Discussion Paper No. 638

THE OPTIMAL RANKING METHOD IS THE BORDA COUNT

by

Donald G. Saari

January 1985

THE OPTIMAL RANKING METHOD IS THE BORDA COUNT

Donald G. Saari
Department of Mathematics
Northwestern University
Evanston, Illinois 60201

January, 1985

Central to social choice is the development of techniques to aggregate individual rankings of N alternatives into a group ranking. Many approaches exist, but if $N \geq 3$, none does what we really what it to do. The difficulty is that although it is common to treat the group's ranking as if it were transitive, it need not be. In this paper, I'll analyze this social choice problem for voting methods to show what can occur. (This is where, for a given set of weights w_1, .., w_N, w_j points are tallied for a voter's j^{TH} place alternative.) For instance, it is standard to claim that plurality voting is among the worse methods that can be used. We support this assertion by characterizing the inconsistencies of its election results. (In a related paper [10], it is shown that the proposed reform method of "approval voting" [2] has features even worse than plurality voting.) Then, I'll propose a resolution for this social choice problem by determining what is the "best" voting method.

To see the problem, consider a hypothetical situation where nine people select a common luncheon beverage. Four of them have the ranking beer (b) over wine (wi) over water (wa) (b>wi>wa), three have the ranking wa>wi>b, and two have the ranking wi>b>wa. By use of the plurality voting scheme (only your first place alternative is tallied), the group's ranking is b>wa>wi. If this ranking were transitive, then, should beer be unavailable, water would be the group's second choice. But, 2/3 of these people prefer wine to water; indeed, a majority of them prefer wine to beer! The rankings of the pairs can be reversed when considered separately; thus the outcomes of plurality voting need not be consistent.

As it is well known, Arrow's theorem [1] asserts for $N \geq 3$ that this phenomenon occurs for <u>any</u> non-dictatorial aggregation technique which satisfies certain standard conditions. It always is possible to find an example of voters' preferences where the group's ranking of $N \geq 3$ alternatives is not consistent with how the same group, using the same (or any other specified) procedure, ranks some pair of alternatives. Universal consistency of the outcome is an impossibility.

Yet, decisions must be made, so individual rankings must be aggregated into a group ranking. Consequently, even though all voting methods are flawed, we need to determine the "best" one. To do this, the goal for the selection of an aggregation technique must be relaxed. Our unrealistic dream was to find a procedure which always yields a transitive ordering; a more realistic objective is to find those voting techniques which minimize the damage to consistency. We will show that <u>for voting methods, the Borda Count is the unique answer.</u> This is where n-j points are tallied for a voter's j^{TH} place alternative. (A related issue arises for certain ranking methods of nonparametric statistics. Again, the resolution is the Borda Count.)

The Borda Count is optimal for several reasons; the first is with respect to the rankings of pairs of alternatives. If some alternative is preferred to all others by majority votes (i.e., it is a Condorcet winner), then it shouldn't be ranked last in the ranking of the N alternatives. But, the introductory example illustrates that a Condorcet winner (wine) can be ranked last by the plurality vote. We show that <u>the Borda Count is the unique method which never ranks a Condorcet winner in last place, nor a Condorcet loser in first.</u>

If a voting method is to be judged superior, it must be decisive; it must admit fewer ranking inconsistencies than any other method. To investigate this question, I'll introduce some natural measures of the inconsistencies permitted by a voting method. Again, it will turn out that the Borda Count is the unique, best solution.

To gain a flavor of the type of measures which will be used, consider the set of four alternatives $\{a_1,a_2,a_3,a_4\}$. This set has one subset of four alternatives, four subsets of three alternatives, and six subsets of two alternatives. For each subset, specify a voting method; that is, specify the number of points which are to be tallied for a voter's j^{TH} ranked alternative where j ranges over the number of alternatives. Let <u>W</u> denote the collection of these eleven balloting methods. Then, given $\underline{W}$ and the voters' profiles, the group's ordinal rankings for each of the eleven subsets is determined. As the voters' profiles vary over all possible choices, we obtain the set, R_W, of all possible ordinal rankings obtained from $\underline{W}$. So, an element of R_W is a listing of the eleven ordinal rankings resulting from some profile of voters. Let $\underline{W}=\underline{B}$ when all of the subsets are ranked by a Borda vector. So, R_B is a listing of all possible Borda rankings.

Clearly, R_W contains all transitive orderings; such an outcome results when all of the voters have an identical ranking of the four alternatives. So, if $|R_W|$ is the cardinality of R_W, then $|R_W|-4!$ is the number of possible results which are not transitive. (For N alternatives, $|R_W|-N!$ is the number of non-transitive outcomes.) In this way, $|R_W|$ is a measure of the inconsistencies admitted by a set of voting methods. We show that <u>the unique, minimum value for</u> $|R_W|$ <u>occurs only if the voting methods for all subsets of alternatives are Borda Counts</u>, i.e., $|R_W|\geq|R_B|$ for all choices of $\underline{W}$.

One difficulty with $|R_W|$ is that it doesn't indicate what are the inconsistent rankings. For instance, it doesn't eliminate the possibility that there are Borda rankings which, in some sense, "violate transitivity" more so than any ranking introduced by some other voting system. This can't happen because R_B is contained in R_W for all choices of $\underline{W}$. <u>Any nontransitive ranking admitted by a Borda Count also is admitted by any other set of voting methods.</u>

A consequence of the above is that the Borda Count admits only those inconsistencies which are unavoidable. Thus, we need to know what they are; we need to characterize R_B. It turns out to be inefficient to catalogue this set; so, we introduce some simple methods which permits one to easily answer questions about R_B. To illustrate the types of results which now are possible, we derive necessary and sufficient conditions for an alternative to be Borda ranked first, last, etc. In keeping with a dominant theme of social choice, these conditions are based upon how the voters ranked the pairs of alternatives.

2. Voting Methods

Let the $N \geq 3$ alternatives be $\{a_1, a_2, \ldots, a_N\}$. Assume that each voter has an ordinal, complete (all alternatives are included in the ranking), transitive ranking of the N alternatives. A listing of the rankings for the voters is called a "profile". A balloting or a voting method is where the group ranking is determined from voters' profiles in the following way: Given $\underline{W}^N = (w_1, w_2, \ldots, w_N)$, w_j points are tallied for a voter's j^{TH} place alternative. Then, the set of alternatives are ordered according to the sum of points each alternative receives. This final ordering can be determined either by asserting that the smaller the total, the higher the ranking (a reversed method), or the larger the total, the higher the ranking (a monotone method). In the latter case, the weights satisfy the conditions that $w_k \geq w_j$ if and only if $k < j$, and that $w_1 > w_N$. For a reversed method, these inequalities are reversed. For example, a plurality vote is a monotone method with $\underline{W}^N = (1, 0, \ldots, 0)$. For simplicity of exposition, assume that the w_k's are all rational numbers. (This doesn't impose any practical limitations. The only theoretical limitation occurs should $\{w_j\}$ be a set of completely irrational numbers; here, certain statements asserting the possibility of election results with indifference among alternatives may not hold. See [7,8] for an explanation of this.)

According to the above, voting methods differ by the choice of the voting vectors used in the tallying process. However, two methods may be equivalent because they always yield the same group ranking. For instance, it is clear that the outcome of an election is the same whether the voters' rankings are tallied with $\underline{W}^N$ or with $a\underline{W}^N$ where a is a nonzero constant. (If a is negative, then one system is a monotone method while the other one is a reversed method.) Likewise, the outcome remains invariant should the preferences be tallied by using $\underline{W}^N+b\underline{E}_N$. Here b is a nonzero scalar and the complete indifference vector E_N is $N^{-1}(1,..,1)$. Consequently,

Definition 1. Two voting vectors $\underline{W}^N{}_1$ and $\underline{W}^N{}_2$ are said to be equivalent if they and the vector $\underline{E}_N$ define a two dimensional linear subspace of R^N.

This defines an equivalence relation and the equivalence classes of voting vectors and methods. In what follows, we exploit this equivalence by normalizing the voting vectors. As a first normalization, we consider only monotone voting methods. The following characterizes an important equivalence class.

Definition 2. A Borda Count over $N\geq 3$ alternatives is where the vote tally vector, $\underline{W}^N$, has the property that w_k-w_{k+1} is the same nonzero constant for $k=1,2,..,N-1$. Denote both a Borda vector and the equivalence class of Borda vectors by $\underline{B}^N$.

Vector $\underline{V}_1=(1,2,..,N)$ is a reversed Borda vector, while $\underline{V}_2=(2N-2,2N-4,...,2,0)$ is a monotone Borda vector. They both belong to the same equivalence class because $2(N^2\underline{E}_N-\underline{V}_1)=\underline{V}_2$.

The $N\geq 3$ alternatives define a family of $2^N-(N+1)$ subsets, each of which has at least two alternatives. For each subset, select a voting vector. (For the subsets of two alternatives, we can assume that the voting methods are the same.) Let $\underline{W}$ denote these $2^N-(N+1)$ voting vectors. Should these vectors all be Borda, denote the combined vector by $\underline{B}$.

A given $\underline{W}$ and a choice of voters' profiles uniquely determines the ordinal rankings for the $2^N-(N+1)$ subsets. Let R_W be the set obtained by varying the voters' profiles over all possible choices. One of the main result of this paper, which establishes the superiority of the Borda Count, is given in the following theorem. (In this and several other statements, we assert that certain conclusions hold for "most" voting systems. "Most" means "almost all" in a measure theoretic sense, or "open-dense" in a topological sense. More precisely, it will mean all vectors $\underline{W}$ except those where the vector components satisfy a specified, strict, algebraic relationship among each other - see Section 5. It is of importance to note that if $\underline{W}$ consists solely of plurality voting methods, then it is in this general class.)

Theorem 1. **Let $N \geq 3$. Consider the family of all $2^N-(N+1)$ different subsets of at least two alternatives, and let $\underline{W}$ represent the collection of voting vectors adopted to rank the different subsets. Then**

2.1 $$R_W \supset R_B.$$

If $\underline{W} \neq \underline{B}$, then the first set strictly contains the second. Moreover, for most choices of $\underline{W}$, R_W contains all possible rankings.

This means that any voting method other than the Borda Count admits more nontransitive rankings than those obtained by the Borda Count. Furthermore, it follows from the strict containment that the Borda Count is the unique balloting resolution for this social choice problem. The following statement extends this conclusion to subfamilies of subsets of alternatives.

Corollary 1.1. **Out of the $2^N-(N+1)$ subsets of at least two alternatives, select a family of K of them. Let $\underline{W}$ be the collection of the voting vectors used to rank this family of subsets, and let R'_W be all possible election outcomes from this family. Then R'_W contains R'_B.**

We do not assert that the first set strictly contains the second one because there are families of subsets where $R'_W=R'_B$ independent of the choice of $\underline{W}$. Often this is characterized by R'_B containing all possible rankings of the subsets of alternatives.

A second important feature of Theorem 1 is that, for most choices of $\underline{W}$, R_W contains everything! Since any type of inconsistency can occur, these are the worse systems which can be used. As asserted, plurality voting is in this general class. The following corollary follows immediately.

Corollary 1.2. For each of the $2^N-(N+1)$ subsets, select, in an arbitrary fashion, some ranking of the alternatives. Then, there exist profiles of voters so that their ranking of each subset is the specified ranking.

Example: There exist profiles of voters so that their plurality rankings change with the number of alternatives; e.g., the group's plurality rankings are $a_1>a_2>a_3>a_4$, but $a_3>a_2>a_1$, $a_4>a_2>a_1$, $a_4>a_3>a_1$, and $a_4>a_3>a_2$, but $a_j>a_k$ iff $j<k$.

Suppose for a voting vector $\underline{W}^N$ there is only one choice of j so that $w_j-w_{j+1}\neq 0$. Such a voting system can distinguish between only two sets of alternatives; e.g., j=1 characterizes the plurality voting vector which distinguishes only between the top ranked alternative and all others. A commonly used method for committee selection is to indicate your "top k ranked alternatives". Again, this system distinguishes between only two subsets. It turns out that if all the voting components of $\underline{W}$ distinguish only between two subsets of alternatives, then $\underline{W}$ is in the general class where any outcome can occur. Thus, the above corollary holds for all of these systems.

Theorem 1 asserts that for most systems there need not be any relationship whatsoever between how the voters rank the various subsets of alternatives. But, for other systems, what type of relationships can be extracted? We start our answer of this question by examining the possible rankings of pairs of alternatives. The motivation is that if the group's outcome were transitive, then the ranking of N alternatives would uniquely establish the group's rankings of the pairs of alternatives. For instance, if the group's ranking is $a_1>a_2>\ldots>a_N$, then, to preserve transitivity, a majority of the voters would prefer a_k to a_j if and only if $k<j$.

