore and construct algorithms to produce symmetry adapted bases, both in general and on S_n explicitly. Then, in Section 3.3 we will detail some theoretical tools used to analyze these algorithms. Finally, in Section 3.4 we will detail our implementation of these algorithms.

3.2 Algorithms to Produce Symmetry Adapted Bases

This section's work is largely based on an algorithm presented in Stiefel and Fässler (1992), Chapter 5. The interested reader should consult this text for more information on the topics of this section. Additional noteworthy references are Clausen and Baum (1993) and Terras (1999).

3.2.1 Theoretical Justification

Let us begin with a theorem proving the *existence* of a symmetry adapted basis for a general complex representation of a finite group. Throughout this section, let G be a finite group, (Θ, V) an n dimensional complex representation, and $\theta_1, \ldots, \theta_M$ a complete set of irreducible, pairwise non-isomorphic representations of G, that are known in the form of matrix tables.

Theorem 3.1. *Given G, (Θ, V) as above, there exists a symmetry adapted basis on V, with respect to $\theta_1, \ldots, \theta_M$.*

Proof. Given the complete set of irreducible representations of G, $\theta_1, \ldots, \theta_M$, we can decompose Θ via

$$\Theta \cong \bigoplus_{\rho=1}^{M} c_\rho \theta_\rho$$

$$V \cong \bigoplus_{\rho=1}^{M} V_\rho$$

$$\cong \bigoplus_{\rho=1}^{M} \bigoplus_{i=1}^{c_\rho} V_\rho^i,$$

where V_ρ corresponds to the c_ρ copies of θ_ρ, with each copy given by V_ρ^i. Let us construct the following, alternative representation for G, $(\Theta', V' = V)$, such that

$$\Theta' = \bigoplus_{\rho=1}^{M} c_\rho \theta_\rho$$

Note that here, Θ' is *equal* to this direct product, not merely isomorphic to it. Consequently, the standard basis for V' is clearly symmetry adapted on Θ', as, by construction, $\Theta'(g)$ will have the desired form for all $g \in G$. But, we can also see that $\Theta' \cong \Theta$. Thus, there exists a vector space isomorphism between V' and V that respects the actions of Θ' and Θ. As V' and V are canonically isomorphic, this isomorphism is merely a change of basis on V. Thus, there exists a basis for V that is symmetry adapted for Θ, as desired. □

We will now use the abstract *existence* of a symmetry adapted basis to justify and drive the creation of an algorithm to produce such a basis. To do so, recall the definition of a Discrete Fourier Transform.

Definition 3.2. Recall that $\mathbb{C}G$ is the *complex group algebra* of G, given by $\mathbb{C}G = \{f : G \to \mathbb{C}\}$. Any $\mathbb{C}$-algebra isomorphism

$$D : \mathbb{C}G \to \mathbb{C}^{n_1 \times n_1} \oplus \cdots \oplus \mathbb{C}^{n_M \times n_M},$$

where the operations on the right are component-wise addition and multiplication, is a *discrete Fourier transform* (DFT) of G.

Recall that for any DFT D of G, the component maps $\theta_1, \ldots, \theta_M$ form a complete set of irreducible representations for G.

Definition 3.3. Let $H \leq G$. We say a DFT $D = \theta_1 \oplus \cdots \oplus \theta_M$ is *subgroup-adapted* or merely *adapted* to the chain $H \leq G$ if, for all $h \in H$ and $1 \leq i \leq M$,

1. $\theta_i(h)$ is block diagonal, with blocks corresponding to irreducible representations of H, and,
2. blocks inside the matrix $\theta_i(h)$ corresponding to equivalent irreducible representations of H are actually equal.

Note that this can be expressed differently: a DFT D is subgroup adapted to $H \leq G$ if the known matrix forms of each irreducible representations θ_i are encoded with respect to a symmetry adapted basis on H.

We see that Definition 3.3 can also be used to describe a DFT adapted to any finite chain of subgroups $H_1 \leq H_2 \leq \cdots \leq H_n \leq G$. Now, we will define an operator that will be integral in constructing a symmetry adapted basis.

Definition 3.4. Let D be a DFT for G, with component maps $\theta_1, \ldots, \theta_M$. Then, define

$$\hat{b}^j_{k,\ell} = \left(\underbrace{(\mathbf{0}), \ldots, (\mathbf{0})}_{k-1}, \; \begin{matrix} \\ k \end{matrix} \overset{\ell}{\begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}}, \underbrace{(\mathbf{0}), \ldots, (\mathbf{0})}_{h-k} \right)$$

and

$$b^j_{k,\ell} = D^{-1}\left(\hat{b}^j_{k,\ell}\right).$$

Note that $\left(\hat{b}^j_{k,\ell}\right)_{k,\ell,j}$ forms a basis for $\mathrm{Im}(D)$, so, as D is an isomorphism, $\left(b^j_{k,\ell}\right)_{k,\ell,j}$ forms a basis for $\mathbb{C}G$ (as a complex vector space). This basis is called the *dual matrix coefficient basis*, and its elements are called *dual matrix coefficients* or *dual matrix coefficient functions*.

Proposition 3.1. *Given the DFT D, the dual matrix coefficients obey the following relationship:*

$$b^j_{k,1} b^j_{1,1} = b^j_{k,1}$$

Proof. As D is an algebra isomorphism, we have that

$$b^j_{k,1} b^j_{1,1} = D^{-1}\left(\hat{b}^j_{k,1}\hat{b}^j_{1,1}\right)$$

But,

$$\begin{aligned}
\hat{b}^j_{k,1}\hat{b}^j_{1,1} &= \left(\underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{j-1},\ k\begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}, \underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{M-j} \right) \times \\
&\quad \left(\underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{j-1}, \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}, \underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{M-j} \right) \\
&= \left(\underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{j-1},\ k\begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}, \underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{M-j} \right) \\
&= \hat{b}^j_{k,1}.
\end{aligned}$$

Thus,

$$\begin{aligned}
b^j_{k,1} b^j_{1,1} &= D^{-1}\left(\hat{b}^j_{k,1}\right) \\
&= b^j_{k,1}.
\end{aligned}$$

□

Recall that we know that Θ decomposes into a direct product of irreducibles via

$$\Theta \cong \bigoplus_{\rho=1}^{M} c_\rho \theta_\rho, \quad \dim(\theta_\rho) = n_\rho.$$

Thus, if $\mathcal{B}$ is a symmetry adapted basis, we see that for all $g \in G$,

$$[\Theta(g)]_{\mathcal{B}} = \begin{array}{c} \\ c_1\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \\ \\ \\ \\ \\ \\ c_M\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \end{array} \begin{array}{c} \begin{array}{cccccccccc} \overbrace{\quad}^{n_1} & \cdots & \overbrace{\quad}^{n_1} & \overbrace{\quad}^{n_2} & \cdots & \overbrace{\quad}^{n_i} & \cdots & \overbrace{\quad}^{n_M} & \cdots & \overbrace{\quad}^{n_M} \end{array} \\ \begin{pmatrix} \theta_1(g) & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_1(g) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \theta_2(g) & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \theta_i(g) & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \theta_M(g) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & \theta_M(g) \end{pmatrix} \end{array}$$

where each $\theta_i(g)$ is a block of size $n_i \times n_i$. Thus, we can see that, for the elements $b^j_{k,\ell}$, where $c_j \neq 0$, we have

$$[\Theta(b^j_{k,\ell})]_{\mathcal{B}} = \begin{pmatrix} \theta_1\left(b^j_{k,\ell}\right) & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_1\left(b^j_{k,\ell}\right) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \theta_2\left(b^j_{k,\ell}\right) & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \theta_i\left(b^j_{k,\ell}\right) & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \theta_M\left(b^j_{k,\ell}\right) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & \theta_M\left(b^j_{k,\ell}\right) \end{pmatrix}$$

$$
= \begin{pmatrix}
\ddots & & & & \\
& \begin{array}{c|cccc} & & & \ell & \\ \hline k & \ddots & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & \ddots \end{array} & & & \\
& & \ddots_{c_j} & & \\
& & & \begin{array}{c|cccc} & & & \ell & \\ \hline k & \ddots & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & \ddots \end{array} & \\
& & & & \ddots
\end{pmatrix}
$$

where the omitted entries are 0 and all nonzero blocks correspond to the c_j different images of θ_j in the direct product expansion of Θ. This operator will be very important, so let us define it as follows.

Definition 3.5. Let $D = \theta_1 \oplus \cdots \oplus \theta_M$ be a DFT for the group G. Then, the *k, ℓ slice operator associated with representation θ_j, $P_{k\ell}^{(j)}$*, is given by the image of the element $b_{k,\ell}^{j} \in \mathbb{C}G$ under the representation Θ. In particular,

$$P_{k\ell}^{(j)} = \Theta\left(b_{k,\ell}^{j}\right).$$

Let us analyze some properties of $P_{k\ell}^{(j)}$. First, note that our decomposition of Θ,

$$\Theta \cong \bigoplus_{\rho=1}^{M} c_\rho \theta_\rho,$$

implies a similar decomposition of V,

$$
\begin{aligned}
V &\cong \bigoplus_{\rho=1}^{M} \bigoplus_{i=1}^{c_\rho} V_\rho^i \\
&\cong \bigoplus_{\rho=1}^{M} V_\rho,
\end{aligned}
$$

where V_ρ^i is the irreducible subspace corresponding to the ith "copy" of θ_ρ. At this point, let us note that any symmetry adapted basis on V must be able to be expressed as a union of bases for the irreducible subspaces V_ρ^i, for all ρ, i. Thus, our symmetry adapted basis $\mathcal{B}$ can be described by a collection of bases for every value j of the index ρ $(1 \leq j \leq M)$, of the form

$$\begin{aligned} V_j^1 &= \left\langle v_{j1}^{1}, v_{j2}^{1}, \ldots, v_{jn_j}^{1} \right\rangle \\ V_j^2 &= \left\langle v_{j1}^{2}, v_{j2}^{2}, \ldots, v_{jn_j}^{2} \right\rangle \\ &\vdots \\ V_j^{n_j} &= \left\langle v_{j1}^{c_j}, v_{j2}^{c_j}, \ldots, v_{jn_j}^{c_j} \right\rangle, \end{aligned} \tag{3.1}$$

With this formulation of the basis $\mathcal{B}$, we can prove the following theorem.

Proposition 3.2. *The k, ℓ slice operator associated with irreducible θ_j obeys the following relationship, for the symmetry adapted basis $\mathcal{B}$:*

$$P_{k\ell}^{(j)}\left(v_{j'\ell'}^{\ m}\right) = \begin{cases} 0 & \text{if } j' \neq j \text{ or } \ell' \neq \ell \\ v_{j'k}^{\ m} = v_{jk}^{m} & \text{if } j' = j, \ell' = \ell. \end{cases}$$

Proof. This can be seen readily from an algebraic perspective, but perhaps easiest is simply to note the (known) matrix form for $P_{k\ell}^{(j)}$ under basis $\mathcal{B}$. We recall

$$[P_{k\ell}^{(j)}]_{\mathcal{B}} = [\Theta(b_{k,\ell}^{j})]_{\mathcal{B}}$$

$$= \begin{pmatrix} \ddots & & & \\ & \begin{array}{c|cccc|} \multicolumn{1}{c}{} & & & \ell & \multicolumn{1}{c}{} \\ \cline{2-5} & \ddots & 0 & 0 & 0 \\ k & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & \ddots \\ \cline{2-5} \end{array} & & \\ & & \ddots_{c_j} & \\ & & & \begin{array}{c|cccc|} \multicolumn{1}{c}{} & & & \ell & \multicolumn{1}{c}{} \\ \cline{2-5} & \ddots & 0 & 0 & 0 \\ k & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & \ddots \\ \cline{2-5} \end{array} \\ & & & & \ddots \end{pmatrix}$$

where the omitted entries are zero. With this formulation, the images of the basis vectors in $\mathcal{B}$ can be read off directly from the columns of $P_{k\ell}^{(j)}$. In particular, we see immediately that for any $j' \neq j$, $P_{k\ell}^{(j)} v_{j'}{}_{\ell'}^{m} = 0$, as $P_{k\ell}^{(j)}$ has columns of zeros for all spaces other than V_j. Further, if $\ell \neq \ell'$, then we see that even when restricted to the mth block of c_j, we will be selecting a column of all zeros. Thus, the only time we ever receive a nonzero output is when $j = j'$ and $\ell = \ell'$. In this case, we receive a column, when restricted to the block corresponding to V_j^m, with all 0s save a 1 at the kth position. Thus, this output is, in a basis independent sense, the vector $v_{j}{}_{k}^{m}$, as desired. □

Corollary. *$P_{k\ell}^{(j)}$ maps the spaces V_ρ to $\{0\}$ for all $\rho \neq j$, and maps V_j into itself. Furthermore, $P_{k\ell}^{(j)}$ has rank c_j.*

Proof. We see that V_ρ is spanned by all basis vectors in $\mathcal{B}$ of the form $v_{\rho}{}_{\ell}^{m}$. If $\rho \neq j$, then $P_{k\ell}^{(j)}$ maps all of these vectors to zero. The first part of the proposition follows. We can also see that $P_{k\ell}^{(j)}$ maps the set of basis vectors $v_{j}{}_{\ell}^{m}$, for $1 \leq m \leq c_j$ to a linearly independent set, and these are the only vectors that are mapped to nonzero results. Thus, the rank of $P_{k\ell}^{(j)}$ is c_j. □

Note a special case of these results: The slice operator $P_{kk}^{(j)}$ maps V_j onto V_j and is a true projection operator as it maps all basis vectors $v_{j}{}_{k}^{m}$ to themselves, and maps the space V into the space spanned by the c_j basis vectors in the kth column of Equation 3.1, which are exactly $v_{j}{}_{k}^{m}$, for $1 \leq m \leq c_j$.

