2005

An Algebraic Approach to Voting Theory

Zajj Daugherty
Harvey Mudd College

An Algebraic Approach to Voting Theory

Zajj Daugherty

Michael Orrison, Advisor

John Alongi, Reader

May 11, 2005

Harvey Mudd College

Department of Mathematics

Abstract

In voting theory, simple questions can lead to convoluted and sometimes paradoxical results. Recently, mathematician Donald Saari used geometric insights to study various voting methods. He argued that a particular positional voting method (namely that proposed by Borda) minimizes the frequency of paradoxes. We present an approach to similar ideas which draw from group theory and algebra. In particular, we employ tools from representation theory on the symmetric group to elicit some of the natural behaviors of voting profiles. We also make generalizations to similar results for partially ranked data.

Contents

Abstract **iii**

Acknowledgments **vii**

1 Introduction **1**
1.1 A Brief Introduction to Voting Theory 1
1.2 New Directions 8

2 Mathematical Framework **9**
2.1 Permutation Modules 9
2.2 Representation Theory 11
2.3 Voting Theory in Algebraic Terms 13

3 Exploring Full Rankings with Algebraic Theory **19**

4 Extending to Partial Rankings **29**

5 Conclusion and Further Work **39**

Bibliography **41**

Acknowledgments

I would like to thank my advisor Michael Orrison for his outstanding guidance throughout this project, as well as my second reader John Alongi for his time and insightful commentary. I would also like to thank the Reed Institute for Decision Sciences for making my preliminary research on this topic possible. Finally, many thanks to my fellow students Grant Clifford, David Gross, Dan Halperin, Jamie Kunkle, Eric Malm, Jason Murcko, and Gwen Spencer for their support, feedback, camaraderie, and general aid in meltdown-prevention.

Chapter 1

Introduction

1.1 A Brief Introduction to Voting Theory

Our society is very familiar with the process of voting. We vote in political elections for mayors, senators, and presidents. Committees vote on which actions they want their group to take. Voting is also a process which we essentially take part in each day. We 'vote' when we go out with our friends and need to decide which restaurant to go to or which movie to see. Voting takes place whenever a group needs to pool preferences to make an overall decision.

Unfortunately, it is not always easy to make a single choice when given the preferences of many voters. Not only are there different ways of voting, but there are also many ways to count those votes. In many political elections, we choose our favorite candidate and give no more information. But there are also situations in which voters might be asked to give a *full ranking* of multiple alternatives, such as in a marketing survey. If we then choose to give a number of points to these items according to where they were placed in each ranking, this would be referred to as a *positional* tally method. For example, if we were to ask voters to rank three candidates, we might tally their votes by giving two points to a candidate every time he or she is ranked as a top choice, and one point to a candidate every time he or she is ranked as a second choice. The candidate with the largest number of points in the end would be the winner.

Each positional tally method can be represented by a *weighting vector*, $\mathbf{w}$, whose entries correspond to the number of points given to a voter's first, second, . . . , last choice. For convenience, we will often normalize these vectors so that their entries range from 1 to 0. For example, if we give one point to a voter's top choice, and no points to any other candidate in their ranking, this would be represented by the weighting vector $\mathbf{w} = [1, 0, \ldots, 0]$. This is similar to the *plurality* voting scheme, where we are essentially asking voters to choose only their top candidate. Later, we will also discuss the *Borda Count*, which has weighting vector $\mathbf{w}_{\text{Borda}} = [1, \frac{n-2}{n-1}, \frac{n-3}{n-1}, \ldots, 0]$, where n is the number of candidates.

Alternatively, we could compare candidates *pairwise*, and give points according to how many times one candidate beats each other candidate. In other words, candidate A wins over candidate B if A is ranked higher than B more times than B is ranked higher than A. There is also *approval* voting, where voters simply give out one point to each candidate of whom they approve.

This naturally leads to the question of which method is the "best"—which method most accurately represents the intention of the voters. Although many methods of voting had been introduced for centuries prior, this debate made its way significantly into academia in the 18$^{\text{th}}$ century. The mathematician and philosopher Marie Jean Antoine Nicolas de Caritat Condorcet brought the pairwise method, first recorded by Ramon Llull of Spain in the 13$^{\text{th}}$ century, to the attention of the French academy (3). In 1785, Condorcet published the paper *Essai sur l'application de l'analyse la probabilite des decisions rendues la pluralite des voix* (Essay on the Application of Analysis to the Probability of Decisions Rendered by a Plurality of Votes). In this paper, he proposed the idea that if a candidate were to win all head-to-head comparisons, then that candidate should win the overall election. He recognized, however, that this method would not always guarantee a winner. He also discussed what has become known as the *Condorcet paradox*: in head-to-head comparisons, a society can prefer candidate A over candidate B, candidate B over candidate C, and candidate C over candidate A. In other words, the societal preferences in pairwise tallies may not be transitive. Because of his influence in bringing pairwise voting into focus, the method of pairwise comparison is also often referred to as the *Condorcet method*. The requirement that a method elects the Condorcet winner (if one exists) is called the *Condorcet criterion*.

Condorcet's paper spurred a debate with another French mathematician by the name of Jean Charles de Borda (3). Borda was known for his work on engineering in the military and his contributions to developing the metric system. He worked with Condorcet, alongside Laplace, Lavoisier and Legendre, on the Commission of Weights and Measures (founded in 1790). Borda challenged Condorcet and Llull's method on the grounds that it was not entirely workable. Often, Condorcet's method will show cyclic preferences, and fail to produce a winner in circumstances which do not intuitively suggest a tie. Instead, Borda proposed a method developed by Nicholas of Cusa in the 15^{th} century. He proposed the positional tally for n candidates which gives n points to a voter's top choice, $n-1$ points to the second choice, and so on (as mentioned previously, we will normalize these values for our purposes). This method was also prone to producing ties, but was more likely to produce a winner. These two mathematicians continued to debate which of these methods was the most 'fair' and productive almost until Condorcet's arrest and death in 1794. Unfortunately, since their debate was mostly philosophical (despite some use of certain axiomatic and probabilistic tools), and since both systems had notable strengths and weaknesses, they were not able to make much headway.

To explicitly illustrate how these two tallying methods might differ, we look at an example. Say we ask voters to rank three candidates—A, B, and C. To represent A beating B, we will write $A > B$ (sometimes we will simply write AB). Suppose 24 voters distribute their votes in the following manner:

votes	ranking	votes	ranking	votes	ranking
3	$A > B > C$	8	$B > A > C$	0	$C > A > B$
8	$A > C > B$	0	$B > C > A$	5	$C > B > A$

In other words, three people rank A above B above C, eight people rank A above C above B, and so on. We would then collect this data into a *profile* $\mathbf{p}$ which represents how many people voted for each ranking of candidates. The profile for this outcome, if we order the rankings lexicographically, is $\mathbf{p} = (3, 8, 8, 0, 0, 5)$. If we use the Borda Count to tally these votes (with weighting vector $\mathbf{w} = [1, 1/2, 0]$), A gets 15 points, B gets 12 points, and C gets 9 points. This results in an overall ranking of $A > B > C$. However, if we choose to calculate the pairwise ranking, we would find the following tally:

points	pair	points	pair
11	$A > B$	13	$B > A$
19	$A > C$	5	$C > A$
11	$B > C$	13	$C > B$

In other words, A ranked higher than B eleven times, where B ranked higher than A thirteen times, which gives that B beat A overall. Similarly, we get that A beats C by fourteen points, and C beats B by two points. Notice that this does not give us a transitive ranking at all. Also, this tally disagrees with our Borda tally on two accounts—the comparison of A and

B, and the comparison of *B* and *C*. This is what is referred to as a *paradox*, because two seemingly fair tallying methods result in differing outcomes.

Following Condorcet and Borda's debate, many people attempted to modify these procedures to improve upon them. For example, in 1876, mathematician Charles Ludwidge Dodgson (better known as Lewis Carroll) proposed a procedure in which a Condorcet winner will be chosen in the case that one exists; if there is no Condorcet winner, then the candidate who needs fewest ballots to be changed to become the Condorcet winner will be chosen. John Kemeny proposed a similar system in 1959, where the winner is the candidate who requires the fewest number of rank pairs being exchanged (flipped) on voters' ballots to make that candidate win by Condorcet's rule. In 1958, Duncan Black constructed a method which blended Condorcet's and Borda's methods. Namely, the Black winner will be the Condorcet winner, unless one fails to exist—otherwise, the Borda winner will be chosen. Each of these theories provided additional strengths to Condorcet methods (i.e. we are more likely to calculate a winner), but many of them also resulted in new paradoxes and weaknesses (for example, the resulting winner may no longer be intuitive) (4).

1.1.1 Arrow's Theorem

A major advancement in voting theory was made in the mid-20th century by the Nobel laureate economist Kenneth Arrow. With contributions from Blau and Murakami, he used axiomatic methods to prove that there can not exist an entirely fair voting system. More specifically, he chose four preferred criteria for a voting scheme to exhibit and showed that they were inconsistent. These four criteria were Universality, Independence of Irrelevant Alternatives, Citizen's Sovereignty, and Non-dictatorship (2). *Universality* (or *unrestricted domain*) means that the election procedure should provide a full ranking (with strict preferences) for all possible sets of data. *Independence of Irrelevant Alternatives* requires that any ranking of a subset of alternatives will be unaffected by changes in rankings of other alternatives. *Citizen's Sovereignty* (or *non-imposition*) requires that the election procedure allows for all possible outcomes. *Non-dictatorship* simply means that more than one person's vote can affect the outcome. Arrow proved that if the first three criteria were true for a given election procedure, then it must be a dictatorship. This was a profound message for voting theorists, because it meant that we could not possibly construct a truly fair system, at least according to Arrow's criteria. There was some debate following this result, questioning the nature of each of these requirements. In particular, the

	K	*B*			*C*	*R*		
		$\mathbf{b_A}$	$\mathbf{b_B}$	$\mathbf{b_C}$		$\mathbf{r_A}$	$\mathbf{r_B}$	$\mathbf{r_C}$
ABC	1	1	0	−1	1	1	−2	1
ACB	1	1	−1	0	−1	1	1	−2
BAC	1	0	1	−1	−1	−2	1	1
BCA	1	−1	1	0	1	1	1	−2
CAB	1	0	−1	1	1	−2	1	1
CBA	1	−1	0	1	−1	1	−2	1

Table 1.1: Saari's basis for three alternatives.

Universality criterion may be a bit strict when a profile would intuitively suggest a tie. However, in each modification, the adjusted rules still proved to be inconsistent. Thus, it is perhaps no longer useful to ask the question "How can we construct a fair system?" Instead, the focus has been shifted to a search for the system which is the 'most' fair, and how we might redefine our concept of fairness.

1.1.2 Donald Saari and Geometric Methods

In the late 20^{th} century, mathematician Donald Saari began to tackle this revised objective using geometric methods. He began describing voting structures as vector spaces. As mentioned before, we can discuss profiles and positional weightings as vectors. We can then naturally begin to study the vector spaces in which these vectors reside. This is precisely what Saari did. He decomposed profile spaces using geometric ideas in order to explain how paradoxes arise, and with what significance. By analyzing the dimensions of various components, he could make statements as to how often certain tally methods would agree with others. In particular, he has concentrated on comparing pairwise and positional methods—modifying the positional methods to agree with the pairwise outcomes as often as possible.