It has been recognized for a long time that transitivity among the rankings of pairs need not exist; cycles can occur. One of the oldest examples, known as the

Condorcet triplet, is where the profiles of three voters are $a_1 > a_2 > a_3$, $a_2 > a_3 > a_1$, and $a_3 > a_1 > a_2$. A simple computation shows that, by votes of 2 to 1, $a_1 > a_2$, $a_2 > a_3$, but $a_3 > a_1$. The insidious effects of such cycles have been illustrated by the practical considerations of agenda manipulation, the effects of "seeding" on the conclusions of tournaments, etc. (While there is a very large literature on this subject, I suggest the references [4,5,9].) The following theorem asserts that there need not be any relationship whatsoever among the rankings of pairs of alternatives; any outcome is possible. (This result is a slight generalization of that given in [9].)

Theorem 2. Consider the $(N;2)=N(N-1)/2$ pairs of alternatives $\{a_k,a_j\}$. For each pair of alternatives, there are three different rankings: $a_k > a_j$, $a_k < a_j$, or $a_k = a_j$ ("a_k is indifferent to a_j"). This defines a set of $3^{(N;2)}$ sequences; each of the $(N;2)$ entries of a sequence designates the ranking for a specific pair of alternatives. Any such sequence can realized; there exist voters' profiles such that for each pair of alternatives the designated choice results by a majority vote for strict preference and a tie vote for indifference. That is, R'_N can be equated with the set of all possible sequences of rankings of pairs of alternatives.

For N=3, the Condorcet cycle is an example of this theorem. For N=4, this theorem means, for instance, that $a_1 > a_2$, $a_2 > a_3$, $a_3 > a_1$, $a_4 > a_1$, $a_2 > a_4$, all are realized by majority votes from the same set of voters profiles and these same voters are evenly split between a_3 and a_4. In general, all possible cycles, subcycles, or anything else can be constructed by means of majority vote.

This result imposes lower bounds on the consistency of voting independent of the choice of the voting method used to rank N alternatives. This can be seen with a profile of voters for the above example. Independent of how these voters rank the set of four alternatives, the outcome must be inconsistent with how the same voters rank at least two pairs of alternatives. (This is because the pairwise rankings of the two subsets $\{a_1,a_2,a_3\}$ and $\{a_1,a_2,a_4\}$ form cycles.) This illustrates Arrow's theorem for voting methods, and it imposes a lower bound on the degree of consistency which can be achieved through voting.

Because there need not be any relationship among the rankings of the pairs of alternatives, it might be argued that the search for consistency should be restricted to those profiles where there is order among the rankings of the pairs. The goal, then, would be to determine whether this relationship is reflected in the ranking of the N alternatives. For instance, cycles need not always occur; there are situations where, by majority votes, certain alternatives emerge as clear favorites, or as clear losers. Such alternatives, which were identified by Condorcet, often are used as the standard for comparison for the consistency of a voting method.

Definition 3. Alternative a_K is called a Condorcet winner if in all possible pairwise comparisons with the other alternatives, a_K always wins by a majority vote. Alternative a_K is called a Condorcet loser if in all possible pairwise comparisons with the other alternatives, a_K always loses by a majority vote.

For consistency, a voting method should rank a Condorcet winner in first place and a Condorcet loser in last place. But, this need not be the case; the introductory example demonstrates, and it follows in general from Corollary 1.2, that plurality voting can rank the Condorcet winner in last place and the Condorcet loser in first place. The next theorem asserts that, with the exception of the Borda Count, this and much worse phenomena can occur for <u>any</u> voting method. Only the Borda Count always reflects the rankings of the pairs of alternatives.

Theorem 3. Let $N \geq 3$ alternatives be given, and let $\underline{W}^N$ be a voting method to rank the N alternatives. Consider the relationship between the rankings of the N alternatives and the (N;2) pairs of alternatives. If $\underline{W}^N \neq \underline{B}^N$, then R'_w contains all possible combinations of the rankings of pairs of alternatives and the rankings of the N alternatives. The Borda vector, $\underline{B}^N$, never ranks a Condorcet winner in last place, nor a Condorcet loser in first place. There is no voting system which always ranks the Condorcet winner in first place and the Condorcet loser in last place.

Thus, with the exception of the Borda Count, there need not be any relationship whatsoever between the rankings of pairs of alternatives and the rankings of the N alternatives. In other words, even when we restrict attention to those profiles where there is order in the rankings of pairs, we don't find added consistency with the ranking of the N alternatives. Indeed, the following statement displays an extreme situation where the pairs do possess order, but it is at odds with the ranking of the N alternatives.

Corollary 3.1. Suppose that $N \geq 3$ and that the adopted voting vector, $\underline{W}^N$, is not a Borda vector. Then there exist profiles of voters so that, by majority votes, the pairs of alternatives are ranked $a_j > a_k$ if and only if $j < k$. Yet, their $\underline{W}^N$ ranking of the N alternatives is the reversal: $a_N > a_{N-1} > \ldots > a_1$.

3. A Vector Space Approach

Although the above statements demonstrate the superiority of the Borda Count over other voting methods, they do not adequately describe R_B nor R_W. To remedy this, we need a more complete description of voting systems. We start by relating cardinal rankings with ordinal rankings.

In the N dimensional space R^N, identify the K^{TH} component x_K with the K^{TH} alternative a_K. A vector $\underline{x}=(x_1,..,x_N)$ can be interpreted as a cardinal ranking of the N alternatives where larger values of x_K denote "stronger" preference for a_K. The hyperplane $x_K=x_J$ divides R^N into three regions; the two open regions denote strict ordinal preference (e.g., $x_K > x_J$ corresponds to where a_K is preferred to a_J), and the hyperplane corresponds to indifference between the two alternatives. By allowing K and j to vary over all possible $(N;2)=N(N-1)/2$ pairs of indices, the $(N;2)$ hyperplanes divide R^N into regions representing all possible ordinal rankings of the N alternatives. A connected open set is a "ranking regions" with strict preferences among alternatives; those regions contained in the hyperplanes are ranking regions with indifference among or between some of the alternatives. The line passing through $\underline{0}$ and $\underline{E}_N$ corresponds to complete indifference among the alternatives; this line is the intersection of the (N;2) "indifference" hyperplanes.

Let A denote the ranking $a_1 > a_2 > .. > a_N$. If $\underline{W}^N$ is a voting method, then, because it is a monotone method, it is in the closure of the ranking region of A. (If any two of the components of $\underline{W}^N$ are the same, this vector is on the boundary; otherwise, it is in the interior.) Vector $\underline{W}^N$ represents the tally for a voter with preference given by A. Denote this dependency by $\underline{W}^N{}_A$. Any other ranking of the N alternatives is a permutation of A, P(A). The tally for the ranking of such a voter is a permutation of $\underline{W}^N{}_A$; denote it by $\underline{W}^N{}_{P(A)}$. ($\underline{W}^N{}_{P(A)}$ is in the closure of the ranking region defined by P(A).) If there are $n'_{P(A)}$ voters with the ranking P(A), then the final tally is

3.1 $$\sum n'_{P(A)} W^N{}_{P(A)}$$

where the summation index, P(A), ranges over N! permutations of A. The group ranking is determined by the ranking region which contains this sum.

The ranking is invariant should the sum be divided by n, the total number of voters. If $n_{P(A)} = n'_{P(A)}/n$ is the fraction of the voters with ranking P(A), then the sum becomes

3.2 $$\sum n_{P(A)} \underline{W}^N{}_{P(A)}.$$

Because the variables $\{N_{P(A)}\}$ are non-negative and sum to unity, Eq. 3.2 can be interpreted as representing a convex combination of the vectors $\{\underline{W}^N{}_{P(A)}\}$. This set is in the affine plane containing these vectors. Our analysis is simplified when this plane is a linear subspace of R^N. This motivates the following.

Vector Normalization Assumption: The sum of the components of a voting vector equals zero.

Examples: a) The standard vector for plurality voting over N alternatives is (1,0,..,0). A normalized vector is (N-1,-1,-1,..,-1).

b) For N=2, we always use (1,-1).

This assumption forces the voting vectors and the sum in Eq. 3.2 to be in the linear subspace of R^N which is orthogonal to $\underline{E}_N$. Denote this N-1 dimensional subspace by E^N. For N=3, the ranking regions of E^3 are given in Figure 1.

For k=2,...,N-1, consider a subset of k alternatives, and let $\underline{W}^k$ be the voting method adopted to rank this subset of alternatives. For any voter's ranking of this subset, the tally of the ballot is given by the appropriate permutation of the components of $\underline{W}^k$. However, this permutation of $\underline{W}^k$ also can be indexed by how this voter ranks all N alternatives, not just this relevant subset. So, for any permutation of A, let $\underline{W}^k{}_{P(A)}$ be the unique permutation of $\underline{W}^k$ which corresponds to how <u>the specified</u> k <u>alternatives are ranked</u> in P(A). For instance, suppose N=4, k=3, and the specified ranking is $a_1 > a_2 > a_3$. There are four choices of P(A) which preserve this ranking -- they are the four ways in which a_4 can be positioned within this ranking of three alternatives. Thus, for exactly four different choices of the subscript P(A), the vectors $\underline{W}^3{}_{P(A)}$ agree and represent the vote tally for the same ranking of the three alternatives.

Let (N;k) represent the usual combinatoric symbol $\binom{N}{k}$. Each ranking of k alternatives is preserved in precisely (N;N-k)=(N;k) different permutations of P(A), so the vector $\underline{W}^k{}_{P(A)}$ is given by (N;k) different subscripts P(A). The group's ranking of these alternatives is given by the ranking region of E^k which contains the vector sum

3.3 $$\sum n_{P(A)}\underline{W}^k{}_{P(A)}.$$

To model how the same voters would rank the N alternatives with the method $\underline{W}^N$ and a subset of k alternatives with the method $\underline{W}^k$, we use the space $E^N \times E^k$. The ranking regions in this product space are given by the product of ranking regions in the component spaces. The outcome is the ranking region which contains the vector sum

3.4 $$\sum n_{P(A)}(\underline{W}^N{}_{P(A)},\underline{W}^K{}_{P(A)}).$$

If $\underline{W}=(\underline{W}^N,\underline{W}^K)$, this can be represented by the sum

3.5 $$\sum n_{P(A)}\underline{W}_{P(A)}$$

This equation has an interpretation similar to that of Eq. 3.2, and the sum is in the convex hull of the vectors $\{\underline{W}_{P(A)}\}$. To understand what nontransitive outcomes can result, we need to know which ranking regions meet this convex set.

A unique linear subspace is spanned by the convex set defined in Eq 3.5. What simplifies our analysis is that both the linear subspace and the convex hull meet the same ranking regions of $E^N x E^K$. (This will be shown in Section 5.) Therefore, the task of determining the elements of R'_W is equivalent to determining which ranking regions of $E^N x E^K$ meet the linear subspace spanned by the vectors $\{\underline{W}_{P(A)}\} = \{(\underline{W}^N{}_{P(A)},\underline{W}^K{}_{P(A)})\}$. Denote this subspace by V_W. Moreover, the dimension of the convex set and V_W are the same, so this common dimension serves as another measure of the number of nontransitive group rankings which can occur.

To illustrate this, Theorem 2 will be expressed in terms of the vector space V_W. For this, and future statements, we impose the following ordering on the listing of the (N;2) pairs of alternatives: A given pair (a_J,a_K) is listed with index $j<k$. The pairs are listed in the order $k=j+1,..,N, j=1,..,N-1$; i.e., (a_1,a_2), ... , (a_1,a_N); (a_2,a_3), ..., (a_{N-1},a_N). Each pair of alternatives (a_J,a_K), $j<k$, is represented by a space E^2 where, because of the ordering, the vector (1,-1) indicates that a_J is preferred to a_K. Thus, the space of all pairs is represented by $(E^2)^{(N;2)}$, and the above imposes an ordering on this space.

Theorem 4. Consider all (N;2) pairs of alternatives, and let $\underline{P}$ be the vector of voting methods. Then V_P is the total space $(E^2)^{(N;2)}$, and it has dimension (N;2).

Since V_P agrees with the space $(E^2)^{(N;2)}$, it meets each of the ranking regions. Thus, Theorem 2 follows. The importance of the dimension of V_P is that it imposes a lower bound on the dimension of V_W when results are compared over all $2^N-(N+1)$ subsets of alternatives. This is because when V_W is computed, it must reflect this freedom from consistency among the rankings of pairs. Thus, <u>a lower bound on the dimension of</u> V_W <u>for the general problem is</u> $(N;2)$.