Let $x \in V$ be a vector such that

$$P_{11}^{(j)} x = x_1 \neq 0 \tag{3.2}$$

(this will exist if and only if $c_j \neq 0$).

By the above, we have that $x_1 \in \left\langle v_{j}{}_{1}^{1}, \ldots, v_{j}{}_{1}^{c_j} \right\rangle$, so

$$x_1 = \alpha_1 v_{j}{}_{1}^{1} + \cdots + \alpha_{c_j} v_{j}{}_{1}^{c_j}.$$

Therefore, if we produce vectors x_μ, $2 \leq \mu \leq n_j$ via

$$x_\mu = P_{\mu 1}^{(j)} x_1,$$

we see by Theorem 3.2 that

$$x_\mu = \alpha_1 v_{j}{}_{\mu}^{1} + \cdots + \alpha_{c_j} v_{j}{}_{\mu}^{c_j}.$$

In particular, $x_\mu \in V_j^m$ and is composed of the *exact* same linear combination of the basis vectors $v_{j_\mu}^1, \ldots, v_{j_\mu}^{c_j}$ as x_1 is of $v_{j1}^1, \ldots, v_{j1}^{c_j}$. Let us note that from a module description, we can realize these vectors x_i in a different way: As $P_{k\ell}^{(j)}$ is merely $\Theta\left(b_{k,\ell}^j\right)$, their action on vectors in V is exactly the same as the action of $b_{k,\ell}^j$ on these vectors from a $\mathbb{C}G$-module perspective. Thus, in this way, we can also realize x_μ as

$$x_\mu = b_{\mu,1}^j \left(b_{1,1}^j x\right)$$

for the fixed x in 3.2. But, by the definition of the action in a module, we see that this reveals

$$\begin{aligned} x_i &= \left(b_{\mu,1}^j b_{1,1}^j\right) x \\ &= b_{\mu,1}^j x \end{aligned}$$

by Theorem 3.1.

As $P_{k\ell}^{(j)}$ has rank c_j, there exist c_j linearly independent vectors x^i that generate analogous, but linearly independent sequences of x_m^i as did the fixed element x. In this way, we generate the following sets:

$$\left\{b_{1,1}^j x^1, b_{2,1}^j x^1, \ldots, b_{n_j,1}^j x^1\right\}, \left\{b_{1,1}^j x^2, b_{2,1}^j x^2, \ldots, b_{n_j,1}^j x^2\right\}, \ldots, \left\{b_{1,1}^j x^{c_j}, b_{2,1}^j x^{c_j}, \ldots, b_{n_j,1}^j x^{c_j}\right\}.$$

Note that as all $n_j c_j$ of these linearly independent elements live in V_j, which has dimension $n_j c_j$, we can say that

$$\mathcal{A}_j = \left\{b_{1,1}^j x^1, b_{2,1}^j x^1, \ldots, b_{n_j,1}^j x^1, b_{1,1}^j x^2, b_{2,1}^j x^2, \ldots, b_{n_j,1}^j x^2, \cdots, b_{1,1}^j x^{c_j}, b_{2,1}^j x^{c_j}, \ldots, b_{n_j,1}^j x^{c_j}\right\}$$

is a basis for V_j. Moreover, we see that the collections

$$\{b_{1,1}^j x^i, \ldots, b_{n_j,1}^j x^i\}$$

span an n_j dimensional space, and there are exactly c_j such spaces. To characterize these spaces, let

$$\mathcal{A}_j^i = \left(b_{1,1}^j x^i, \ldots, b_{n_j,1}^j x^i\right)$$

denote our resulting basis, and let

$$U_j^i = \left\langle \mathcal{A}_i^j \right\rangle.$$

Performing this procedure for all j generates a basis for V, and analogous bases $\mathcal{A}_j^i$ and spaces U_j^i, for all j. Let us denote this total basis via

$$\mathcal{A} = \bigcup_{j=1}^{M} \mathcal{A}_j \tag{3.3}$$

with

$$\langle \mathcal{A} \rangle = V.$$

Let us analyze the images of these basis vectors for these U_j^i under $\Theta(g)$, for an arbitrary $g \in G$. We recall that, from a module theoretic perspective, the action $\Theta(g)x$ is given by gx, so we can see that

$$\begin{aligned}\Theta(g)\left(b_{\mu,1}^{j} x^{i}\right) &= g\left(b_{\mu,1}^{j} x^{i}\right) \\ &= g b_{\mu,1}^{j} x^{i}\end{aligned}$$

Now, we note that the product $gb_{\mu,1}^j$ has a very specific form. We see,

$$\begin{aligned} D\left(gb_{\mu,1}^{j}\right) &= (\theta_1(g), \ldots, \theta_h(g)) \times \left(\underbrace{(\mathbf{0}), \ldots, (\mathbf{0}),}_{k-1} \ \mu \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}, \underbrace{(\mathbf{0}), \ldots, (\mathbf{0})}_{h-k} \right) \\ &= \left(\underbrace{(\mathbf{0}), \ldots, (\mathbf{0}),}_{k-1} \begin{pmatrix} | & | & \cdots & | \\ \theta_j(g)_{\vec{\mu}} & 0 & \cdots & 0 \\ | & | & \cdots & | \end{pmatrix}, \underbrace{(\mathbf{0}), \ldots, (\mathbf{0})}_{h-k} \right) \end{aligned}$$

where by $\theta_j(g)_{\vec{\mu}}$, we mean the μth column of $\theta_j(g)$ (which, we recall, is a known matrix). But, we see that this is

$$D\left(gb^j_{\mu,1}\right) = \left(\underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{j-1}, \begin{pmatrix} \theta_j(g)_{1\mu} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \theta_j(g)_{\mu\mu} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \theta_j(g)_{n_j\mu} & \cdots & 0 & \cdots & 0 \end{pmatrix}, \underbrace{(\mathbf{0}),\ldots,(\mathbf{0})}_{M-j}\right)$$
$$= \sum_{m=1}^{n_j} \theta_j(g)_{m\mu}\hat{b}^j_{m,1}.$$

Thus, as D is a ring isomorphism,

$$gb^j_{\mu,1} = \sum_{m=1}^{n_j} \theta_j(g)_{m\mu} b^j_{m,1}.$$

So,

$$\Theta(g)b^j_{\mu,1} = \sum_{m=1}^{n_j} \theta_j(g)_{m\mu} b^j_{m,1}x^i$$

We recall that the space U^i_j is given by

$$U^i_j = \left\langle b^j_{1,1}x^i, \ldots, b^j_{n_j,1}x^i \right\rangle,$$

so,

$$\Theta(g)\left(b^j_{\mu,1}x^i\right) = \sum_{m=1}^{n_j} \theta_j(g)_{m\mu} b^j_{m,1}x^i \in U^i_j.$$

As g is arbitrary, and $b^j_{\mu,1}x^i$ an arbitrary basis vector of U^i_j, this shows that U^i_j is is invariant under Θ. Furthermore, let us express $\Theta(g)$ in matrix form under the given basis of U^i_j. We recall from basic linear algebra that

$$\left[\Theta(g)|_{U^i_j}\right]_{\mathcal{A}^i_j} = \left(\left[\left(\Theta(g)|_{U^i_j}\right) b^j_{1,1}x^i\right]_{\mathcal{A}^i_j} \quad \left[\left(\Theta(g)|_{U^i_j}\right) b^j_{2,1}x^i\right]_{\mathcal{A}^i_j} \quad \cdots \quad \left[\left(\Theta(g)|_{U^i_j}\right) b^j_{n_j,1}x^i\right]_{\mathcal{A}^i_j} \right)$$

where each entry in this matrix denotes a column vector. But, As U^i_j is invariant under Θ, we have already computed these products, with

$$\Theta(g)\left(b^j_{\mu,1}x^i\right) = \sum_{m=1}^{n_j} \theta_j(g)_{m\mu} b^j_{m,1}x^i.$$

Thus, the μth column vector of $(\Theta(g)|_{U_j^i})$ is given by

$$\begin{aligned}
\left[\left(\Theta(g)|_{U_j^i}\right) b^j_{\mu,1} x^i\right]_{\mathcal{A}_j^i} &= \left[\sum_{m=1}^{n_j} \theta_j(g)_{m\mu} b^j_{m,1} x^i\right]_{\mathcal{A}_j^i} \\
&= \sum_{m=1}^{n_j} \theta_j(g)_{m\mu} \left[b^j_{m,1} x^i\right]_{\mathcal{A}_j^i} \\
&= \begin{pmatrix} \theta_j(g)_{1\mu} \\ \theta_j(g)_{2\mu} \\ \vdots \\ \theta_j(g)_{n_j\mu} \end{pmatrix}.
\end{aligned}$$

So,

$$\begin{aligned}
\left[\Theta(g)|_{U_j^i}\right]_{\mathcal{A}_j^i} &= \begin{pmatrix} \theta_j(g)_{11} & \theta_j(g)_{12} & \cdots & \theta_j(g)_{1n_j} \\ \theta_j(g)_{21} & \theta_j(g)_{22} & \cdots & \theta_j(g)_{2n_j} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_j(g)_{n_j1} & \theta_j(g)_{n_j2} & \cdots & \theta_j(g)_{n_jn_j} \end{pmatrix} \\
&= \theta_j(g).
\end{aligned} \tag{3.4}$$

Note that this holds *for all* i. In particular, we see that the spaces U_j^i correspond *exactly* to the c_j copies of irreducible representation θ_j, with isomorphic copies each equal to $\theta_j(g)$. This reveals the following theorem.

Theorem 3.2. *Let G, (Θ, V), be given as before. Then, the basis $\mathcal{A}$ constructed in Equation 3.3 is symmetry adapted.*

Proof. Given the argument in the preceding paragraph, this is immediate. By definition, a basis for a representation (Θ, V) is symmetry adapted if it renders Θ in block diagonal form, with blocks corresponding to isomorphic irreducible representations actually equal. But, this is precisely what was shown in 3.4. □

Note that this construction relies *explicitly* on the ability to actually construct the primitive idempotents. To do so, recall the following theorem (See Clausen and Baum (1993), page 84).

Theorem 3.3. *Let G be a finite group with irreducible representations $\theta_1, \ldots, \theta_M$, each of dimension $n_1, \ldots, n_M$. Then, if $a \in \mathbb{C}G$,*

$$a(g) = \frac{1}{|G|} \sum_{i=1}^{M} n_i \operatorname{tr}\left(\theta_i(g^{-1})\theta_i(a)\right).$$

With this theorem, we can directly compute the dual matrix coefficients. We see

$$
\begin{aligned}
b^j_{k,\ell}(g) &= \frac{1}{|G|}\sum_{i=1}^{M} n_i \operatorname{tr}\left(\theta_i(g^{-1})\theta_i(b^j_{k,\ell})\right) \\
&= \frac{n_j}{|G|}\operatorname{tr}\left(\theta_j(g^{-1})\theta_j(b^j_{k,\ell})\right) \\
&= \frac{n_j}{|G|}\left(\theta_j(g^{-1})\right)_{\ell,k}.
\end{aligned}
$$

Thus,

$$
b^j_{k,\ell} = \frac{n_j}{|G|}\sum_{g\in G}\left(\theta_j(g^{-1})\right)_{\ell,k} g.
$$

With this, we now know how to construct the dual matrix coefficient functions, and hence the adapted basis $\mathcal{A}$. We can thus now present the following algorithm. Note that, for ease of use computationally, we will focus on the slice operators $P^j_{k,\ell} = \Theta(b^j_{k,\ell})$.

3.2.2 Final Algorithm

Input A finite group G, such that the M irreducible, pairwise inequivalent complex representations of G, denoted θ_ρ, are known and presented in matrix form. A complex representation (Θ, V) is also given such that Θ is of dimension n and $\Theta : g \mapsto \Theta(g)$. We also presume that the decomposition of Θ is known:

$$
\Theta = \bigoplus_{\rho\leq M} c_\rho\theta_\rho.
$$

This decomposition also decomposes V into invariant subspaces

$$
V \cong \left(V^1_1 \oplus V^1_2 \cdots V^1_{c_1}\right) \bigoplus \left(V^2_1 \oplus \cdots V^2_{c_2}\right) \bigoplus \cdots \bigoplus \left(V^M_1 \oplus V^M_{c_M}\right),
$$

where $V^j_i \cong V^j_k$ for all i, k, j. Now, perform the following steps for each irreducible representation θ_j, $j \leq M$. As these steps are independent, throughout the rest of the algorithm, let j be a fixed, but arbitrary index of the irreducible representation θ_j. This will enable some useful notational freedom, by omitting notational dependence on j of objects that, in fact, do depend on j.

Step 1—The Projection Matrices Compute the projection matrices

$$\pi^{(j)} = \frac{|G|}{n_j}\Theta(b^j_{1,1})$$
$$= \sum_{g\in G}\left(\theta_j(g^{-1})\right)_{1,1}\Theta(g)$$

Step 2—Extracting the Columns The image of the matrix $\pi^{(j)}$ is a c_j dimensional space, spanned by the columns of $\pi^{(j)}$. *Choose a basis for this space.* Let this basis be denoted by $v^1_1, v^2_1, \ldots, v^{c_j}_1$.

Step 3—The Remaining Sections Compute the slice operators

$$P^{(j)}_{\mu,1} = \Theta\left(b^j_{\mu,1}\right)$$
$$= \frac{n_j}{|G|}\sum_{g\in G}\left(\theta_j(g^{-1})\right)_{1,\mu}\Theta(g)$$

for all $\mu \in \{2, 3, \ldots, n_j\}$.