In three major papers (6; 7; 8), Saari divided profile spaces into four subspaces: the *Kernel* (the all ties space, generated by the all-ones vector), the *Basic* space (containing profiles for which the positional and Condorcet tallies agree), the *Condorcet* space (containing profiles which only influence Condorcet outcomes), and the *Reversal* space (containing profiles which only influence positional outcomes).

For example, Table 1.1 outlines Saari's basis for a full ranking of three alternatives. The columns correspond to the four subspaces outlined above: the Kernel (K), the Basic space (B), the Condorcet space (C), and the Rever-

	K	B			C	R		
		$\mathbf{b_A}$	$\mathbf{b_B}$	$\mathbf{b_C}$		$\mathbf{r_A}$	$\mathbf{r_B}$	$\mathbf{r_C}$
A	$2+2t$	2	−1	−1	0	$2-4t$	$-1+2t$	$-1+2t$
B	$2+2t$	−1	2	−1	0	$-1+2t$	$2-4t$	$2-4t$
C	$2+2t$	−1	−1	2	0	$-1+2t$	$-1+2t$	$-1+2t$
AB	3	2	−2	0	1	0	0	0
BA	3	−2	2	0	−1	0	0	0
AC	3	2	0	−2	−1	0	0	0
CA	3	−2	0	2	1	0	0	0
BC	3	0	2	−2	1	0	0	0
CB	3	0	−2	2	−1	0	0	0

Table 1.2: Tallies for Saari's basis for three alternatives.

sal Space (R). For both B and R, each of the three vectors corresponds to a particular candidate, and any two of the three vectors generates the third. For example, in B, the first vector corresponds to the first candidate, A: 1 vote is made for rankings where A is top-ranked (such as $A > B > C$), -1 vote is made for rankings where A is bottom-ranked (such as $B > C > A$), and no votes are made otherwise, giving $\mathbf{b_A} = (1, 1, 0, -1, 0, -1)^T$. The three basic vectors are linearly dependent (namely, $\mathbf{b}_A + \mathbf{b}_B + \mathbf{b}_C = 0$, so we can choose any two for our basis. Similarly, $\mathbf{r}_A$ is constructed by giving 1 point wherever A is top or bottom ranked, and 2 points otherwise, giving $\mathbf{r_A} = (1, 1, -2, 1, -2, 1)^T$.

Remark: It may be useful to note here that although the concept of 'negative votes' is unintuitive, it is acceptable for our purposes (and relatively essential). Notice that since the all-ones vector does not affect the outcome of either procedure, we can simply shift a profile by multiples of this vector to recover an intuitive profile. So, for example, the profile $\mathbf{p}' = (-1, 4, 4, -4, -4, 1)$ elicits the same results under both maps as $\mathbf{p} = (3, 8, 8, 0, 0, 5)$ did in our previous example. Similarly, scaling a profile by any positive constant will also produce similar results. This is to say, if a candidate wins in one profile, he or she may win by more or fewer votes in absolute terms, but will still win. So, again, $\mathbf{p}' = (6, 16, 16, 0, 0, 10)$ would produce the same result under both maps as $\mathbf{p}$ above.

Table 1.2 outlines the image (or outcomes) of the basis profiles under the positional and pairwise tallies. The first row shows how many points are given to each candidate under the positional scheme, where our weighting vector is $\mathbf{w} = [1, t, 0]$. For example, the basis vector for K will result in all three candidates receiving $2 + 2t$ points under the positional map. Notice, as described, vectors in K will give rise to a tie, regardless of our choice of t. The Basic vectors will always lead to one candidate winning, and the rest tying (for the Basic vector corresponding to candidate X, X will win). For example, $\mathbf{b}_A$ gives A 2 points, B -1 points, and C -1 points, yielding an overall ranking of $A > B \sim C$ (where '$\sim$' indicates a tie). The Condorcet vector will not contribute at all to the outcome, as it is all zeros, again regardless of our choice of t. The Reversal space, however, will contribute depending on our choice of t. For the Reversal vector corresponding to candidate X, X will win if $t < 1/2$ and X will lose if $t > 1/2$. However, if $t = 1/2$, as in the Borda Count, the Reversal space will no longer contribute to the final outcome. In other words, if $t \neq 1/2$, then if a profile vector has a non-trivial projections into the Reversal space, we will see a difference between the positional and pairwise outcomes. Therefore, a larger percentage of profiles have the potential to elicit this kind of paradox.

The second row of Table 1.2 shows the image of these basis vectors under the pairwise method. Just as was the case under the positional tally, vectors in K will give rise to a tie, and the Basic vectors will always lead to one candidate winning, and the rest tying (for the Basic vector corresponding to candidate X, X will win). However, as projected for the pairwise tally, the Condorcet space will contribute, contradicting the positional tally. Therefore, there is no t where this procedure agrees with the positional procedure for all profiles. Also, the Reversal space will not contribute under any circumstances. However, if we choose $t = 1/2$, each of the results for the positional outcome evaluate to zero in the positional space, thus providing fewer dimensions of conflict with the pairwise outcome. In fact, for the three-candidate case, if we use the Borda Count, the only circumstances in which the two tally methods will contradict each other are when the profiles have non-trivial projections onto the vector generating C ($\mathbf{c} = (1, -1, -1, 1, 1, -1)^T$). This is what we mean when we say that the Borda Count reduces the amount of conflict between the positional and pairwise tallies.

This was, in fact, one of Saari's main contributions to the field—proving mathematically that the positional tallying method which agrees most often with pairwise tallies for full rankings is the Borda Count, and that this result is unique.

1.2 New Directions

One possible natural extension of Saari's work is to provide similar decompositions using algebraic techniques. This thesis will contain the necessary algebraic background and translations of voting theory into algebraic terms to make use of these techniques. For example, we have already seen how profiles can be represented as elements of a vector space. With this idea there come many tools from linear algebra. Abstract algebra lends itself well to working with both the vector spaces of profiles, as well as with maps associated with them. In particular, we can describe various rankings of candidates as permutations—elements of the symmetric group. Tallying these profiles can be described as mapping their profile vectors from one space to another. As we will show later, these maps are $\mathbb{Q}S_n$-module homomorphisms. So in particular, representation theory will prove to be quite useful. We will see how representation theory can help us to study the profile space in terms of the maps we use to tally the votes.

Another interesting extension which we will be exploring here is an analogous study of partially ranked data. It is not always practical to ask for a fully ranked list from voters (for just ten candidates, this would give voters $3,620,800$ choices!). A *partial ranking* calls for voters to place candidates into ranked sets. For example, we might ask voters to tell us, out of six candidates, their top choice, then their next two favorite, and then their three least favorite, without making any distinctions within these categories. For the most part, we will be concentrating on the partial ranking of n alternatives, where voters are asked to fully rank their favorite k alternatives. Representation theory can provide useful tools for analyzing these kinds of votes as well. A significant amount of theory has been developed for objects which we will use to represent partially ranked data (11). This theory will prove useful for making generalization to partially ranked data.

This thesis proceeds as follows. In Chapter 2, we describe the mathematical framework in which we study voting theoretical objects. In particular, we introduce the reader to key tools for working with the symmetric group and permutation modules. Also, we devote some time to discussing how we treat familiar voting structures within our algebraic context. In Chapter 3, we revisit the case of fully ranked data, but within this algebraic context. In Chapter 4 we apply the tools developed in the previous two chapters to partially ranked data, achieving analogous results. In the final chapter, Chapter 5, we discuss possible further directions that algebraic techniques could take voting theory.

Chapter 2

Mathematical Framework

We have already reviewed many of the economic terms necessary to understand the framework in which we are examining voting. In this chapter, we will discuss more of the algebraic background necessary to understand our approach. The theory behind the first two sections of this chapter can be found in Dummit and Foote's *Abstract Algebra* (1) and Sagan's *The Symmetric Group* (11). We will also further discuss how to translate voting-related concepts into algebraic language in the third section.

2.1 Permutation Modules

First, we examine the tools which we use to represent the data itself. Recall that the *symmetric group on n elements, S_n*, is defined to be the set of all permutations of n objects, together with the binary operation of (function) composition. In the case of asking voters to provide a full ranking on n candidates, we can consider each of the possible rankings as elements of S_n. Recall that a *group ring RG* is simply the ring of formal sums $\sum_{g \in G} r_g g$ ranging over a finite multiplicative group G, where the coefficients r_g are coming from a commutative ring R with identity $1 \neq 0$. Thus, we can think of any profile as an element of the group ring $\mathbb{Q}S_n$ (in this case each profile only has integer coefficients, but $\mathbb{Q}$ allows for scaling as described in Chapter 1). However, if we wish to study the outcomes of partially ranked procedures, we can draw on the following generalized analog to the symmetric group.

As mentioned in Chapter 1, we can think of partially ranking candidates as placing candidates into ranked sets. One way of expressing these ranked sets is in terms of objects called *tabloids*.

Definition 2.1. $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ is a *(combinatorial) composition* of n, written $\lambda \vDash n$, if each λ_i is a positive integer and $\sum_{i=1}^{l} \lambda_i = n$. λ is a *partition* of n if it is a composition, and for all $i < j$, $\lambda_i \geq \lambda_j$. The *Ferrers diagram*, or *shape*, of λ is then an array of n dots having l left-justified rows with row i containing λ_i dots for $1 \leq i \leq l$.

For example, $\lambda = (2, 2, 1) \vDash 5$. The Ferrers diagram for $\lambda = (2, 2, 1)$ is

$$\begin{array}{ll} \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \end{array}.$$

Definition 2.2. Given composition $\lambda \vDash n$, a *Young diagram of shape* λ is an array t^λ obtained by replacing the dots of the Ferrers diagram of shape λ with the numbers $1, 2, \ldots, n$ bijectively. Two diagrams are equivalent if the sets of elements in each corresponding row are equal. Each of these equivalancy classes is a *Young tabloid of shape* λ (or a λ*-tabloid*).

For our example of $\lambda = (2, 2, 1)$, a few possible *non-equivalent* λ-tabloids would be

$$\begin{array}{ll} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & \\ \hline \end{array} \qquad \begin{array}{ll} \hline 1 & 3 \\ \hline 2 & 4 \\ \hline 5 & \\ \hline \end{array} \qquad \begin{array}{ll} \hline 2 & 5 \\ \hline 1 & 4 \\ \hline 3 & \\ \hline \end{array}.$$

The lines between the rows are to emphasize the equivalence within rows. The following tabloids are a complete set of equivalent tabloids of a particular configuration:

$$\begin{array}{ll} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & \\ \hline \end{array} \sim \begin{array}{ll} \hline 2 & 1 \\ \hline 3 & 4 \\ \hline 5 & \\ \hline \end{array} \sim \begin{array}{ll} \hline 1 & 2 \\ \hline 4 & 3 \\ \hline 5 & \\ \hline \end{array} \sim \begin{array}{ll} \hline 2 & 1 \\ \hline 4 & 3 \\ \hline 5 & \\ \hline \end{array}.$$

In total, there are $\frac{5!}{2!2!1!} = 30$ of these tabloids of shape $\lambda = (2, 1, 1)$. In general, there are $\frac{n!}{\lambda_1!\lambda_2!\ldots\lambda_l!}$ tabloids of shape $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l) \vDash n$.