When different sets of alternatives are ranked, the subspace V_W can vary depending on the scalar normalizations adopted for the voting components of $\underline{W}$. For instance, the space spanned by the permutations of the voting vectors $(3,1,-1,-3;1,0,-1)$ and $(3,1,-1,-3;5,0,-5)$ differ even though in both cases the set of 4 and the set of 3 alternatives are ranked by Borda Counts. (As we have shown, the ranking regions which meet these two subspaces are the same.) So, to compare vector spaces, we need to impose a scalar normalization. Because other voting vectors will be compared with the Borda Count, the only standards we impose are for $N=2$ and for the Borda Count; the normalization for the other vectors will be determined as needed.

Scalar Normalization Assumption: For $N \geq 3$ alternatives, the normalized Borda vector is $(N-1,..,N+1-2i,..,1-N)$. The voting vector used for $N=2$ is $(1,-1)$.

Example: For $N=4$, the Borda vector is $(3,1,-1,-3)$.

By taking a vector approach and by standarizing the Borda Count, sharper conclusions are possible. To illustrate this, an improvement of Theorem 3 follows. Here we are comparing the ranking of the N alternatives with the rankings of the pairs of alternatives, so the space is $E^N \times (E^2)^{(N;2)}$. The first voting component of $\underline{W}=(\underline{W}^N,\underline{P})$ ranks the N alternatives. The remaining voting components rank the pairs of alternatives where $\underline{P}$ is the vector in Theorem 4.

Define the vectors $\{\underline{Z}_{NK}\}$, $k=1,..,n$, in $E^N \times (E^2)^{(N;2)}$ in the following way. The E^N component of $\underline{Z}_{NK}$ has the value $-(N-1)/N$ in the K^{TH} component, and $1/N$ in all others. The E^2 component of $\underline{Z}_{NK}$ is zero if this component space E^2 does not represent a pair which includes a_K. If E^2 represents (a_K,a_J), the component is $(1,-1)$ if $K<j$, otherwise it is $(-1,1)$. This choice reflects that a_K is the preferred alternative.

These vectors can be interpreted in the following manner. The components in $(E^2)^{(N;2)}$ designate that a_K is a Condorcet winner. The E^N component designates that a_K is ranked in last place while all other alternatives are tied for first. For N=3, these vectors are $\underline{Z}_{3,1}=(-2/3,1/3,1/3;1,-1;1,-1;0,0)$, $\underline{Z}_{3,2}=(1/3,-2/3,1/3;-1,1;0,0;1,-1)$, and $\underline{Z}_{3,3}=(1/3,1/3,-2/3;-1,1;0,0;-1,1)$.

Theorem 5. Assume the hypothesis of Theorem 3. If $\underline{W}^N \neq \underline{B}^N$, then V_W is the total space $E^N x(E^2)^{(N;2)}$. The space V_B is a $(N;2)$ dimensional space characterized by the normal vectors $\{\underline{Z}_{NK}\}$..

It is remarkable that the dimension of V_B equals the theorectic lower bound of $(N;2)$! It is impossible to do better without eliminating the rankings of pairs, so this is another argument supporting the superiority of the Borda Count.

Because we can't do better than the Borda Count, we need to know these rankings which cannot be avoided. Any such ranking defines a set of vectors from a ranking region. This ranking is Borda admissible if and only if at least one vector from this set is orthogonal to <u>all</u> of the $\underline{Z}_{NK}$ vectors. The proof of the following statement illustrates this. The first assertion improves upon Theorem 3 because it relaxes the condition that an alternative must be a Condorcet winner to avoid being Borda ranked last. The second statement illustrates how Theorem 5 can be used to find sufficient conditions for an alternative to be Borda ranked first. Related results are easily derived.

Corollary 5.1. a) Let $f(k,j)$ be the difference between the fractions of the voters preferring a_k to a_j and those preferring a_j to a_k. If $F^N(k)=\sum_j f(k,j)$ is positive, a_k will not be Borda ranked in last place, nor tied for last place. If $F^N(k)$ is negative, a_k will not be Borda ranked in first place, nor tied for first place.

b) If $F^N(1)>N-2$, then a_1 is Borda ranked in first place. If $F^N(1)<2-N$, then a_1 is Borda ranked in last place.

If a_k is a Condorcet winner, then $f(k,j)>0$ for all choices of j. Trivially, $F^N(k)>0$, so a_k cannot be Borda ranked last. However, it is easy to construct examples where an alternative a_k has $F(k)>0$ even though it isn't a Condorcet winner. Thus, this result significantly improves upon Theorem 3. These inequalities are reversed if a_k is a Condorcet loser.

What we really want are necessary and sufficient conditions for an alternative to be Borda ranked in k^{TH} place, $k=1,..,N$, based upon how the voters rank the pairs of alternatives. This can't be done based solely upon the ordinal rankings of the pairs, but it can with the added information of how decisively each alternative won or lost in the pairwise comparisions. The following statement describes the close link between the Borda Count and the rankings of the pairs.

Corollary 5.2. Given a profile of voters, compute $F^N(k)$, $k=1,..,N$. The Borda tally is $(F^N(1),F^N(2),..,F^N(N))$. Thus, the algebraic ranking of $\{F^N(k)\}$ determines the Borda ranking of $\{a_k\}$.

The Borda ranking reflects how decisively an alternative fares in the pairwise comparisons with the other alternatives. From this, a case can be made that <u>the Borda winner is preferable to a Condorcet winner</u>. One supporting argument is that a Borda outcome is robust while a Condorcet winner need not be. For instance, it is easy to construct examples where a_1 is the Condorcet winner by virtue of barely winning majority votes over a_2 and a_3, yet a_2 wins decisively over a_3. Here, a_2 emerges as the Borda winner. Now, a slight change in the voters' rankings of a_1 and a_2 would change the Condorcet winner to a_2, but it wouldn't affect the Borda ranking. The reason, of course, is that a Condorcet winner is determined by ordinal rankings while a Borda outcome reflects the strength of a pairwise victory. A similar type of robustness argument characterizes the situations when a Condorcet winner isn't Borda ranked first.

Proof of the Corollaries. Proof of Corollary 5.2. For a given profile of voters, the outcome over the set of N alternatives and the (N;2) pairs of alternatives is given by the ranking region which contains the sum

3.6 $$\sum n_{P(A)}\underline{B}_{P(A)}$$

where $\underline{B}=(\underline{B}^N,\underline{P})$ is the normalized Borda vector in $E^N \times (E^2)^{(N;2)}$. Let the E^N component of this vector (the Borda outcome) be given by $(x_1,..,x_N)$. Because Eq. 3.6 is an admissible outcome, this vector sum is orthogonal to $\underline{Z}_{NK}$, $k=1,..,N$. Take this scalar product. The value of that part of the scalar product resulting from the E^N components is $(-x_K(N-1)/N)+(\sum_{j\neq K}(x_J/N)$. But, because of the vector normalization, $\sum x_S=0$, so $\sum_{j\neq K} x_J/N=-x_K/N$. Thus, the contribution to the scalar product from the E^N components is $-x_K$.

Because of the form of $\underline{Z}_{NK}$, the part of the scalar product corresponding to the space $(E^2)^{(N;2)}$ is $F^N(k)$. Thus, the orthogonality condition leads to the desired conclusion

3.7 $$x_K = F^N(k),$$

or

3.8 $$(x_1,\ldots,x_N) = (F^N(1),\ldots,F^N(N)).$$

This completes the proof of the Corollary 5.2.

Proof of Corollary 5.1. Part a. Because $f(j,k)=-f(k,j)$, $\sum F^N(k)=0$. So, either all of the $F^N(k)$'s are zero (the group's ranking is complete indifference), or there are some which are positive and some which are negative. In the latter case, if $F^N(k)\geq 0$, then it follows from Corollary 5.2 that a_K can't be Borda ranked last, nor tied for last. Similarly, if $F^N(k)\leq 0$, a_K can't be Borda ranked first, nor tied for first.

Part b. This proof involves nothing more than showing that $F^N(1)>N-2$ implies that $F^N(1)>F^N(K)$. To do this, we need to determine the maximum values for the x_K's. To illustrate the ideas, we restrict attention to $N=3$; the proof for $N\geq 3$ is similar.

Assume that $F^3(1)>1$. The vector outcome in Eq. 3.8 must be in the convex hull of the 6 permutations of the Borda vector $(2,0,-2)$. Trivially, the maximum values for the x_K's occur on the boundary of this set. Because $F^3(1)>0$, a_1 must be ranked either first or second, and either a_2 or a_3 occupies the other top two positions. Assume without loss of generality that a_1 and a_2 are the top two rated alternatives. This assumption determines an edge of the convex hull: $t(2,0,-2)+(1-t)(0,2,-2) = (2t,2-2t,-2)$, where $0\leq t\leq 1$. The assumption $F^3(1)>1$ forces $t>1/2$, which in turn forces the ranking to have a_1 in first place. This completes the proof. For $N>3$, other results follow by using the surfaces of the convex hull rather than just the edges.

We end this section with our main result.

Theorem 6. **Let $N\geq 3$ alternatives be given. Consider the family of all $2^N-(N+1)$ subsets of at least two alternatives. Let T be the space $E^N\times..\times(E^K)^{(N;K)}\times..\times(E^2)^{(N;2)}$. For each subset of alternatives, select a voting method, and let $\underline{W}$ in T be the vector consisting of all of the voting methods.**

a) Any $\underline{W}$ has a normalization so that V_B is a linear subspace of V_W. If $\underline{W}\neq\underline{B}$, then V_B is a proper subspace.

b) V_B is a $(N;2)$ dimensional linear subspace of T. The normal vectors for V_B are found in the following manner. For each subset of K alternatives, the vectors $\{\underline{Z}_{KJ}\}$ are defined. These vectors can be extended to T by allowing the new vector components to be the zero vectors. The set of all such vectors span the linear space which is normal to V_B.

c) For most choices of $\underline{W}$, $V_W=T$.

Theorem 1 and Corollary 1.1 are special cases of Statement a. Theorems 4 and 5 have imposed a lower bound on the dimension of V_B; the remarkable fact is that even though we are considering all possible subsets of alternatives, the dimension of V_B has not changed; it still equals (N;2). This again demonstrates the efficiency of the Borda vectors. Because the normal vectors to V_B, are specified, the elements in R_B can be computed in manner similar to that given above. This, then, constitutes a simple tool to determine possible Borda rankings.

If for some subset of K alternatives the appropriate component of $\underline{B}$ is replaced with another voting vector, then the dimension of the new vector space increases by (K-1). Essentially, the new vector space is V_B augmented by E^K in the appropriate space. This is one way in which the V_W spaces come about. A second way, which will be discussed in the section on proofs, is where there is a a linear combination between the voting methods at different levels which are of a very specific type.

In Section 2, we stated that plurality voting scheme is in this general class of "most" voting systems. Thus

Corollary 6.1. Assume the hypothesis of Theorem 6. If all of the voting components of $\underline{W}$ correspond to the plurality voting scheme, then V_W = T.

From this corollary, it follows that there exist voters' profiles leading to, say, the plurality rankings $a_1 \rangle a_2 \rangle a_3 \rangle a_4$, $a_4 \rangle a_3 \rangle a_2$, $a_1 \rangle a_2 \rangle a_4$, $a_2 \rangle a_3 = a_4$, $a_4 \rangle a_1 \rangle a_3$, while the pairs of alternatives are ranked as given in the example following Theorem 2. Of course, this same conclusion holds for any $\underline{W}$ where the component voting vectors distinguish only between two subsets of alternatives.

If N=3, all of the Borda rankings can be obtained by use of Theorem 5. If N=4, then the normal space to V_B is nine dimensional. This increased dimension means that there are a large number of inconsistent rankings which are not Borda admitted. Indeed, the numbers are so large, that a simple listing would not be reasonable. But, qualitative results of the nature given in Corollaries 5.1 and 5.2 are possible by using the same type of methods.

Corollary 6.2. For a given profile of voters, the Borda rankings of a subset of k alternatives C is given by the algebraic rankings of $F^k{}_C(j)=\sum f(j,i)$ where the summation is over $i\neq j$, $i,j\in C$.