Step 4—Forming the Basis Form an adapted basis for each n_j dimensional irreducible subspace V^i_j, where $i \leq c_j$, by acting on the vectors v_1 with the operators $P^{(j)}_{\mu,1}$ for all $\mu \in \{2, \ldots, n_j\}$. Specifically, let

$$v^i_\mu = P^{(j)}_{\mu,1}v^i_1.$$

for all $\mu \in \{2, \ldots, n_j\}$. Then, the spaces V^j_i are spanned by the vectors $\{v^i_1, v^i_2, \ldots, v^i_{n_j}\}$.

Output The bases produced in this way for all irreducible subspaces indexed by each $j \leq \rho$, arranged in any order.

The proof of correctness of this algorithm is immediate as it follows directly from the previous two sections. Note that this algorithm is presented in a different formulation in Stiefel and Fässler (1992).

It is important to note that in **Step 2** there is an aspect of unspecified *choice*. Specifically, any mechanism of finding a basis for the span of the columns of $\pi^{(j)}$ will still yield an adapted basis, though some of those mechanisms may result in bases that enable much faster conversion than others. In order to counter this ambiguity in this work, we produced code to analyze a variety of choices and the resultant bases. In particular, even when using Gram-Schmidt to produce an orthonormal basis of the column space of $\pi^{(j)}$, the initial ordering of the delta basis $\mathcal{C}$ of $\mathbb{C}X^{\lambda}$ changes the final output. A large number of possible permutations of this choice were tested; none made any significant difference on the speed of the resulting transformation.

Recall that it is this adapted basis that we will use to decompose our dataset into meaningful parts. With this algorithm, we now know how to compute the final basis for our transformation, and thus the actual decomposition of the space can be performed using standard, linear algebra, change of basis transformations. In the remainder of this work, we focus on how to make these transformations faster.

3.3 Theoretical Tools and Transformation Efficiency

The algorithm concluding the previous section will enable us to compute a symmetry adapted basis for *any* S_n and $\mathbb{C}X^{\lambda}$. However, our primary concern here is not whether or not we can find a symmetry adapted basis, but rather how quickly we can transform from the delta basis to a symmetry adapted basis. The transformation matrix from $\mathcal{C}$ to $\mathcal{A}$ immediately, in "one jump," is fixed as soon as $\mathcal{A}$ is specified. However, we do not necessarily need to compute this change of basis in just one jump. In particular, the strategy we will take here is to use the fact that the DFT we use here is naturally adapted to the chain $S_1 \leq S_2 \leq \cdots \leq S_n$ to produce intermediate, adapted bases $\mathcal{A}_j$ which is symmetry adapted for S_j, and then compute the transformation factors given by the change of basis matrices from $\mathcal{A}_{j-1}$ to $\mathcal{A}_j$.

$$F_j(\lambda) = M_{\mathcal{A}_j \leftarrow \mathcal{A}_{j-1}}$$

It will almost always be the case that due to the only incremental differences between S_j and S_{j+1} that $F_j(\lambda)$ (often denoted F_j if the context is clear) will be very sparse. In order to prove real bounds on this sparseness, consider the following.

Definition 3.6. Given $G, D, \left(b^k_{i,j}\right)$ as before, the *primitive idempotents* for $\mathbb{C}G$ are given by the elements $b^k_{i,i}$ for all applicable i, k. Note that these are true idempotents, as their matrix images square to themselves.

Let X be a finite set on which G acts transitively. Then, consider the collection $\{b^k_{i,i}\mathbb{C}X\}_{i,k}$. We see that this must span all of $\mathbb{C}X$ as $\sum_{i,k} b^k_{i,i} =$

$1_{\mathbb{C}G}$. But, further, as $b^k_{i,i} b^d_{j,j} = \delta_{i,j}\delta_{k,d}$, we see that these actually separate the space into a collection of disjoint subspaces. Let us thus make the following definition.

Definition 3.7. Given $G, X, \left(b^k_{i,i}\right)_{i,k}$ as above, we will refer to the component $b^k_{i,i}\mathbb{C}X$ as the *frequency space* of $\mathbb{C}X$ associated to $b^k_{i,i}$.

Theorem 3.4. *Let X be a finite set and G be a finite group such that G acts on X transitively. Let $H \leq G$. Then the frequency spaces of ${}_{\mathbb{C}H}\mathbb{C}X$ are direct sums of the frequency spaces of ${}_{\mathbb{C}G}\mathbb{C}X$.*

Proof. Let us first clarify some notation. We will let $\theta_1, \ldots, \theta_M$ be a complete set of irreducible representations of G, such that the associated G-DFT $D = \theta_1 \oplus \cdots \oplus \theta_M$ is adapted to the chain $H \leq G$. Let $b^k_{i,i}$ be the associated primitive idempotents. Let $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_{\widetilde{M}}$ be a complete set of irreducible representations of H, with $\widetilde{D}$ the associated H-DFT, and $\widetilde{b}^k_{i,i}$ the primitive idempotents. We will presume that the matrix forms for $\theta_\rho, \widetilde{\theta}_\rho$ are known for all ρ

With that notation in mind, we begin by expressing the frequency spaces of ${}_{\mathbb{C}H}\,\mathbb{C}\mathrm{X}$. We see that they are given by $\widetilde{b}^k_{i,i}\,\mathbb{C}\mathrm{X}$, for some choice i, k. What is $\widetilde{b}^k_{i,i}$? It is an element such that

$$\widetilde{\theta}_k\left(\widetilde{b}^k_{i,i}\right) = \begin{array}{c} \\ i \\ \\ \end{array} \overset{\qquad\qquad\quad i}{\begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}}.$$

Recall that the adaptivity of the DFT D implies that $\theta_j\big|_{\mathbb{C}H}$ is in block diagonal form with blocks corresponding to irreducible representations of H such that all blocks associated with representation $\widetilde{\theta}_\rho$ are given by $\widetilde{\theta}_\rho(h)$, for all $h \in H$.

Throughout the rest of the proof, let k represent a fixed, but arbitary, index of the irreducible representation $\widetilde{\theta}_k$ of H and i represent a fixed, but arbitrary index of a primitive idempotent with respect to $\widetilde{\theta}_k$. We fix these values for notational convenience; we will now often omit any notational dependence on k and i for certain constants that, in fact, do depend on which primitive idempotent of H we are examining. But, as k and i remain arbitrary, this can be done without loss of generality, and result will still follow from a much cleaner presentation.

Then, the adaptivity of D implies

$$\theta_j\left(b_{h\,i,i}^{\;k}\right) = \begin{pmatrix} \ddots & 0 & 0 & 0 & 0 \\ 0 & \widetilde{\theta}_\ell\left(\widetilde{b}_{i,i}^{k}\right) & 0 & 0 & 0 \\ 0 & 0 & \widetilde{\theta}_m\left(\widetilde{b}_{i,i}^{k}\right) & 0 & 0 \\ 0 & 0 & 0 & \widetilde{\theta}_n\left(\widetilde{b}_{i,i}^{k}\right) & 0 \\ 0 & 0 & 0 & 0 & \ddots \end{pmatrix}$$

$$= \begin{pmatrix} \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{\begin{matrix} \ddots & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{matrix}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{\begin{matrix} \ddots & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{matrix}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots \end{pmatrix}$$

where each nonzero block corresponds to $\widetilde{\theta}_k$ and the singular 1 in each nonzero block appears at position (i,i) within that block. If θ_j, when viewed as a representation of H by restriction, contains $\widetilde{c}_j$ copies of $\widetilde{\theta}_\rho$, then there will be exactly $\widetilde{c}_j$ such blocks. As all blocks are square, the 1s contained in each block also appear on the diagonal of the overall matrix $\theta_j\left(\widetilde{b}_{i,i}^{k}\right)$. Let the indices of these 1s with respect to the full matrix be given by $\left(I_1^j, I_1^j\right), \ldots, \left(I_{\widetilde{c}_j}^j, I_{\widetilde{c}_j}^j\right)$. Note these indices exist for all j such that $\widetilde{c}_j > 0$; if $\widetilde{c} = 0$, we will merely presume the set of such indices is empty, and any sums over the indices in this case would return zero. Then, we can see that

$$\theta_j\left(\widetilde{b}_{i,i}^{k}\right) = \theta_j\left(b^j_{I_1^j, I_1^j} + \cdots + b^j_{I_{\widetilde{c}_j}^j, I_{\widetilde{c}_j}^j}\right).$$

Note the following. By the definition of the primitive idempotents,

$$\theta_j\left(b^{\rho}_{\ell,\ell}\right) = 0,$$

for all $\rho \neq j$. Therefore, we can also write

$$\theta_j\left(\widetilde{b}^k_{i,i}\right) = \theta_j\left(\sum_{\rho=1}^{M}\left(b^{\rho}_{I^{\rho}_1,I^{\rho}_1} + \cdots + b^{\rho}_{I^{\rho}_{\tilde{c}_\rho},I^{\rho}_{\tilde{c}_\rho}}\right)\right).$$

Note that this identification will hold for all j. Thus, if we examine the image of $\tilde{b}^k_{i,i}$ under the G-DFT D, we see

$$D\left(\widetilde{b}^k_{i,i}\right) = D\left(\sum_{\rho=1}^{M}\left(b^{\rho}_{I^{\rho}_1,I^{\rho}_1} + \cdots + b^{\rho}_{I^{\rho}_{\tilde{c}_\rho},I^{\rho}_{\tilde{c}_\rho}}\right)\right).$$

But, D is an isomorphism. Thus,

$$\widetilde{b}^k_{i,i} = \sum_{\rho=1}^{M}\left(b^{\rho}_{I^{\rho}_1,I^{\rho}_1} + \cdots + b^{\rho}_{I^{\rho}_{\tilde{c}_\rho},I^{\rho}_{\tilde{c}_\rho}}\right).$$

But, by the linearity of the action of $\mathbb{C}H$ or $\mathbb{C}G$ on $\mathbb{C}X$, this implies

$$\widetilde{b}^k_{i,i}\,\mathbb{C}X \cong \bigoplus_{\rho=1}^{M}\left(b^{\rho}_{I^{\rho}_1,I^{\rho}_1}\,\mathbb{C}X + \cdots + b^{\rho}_{I^{\rho}_{\tilde{c}_\rho},I^{\rho}_{\tilde{c}_\rho}}\,\mathbb{C}X\right).$$

As k and i were arbitrary, this shows that the frequency spaces of H are direct sums of those with respect to G. □

Note that this theorem implies that in order to transform from an adapted basis for H to an adapted basis for G, we only need to transform all of the frequency spaces of H! This follows directly from the fact that the frequency spaces over H are direct sums of those over G, and the frequency spaces naturally contain both the adapted basis for H and for G, by 3.3. This reveals the following theorem.

Theorem 3.5. *Let G be a finite group with subgroup H and let D be a DFT for G adapted to the chain $H \leq G$. Let X be a finite set of options, such that G acts transitively on X. Let the frequency spaces of $\mathbb{C}X$ with respect to H have dimensions $\alpha_1, \ldots, \alpha_f$. Then, given an adapted basis $\mathcal{B}_H$ over H and an adapted basis $\mathcal{B}_G$, the change of basis matrix $M_{\mathcal{B}_H \leftarrow \mathcal{B}_G}$ requires at most $\sum_{i=1}^{f} \alpha_i^2$ nonzero entries, or at most $\max_{1\leq i\leq f} \alpha_f$ nonzero entries per row/column.*

Proof. This stems directly from the previous theorem and our observation. By the definition of an adapted basis, we see that $\mathcal{B}_G$ and $\mathcal{B}_H$ can be written as unions of bases for the frequencey spaces of $\mathbb{C}X$ with respect to G or H, respectively. As the frequency spaces of G are direct sums of those of H, to perform a change of basis from $\mathcal{B}_H$ to $\mathcal{B}_G$ will never require more work than would performing a change of basis from the frequency spaces of H to the frequency spaces of G locally. These have dimension $\alpha_1, \ldots, \alpha_f$ here, and thus performing the necessary f change of basis operations on these spaces requires at most $\sum_{i=1}^{f} \alpha_i^2$ nonzero entries per row/column. It is clear that this also implies that it requires no more than $\max_{1 \leq i \leq f} \alpha_f$ nonzero entries per row/column. □

These theorems will lay the groundwork for our investigations into the complexity of computing the various factors F_j in the full decomposition with which we work. In particular, there we will mostly task ourselves with trying to compute the dimensions of the frequency spaces over the decomposition of S_{j-1} in S_j. To make this result even more useful, let us characterize exactly what the dimensions of these frequency spaces are.