Just as it was natural to express full rankings as elements of S_n, we can now express partial rankings as tabloids in X^λ, where X^λ is defined as the set of all λ-tabloids. Similarly, just as we can express profiles for fully-ranked data as elements of $\mathbb{Q}S_n$, we can now express profiles for partially-ranked data as elements of $\mathbb{Q}X^\lambda$. Again, this is simply the ring of formal sums $\sum_{t^\lambda \in X^\lambda} qt^\lambda$ with coefficients $q \in \mathbb{Q}$. Notice now that we can also associate an action of S_n on these tabloids. Namely, the action of S_n on n

elements induces an action on a tabloid of shape $\lambda \vDash n$ with the same set of elements for entries. For example, if we have the partial ranking

$$\begin{array}{cc} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & \\ \hline \end{array},$$

we can permute it on the left, say, with the transposition $(13) \in S_n$ (swap 1 and 3):

$$(13) \cdot \begin{array}{cc} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & \\ \hline \end{array} = \begin{array}{cc} \hline 2 & 3 \\ \hline 1 & 4 \\ \hline 5 & \\ \hline \end{array}.$$

We call $\mathbb{Q}X^{\lambda}$ the *permutation module corresponding to* λ, and denote it by M^{λ}.

Notice, for example, that for $\lambda = (n)$, M^{λ} is isomorphic as a $\mathbb{Q}S_n$ module to $\mathbb{Q}$. This would be equivalent to simply asking voters to rank everyone the same. The profile, then, would record exactly the number of people who voted. Alternatively, if $\lambda = (1, 1, \ldots, 1)$, then M^{λ} is isomorphic as a $\mathbb{Q}S_n$ module to $\mathbb{Q}S_n$ itself. This would be equivalent to asking voters to give a full ranking of n candidates. Thus, any analysis for the case of full rankings can also be done in terms of these permutation modules.

Remark: Note that for any $\lambda \vDash n$, M^{λ} and $M^{\sigma(\lambda)}$ are isomorphic as modules, where σ permutes λ. This is to say, $M^{(n-3,1,1,1)}$ is isomorphic as a $\mathbb{Q}S_n$-module to $M^{(1,1,1,n-3)}$. Thus, analyzing tabloids of shape $(n-3,1,1,1)$ is equivalent to analyzing tabloids of shape $(1,1,1,n-3)$. For the sake of convenience, we assume that each composition λ is a partition.

We will now move on to review some of the larger algebraic framework in which we will be analyzing these permutation modules.

2.2 Representation Theory

In this section, we discuss some of the algebraic tools necessary to analyze our profile spaces. First, we will examine methods for separating a space into useful subspaces.

Recall that if U and V are subspaces of vector space W which intersect trivially at $\{\mathbf{0}\}$, then their *(internal) direct sum*, $U \oplus V$, is defined as

$$U \oplus V = \{u + v \mid u \in U, v \in V\}.$$

The phrase "decomposing a space" then refers to expressing it as a direct sum of subspaces.

Definition 2.3. If R is a ring with identity, then a *(left) R-module* is defined to be an abelian group M together with an action of R on M such that, for all $r, s \in R$, $m, n \in M$, the following four relationships hold:

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

An *R-submodule of M* is a subgroup N of M which is closed under the action of R. Moreover, if M only contains $\{\mathbf{0}\}$ and M as R-submodules, then M is *irreducible*.

The permutation modules described in section 2.1 are $\mathbb{Q}S_n$-modules. Note that if R is also a field, which is certainly the case with $\mathbb{Q}$, saying that M is an R-module is equivalent to saying that it is a *vector space over R*. Thus, every permutation module M^λ, in addition to being a $\mathbb{Q}S_n$-module, is also a vector space over $\mathbb{Q}$. We will sometimes use these terms interchangeably, as it is easy to think of profiles as vectors in vector spaces.

Often, we will want to decompose our permutation modules into irreducible submodules. The following theorem guarantees that we can do so in our case, when permutation modules are associated with the ring $\mathbb{Q}S_n$.

Theorem 2.1. ***(Maschke's Theorem)*** *Let G be a finite group, and F be a field whose characteristic (the additive order of 1_F) does not divide $|G|$. If M is a nontrivial FG-module, then M can be expressed as the direct sum of irreducible submodules, i.e.,*

$$M \cong N_1 \oplus \cdots \oplus N_k,$$

where each N_i is an irreducible FG-module.

Note that in our case, $F = \mathbb{Q}$. The characteristic of a field is defined to be 0 if 1_F does not have finite order, so the characteristic of $\mathbb{Q}$ does not divide the order of any S_n. Thus Maschke's theorem applies. However, this decomposition may not be unique. Since our goal is to separate a space into particular subspaces based on their behavior under various procedural maps (as Saari did for the fully-ranked case), it would be useful to understand in what ways this decomposition may be unique. An *isotypic component* of M is the direct sum of all isomorphic copies of a particular irreducible submodule in a given decomposition of M. For example, if $M \cong N_1 \oplus N_1 \oplus N_2 \oplus N_3 \oplus N_3$, where N_1, N_2, and N_3 are distinct non-isomorphic irreducible submodules, then the isotypic components of M are $2N_1$, N_2, and $2N_3$. An *isotypic decomposition* then is a decomposition of a module M as in Maschke's theorem, but with irreducible submodules grouped into isotypic components.

Theorem 2.2. *If U_i are isotypic components of M, then the decomposition*

$$M = U_1 \oplus \cdots \oplus U_k$$

is unique up to order of the U_is.

This statements is valuable largely because it gives us a stable decomposition of any profile space.

Definition 2.4. If R is a ring, and M and N are R-modules, an *R-module homomorphism* is defined as a map $\varphi : M \to N$ such that

1. $\varphi(m_1 + m_2) = \varphi(m_1) + \varphi(m_2)$ for all $m_1, m_2 \in M$, and
2. $\varphi(rm) = r\varphi(m)$, for all $r \in R, m \in M$.

The last theorem in this section pertains directly to the behavior of irreducible subspaces under R-module homomorphic maps. In particular, we will be exploring voting procedures as *tally maps*. Thus the following theorem will allow us to discuss irreducible subspaces in terms of their relation to a particular procedure.

Theorem 2.3. ***(Schur's Lemma)*** *If M_1 and M_2 are two irreducible R-modules, then any nonzero R-module homomorphism from M_1 to M_2 is an isomorphism.*

Essentially, this says that for any portion of an irreducible subspace of our profile space to contribute to the outcome of a particular procedure, the entire subspace must contribute. From what we have already seen, we might guess then that, for example, the Reversal space in Saari's decomposition must be the direct sum of irreducible subspaces, as it seems to behave in this manner. However, we must first know if our procedural functions are indeed $\mathbb{Q}S_n$-module homomorphisms.

2.3 Voting Theory in Algebraic Terms

Now that we have built up some of the mathematical background for our approach, we can begin to apply these concepts to voting. First, let's take a critical look at how we might represent our tally procedures in an algebraic framework.

Just as we treat our profiles and outcomes as vectors, we can think of the process of tallying votes as mapping from one vector space to another. If we think of the profile space for rankings of the shape $\lambda \vdash n$ as the permutation module M^λ (as defined in section 2.1), we can also think of the outcome space for the pairwise procedure as the permutation module $M^{(n-2,1,1)}$. This is because each tabloid in $M^{(n-2,1,1)}$ corresponds to a given ordered pair. For example, the pair $A > B$ can be depicted as

$$\begin{array}{l} \hline A \\ \hline B \\ \hline CD\ldots \\ \hline \end{array}.$$

Thus, we can think of the pairwise procedure as a function from M^λ to $M^{(n-2,1,1)}$. Similarly, the positional procedure can be represented as a function from M^λ to $M^{(n-1,1)}$. However, as we have seen, if we fix an order on X^λ, we can also represent profile vectors as elements of $\mathbb{Q}^{|X^\lambda|}$. Thus, we can write these tally maps as matrices.

For example, in the case of fully ranking three candidates, the pairwise function P can be written as

$$P = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Here, we have fixed the ordering of both the full rankings and the ordered pairs lexicographically. The columns correspond to the possible rankings of candidates, where the rows correspond to the ordered pairs. For example, the element in the first column and the first row is a 1 because each vote for the full ranking $A > B > C$ gives one point to the pair $A > B$. However, the element in the first column, but in the last row is a 0 because C does not rank above B in the same full ranking. Note that there may be some ambiguity as to how we might give points in the case of ties in partial rankings. We will address this in Chapter 4, treating the general case of giving t points for ties.

Similarly, if we let $\mathbf{w} = [1, t, 0]$ be our weighting vector, then we can write the positional function $T_{\mathbf{w}}$ as

$$T_{\mathbf{w}} = \begin{pmatrix} 1 & 1 & t & 0 & t & 0 \\ t & 0 & 1 & 1 & 0 & t \\ 0 & t & 0 & t & 1 & 1 \end{pmatrix}.$$

Here, the columns still correspond to the full rankings, but now the rows correspond to each individual candidate. So, in the first column, each vote for $A > B > C$ gives one point to A, t points to B, and zero points to C.

Now, if we wish to apply Schur's lemma, we must first ensure that these maps are indeed $\mathbb{Q}S_n$-module homomorphisms.

Theorem 2.4. *All pairwise and positional maps are $\mathbb{Q}S_n$-module homomorphisms.*

Proof. We begin by describing a general map between permutations modules which will reduce easily to both the positional and pairwise maps. Let M^λ and M^μ be permutation modules for tabloids of shape $\lambda = (\lambda_1, \ldots, \lambda_l), \mu = (\mu_1, \ldots, \mu_m) \vdash n$, such that the partition μ is formed by combining consecutive values of λ_i. For example, if $\lambda = (n-4, 1, 1, 1, 1)$, we might have $\mu = (n-3, 2, 1)$. Let $|X^\lambda| = L$ and $|X^\mu| = N$.

Our restriction of μ simply implies that we are mapping partial rankings into a space which represents examining subrankings. For example, mapping from $\lambda = (n-3, 1, 1, 1)$ to $\mu = (n-2, 1, 1)$ represents counting up rankings of three candidates from n by comparing subsets ranking two candidates from n. This is of course the pairwise map.

Let $\mathbf{v} = (v_1, \ldots, v_N)$ be a vector in $\mathbb{Q}^N$ with to following restriction. Recall that a natural basis of M^λ is $X^\lambda = \{t_i^\lambda\}_{i=1}^L$, the set of all tabloids of shape λ. Fix one of these tabloids and call it t_1^λ. For every $\sigma \in S_n$ such that $\sigma t_1^\lambda = t_1^\lambda$ (σ is in the stabilizer of t_1^λ), if $\sigma t_i^\mu = t_j^\mu$, then $v_i = v_j$. This vector will be used to give points to subsets of candidates.