Example. To illustrate this, we first show how a Condorcet winner over a specific subset of alternatives fares over other subsets. Suppose that N=4 and that $a_1 \rangle a_j$ for j=2,3 by majority votes. By the above results, a_1 can't be Borda ranked last in subset $\{a_1,a_2,a_3\}$. But, just from the knowledge that $a_1 \rangle a_3$, it follows that independent of how the group ranks a_1 and a_4, a_1 can't be Borda ranked in first place in $\{a_1,a_2,a_4\}$ while Borda ranked last in the total set. This is because the first condition implies that $F^3(1)=f(1,2)+f(1,4)\rangle 0$, while the second condition implies that $F^4(1)=F^3(1)+f(1,3)\langle 0$. Because $f(1,3)\rangle 0$, this is a contradiction.

We conclude this section with a comment concerning the probability that an inconsistent ranking occurs. The profile of voters are represented by the sets $\{n_{P(A)}\}$. Thus, because they sum to unity, they can be identified with the rational points in the positive orthant of a N!-1 dimensional space. Assume that the profiles of voters are distributed in such a manner that the rational points in any open set in this space has positive probability of occurring. Then, it turns out that for any choice of $\underline{W}$, <u>any admissible group rankings with strict preference between the alternatives has a positive probability of occurring</u>.

This can be proved by a simple vector analysis argument similar to that given above but with $\underline{W}$ instead of $\underline{B}$. An alternative, geometric approach is to note that the outcomes are given by convex combinations of the vectors $\{\underline{W}_{P(A)}\}$ where the (rational) coefficients indicate the number of voters with each ranking of the alternatives. As we will see, if this convex hull on V_W meets a ranking region with strict preferences, then this intersection forms an open subset of V_W. From this it follows immediately that 1) there are an infinite number of examples, indeed, the examples correspond to the rational points in an open set of N!-1 space, 2) the examples need only satisfy inequality constraints, and 3) as the number of voters, n, approaches infinity, then the probability that such a ranking occurs approachs a positive limit (which is determined by this open set in N!-1 space).

4. Some Extensions

The purpose of this section is to extend the above results in two different directions. The first is to admit additional voting methods over a fixed set of alternatives to determine how adverse of an effect this has on consistency. The second is to determine whether there are families of subsets which admit more consistency among the rankings than suggested above.

In Section 2, a standard equivalence relation for voting methods was given. The basic idea was that two methods are equivalent should they yield the same group ranking for any choice of voters' profiles. But, is this the best one can do; does this definition capture all of the relationships which preserve this invariance of group outcome? We show that if two voting methods are not equivalent according to this definition, then there exist profiles of voters where the outcomes differ. Indeed, much more can occur; should there be several voting methods which cannot be cannot be expressed in terms of each other, then the same profile of voters can lead to totally unrelated group rankings.

Definition 4 [8]. Let $\{\underline{W}^N{}_j\}$, j=1,..,k, be a set of k voting vectors used to rank N alternatives. They are said to be "completely different" if they and the vector $\underline{E}_N$ are linearly independent.

If the voting vectors are vector normalized, then we don't need to use $\underline{E}_N$. When k=2, the assertion that two voting methods are completely different means that they are not equivalent in the sense of Definition 1. The next theorem asserts that if there are K completely different voting vectors, there need not be any relationship among the same group's rankings of the same alternatives. Our assertion that Definition 1 captures all of the relationships leading to invariance of outcomes follows for K=2.

Theorem 7 [8]. Let $\{\underline{W}^N_j\}$, $j=1,2,..,K<N$ be a set of K completely different voting methods to rank a set of $N\geq 3$ alternatives. Let T′ be the space $(E^N)^K$ and let $\underline{W}$ have $\underline{W}^N_j$ as its j^{TH} vector component, $j=1,..,K$. Then $V_W=T'$. That is, select any K ordinal rankings of the N alternatives. Then there exist profiles of voters so that when the same voters rank the N alternatives by using the j^{TH} voting method, the outcome is the j^{TH} selected ranking, $j=1,..,K$.

Even the Borda vector doesn't provide any a[illegible]. This is beca[illegible] the Borda vector derives its power from its sensitivity to interaction effects over subsets of alternatives; here we are considering only one subset of alternatives. As an example, there exist profiles of voters so that their Borda ranking is $a_1>a_2>a_3>a_4$, their plurality ranking is $a_4>a_3>a_2>a_1$, and their ranking by designating the top two alternatives (weight vector (1,1,0,0) with a normalization of (1,1,-1,-1)) is $a_3>a_4>a_1>a_2$.

In the previous section, the value (N;2) arose both as the dimension of the subspace of pairs of alternatives and as the dimension of V_B. From this and Theorem 5, it may appear that the social choice problems are caused by the inconsistencies in the rankings of pairs of alternatives. Should this be so, then it would be natural to ignore the binaries; namely, in the interest of finding added consistency, perhaps the usual binary relevancy condition should be replaced with a K-fold relevancy condition. However, it turns out that there is no advantage in doing this. For instance, if attention is restricted only to the subsets of K alternatives, the minimal dimension value of (N;2) still is obtained by a Borda Count, and it is larger for any other voting method. However, there are families where the dimension of V_B is smaller than (N;2). They are characterized at the end of this section.

Theorem 8. Let $N>3$ alternatives be given, and let K be such that $2<K<N$. Consider the $(N;K)$ subsets of K alternatives, and let $\underline{W}^K{}_j$ be the voting method used to rank the j^{TH} set, $j=1,..,(N;K)$. Let $\underline{W}$, a vector in $T'=(E^K)^{(N;K)}$, have $\underline{W}^K{}_j$ as its j^{TH} vector component. Then,

a) V_B is a $(N;2)$ dimensional subspace of T'.

b) If $\underline{W}\neq\underline{B}$, then V_B is a proper subset of V_W.

By use of the methods developed in Section 5, it isn't difficult to show, for instance, that if $K=3$ and if at most one of the components of $\underline{W}$ is a Borda vector, then $V_W=T'$. A similar type of result extends for other choices of K. This means that if the descriptions of "agendas" or "tournaments" are extended from being based upon binary comparisons to being based upon K-fold comparisons, then the new setting inherits all of the well known "seeding" problems of agenda manipulation, etc. Moreover, if the K-fold rankings aren't done by the Borda Count, the damage to consistency is much worse than if the comparisons were made with the binaries. (The dimension of V_W is larger than $(N;2)$.)

Theorems 7 and 8 can be combined to show that the K alternatives in each of the $(N;K)$ subsets can be ranked in any number of ways. However because we are considering different subsets of alternatives, a hidden effect of the Borda Count manifests itself.

Definition 5. A set of voting vectors $\{\underline{W}^N{}_j\}$ are "Borda independent" if the subspace spanned by them and $\underline{E}_N$ doesn't include any Borda vectors.

Again, if voting vectors are vector normalized, we can exclude $\underline{E}_N$ from this definition. It follows immediately that for N alternatives, one can define a set of N-2 completely different, Borda independent voting methods.

Corollary 8.1. Let k be as defined in Theorem 8. For each subset of k alternatives, choose k-2 completely different, Borda independent voting methods. Then, for most choices of the voting methods, V_w is the total space $((E^k)^{k-1})^{(N;k)}$. That is, for each subset of alternatives, select k-2 rankings of the alternatives. Then there exist profiles of voters so that when the same voters rank the i^{TH} subset of alternatives with the j^{TH} voting method, the outcome is the j^{TH} selected ranking of the alternatives, i=1,..,(N;k), j=1,..,k-2.

For example, if N=5, k=4, then there exist profiles of voters so that their plurality rankings ($\underline{W}$=(1,0,0,0)) are $a_1 > a_2 > a_3 > a_4$, $a_5 > a_3 > a_1 > a_2$, and $a_4 > a_2 > a_5 > a_1$. However, when these same voters use the slightly perturbed voting system (1,1/100,0,0), the ranking for each subset is reversed.

This concept of Borda independent voting methods plays a role in the ranking of all possible subsets of the N alternatives.

Corollary 8.2. Consider the family of all $2^N-(N+1)$ subsets of N alternatives. For each subset of k alternatives, select k-2 completely different, Borda independent voting methods, and let $\underline{W}$ be the combined voting vector. Then, in general, V_w is the total space $(E^2)^{(N;2)}\times\ldots\times((E^k)^{k-1})^{(N;k)}\times\ldots\times(E^N)^{N-1}$.

We might expect an equivalent result to hold if for each subset of k alternatives, we select k-1 completely different vectors, none of which is a Borda vector. But a hidden effect of the Borda vector appears; this set of vectors can be re-expressed as k-2 Borda independent vectors and the Borda vector. As we now know, the Borda vector leads to a dimensional saving, and so the dimension of V_w is reduced accordingly. The same type of argument as given above explains why "Borda independent" is part of the hypothesis for the above Corollary.

Next, we present a situation where the Borda vector does not offer any savings even though we are considering several subsets of alternatives.

Theorem 9 [8]. Let $N \geq 3$, and let F be a family of N-1 nested subsets, S_j, $j=2,..,N$, where $|S_j|=j$ and S_j is a proper subset of S_{j+1}. For each subset, select a voting method, and let $\underline{W}$ be the collection of voting vectors. Then, for any choice of $\underline{W}$, $V_{\underline{W}}$ has dimension (N;2) and equals the total space $E^N x \ldots x E^2$. That is, for any choice of rankings from each of the sets, there exist choices of voters' profiles so that by using the adopted voting method to rank S_j, the outcome is the chosen ranking, $j=2,..,n$.

The Borda vector provides no advantage here; it admits all of the inconsistencies admitted by any other voting method. As such, Theorem 7 can be used to extend the result without worry whether the vectors are "Borda independent".

Corollary 9.1. For each of the subsets S_j in Theorem 8, select j-1 completely different voting methods. For each subset, select j-1 ranking of the alternatives. There exist profiles of voters so that when the same voters rank S_j with the k^{TH} voting method, the outcome is the k^{TH} selected ranking of the alternatives, $k=1,..,j-1$, $j=2,..,n$.

Example. Let $S_j=\{a_1,..,a_j\}$, $j=2,..,N$. Then, there exist profiles of voters so that the Borda ranking of S_j is $a_1 \rangle a_2 \rangle \ldots \rangle a_j$ if j is even, but the reverse of this if j is odd. Moreover, for $j \geq 3$, the plurality ranking of S_j is the same as the Borda ranking if j is a multiple of 3, it is the reverse of Borda ranking if j+1 is a multiple of 3, and it is $a_1 \rangle a_j \rangle a_2 \rangle a_{j-1} \rangle ..$ if j+2 is a multiple of 3.

Theorem 9 asserts that if specified subsets of alternatives are selected, then no method allows for consistency among the various rankings. (Thus, any such rankings are admitted in R_B.) This may suggest that methods based upon dropping alternatives, such as the Hare method, can lead to serious difficulties. This is true; now it is easy to show that inconsistencies result from such hierarchical methods. But, the relevant theorem to use is Theorem 6, not Theorem 9. The reason is that family of subsets defined in Theorem 9 specifies <u>in advance</u> which alternatives are to be dropped. Theorem 6 specifies what can occur over all possible subsets. (On the other hand, Theorem 9 can be used to show that the nested property and strict formulas which characterize such methods, such as always dropping the last place alternative, can lead to unexpected surprises.)

As stated above, the conclusion of Theorem 9 holds whether or not Borda methods are used. But note, the dimension of this space is (N;2), and the dimension of V_B in this space also is (N;2). Restricting attention to certain subsets of alternatives is equivalent to projecting V_B into the appropriate coordinate subspaces of T. Theorems 8 and 9 demonstrate that over these important, natural subspaces, this projection retains its dimension, so no savings in added consistency of rankings is achieved.

The basic question remains; is there some way we can reduce the types of inconsistencies which can occur by a clever choice of a collection of the subsets of the N alternatives? The mathematical idea for this is the following: The vector space V_B is a (N;2) dimensional subspace of T. Restricting attention a family of subsets, F, is equivalent to projecting V_B onto the subspace of T, T_F, which models this family. Now, if this family is selected properly, perhaps the projected dimension of V_B will be reduced, which in turn means that a lower number of inconsistencies occur. For example, consider the plane x=y in R^3; its projection onto the x-z plane or the y-z plane is two dimensional, but its projection onto the x-y plane is one dimensional.

Mathematically, this ranking problem is to determine whether there exist families of subsets, F, so that the projection of V_B into T_F has a lower dimension. The somewhat surprising result is that this can occur! Such families are characterized by the following statement which follows immediately from linear algebra.