Theorem 3.6. *Let G be a finite group, with $H \leq G$ and D a G-DFT adapted to the chain $H \leq G$. Let H have irreducible representations given by $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_{\widetilde{M}}$, with associated irreducible modules given by N_i. Suppose G acts transitively on a finite set X. Then let $X = X_1 \cup X_2 \cup \ldots \cup X_\omega$ be the orbital decomposition of X with respect to H. This orbital decomposition inherently produces a module decomposition of ${}_{\mathbb{C}H}\mathbb{C}X$ via*

$${}_{\mathbb{C}H}\mathbb{C}X = \bigoplus_{i=1}^{\omega} {}_{\mathbb{C}H}\mathbb{C}X_i\,.$$

Let the irreducible decomposition of each of these modules be given by

$${}_{\mathbb{C}H}\mathbb{C}X_i \cong \bigoplus_{\rho=1}^{\widetilde{M}} \alpha_\rho^i N_\rho.$$

Then, the dimension of the frequency space $\widetilde{b}_{i,i}^k\,\mathbb{C}X$ is given by

$$\dim\left(\widetilde{b}_{i,i}^k\,\mathbb{C}X\right) = \sum_{i=1}^{\omega} \alpha_k^i.$$

Furthermore, there are exactly $\dim(N_k)$ *copies of this frequency space.*

Proof. The proof of this theorem is relatively straightforward. Let us expand the definition of the frequency space:

$$\begin{aligned}
\widetilde{b}^{k}_{i,i}\,\mathbb{C}X &= \widetilde{b}^{k}_{i,i} \bigoplus_{i=1}^{\omega} {}_{\mathbb{C}H}\,\mathbb{C}X_i \\
&= \bigoplus_{i=1}^{\omega} \widetilde{b}^{k}_{i,i\,\mathbb{C}H}\,\mathbb{C}X_i \\
&= \bigoplus_{i=1}^{\omega} \widetilde{b}^{k}_{i,i} \left(\bigoplus_{\rho=1}^{\widetilde{M}} \alpha^{i}_{\rho} N_\rho \right) \\
&= \bigoplus_{i=1}^{\omega} \widetilde{b}^{k}_{i,i} \left(\bigoplus_{\rho=1}^{\widetilde{M}} \bigoplus_{m=1}^{\alpha^{i}_{\rho}} N_\rho \right) \\
&= \bigoplus_{i=1}^{\omega} \bigoplus_{\rho=1}^{\sim} \bigoplus_{m=1}^{\alpha^{i}_{\rho}} \widetilde{b}^{k}_{i,i} \left(N_\rho \right)
\end{aligned}$$

where in the last step we recognize the symbolic meaning of the product $\alpha^{i}_{\rho} N_\rho$. But, by the construction of the $b_{h}{}^{k}_{i,i}$, we see that $b_{h}{}^{k}_{i,i} N_\rho = 0$ if $\rho \neq k$ and is a one dimensional space otherwise. Thus,

$$\widetilde{b}^{k}_{i,i}\,\mathbb{C}X = \bigoplus_{i=1}^{\omega} \bigoplus_{m=1}^{\alpha^{i}_{k}} \widetilde{b}^{k}_{i,i} N_k$$

But, as this is a collection of one-dimensional spaces, we can find its dimension immediately. We see that it has dimension

$$\dim \left(b_{h}{}^{k}_{i,i}\,\mathbb{C}X \right) = \sum_{i=1}^{\omega} \alpha^{i}_{k},$$

as desired. The second statement in the theorem is merely a recognition that there are exactly $\dim \left(N_\rho \right)$ primitive idempotents associated with N_ρ, namely $\widetilde{b}^{\rho}_{1,1}\,\mathbb{C}X$ through $\widetilde{b}^{\rho}_{\dim\left(N_\rho\right),\dim\left(N_\rho\right)}\,\mathbb{C}X$. This proves the theorem. □

Note that all of these theorems about the decomposition of $H \leq G$ can be extended to arbitrary finite chains, $G_0 \leq G_1 \leq \cdots \leq G_n \leq G$, as the orbits of each G_i will nest within those of G_{i+1}, so that we can always restrict our attention to a single orbit of G_{i+1}, and then examine the decomposition of G_i restricted to this orbit exactly as in these theorems.

We now recall that the dimensions and multiplicities of the irreducible representations of S_n are both well understood, through standard tableaux and Kostka numbers, respectively. As such, Theorems 3.5 and 3.6 give us an immediate way to bound the complexity of transforming an adapted basis from S_j to S_{j+1}, within the chain $S_1 \leq S_2 \leq \cdots \leq S_n$ over X^λ for $\lambda \vdash n$. In particular, we see that we will obtain a rough bound for the complexity required by any factor $F_j = M_{\mathcal{B}_j \leftarrow \mathcal{B}_{j-1}}$ as follows.

First, we decompose the space $\mathbb{C}X^\lambda$ into orbits of S_j. Recall from Theorem 2.2 that this decomposition is given by

$$_{\mathbb{C}S_j}\mathbb{C}X^\lambda \cong \bigoplus_{\eta \in \lambda^{*n-j}} {}_{\mathbb{C}S_j}\mathbb{C}X^\eta$$

Each of these orbits is thus described by the transitive action of S_j on the set X^η. Further, we see that each space $_{\mathbb{C}S_j}\mathbb{C}X^\eta$ decomposes further into orbits of S_{j-1}, via

$$_{\mathbb{C}S_{j-1}}\mathbb{C}X^\eta \cong \bigoplus_{\nu \in \eta^*} {}_{\mathbb{C}S_{j-1}}\mathbb{C}X^\nu .$$

The union of these trivially intersecting submodules of S_{j-1} will yield the total module $_{\mathbb{C}S_{j-1}}\mathbb{C}X^\lambda$. But, we know that our adapted basis for S_{j-1} will respect its orbital decomposition, as the orbits of S_{j-1} span invariant subspaces of $_{S_{j-1}}\mathbb{C}X^\lambda$, so we can describe any adapted basis for S_{j-1} as a collection of adapted bases for each orbit of S_{j-1}. The same holds for S_j. Thus, the desired change of basis operator, F_j, can be realized as a direct sum of operators on each orbit of S_j. But, each of these transformations can be bounded with the frequency space dimensions, via Theorem 3.5. These frequency space dimensions, by Theorem 3.6, will be given exactly by sums of Kostka numbers across the orbits of S_{j-1} that are associated to the same irreducible representation. As these Kostka numbers can, in principle, be computed directly, this gives a mechanism to provide a total complexity bound for using this subgroup chain factorization of a change of basis over $\mathbb{C}X^\lambda$ from the data collection basis to an adapted basis, in particular, basis $\mathcal{A}$. We will explore this further in Chapter 4, and, in particular, in Theorem 4.1.

3.4 Implementation

This algorithm permits us the ability to decompose our data space with respect to each $S_j \leq S_n$ first, then doing incremental transformations to the data rather than a full transform all at once. But, in order to see how fast it truly performs, we must run experiments. To gather this experimental data about the complexity of the factors F_j and to test various other strategies with the algorithm in general, code was written in `C++`, first providing merely an implementation of the algorithm discussed in Section 3.2, then extending it in several ways that enabled greater experimentation. Here, we will give a brief introduction to the code base.

3.4.1 Object Design

To implement these algorithm, code was written in `C++`, providing functionality to express and manipulate elements of S_n, to create and represent tabloids for any $\lambda \vdash n$, to explore the actions of S_n on $\mathbb{C}X^\lambda$, to produce the matrix forms of the irreducible representations θ_ρ of S_n, and to produce a final, symmetry adapted basis for all $j \leq n$ in the context of a given space $\mathbb{C}X^\lambda$. This code has three major classes that simplify the problem:

`SymGpElm` Class This class provides access to a symmetric group element object, capable of multiplying against other symmetric group element objects, being decomposed to a product of adjacent transpositions, being extended to an element of a larger symmetric group or restricted to an element of a smaller symmetric group. Finally, and most importantly, these objects can act on elements of the `LambdaTableau` class to produce other elements of the `LambdaTableau` class. In this way, we explicitly specify the action of S_n on X^λ.

Elements $\sigma \in S_n$ are realized in implementation by explicit storage of n and an explicit map from $\{1, \ldots, n\} \to \{1, \ldots, n\}$, stored through a permuted list such that `map[i] = `σ`(i)`. It is important to note that, though relevant, these implementation details are far from the driving cost on the efficiency of this code and as such are less important than the overall implementation details of the algorithm.

`LambdaTablaeu` Class This class provides access to tabloids of shape $\lambda \vdash n$. Note that this is a strict requirement; the class *does not* provide access to tabloids over tabloids of shape $\lambda \vDash n$. Note also that this implementation does not provide a mechanism to take formal linear combinations of tabloids. This may seem counterintuitive, as $\mathbb{C}X^{\lambda}$ is composed of exactly formal linear combinations of elements of X^{λ} with coefficients in $\mathbb{C}$; however, as all of our efforts can restrict their attention to the basis of $\mathbb{C}X^{\lambda}$ given by X^{λ} whenever it needs to deal with real tabloids and as such we do not need any explicit implementation of $\mathbb{C}X^{\lambda}$. There are several common computational representational strategies for tabloids; in this work, tabloids are represented by a vector of length n cataloging the rows in which various numbers appear. More explicitly, in the code, the tabloid $t \in X^{(3,2)}$ given by

$$t = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 & \\ \cline{1-2}\end{array}$$

would be represented by the map

```
t.pos_ = [1,1,1,2,2].
```

`Matrix` **Class** This class merely enables basic linear algebra constructions, Gram-Schmidt orthogonalization, and some documentation features. The code previously used the Boost Basic Linear Algebra library for these computations, but this has been phased out of the system completely and is no longer required.

3.4.2 Algorithm Flow

Elements of the design are only interesting and useful in so much as they aid the global algorithm design. The current version of the code implements the following algorithm:

Input A choice $n \in \mathbb{N}$, $\lambda \vdash n$ which indicate that we seek a change of basis of the space $\mathbb{C}X^{\lambda}$ which respects the action of S_n.

Step 1 Compute the final, symmetry adapted basis for S_i for all $i \leq n$. This step is done following the algorithm described in Section 3.2 and Stiefel and Fässler (1992) to construct a symmetry adapted basis given a finite group. Specifically, however, we compute these bases such that they are also orthogonal, by modifying the “choose” step by using Gram-Schmidt to select our c_j linearly independent columns. One amendment made to ease experimentation is that there is a mechanism in the code to permute all possible starting orderings, of the set X^{λ}, so as to artificially enable *different choices* to be made at the step of unspecified *choice*.

Step 2 Compute the change of basis matrices $S_{i-1} \to S_i$ for all $2 \leq i \leq n$. In this way, we choose the decomposition of the global change of basis matrix $S_1 \to S_n$ by factorizing it as $(S_{n-1} \to S_n) \cdots (S_2 \to S_3)(S_1 \to S_2)$.

Output The basis found in **Step 1**, factorization found in **Step 2**, and the number of total nonzero entries across all factors of the factorization in **Step 2**.

Note that, though important for the viability of experimentation, the complexity of this algorithm to *produce* this final basis and decomposition is not of prime importance to this work; our focus is instead on minimizing the complexity of changing from the starting basis to the final basis, measured in particular by the number of nonzero entries in the final decomposition.

Chapter 4

Results

Thus far, we have devised an algorithm to perform fast data analysis on partially ranked data, but performing a change of basis on the data space, from a basis natural for data collection, to a basis natural for data analysis. We factored this change of basis transformation into a product of smaller transformations, each of which captured more symmetry on the space. Now, we are ready to begin analyzing the complexity of these algorithms, both experimentally and theoretically, which will show their viability for real-world applications.

4.1 General Tabloids

We begin with a definition, that will guide our discussion onto general complexity:

Definition 4.1. Let λ be partition of n. Then, let

$O(\lambda) = \{\rho | \rho \text{ is obtainable by removing any number of boxes from any rows of } \lambda\}$.

We can see that $O(\lambda)$ is the set of all tabloids describing orbits or sub-orbits of λ. Next, define $K(\lambda)$ be the maximum over all tabloids in $O(\lambda)$ of the maximum sum of the Kostka numbers associated to the irreducible decomposition of the orbits of the given tabloid shape in $O(\lambda)$. In particular,

$$K(\lambda) = \max_{\mu \in O(\lambda)} \left(\max_{\rho \trianglerighteq \mu^*} \left(\sum_{t \in \mu^*} \kappa_{\rho,t} \right) \right)$$

where the '$\trianglerighteq$' symbol here denotes the dominance order. We recall that the notation μ^* denotes the set of all orbits of μ under the action of $S_{\sum \mu - 1}$. Or, equivalently, $\mu^* = \{\mu^i | \mu_i > 0\}$. Further, we use notation $\rho \trianglerighteq \mu^*$ to mean $\rho \trianglerighteq \mu'$, for some $\mu' \in \mu^*$. Finally, we recall that $\kappa_{\lambda_1,\lambda_2} = 0$ if $\lambda_1 \not\trianglerighteq \lambda_2$.

Note that another way to define $K(\lambda)$ is as the maximum dimension of a frequency space associated any orbit or sub-orbit of λ over any group S_j, $j \leq n$.

With this definition, we will prove the following theorem that will simplify our investigations into specific tabloid decompositions.

