UNIVERSITY OF CALIFORNIA,
IRVINE

Geometric Ways of Understanding Voting Problems

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mathematical Behavioral Sciences

by

Tomas J. McIntee

Dissertation Committee:
Distinguished Professor Donald G. Saari, Chair
Professor Louis Narens
Professor William Batchelder

2015

DEDICATION

For my father.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank...

...my advisor and my dissertation committee members, Drs. Donald Saari, Louis Narens, and William Batchelder;
...Drs. Jennifer Trueblood and Li-Sheng Tseng, who served on my advancement committee;
...Dr. Rick Klima, who introduced me to the mathematics of voting theory;
...Dr. Jeffry Hirst, who got me to want to do mathematical research in the first place;
...fellow "Don Squad" members Jonathan Cook, Santiago Giusasola, Daniel Jessie, Ryan Kendall, Reuben Kline, Jason Kronewetter, and Heidi Tucholski for feedback and inspiration;
...Tony Gielow, Alex Gord, and Mark McIntee for additional error-checking, proof-reading, and feedback on this document;
...my brother, Shea McIntee, for motivation;
...Elsevier fo the necessary permission to include Chapter 2;
...those who provided funding that supported this research - teaching assistantships from the School of Social Sciences, travel and summer funding from the Institute of Mathematical Behavioral Sciences, and the Social Sciences Associate Dean's Fellowship.

CURRICULUM VITAE

Tomas J. McIntee

EDUCATION

Doctor of Philosophy in Mathematical Behavioral Sciences **2015**
University of California-Irvine *Irvine, CA*

Master of Arts in Mathematics **2009**
Appalachian State University *Boone, NC*

Bachelor of Science in Physics **2007**
Appalachian State University *Boone, NC*

Bachelor of Science in Mathematics **2007**
Appalachian State University *Boone, NC*

Bachelor of Arts in Philosophy & Religion **2007**
Appalachian State University *Boone, NC*

ABSTRACT OF THE DISSERTATION

Geometric Ways of Understanding Voting Problems

By

Tomas J. McIntee

Doctor of Philosophy in Mathematical Behavioral Sciences

University of California, Irvine, 2015

Distinguished Professor Donald G. Saari, Chair

General conclusions relating pairwise tallies with positional (e.g., plurality, antiplurality ("vote-for-two")) election outcomes were previously known only for the Borda Count. While it has been known since the eighteenth century that the Borda and Condorcet winners need not agree, it had not been known, for instance, in which settings the Condorcet and plurality winners can disagree, or must agree. Results of this type are developed here for all three-alternative positional rules. These relationships are based on an easily used method that connects pairwise tallies with admissible positional outcomes; e.g., a special case provides the first necessary and sufficient conditions ensuring that the Condorcet winner is the plurality winner; another case identifies when there must be a profile whereby *each candidate* is the "winner" with some positional rule. Previous work relating the probability of positional and pairwise tallies have used specific selected distributions (primarily the Impartial Culture and Impartial Anonymous Culture assumptions) and specific voting rules (particularly plurality). Techniques are developed here that can be applied to analyzing the probability of conflict between all different positional methods, and between combinations of pairwise tallies with positional results. Results are given for several broad categories of probability distribution, along with a qualitative analysis of the relationship between probability distributions over voter profiles and the likelihood of voting paradoxes. A method of geometrically comparing multiple-stage and single-stage elections is developed, which shows that multiple stage elec-

tions are not necessarily more vulnerable to being manipulated, but less vulnerable when all rank-order outcomes matter, and specifically only similar when an election only identifies a first-place winner. In the case where results are defined in terms of a singular winner, a plurality vote is identified as less manipulable in a single stage than in multiple stages, while an antiplurality vote is identified as more vulnerable in a single stage than in multiple stages.

Chapter 1

Overview

After a several centuries of mathematical study of voting systems, it is clear that determining which voting method most accurately reflects the views of the voters in some particular set of circumstances is a surprisingly subtle challenge. In the last half century, it has become widely accepted that there is *no* perfect voting system when there are more than two candidates in an election. Several lists of unattainable combinations of properties for elections with more than two candidates or options are now famous impossibility theorems for voting problems. This includes Arrow's theorem (Pareto unanimity and independence of irrelevant alternatives), Sen's theorem (Pareto unanimity and liberal rights), and the Gibbard-Satterthwaite theorem (non-manipulable and deterministic). Efforts to determine which voting system is best have continued in spite of acceptance that no system is perfect, primarily but not exclusively by the use of examples constructed to show differences between voting systems.

Even the oldest puzzles surrounding voting theory remain of interest. The most famous puzzle is perhaps the Condorcet paradox, the namesake of which was alive during the French Revolution; results are still being published in voting theory relating to the nature and importance of the Condorcet paradox. The Condorcet parodox is not a result that relates to the outcome of voting rules directly; it is a problem that occurs as the result of preference structures alone, and looks at inconsistences within pairwise comparisons. The fact that a Condorcet (pairwise) winner may not exist, even when all voters have transitive and complete preferences, is an intrinsic property of profiles of transitive preferences. The frequency with which Condorcet paradoxes occur remains a matter of current interest; this question is answered with new thoroughness in Chapter 3.

More generally, the Condorcet paradox is a paradox of the relationship between profiles and pairwise majority votes. A complete qualititative understanding of the full relationship between profiles as collections of transitive preferences, and the pairwise tallies related to those collections of transitive preferences, has been elusive until fairly recently (see Chapter 2). That understanding is in turn necessary for a full understanding of the context of the Condorcet paradox and its true importance in Chapter 3).

However, the puzzle of the profile is only half of the problem. The effect of Condorcet-type components (i.e., the general component of preference profiles that lead to a Condorcet paradox) on particular voting systems, and how different voting systems treat Condorcet winners and Condorcet losers, brings us to a second class of puzzle. There are almost certainly questions which we have not thought to even ask about the different behaviors of particular voting systems, which are nevertheless of some importance. Additionally, there exist a number of questions whose answers are known but whose significance is not necessarily well understood.

Several examples: Will the system elect a Condorcet winner if one exists? Can it elect a Condorcet loser? Is it monotone? Prior research has led to a library of criteria for distinguishing between voting systems has been developed; many of which, though not all, are related to specific voting paradoxes or impossibility theorems. This work adds to the understanding of the significance of impossibility theorems and of criteria violations.

Most of these criteria are binary in nature. That is to say, the criteria is either satisfied or not. The questions they pose are *yes* or *no* questions, such as the above answers. For example, a Condorcet winner can lose in a plurality vote, a Borda count, and most other voting systems that do not explicitly check for a Condorcet winner. A Condorcet loser cannot win a Borda count, but can win in *any* other positional method, and any multi-stage elimination method consisting of two or more positional votes over three or more candidates has the potential for non-monotonicity. In regard to this vein of research, the present work aims to answer the question of how *likely* a voting paradox is to occur. Prior research on this has involved a fairly narrow set of conditions, or a small set of actual elections (e.g., as in Gehrlein and Lepelley [2011]); Chapter 3 attempts to put the question of the probability of a Condorcet paradox to rest for the general setting, inasmuch as any theoretical work can do so. (Empirical measurement of the actual frequency of a Condorcet paradox within a given population is beyond the scope of any theoretical work.)

Chapter 4 examines the problem of manipulability, and the sensitivity of multiple-stage elections to manipulation. It is known that all multi-stage systems can exhibit non-monotone behavior, which produces a strong incentive for strategic manipulation; does this mean that a multi-stage election is necessarily more vulnerable to strategic manipulation than a single-stage system? Not necessarily; in fact, introducing additional stages to a voting system tends to increase its stability.

The same tools which allow us a complete understanding of how likely the Condorcet paradox is under a variety of large classes of probability distributions can also be used to explore a variety of other events - such as the probability that a plurality vote elects someone other than a Condorcet winner. By tying together these sorts of different results, we can better see the larger picture of the relationships between voting systems and voting paradoxes.

The primary method used in this work is decomposition. To better understand a problem that exists in $n!$ dimensions, we reduce it to a more tractable and understandable combination of lower-dimensional subspaces. The essential profile; the representation cube; result space; reversal components; all of these are lower-dimensional spaces that display salient features of interest. This is not to say that we need to, or should, ignore the full dimensionality of the problem; but by dealing with different sets of dimensions separately, results can be made simpler, more understandable, and more general.

In Chapter 2, the 6-dimensional simplex of possible profiles over $n = 3$ candidates is broken up between an *essential* profile - corresponding precisely to a point on the representation cube - and the reversal components which produce all differences between different positional methods. Understanding the relationships between the magnitude of those two spaces is simple: The closer to the surface of the representation cube, the smaller the supporting space of reversal components is, and the smaller the class of binary-equivalent profiles.

This simple relationship lets us go from the essential profile to the magnitude of variation between the different possible positional rules (for $n = 3$ candidates, rules which give $(1, s, 0)$ points to first, second, and third place on each ballot); and from there we have explicit equations laying out when it is possible for various events to happen, e.g., a plurality winner and a Borda winner disagreeing or a plurality winner and Condorcet winner being the same. Examining the magnitude of the spaces provides qualitative information about the probabilties of these events. Without knowledge of the underlying probability distribution of profiles, this qualitative information is very useful; and allows for rapid estimation of the relative magnitude of probabilities once a particular probability distribution is selected.

In Chapter 3, the essential profile is placed directly and explicitly on the representation cube (and a variant thereof) in order to display the relationship between pairwise majority tallies and probability distributions over the space of possible profiles. This also allows direct specification and further understanding of the types of probability distributions we may be dealing with on the representation cube itself. By integrating separately over the supporting space, a probability density distribution in the original profile space can be transformed into a probability density distribution on either the representation cube or another similar space.

Having the probability distribution function over the representation cube, instead of profile space, allows us to make easier quantitative calculations of the probability various voting paradoxes or events of interest. It also shows, qualitatively, what happens to the relationship between pairwise majority tallies and positional results as the probability distribution function over profile space is varied. This in turn provides an understanding of how these phenomena vary over the space of possible probability distributions. (Or rather, over part of that space; the space of *all* possible probability distributions is very large.)