Proposition. Let $N \geq 3$ alternatives be given. Choose a family of $K \leq 2^{(N,2)}-(N+1)$ subsets of alternatives, and let T_F be the subspace of T representing this family. If there are sets of L (but not L+1) linearly independent vectors in V_B (in T) which are normal to T_F, then the dimension of the projection of V_B into T_F is (N;2)-L.

The problem is to determine when a family of subsets satisfies the condition of the Proposition. The governing factor is a subtle symmetry group property which turns out to be related to the cycles and the Condorcet triplet described in Section 2. If the indices of the subsets in the family admits this symmetry property, then the projection of V_B is of lower dimension.

Definition 6. Let F be a family of subsets and let D be the set $\{d_{jk}\}$, $1 \le j < k \le N$, of scalar constants which are not all zero. Family F is said to satisfy the cycle symmetry property with set D if for each subset B in F

4.1 $$\sum d_{jk}(\underline{e}_j - \underline{e}_k) = \underline{0},$$

where the summation is over the indices of all pairs of alternatives in B.

Let $\underline{D}$ be the $(N;2)$ dimensional vector defined by set D.

Theorem 10. Consider a family, F, of subsets of the N alternatives, and let T_F be the subspace of T corresponding to this family. A necessary and sufficient condition that the projection of V_B into T_F has dimension $(N;2)$ is that F does not admit a cycle symmetry property.

Consider all possible sets D which define a cycle symmetry for F. If the corresponding set of vectors defines an L dimensional space, then the projection of V_B into T_F has dimension $(N;2)-L$.

This theorem highlights a major, mathematical cause of the difficulties in social choice and in voting methods. Each subset of alternatives defines certain algebraic symmetry groups - these are the different ways in which the alternatives can be permuted. But, in an attempt to make voting "fair" and "consistent", we add extra mathematical requirements. These extra conditions impose the constraint that the results must be considered (and hopefully, consistent) over several different subsets of alternatives. The paradoxes, impossibility theorems, and inconsistencies which are common to this area, then, are manifestations of the fact that the different symmetry groups associated with these subsets of alternatives aren't compatible with each other. Occasionally, there are family of subsets where the symmetry groups associated with the subsets do admit some compatible subgroups; in these situations additional consistency results. This cycle symmetry condition is a way to determine whether a particular family of subsets admits this symmetrical permutation condition, and to extract the symmetry subsets.

Examples: 1. It follows from this theorem that the families defined in Theorem 9 do not satisfy the cycle symmetry property. For example, let the family be $S_j=\{a_1,..a_j\}$, $j=3,..,N$, and let S_2 be a binary contained in S_3, then the projection of V_B is a $(N;2)$ dimensional subspace. Now, define a new and larger family F′ by replacing S_2 with all binaries except $\{a_1,a_2\}$, $\{a_1,a_3\}$, and $\{a_2,a_3\}$. Let set D be defined by $d_{12}=1$, $d_{13}=-1$, $d_{23}=1$, and all other d_{jk}'s equal zero. Then, F′ satisfies the cycle symmetry property with D. Thus, the dimension of the projection of V_B into this larger dimensional subspace of T is smaller. This example illustrates the importance of the condition in Theorem 9 that the sets S_j are nested.

Notice that this choice of D defines a cycle $a_1 \succ a_2 \succ a_3 \succ a_1$ similar to that defined by the Condorcet triplet. A related explanation holds for all cycle symmetry sets D.

2. The nested condition in Theorem 9 can be relaxed. For example, for N=6, the family $\{\{a_1,a_2\}, \{a_1,a_2,a_3\}, \{a_2,a_3,a_4,a_5\}, \{a_1,a_2,a_3,a_5,a_6\}, S_6\}$ does not satisfy the nested set property, nor does it admit a cycle symmetry property. Therefore, the projection of V_B into T_F equals T_F. A complete generalization of Theorem 9 is given by Theorem 10.

3. Let F be the family of all subsets which do not include the pair $\{a_1,a_2\}$. Then F satisfies the cycle symmetry property with $d_{12}=1$, $d_{jk}=0$ for all other pairs. This illustrates that not only cycles are involved in the choice of D sets.

4. Let N=6, and let the family be $\{\{a_3,a_5\}, \{a_4,a_5,a_6\}, S_4=\{a_1,a_2,a_3,a_4\}, \{a_1,a_2,a_3,a_4,a_6\}, S_6\}$. From S_4, five D sets can be defined which correspond to the five cycles $a_1 \succ a_2 \succ a_3 \succ a_1$,

$a_1 > a_2 > a_4 > a_1$, $a_1 > a_3 > a_4 > a_1$, $a_2 > a_3 > a_4 > a_2$ and $a_1 > a_2 > a_3 > a_4 > a_1$. It is elementary to show that the five vectors corresponding to these D sets are linearly dependent, but they have a basis of dimension 4. Thus, $L \geq 4$. That $L=4$ is a simple computation. This example illustrates that cycles greater than those of a three tuple are involved; the basis could be defined by the leaving out the cycle symmetry corresponding to the third cycle. It also illustrates what can occur if some member of a family, as S_4 above, is strategically chosen so that each of the other members of the family either contains this set, or does not contain a pair from this set. Then, all the symmetries of this set define cycle symmetries for F. Notice that one D cyclic symmetry class for F is where $d_{12}=-d_{13}=1$, $d_{23}=d_{24}=-d_{34}=1/3$. This demonstrates that there are symmetry sets where not all of the magnitudes of the d_{JK}'s are either the same constant or zero.

5. Let N=4, and let F be $[S_4, \{a_1,a_2\}]$. Two D sets with independent $\underline{D}$ vectors are given by the cycles $a_2 > a_3 > a_4 > a_2$ and $a_1 > a_3 > a_4 > a_1$. So, the projection of V_B has dimension $(4;2)-2=4$, which is the dimension of T_F. Thus, all possible rankings are possible with any voting vector.

Of course, the obvious extensions of Theorem 10 by using systems of either completely different, Borda independent voting vectors (the dimension of the projection of V_B is less than that of T_F), or completely different voting vectors also hold.

5. Proofs of the theorems

All of the questions considered in this paper concern a vector sum of the form

5.1 $$\sum n_{P(A)} \underline{W}_{P(A)}.$$

Assume that the vector components of $\underline{W}$ are in vector normalized form. Then, $\{\underline{W}_{P(A)}\}$ is in T' where T' is the cartesian products of spaces E^K which represents this family of subsets.

As it has been noted earlier, $n_{P(A)}$ is non-negative for all P(A) and

5.2 $\sum n_{P(A)} = 1.$

This means that these coefficients define a simplex in the positive orthant of a N! Euclidean space. Consequently, if

$$Si(k) = \{ \underline{x} \in R^K \mid x_j \geq 0, \sum x_j = 1 \},$$

then Eq. 5.1 can be interpreted as being a restriction of the mapping

5.3. $F: Si(N!) \text{-------}> T'.$

More precisely, define F as the mapping in Eq. 5.1 where the coefficients $\{n_{P(A)}\}$ are elements in Si(N!). For this mapping to correspond to a vote tally, the components of $\{n_{P(A)}\}$ must be rational numbers. (A common denominator is the number of voters while the numerator of each component represents the number of voters with a particular ranking of the alternatives.) Thus, in order to interpret the image of F as a vote tally, the domain point must be a rational point in Si(N!).

Let V_W be the vector space in T′ spanned by the vectors $\{\underline{W}_{P(A)}\}$. It follows that

5.4. $F: Si(N!) \text{-------}> V_W \quad T'.$

What we show next is that if V_W meets a ranking region of T′, then this is the group ranking for some set of voters′ profiles. As the first step, note that by interpreting Eq. 5.1 as a convex combination of the vectors, it follows that the image set of F is the convex hull of the vectors $\{\underline{W}_{P(A)}\}$ in V_W.

The "indifference" point $\underline{E}_{N*}$ in Si(N!) corresponds to where the preferences of voters are equally split among the N! permutations of the alternatives. From this it follows immediately that

5.5 $F(\underline{E}_{N*}) = \underline{0}.$

Namely, with this even division among the voters, the group ranking of any subset of alternatives is complete indifference. For the normalized vectors, this is $\underline{0}$.

Next, <u>assume that the Jacobian of F has rank equal to the dimension of the</u>

<u>linear subspace</u> V_W. (This will be verified later.) A consequence of this assumption is that F maps an open neighborhood, U, of $\underline{E}_{N*}$ in Si(N!) to an open neighborhood of $\underline{0}$ in V_W. But $\underline{0}$ is a boundary point of all of the ranking regions of T′. Therefore, F(U) and V_W meet the same ranking regions of T′. It remains to show that for each such ranking region, there is a rational point in U which is mapped to this region.

Consider the ranking regions where all rankings have strict preference between alternatives. These ranking regions are open sets in T′. Consequently, they are open sets in V_W. But, because F is linear and has maximal rank, it is an open mapping. Thus, the intersection of F(U) and such a ranking region is an open set in V_W. The continuity of F ensures that the inverse image of this new open set is an open set, U′, in Si(N!). The rational points are dense in Si(N!), so there exist rational points which are mapped by F to the appropriate ranking region.

Consider those ranking regions which admit indifference among or between some alternatives in some subset of alternatives. Then, in some component(s) of T′, the ranking region is a part of a hyperplane; the boundaries are defined by other hyperplanes passing through $\underline{0}$. (These are the hyperplanes corresponding to those other alternatives which are ranked with strict preference.) The intersection of this ranking region with V_W is a part of a linear subspace of V_W. Because this set has $\underline{0}$ as a boundary point, the intersection of F(U) with this space is nonempty, so the inverse image is a portion of a hyperplane in Si(N!). This hyperplane in Si(N!) is characterized by the voting vectors which involve rational components. Thus, this hyperplane contains rational points.

We have established that if the assumption on the Jacobian of F is true, then the set of rankings resulting from a set of voting methods, $\underline{W}$, corresponds to the set of ranking regions which meet V_W. Most of the remainder of this section is devoted to verifying the assumption.

Proof of Theorem 7. The basic ideas are illustrated with the proof of Theorem 7. That is, for N alternatives and N-1 completely different voting methods $\underline{W} = (\underline{W}^N_1,..,\underline{W}^N_{N-1})$, we will show

1) that the Jacobian of F has the rank equal to the dimension of V_W, and
2) that V_W equals $T'=(E^N)^{N-1}$.

Let $M_{JK}(\underline{x})$, $1\leq j<k\leq n$, be the permutation mapping from E^N back to itself which interchanges the j^{TH} and the k^{TH} components of $\underline{x}$. If $\underline{X}=(\underline{x}_1,...,\underline{x}_{N-1})$ is in $(E^N)^{N-1}$, then let

5.6) $M_{KJ}(\underline{X}) = (M_{KJ}(\underline{x}_1),...,M_{KJ}(\underline{x}_{N-1}))$.

Let G be the group of permutations generated by the (N;2) mappings $\{M_{JK}\}$, and define

5.7) $L(G) = \{V|V$ is a linear subspace of T' which is invariant under $G\}$.

That is, if M is a permutation from G, then M maps V back into itself. Because such a mapping M just permutes the components of the vectors, it follows that M maps V_W back into itself. Consequently, V_W is in L(G). To complete the theorem, we characterize the subspaces in L(G). From this it will follow that $V_W=T'$.

To characterize L(G), we determine the eigenvalues and eigenvectors of M_{JK}. A simple computation yields that the eigenvalues are -1 with multiplicity N-1 and +1 with multiplicity (N-2)(N-1). A set of (N-1) eigenvectors corresponding to the eigenvalue -1 are $(\underline{e}_J-\underline{e}_K,\underline{0},..,\underline{0})$, ..,$(\underline{0},...,\underline{e}_J-\underline{e}_K)$ where $\underline{e}_s$, s=1,..,N, is the unit vector in R^N with unity in the s^{TH} component and zero in all others. Call the subspace spanned by these N-1 vectors, the -1 eigenspace for (j,k). Notice that the -1 and the +1 eigenspaces for (j,k) are orthogonal to each other.

Claim 1: Let V be in L(G). The projection of V into the -1 eigenspace for (j,k) is a linear subspace of V.

Proof of the claim. Clearly, the projection is a linear subspace of the -1 eigenspace for (j,k). We must show that this linear subspace also is a subspace of

V.

To prove this, it suffices to show that if $\underline{v}_1$ is the projecton of $\underline{v}$ in V, then $\underline{v}_1$ is in V. By using the orthogonality of the two eigenspaces, it follows that $\underline{v}$ has a unique representation $\underline{v}_1+\underline{v}_2$ where $\underline{v}_2$ is in the +1 eigenspace for (j,k). By the invariance assumption, the vector $M_{jk}(\underline{v}) = M_{jk}(\underline{v}_1+\underline{v}_2) = -\underline{v}_1+\underline{v}_2$ is in V_w. Therefore

5.8 $\quad \underline{v} - M_{jk}(\underline{v}) = 2\underline{v}_1$

is in V_w. This completes the proof.