Definition 4.2. Let $F_j(\lambda)$ represent the jth factor in the overall S_n change of basis. Then, let $c_j(\lambda)$ denote the average number of the nonzero entries per row/column in $F_j(\lambda)$. In particular, define

$$c_j(\lambda) = \frac{\text{The number of nonzero entries in } F_j(\lambda)}{\dim\left(\mathbb{C}X^\lambda\right)}.$$

Theorem 4.1. *Let $\lambda \vdash n$ be given. Then, $c_j(\lambda) \leq K(\lambda)$.*

Proof. Though quite general in statement, the proof of this theorem is surprisingly straightforward. We see that ${}_{\mathbb{C}S_j}\mathbb{C}X^\lambda$ decomposes into a collection of orbits. In particular,

$${}_{\mathbb{C}S_j}\mathbb{C}X^\lambda \cong \bigoplus_{\mu \in \lambda^{*^{n-j}}} \mathbb{C}X^\mu.$$

But, each of these orbits in turn decomposes into at most k orbits over S_{j-1}. In particular, for any given $\mu \in \lambda^{*^{n-j}}$, we have

$${}_{\mathbb{C}S_{j-1}}\mathbb{C}X^\mu \cong \bigoplus_{\rho \in \mu^*} \mathbb{C}X^\rho. \tag{4.1}$$

But, the bases $\mathcal{B}_j$ and $\mathcal{B}_{j-1}$ are, by construction, adapted, so they will still break apart into our orbit decomposition. But, this means that our factors $F_j(\lambda)$ will be expressible as a direct sum of change of basis operators over each orbit. In particular, we have that

$$F_j(\lambda) \cong \bigoplus_{\mu \in \lambda^{*^{n-j}}} F_j^\mu(\lambda).$$

But, by the orbit decomposition expressed in Equation 4.1, each orbit-operator is a map described via $F_j^\mu(\lambda) : \left(\bigoplus_{\rho \in \mu^*} \mathbb{C}X\rho\right) \to \mathbb{C}X^\mu$. So, each $F_j^\mu(\lambda)$

describes a transformation over a transitive set ($\mathbb{C}X^\mu$) from the set of orbits of $\mu \in O(\lambda)$. Therefore, the irreducible decomposition of each of the spaces $\mathbb{C}X^\rho$ has maximum sum of Kostka number bounded above by $K(\lambda)$. Said alternatively, the maximum frequency space dimension transformed by $F_j^\mu(\lambda)$ is bounded above by $K(\lambda)$. Expressed differently, this says that the maximum number of nonzero entries per row/column in the transformation of $F_j^\mu(\lambda)$ is bounded above by $K(\lambda)$. But, this applies to all $\mu \in \lambda^{*^{n-j}}$, which therefore means that this bound holds for $F_j(\lambda)$ in total. As j is general, this proves the theorem. □

Corollary. *Let κ be the maximum Kostka number attained in the decomposition of any tabloid $\mu \in O(\lambda)$. Then, $K(\lambda) \leq |\lambda| \cdot \kappa$*

Proof. We see the following.

$$\begin{aligned} K(\lambda) &= \max_{\mu \in O(\lambda)} \left(\max_{\rho \trianglerighteq \mu^*} \left(\sum_{t \in \mu^*} \kappa_{\rho,t} \right) \right) \\ &\leq \max_{\mu \in O(\lambda)} \left(\max_{\rho \trianglerighteq \mu^*} \left(\sum_{t \in \mu^*} \kappa \right) \right) \\ &\leq \max_{\mu \in O(\lambda)} \left(\max_{\rho \trianglerighteq \mu^*} |\lambda|\kappa \right) \\ &= |\lambda|\kappa \end{aligned}$$

as desired. □

Corollary. *Given any tabloid of shape $\lambda \vdash n$, with $|\lambda| \leq k$, the full factorization of change of basis matrix requires no more than*

$$\dim\left(\mathbb{C}X^\lambda\right)(n-1)K(\lambda) \leq \dim\left(\mathbb{C}X^\lambda\right)(n-1)|\lambda|\kappa$$

nonzero entries, as compared to potentially $\left(\dim\left(\mathbb{C}X^\lambda\right)\right)^2$ for the naïve, full transformation.

Proof. By the preceding theorem, in the subgroup decomposition, each factor has at most $K(\lambda)$ nonzero entries per row/column, and each is of size $\dim\left(\mathbb{C}X^\lambda\right)$. Thus, each factor contributes at most $K(\lambda)\dim\left(\mathbb{C}X^\lambda\right)$ nonzero entries, and there are $(n-1)$ such factors. This yields a total of $\dim\left(\mathbb{C}X^\lambda\right)(n-1)K(\lambda)$ nonzero entries across the factorization. But, by the previous corollary, this is at most $\dim\left(\mathbb{C}X^\lambda\right)(n-1)|\lambda|\kappa$. The naïve algorithm potentially can have a completely full transformation matrix, which requires at most $\left(\dim\left(\mathbb{C}X^\lambda\right)\right)^2$ nonzero entries. □

We can use this theorem to provide nearly immediate bounds on various, specific tabloid shapes. In particular, consider the following theorem, which provides a bound on the overall complexity of the factor F_j for an arbitrary r-rowed tabloid.

Theorem 4.2. *Let $\lambda = (\lambda_1, \ldots, \lambda_r) \vdash n$. Let $u(\lambda)$ be given via*

$$u(\lambda) = \left(\!\!\binom{\lambda_r + 1}{r-2}\!\!\right) \left(\!\!\binom{\lambda_{r-1} + 1}{r-3}\!\!\right) \cdots \left(\!\!\binom{\lambda_4 + 1}{2}\!\!\right) \left(\!\!\binom{\lambda_3 + 1}{1}\!\!\right).$$

Then,

$$K(\lambda) \leq \sum_{\mu \in \lambda^*} u(\mu).$$

Proof. To prove this, first note that $u(\lambda)$ is an increasing function λ_i for all applicable i. Given that we regard the orbits of λ as compositions of n, with row lengths allowed to be zero, then this shows that $u(\lambda)$ is increasing over all orbits and sub-orbits of λ, yielding maximal behavior at λ itself. Thus, if we show that this inequality holds for only the top-most level of orbits of λ, this will establish the inequality in general. To do this, let us first note that

$$\max_{\mu \trianglerighteq \lambda} \kappa_{\mu,\lambda} \leq u(\lambda).$$

This can be seen from a straightforward combinatorics argument. We note that any Kostka filling of μ must completely fill up the rth row of μ with rs, as using any other number would violate the Kostka constraints on the columns. Thus, when we ask the question, "in how many ways can we place the rs in μ?" we intrinsically have at most $r-1$ rows in which to place them. But, distributing λ_r things among $r-1$ rows is a clear combinatorics question. We see that there are $\left(\!\binom{\lambda_r+1}{r-2}\!\right)$ ways to distribute λ_r things among $r-1$ slots. Once we have done this, we see that any remaining slots in the $r-1$st row must be filled with $r-1$s, just as the rth row was mandated. Thus, we similarly have at most $r-2$ rows in which to place the $r-1$s for any placement of the rs. So, the total number of fillings is bounded above by the product of the number of individual fillings, given by

$$\left(\!\!\binom{\lambda_r + 1}{r-2}\!\!\right) \left(\!\!\binom{\lambda_{r-1} + 1}{r-3}\!\!\right).$$

We see that this pattern will continue, all the way until we are placing the 3s. We see that these will be the last entry we'll need to place, as follows.

Note that the positions of the 1s are mandated in any Kostka filling of μ; they must go in the left-most positions of the first row. Thus, once we place the 3s, the only number remaining that must be placed are the 2s. But, if there is only one number left to place, it simply must fill up every available slot. So, there will only ever be one way to fill the twos if we have placed all numbers at least 3. Thus, we see that the termination point for our product is at $\lambda_3 + 1$, so we find, as desired,

$$\max_{\mu \trianglerighteq \lambda} \kappa_{\mu,\lambda} \leq \left(\!\!\binom{\lambda_r + 1}{r-2}\!\!\right) \cdots \left(\!\!\binom{\lambda_3 + 1}{1}\!\!\right)$$

We take convention that if $r \leq 2$, then this product is an empty product, and hence $u(\lambda) = 1$.

Now that we have shown that this bounds the Kostka number on λ, the result is immediate. The function u can therefore bound the Kostka numbers on each orbit of λ, which is exactly what is encapsulated by

$$K(\lambda) \leq \sum_{\mu \in \lambda^*} u(\mu).$$

This completes the proof. □

Corollary. *Given $u(\lambda)$ as in Theorem 4.2, the full factorization over $\mathbb{C}X^\lambda$ requires no more than*

$$\dim\left(\mathbb{C}X^\lambda\right)(n-1)\left(\sum_{\mu \in \lambda^*} u(\mu)\right)$$

nonzero entries.

Proof. This is an exact application of Theorems 4.1 and 4.2. □

Corollary. *If we only let the first two rows of a tabloid shape grow towards infinity, then this factorization method has complexity*

$$O\left((n-1)\dim\left(\mathbb{C}X^\lambda\right)\right)$$

versus

$$O\left(\dim\left(\mathbb{C}X^\lambda\right)^2\right)$$

Proof. We see that $K(\lambda) \leq ru(\lambda)$ and that $ru(\lambda)$ is constant relative to λ_1, λ_2. Thus, letting these approach infinity along any path doesn't change the value of $ru(\lambda)$. The result follows. □

4.2 Two-Rowed Tabloids

4.2.1 Theoretical Complexity Bounds

First, we note that the bound in Theorem 4.2 predicts that two-rowed tabloids should never have *any* factor that needs more than two nonzero entries per row/column. In fact, this is the best bound obtainable for two-rowed tabloids using this method. To showcase this, consider the following example

Example 4.1. Let $\lambda_n = \left(\frac{n}{2}, \frac{n}{2}\right)$ for all even n. Then, consider the last factor $F_n(\lambda_n)$. We can bound its number of non-zeros by considering the decomposition of the module given by the action of S_{n-1} on each orbit of λ_n. We see that these orbits are described by the compositions of n given by $\left(\frac{n}{2} - 1, \frac{n}{2}\right)$ and $\left(\frac{n}{2}, \frac{n}{2} - 1\right)$. As these are merely reordered versions of each other, in fact these two orbits are isomorphic as modules. Thus, they have the same decomposition. As every Kostka number of an irreducible representation over a two-rowed space is 1, but each of them appears in both orbits, we see that *all* frequency spaces have dimension 2. Therefore, if we only bound based on the dimension of the frequency space, for a general two-rowed tabloid, we will never be able to bound lower than 2 nonzero entries per row/column.

4.2.2 Experimental Data:

We also tested these bounds via the implementation discussed in Section 3.4 and collected data about the resulting factorizations. For 2-rowed tabloids, in particular, we found that the bound of no more than 2 nonzero-entries per row/column was respected in all cases, with additional behavior observed in specific cases. In particular, consider the following sequences:

Let us first examine $\lambda = (n-1, 1)$. Decompositions over these shapes were very clearly structured. See, for example, the nonzero entries in the decomposition over $n = 9, \lambda = (8, 1)$ in Figure 4.1.

This decomposition clearly obeys the rule of no more than 2 per row/column; in fact, it typically has an average much closer to 1. A plot of the average number of nonzero entries per factor as a function of n is shown in Figure 4.2

Given this structure, it might seem that we could do even faster in this case. In fact, for this simple example, as the dimension of the space is so small (n), it is faster to simply run the full factorization. We can compare

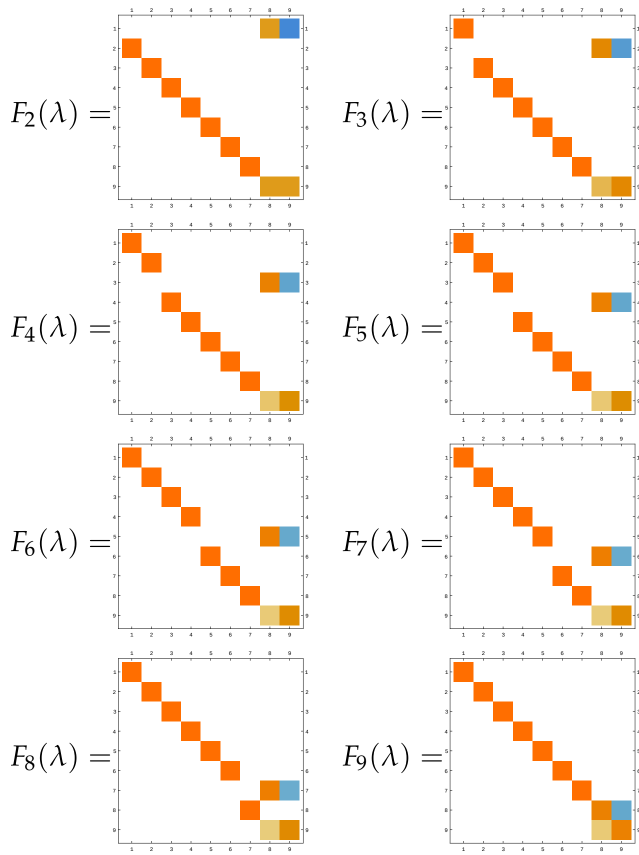


Figure 4.1 Matrix plots of the factors of the decomposition of the space $\mathbb{C}X^{(8,1)}$.

the total number of nonzero entries required across the factorization with those only required for the standard transformation via Figure 4.3.

For a more interesting example, let us examine the sequence $(n-3,3)$, for $n = 6,7,8$. We see from Figure 4.4 that its average obeys our bound and is a decreasing function. Further, it is clear that even in this simple case, the factorization performs much better than does the naïve, full decomposition. We can also examine a sample decomposition, such as that over the space $\mathbb{C}X^{(5,3)}$, shown in Figure 4.5

We can also use this experimental data to analyze the distribution of nonzero entries among the factors. Though yet unproven, the experimental data on two and three-rowed tabloids indicates that the number of nonzero entries for each factor forms a non-decreasing sequence. To illustrate this

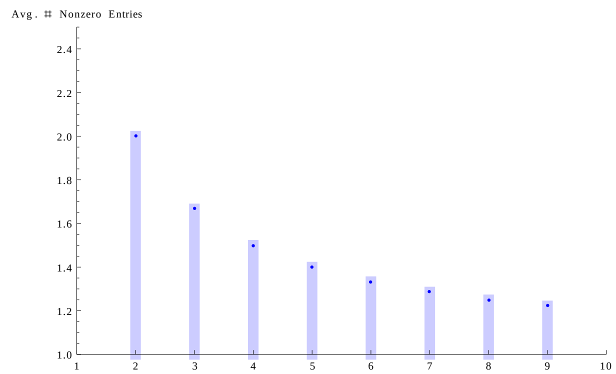


Figure 4.2 The average number of nonzero entries per row/column across all factors of $\lambda = (n-1,1)$ for n from 2 to 9.

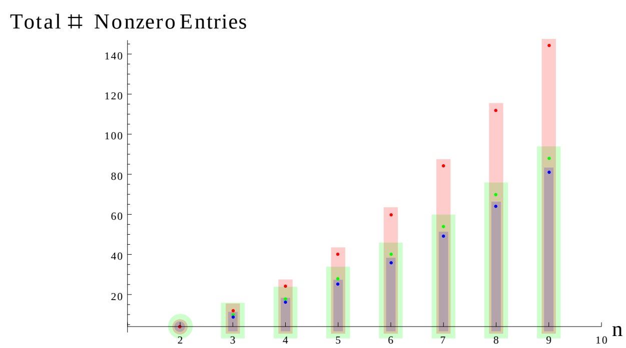


Figure 4.3 The total number of nonzero entries across all factors of the decomposition of $\lambda = (n-1,1)$ (orange), compared to dim $\left(\mathbb{C}X^{\lambda}\right)^2$ (blue) for n from 2 to 9.

for two-rowed tabloids, consider Figure 4.6. Here, we see this trend, that the number of nonzero entries is greatest in the last factor.