In the event that a Condorcet paradox occurs, for example, it does not matter what the variance of the probability distribution is on the representation cube, simply the overall shape; this leads to easy calculation of the probability of a Condorcet paradox. In the event that a plurality winner agrees with the Condorcet winner, by contrast, this probability is much more sensitiv to variance on the representation cube. The variance on the representation cube is the more important factor because close to the origin, the large variation within the supporting space becomes more important; and in particular, the relationship between the variance on the cube and the variance on the supporting space becomes critical.

Chapter 4 takes the reduction as far as possible and make use of the *result space*, which is to say the space of possible point totals - a reduction from $n!$ dimensions to n dimensions. Using this space allows for the comparison of methods which are not simply positional or multiple positional methods. In particular, this method can be used to analyze multi-round and single-round versions of positional and multiple positional methods.

In spite of the limited information the space of results provides, this turns out to be entirely adequate for discussing *stability* of the results against some small change in the votes cast - some small amount of counting error, fraud, or other form of manipluation. The details of the supporting profile that produce a result are only relevant as part of the underlying probability distribution; the two factors which matter are the density distribution of subtotal results, and the structure of the votes that can be added or subtracted.

The present work aims to take a general and conceptual approach. While there are some specific calculations offered on probabilities under particular probabilistic assumptions (particularly in Chapter 3), the only specific figure offered that is particularly important is the range of probabilities for a Condorcet paradox under a broad range of assumptions: Between 6.25% and 8.8%. Within that range, it is very difficult to empirically distinguish between different probability assumptions; evidence falling outside of that range, on the other hand, implies that the underlying distribution of preferences is very unlike the sorts of distributions usually assumed. Most other probabilistic calculations are important only qualitatively; since the true distribution of preferences is not known, and may not be the same in all populations, it is important to understand qualitatively what is going on.

It is often said that a picture is worth a thousand words. What follows in Chapters 2-4 in the body of this work is an expression of agreement with that statement. Geometric methods make full use of this asymmetry between words and pictures, because they *can* be readily illlustrated, leading to easier conceptual understanding of paradoxes, voting criteria, and the qualitative differences in different voting systems.

Chapter 2

Connecting Pairwise and Positional Election Outcomes

This chapter consists of joint work with D. Saari. A version of this chapter has been published in Mathematical Social Sciences (Vol. 6 #2, p140-151).

2.1 Introduction

After a quarter of a millennium of study, it is clear that the objective of determining which voting method most accurately reflects the views of the voters is a surprisingly subtle, major challenge. The complexity of this issue has forced researchers to adopt secondary measures, such as seeking properties of specific rules or probability estimates of paradoxical events. While providing useful information, these approaches remain surrogates for the true intent of identifying which profiles cause different kinds of election outcomes. Rather than determining the likelihood of particular paradoxical outcomes, for instance, a preferred outcome would be to identify all profiles that cause these difficulties.

To advance our understanding of which profiles create various conclusions, the approach introduced here identifies all three-alternative profiles that support specified paired majority vote tallies. An advantage of knowing all possible supporting profiles is that it makes it possible to determine all of the associated positional outcomes.

To illustrate the variety of new questions that can be answered, suppose all we know about a profile is that its majority vote pairwise comparisons are

A beats B by 70:30, A beats C by 60:40, and B beats C by 55:45.

Here A is the Condorcet winner (she beats all other candidates) and C is the Condorcet loser (she loses to everyone). Just from these tallies, where the two involving the Condorcet winner A are of "landslide proportions" (winning 60% or more of the vote), the goal is to determine all admissible plurality (vote-for-one), antiplurality (a "vote-for-two" is equivalent to a "vote against one"), Borda (assign two and one points, respectively, to a ballot's first and second positioned candidate) and other positional outcomes. Even though Condorcet winner A badly defeats the Condorcet loser C, are there profiles with these majority votes where C, rather than A, is the plurality winner? Could C be an antiplurality winner? Could middle-ranked B win a plurality election? Do the Borda and Condorcet winners agree or differ? (Complete answers are in Sect. 2.4.2.)

The easily used approach developed here connects majority votes with positional outcomes, so this method becomes a central tool to answer all such questions. As our intent is to develop relationships between positional methods and pairwise votes, only sincere voting is considered. (Proofs are in Appendix A.1, but the basic ideas are developed in Sects. 2.3.1, 2.3.2)

2.1.1 Basic definitions

A profile lists each voter's ranking of the alternatives where it is assumed that each voter has complete, strict (no ties), transitive preferences. As for outcomes, assume in an $\{A, B\}$ majority vote that A always has at least as many votes as B, and, in a $\{B, C\}$ majority vote, B has at least as many as C. With this assumption and by denoting a strict preference by "$\succ$," a strict transitive outcome of these paired comparisons is $A \succ B \succ C$. So, if there is a Condorcet winner in what follows, it always is A, and C always is the Condorcet loser. If the rankings define a cycle, it has the $A \succ B, B \succ C, C \succ A$ form. A "name change" converts any other situation into our setting, so this assumption does not affect the generality of our conclusions.

Rather than using the actual tallies, the differences between majority vote tallies turns out to be a more useful way to analyze these issues.

Definition 2.1. *For an $\{X, Y\}$ majority vote election with N voters, let*

$$P(X, Y) = \{X's \text{ majority vote}\} - \{Y's \text{ majority vote}\}. \tag{2.1}$$

Illustrating with the introductory example, $P(A, B) = 70 - 30 = 40$, $P(A, C) = 20$, and $P(B, C) = 10$; e.g., the larger the $P(X, Y)$ value, the better X does against Y. Also, $P(X, Y) = -P(Y, X)$; e.g., in the introductory example, $P(B, A) = 30 - 70 = -40$. This notation converts our $A \succeq B, B \succeq C$ assumption about the paired elections (where "$\succeq$" has the obvious "preferred or indifferent to" meaning) into the equivalent $P(A, B) \geq 0$, $P(B, C) \geq 0$ condition.

With N voters, because $N = \{X's \text{ vote}\} + \{Y's \text{ vote}\}$, it follows that

$$\{X's \text{ vote}\} = \frac{1}{2}[N + P(X, Y)]. \tag{2.2}$$

So, with $N = 60$ voters, a $P(A,B) = 10$ outcome means that A received $\frac{1}{2}[60+P(A,B)] = 35$ votes while B received $\frac{1}{2}[60 + P(B,A)] = \frac{1}{2}[60 - 10] = 25$ votes. Because X's tally is an integer, it follows from Eq. 2.2 that N and all three $P(X,Y)$'s must have the same parity; i.e., either all are odd integers, or all are even integers. This parity agreement is used throughout the paper.

Different profiles can yield the same pairwise tallies, so the following definition is introduced to collect all of them into one class.

Definition 2.2. *Two profiles* $\mathbf{p}_1$ *and* $\mathbf{p}_2$ *are "binary equivalent" (*$\mathbf{p}_1 \sim_{BE} \mathbf{p}_2$*) if they have identical* $P(A,B), P(B,C), P(A,C)$ *values.*

The $\sim_{BE}$ connection is an equivalence relationship. (That is, for each $\mathbf{p}_i$, $\mathbf{p}_i \sim_{BE} \mathbf{p}_i$; if $\mathbf{p}_1 \sim_{BE} \mathbf{p}_2$, then $\mathbf{p}_2 \sim_{BE} \mathbf{p}_1$; and, finally, if $\mathbf{p}_1 \sim_{BE} \mathbf{p}_2$ and $\mathbf{p}_2 \sim_{BE} \mathbf{p}_3$, then it must be that $\mathbf{p}_1 \sim_{BE} \mathbf{p}_3$.) Thus $\sim_{BE}$ partitions the space of profiles into equivalence classes; each class consists of all profiles with the same $P(X,Y)$ values, $X, Y = A, B, C$. To be useful, a trait must be found to identify which profiles belong to a particular $\sim_{BE}$ class. As described in Thm. 2.5, this is the profile's "essential part" – the unique portion of a profile that determines the $P(X,Y)$ values.

Answers for the above questions are found by characterizing how profiles in a $\sim_{BE}$ class differ. As a preview of what will be discovered, although profiles from the same class have identical $P(X,Y)$ values, their plurality and other positional rankings can differ.

Definition 2.3. *A* positional voting rule *tallies a ballot by assigning specific points to the three alternatives according to how they are positioned on a ballot. Here,* w_j *points are assigned to the* j^{th} *positioned choice. The conditions are that not all* w_j *values are equal, and* $w_1 \geq w_2 \geq w_3 = 0$*. The* normalized positional rule *is obtained by dividing all* w_j *values by* w_1 *to create* $\mathbf{w}_s = (1, s, 0)$ *where* $s = \frac{w_2}{w_1}$ *represents the number of second ranked points.*

The plurality vote is $\mathbf{w}_0 = (1,0,0)$, the anti-plurality vote is $\mathbf{w}_1 = (1,1,0)$, and the Borda Count (normally defined by $(2,1,0)$) has the $\mathbf{w}_{\frac{1}{2}} = (1,\frac{1}{2},0)$ normalized form. Although we discuss only positional methods, all results extend immediately to other rules, such as Approval Voting and Cumulative Voting, by using the techniques developed in Saari [2010].

A technical assumption, which is needed to separate the different cases, follows:

Definition 2.4. *A profile satisfies the* strongly non-cyclic *condition if*

$$P(A,C) \geq \min(P(A,B), P(B,C)). \tag{2.3}$$

To explain Def. 2.4, if the pairwise rankings define a transitive $A \succ B \succ C$ outcome, the strongly non-cyclic condition (Eq. 2.3) just means that Condorcet winner A's victory over the Condorcet loser C is at least as decisive as A's victory over B, or B's victory over C. With the $P(A,B), P(B,C) \geq 0$ assumption, Eq. 2.3 requires $P(A,C) \geq 0$, so it precludes cycles; this leads to its "non-cyclic" name. The "strongly" modifier refers to the fact that even if weaker $P(A,C)$ values fail Eq. 2.3, they do not define cycles if $P(A,C) \geq 0$. An $N = 100$ example is where $P(A,B) = 20$, $P(B,C) = 14$ and $P(A,C) = 10$. (To create all profiles with these pairwise values, see Thm. 2.5 and Sect. 2.3.)

2.1.2 Sample of outcomes

Our approach decomposes a profile into the part that determines the $P(X,Y)$ paired comparison values – the portion essential for a profile to be in the $\sim_{BE}$ equivalence class – and the part that affects only positional outcomes. This decomposition simplifies discovering and proving new conclusions. Samples of the kinds of results that can be found are given in the following four theorems. (Proofs are in Appendix A.1, but intuition and partial proofs are given in Sects. 2.3.1, 2.3.2.)