A consequence of this theorem is that V can be expressed as the direct sum of a vector space from the -1 eigenspace and the +1 eigenspace for (j,k). The next statement shows that as the j and k vary, the subspaces obtained by the projection of V are related. We do this by showing how they are all related to some one space.

<u>Claim</u> <u>2</u>: For V in L(G) let V^{jk} be the subspace obtained by projecting V into the -1 eigenspace of (j,k). Then $M_{2j}(M_{1k}(V^{jk}))=V^{12}$. Both subspaces have the same dimension.

Proof of the claim. Notice that

5.9) $\quad M_{sk}(\underline{e}_j-\underline{e}_k) = \underline{e}_j-\underline{e}_s$,

and that this is a -1 eigenvector for M_{js}. From this it follows that $M_{sk}(V^{jk})$ is a linear subspace of the -1 eigenspace for (s,j) which has the same dimension as V^{jk}. Because V is in L(G), $M_{sk}(V^{jk})$ is a linear subspace of V; hence it is in V^{js}.

A similar argument shows that $M_{sk}(V^{js})$ is a subset of V^{jk}. (M_{sk} is an involution.) Because M_{sk} preserves dimension, it follows that $M_{sk}(V^{js}) = V^{sk}$ and that the dimensions of both linear subsets agree. This completes the proof.

The above two claims will be used to characterize any V in L(G).

<u>Claim</u> <u>3</u>: Let V in L(G) be such that V^{12} is j dimensional. Then V is spanned

by the sets V^{1K}, $K=2,..,N$, and the dimension of V is $(N-1)j$.

Proof of the claim: A basis for E^N is $\{\underline{e}_1-\underline{e}_K\}$ where K ranges from 2 to N. Therefore, a basis for $T'=(E^N)^{N-1}$ is the standard one of extending the basis for each component space to the product space.

Assume that V^{12} is j dimensional. According to Claim 2, V^{1K}, $K=2,..,N$ is a j dimensional subspace of V. Thus V contains the span of these vector spaces. Moreover, it follows from our choice of a basis for V that the basis for the subspaces form a linearly independent set of vectors. Thus, the dimension of V is bounded below by $(N-1)j$.

If the dimension of V is greater than $(N-1)j$, then there is a vector $\underline{v}$ in V which cannot be expressed as the linear combination of vectors from the spaces V^{1K}, $K=2,..,N$. But, by our choice of a basis, this means that for some choice of K, the projection of $\underline{x}$ into the -1 eigenspace of $(1,K)$ is not in V^{1K}. This contradicts the definition of V^{1K}.

Completion of the proof of Theorem 7. Because V_W is in $L(G)$, its dimension is $(N-1)j$ where j is the dimension of V^{12}_W, the projection of V_W into the -1 eigenspace of $(1,2)$. We show that if the voting vectors are completely different, then $j=N-1$.

Assume that V^{12}_W is j dimensional. This means that a basis for V^{12}_W is given by $\{\underline{c}_s\}$, $s=1,..,j$. Here, each $\underline{c}_s$ is a linear combination of the N-1 vectors $(\underline{e}_1-\underline{e}_2,\underline{0},..,\underline{0})$, ..., $(\underline{0},..,\underline{0},\underline{e}_1-\underline{e}_2)$. By a standard row reduction argument (and perhaps by a reassignment of the order of the voting vectors in $\underline{W}$) we can assume that the basis $\{\underline{c}_s\}$ is replaced with an equivalent basis $\{\underline{d}_s\}$ where

5.10) $\underline{d}^{12}_s = (a^s_1(\underline{e}_1-\underline{e}_2);a^s_2(\underline{e}_1-\underline{e}_2);\ldots;a^s_N(\underline{e}_1-\underline{e}_2)$,

$$\text{where } a^s_K \begin{cases} = 1 \text{ for } s=K,\ s=1,..,j; \\ = 0 \text{ for } K\le j,\ K\neq s; \end{cases}$$

It follows from Eq. 5.9 and Claim 2 that a basis for $M_{2K}(V^{12}_W) = V^{1K}_W$ is

$\{\underline{d}^{1K}{}_s\}$ where $\underline{d}^{1K}{}_s$ is $(a^s{}_1(\underline{e}_1-\underline{e}_K,..,a^s{}_N(\underline{e}_1-\underline{e}_K))$. Thus, a basis for V_W is given by

5.11) $\{\underline{d}^{1K}{}_s\}$, $K=2,..,N$; $s=1,..,j$.

The vector $\underline{W}=(\underline{W}^N{}_1,..,\underline{W}^N{}_{N+1})$ is in the space V_W, so it can be expressed as a linear combination of the vectors from Eq. 5.11. Each voting component $\underline{W}^N{}_s$ is in a space E^N which is spanned by the vectors $(\underline{e}_1-\underline{e}_K)$. Therefore, it has a unique linear representation in terms of this basis. But because of the row reduced form of the vectors in Eq. 5.11, this means that the components $\underline{W}^N{}_K$, $K\le j$, uniquely determine the representation of $\underline{W}$ in terms of the basis 5.11. In particular, the linear combination used to determine $\underline{W}^N{}_K$ in E^N uniquely determines the linear combination of the $\underline{d}^{1K}{}_s$ vectors required to represent the K^{TH} component of $\underline{W}$, $K=2,..,N$.

Assume that $j<N-1$, and let s be such that $j<s\le N-1$. Because of the row reduced form of the basis vectors, it follows that the linear combination required to represent the K^{TH} vector component of $\underline{W}$ yields $a^K{}_s\underline{W}^{NK}$ in the s^{TH} component, $s= j+1,..,N-1$, $K=1,..,j$. For this to hold, the voting vector $\underline{W}^N{}_s$ must be a linear combination of the vectors $(\underline{W}^N{}_K)$, $K=1,..,j$. This contradiction to the assumption that the voting vectors are completely different force $j=N-1$ and completes the proof.

The key idea in the above proof is to use the mappings M_{KJ} to determine a basis for V_W. Once this is done, then the rest of the proof is simple vector analysis. This basic theme persists in what follows, but there are some significant differences. The most serious one is that the permutation map, M_{JK}, doesn't admit an extension of the type given in Eq 5.7 when more than one subset of alternatives are considered. To see this, let N=3 and consider M_{12} on the space $T'=E^3x(E^2)^3$. The vector $(2,0,-2;1,-1;1,-1;1,-1)$ is the Borda vector $\underline{B}_A$ where A is the ranking $a_1>a_2>a_3$. Now, if $M_{12}(\underline{B}_A)$ were defined as in Eq.

5.7, then the outcome would be (0,2,-2;-1,1;-1,1;-1,1). But this corresponds to the inconsistent ranking $a_2 > a_1 > a_3$; $a_2 > a_1$, $a_3 > a_1$, and $a_3 > a_2$.

It is clear what properties we want the extension of a mapping, such as M_{12}, to have. We wish 1) to interchange the a_1 and the a_2 alternatives in the ranking, and 2) to map elements from $\{\underline{W}_{P(A)}\}$ back into this set. It can be seen from the above and other examples that to preserve consistency, the permutation mapping on subsets of alternatives is determined <u>both</u> by the choice of permutation mapping on the total set of alternatives and by the choice of the ranking of the alternatives. This leads to the following definition.

Definition 7. Let C be a ranking of the N alternatives, and let $\underline{W}^C$ in T be the voting vector corresponding to C. Define $M_{KJ}(\underline{W}_C) = \underline{W}_{P(C)}$ where P(C) is the ranking obtained from C by interchanging the positions of a_K and a_J. Let

5.13 $M_{KJ}(\sum d_C \underline{W}_C) = \sum d_C M_{KJ}(\underline{W}_C)$,

where the summation index, C, is over all rankings of the N alternatives and where d_C is a scalar.

Example: Let N=3. When the mapping M_{12} is defined over the space T and when the ranking it is operating on is A, then $M_{12}=(M_{12},M_{12},E,E)$ where E is the identity mapping. On the other hand, if the ranking is $a_1 > a_3 > a_2$, then $M_{12}=(M_{12},M_{12},M_{12},M_{12})$. When there is only one set of alternatives, as in Theorem 7, this definition reduces to Eq. 5.7.

This type of structure, where a collection of mappings is defined over a product space and each component mapping depends upon what was the mapping and the base point in some previous component space, is called a <u>wreath product</u>. (See, for example, [3,4].) Thus, the permutation mappings over the space of rankings of subsets of alternatives defines a wreath product of permutation mappings. (Incidently, as it will be shown elsewhere, many of the chronicalled difficulties of social choice can be explained and extended by use of this wreath product.)

(As a brief aside, I would like to point out that this wreath product captures the symmetry properties of the simplex. This can be seen by using a normalization where the components, w_j, in the voting vectors are all non-negative and sum to unity. This changes the underlying space from E^N to the simplex $Si(N)$ where subsets of alternatives correspond to faces and edges of this simplex. Then it is easy to see that the above wreath product coordinates symmetry actions on the faces and edges of the simplex with symmetry action inside the simplex.)

By use of Eq. 5.13, it is easy to establish that V_w is in $L(G)=\{V|V$ is a linear subspace which is invariant with respect to G, the group generated by the permutation mappings $M_{IJ}\}$. On the other hand, the extension of M_{KJ} in Definition 7 does not define a linear mapping. (There are choices of vectors $\underline{v}$ and $\underline{w}$ so that

5.14 $M_{JK}(\underline{v}-\underline{w})\neq M_{JK}(\underline{v})-M_{JK}(\underline{w})$.)

Thus, because the proof of Theorem 7 depends upon the linearity of M_{JK}, it does not extend directly to the general setting. The purpose of most of what follows is to overcome the effects of Eq. 5.14.

Let T be as defined earlier; i.e., $T=E^N \times (E^{N-1})^{(N;N-1)} \times .. \times (E^2)^{(N;2)}$.

Claim 4. Consider the set of vectors $\{\underline{V}^C_{JK}\}$ in T which are constructed in the following way. For each choice of K and j, K>j, consider a component subspace, C, of T which represents a subset of alternatives which includes a_J and a_K. Then, the vector component of $\underline{V}^C_{JK}$ corresponding to the subspace C is $\{e_J-e_K\}$; all other vector components are $\underline{0}$. The set of these vectors, as (K,j) vary over all possible choices of indices and as C ranges over all component subspaces of T, span T.

The validity of this claim is obvious. A counting argument shows that this set of vectors is not linearly independent.

If $\underline{V}$ is in V_w, then so is the vector $\underline{V}-M_{JK}(\underline{V})$. (This is because V_w is in L(G).) Vector differences of this type replace the role of the "-1 eigenspace of

(j,k)" when we determine a basis for V_W. To illustrate the ideas in a simplier setting, we first prove Theorem 9.

Proof of Theorem 9. Without loss of generality, assume that $S_j=\{a_1,..,a_j\}$, $j=2,..,N$. Let $\underline{W}=(\underline{W}^N,\underline{W}^{N-1},..,(1,-1))$ be the collection of normalized voting vectors selected to rank the alternatives. Because $\underline{W}^N$ is a voting vector, there is some adjacent pair of weights which are not equal. Let j be the first index where $w_j>w_{j+1}$.

Consider

5.15 $\qquad \underline{W}_{P(A)} - M_{kN}(\underline{W}_{P(A)})$,

$k=1,..,N-1$, where $\underline{W}_{P(A)}$ is defined in terms of k and j in the following way. For given value of k, P(A) is a ranking of the N alternatives where a_k is in j^{TH} place while a_N is the $(j+1)^{TH}$ place. The remainder of the alternatives can be ranked in any way. Now, because a_k and a_N are adjacent in the ranking and because a_N isn't in any of the sets S_j for $j<N$, it follows that if these two alternatives are transposed, it won't affect the rankings in any other subset S_j, $j<N$. Consequently, the vector in 5.15 is the positive scalar multiple (w_j-w_{j-1}) of

5.16 $\qquad (\underline{e}_k-\underline{e}_N;\underline{0},..,\underline{0})$,

$k=1,..,N-1$. By construction, these vectors are in V_W. Moreover, they form a basis for the subspace $E^N\times\underline{0}\times..\times\underline{0}$. Therefore, this subspace is contained in V_W.