4.3 Three-Rowed Tabloids

For three-rowed tabloids, we can first form a bound based on Theorem 4.2,

Theorem 4.3. *Let $\lambda \vdash n$ be given by $\lambda = (a,b,c)$, $a \geq b \geq c > 0$, $a+b+c = M$. Then, $K(\lambda) \leq 3c+2$.*

Proof. We can see that $u((a,b,c,)) = c+1$. But, when summed over the three, top-level orbits of λ, namely $\lambda^1 = (a-1,b,c)$, $\lambda^2 = (a,b-1,c)$, and

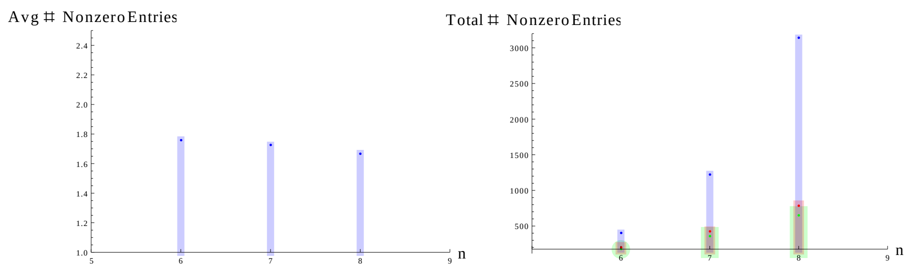


Figure 4.4 Examining the factorization of the transformation when $\lambda = (n-3, 3)$, for $n = 6, 7, 8$. *(left)* The average number of nonzero entries across all factors. *(right)* The total number of nonzero entries across all factors (green), compared to the theoretical bound $2(n-1)\dim\left(\mathbb{C}X^{\lambda}\right)$ (red) and $\dim\left(\mathbb{C}X^{\lambda}\right)^2$ (blue) for $n = 6, 7, 8$.

$\lambda^3 = (a, b, c-1)$, we see that this bound becomes

$$c + 1 + c + 1 + c = 3c + 2,$$

as desired. □

Given this theorem, we can now see that any 3-rowed decomposition requires at most $(3c+2)(n-1)\dim\left(\mathbb{C}X^{\lambda}\right)$ nonzero entries per row/column, where c is the length of the last row.

However, there is one specific case in which we can do better. For example, we form the following theorem:

Theorem 4.4. *Let $\lambda = (n-2, 1, 1)$, $n \geq 3$. Then, the jth factor $F_j(\lambda)$ requires no more than 4 nonzero entries per row/column.*

Proof. Here, let us examine the decomposition of ${}_{S_n}\mathbb{C}X^{\lambda}$ into orbits over S_{n-1}. We see that it is a union of three orbits, each of the following form:

Orbit 1: $\lambda^1 = (n-3, 1, 1)$**:** In this orbit, we work again with a three-rowed tabloid, described by shape $(n-3, 1, 1)$. This shape breaks down via

$$\mathbb{C}X^{\lambda^1} \cong S^{(n-1)} \bigoplus 2S^{(n-2,1)} \bigoplus S^{(n-3,1,1)}.$$

Orbit 2: $\lambda^2 = (n-2, 0, 1) = (n-2, 1)$**:** Here, we act over the composition described by λ^2, which we identify with the partition $(n-2, 1)$ (as the action of the symmetric group knows nothing about the row-lengths of the tabloids on which it acts). This shape breaks down via

$$\mathbb{C}X^{\lambda^2} \cong S^{(n-1)} \bigoplus S^{(n-2,1)}.$$

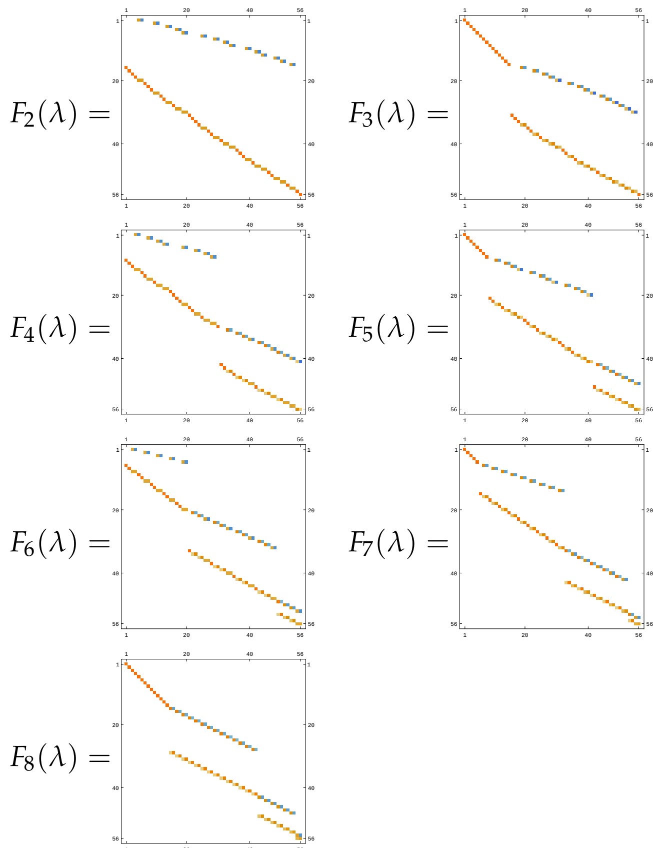


Figure 4.5 Matrix plots of the factors of the decomposition of the space $\mathbb{C}X^{(5,3)}$.

Orbit 3: $\lambda^3 = (n-2,1,0) = (n-2,1)$**:** Here, we act over the composition described by λ^3, which we identify with the partition $(n-2,1)$ as above. This shape breaks down via

$$\mathbb{C}X^{\lambda^3} \cong S^{(n-1)} \bigoplus S^{(n-2,1)}.$$

Thus, across all orbits, we see that the decomposition is

$$\begin{aligned}\mathbb{C}X^{\lambda} &\cong \mathbb{C}X^{\lambda^1} \bigoplus \mathbb{C}X^{\lambda^2} \bigoplus \mathbb{C}X^{\lambda^3} \\ &\cong 3S^{(n-1)} \bigoplus 4S^{(n-2,1)} \bigoplus S^{(n-3,1,1)}\end{aligned}$$

Thus, we see that the transformation over any transitive space described by the action of S_k on a tabloid $(k-2,1,1)$, from the $k-1$ level to the kth

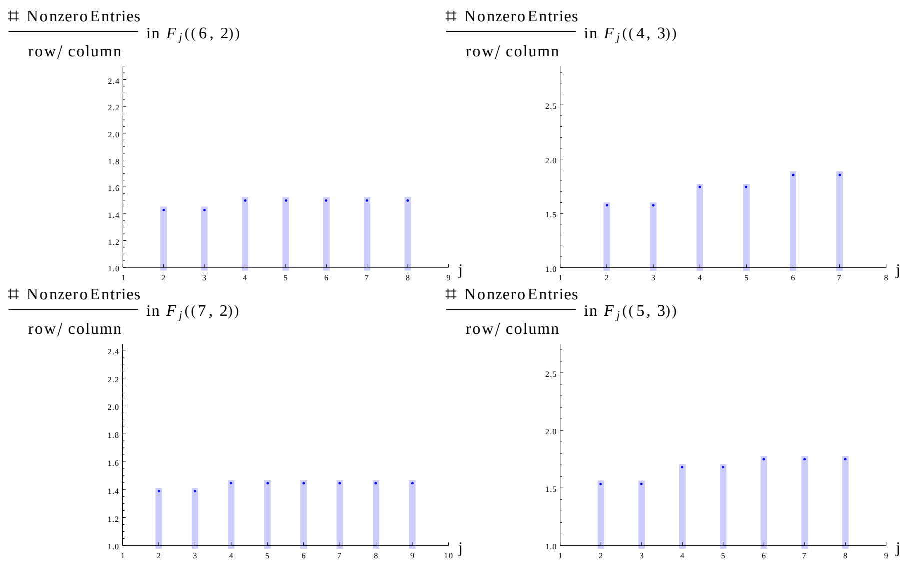


Figure 4.6 Examining the number of nonzero entries in each factor of the transformation when λ is given by *(upper left)* $(5,1,1)$, *(lower left)* $(6,1,1)$, *(upper right)* $(2,2,1)$, and *(lower right)* $(3,2,1)$.

level, requires no more than 4 nonzero entries per row/column. But, every level is described as a collection of orbits describable either by tabloids with fewer than 3 rows, or by a 3 rowed tabloid of shape $(m-2,1,1)$ for some m. Thus, though this analysis only directly examines factor F_n, its conclusion extends to factors F_j for all $j \leq n$. This proves the theorem. Why does this differ from the $K(\lambda)$ bound found above? In that bound, we used a general $\lambda = (a,b,c)$. In fact, this analysis shows by a direct decomposition that for $\lambda = (n-2,1,1)$, $K(\lambda) \leq 4$. In fact, with a slightly more nuanced investigation int $K(\lambda)$, this can be seen as well, but a full decomposition also gives you access to the number of appearances of every frequency space of each dimension, which would be very helpful for any large, industry level system. □

We summarize by describing three examples of 3-rowed tabloids and the implied bounds.

$\lambda = (n-2,1,1)$**:** This shape could be used to select one's ranked two favorites of n objects. Here, we have that $\mathbb{C}X^{\lambda}$ can be transformed with at most $4(n-1)n(n-1)$ entries, as opposed to $(n(n-1))^2$, as in the naïve bound. This case is truly a distinct bound from the next two.

$\lambda = (n-k-1, k, 1)$**:** This shape could be used to select a committee of size k and a chairperson out of n candidates. In this case, we have that $\mathbb{C}X^{\lambda}$ can be transformed with at most $5(n-1)\binom{n}{k,1}$ entries, as opposed to $\binom{n}{k,1}^2$. This case is an evaluation of the bound for the general case.

$\lambda = (n-k-\ell, k, \ell)$**:** In this full general case, we have that $\mathbb{C}X^{\lambda}$ can be transformed with at most $(3\ell+2)(n-1)\binom{n}{k,\ell}$ entries, as opposed to $\binom{n}{k,\ell}^2$.

4.3.1 Experimental Data:

Given the increased complexity of computing 3-rowed decompositions, we only present limited examples here. It is also when we begin experimenting on 3-rowed spaces that we begin to encounter computational round off error. In particular, as the spaces grow larger, the size of typical, erroneous round-off calculations computed in this decomposition increases. Simultaneously, the true size of the entries in these matrices decreases. It is in complex, 3-rowed tabloids that these start to approach similar ranges, and thus become indistinguishable. However, we can still gleam much structural information regarding these decompositions despite the occasional erroneous entry. In particular, let us first examine the sequence $(n-2,1,1)$, for n between 3 and 8. Figure 4.7 illustrates the average and total behavior of the number of non-zeros given by this decomposition.

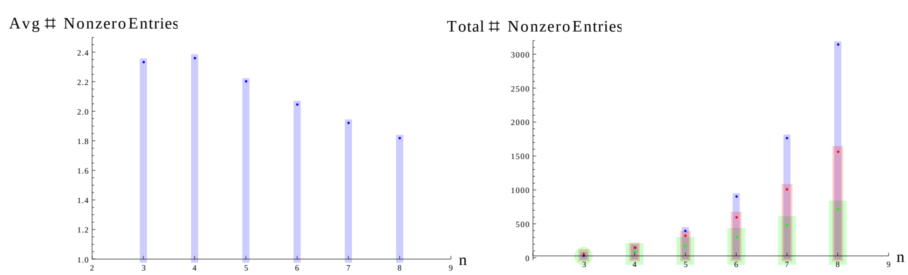


Figure 4.7 Examining the factorization of the transformation when $\lambda = (n-2,1,1)$, for n between 3 and 8. *(left)* The average number of nonzero entries across all factors. *(right)* The total number of nonzero entries across all factors (green), compared to the theoretical bound $4(n-1)\dim\left(\mathbb{C}X^{\lambda}\right)$ (red) and $\dim\left(\mathbb{C}X^{\lambda}\right)^2$ (blue) for n between 3 and 8.

Here, we can also examine the structure of the factors within a particular instance as well. Figure 4.8 shows each factor over the overall transformation for the space $\mathbb{C}X^{\lambda}$, with $\lambda = (6, 1, 1)$.