To introduce the first theorem, it is reasonable to wonder whether the plurality and Condorcet winners agree. The first two parts of Thm. 2.1 ensure that this always is true for at least one profile that supports the pairwise tallies. But the next two parts show that the pairwise tallies must satisfy exacting conditions in order for the Condorcet winner A to be the plurality winner for *all supporting profiles.*

Theorem 2.1. *With the strongly non-cyclic condition (Eq. 2.3), the following are true:*

1. *A profile supporting the paired outcomes always exists where the Condorcet winner A is a plurality winner. She may, however, be tied with another candidate.*

2. *If* $P(A,B), P(A,C) > 0$*, there is at least one profile supporting the paired tallies where the Condorcet winner A is the sole plurality winner.*

3. *If with alternative* Y*,* $P(A,Y) > 0$ *is the largest pairwise victory, then a necessary and sufficient condition for all supporting profiles to have the Condorcet winner A as the sole plurality winner (where X is the third alternative) is*

$$2P(A,X) + P(A,Y) > n. \tag{2.4}$$

   If Eq. 2.4 is an equality, there are profiles where A is tied with X. If the Eq. 2.4 inequality is reversed, then some profiles have X as the sole plurality winner.

4. *For* $P(B,C) > \max(P(A,B), P(A,C))$*, a necessary and sufficient condition for all profiles to have A as the sole plurality winner is*

$$2P(A,B) + P(B,C) > n. \tag{2.5}$$

Statement 1 is required for completeness. The ten-voter profile where five voters prefer

$A \succ B \succ C$ and five prefer $C \succ A \succ B$, for instance, has:

$$P(A, B) = 10, P(A, C) = P(B, C) = 0 \tag{2.6}$$

so it satisfies Def. 2.4 but not the part 2 condition. It will turn out (Thm. 2.5) that this is the only ten-voter profile supporting these $P(X, Y)$ values; its plurality $A \sim C \succ B$ outcome (where "$\sim$" means a tie) is as stated in Thm. 2.1, part 1.

Part (2) asserts that *some supporting profiles* must elect A, but not necessarily all of them. For instance, the fifteen voter profile where eight prefer $A \succ B \succ C$, five prefer $C \succ A \succ B$ and two prefer $C \succ B \succ A$ has the $P(A, B) = 13, P(B, C) = P(A, C) = 1$ values with the plurality $A \succ C \succ B$ outcome. In contrast, the profile where six prefer $A \succ B \succ C$, seven prefer $C \succ A \succ B$ and two prefer $B \succ A \succ C$ has identical $P(X, Y)$ values but a different plurality winner with the $C \succ A \succ B$ plurality ranking.

This example motivates Eq. 2.4, which identifies all possible settings where the plurality winner always is the Condorcet winner. With $P(A, C) = 10$ and $P(A, B) = 8$ (so $Y = C, X = B$), for instance, the Condorcet winner A must be the plurality winner if and only if the number of voters satisfies $2(8) + 10 = 26 > n$. Thus, A is both the plurality and Condorcet winner for all supporting profiles if and only if there are no more than 24 voters. With 26 voters, a profile exists with these $P(A, B), P(A, C)$ values where A and B are tied in a plurality election; with 28 voters, a supporting profile has B as the plurality winner. (Equations 2.4 and 2.5 differ slightly because the pairwise middle-ranked B, rather than the Condorcet winner A, defines the largest pairwise victory.) In this way, Thm. 2.1 provides new, general results about the number of voters required for various assertions.

The lower bounds for Eqs. 2.4, 2.5 involve N, which can require huge $P(X, Y)$ victory margins to ensure agreement between the plurality and Condorcet winners. To satisfy Eq. 2.4, for instance, it must be that $P(A, Y) > \frac{1}{3}n$, which means that the Condorcet winner A

receives more than two-thirds of the vote when compared with Y.[1] In other words, unless the Condorcet winner A exhibits exceptional dominance over the other alternatives in paired comparisons, expect the Condorcet and plurality winners to differ.

This required level of dominance exceeds even the landslide proportions of the introductory example! To use Thm. 2.1 to analyze this case, because $P(A,B) = 40$ is the strongest pairwise victory, $Y = B$, $X = C$, and $N = 100$. These pairwise tallies define a strict transitive ranking, so (Thm. 2.1, part 2) some supporting profiles have A as the sole plurality winner. But $2P(A,C) + P(A,B) = 80 < 100$ reverses the Eq. 2.4 inequality, so other supporting profiles have the Condorcet loser C as the sole plurality winner. (As developed later, there are no supporting profiles that elect B.)

If the strongly non-cyclic condition is not satisfied, the requirements for A to always be the plurality winner are slightly more complicated, but similar in form.

Theorem 2.2. *Suppose Eq. 2.3 is not satisfied (so $P(A,C) < \min(P(A,B), P(B,C))$). Whether there is, or is not, a cycle, a necessary and sufficient conditions for A to always beat B, and for A to always beat C, in a plurality vote are, respectively,*

$$2P(A,B) + P(B,C) > n, \quad P(A,B) + 2P(B,C) > n. \tag{2.7}$$

Both inequalities must be satisfied for A to always be the plurality winner. A benefit of Thm. 2.2 is that it also applies to cycles, such as $P(A,B) = 6, P(B,C) = 14, P(A,C) = -4$. Although C beats A in their pairwise vote, it follows from Eq. 2.7 that A always is plurality ranked over C if and only if $6 + 2(14) = 34 > n$, or if there are no more than 32 voters. It follows from the first Eq. 2.7 inequality that for A to always beat B (and to be the sole plurality winner for all supporting profiles), there can be no more than 24 voters.

[1] The best case for Y is if $P(A,X) = P(A,Y)$ where $P(A,Y) > \frac{n}{3}$. According to Eq. 2.2, A receives $\frac{1}{2}[n + P(A,Y)] > \frac{1}{2}[n + \frac{n}{3}] = \frac{2}{3}n$ votes.

These stringent conditions make it difficult for the Condorcet and plurality winners to always agree. This suggests exploring whether other positional rules enjoy more relaxed requirements. But as asserted next (Thm. 2.3), this is not the case for the antiplurality rule; instead, conditions ensuring that the Condorcet and antiplurality winners always agree impose more demanding constraints on the paired victories.

Theorem 2.3. *With the strongly non-cyclic condition (Eq. 2.3), the following are true:*

1. *If with alternative Y, $P(A,Y)$ is the largest paired victory, a necessary and sufficient condition for A to be the only antiplurality winner for all supporting profiles is*

$$2P(A,Y) > n + P(B,C). \tag{2.8}$$

2. *If $P(B,C) > \max(P(A,B), P(A,C))$, A cannot be the only antiplurality winner. If $2P(B,C) > n + P(A,C)$, then A never can be an antiplurality winner.*

With the $P(B,C) \geq 0$ value on the right-hand side, Eq. 2.8 requires a stronger pairwise victory for A over some alternative than needed for the plurality vote (Eq. 2.4). Illustrating with the introductory example, where $2P(A,B) = 80 < 100 + P(B,C)$ violates Eq. 2.8, it follows that some supporting profiles elect someone other than A as the antiplurality winner. The Eq. 2.8 condition also demonstrates how difficult it is for A to be the sole antiplurality winner. Even in the extreme case where B and C tie (so $P(B,C) = 0$, which means from Eq. 2.8 that $P(A,Y) > \frac{n}{2}$), it follows from Eqs. 2.2, 2.8 that A's victory over some candidate must give her more than 75% of the vote!

The last assertion shows that a strong pairwise victory of middle-ranked B over the Condorcet loser C jeopardizes A's antiplurality standing. With $N = 60$ voters, if $P(B,C) = 40$ and $P(A,B) = P(A,C) = 16$, then the Condorcet winner A can never be an antiplurality winner! In contrast, all supporting profiles have A as the sole plurality winner (Eq. 2.5).

A last illustration of the kind of results that can be derived from our approach compares the Condorcet and Borda winners.

Theorem 2.4. *In all cases (that is, independent of whether Eq. 2.3 is satisfied), if A is a Condorcet winner, then a necessary and sufficient condition for the Borda and Condorcet winners to agree is the more relaxed:*

$$2P(A,B) + P(A,C) > P(B,C). \tag{2.9}$$

Illustrating with the introductory example, as $2P(A,B) + P(A,C) = 100 > P(B,C) = 10$, it follows from Eq. 2.9 that A is the Condorcet and Borda winner.

By avoiding a lower bound with N, Eq. 2.9 is a significantly more relaxed condition than required for the plurality and antiplurality rules. This means it is easier and far more likely for the Condorcet and Borda winners to agree than, say, the Condorcet and plurality winners. But to indicate how to extract new statements from these inequalities, notice that Eq. 2.9 imposes an upper bound on $P(B,C)$ while Eq. 2.5 does not. This difference allows profiles to satisfy Eq. 2.5 (so A is the Condorcet and plurality winner) but not Eq. 2.9 (B, not the Condorcet winner A, is the Borda winner). As an example requires a large $P(B,C)$ value, with $N = 2k + 1 \geq 5$ voters, let $k + 1$ of them prefer $A \succ B \succ C$ while k prefer $B \succ C \succ A$. Because $P(B,C) = n$ and $P(A,B) = P(A,C) = 1$, Eqs. 3, 2.5 are satisfied, as they must because A is the plurality and Condorcet winner. But Eq. 2.9 is not satisfied, so B is the Borda winner. (With this profile, B's superiority over A is obvious.) More generally, it follows from these inequalities that if $P(B,C) > P(A,C) \geq P(A,B)$, then a necessary (but not sufficient) condition for A to be the sole plurality, Borda, and Condorcet winner is $4P(A,B) + P(A,C) \geq n + 2$.

The form of Eq. 2.9 also suggests that it can be difficult to find actual elections where the Condorcet and Borda winners differ. This is because to violate Eq. 2.9 (to make B

the Borda winner), the Condorcet winner A must experience narrow victories over both the Condorcet loser C and the middle-ranked B, while the middle-ranked B must enjoy a relatively substantial victory over the Condorcet loser C. (Also see Saari [1999].)