We know, since S_n acts transitively on X^λ, there is some $\sigma_i \in S_n$ such that $t_i^\lambda = \sigma_i t_1^\lambda$ for all $t_i^\lambda \in M^\lambda$. Let $T_{\mathbf{v}}(t_i^\lambda)$ be defined as follows:

$$T_{\mathbf{v}}(t_i^\lambda) = \sum_{j=1}^{N} v_j \sigma_i t_j^\mu.$$

First, we must check that $T_{\mathbf{v}}$ is well defined. Thus, let $\alpha, \beta \in S_n$ have the property that $\alpha t_1^\lambda = \beta t_1^\lambda = t_i^\lambda$. We know that β can be described as the product of α and an element of the stabilizer of t_1^λ. In other words, $\beta = \alpha\sigma$, where $\sigma t_1^\lambda = t_1^\lambda$. Therefore,

$$\begin{aligned} T_{\mathbf{v}}(\beta t_1^\lambda) &= \sum_{j=1}^{N} v_j \beta t_j^\mu \\ &= \sum_{j=1}^{N} v_j \alpha\sigma t_j^\mu \\ &= \sum_{j=1}^{N} v_j \alpha t_j^\mu \ \text{ (by our restraint on } \mathbf{v}) \\ &= T_{\mathbf{v}}(\alpha t_1^\lambda) \end{aligned}$$

Thus $T_{\mathbf{v}}$ is well defined.

Note now that the following equation holds:

$$T_{\mathbf{v}}(t_1^\lambda) = \sum_{j=1}^{N} v_j t_j^\mu.$$

Recall that $T_{\mathbf{v}}$ is a $\mathbb{Q}S_n$-module homomorphism if

1. $T_{\mathbf{v}}(\mathbf{p}_1 + \mathbf{p}_2) = T_{\mathbf{v}}(\mathbf{p}_1) + T_{\mathbf{v}}(\mathbf{p}_2)$ for all $\mathbf{p}_1, \mathbf{p}_2 \in M^\mu$, and
2. $T_{\mathbf{v}}(\alpha \cdot \mathbf{p}) = \alpha \cdot T_{\mathbf{v}}(\mathbf{p})$, for all $\alpha \in \mathbb{Q}S_n, \mathbf{p} \in M^\mu$.

The first follows directly from the linear nature by which we defined $T_{\mathbf{v}}$. To illuminate the second, we rewrite $T_{\mathbf{v}}(t_i^\lambda)$ as follows:

$$\begin{aligned} T_{\mathbf{v}}(t_i^\lambda) &= \sum_{j=1}^{N} v_j \sigma_i t_j^\mu \\ &= \sigma_i \sum_{j=1}^{N} v_j t_j^\mu \\ &= \sigma_i T_{\mathbf{v}}(t_1^\lambda). \end{aligned}$$

Similarly for any $\gamma \in S_n$,

$$\begin{aligned} T_{\mathbf{v}}(\gamma t_i^\lambda) &= \sum_{j=1}^{N} v_j \gamma \sigma_i t_j^\mu \\ &= \gamma \sum_{j=1}^{N} v_j \sigma_i t_j^\mu \\ &= \gamma T_{\mathbf{v}}(t_i^\lambda). \end{aligned}$$

Extending linearly to all vectors, we see that $T_{\mathbf{v}}(\alpha \cdot \mathbf{p}) = \alpha \cdot T_{\mathbf{v}}(\mathbf{p})$, for all $\alpha \in \mathbb{Q}S_n, \mathbf{p} \in M^\mu$. Therefore $T_{\mathbf{v}}$ is a $\mathbb{Q}S_n$-module homomorphism.

Notice that $T_{\mathbf{v}}$ can be described as a matrix whose columns are permutations of the vector $\mathbf{v}$ according to how the corresponding σ_i permutes the basis elements of M^μ.

In particular, for the positional method, $\mathbf{v}$ is precisely the weighting vector $\mathbf{w}$, λ is the shape of rankings in the profile space, and $\mu = (n-1, 1)$. Our restriction above on $\mathbf{v}$ simply translates to giving candidates who tie the same number of points. Therefore, all positional maps are also $\mathbb{Q}S_n$-module homomorphisms.

Finally, in the case of the pairwise map, we take $\mathbf{v}$ to be the first column vector of the pairs matrix. Again, if σ stabilizes the first ranking, then it only permutes candidates who tie. Thus all pairwise relationships are preserved by swapping these candidates in the pairwise space. For example, if A and B tie in a ranking, then C beats A only if C also beats B. Similarly λ is the shape of rankings in the profile space, and $\mu = (n-2, 1, 1)$. Therefore, the pairwise map is also a $\mathbb{Q}S_n$-module homomorphism.

□

Now that we know that Schur's lemma is applicable to our problem, the next question is how exactly it applies to our permutation modules. To answer this, we will need to have some understanding of how these permutation modules decompose into irreducible subspaces. As discussed in Sagan (11), there is an injection from partitions of n to cyclic modules. For each partition $\mu \vdash n$, we will denote the corresponding module as S^μ. Moreover, when the field from which we draw scalars is $\mathbb{Q}$ (which is true in our case, as asserted previously), these modules are irreducible. Also, the S^μ for $\mu \vdash n$ form a complete list of irreducible $\mathbb{Q}S_n$-modules over $\mathbb{Q}$. Sagan provides a relatively simple algorithm for computing the decomposition of

composition λ	decomposition of M^λ
$(n-1,1)$	$M^\lambda \cong S^{(n)} \oplus S^{(n-1,1)}$
$(n-2,1,1)$	$M^\lambda \cong S^{(n)} \oplus 2S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus S^{(n-2,1,1)}$
$(n-3,1,1,1)$	$M^\lambda \cong S^{(n)} \oplus 3S^{(n-1,1)} \oplus 3S^{(n-2,2)} \oplus 3S^{(n-2,1,1)}$ $\oplus S^{(n-3,3)} \oplus 2S^{(n-3,2,1)} \oplus S^{(n-3,1,1,1)}$
$(1,1,1)$	$M^\lambda \cong S^{(3)} \oplus 2S^{(2,1)} \oplus S^{(1,1,1)}$
$(1,1,1,1)$	$M^\lambda \cong S^{(4)} \oplus 3S^{(3,1)} \oplus 2S^{(2,2)} \oplus 3S^{(2,1,1)} \oplus S^{(1,1,1,1)}$
$(1,\ldots,1) \vdash n$	$M^\lambda \cong S^{(n)} \oplus (n-1)S^{(n-1,1)} \oplus 2(n-3)S^{(n-2,2)}$ $\oplus \frac{1}{2}(n-1)(n-2)S^{(n-2,1,1)} \oplus \cdots \oplus S^{(1,\ldots,1)}$

Table 2.1: A decomposition of some important permutation modules.

M^μ into these irreducible modules. Though we will not be going into this algorithm in depth here, Table 2.3 provides some important examples of decompositions of common spaces.

Again, according to Theorems 2.1 and 2.2, these direct sum decompositions are unique up to isotypic components. Also, note that each of these decompositions has one isomorphic copy of $S^{(n)}$. This irreducible module will correspond the the subspace spanned by the $\mathbf{1}_N$ vector, the all ties space under any tally map For example, let's return to the 3-alternative case discussed on Chapter 1. As mentioned in section 2.1, we can represent the profile space for full rankings on three candidates as $M^{(1,1,1)}$. As we can see from table 2.3, this has the following decomposition:

$$M^{(1,1,1)} \cong S^{(3)} \oplus 2S^{(2,1)} \oplus S^{(1,1,1)}.$$

Again, $S^{(3)}$ corresponds to the all-ones space, or Saari's Kernel (we must be careful here, as we mean something very different by 'kernel'—so we will differentiate through context and capitalization). The dimension of $(S^{(2,1)})$ is 2 and the dimension of $(S^{(1,1,1)})$ is 1. If we look back at Table 1.1, we can quickly see that Saari's Condorcet space must be an irreducible subspace, as it is the only portion of the profile space which is in the kernel of one map, and not in the kernel of the other. Thus, since this is the only subspace of dimension 1 which is not the all-ones space, the Condorcet space is isomorphic to $S^{(1,1,1)}$. Now we are only left with the Basic space and the Reversal space, both of dimension 2. This implies that they must both subspaces of the remaining isotypic, and therefore must both be isomorphic to $S^{(2,1)}$. In the following chapters, we will further explore these kinds of dimensional and homomorphism-related arguments.

Chapter 3

Exploring Full Rankings with Algebraic Theory

Armed with the tools to discuss voting structures in an algebraic context, we will begin with the case of fully ranked data in an attempt to recover some of the previously-established results in the field of voting theory. Namely, we will uncover an algebraic explanation for why the Borda Count is the unique positional tally yielding the fewest dimensions of conflict with the pairwise tally on fully ranked data.

First, as discussed in Sections 2.1 and 2.3, we can represent fully-ranked data on n candidates as elements of the $\mathbb{Q}S_n$-module $M^{(1,\dots,1)}$. Similarly we can represent the pairs and positional tally spaces as $M^{(n-2,1,1)}$ and $M^{(n-1,1)}$, respectively. Recall that both of these tally maps are $\mathbb{Q}S_n$-module homomorphisms. Thus we can apply Schur's lemma (Theorem 2.3) to these maps and immediately reveal from the decompositions in Table 2.3 which portions of this profile space even have the potential of contributing to either the pairwise or positional outcomes. For the pairwise procedure, the map is as follows:

$$\begin{aligned}
\mathbf{P} : \mathbf{M}^{(\mathbf{1},\dots,\mathbf{1})} \\
&\cong S^{(n)} \oplus (n-1)S^{(n-1,1)} \oplus 2(n-3)S^{(n-2,2)} \oplus \binom{n-1}{2} S^{(n-2,1,1)} \oplus \dots \\
\rightarrow \mathbf{M}^{(\mathbf{n-2,1,1})} &\cong S^{(n)} \oplus 2S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus S^{(n-2,1,1)}.
\end{aligned}$$

Therefore, we know that anything after the $S^{(n-2,1,1)}$ isotypic in $M^{(1,\dots,1)}$ must be in the kernel of P since none of the remaining irreducible submodules have corresponding isomorphic copies in the pairs space. So the only portions of the profile space which could potentially contribute to the pairwise outcomes are irreducible subspaces isomorphic to $S^{(n)}$, $S^{(n-1,1)}$, $S^{(n-2,2)}$, or $S^{(n-2,1,1)}$. Similarly, for the positional procedure, we have

$$\begin{aligned} \mathbf{T_w} : \mathbf{M^{(1,\dots,1)}} &\cong S^{(n)} \oplus (n-1)S^{(n-1,1)} \oplus \dots \\ \rightarrow \mathbf{M^{(n-1,1)}} &\cong S^{(n)} \oplus S^{(n-1,1)}. \end{aligned}$$

So any profile not in either the $S^{(n)}$ or $S^{(n-1,1)}$ isotypic components in $M^{(1,\dots,1)}$ must be in the kernel of $T_{\mathbf{w}}$ since none of the remaining irreducible submodules have corresponding isomorphic copies in the positional space.