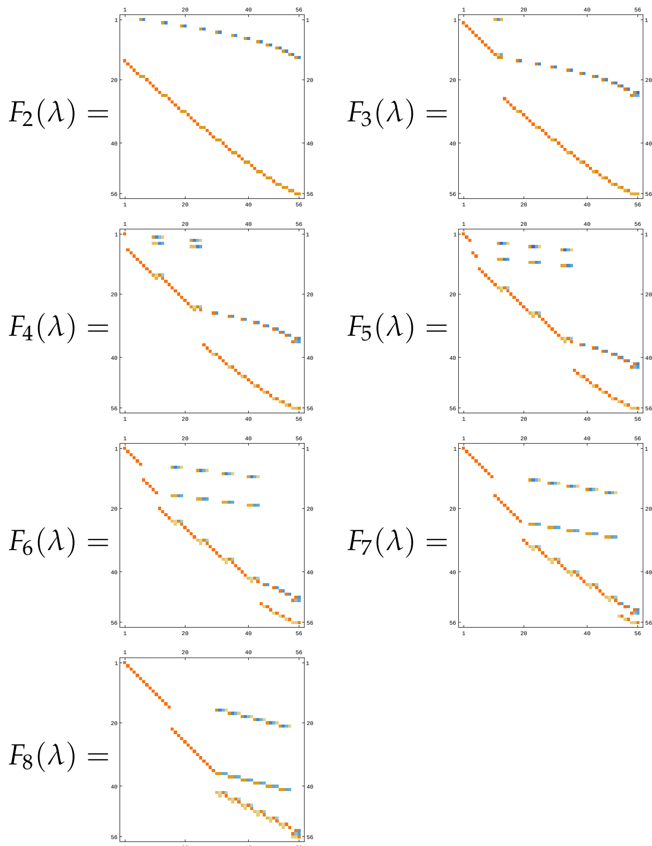


Figure 4.8 Matrix plots of the factors of the decomposition of the space $\mathbb{C}X^{(6,1,1)}$.

We can observe more complicated behavior with tabloid shape $\lambda = (n-3, 2, 1)$ or $(n-4, 2, 2)$. For $(n-3, 2, 1)$, average and total distributions are shown in Figure 4.9 for n from 5 to 7. For $(n-4, 2, 2)$, the total factorizations are illustrated for $n = 6$ and $n = 7$ in Figure 4.10.

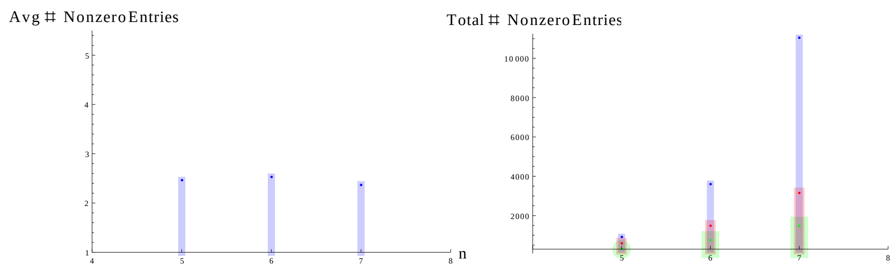


Figure 4.9 Examining the factorization of the transformation when $\lambda = (n-3,2,1)$, for $n = 5,6,7$. *(left)* The average number of nonzero entries across all factors. *(right)* The total number of nonzero entries across all factors (green), compared to the theoretical bound $5(n-1)\dim\left(\mathbb{C}X^{\lambda}\right)$ (red) and $\dim\left(\mathbb{C}X^{\lambda}\right)^2$ (blue) for n between 5 and 7.

Like in the two-rowed case, here we also find that the last factor always has the most nonzero-entries. This happened in all cases tested; several relevant cases are illustrated in Figure 4.11.

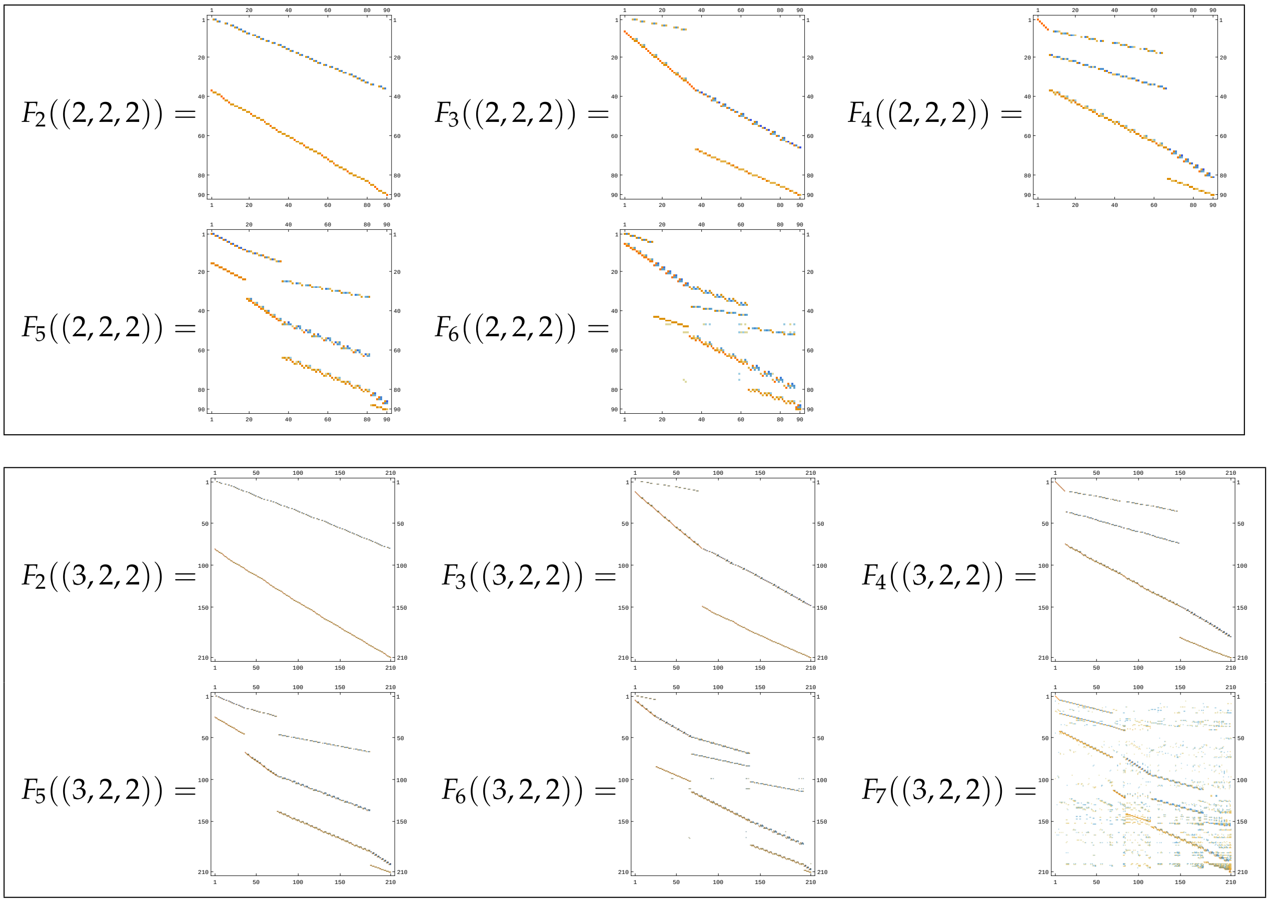


Figure 4.10 Matrix plots of the factors of the decomposition of the space *(top)* $\mathbb{C}X^{(2,2,2)}$ and *(bottom)* $\mathbb{C}X^{(3,2,2)}$. Note the accumulation of round off error in $F_7(3,2,2)$; the matrix appears to have lost most of its structure and clearly violates the theoretical bounds. This also appears in $F_6((2,2,2))$, though less prominently.

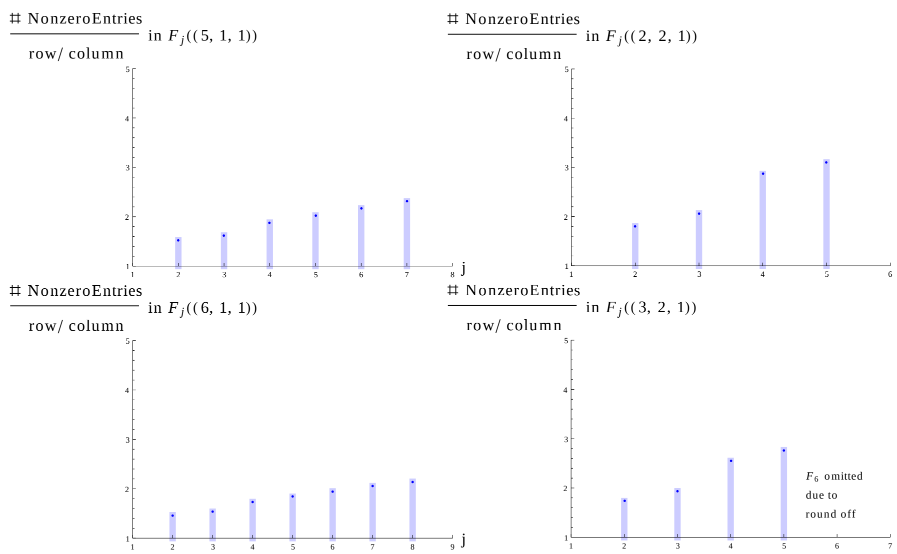


Figure 4.11 Examining the number of nonzero entries in each factor of the transformation when λ is given by *(upper left)* $(5, 1, 1)$, *(lower left)* $(6, 1, 1)$, *(upper right)* $(2, 2, 1)$, and *(lower right)* $(3, 2, 1)$. Note that we omit analysis of the transformation $F_6(\lambda)$ so that the round-off error encountered in that case does not bias the results.

Chapter 5

Future Work

5.1 Theoretical Expansions

The theoretical endeavors thus far represent attempts to bound the algorithmic complexity of the change of basis transformations F_j, given Theorems 3.5 and 3.6, and the relationships between Kostka numbers and irreducible multiplicities for the symmetric group. But, in fact, our bound $K(\lambda)$ is a significant overshoot; as experimental data demonstrates, specifically Figure 4.7, our bounds consistently overestimate the true complexity to perform transformation F_λ. Why is this, and how can we improve it? There are two reasons our bounds consistently overestimate the complexity of F_j. First and foremost, 3.5 is in and of itself an upper bound; we have no theorem that says these transformations will take *at least* some number of operations, only *at most*. Secondly, however, we are not even crafting the tightest bounds for our tabloids that Theorem 3.5 allows. Recall that our guiding object $K(\lambda)$ is given by

$$K(\lambda) = \max_{\mu \in O(\lambda)} \left(\max_{\rho \trianglerighteq \mu^*} \left(\sum_{t \in \mu^*} \kappa_{\rho,t} \right) \right).$$

What bound does Theorem 3.5 predict? Let G be a finite group with subgroup $H \leq G$ and adapted DFT D. Let H have f frequency spaces, of dimension $\alpha_1, \ldots, \alpha_f$. Then, Theorem 3.5 predicts that a change of basis from an adapted basis of H to one of G requires no more than $\sum_{i=1}^{f} \alpha_i^2$. We recall that the frequency spaces of H have dimensions specified exactly by the orbital decomposition of H over X and the resulting irreducible decomposition of each of the $\mathbb{C}$H modules spanned by its orbits. Let us interpret these insights in the context of the S_n and X^λ. If we wish to use Theorem 3.5 to determine the tightest possible bound on the complexity of factor F_j, we first must decompose the space $\mathbb{C}X^\lambda$ according to the orbits of S_j on X^λ. We recall from Theorem 2.2 that

$$_{\mathbb{C}S_j}\mathbb{C}X^\lambda \cong \bigoplus_{\eta \in {\lambda^*}^{n-j}} {}_{\mathbb{C}S_j}\mathbb{C}X^\eta .$$

Now, we must decompose each of these orbits with respect to $\mathbb{C}S_{j-1}$. Let us restrict ourselves to one particular orbit $\eta \in {\lambda^*}^{n-j}$. Then, we see that

$$_{\mathbb{C}S_{j-1}}\mathbb{C}X^\eta \cong \bigoplus_{\nu \in \eta^*} {}_{\mathbb{C}S_{j-1}}\mathbb{C}X^\nu .$$

Each $\mathbb{C}X^\nu$ then decomposes into irreducibles via

$$\mathbb{C}X^\nu \cong \bigoplus_{\mu \trianglerighteq \nu} \kappa_{\mu,\nu} S^\nu .$$

Let us use notation $\mu \trianglerighteq \eta^*$ to mean that there exists a $\nu \in \eta^*$ such that $\mu \trianglerighteq \nu$. To collapse irreducibles of the same type across orbits, we can then write

$$_{\mathbb{C}S_{j-1}}\mathbb{C}X^\eta \cong \bigoplus_{\mu \trianglerighteq \eta^*} \left(\sum_{\nu \in \eta^*} \kappa_{\mu,\nu} \right) S^\mu .$$

Thus, according to Theorem 3.6, the transformation restricted to the orbit of S_j described by η requires no more than

$$\sum_{\mu \trianglerighteq \eta^*} \left(\sum_{\nu \in \eta^*} \kappa_{\mu,\nu} \right)^2 \dim(S^\mu)$$

nonzero entries. To unpack this, recall that Theorem 3.6 reveals that the dimension of a frequency space associated to the jth irreducible representation is precisely the sum of the multiplicities of this irreducible representation in the decompositions of the orbital spaces of the subgroup in question. This yields our term $\sum_{\nu \in \eta^*} \kappa_{\mu,\nu}$. Further, Theorem 3.6 also reveals that there are precisely $\dim(\theta_\rho)$ frequency spaces associated with irreducible representation θ_ρ. Thus, we will have exactly $\dim(S^\mu)$ frequency spaces of dimension $\sum_{\nu \in \eta^*} \kappa_{\mu,\nu}$. But, this is only over one orbit. To compute the total number of nonzero entries in F_j, we sum over all orbits. Thus, we need no more than

$$\sum_{\eta \in {\lambda^*}^{n-j}} \left(\sum_{\mu \trianglerighteq \eta^*} \left(\sum_{\nu \in \eta^*} \kappa_{\mu,\nu} \right)^2 \dim(S^\mu) \right)$$

nonzero entries, for a maximum average number of nonzero entries per row/column of

$$\begin{aligned} c_j \leq C_j(\lambda) &= \frac{1}{\dim(\mathbb{C}X^\lambda)} \sum_{\eta \in {\lambda^*}^{n-j}} \left(\sum_{\mu \trianglerighteq \eta^*} \left(\sum_{\nu \in \eta^*} \kappa_{\mu,\nu} \right)^2 \dim(S^\mu) \right) \\ &= \frac{1}{|X^\lambda|} \sum_{\eta \in {\lambda^*}^{n-j}} \left(\sum_{\mu \trianglerighteq \eta^*} \left(\sum_{\nu \in \eta^*} \kappa_{\mu,\nu} \right)^2 \dim(S^\mu) \right). \end{aligned} \tag{5.1}$$

This bound is much more accurate than the bounds shown in Chapter 4, but is also much more difficult to interpret. For example, though Equation 5.1 is more nuanced than is $K(\lambda)$, it is really only useful when looking at a particular tabloid λ, as opposed to a class of tabloids, such as $(n-2,1,1)$ as $n \to \infty$ and so on. As this function is not obviously increasing with j, in these situations we cannot, for example, simply evaluate this bound at the top-most layer of orbits, like we did for many of the Kostka number bounds such as $u(\lambda)$ from Theorem 4.2. Granted, it may be the case that $C_j(\lambda)$ is increasing with respect to j and the component parts of λ; this just remains unproven. It is certainly true that in *all* experimental evidence we have, the last factor, F_n contributes the most nonzero entries to the entire factorization. On the other hand, these bounds are not increasing functions when restricted to the average number of nonzero entries within any particular orbit of S_j. Certain orbits of $S_j \leq S_n$, with j fixed as $n \to \infty$, maintain average nonzero counts that do not decay to one with n, whereas it can be shown that the complexity of the last factor for two- and three-rowed tabloids decays to 1 as $n \to \infty$. However, it is possible that the orbits in which this occurs only represent a small, decaying fraction of all orbits of S_j, and thus the factor on the whole could still have a decaying complexity cost even while certain small orbits maintain relatively high nonzero densities. Given this confusion, a key area of future investigation with this project is to determine whether or not $C_j(\lambda)$ is an increasing function with j; more generally, to determine whether or not the last factor truly always has the most nonzero entries.