2.1.3 Contributions to the literature

It is often stated that the Borda and Condorcet winners need not agree. To appreciate whether this comment has any significance, it is necessary to identify the settings in which the two winners disagree. Theorem 2.4 does this; it specifies precisely where agreement can, and cannot, happen. For the winners to differ, the profile must be of a special, perhaps unusual type. (How unusual is explored in Section 2.3.)

Beyond settling this question for the Borda and Condorcet winners, it would be useful to derive similar relationships for other positional rules. As the plurality vote is so widely used, for instance, a valuable result would specify when the plurality and the Condorcet winners must agree, and when they can disagree. Prior to this paper and Thm. 2.1, general conditions in terms of the paired comparison tallies were not known.

The plurality and antiplurality rules fare so poorly with respect to the Condorcet winner that the next step is to explore whether this burden extends to other positional voting rules. More relaxed requirements do hold for the Borda Count (Thm. 2.4), but it is not clear what conditions are needed, for example, to ensure that the $(3, 1, 0)$ winner (i.e., $\mathbf{w}_{\frac{1}{3}}$) always is the Condorcet winner. Necessary and sufficient conditions of this kind are developed (Thms. 2.6, 2.7) for all three-alternative positional voting rules. They prove that as $s \to \frac{1}{2}$ (i.e., as the normalized form of the positional rule approaches that of the Borda Count), requirements ensuring that the $\mathbf{w}_s$ and Condorcet winners agree become more relaxed. Thus consistency is more likely to occur with positional methods that more closely resemble the Borda Count.

Another contribution of this paper is to introduce an easily used method to construct profiles. To illustrate, while various $P(X, Y)$ values are specified in the above discussion, it need not be obvious whether profiles exist to support them. This is particularly true with lopsided values such as $P(A, B) = 20, P(B, C) = 40, P(A, C) = -30$. But should the $P(X, Y)$ values satisfy a minimal requirement (Cor. 1), profiles always exist. A way to construct them is developed in Thm. 2.5 and Sect. 2.3.1. This result makes it simple to find, say, all possible 100 voter profiles with specified $P(X, Y)$ values. This new ability to identify all possible supporting profiles is our central tool.

As this method also permits us to find all possible profiles that generate specified conflicting outcomes, it allows new concerns to be addressed. For instance, while it has been known since the eighteenth century that the Borda and Condorcet winners can disagree, to the best of our knowledge the likelihood of this disagreement has not been computed. Similarly, it is worth determining the likelihood that the Condorcet and plurality winners disagree (Eq. 2.4). By using results and the approach developed here, issues of this kind are addressed in a companion paper.

2.1.4 Finding all outcomes

The above theorems, and the following extensions, are derived by identifying all profiles that support specified paired comparison tallies. That is, all profiles in a $\sim_{BE}$ equivalence class are found, which means that all associated positional outcomes can be determined. This emphasis on which positional election outcomes can accompany specified paired majority vote tallies makes our results a converse to the (Sieberg and McDonald [2011]) contribution of identifying which majority vote outcomes can accompany specified plurality tallies; e.g., they examined when a plurality tally ensures whether a Condorcet winner, or cycle, can arise.

Table 2.1: Rankings for Fig. 2.1a

No.	Ranking	No.	Ranking	No.	Ranking
1	$A \succ B \succ C$	2	$A \succ C \succ B$	3	$C \succ A \succ B$
4	$C \succ B \succ A$	5	$B \succ C \succ A$	6	$B \succ A \succ C$

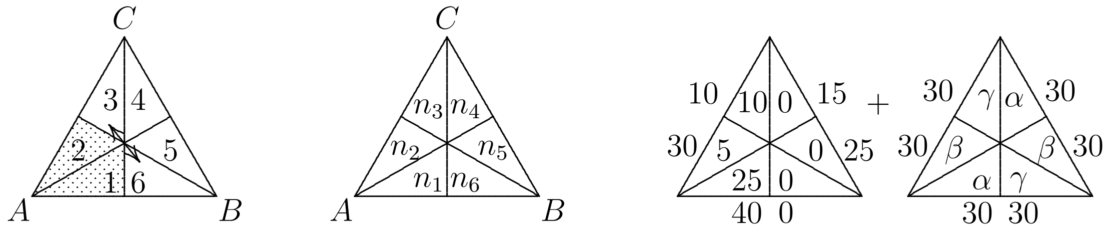
(2.10)

Our emphasis on tallies also distinguishes our results from the literature; e.g., for any number of candidates, it is known which majority vote and positional *rankings* can accompany each other (Saari [1989, 2000, 2008]). With three candidates, for instance, anything can happen with a non-Borda method (i.e., for $\mathbf{w}_s$, $s \neq \frac{1}{2}$). Namely, select any ranking for each pair of candidates and any ranking for the triplet; there exists a profile with these majority and $\mathbf{w}_s$ vote outcomes. (To construct illustrating profiles, see Saari [1999] or [Saari, 2008, Chap. 4].) The missing refinement (which is developed here) is to connect majority vote tallies with all associated positional outcomes. Our approach depends upon a profile decomposition (Saari [1999, 2008]) that identifies the precise profile portions that cause all possible differences between, say, the plurality and antiplurality election outcomes, or with paired comparisons.

2.2 Paired comparisons

Central to our approach is a geometric way to tally ballots (developed in Saari [1995, 2011]; for applications to actual elections, see Nurmi [2002]). Start by assigning each alternative, A, B, C, to a vertex of an equilateral triangle. The ranking assigned to a point in the triangle is determined by its distance to each vertex where "closer is better." Thus, points in the Fig. 2.1a labeled regions have the rankings: For a given profile, let N_j be the number of voters with the j^{th} preference ranking; place N_j in the j^{th} ranking region as indicated in Fig. 2.1b. The geometry conveniently separates these values in a manner to simplify the tallying of ballots. For example, to compute pairwise votes, just sum the numbers on each side of an edge's perpendicular bisector as indicated in the first Fig. 2.1c triangle; the

Figure 2.1: Computing tallies



a. Profiles and tallying **b.** Example **c.** Initial example

paired comparison tallies are listed below the appropriate edge. For instance, the vertical line separates preferences where $A \succ B$ (on the left) from $B \succ A$ (on the right) leading to the 40:0 tally supporting $A \succ B$. Similarly, $A \succ C$ by 30:10 and $B \succ C$ by 25:15.

2.2.1 Computing positional outcomes

A candidate's plurality tally is the sum of entries in the two regions sharing her vertex. For A, this tally is the sum of entries in the two Fig. 2.1a shaded regions. A candidate's $\mathbf{w}_s$ tally is

> {her plurality tally} plus {s times the number of voters who have her second ranked}.

For A with Fig. 2.1b values, add to her plurality tally s times the sum of entries in the two Fig. 2.1a regions with arrows in them; e.g., the $\mathbf{w}_s$ positional tallies for A and B are, respectively, $(n_1+n_2)+s(n_3+n_6)$ and $(n_5+n_6)+s(n_1+n_4)$. For the first Fig. 2.1c triangle, the positional tallies for

$$A : B : C \text{ are, respectively, } 30 + 10s : 25s : 10 + 5s. \tag{2.11}$$

With these tallies, the Condorcet winner A wins with any $\mathbf{w}_s$, the Condorcet loser C is second ranked for $\mathbf{w}_s$, $0 \leq s < \frac{1}{2}$, but B advances to second position for $\frac{1}{2} < s \leq 1$.

To list the tallies as a point in $\mathbb{R}^3$, let $V_s(\mathbf{p})$ be profile $\mathbf{p}$'s $\mathbf{w}_s$ tallies listed in the (A, B, C) order. So with Eq. 2.11 ($\mathbf{p}$ from the first Fig. 2.1c triangle), $V_s(\mathbf{p}) = (30+10s, 25s, 10+5s)$. Notice that $V_s(\mathbf{p}) = (1-s)(30,0,10)+s(40,25,15) = (1-s)V_0(\mathbf{p})+sV_1(\mathbf{p})$ defines a straight line connecting the profile's plurality and antiplurality tallies. This line (which is used in Sect. 2.4) represents a general behavior: For any profile $\mathbf{p}$,

$$V_s(\mathbf{p}) = (1-s)V_0(\mathbf{p}) + sV_1(\mathbf{p}), \quad 0 \leq s \leq 1. \tag{2.12}$$

The Eq. 2.12 line segment in $\mathbb{R}^3$ is called the *procedure line* (Saari [1995]); the point s^{th} of the way along this line (from the plurality to the antiplurality tally) is the $\mathbf{w}_s$ tally for $\mathbf{p}$.

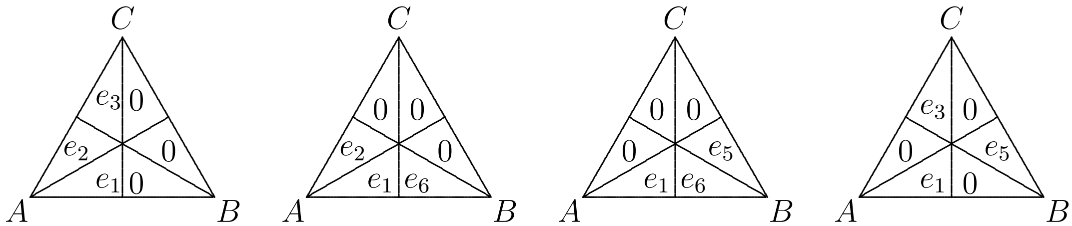
2.2.2 The "essential profile"

Because the $P(X,Y)$ values in the first Fig. 2.1c triangle agree with those of the initial example, any profile supporting the initial example and this Fig. 2.1c profile belong to the same $\sim_{BE}$ equivalence class. Even stronger, it will follow from the next theorem that the Fig. 2.1c profile is the *essential profile*, denoted by $\mathbf{p}_{ess}$, for the original choice. What makes this profile "essential" is that, as developed below, all possible profiles that support these $P(X,Y)$ values build upon $\mathbf{p}_{ess}$; i.e., the essential profile characterizes all profiles in its $\sim_{BE}$ equivalence class.