When we ask the question of how to minimize conflict between the two tally procedures, embedded in that question is one of how to maximize overlap amongst the kernels of each map. In other words, we want it to be the case that if **p** is in the kernel of one map, then it should be in the kernel of both. Otherwise, as in our example outlining Saari's three-alternative case in Chapter 1, we will have unnecessary conflict where profiles tie under one map, but elicit a winner under another. Also, if a profile **p** completely ties under both maps, and is in neither of their kernels, then it must be in the all-ones space ($S^{(n)}$) of the profile space. If it is in neither of their kernels, and is not a complete tie, then the only other possibility is that **p** is in the isotypic $(n-1)S^{(n-1,1)}$. Moreover, if we want to maximize the "number" of these kinds of profiles, it is clear that this will occur when the same copy of $S^{(n-1,1)}$ is in the orthogonal complement to the kernels of both maps.

But how many isomorphic copies of $S^{(n-1,1)}$ could possibly allow for such an agreement? Certainly, the pairs map is fixed in the case of full rankings. So this overlap is entirely dependent on our choice of weighting vector. The decomposition of the pairs space in Table 2.3 indicates that there might be two copies of $S^{(n-1,1)}$ in the image of the pairs map, which could possibly give us more freedom in our choice of weighting vectors. However, this is not the case.

Theorem 3.1. *If $P : M^{(1,\dots,1)} \rightarrow M^{(n-2,1,1)}$ is the pairs map, then*

$$img(P) \cong S^{(n)} \oplus S^{(n-1,1)} \oplus S^{(n-2,1,1)}.$$

In particular, P is not surjective.

Proof. We will show this using a dimension argument. As we will show later (in Theorem 4.1), the pairs map has rank $\binom{n}{2} + 1$. In short, this is because for any arbitrary pair AB, we can recover how many points will be assigned to $B > A$ from the number of points assigned to

$A > B$ along with the total number of votes. Thus, we only need one vector corresponding to each pair (for example, showing one point for $A > B$ and no points for any other ordered pair), along with the all-ones vector to reconstruct the entire image of P. Every other vector in the image of P can be expressed as a sum of these basis vectors.

Now, we will note again that

$$M^{(n-2,1,1)} \cong S^{(n)} \oplus 2S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus S^{(n-2,1,1)}.$$

By Schur's lemma, we know that the image of the pairs map will be a direct sum of a subset of these irreducible submodules. Thus we only need find which integer coefficients a, b, c, and d will give

$$\begin{aligned} \operatorname{rank}(P_t) &= a \cdot \dim(S^{(n)}) + b \cdot \dim(S^{(n-1,1)}) \\ &\quad + c \cdot \dim(S^{(n-2,2)}) + d \cdot \dim(S^{(n-2,1,1)}). \end{aligned} \tag{3.1}$$

These irreducible modules have the following dimensions:

$$\begin{aligned} &\dim S^{(n)} = 1 \\ &\dim S^{(n-1,1)} = n-1 \\ &\dim S^{(n-2,2)} = \tbinom{n-1}{2} - 1 \\ &\dim S^{(n-2,1,1)} = \tbinom{n-1}{2}. \end{aligned}$$

Substituting into equation 3.1 we get the following equation:

$$\binom{n}{2} + 1 = a \cdot 1 + b \cdot (n-1) + c \cdot \left(\binom{n-1}{2} - 1\right) + d \cdot \binom{n-1}{2}.$$

Using the identity $\binom{x}{2} = \frac{1}{2}(x)(x-1)$, and expanding both sides, this yields

$$\begin{aligned} (1/2)n(n-1)+1 &= \\ (1/2)(n^2-n+2) &= a + b(n-1) + c((1/2)(n-1)(n-2)-1) \\ &\quad + d((1/2)(n-1)(n-2)) \\ &= a + nb - b + c((1/2)n^2 - 3/2n) \\ &\quad + d((1/2)n^2 - \frac{3}{2}n + 1) \\ &= (1/2)((c+d)n^2 + (2b-3c-3d)n \\ &\quad + (a-b+2d)). \end{aligned}$$

Since n is arbitrary, this gives us the following system of equations:

$$\begin{aligned} 2a - 2b + 2d &= 2 \\ 2b - 3c - 3d &= -1 \\ c + d &= 1. \end{aligned}$$

This system then has the solutions $b = 1$, $a + d = 2$ and $c + d = 1$. But d can only take on the values 1 and 0. If $d = 0$, then $a = 2$. But a can only take a value of 1 or 0 as well. Thus $d = 1$, which gives $a = 1, b = 1, c = 0$. Thus $\text{img}(P_t) \cong S^{(n)} \oplus S^{(n-1,1)} \oplus S^{(n-2,1,1)}$.

□

Thus, we are forced to construct a positional map which includes one particular copy of this module in the orthogonal complement to the kernel in order to fulfill our maximization criterion above. The following theorem shows that this fact forces a unique weighting vector $\mathbf{w}$.

Theorem 3.2. *Let $T_{\mathbf{w}}$ be the positional tally map on n candidates with weighting vector $\mathbf{w}$. Then $\ker T_{\mathbf{w}} = \ker T_{\mathbf{w}'}$ if and only if $\mathbf{w} = \mathbf{w}'$*

Proof. The fact that $\mathbf{w} = \mathbf{w}'$ implies $\ker T_{\mathbf{w}} = \ker T_{\mathbf{w}'}$ is trivial, so we will simply explore the reverse implication. Let $T^n_{\mathbf{w}}$ be defined as the positional map on n candidates, and let $\mathbf{w} = [1, w_2, \ldots, w_{n-1}, 0]$ be a normalized weighting vector. First, examine the case of maps on two candidates ($n = 2$). This is simply the identity matrix on two elements:

$$T^2_{\mathbf{w}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Vacuously, this theorem holds for $n = 2$, as there is only one weighting vector.

However, this will not prove to be a very useful base for induction, so we look also at the positional map on $n = 3$ candidates with weighting vector $\mathbf{w} = [1, t, 0]$:

$$T^3_{\mathbf{w}} = \begin{pmatrix} 1 & 1 & t & 0 & t & 0 \\ t & 0 & 1 & 1 & 0 & t \\ 0 & t & 0 & t & 1 & 1 \end{pmatrix}.$$

If $t \neq 1, 0$, the null space for this matrix is

$$\mathrm{NS}(T^3_{\mathbf{w}}) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ -(1-t) \\ -1 \\ 1-t \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1-t \\ -(1-t) \\ -t(1-t) \\ 0 \\ t(1-t) \\ 0 \end{pmatrix}, \begin{pmatrix} t^2-t+1 \\ 1-t \\ t \\ 0 \\ 0 \\ -t(1-t) \end{pmatrix} \right\}.$$

This can be verified through simple calculations. For example, take the first of those and take the dot product of it with the first row of $T^3_{\mathbf{w}}$ we get

$$\begin{aligned} 1 \cdot 1 - 1 \cdot (1-t) - t \cdot 1 + 0 \cdot (1-t) + t \cdot 0 + 0 \cdot 0 \\ = 1 - 1 + t - t + 0 = 0. \end{aligned}$$

Also, as long as $t \neq 0, 1$, it is easy to see that these vectors are linearly independent.

If we take another weighting vector $\mathbf{w} = (1, t', 0)$, the only case in which these null spaces would be equal is when $t = t'$. This can be seen by taking the first vector given in the null space above, and substituting t' for t, and multiplying by our original positional matrix. For example, the dot product of this vector with the first row of $T^3_{\mathbf{w}}$ is

$$\begin{aligned} 1 \cdot 1 - 1 \cdot (1-t') - t \cdot 1 + 0 \cdot (1-t') + t \cdot 0 + 0 \cdot 0 \\ = 1 - 1 + t' - t + 0 = t - t'. \end{aligned}$$

Thus, this product is only zero if $t = t'$, and so this vector is only in the null space of our original positional map if $t = t'$. Also, if we calculate the individual null spaces for the cases in which $t = 1$ and $t = 0$, these are also unique from the spaces given above. Therefore, $\ker T^3_{\mathbf{w}} = \ker T^3_{\mathbf{w}'}$ only if $\mathbf{w} = \mathbf{w}'$.

Now, assuming that the theorem holds for maps on $n-1$ candidates, examine the case of maps on n candidates. Let $T^n_{\mathbf{w}}|(A, i)$ be defined as the submatrix of $T^n_{\mathbf{w}}$ formed by the columns where candidate A is ranked in the i^{th} place. For example, observe that the positional map on 4 candidates, where $\mathbf{w} = [1, s, t, 0]$, is as follows:

$$T^4_{\mathbf{w}} = \begin{pmatrix} 1\,1\,1\,1\,1\,1 & s\,s\,t\,0\,t\,0 & s\,s\,t\,0\,t\,0 & s\,s\,t\,0\,t\,0 \\ s\,s\,t\,0\,t\,0 & 1\,1\,1\,1\,1\,1 & t\,0\,s\,s\,0\,t & t\,0\,s\,s\,0\,t \\ t\,0\,s\,s\,0\,t & t\,0\,s\,s\,0\,t & 1\,1\,1\,1\,1\,1 & 0\,t\,0\,t\,s\,s \\ 0\,t\,0\,t\,s\,s & 0\,t\,0\,t\,s\,s & 0\,t\,0\,t\,s\,s & 1\,1\,1\,1\,1\,1 \end{pmatrix}.$$

Columns number $7, 8, 13, 14, 19$, and 20 of $T^4_{\mathbf{w}}$ all correspond to A being placed in second place (where A gets s points), so

$$T^4_{\mathbf{w}}|(A,2) = \begin{pmatrix} s & s & s & s & s & s \\ 1 & 1 & t & 0 & t & 0 \\ t & 0 & 1 & 1 & 0 & t \\ 0 & t & 0 & t & 1 & 1 \end{pmatrix}.$$

If we restrict our attention to this submatrix, we notice that the rows are comprised of a constant vector (in the row corresponding to candidate A), and row vectors consistent with those from the positional map on 3 candidates.

In general, if $T^n_{\mathbf{w}}$ has weighting vector $\mathbf{w} = [1, w_2, \ldots, w_{n-1}, 0]$, then the submatrix $T^n_{\mathbf{w}}|(A,i)$ is of the form $T^{n-1}_{\mathbf{w}}$ with weighting vector $\mathbf{w} = [1, w_2, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{n-1}, 0]$ with an extra constant row of weight w_i.