Next, the mappings M_{kN-1} are applied in the same way where j is selected to be the first index where two adjacent components of $\underline{W}^{N-1}$ are not the same. Then, Eq 5.15 yields vectors of the form

5.17 $\quad (b_j(\underline{e}_k-\underline{e}_{N-1}),(w_j-w_{j-1})(\underline{e}_k-\underline{e}_{N-1}),\underline{0},..,\underline{0})$.

But, because Eq. 5.16 forms a basis for $E^N\times..\times\underline{0}$, for each choice of k, there is a combination of these vectors which equals $(-b_j(\underline{e}_k-\underline{e}_{N-1}),\underline{0},..,\underline{0})$.Thus, it follows that the vectors $(\underline{0},\underline{e}_k-\underline{e}_{N-1},\underline{0},..,\underline{0})$, $k=1,..,N-2$, are in V_W. This forms a basis for $\underline{0}\times E^{N-1}\times\underline{0}\times..\times\underline{0}$.

Continuing with the obvious induction argument, it follows that V_W contains $E^N x E^{N-1} x..x E^2$. This completes the proof.

The key point of this proof is that the symmetry properties of permutations of S_J differ significantly from those of S_K, $j \neq k$. This difference is significant enough so that an element in L(G) cannot have a nontrivial, proper subspace of $E^J x E^K$. Because this symmetry incompatibility is based on permutations of the alternatives, it persists when j-1 completely different voting methods are used to rank E^J. Thus, the proofs of Theorem 7 and the above combine to prove this special case of Corollary 9.1.

Proof of Theorem 4. On the space $T' = (E^2)^{(N;2)}$ consider

5.18 $\underline{W}_{P(A)} - M_{JK}(\underline{W}_{P(A)})$

where j and k range through all possible pairs in the order k=j+1,..,N, j=1,..,N. For each pair of indices (j,k), P(A) is a ranking where a_J is the top ranked alternative and a_K is the second ranked alternative. Because these alternatives are adjacent in P(A), and because this pair doesn't appear in any other set of alternatives, their transposition in P(A) doesn't effect the ranking of any other pair. Therefore, the vector difference in Eq. 5.18 has (2,-2) in the one component corresponding to the pair (a_J, a_K), and $\underline{0}$ in all others. These (N;2) vectors form a (N;2)x(N;2) matrix with two dimensional vectors as the entries. Because of the ordering of the indices, the entries along the diagonal of this matrix are (2,-2) and $\underline{0}$ off the main diagonal. That these (N;2) vectors are linearly independent is immediate. Thus V_P has vector dimension (N;2) and it contains (and hence equals) T'.

To simplify the notation, if $\{a_J, a_K\}$ is contained in a subset defining a component space of T, call it "an (j,k) component space".

Proof of Theorem 6. First we show that V_B has dimension (N;2). To do this, for each j and k, consider Eq. 5.18 where $\underline{B}$ replaces $\underline{W}$ and where P(A) ranges through all rankings where a_j is the i^{TH} ranked alternative, a_k is the $(i+1)^{TH}$ ranked alternative, i=1,..,N-1. Now, in a (j,k) component space of T, the vector difference from Eq. 5.18 is $(w_i-w_{i+1})(\underline{e}_j-\underline{e}_k)$. But, for a Borda Vector, $(w_i-w_{i+1})=2$ for all choices of i. Since the voting vector for N=2 is (1,-1), this difference is 2(1,-1). Therefore, Eq. 5.18 is independent of i, and it defines a vector $2\underline{V}_{jk}$ where $\underline{V}_{jk}$ has $(\underline{e}_j-\underline{e}_k)$ in any (j,k) component space, and $\underline{0}$ in all others.

The set $\{\underline{V}_{jk}\}$, j<k, has (N;2) vectors which are linearly independent. (The linear independence follows the proof of Theorem 4; the vectors used in this proof are the last (N;2) vector components of the vectors $\underline{V}_{jk}$.) To prove the theorem, it suffices to show that any vector $\underline{B}_{P(A)}$ can be expressed as a linear combination of the $\underline{V}_{jk}$ vectors.

First we show that

5.19 $$\underline{B}_A = \sum \underline{V}_{jk},$$

where the summation is over $1 \leq j < k \leq N$. Consider a subset, D, of $s \leq N$ alternatives. Let j be such that a_j is in D. Then, for exactly s-1 choices of $k \neq j$, $\underline{V}_{jk}$ has a non-zero vector component in the space corresponding to the set D. Each of these vectors has a non-zero component in the direction corresponding to a_j; it is +1 if j<k, it is -1 if k<j. Therefore, in the component corresponding to the subset D, the sum in Eq. 5.19 is (s-1, s-3,..,s+1-2i,..,1-s). Thus, the sum is the Borda vector.

Next, we show that $\underline{B}_{P(A)}$ can be expressed as a linear combination of the $\underline{V}_{jk}$ vectors for any choice of P(A). But, any P(A) can be expressed as the compositions of transpositions of the ranking A. So, it suffices to show that if $\underline{B}_{P(A)}$ can be expressed by such a combination, then so can $\underline{B}_C$ where C is a ranking obtained by a transposition of some two alternatives which are adjacent in the ranking $P(A)$. According to Eq. 5.18, this is given by $\underline{B}_{P(A)}+2\underline{V}_{JK}$ for the choice of j and k determined by the transposition.

The above demonstrates that V_B is spanned by $\{\underline{V}_{JK}\}$, $j<k$. That the $\underline{Z}$ vectors are orthogonal to the space V_B follows from the simple computation that each such vector is orthogonal to the basis. That the $\underline{Z}$ vectors determine a basis for the normal space of V_B follows from a counting argument. (In this counting argument, note that the $\underline{Z}$'s determined for the subset of alternatives D are linearly dependent. There are s such vectors, but any s-1 of them are linearly independent. This is a consequence of our normalization.) This completes the proof of part b.

Proof of part a. Let $\underline{W}$ be the collection of voting vectors. If each voting method distinguishes between the top and the second ranked alternative, then let the normalization of the voting vectors be $w_1-w_2=2$. For each choice of $j<k$, let $P(A)$ be the ranking where a_J is the top ranked alternative and a_K is the second ranked alternative. Then, Eq. 5.18 yields $2\underline{V}_{JK}$, $j<k$. Because these vectors form a basis for V_B, it follows that V_B is a subspace of V_W.

Assume that there are some voting methods reflected in $\underline{W}$ which do not distinguish between the top two ranked alternatives. Because these components are voting vectors, they must distinguish between some two rankings. Fix j and k, and consider all possible rankings of the alternatives where a_J is the i^{TH} ranked alternative while a_K is the $(i+1)^{TH}$ ranked alternative for each $i=1,..,N-1$. For each such ranking, in each (j,k) component space Eq. 5.18 has a non-negative multiple of $(\underline{e}_J-\underline{e}_K)$. When $i=1$, this multiple is w_1-w_2; and it occurs $(N-2)!$ times. (This is the number of rankings of A satisfying this condition.) When $i>1$, it is a combinatoric problem to determine the number of different rankings of the N alternatives where $i-(s+1)$ elements of D and s elements not in D can be ranked ahead of a_J and a_K. This will determine how often in Eq. 5.18 the scalar multiple $w_{I-s}-w_{I-1-s}$ will occur. Notice that these numbers are independent of the choice of j and k; they only depend upon the choice of the voting vector, the number of elements in D, and N.

Add all of these vector differences from Eq. 5.18 together. The resulting vector has a positive scalar multiple $(\underline{e}_j-\underline{e}_k)$ in each component representing a (j,k) subset of alternatives. This scalar depends only upon the choice of the voting vector and the number of elements in this subset, not on the choice of j,k, or i. Normalize the voting vectors so that this vector is a scalar multiple of $\underline{V}_{JK}$. (To avoid renormalizing the voting vector used to rank pairs of alternatives, the different voting vectors are scaled so that this sum of vectors is $2[(N-2)!]^{N-2}\underline{V}_{JK}$.) Thus, $\underline{V}_{JK}$ is in the corresponding V_W. Because the normalization doesn't depend upon the choice of j and k, this statement is true for all choices of j and k. From this it follows that V_B is a subspace of V_W.

Next, we must show that if $\underline{W}$ is not a Borda vector, then V_B is a proper subset of V_W. To do this, we examine the last argument more closely. Since the choice of j and k only influences the choices of subsets of alternatives being considered, we start by considering the indices 1 and 2, and later we indicate what changes must be made to obtain the general proof.

Consider the vector differences $\underline{W}_{P(A)} - M_{12}(\underline{W}_{P(A)})$ where only the choice of P(A) varies. Furthermore, to simplify the notation, let $w*_s$ denote w_s-w_{s+1}. When A is the standard ranking, this vector difference is $w*_1(\underline{e}_1-\underline{e}_2)$ in each 1,2 component space. Next, we consider rankings where a_1 is the second ranked alternative and a_2 is the third ranked alternative. The only alternatives which concern us are those ranked above a_1 and a_2. So, consider the N-2 rankings obtained in the following order: The j-2 ranking has a_j as the top ranked alternative, the ranking of the alternatives in the k^{TH}, k=4,..,N, position is determined in some arbitrary fashion. Then, the vector components of the difference have a scalar multiple $w*_2$ for those (1,2,j) component spaces, and $w*_1$ for the remaining (1,2) component spaces. (Notice that this vector is independent of what is ranked in the k^{TH} position, k=4,..,N.)

For each $i=3,..,N-2$, continue in the same fashion. Let a_1 be the i^{TH} ranked alternative, and let a_2 be the $(i+1)^{TH}$ ranked alternative. Consider the $(N-2;i-1)$ rankings obtained by choosing sets of $(i-1)$ alternatives ranked above a_1 and a_2. The ordering of these alternatives isn't important, just the choice of the set. For each of these rankings, determine the vector difference. The resulting scalar multiple, w^*_s, in each (1,2) component space depends upon how many of the selected $(i-1)$ elements are also in this set. Notice that the set of vectors obtained in this way includes all differences of Eq. 5.18 where $(j,k)=(1,2)$ and $P(A)$ has a_1 and a_2 adjacent in the ranking.

Now consider these vectors where in each component space we use the appropriate w^*_s (rather than its numeric value). Given any such vector, we can determine the value of i and the elements of the selected set of $(i-1)$ elements. Indeed, this already can be determined by the (1,2) component subspaces of $N-1$ elements. Namely, because only one element is left out of each these subsets, the largest subscript s of w^*_s from these components is the value of i. To determine the set, we compute its complement. For each component subspace with multiple w^*_{I-1}, the element which is in the total set but not in this subset is in the complement. In fact, this same analysis can be done over all of the (1,2) component subspaces of K elements where $K \geq 3$. A consequence of this is that no two of these vectors are the same.

These vectors span a subspace. (This subspace plays the same role as the subspace in the -1 eigenspace for (1,2) in the proof of Theorem 7.) We already know that a linear combination of these vectors equals $\underline{V}_{12}$. If any of these vectors differs from a scalar multiple of $\underline{V}_{12}$, then this subspace is at least of dimension 2. (By comparing the component spaces for the binary pair (a_1,a_2), it follows that this scalar multiple must be 2.) This would be sufficient to prove that V_W properly contains V_B independent of what happens for a similar analysis for the other choices of j and k.

Now suppose each such vector equals $2\underline{V}_{12}$. When the ranking is A, this means that for all (1,2) component spaces, the voting vector has $w_1-w_2=2$. By comparing the vector from all (j,1,2) component spaces, it follows that $w_2-w_3=2$. As j varies, this captures all of the voting vectors for (1,2) component spaces. Continuing in the same fashion through all choices of i, it follows that all of the voting vectors for (1,2) component spaces are Borda Vectors. Applying the same analysis for all choices of j<k proves part a of the Theorem 6.

Proof of part c. For each subset of s alternatives, there are s-1 choices of the differences w^*_s in the definition of the voting vector. The scalar normalization reduces one of these degrees of freedom, so there are s-2 degrees of freedom in the choice of a voting vector. Let $d=\sum_i (N-2;i-2)$. A simple counting argument shows that there are d different subsets of alternatives which contain the pair (a_1,a_2). Therefore, in the choice of the voting vectors to rank these subsets of alternatives, there are $\sum_i (N-2;i-2)(i-2)$ degrees of freedom. The above construction defined d vectors. Now, if these vectors are linearly independent, then the linear space defined by them is of full dimension d. That is, for each (1,2) component space, the vector which has $(\underline{e}_1-\underline{e}_2)$ in this component and $\underline{0}$ in all others is in V_w. This means that V_w contains this subspace. But, linearly dependent means the vectors must satisfy an algebraic condition (from a vanishing determinent). On the other hand, being independent is an "open" condition; if there is at least one vector leading to independence, then this is the standard, generic condition.