Definition 2.5. *For specified $P(A,B), P(A,C), P(B,C)$ values, the* essential profile *is the profile with the smallest number of voters that have the specified $P(X,Y)$ values.*

As it will become clear, each profile in a $\sim_{BE}$ equivalence class includes the class's essential profile as a component. For strongly non-cyclic settings, a way to derive $\mathbf{p}_{ess}$ is to take the

Figure 2.2: The four essential profiles



a. $P(A,B)$ dominates **b.** $P(A,C)$ dominates **c.** $P(B,C)$ dominates **d.** Cyclic

largest $P(X,Y) > 0$ value, and subtract Y's tally from each candidate's tally in each pairwise comparison. Illustrating with the initial example, while the original A:B tallies were 70:30, the reduced tallies are A:B by 40:0, B:C by 25:15, and A:C by 30:10. It follows from Thm. 2.5 (below) that these reduced tallies uniquely define the essential profile.

Theorem 2.5 identifies all essential profiles (four of them) and the corresponding number of $\mathbf{p}_{ess}$ voters. In defining an essential profile, the N_j value is given by e_j. But an essential profile has the smallest number of voters, so certain N_j terms must be expected to equal zero. This is the case; for each $\mathbf{p}_{ess}$, Thm. 2.5 specifies which N_j values are set equal to zero. If $N_j = e_j$ need not equal zero, the e_j value is determined by its Fig. 2.1b ranking; i.e. if the j^{th} ranking is $X \succ Y \succ Z$, then:

$$e_j = \frac{1}{2}[P(X,Y) + P(Y,Z)]. \tag{2.13}$$

Illustrating with e_1 that represents $A \succ B \succ C$, it follows from Eq. 2.13 that $e_1 = \frac{1}{2}[P(A,B) + P(B,C)]$. Similarly, $e_5 = \frac{1}{2}[P(B,C) + P(C,A)] = \frac{1}{2}[P(B,C) - P(A,C)]$. A useful relationship when computing outcomes with Eq. 2.13 is that $e_{j+3} = -e_j$, $j = 1,2,3$. (So if a computation involves $e_1 - e_6$, it could be computed as $e_1 + e_3$ with Eq. 2.13 values.)

Figure 2.2 represents the four essential profiles; the first three correspond to strongly non-cyclic settings. The last one, Fig. 2.2d, represents where the strongly non-cyclic condition is violated. Recall from the comments following Def. 2.4 that Fig. 2.2d includes cyclic and those non-cyclic outcomes that fail to qualify as being "strongly non-cyclic."

Theorem 2.5. *A specified set of $P(X,Y)$ values defines a unique essential profile: The values of $N_j = e_j$ terms that need* not *equal zero are given by Eq. 2.13. With the standard assumptions that $P(A,B) \geq 0$, $P(B,C) \geq 0$, there are four essential profiles.*

1. *With the strongly non-cyclic condition, the essential profile has one of the Fig. 2.2a, b, c forms where the choice is defined by the largest $P(X,Y) > 0$ value; call it the XY essential profile. The XY $\mathbf{p}_{ess}$ has $P(X,Y)$ voters. The three essential profiles are:*
   - *(a) If $P(A,B)$ is the largest value (Fig. 2.2a), then the AB essential profile has $N_4 = n_5 = n_6 = 0$.*
   - *(b) If $P(A,C)$ is the largest value (Fig. 2.2b), then the AC essential profile has $N_3 = n_4 = n_5 = 0$.*
   - *(c) if $P(B,C)$ is the largest value (Fig. 2.2c), then the BC essential profile has $N_2 = n_3 = n_4 = 0$.*
2. *If the $P(X,Y)$ values do not satisfy the strongly non-cyclic condition, the "cyclic essential profile" is given by Fig. 2.2d where $N_2 = n_4 = n_6 = 0$. The number of $\mathbf{p}_{ess}$ voters is $\{P(A,B) + P(B,C) - P(A,C)\}$.*

A direct computation using Thm. 2.5 with $P(X,Y)$ values from the introductory example proves that the first Fig. 2.1c triangle is its essential profile. To illustrate a different aspect of Thm. 2.5, if three arbitrarily selected integer values with the same parity are specified as tentative $P(X,Y)$ values, it need not be clear whether there is a supporting profile. But an immediate corollary of Thm. 2.5 is that such a profile always exists.

Corollary 1. *For any three integers with the same parity, I_1, I_2, I_3, there exists a profile supporting the values $P(A,B) = I_1$, $P(B,C) = I_2$, $P(A,C) = I_3$*

If I_1 and/or I_2 are negative, changing the candidates' names converts everything to our setting. (For instance, $P(A, B) = -7, P(B, C) = -11$ is equivalent to $P(C, B) = 11$ and $P(B, A) = 7$, so rename C as A^*, B as B^* and A as C^*; the A^*, B^*, C^* names satisfy the required $P(A^*, B^*), P(B^*, C^*) \geq 0$.) Thus, assume that $I_1, I_2 \geq 0$. The I_k values both identify the appropriate essential profile and define the e_j's (Eq. 2.13).

If, for instance, $P(A, B) = 10$, $P(B, C) = 6$, and $P(A, C) = 8$, the AB essential profile is given by $e_1 = \frac{1}{2}[P(A, B) + P(B, C)] = 8$, $e_2 = \frac{1}{2}[P(A, C) + P(C, B)] = \frac{1}{2}[8 - 6] = 1$, $e_3 = \frac{1}{2}[P(C, A) + P(A, B)] = \frac{1}{2}[-8 + 10] = 1$, and $N_4 = n_5 = n_6 = 0$. This $\mathbf{p}_{ess}$ has ten voters; all other supporting profiles have an even number of voters where $N > 10$.

2.3 Positional voting outcomes

All possible supporting profiles for given $P(X, Y)$ values are found by adding to $\mathbf{p}_{ess}$ profile components that never affect $P(X, Y)$ values; all ways this can be done identify all profiles in a given $\sim_{BE}$ equivalence class. With the original $N = 100$ voter example, for instance, its AB essential profile has 40 voters, so 60 voters must be added in appropriate ways. The different choices of doing this is what creates the different positional outcomes.

One approach to add terms without affecting $P(X, Y)$ values is to use "reversal pairs;" e.g., a voter preferring $A \succ B \succ C$ has a companion who prefers the reversed $C \succ B \succ A$. As a reversal pair defines a tied majority vote for all pairs of candidates, it does not change $P(X, Y)$ values. This is illustrated in the second Fig. 2.1c triangle where α voters have $A \succ B \succ C$ preferences and another α voters have the reversed $C \succ B \succ A$ preferences. The other reversal pairs are indicated by the β and γ terms in this triangle. Notice, each candidate in each paired comparison receives $\alpha + \beta + \gamma$ votes to create complete ties.

What is not obvious is that, as proved in Saari [1999, 2008], *adding reversal pairs to a*

$\mathbf{p}_{ess}$ *is the only way to preserve all* $P(X,Y)$ *values.*[2] This result is what makes it possible to identify all possible supporting profiles; just add reversal pairs to $\mathbf{p}_{ess}$. Conversely, to compute the $\mathbf{p}_{ess}$ for a specified profile, *remove* as many reversal pairs as possible. Illustrating with the 40 voter profile $(10,8,2,6,11,3)$, removing pairs $(6,0,0,6,0,0)$, $(0,8,0,0,8,0)$, and $(0,0,2,0,0,2)$ creates the eight-voter BC essential profile $(4,0,0,0,3,1)$.

The reason "pairs" can be added to $\mathbf{p}_{ess}$ to create N-voter profiles is that Eq. 2.2 ensures that N and each $P(X,Y)$ have the same parity. Thus (Thm. 2.5) for any admissible N, the number of voters not in $\mathbf{p}_{ess}$ is an even integer $2q$, which allows q pairs to be added. In particular, it follows from Thm. 2.5 that with the strongly non-cyclic condition and if $P(X,Y)$ is the maximum pairwise victory, then

$$q = \frac{1}{2}[n - P(X,Y)] \tag{2.14}$$

pairs are added. If the strongly non-cyclic condition is not satisfied, the number becomes

$$q = \frac{1}{2}[n - (P(A,B) + P(B,C) + P(C,A))]. \tag{2.15}$$

In all cases, the reversal pairs (see the second Fig. 2.1c triangle) must satisfy the equality

$$\alpha + \beta + \gamma = q. \tag{2.16}$$

2.3.1 Finding new results

To use Thm. 2.5 to discover and prove new results, first determine which Fig. 2.2 choice applies. Next, add the α, β, γ terms (as in the second Fig. 2.1c triangle) and then just

[2]This statement is not true for four or more alternatives. Thus our results do not extend in a simple way to settings with more than three alternatives.

compute and compare tallies. To illustrate by developing an explanation why B can never be a plurality winner for the initial example, notice that answers depend on the value of q (the number of reversal pairs to be added to the essential profile). A large q value provides a rich assortment of positional outcomes. But large q values correspond to small $P(X, Y)$ values, which require more competitive, closer pairwise election outcomes. So, expect a wealth of positional outcomes to accompany competitive paired comparison elections. A listing of what can happen with plurality winners follows:

Corollary 2. *With the strongly non-cyclic condition, if $P(X, Y)$ is the largest paired victory, a necessary and sufficient condition to have at most two different plurality winners is*

$$3P(X, Y) > n - 4. \tag{2.17}$$

If this largest $P(X, Y)$ satisfies $3P(X, Y) \leq n - 4$, then, for each of the three alternatives, a supporting profile exists where that alternative is the plurality winner.

If the strongly non-cyclic condition is not satisfied, then a necessary and sufficient condition that A is the sole plurality winner is

$$\min(2P(A, B) + P(B, C),\ 2P(B, C) + P(A, B)) > n. \tag{2.18}$$

Furthermore, a necessary and sufficient condition for there to be at least one profile where alternative X is the sole plurality winner is

$$n - 4 \geq 3P(Y, X) \tag{2.19}$$

where Y is the alternative immediately preceding X in the listing A, B, C, A.

Equations 2.18 and 2.19 hold whether there is, or is not, a cycle. The reason B cannot be a plurality winner with the initial example follows from Eq. 2.17; its left-hand side is 120 while the right is 96. These pairwise outcomes, then, allow at most two different plurality winners. But it already has been established that the Condorcet winner and loser, A and C, can be plurality winners, so B cannot.

Corollary 2 identifies interesting properties about settings with small number of voters. With $N = 8$, for instance, Eq. 2.17 is satisfied should any candidate win a majority election. (As N is an even integer, a victory requires $P(X, Y) \geq 2$.) So, with the strongly non-cyclic condition and no more than eight voters, some alternative never is the plurality winner.