We now take two positional maps on n candidates with weighting vectors $\mathbf{w} = [1, w_2, \ldots, w_{n-1}, 0]$ and $\mathbf{w}' = [1, w'_2, \ldots, w'_{n-1}, 0]$, and assume that their kernels are identical. Now, for each vector in the null space of $T^{n-1}_{\mathbf{w}}$, we will construct a corresponding vector in the null space of $T^n_{\mathbf{w}}$. First notice that if we take any vector in the null space of $T^{n-1}_{\mathbf{w}}$, then it is also in the null space of $T^n_{\mathbf{w}}|(A,i)$. It is trivial to see that a vector which is in the null space of $T^{n-1}_{\mathbf{w}}$ has a dot product of zero with any row in $T^n_{\mathbf{w}}|(A,i)$ which corresponds to $T^{n-1}_{\mathbf{w}}$. However, it only takes summing the rows of $T^{n-1}_{\mathbf{w}}$ to see that the all-ones vector is in the row space of this map, and is thus orthogonal to everything in the kernel. Thus the dot product of any vector in the kernel of $T^{n-1}_{\mathbf{w}}$ with the all-w_i vector is also zero, so this vector is also in the kernel of $T^n_{\mathbf{w}}|(A,i)$. Therefore, we can treat each of these vectors as though they had been restricted by (A,i), and then expand them by assigning zeros to all of the missing entries. By our assumption, the null spaces of both positional maps ($T^n_{\mathbf{w}}$ and $T^n_{\mathbf{w}'}$) are identical, thus all of these specifically constructed vectors are in the null space of both maps. Thus, $\ker T^{n-1}_{\mathbf{w}'} = \ker T^{n-1}_{\mathbf{w}'}$.

For example, the vector $\mathbf{v} = (1, t-1, -1, 1-t, 0, 0)^T$ is in the null space of $T^3_{\mathbf{w}=[1,t,0]}$, so it is also in the null space of $T^4_{\mathbf{w}=[1,s,t,0]}|(A,2)$. Thus

$$\bar{\mathbf{v}} = (0,0,0,0,0,0,\ 1, t-1, 0,0,0,0,\ -1, 1-t, 0,0,0,0,\ 0,0,0,0,0,0)^T$$

is in the null space of $T^4_{\mathbf{w}=[1,s,t,0]}$. Finally, restricting again so the same columns of $T^4_{\mathbf{w}'=[1,s',t',0]}$, $\mathbf{v}$ is in the null space of $T^3_{\mathbf{w}'=[1,t',0]}$. Since this can be done for all vectors in the null space of either map on 3 candidates, it follows that $\ker T^3_{\mathbf{w}'=[1,t',0]} = \ker T^3_{\mathbf{w}'=[1,t',0]}$.

But, by our inductive hypothesis, this means that the corresponding submaps must be identical. Rather, $\ker T^{n-1}_{\mathbf{w}'} = \ker T^{n-1}_{\mathbf{w}'}$ implies that these reduced weighting vectors are equal. By the nature of our restriction to $T^n_{\mathbf{w}}|(A, i)$, this ensures that $w_j = w'_j$ for all $j \neq i$. In our example, we showed specifically that $t = t'$. But since i was arbitrary within the bounds of $2 \leq i \leq n-1$, this implies that $w_j = w'_j$ for all $2 \leq j \leq n-1$. Thus $\mathbf{w} = \mathbf{w}'$ for the positional map on n candidates.

□

From this result, we now have that if there is a weighting vector which could minimize conflict between the pairwise and positional tallies in this way (in terms of compatible kernels), then it is unique. Thus we are only left with the task of finding the unique weighting vector satisfying this requirement.

Notice, if we construct some $\mathbb{Q}S_n$-module homomorphism

$$\varsigma : M^{(n-2,1,1)} \to M^{(n-1,1)},$$

we can compare the kernels of the pairs and positional maps through their images. Namely, we will try to construct ς such that

$$\varsigma \circ P(\mathbf{p}) = T_{\mathbf{w}}(\mathbf{p}),$$

for all profiles $\mathbf{p} \in M^{(n-1,1)}$. In other words, we want the diagram in Figure 3.1 to commute.

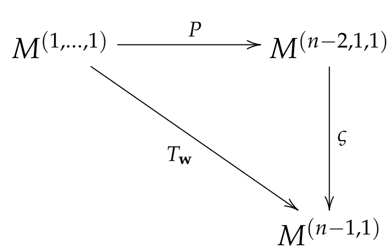

Figure 3.1: Diagram of relationships between profile and tally spaces for fully ranked data.

This is useful because if we can find some ς for which $\varsigma \circ P$ is a surjective map, then it will have a kernel isomorphic to that of $T_{\mathbf{w}}$. However, these kernels will be equal *only* if the kernel of P is contained in the kernel of $T_{\mathbf{w}}$. We can see this by taking some profile $\mathbf{p}$. If $\mathbf{p}$ is in the kernel of P, then it is certainly in the kernel of $\varsigma \circ P$. Therefore, if these maps commute with $T_{\mathbf{w}}$, then $\mathbf{p}$ is also in the kernel of this positional map. Similarly, if $\mathbf{p}$ is in the kernel of $T_{\mathbf{w}}$, then it must also be in the kernel of $\varsigma \circ P$, and in turn, P.

One such $\varsigma : M^{(n-2,1,1)} \to M^{(n-1,1)}$ would be to map an outcome in $M^{(n-2,1,1)}$, to, for each candidate A, the sum of the points for $A > X$ (over all other candidates X). For the sake of normalizing results, we will then divide this result by $n-1$. For example, if we had an outcome which gave one point for the pair $A > B$, and one point for $A > C$, this would map under ς to two points for A, and no points elsewhere (all divided by $n-1$). This map must have the desired kernel, as $\varsigma \circ P$ is surjective. And indeed, this is true for the Borda Count!

Theorem 3.3. *If we take P to be the standard pairs map, ς to be defined as above, and $T_{\mathbf{w}}$ to be the positional tally defined as the Borda Count, then $\varsigma \circ P(\mathbf{p}) = T_{\mathbf{w}}(\mathbf{p})$, for all profiles $\mathbf{p} \in M^{(n-1,1)}$.*

Proof. Let's take the basis vector $e_1 = (1, 0, \ldots, 0)$, which is equivalent to one vote for the ranking $X_1 > X_2 > ... > X_n$ and no votes otherwise. Under the pairs map, this profile gets tallied as one point for each pair $X_i > X_j$ for all $i > j$, and no points otherwise. Under ς, this then maps to the positional outcome of $\frac{n-i}{n-1}$ points for each candidate X_i. First note that this is clearly not a tie, so must have a non-zero projections into $S^{(n-1,1)}$. Since these irreducible submodules are cyclic, it must be that $\varsigma \circ P$ is indeed surjective. Now, if we take our original profile and map it directly to the positional space using the Borda Count, we clearly get the same result. Thus $T_{\mathbf{w}}(e_1) = \varsigma \circ P(e_1)$. Since these are $\mathbb{Q}S_n$-module homomorphisms, if this is true for one elementary basis vector, this must be true for all elements of the profile space. Thus, for all $\mathbf{p} \in M^{(1,\ldots,1)}$,

$$T_{\mathbf{w}}(\mathbf{p}) = \varsigma \circ P(\mathbf{p}).$$

Thus, the above diagram commutes.

□

Theorem 3.4. *The Borda Count is the unique positional weighting scheme which minimizes conflict with the pairwise map for fully-ranked data.*

This concludes our exploration of the fully-ranked case. However, not only does this result agree with previous work done by Saari, but it generalizes nicely to partially ranked data. In the following chapter, Chapter 4, we will use similar techniques to find an analog to the Borda Count using similar maximization criteria.

Chapter 4

Extending to Partial Rankings

Although much theory has been developed around voting on full rankings, it is not always practical to ask voters for that much information. For example, as mentioned before, if there were even just ten candidates, asking for a full ranking would give voters $10! = 3,620,800$ choices. Thus, it may prove useful to extend the theory to include partial rankings.

As mentioned in Section 2.3, we can view profile spaces as $\mathbb{Q}S_n$-modules. Namely, if we ask for data of a shape λ, then we can represent the profile space for this data as the permutation module M^λ. Of course, there are partial rankings which cannot be represented by a simple combinatorial composition. However, many of the natural partial ranking structures *can* be represented this way—and we are restricting our studies to this kind of data. In fact, we concentrate specifically on data of the shape $\lambda = (n-k, 1, \ldots, 1)$—ranking k candidates out of n.

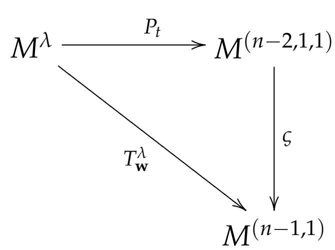

Figure 4.1: Diagram of relationships between profile and tally spaces for partially ranked data.

Drawing on the ideas developed in Chapter 3, we begin with a similar diagram of our primary tally maps, the diagram seen in Figure 4.1. In this diagram, $M^{(n-2,1,1)}$ and $M^{(n-1,1)}$ are still the pairwise and positional tally spaces as in Chapter 3, and ς is the same map from $M^{(n-2,1,1)}$ to $M^{(n-1,1)}$. The map $T_{\mathbf{w}}^{\lambda} : M^{\lambda} \to M^{(n-1,1)}$ is exactly the positional map that we expect it to be—simply giving the same amount of points to every candidate ranked at the same level, according to the weighting vector **w**. For example, if we were ranking two candidates out of ten, our weighting vector might be $\mathbf{w} = [1, 1/2, 0]$—this represents giving 1 point to the first place candidate in each ranking, 1/2 points to the second-place candidate, and no points to any of the others.

The general pairs map, P_t, needs a bit more attention at this stage however. Namely, we must define how the pairs map treats ties. One might first think, in the case that two candidates tie, to give each of them a point. Or, perhaps we should give neither of them any points. Just as easily, we might choose to let them split the point. However, just as was the case for the positional map, a decision that might first appear as frivolous may end in significant discrepancies. For example, as we will see in Theorem 4.1, the rank of the general pairs map depends on our choice of this new parameter. Thus, by changing this parameter, we have the potential to vary the amount of freedom we have in choosing a compatible weighting vector. So, let us define the general pairs map P_t to give an ordered pair AB 1 point when A is ranked strictly higher than B, t points when A ties with B, and 0 points when A strictly loses to B.

Theorem 4.1. *For data of the shape $\lambda = (n-k, 1, \ldots, 1)$, the general pairs map P_t has rank $\binom{n}{2} + 1$ when $n = k+1$ or $t = 1/2$. Otherwise, it has full rank of $n(n-1)$.*

Proof. Every entry in a vector **v** in the image of P_t corresponds to some ordered pair of candidates, say $A > B$. The number of points given to this ordered pair is is the number of votes ranking A strictly above B (denote this x_{AB}) plus t times the number of votes ranking A the same as B (denote this y_{AB}). Rather,

$$\mathbf{v}_{A>B} = x_{AB} + t y_{AB}.$$

If we know only this value for any one pair, we cannot reconstruct any other value in **v**. Specifically, the number of points assigned to any ordered pair $A > B$ is independent of the number of points assigned to any other ordered pair besides $B > A$, even if we also know the total number of votes. For example, if we know $\mathbf{v}_{A>B}$, we have no way of calculating $\mathbf{v}_{A>C}$. Therefore, the rank of P_t must be at least $\binom{n}{2}$. However, if we also know the total number of people who voted, m, there are circumstances in which we will be able to recover the value assigned to $B > A$: $\mathbf{v}_{B>A} = x_{BA} + ty_{BA}$. When this is the case, we would only need $\binom{n}{2} + 1$ vectors to reconstruct anything in the image of P_t—one vector corresponding to each (unordered) pair of candidates, and one representing the number of votes cast (represented by some some scaling of the all-ones vector). This is not to say that we can recover the number of votes cast from a result without any other information, but that the ones vectors, which can serve to count this, is in the image of the pairs map. If we could not generically calculate the value $\mathbf{v}_{B>A}$ from the value $\mathbf{v}_{A>B}$ and another generic piece of information, then the image of the pairs map would be the entire space (all of the entries in $\mathbf{v}$ would be independent of each other), implying that P_t would have a rank of $n(n-1)$.