To show that most vectors $\underline{W}$ lead to independent vectors in this (1,2) space, it suffices to show that the open set is nonempty. Thus, we only need to show that if $\underline{W}$ corresponds to plurality voting, then all of these vectors are linearly independent. But, for a plurality voting vector, $w^*_1>0$, $w^*_s=0$ for $s>1$. To show independence, we use the above construction, starting with the last ranking and working forward.

If a_1 and a_2 are the bottom two ranked alternatives ($i=N-1$), then only in the subset of two alternatives do they emerge as the top two alternatives. Thus, the vector difference has $\underline{0}$ in all components except in the component space for the pair (a_1,a_2) where the multiple is 2. Next, consider $i=N-2$ where a_j is the bottom ranked alternative. In this vector difference, only the components for the subsets $\{a_1,a_2,a_j\}$ and $\{a_1,a_2\}$ have positive multiples. But, the first stage of $i=N-1$ can be used to obtain $(N-3)$ vectors with zero in all but the $(1,2,j)$ component space, where the entry is $(\underline{e}_1-\underline{e}_2)$. The same argument is continued over the various values of i and the different subsets of alternatives. This completes the proof for (1,2). The same argument holds for all choices of indices $j<k$. This completes the proof of Theorem 6 and it shows that if $\underline{W}$ consists of the plurality voting scheme for all subsets, then $V_W=T$.

Several places in this paper there are informal comments asserting that for certain types of voting methods, certain conclusions hold. For economy of exposition, proofs of these statements aren't supplied, but it is clear from the above proof that these conditions involve showing that a certain number of the $\underline{W}_{P(A)}$ vectors are independent. To assist the reader interested in verifying the comments made here, we outline a geometric proof which simplifies the analysis.

For $N=4$, we indicate the algebraic conditions the vector components of $\underline{W}$ must satisfy in order that $V_W\neq T$. It follows from the proof of Theorem 6 that we want to find the dimension of the subspace spanned by $\underline{W}_{P(A)}-M_{JK}(\underline{W}_{P(A)})$ for each pair $j<k$, and where $P(A)$ varies over all permutations of A such that a_j is the i^{TH} ranked alternative while a_k is the $(i+1)^{TH}$ ranked alternative. So, consider $j=1$, $k=2$. The components for the spaces (a_1,a_2), (a_1,a_2,a_3), (a_1,a_2,a_4), and S_4 in this difference define the vectors

$$\begin{array}{llll}
(2, & w^*_1, & w^*_1, & w^*_1) \\
(2, & w^*_2, & w^*_1, & w^*_2) \\
(2, & w^*_1, & w^*_2, & w^*_2) \\
(2, & w^*_2, & w^*_2, & w^*_3)
\end{array}$$

where the components $w^*_j = w_j - w_{j+1}$ for the voting vectors of the respective subsets. (We suppress the notation which indicates that the choice of the w_j's can vary with the subsets.) The general situation gives a pattern similar to the middle 2x2 matrix. Blocks appear which assume the form of a square matrix with a dominant diagonal term. This is a consequence of the symmetry of the wreath product.

We want this set to have a minimal number of linearly independent vectors. It has dimension one if and only if they are all the same, but this corresponds to where the differences $w_j - w_{j+1}$ are the same constant for all j. These are the Borda vectors. This set of vectors can define a two, three, or four dimensional subset when some of the Borda components are replaced with other voting methods.

Assume now that none of the vectors are Borda, and that for each set, $w_{12} \neq 0$. Because it is the last three components of each vector which determine the independence, we plot these four vectors in three space. The first vector determines the distant vertex of a rectangle, while the next two determine points on two of the three faces of the rectangle containing this vertex. For a minimal dimension of independence, these three points and the origin of the rectangle must lie on the same plane. This imposes a strong algebraic condition on the weights. Moreover, the last point must also lie on this plane.

From this we see why voting systems which distinquish between only two subsets can't satisfy such stringent vector conditions. This is because they can be normalized so that they define a unit cube in the appropriate dimensional space. However, because all but one of the components w_{jk} are zero, the vectors end up on axes or lower dimensional coordinate planes, and the symmetry properties change the locations. Thus, they cannot be on planes of the type described above.

Proof of Theorem 5. For this theorem, the space is $T'=(E^N)x(E^2)^{(N;2)}$. The obvious modification of the above shows that V_B has dimension (N;2). Thus, all we need to show is that if $\underline{W}^N \neq \underline{B}^N$, then $V_W=T'$. But, $\underline{W}^N$ not being a Borda Vector means that there are at least two choices of s, say j and k, where $w*_j \neq w*_k$. Consequently, there are choices of P(A) so that the vector difference in Eq. 5.20 has a $w*_s$ as the multiple of $\underline{e}_1-\underline{e}_2$ in the first component, $2(\underline{e}_1-\underline{e}_2)$ as the component for the binary, and $\underline{0}$ in all other components, s=j,k. It is obvious that these two vectors are linearly independent, so they fill the maximum dimesion for this (1,2) space. This is true for all choices of j<k. Thus, V_W contains T'. This completes the proof.

Proof of Theorem 10. Let T_F be the subspace of T corresponding to the family F. A normal vector to T_F, $\underline{N}$, is in V_B if and only if it can be expressed as a linear combination of the basis vectors. That is, $\underline{N}$ has the desired properties if and only if there exist scalars $D=\{d_{JK}\}$ such that

5.20 $$\underline{N} = \sum d_{JK}\underline{V}_{JK}.$$

Those vector components of $\underline{N}$ which correspond to members of F must be $\underline{0}$; otherwise $\underline{N}$ wouldn't be a normal vector to T_F. Because of the form of the vectors $\underline{V}_{JK}$, this means that for each member B of F, the sum $\sum d_{JK}(\underline{e}_J-\underline{e}_K)=\underline{0}$ where the summation is over the indices with pairs of alternatives in B.

Suppose that there are several sets D which define a cyclic symmetry property for F. Each set defines a vector in the (N;2) space V_B. Now, since all normal vectors to T_F which are in V_B has such a representation, the dimension L is given by the dimension of the space spanned by the D vectors. This completes the proof of Theorem 10.

(Second) Proof of Theorem 9. Assume that F is the family of nested subsets defined in the statement of Theorem 9. As before, assume that $S_j=\{a_1,a_2,..,a_j\}$, $j=2,..,N$. Moreover, assume that there exists a set $D=\{d_{jk}\}$ which defines a cyclic symmetry property for F. By induction, we will show that all of the d_{jk}'s must be zero, which is a contradiction to the definition of D.

Because S_2 is a member of this family, $d_{12}(\underline{e}_1-\underline{e}_2)=\underline{0}$, or $d_{12}=0$.

Assume that for $s\geq 3$, $d_{jk}=0$ if $j<k\leq s-1$. We now show this is true for $j<k\leq s$. But, because S_s is a member of F, the sum Eq 5.20 must hold for all $j<k\leq s$. By the induction hypothesis, this sum is $\sum d_{js}\underline{e}_j-(\sum d_{js})\underline{e}_s=\underline{0}$. Because the vectors $\{\underline{e}_k\}$ are linearly independent, the coefficients for $\underline{e}_j$ must be zero. That is, the terms $d_{js}=0$ for $j<s$. This completes the proof.

Proof of Theorem 8. Let F be the family of subsets defined in the statement of the Theorem. First, assume that $k=N-1$ and that this family has the cyclic symmetry property with set D. By definition, not all of the terms d_{jk} equal zero, so assume without loss of generality that $d_{12}\neq 0$. We use an induction argument to show that $d_{12}=d_{1j}$ for $j=3,..,N$. First, this is shown for $j=3$.

Consider the only member of F which is missing a_3. Eq 5.20 must hold for the associated set of pairs of alternatives. Thus, in order that the $\underline{e}_1$ term vanishes, it follows that

5.21 $$d_{12} = -\sum d_{1j}, \quad j>3.$$

Thus, the summation on the right hand side must be nonzero. Now consider the subset where a_2 is the missing alternative. Again it follows from Eq. 5.20 and from the vanishing of the $\underline{e}_1$ coefficient that

5.22 $$d_{13} = -\sum d_{1j},\ j>3,$$

so $d_{13}=d_{12}$.

Assume that $d_{1j}=d_{12}$ for $j<s$. We now show that $d_{1s}=d_{12}$. To do this, consider the member of F where a_s is the missing alternative. Then, it follows

from Eq. 5.20 that

5.23 $$d_{12} = -\sum d_{1j} \quad \text{where } j>2,\ j\neq s.$$

By using Eq. 5.20 for the subset which is missing a_s, we have that d_{1s} equals the summation on the right hand side of Eq. 5.23. Thus, $d_{12}=d_{1s}$, and the induction proof is completed.

Now, consider any member of F. The $\underline{e}_1$ coefficient from Eq. 5.20 is $(N-1)d_{12}$. Because this coefficient must be zero, it follows that $d_{12}=0$. This contradiction completes the proof for $K=N-1$.

Let K be such that $2<K<N-1$. Assume that F is cyclic symmetric with set D where, without loss of generality, $d_{12}\neq 0$. Consider the the subfamily of F consisting of the K+1 subsets of K elements which can be constructed from the elements $\{a_1,..,a_{K+1}\}$. Because this subfamily is contained in F, it must be cyclic symmetric with respect to D. But, the above argument then shows that $d_{12}=0$. This contradiction completes the proof.

Lemma. Consider the two linear subspaces of T, V_W and V_V, where the second is a proper subspace of the first. Then, there are ranking regions of T which meet V_W but do not meet V_V.

Proof. Let $\underline{N}$ be a vector in V_W which is normal to V_V. Because $\underline{N}$ is in T, each vector component of $\underline{N}$ is normalized, and it corresponds to a ranking of this particular subset of alternatives. It suffices to show that V_V does not meet the ranking region corresponding to the listing of rankings defined by $\underline{N}$. To do this, it suffices to show that any boundary vector for this ranking region is not orthogonal to $\underline{N}$.

The boundary surfaces for this ranking regions are given by the various indifference planes in the component spaces of T. That is, if in some component space, we have $n_j>n_K$, then the bounding plane is given by $x_j=x_K$. Choose the normal vector to this plane which points to the interior of the ranking region, i.e., the vector $\underline{e}_J-\underline{e}_K$. The scalar product with $\underline{N}$ and such a normal vector to the boundary of the ranking region has, for each component space, a positive value $(n_J-n_K>0)$. Thus, all of the boundaries of the ranking region form an angle of less than 90° degrees, so this ranking region cannot meet V_V.

Proof of the theorems in Section 2. We have shown that V_W contains V_B. These theorems follow from the above lemma.

The proofs of the Corollaries 8.1 and 8.2 are obvious combinations of the proofs of Theorems 7 and 8. If the projection of V_B is the total space, then there is no interaction effect among the subsets of the alternatives, and so the proof of Theorem 7 can be used directly. If the projection of V_B isn't equal to the total space, then there is a modification to take into account the interaction effect among the subsets of alternatives given by the Borda vectors. This is the source of the Borda independence condition.

Acknowledgements

This research was supported in part by NSF Grant IST 8415348. Some of this work was done while I was at the System and Decision Sciences Division of IIASA at Laxemberg, Austria, during August and September of 1984.

References

1. Arrow, K., Social Choice and Individual Values, Cowles Foundation for Research in Econ., Monograph 12, New Haven, 1963.

2. Brams, S. and P. Fishburn, Approval Voting, Birkhauser, Boston, 1982.

3. Ginsburg, S., Algebraic Theory of Automata, Academic Press, New York, 1968.

4. Krohn, K., and J. Rhodes, Algebraic theory of machines, Trans. Amer. Math. Soc. (1965), 450-464.

5. McKelvey, R.D., General conditions for global intransititivities in formal voting models: Some implications for agenda control, Econometrica 47 (1979), 1085-1112.

6. Plott, C.R. and M.E. Levine, A model of agenda influence on committee decisions, The American Economic Review 68 (1978), 146-160.

7. Saari, D.G., Inconsistencies of weighted voting systems, Math of Operations Research 7 (1982), 479-490.

8. Saari, D.G., The ultimate of chaos resulting from weighted voting systems, Advances in Applied Mathematics, 5 (1984), 286-308.

9. Saari, D.G., The source of some paradoxes from social choice and probability, NU Center for Math Studies in Econ Discussion paper 609, June 1984.

10. Saari, D.G., and J. von Newenhizen, A case against bullet, approval and plurality voting, NU Discussion paper.