With more voters, the situation changes. To ensure for $N = 100$ that, for each candidate, a profile can be constructed where that candidate is the plurality winner, it follows from Eq. 2.17 that this happens if each $P(X, Y) \leq 32$. In other words, each candidate is the plurality winner with some supporting profile if no candidate receives more than 66 of the 100 votes in paired comparisons. Even landslide outcomes, such as where A beats B by 62:38, A beats C by 66:34, and B beats C by 64:36, admit enormous flexibility in the associated plurality winners; with this example, for each candidate there is a supporting profile where she is the plurality winner.

Intuition and proofs of parts of Thm. 2.1 and Cor. 2.

To illustrate how to use Thm. 2.5, Cor. 2 is proved for the strongly non-cyclic case where $P(A, B)$ has the largest victory (as with the initial example). The structure builds upon the AB essential profile, so the plurality tallies for A, B, and C are, respectively,

$$e_1 + e_2 + \alpha + \beta, \quad \beta + \gamma, \quad e_3 + \alpha + \gamma. \tag{2.20}$$

For only A to be the plurality winner, A must always have the largest tally; that is, it must always be that $e_1 + e_2 + \alpha + \beta > \beta + \gamma$ (for A to always beat B) and $e_1 + e_2 + \alpha + \beta > e_3 + \alpha + \gamma$

(for A to always beat C). These inequalities reduce, respectively, to $e_1 + e_2 > \gamma - \alpha$ and $P(A,C) = e_1 + e_2 - e_3 > \gamma - \beta$.

The worse case scenario threatening A's status is where $\alpha = \beta = 0$ and $\gamma = q$; these inequalities are satisfied if $P(A,C) > \gamma = q = \frac{1}{2}[n - P(A,B)]$. Collecting terms leads to $P(A,B) + 2P(A,C) > n$, which is Eq. 2.4 in Thm. 2.1. With equality, this $\gamma = q$ value creates a profile with an $A \sim C$ tie; if the inequality is reversed, then a profile exists where C is the plurality winner.

Again with the AB essential profile, to determine what it takes for *each candidate* to be a plurality winner with some supporting profile, because (as just shown) it is easier with an AB essential profile for C to be a plurality winner, just find conditions where B can be the plurality winner. Using Eq. 2.20, this requires finding a profile where in a plurality election B beats both A and C. Computing these tallies leads, respectively, to the inequalities $\beta + \gamma > e_1 + e_2 + \alpha + \beta$, $\beta + \gamma > e_3 + \alpha + \gamma$, or $\gamma > e_1 + e_2 + \alpha$ and $\beta > e_3 + \alpha$. Thus, minimal conditions for such a profile are where $\alpha = 0, \gamma \geq \gamma_{min} = e_1 + e_2 + 1$, and $\beta \geq \beta_{min} = e_3 + 1$. Such values exist if and only if Eq. 2.16 is satisfied, or if:

$$\beta_{min} + \gamma_{min} = (e_3 + 1) + (e_1 + e_2 + 1) = P(A,B) + 2 \leq q = \frac{1}{2}[n - P(A,B)] \qquad (2.21)$$

Collecting terms leads to the $3P(A,B) \leq n - 4$ condition. This inequality establishes necessary and sufficient conditions for each of the three candidates to be the plurality winner with some supporting profile, so it is equivalent to Eq. 2.17.

Illustrating with $P(A,B) = 6$, $P(A,C) = P(B,C) = 4$ (so $\mathbf{p}_{ess}$ has the Fig. 2.2a form), it follows that each candidate can be a plurality winner with some profile as long as $3P(A,B) = 18 \leq n - 4$. Thus, with any even number of voters where $N \geq 22$, these pairwise outcomes allow such profiles to be constructed. For large N values, then, expect this behavior to occur if the winner of each pairwise election receives less than two-thirds of the vote. Stated in a different manner, to avoid this behavior, some candidate must have an exceptionally strong pairwise victory. All remaining assertions in Thms. 2.1 - 4 and Cor. 2 are proved in this same manner by applying elementary algebra to the admissible tallies.

2.3.2 Results for other positional rules

With this approach, Thms. 2.1 - 4 can be extended to all $\mathbf{w}_s$ rules. To see how to do this, the strongly non-cyclic assumption makes $\mathbf{p}_{ess}$ one of Fig. 2.2 a, b, or c; e.g., if $P(A, B)$ has the largest value, the AB $\mathbf{p}_{ess}$ is given by Fig. 2.2a. To Fig. 2.2a, add the α, β, γ values and compute the $\mathbf{w}_s$ tallies. To ensure it always is true that $A \succ C$ with $\mathbf{w}_s$, for instance, A's tally must always be larger so that:

$$(e_1 + e_2 + \alpha + \beta) + s(e_3 + 2\gamma) > (e_3 + \alpha + \gamma) + s(e_2 + 2\beta) \tag{2.22}$$

Collecting terms leads to:

$$e_1 + (1 - s)(e_2 - e_3) > (1 - 2s)\gamma + (2s - 1)\beta. \tag{2.23}$$

The $(1 - 2s)$ term (on the right-hand side with reversal pairs) differs in sign depending on whether $s > \frac{1}{2}$ or $s < \frac{1}{2}$. To make it difficult for A by enhancing C's tally, let $\beta = 0$ and $\gamma = q$ for $s < \frac{1}{2}$, and let $\beta = q$ and $\gamma = 0$ for $s > \frac{1}{2}$. This means that the right-hand side of Eq. 2.23 becomes $|1 - 2s|q$. Using the Eqs. 2.13 and 2.14 values leads to the assertion that, in this setting, A always beats C in a $\mathbf{w}_s$ election if and only if

$$(s + |1 - 2s|)P(A, B) + sP(B, C) + 2(1 - s)P(A, C) > |1 - 2s|n. \tag{2.24}$$

Similarly, by writing down the tallies and collecting terms, it follows that A always beats B in a $\mathbf{w}_s$ election if and only if $(1-s)e_1 + e_2 + se_3 > |1-2s|q$, or

$$(1 + |1-2s|)P(A,B) + (1-s)P(A,C) > |1-2s|n + sP(B,C). \tag{2.25}$$

For A to always be the sole winner, both inequalities must be satisfied. The two extremes of $s = 0, 1$, capture and prove Thms. 2.1 and 3 statements.

It is interesting how these two equations capture the transition from Eq. 2.4 for $s = 0$ to Eq. 2.8 for $s = 1$. Here, Eq. 2.24 is the more demanding for $s = 0$, while Eq. 2.25 is the more demanding for $s = 1$. It also follows from this computation that if either Eq. 2.24 or 2.25 is an equality, the above choices ($\beta = q$ or $\gamma = q$) create a profile with A being tied with the appropriate candidate; if the inequality is reversed, the construction shows how to select reversal terms to create profiles where the appropriate candidate is the $\mathbf{w}_s$ winner.

An interesting Borda Count feature is how the $s = \frac{1}{2}$ value forces Eq. 2.23 to drop the right-hand side, which consists of reversal terms. (This also happens with Eq. 2.25.) This demonstrates the known fact (e.g., Saari [1999, 2000] and [Saari, 2008, Chap. 4]) that the Borda Count is the only positional method never affected by reversal terms. These equations also illustrate why conditions for the Borda Count are more relaxed and do not involve the N value. The Eq. 2.23 condition for A to be the Borda winner is $P(A,B) + P(B,C) + 2P(A,C) > 0$, which, unless there is a complete tie, must always be satisfied. This statement is a special case of the following (also see Thm. 2.4 and Saari [1999]):

Corollary 3. *With the strongly non-cyclic condition and an alternative Y where the largest pairwise victory is A over Y, then A is both the Borda and Condorcet winner. .*

The following theorem, which generalizes Thms. 2.1 - 4 to all $\mathbf{w}_s$ rules, is proved in the same manner as Eqs. 2.24, 2.25. Each strongly non-cyclic essential profile defines a pair of conditions. The minor differences in the expressions for the pairs reflect subtle differences in the essential profiles.

Theorem 2.6. *With the assumption that the strongly non-cyclic condition holds:*

1. *If $P(A,B)$ is the strongest pairwise victory, then*

   (a) *A always beats B in a $\mathbf{w}_s$ election if and only if*

$$(1+|1-2s|)P(A,B)+(1-s)P(A,C) > |1-2s|n + sP(B,C) \tag{2.26}$$

   (b) *and A always beats C in a $\mathbf{w}_s$ election if and only if*

$$(s+|1-2s|)P(A,B)+sP(B,C)+2(1-s)P(A,C) > |1-2s|n. \tag{2.27}$$

2. *If $P(A,C)$ is the strongest pairwise victory, then*

   (a) *A always beats B in in a $\mathbf{w}_s$ election if and only if*

$$2(1-s)P(A,B)+(s+|1-2s|)P(A,C) > |1-2s|n + sP(B,C), \tag{2.28}$$

   (b) *and A always beats C in a $\mathbf{w}_s$ election if and only if*

$$(1-s)P(A,B)+sP(B,C)+(1+|1-2s|)P(A,C) > |1-2s|n. \tag{2.29}$$

3. *If $P(B,C)$ is the unique strongest pairwise victory, then*

   (a) *A always beats B in a $\mathbf{w}_s$ election if and only if*

$$2(1-s)P(A,B)+sP(A,C) > |1-2s|n + (s-|1-2s|)P(B,C), \tag{2.30}$$

(b) and A always beats C if and only if

$$(1-s)P(A,B)+(1-s+|1-2s|)P(B,C)+2sP(A,C)>|1-2s|n. \quad (2.31)$$

For the Condorcet winner A to always be the $\mathbf{w}_s$ winner, both (a,b) conditions in a pair must be satisfied. Notice that, with the possible exception of Eq. 2.30, the Borda Count ($s=\frac{1}{2}$) satisfies all of these conditions, so these favorable conclusions always hold for Borda's method. For $s=\frac{1}{2}$, Eq. 2.30 reduces to Eq. 2.9, which guarantees that A is the Borda winner. Incidentally, the "uniqueness condition" prior to Eq. 2.30 is imposed only to ensure that the Fig. 2.2c essential profile, and only this $\mathbf{p}_{ess}$, is the relevant one. Without uniqueness, one of the other Thm. 2.6 conditions would apply.