The case when $\mathbf{v}_{B>A}$ is recoverable from $\mathbf{v}_{A>B}$ and m is exactly when we can find non-trivial values of a, b, and c such that

$$a(x_{AB} + ty_{AB}) + b(x_{BA} + ty_{BA}) + c(m) = 0.$$

Notice first that $y_{AB} = y_{BA}$, as this is just the number of times A and B tie, which is symmetric. Also, note that $x_{AB} + x_{BA} + y_{AB} = n$. In other words, all of the votes can be accounted for by counting those ranking A above B, those ranking B above A, and those ranking them the same. These equalities imply that

$$x_{AB}(a + c) + x_{BA}(b + c) + y_{AB}(ta + tb + c) = 0.$$

Since x_{AB}, x_{BA}, and y_{AB} are arbitrary (depending only on the profile of votes which generated them), we get the system of equations

$$a + c = 0 \qquad b + c = 0 \qquad ta + tb + c = 0.$$

This, in turn, gives us that either $a = b = c = 0$, or $t = 1/2$. However, if, instead, we fix $y_{AB} = 0$ (which is always the case only when we are asking for full rankings, i.e., $n = k + 1$), we get that $a = b = -c$, for any value of a. Thus, $\text{rank} P_t = \binom{n}{2} + 1$ when $t = 1/2$ or $\lambda = (1, \ldots, 1)$, and $\text{rank} P_t = n(n-1)$ otherwise.

□

In order to build some intuition for how to choose a value of t for P_t, we will relate it to P and the *full* ranking case. Namely, we will create an

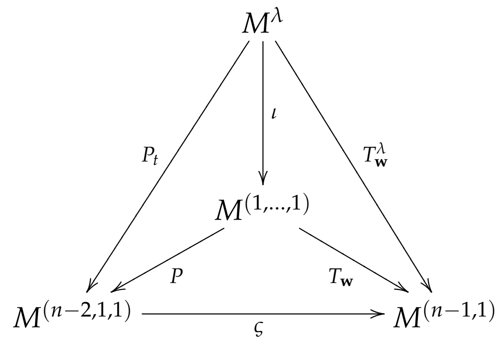

Figure 4.2: Complete diagram of relationships between profile and tally spaces.

embedding (or lift) $\iota : M^{\lambda} \to M^{(1,\dots,1)}$. This relationship allows us to define P_t (and eventually $T^{\lambda}_{\mathbf{w}}$) in terms of the maps on fully ranked data, with which we are very familiar. The diagram in Figure 4.2 serves to summarize these relationships.

In this diagram, the lower triangle corresponds to our study of the fully ranked case pictured in 3.1. The outer triangle is the analogous correspondence for the partially ranked case, as pictured in 4.1. As shown, the surjection from $M^{(n-2,1,1)}$ to $M^{(n-1,1)}$ is the same for both. Again, the map $\iota : M^{\lambda} \to M^{(1,\dots,1)}$ is just the embedding of the partially ranked data into the fully ranked data. The goal then is to construct our maps such that every subdiagram in this larger picture commutes. Namely, given a specific embedding, we may be able to induce our choices of t and $\mathbf{w}$ for the partially ranked mappings from our results on fully ranked data.

One natural choice for ι is to take the linear extensions of the partial rankings. This is to say that every partial ranking gets mapped to the sum of the full rankings which preserve the order imposed by the partial ranking. To allow for normalization, we will then divide through by the total number of elements in the sum. For example, if we have the partial ranking

$$\begin{array}{c}\hline A\\ \hline BC\\ \hline DE\\ \hline\end{array},$$

this would get mapped to the following direct sum of full rankings

$$1/4 \cdot \left(\begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix} \quad + \quad \begin{matrix} A \\ B \\ C \\ E \\ D \end{matrix} \quad + \quad \begin{matrix} A \\ C \\ B \\ D \\ E \end{matrix} \quad + \quad \begin{matrix} A \\ C \\ B \\ E \\ D \end{matrix} \right).$$

In what follows, we concentrate on data of the shape $\lambda = (n-k, 1, \ldots, 1)$ (ie, ranking the top k candidates out of n).

Now that we have an embedding, we focus our attention back on the general pairs map, P_t. To induce a value of t, we will ask that the left-most diagram in Figure 4.2 commute—or, rather, that Figure 4.3 commutes.

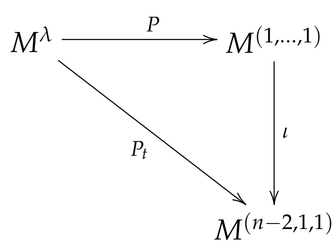

Figure 4.3: Pairs map diagram.

Another way of stating this is that we require that, for all $\mathbf{p} \in M^\lambda$, $\iota \circ P_t(\mathbf{p}) = P(\mathbf{p})$. And, in fact, this happens only when we let $t = 1/2$.

Theorem 4.2. *For partially ranked data of the shape $\lambda = (n-k, 1, \ldots, 1)$, we have $P_t(\mathbf{p}) = \iota \circ P(\mathbf{p})$, for all $\mathbf{p} \in M^\lambda$ if and only if $t = 1/2$.*

Proof. Let $\mathbf{e}_{t^\lambda}$ be a standard basis element in M^λ, namely a profile which has one vote for a single ranking t_λ, and no votes for any other ranking. For every pair of candidates A and B, if $A > B$ in $\mathbf{e}_{t^\lambda}$, then $A > B$ also in every permutation summand in the corresponding element of $M^{(1,\ldots,1)}$. Thus, one point will be given to $A > B$ in the pairs space regardless of whether we map it by P_t or by ι and then by P. Similarly, if $B > A$, then $A > B$ would get zero points in the pairs space by either map. However, if $A \sim B$ in the T^λ, then it is easy to see that A will beat B in exactly half of the summands of the linear extension of t^λ, and B will beat A in the other half. Thus both $A > B$ and $B > A$

will get $1/2$ of a point if we map $\mathbf{e}_{t^\lambda}$ by way of $P \circ \iota$. Thus, in order to force $P_t(\mathbf{e}_{t^\lambda}) = P \circ \iota(\mathbf{e}_{t^\lambda})$, we must define $t = 1/2$. Recall that $\mathbb{Q}S_n$ acts transitively on M^λ—meaning that we can generate all of M^λ from one basis element by acting on it by elements of $\mathbb{Q}S_n$—and that all three of these maps are $\mathbb{Q}S_n$-module homomorphisms. Therefore, since we have proven that t must be equal to $1/2$ for one basis element, it must be true that $t = 1/2$ in general. In short, $\iota \circ P_t(\mathbf{p}) = P(\mathbf{p})$, for all $\mathbf{p} \in M^\lambda$ if and only if $t = 1/2$.

□

Now, from Theorem 4.1, we know that the rank of $P_{1/2}$ is equal to $\binom{n}{2} + 1$. In turn, by the same dimensions argument used in Theorem 3.1, we have that

$$\text{img}(P_{1/2}) \cong S^{(n)} \oplus S^{(n-1,1)} \oplus S^{(n-2,1,1)}.$$

Recall that Theorem 3.2 states that $\ker T_{\mathbf{w}} = \ker T_{\mathbf{w}'}$ if and only if $\mathbf{w} = \mathbf{w}'$ for full rankings. Similar arguments can be used to show that this result holds for partial rankings as well.

Theorem 4.3. *Let $T^\lambda_{\mathbf{w}}$ be the positional tally map on data of the shape $\lambda = (n - k, 1, \ldots, 1)$ with weighting vector $\mathbf{w}$. Then $\ker T^\lambda_{\mathbf{w}} = \ker T^\lambda_{\mathbf{w}'}$ if and only if $\mathbf{w} = \mathbf{w}'$.*

Proof. The following proof will largely mimic that of Theorem 3.2. If we wish to induct on the number of candidates n, our base case for $\lambda = (n - k, 1, \ldots, 1)$ is bounded below by $n = k + 1$, as any smaller quantity has no meaning. However, this implies that the base case for induction will be on data of the shape $\lambda_{\text{base}} = (1, \ldots, 1)$, which is a full ranking. By Theorem 3.2, $\ker T^{\lambda_{\text{base}}}_{\mathbf{w}} = \ker T^{\lambda_{\text{base}}}_{\mathbf{w}'}$ if and only if $\mathbf{w} = \mathbf{w}'$. We now fix k and induct on n—finding a submatrix of the positional map on n candidates that mimics the map on $n - 1$ candidates.

Assume $\ker T^{(n-k,1,\ldots,1)}_{\mathbf{w}} = \ker T^{(n-k,1,\ldots,1)}_{\mathbf{w}'}$. In this case, we pick out the rows which correspond to one candidate being bottom ranked. Rather, our restriction will be $T^{(n-k,1,\ldots,1)}_{\mathbf{w}}|(A, n)$. For example, the positional map for $\lambda = (2, 1, 1)$ is as follows:

$$T^{(2,1,1)}_{\mathbf{w}} = \begin{pmatrix} 1 & 1 & 1 & t & 0 & 0 & t & 0 & 0 & t & 0 & 0 \\ t & 0 & 0 & 1 & 1 & 1 & 0 & t & 0 & 0 & t & 0 \\ 0 & t & 0 & 0 & t & 0 & 1 & 1 & 1 & 0 & 0 & t \\ 0 & 0 & t & 0 & 0 & t & 0 & 0 & t & 1 & 1 & 1 \end{pmatrix}.$$

Then $T^{\lambda}_{\mathbf{w}}|(A,n)$ is the submatrix of $T^{\lambda}_{\mathbf{w}}$ restricted to the columns where A is ranked last. Then

$$T^{(2,1,1)}_{\mathbf{w}}|(A,4) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & t & 0 & t & 0 \\ t & 0 & 1 & 1 & 0 & t \\ 0 & t & 0 & t & 1 & 1 \end{pmatrix}.$$

The rows of $T^{(n-k,1,\dots,1)}_{\mathbf{w}}|(A,n)$ are precisely the same as the rows of $T^{(n-1-k,1,\dots,1)}_{\mathbf{w}}$ with an extra row of zeros. Just as in Theorem 3.2, we can construct a profile in the kernel of $T^{(n-k,1,\dots,1)}_{\mathbf{w}}$ for each profile in the kernel of $T^{(n-k,1,\dots,1)}_{\mathbf{w}}|(A,n)$ (and the same for $T^{(n-k,1,\dots,1)}_{\mathbf{w}'}$ and $T^{(n-k,1,\dots,1)}_{\mathbf{w}'}|(A,n)$). We simply extend the profile by assigning zeros to all of the missing entries. Since we assumed that $\ker T^{(n-k,1,\dots,1)}_{\mathbf{w}} = \ker T^{(n-k,1,\dots,1)}_{\mathbf{w}'}$, by taking profiles in the kernel of $T^{(n-1-k,1,\dots,1)}_{\mathbf{w}}$, extending them, and then restricting again them for $T^{(n-1-k,1,\dots,1)}_{\mathbf{w}'}$, this forces that $\ker T^{(n-1-k,1,\dots,1)}_{\mathbf{w}} = \ker T^{(n-1-k,1,\dots,1)}_{\mathbf{w}'}$. Since we did not drop any weights in this case, this immediately implies that $\mathbf{w} = \mathbf{w}'$.