This question is particularly valuable because, were it the case that the last factor always bounded the complexity of the overall transformation, one could rapidly cull the computation of certain, unnecessary Kostka numbers from the calculation of the overall complexity. Consider the following. Given a sequence $\{\lambda_n \vdash n\}$, there exist sequences $\{\mu_n \vdash n-1\}$, with $\mu_n \trianglerighteq \lambda_n^*$, such that $\lim_{n\to\infty} \frac{\dim(S^{\mu_n})}{|X^{\lambda_n}|} = 0$. For example, let λ_n be anything but the all partition (n) and let $\mu_n = (n-1)$ for all n. Then, S^{μ_n} is the trivial module, and has dimension n, while $|X^{\lambda_n}|$ will have some dependence on n, and thus the limit will tend to 0. Why are such sequences interesting? Note that we have, via Equation 5.1, the following bound on the last factor c_n,

$$\begin{aligned} c_n &\leq \frac{1}{|X^{\lambda_n}|} \sum_{\mu \trianglerighteq \lambda^*} \left(\sum_{\nu \in \lambda^*} \kappa_{\mu,\nu} \right)^2 \dim(S^{\mu}) \\ &= \sum_{\mu \trianglerighteq \lambda^*} \left(\sum_{\nu \in \lambda^*} \kappa_{\mu,\nu} \right)^2 \frac{\dim(S^{\mu})}{|X^{\lambda_n}|}. \end{aligned}$$

If λ is given, for each n, by $\lambda_n \vdash n$, then we have

$$\lim_{n\to\infty} c_n \leq \lim_{n\to\infty} \sum_{\mu \trianglerighteq \lambda_n^*} \left(\sum_{\nu \in \lambda_n^*} \kappa_{\mu,\nu} \right)^2 \frac{\dim(S^{\mu})}{|X^{\lambda_n}|}.$$

For any sequence λ^n, μ^n as given above, we see that

$$\begin{aligned} &\lim_{n\to\infty} \left(\sum_{\mu \trianglerighteq \lambda^*} \left(\sum_{\nu \in \lambda^*} \kappa_{\mu,\nu} \right)^2 \frac{\dim(S^{\mu})}{|X^{\lambda_n}|} \right) \\ &\quad = \lim_{n\to\infty} \left(\left(\sum_{\nu \in \lambda_n^*} \kappa_{\mu_n,\nu}^2 \right) \frac{\dim_{S^{\mu_n}}}{|X_n^{\lambda}|} + \sum_{\mu \trianglerighteq \lambda_n^*, \mu \neq \mu_n} \left(\sum_{\nu \in \lambda_n^*} \kappa_{\mu,\nu} \right)^2 \frac{\dim(S^{\mu})}{|X^{\lambda_n}|} \right) \\ &\quad = \lim_{n\to\infty} \left(\left(\sum_{\nu \in \lambda_n^*} \kappa_{\mu_n,\nu}^2 \right) \frac{\dim_{S^{\mu_n}}}{|X_n^{\lambda}|} \right) + \lim_{n\to\infty} \left(\sum_{\mu \trianglerighteq \lambda_n^*, \mu \neq \mu_n} \left(\sum_{\nu \in \lambda_n^*} \kappa_{\mu,\nu} \right)^2 \frac{\dim(S^{\mu})}{|X^{\lambda_n}|} \right). \end{aligned}$$

Now, let us presume that λ_n simply reflects growth in the first two rows of some base partition $\lambda \vdash n$, as is the case with all asymptotics investigated in this report. Then κ_{μ_n,λ_n} is bounded over all n, via the bound $u(\lambda)$, which is independent from the lengths of the first two rows. But, this implies that

$$\lim_{n\to\infty} \left(\left| \left(\sum_{\nu \in \lambda_n^*} \kappa_{\mu_n,\nu}^2 \right) \frac{\dim_{S^{\mu_n}}}{|X_n^{\lambda}|} \right| \right) \leq M \lim_{n\to\infty} \frac{\dim_{S^{\mu_n}}}{|X_n^{\lambda}|}$$

for some constant M. But, by the construction of the μ_n, this reveals

$$\lim_{n\to\infty}\left(\left(\sum_{\nu\in\lambda_n^*}\kappa^2_{\mu_n,\nu}\right)\frac{\dim_{S^{\mu_n}}}{|X_n^\lambda|}\right)=0.$$

Thus,

$$\lim_{n\to\infty} c_n \leq \lim_{n\to\infty}\left(\sum_{\mu\trianglerighteq\lambda_n^*,\mu\neq\mu_n}\left(\sum_{\nu\in\lambda_n^*}\kappa_{\mu,\nu}\right)^2\frac{\dim(S^\mu)}{|X^{\lambda_n}|}\right).$$

But, we can repeat this procedure for any other sequence $\{\lambda_n', \mu_n'\}$, continually adjusting the bound for c_n to remove any additional sequence that has this property. This shows that, in the limit as $n \to \infty$, c_n does not meaningfully depend on any irreducible components whose dimensions decay relative to the size of the X^λ. Therefore, in computing the limiting behavior of c_n, it suffices to only compute the Kostka numbers associated to irreducible representations whose dimensions are on the order of the size of X^λ.

If it is true that the last factor is an upper bound for the complexity of all factors, this result would be even more powerful, because then one could make limiting statements about all factors by only considering a very select, and often quite small, set of Kostka numbers and their associated irreducible representations.

5.2 Experimental Enhancements

There are a number of experimental enhancements that could be made to improve the process of experimenting and testing various hypotheses. Here, we focus on three specific prospects: round off error, implementing sparse arrays, and randomized initial configuration permutations. Round off error will be examined in Section, 5.2.1, but we will discuss the ideas behind sparse arrays and randomized initial configuration permutations here. Most matrices examined and produced by this algorithm are *very* sparse, yet are still stored in memory and manipulated as standard arrays of arrays. This is costly both in the space and time complexity of experimentation; however, as Boost proved so ineffectual, this idea has not yet been explored. The other idea espoused here revolves around truly randomly permuting the initial configuration order of X^λ prior to computing any bases. However, X^λ is often prohibitively large, forcing us to, until now, merely permute the first k columns for some small k. However, it should be possible to gain nearly truly random behavior over all of S_{12}, if nothing else than by perhaps repeated shuffling. Both this and the sparse array options are very theoretical options that have yet to be explored.

5.2.1 Round Off Error

Right now, the code cannot run on especially complex tabloid shapes or over large n due to an overwhelming tendency towards round off error. See, for example, Figure 5.1.

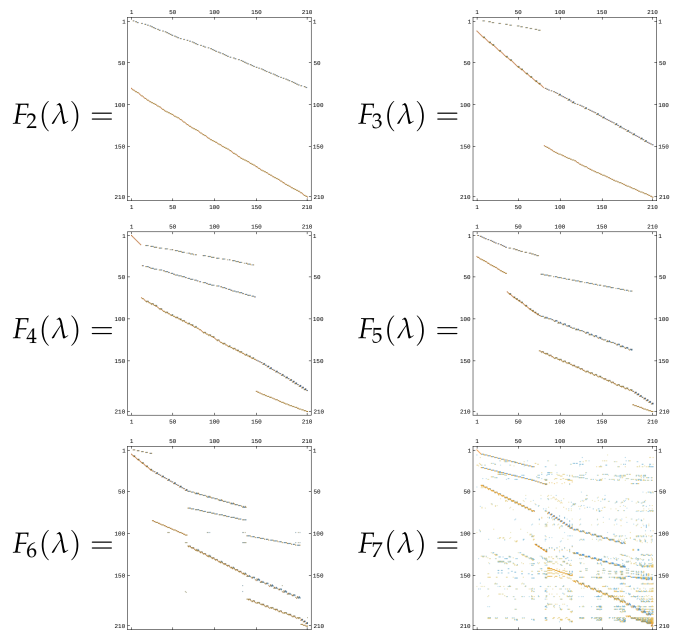


Figure 5.1 Matrix plots of the factors of the decomposition of the space $\mathbb{C}X^{(7,2,1)}$ demonstrate the system's propensity towards unchecked round-off error. The seemingly unstructured rows of entries are erroneous fragments. Later versions of the code corrected this error; however, all current measures are merely patches. A more systemic solution would be to use algebraically precise specialty number classes, such as a RATIONALS class and a SQRT class.

Currently, the code uses a very high precision setting and aggressive round-off termination. However, as the size of the spaces grows larger, the matrices grow more complicated and use smaller and smaller numbers, rendering erroneous output harder to distinguish from correct output. The only true solution would be to use an algebraically infinite precision format, such as a combination of a RATIONALS and SQRT class. Obviously, these would still be bounded by the machine size limits on integers (which could again be extended should the need be) but right now the problem is much closer to the precision, not range. It is important to note that the irreducible representations of S_n can actually be seen as living in the field of rational numbers and their square-roots, not just the complex numbers. Thus, these are actually mathematically feasible ideas.

5.3 Additional Resources

Beyond the many works suggested throughout this text, there are several resources available to the interested reader wishing to further pursue these ideas. Of particular note, on the experimental side, all code I produced for this work is available via github, at https://github.com/mmcdermott/symGpFactorizer. Additionally, there are several other, similar libraries for these tasks, such as the $S_n$ob library, produced by Risi Kondor, available at http://people.cs.uchicago.edu/~risi/SnOB/index.html. On the theoretical side, the books Clausen and Baum (1993) and Terras (1999) are both excellent introductions to Harmonic Analysis in general.

Bibliography

Clausen, Michael, and Ulrich Baum. 1993. *Fast Fourier transforms*. Wissenschaftsverlag.

Daugherty, Zajj, Alexander K. Eustis, Gregory Minton, and Michael E. Orrison. 2009. Voting, the symmetric group, and representation theory. *The American Mathematical Monthly* 116(8):667–687. URL http://www.jstor.org/stable/40391194.

Dummit, David Steven, and Richard M. Foote. 2004. *Abstract Algebra*. John Wiley & Sons Canada, Limited.

Fulton, William. 1996. *Young Tableaux: With Applications to Representation Theory and Geometry*. Cambridge, England ; New York: Cambridge University Press.

Fulton, William, and Joe Harris. 1991. *Representation Theory: A First Course*. Springer.

Huang, Jonathan. 2008. Adaptive Fourier domain inference on the symmetric group. URL http://www.gatsby.ucl.ac.uk/~risi/AML08/.

Huang, Jonathan, Carlos Guestrin, and Leonidas Guibas. 2008. Efficient inference for distributions on permutations. In *Advances in Neural Information Processing Systems 20*, 697–704. Cambridge, MA: MIT Press. URL http://books.nips.cc/nips20.html.

James, G. Gordon Douglas, and M. Martin W. Liebeck. 1993. *Representations and Characters of Groups*. Cambridge University Press.

James, Gordon, and Adalbert Kerber. 1982. *Representation Theory of the Symmetric Group*, *Encyclopedia of Mathematics and its Applications*, vol. 16. Addison-Wesley Publishing Group, 1st ed.

Kondor, Risi, and Marconi Barbosa. 2010. Ranking with kernels in Fourier space. In *COLT 2010*. Haifa, Israel. URL http://people.cs.uchicago.edu/~risi/papers/KondorBarbosaCOLT10.pdf.

Kondor, Risi, Andrew Howard, and Tony Jebara. 2007. Multi-object tracking with representations of the symmetric group. In *In AISTATS*.

Sagan, Bruce Eli. 1991. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*. Wadsworth & Brooks/Cole Advanced Books & Software.

Stiefel, E., and A. Fässler. 1992. *Group Theoretical Methods and Their Applications*. Birkhäuser, 1992nd ed.

Terras, Audrey. 1999. *Fourier Analysis on Finite Groups and Applications*. Cambridge University Press.