As an example with $P(A,B)=30, P(B,C)=10, P(A,C)=20$, and $N=100$, Eq. 2.26 is not satisfied for the plurality vote ($s=0$) because $2P(A,B)+P(A,C)=80$ is smaller than $N=100$, nor for the antiplurality vote ($s=1$) because 60 is smaller than 110. Thus these $P(X,Y)$ values admit supporting profiles where B is plurality ranked over A and supporting profiles where B is antiplurality ranked over A. But as noted, Eq. 2.26 is satisfied by the Borda Count ($s=\frac{1}{2}$), so all supporting profiles have A Borda ranked above B. The fact Eq. 2.26 is satisfied for one positional rule (the Borda Count) motivates the goal of finding all s values for which A always beats B. With these $P(X,Y)$ values and $0\leq s\leq\frac{1}{2}$, Eq. 2.26 becomes $80(1-s)>(1-2s)100+10s$, so Eq. 2.26 is satisfied if $\frac{2}{11}<s\leq\frac{1}{2}$. With $\frac{1}{2}\leq s\leq 1$, Eq. 2.26 becomes $60s+20(1-s)>100(2s-1)+10s$, or $\frac{12}{17}>s$. Thus, for $\frac{2}{11}<s<\frac{12}{17}$ and these $P(X,Y)$ values, A always is $\mathbf{w}_s$ ranked above B.

This example suggests that the positional $\mathbf{w}_s$ rules that satisfy certain Thm. 2.6 conditions can cluster around the Borda Count. The following corollary asserts that this clustering effect holds for most of the Thm. 2.6 conditions.

Corollary 4. *With the possible exception of Eq. 2.30, if the $P(X,Y)$ values satisfy a par-*

ticular Thm. 2.6 inequality for s_1 in $0 \leq s_1 \leq \frac{1}{2}$ and for s_2 in $\frac{1}{2} \leq s_2 \leq 1$, then they satisfies the inequality for all s in $s_1 \leq s \leq s_2$. So, if a condition holds for the plurality vote, it must hold for at least all s in $0 \leq s \leq \frac{1}{2}$. If it holds for the plurality and antiplurality rules, it holds for all positional rules. These statements extend to Eq. 2.30 if Eq. 2.9 holds.

Incidentally, Eq. 2.30 (which extends Eq. 2.9) proves that if $P(B, C)$ has the largest value in a strongly non-cyclic setting, then A cannot be the sole antiplurality winner (which proves the last statement of Thm. 3). This is because Eq. 2.30 would require the impossible $P(A, C) > n$; i.e., A would need more than all of the votes in an $\{A, C\}$ election.

The next result specifies what happens without the strongly non-cyclic condition.

Theorem 2.7. *If $P(A, C) < \min(P(A, B), P(B, C))$ (the outcome is not strongly non-cyclic), then*

1. *A always is $\mathbf{w}_s$ ranked above B if and only if*

$$(1+|1-2s|)P(A,B)+(-s+|1-2s|)P(B,C)+(s-1+|1-2s|)P(C,A) > |1-2s|n. \quad (2.32)$$

2. *and A always is $\mathbf{w}_s$ ranked above C if and only if*

$$(s+|1-2s|)P(A,B)+(1-s+|1-2s|)P(B,C)+(-1+|1-2s|)P(C,A) > |1-2s|n. \quad (2.33)$$

Notice the added burden (from Eqs. 2.26, 2.28, 2.30 and Eq. 2.32 when there is a Condorcet winner) for the Condorcet winner A to always $\mathbf{w}_s$ beat the middle pairwise ranked B; A's pairwise victories must be sufficiently dominant to overcome the $P(B, C)$ terms that make the inequalities more stringent. But, the Eq. 2.32 condition for A beating B need not ensure that with a Condorcet winner (remember, if Eq. 2.3 is not satisfied, the paired outcomes could define a cycle, or a Condorcet winner), she is the $\mathbf{w}_s$ winner; e.g., for smaller values of s, there may be profiles where C beats A. This reflects the fact that with each pair of equations from Thms. 2.6, 2.7, different s values make one inequality more difficult to meet than the other. Illustrating with the AB setting of Eqs. 2.26, 2.27, Eq. 2.27 is the more difficult to meet for $s = 0$ while Eq. 2.26 is more demanding for $s = 1$.

Indeed, the first of two variables in play is the choice of the positional $\mathbf{w}_s$ voting rule. It follows immediately from Thms. 2.6, 2.7 that the larger the $|s - \frac{1}{2}|$ value, the more difficult it becomes to satisfy the conditions. Stated in another manner, the closer a $\mathbf{w}_s$ rule resembles the Borda Count, the easier it is for the $\mathbf{w}_s$ winner to be the Condorcet winner (Cor. 4); conversely, the more removed a positional rule is from the Borda Count, the more freedom there is to admit profiles forcing the $\mathbf{w}_s$ and Condorcet winners to differ.

This added freedom to create conflicts is captured by the second variable; this is the q term (from Eqs. 2.14, 2.15) that determines the number of reversal pairs (Eq. 2.16) to add to an essential profile: The smaller the number of reversal pairs that can be added, the smaller the variance in the possible positional outcomes. But small q values require large $P(X, Y)$ values, so unless the pairwise victories are very decisive, expect differences between the positional and Condorcet winners.

2.3.3 Smaller victories

If sizable pairwise victories are required to have consistency between the $\mathbf{w}_s$ and Condorcet winners, it is reasonable to wonder what happens with smaller, more common $P(X, Y)$ values. (Many other new results can be discovered in the following manner, so our emphasis is to show how to find them.) The question we explore whether there exists *a single profile* that allows *each candidate* to win with an appropriate $\mathbf{w}_s$ rule.

To illustrate with $N = 100$ where Condorcet winner A beats B by 55:45; B beats C by 58:42, and A beats C by 60:40, as $P(A, B) = 10, P(A, C) = 20, P(B, C) = 16$, all supporting profiles are created by adding $q = 40$ reversal pairs to the AC essential profile. The objective is to determine whether there is a profile where *each candidate* is the winner with an appropriate $\mathbf{w}_s$ rule. As computed next, it is *not* possible, but only barely.

The AC essential profile's positive entries are

- $e_1 = \frac{1}{2}[10 + 16] = 13$
- $e_2 = \frac{1}{2}[20 - 16] = 2$
- $e_6 = \frac{1}{2}[-10 + 20] = 5.$

As A is the Borda winner (Cor. 3), our objective requires some non-Borda positional method to elect B while another elects C. It follows from the procedure line (Eq. 2.12) that if this can be done, the easiest way to do so is with the plurality and antiplurality rules. The most undemanding approach is to have A as the Borda winner, B the plurality winner (Thm. 2.1, part 3), and C the antiplurality winner.

By substituting the e_j and α, β, γ values into Fig. 2.2b and computing the plurality tallies, it follows that a necessary and sufficient condition for B to be the plurality winner is $\gamma > 10 + \alpha$. Setting the antiplurality tallies so that C beats A, it follows after collecting terms that this occurs if and only if $\beta > 18 + \gamma$. The minimal values satisfying these expressions are $\alpha_{min} = 0, \gamma_{min} = 11$, and $\beta_{min} = 30$, which must satisfy (Eq. 2.16) $\alpha_{min} + \beta_{min} + \gamma_{min} \leq q = 40$. But the sum is 41, which barely violates Eq. 2.16, so no profile has this property. However, a slightly tighter $\{A, C\}$ election, say 59:41 instead of 60:40, has the larger $q = 41$ value, so these computations show how to construct such a profile.

Using this kind of analysis, the following result is proved. This theorem indicates how easy it is to have a variety of positional election outcomes.

Theorem 2.8. *With the strongly non-cyclic conditions and an alternative Y where $P(A, Y)$*

is the maximum paired victory, a necessary and sufficient condition for a profile to exist where the plurality, Borda, and antiplurality winners all differ is

$$2P(A,Y) + 4P(A,X) + P(X,Y) \leq n - 6. \tag{2.34}$$

Examples, then, never exist for $N \leq 7$. To explore what Thm. 2.8 allows, Eq. 2.34 is satisfied with $N = 100$ voters even if each paired election winner receives around 56 votes. (For instance, if $P(A,C) = 14$ (so A receives 57 votes) and $P(A,B) = P(B,C) = 12$, then $Y = C$, $X = B$ and $2(14) + 4(12) + 12 = 88 < 100 - 6 = 94$.) As such, rather than being an unusual event, Eq. 2.34, which allows almost anything to happen, is satisfied with even decisive paired election outcomes. Also notice how if the strongest pairwise outcome is $P(A,B)$, then Eq. 2.34 is easier to satisfy because $P(X,Y) = P(C,B) = -P(B,C) \leq 0$. Related results using the other XY essential profiles are proved in the same manner.

2.3.4 General statement

The above results connect paired tallies to positional election outcomes. The following more general conclusions (Thm. 2.9) completely specify all possible $\mathbf{w}_s$ rankings that can accompany specified $P(X,Y)$ values.

As reversal pairs are central to what follows, let $\mathbf{r}_{X,Y} = \{X \succ Z \succ Y, Y \succ Z \succ X\}$; i.e., the subscript names the pair's two top-ranked candidates. For example, $\mathbf{r}_{B,C} = \{B \succ A \succ C,\ C \succ A \succ B\}$. The number of pairs of voters with $\mathbf{r}_{A,C}, \mathbf{r}_{A,B}, \mathbf{r}_{B,C}$ preferences are given, respectively, by α, β, γ where $\alpha + \beta + \gamma = q$.

Theorem 2.9. *a. The* $\mathbf{w}_s$ *tally for* $\mathbf{r}_{X,Y}$ *assigns one point to* X *and to* Y*, and* $2s$ *points to*

Z. Thus

$$V_s(\mathbf{r}_{A,C}) = (1, 2s, 1), \quad V_s(\mathbf{r}_{A,B}) = (1, 1, 2s), \quad V_s(\mathbf{r}_{B,C}) = (2s, 1, 1). \tag{2.35}$$

b. For N voters, all possible supporting profiles for specified $P(X, Y)$ values are given by

$$\{\mathbf{p}_{ess} + \alpha\mathbf{r}_{A,C} + \beta\mathbf{r}_{A,B} + \gamma\mathbf{r}_{B,C} \,|\, \alpha, \beta, \gamma \textit{ are non-negative integers; } \alpha + \beta + \gamma = q\}. \tag{2.36}$$

c. The set of all associated $\mathbf{w}_s$ tallies is

$$V_s(\mathbf{p}_{ess}) + \alpha(1, 2s, 1) + \beta(1, 1, 2s) + \gamma(2s, 1, 1), \quad \alpha + \beta + \gamma = q. \tag{2.37}$$

According to Thm. 2.9, all possible $\mathbf{w}_s$ tallies for the introductory example are given by

$$(30 + 10s, 25s, 10 + 5s) + (\alpha + \beta + 2s\gamma, \beta + \gamma + 2s\alpha, \alpha + \gamma + 2s\beta);\ \alpha + \beta + \gamma = 30. \tag{2.38}$$

All corresponding election rankings are computed in Sect. 2.4.2 and Eq. 2.40.