□

Thus, by the same reasoning that we applied to full rankings, there is one particular unique weighting vector which will serve to complement the pairs map.

In changing focus to the positional maps, we will have to address a fine-point of commutativity. As we will see, if we take some basis element in $M^{(n-k,1,\dots,1)}$, and map it through $T_{\mathbf{w}} \circ \iota$, we will get non-zero points for the bottom-ranked candidates in our original partial ranking. However, if we recall that adding multiples of the all-ones vector and scaling both do not affect the outcome of an election, we can still design a weighting vector which will coincide nicely with the functional composition of ι and $T_{\mathbf{w}}$. Specifically, we want a weighting vector $\mathbf{w}$ such that $T_{\mathbf{w}_{\text{Borda}}} \circ \iota(\mathbf{p}) = T^{\lambda}_{\mathbf{w}}(\mathbf{p}) - q \cdot \mathbf{1}_n$, where $q \cdot \mathbf{1}_n$ is some rational multiple of the all-ones vector in the positional space.

Again take a partial ranking from $M^{(n-k,1,\dots,1)}$. If we pass it through ι and then $T_{\mathbf{w}}$, we will get a Borda-like progression for the top k candidates– the first place candidate will get 1 point, and the next $k-1$ candidates after this will get sequentially $\frac{1}{n-1}$ fewer points. If we consider then that every candidate who was in the bottom tier will occupy each of the lower rankings the same number of times in the summands, it becomes clear that the number of points each of these bottom-ranked candidates will get will be equal to

$$b \equiv \frac{1}{n-k} \cdot \sum_{i=1}^{n-k} \frac{i-1}{n-1}.$$

For example, observe what happens to a profile which gives one vote to the ranking,

$$t^\lambda = \frac{\overline{A}}{\underline{\overline{B}}}\atop{\underline{C\ D\ E}}.$$

If we map $\mathbf{e}_{t^\lambda}$ into $M^{(1,\ldots,1)}$, we get

$$\iota(\mathbf{e}_{t^\lambda}) = 1/6 \cdot \begin{pmatrix} A & & A & & A & & A & & A & & A \\ B & & B & & B & & B & & B & & B \\ C & + & C & + & C & + & C & + & C & + & C \\ D & & D & & E & & E & & F & & F \\ E & & F & & E & & F & & E & & F \\ F & & E & & F & & E & & F & & E \end{pmatrix}.$$

Then, applying the Borda Count to this profile,

$$T_{\mathbf{w}} \circ \iota(\mathbf{e}_{t^\lambda}) = \begin{pmatrix} 1 \\ 4/5 \\ 3/5 \\ 1/6(2 \cdot 2/5 + 2 \cdot 1/5 + 2 \cdot 0) \\ 1/6(2 \cdot 2/5 + 2 \cdot 1/5 + 2 \cdot 0) \\ 1/6(2 \cdot 2/5 + 2 \cdot 1/5 + 2 \cdot 0) \end{pmatrix} = \begin{pmatrix} 1 \\ 4/5 \\ 3/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{pmatrix}.$$

As we can see, this vector has non-zero values assigned to the bottom-ranked candidates. This might imply that there would be no weighting vector which would commute with $P \circ \iota$. However, as mentioned previously, we can subtract multiples of the all-ones vector ($1/5 \cdot \mathbf{1}$ in our example), and scale (again, by $\frac{1}{1-1/5}$ in this example). The result for our example would then be that $T_{\mathbf{w}} \circ \iota(\mathbf{e}_{t^\lambda})$ is equivalent to $(1, 3/5, 2/4, 0, 0, 0)^T$.

We now solve for these values in general. First, we simplify b:

$$\begin{aligned} b &= \frac{1}{n-k} \cdot \sum_{i=1}^{n-k} \frac{i-1}{n-1} \\ &= \frac{1}{(n-k)(n-1)} \cdot \left[\sum_{i=1}^{n-k} i + \sum_{i=1}^{n-k} 1 \right] \\ &= \frac{n-k+1}{2(n-1)} - \frac{1}{n-1} \\ &= \frac{1}{2}\left(1 - \frac{k}{n-1}\right) \end{aligned}$$

Thus, if we take the Borda Count on n candidates for the first k candidates, and normalize this by subtracting b from every value, and then dividing through by $1-b$, we get the i^{th} value (where $1 \leq i \leq k$) to be

$$\begin{aligned} w_i &= \frac{\frac{n-i}{n-1} - \frac{1}{2}\left(1 - \frac{k}{n-1}\right)}{1 - \frac{1}{2}\left(1 - \frac{k}{n-1}\right)}, \\ &= 1 - \frac{2(i-1)}{n+k-1} \end{aligned}$$

As stated above, we can expect only one such weighting vector (up to normalization). Thus, if we choose the embedding of linear extensions, this proves the following theorem.

Theorem 4.4. *The unique analog to the Borda Count for data of the shape $\lambda = (n-k, 1, \ldots, 1)$ is a positional map which gives the i^{th} place candidate $1 - \frac{2(i-1)}{n+k-1}$ points for $1 \leq i \leq k$ and zero points for all last-place candidates.*

Chapter 5

Conclusion and Further Work

We have been able to construct an approach to comparing voting methods using algebraic techniques. Using this approach, we were able to recover established results pertaining to the fully-ranked case. Specifically, we found that among the positional methods, the Borda count elicits the least amount of conflict with the pairwise method of tallying votes. We have also been able to derive similar results for a particular shape of partially ranked data. Namely, for data which ranks k candidates from n, the analagous positional method is a scaled linear modification of the Borda Count as given in Theorem 4.4.

There is still much work to be done in this area, particularly with respect to partially ranked data. Firstly, one could explore what arises in the case that the general pairs map gives some weight other than $\frac{1}{2}$ to ties. As given in Chapter 4, the rank of the pairs map drastically increases with any such change. Due to this increase in rank, preliminary computational results have implied that other general pairs maps allow for more freedom in a choice of positional map. Since the image of the pairs map has one more isomorphic copy of the irreducible $\mathbb{Q}S_n$-module $S^{(n-1,1)}$, it appears as though there is a one-parameter family of weighting vectors which will elicit the least amount of conflict with the pairs map. For example, if we choose to rank 3 from 5 candidates, with a weighting vector of $\mathbf{w} = [1, s, t, 0]$, we get a relationship of $t = 2s - 1$ guaranteeing maximal agreement with the pairs map, with varying intervals of s.

Secondly, it would be a natural extension of this work to apply these methods to other shapes of data. In particular, it may be of interest to consider pairs maps which treat separate classes of ties differently. For example, suppose we ask for rankings in the shape $\lambda = (1, 2, n-3)$. If there is a vote for

$$\begin{array}{l} \hline A \\ \hline BC \\ \hline DE\ldots \\ \hline \end{array},$$

we might consider a pairs map which would give each of B and C $\frac{3}{4}$ of a point, but would give each of D and E $\frac{1}{2}$ of a point. In this light, we could assign the pairs map a similar kind of weighting vector to that of the positional map, which designates how it will treat various classes of ties. However, if we continue to focus on partial rankings which correspond to compositions, and use linear extensions to relate the full and partial rankings, the pairs map is still restricted to giving each candidate in any class of tie $\frac{1}{2}$ of a point.

Finally, it may also be of interest to attempt to apply similar techniques to other voting methods. For example, it may be possible to express approval voting using similar algebraic tools. However, after short consideration it becomes apparent that methods which require stages of tallying or elimination of candidates (such as instant run-off elections) are more difficult to approach from our perspective. This is not to say, though, that algebraic techniques will not prove useful in approaching these methods.

Regardless of the particular methods which we choose to study, we hope that the perspective portrayed here will help to facilitate a new discussion about voting. The study of voting theory in the context of mathematics allows the more quantitatively-minded to understand how voting methods compare. Beginning to develop this study in the language of algebra and representation theory opens the doors for a new branch of mathematicians.

Bibliography

[1] David S. Dummit and Richard M. Foote. *Abstract Algebra*. Upper Saddle River, N.J., Prentice Hall, 1999. *Algebra textbook*.

[2] Lewis Lum and David J. Kurtz. Voting made easy: A mathematical theory of election procedures, 1989. *A quick guide to Arrow's theorem*.

[3] Iain McLean and Arnold B. Urken. *Classics of Social Choice*. The University of Michigan Press, 1995. *Contains translated reprints of major papers from early voting theorists and commentary on their context and content.*

[4] Thomas C. Ratliff. A comparison of Dodgson's method and the Borda count. *Econom. Theory*, 20(2):357–372, 2002. *This paper takes a similar geometric approach to that of Saari, but compares Borda to a modification of Condorcet's method. In particular, it points out that more problems may arise when we seemingly improve upon tallying methods in one way or another.*

[5] Donald G. Saari. Symmetry, voting, and social choice. *Math. Intelligencer*, 10(3):32–42, 1988. *This article is a colloquial introduction to Saari's approach to voting theory. It contains many good examples of various kinds of paradoxes that arise from different methods of tallying votes.*

[6] Donald G. Saari. Explaining all three-alternative voting outcomes. *J. Econom. Theory*, 87(2):313–355, 1999. *This artical mostly concentrates on decomposing the voting space in the case of 3 candidates. This is a good introductory article to Saari's geometric methods in that it stays in the context of an example to help simplify the ideas*.

[7] Donald G. Saari. Mathematical structure of voting paradoxes. I. Pairwise votes. *Econom. Theory*, 15(1):1–53, 2000. *This is the first of two papers which go into depth in comparing pairwise and positional tally methods. They give a thorough description of voting spaces for fully ranked data using geometric methods.*

[8] Donald G. Saari. Mathematical structure of voting paradoxes. II. Positional voting. *Econom. Theory*, 15(1):55–102, 2000. *This is the second of two major papers on analyzing voting methods. See "Mathematical structure of voting paradoxes. I. Pairwise votes".*

[9] Donald G. Saari. *Decisions and Elections: Explaining the Unexpected*. Cambridge University Press, 2001. *This book spends extra time building intuition for why our expectations might be as they are. It was written in a similar tone to his article "Symmetry, Voting, and Social Choice", in that it is geared for the educated layman. It does, however, also cover Saari's geometric methods.*

[10] Donald G. Saari and Fabrice Valognes. Geometry, voting, and paradoxes. *Math. Mag.*, 71(4):243–259, 1998. *This paper is similar to the three outcome paper in that it takes a simple approach to Saari's geometric methods. It is a good introductory paper for getting a grasp on the geometric approach to voting.*

[11] Bruce Sagan. *The Symmetric Group : Representations, Combinatorial Algorithms, and Symmetric Functions*. Springer-Verlag New York, Inc., 2001. *A significant amount of theory for tabloids and irreducible subgroups for partially ranked data can be found here.*