For an $N = 20$ voter cyclic outcome example where $P(A, B) = 4, P(B, C) = 10, P(A, C) = -2$, the essential profile (Thm. 2.5) is $(7, 0, 3, 0, 6, 0)$. Because $N = 20$ and the essential profile has 16 voters, so $2q = 4$ and only two $\mathbf{r}_{X,Y}$ pairs can be added. Thus all six supporting profiles are:

- $(9, 0, 3, 2, 6, 0)$
- $(7, 2, 3, 0, 8, 0)$
- $(7, 0, 5, 0, 6, 2)$
- $(8, 1, 3, 1, 7, 0)$

- $(8,0,4,1,6,1)$
- $(7,1,4,0,7,1)$.

As $V((7,0,3,0,6,0)) = (7+3s, 6+7s, 3+6s)$, the $\alpha = 2$ tally is $(7+3s, 6+7s, 3+6s) + 2(1, 2s, 1) = (9+3s, 6+11s, 5+6s)$.

2.4 The associated positional rankings

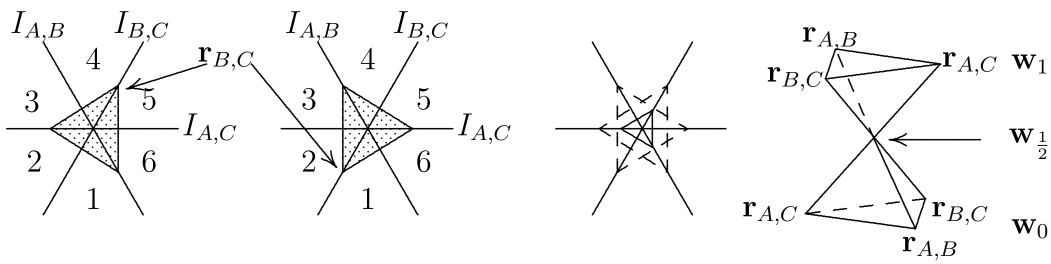
Although Thm. 2.9 completes the problem by identifying all profiles and $\mathbf{w}_s$ tallies associated with specified $P(X,Y)$ values, the algebraic computations required to find new results and determine positional rankings can be messy. In this section, a geometric tool is developed to simplify the analysis of the $\mathbf{w}_s$ rankings. This tool plays a central role when computing probabilities of paradoxical behavior.

2.4.1 Geometry

To simplify the geometry of the $V_s(\mathbf{p})$ points in $\mathbb{R}^3$, let $I_{\{X,Y\}}$ be the indifference plane of points where X and Y are tied; e.g., $I_{\{A,B\}} = \{(x,y,z) \,|\, x = y\}$. Figure 2.3 represents how the three indifferences planes intersect a plane given by $x+y+z=c$, where constant c is described below. The numbers (from Fig. 2.1a) between indifference lines identify the ranking of a tally that lands in the region. The three indifference planes intersect along the main diagonal $t(1,1,1)$ of $\mathbb{R}^3$, where all alternatives are tied. This diagonal (a point in each of the above triangles) is the line of *complete indifference.*

For a specified q, the set of all $\mathbf{w}_s$ tallies, $V_s(\alpha\mathbf{r}_{A,C} + \beta\mathbf{r}_{A,B} + \gamma\mathbf{r}_{B,C})$, is in a convex hull defined by the three extreme tallies $\{V_s(q\mathbf{r}_{A,C}), V_s(q\mathbf{r}_{A,B}), V_s(q\mathbf{r}_{B,C})\}$. According to Thm. 2.9 and Eq. 2.35, these $V_s(q\mathbf{r}_{X,Y})$ points are the vertices of an equilateral triangle in the

Figure 2.3: Positional outcomes



a. $\mathbf{w}_0$ outcomes **b.** $\mathbf{w}_1$ outcomes **c.** $\mathbf{w}_s$ **d.** Pyramidal cone, $0 \leq s \leq 1$

plane defined by $c = (2 + 2s)q$; the triangle's center point is on the complete indifference line. Denote this triangle by $T_s(q)$.

Figure 2.3a depicts the $T_0(q)$ triangle defined by the vertices $(q, 0, q)$, $(q, q, 0)$, $(0, q, q)$, which are the $V_0(q\mathbf{r}_{X,Y})$ plurality tallies. Similarly, Fig. 2.3b represents the antiplurality $T_1(q)$ triangle in the $c = 4q$ plane; the three vertices are $(q, 2q, q)$, $(q, q, 2q)$, $(2q, q, q)$. When applied to reversal terms, the plurality and antiplurality rankings reverse each other; e.g., the plurality ranking for $q\mathbf{r}_{X,Y}$ is $X \sim Y \succ Z$, while the antiplurality ranking is the reversed $Z \succ X \sim Y$. This ranking reversal is indicated by the arrows between Figs. 2.3a and 2.3b for the $\mathbf{r}_{B,C}$ ranking.

The diagram depicting all V_s tallies is Fig. 2.3d; the vertical direction depicts different $c = (2+2s)q$ values, $0 \leq s \leq 1$. (This direction is along the complete indifference diagonal.) Think of this figure as placing Fig. 2.3b (on the $c = 4q$ plane) above Fig. 2.3a (on the $c = 2q$ plane). Next, connect with straight lines the $V_0(q\mathbf{r}_{X,Y})$ and $V_1(q\mathbf{r}_{X,Y})$ vertices (in the plurality and antiplurality planes); that is, with the procedure lines (Eq. 2.12) defined by the reversal pairs. The pyramid is the region defined by these lines. All three procedure lines (actually, all procedure lines of reversal terms) cross at a common point on the $c = 3q$ plane, which is the $\mathbf{w}_{\frac{1}{2}}$ complete tie outcome. (This condition reflects the requirement (Eq. 2.35) that the Borda outcome ($s = \frac{1}{2}$) of reversal terms must be a complete tie.)

The $T_s(q)$ triangle is the intersection of the pyramid with the $c = q(2+2s)$ plane; the vertices are $(q, 2qs, q)$, $(q, q, 2qs), (2qs, q, q)$. Notice how $T_{\frac{1}{2}+t}(q)$ and $T_{\frac{1}{2}-t}(q)$, $0 < t \leq \frac{1}{2}$, have the same size; they differ only in orientation where one is the reversal of the other. This geometry reflects how the $\mathbf{w}_{\frac{1}{2}+t}$ and $\mathbf{w}_{\frac{1}{2}-t}$ rankings of an $\mathbf{r}_{X,Y}$ reverse each other. Also, the $T_s(q)$ triangles are embedded in the manner indicated by Fig. 2.3s with centers along the complete indifference line; the dashed lines represent the $T_0(q)$ and $T_1(q)$ triangles. As the Fig. 2.3c shaded $T_s(q)$ has the same orientation as $T_0(q)$, it represents an s value where $0 < s < \frac{1}{2}$. For $\frac{1}{2} < s < 1$ s values, the $T_s(q)$ orientation is that of $T_1(q)$.

The above figure and descriptions, Thm. 2.9, and Eq. 2.37 provide all tools needed to describe what $\mathbf{w}_s$ rankings could possibly accompany specified $P(X, Y)$ values. In particular, no matter what reversal pairs are added to $\mathbf{p}_{ess}$, the Borda ranking always agree with the $V_{\frac{1}{2}}(\mathbf{p}_{ess})$ ranking. All V_s rankings now are given by rankings in the set

$$V_s(\mathbf{p}_{ess}) + T_s(q).$$

2.4.2 The initial example

To conclude, these methods are illustrated by finding all $\mathbf{w}_s$ rankings for the initial example. They are determined by

$$V_s(\mathbf{p}) = (30 + 10s, 25s, 10 + 5s) + T_s(30), \quad 0 \leq s \leq 1. \tag{2.39}$$

So the Borda ranking ($s = \frac{1}{2}$) is $A \succ B \sim C$.

To find the admissible plurality rankings, if the translated $T_0(30)$ crosses any Fig. 2.3 ranking region, it intersects the triangle's boundary. Thus, to identify all admissible plurality rankings, just analyze what happens with the $s = 0$ extreme points $(30, 0, 10) + (30, 0, 30) =$

$(60, 0, 40), (60, 30, 10)$, and $(30, 30, 40)$; the boundary lines between these vertices give transition rankings. To illustrate, the vertices define, respectively, the rankings $A \succ C \succ B, A \succ B \succ C$, and $C \succ A \sim B$. Then our three boundary lines are:

- The line between the first two vertices, $(1-t)(60, 0, 40) + t(60, 30, 10) = (60, 30t, 40 - 30t)$. This boundary admits a tie when $30t = 40 - 30t$, or $t = \frac{2}{3}$. Thus the only new ranking is the transition ranking $A \succ B \sim C$ that lies between the first two strict rankings.

- The line between the first and third vertex, $(60 - 30t, 30t, 40)$; it admits a tie only at the $t = 1$ endpoint, so no new rankings are created.

- Finally, $(60 - 30t, 30, 10 + 30t)$, which admits a $60 - 30t = 10 + 30t$, or $t = \frac{5}{6}$ point with an $A \sim C \succ B$ ranking transitioning from the first ranking to a $C \succ A \succ B$ plurality outcome.

Thus, the only admissible plurality rankings by profiles supporting the initial $P(X, Y)$ values are $A \succ C \succ B, A \succ B \succ C, C \succ A \sim B$ and the obvious need for the transition rankings $A \succ B \sim C$ and $A \sim C \succ B$.

To analyze this example for any s, the vertices are equal to $(30 + 10s, 25s, 10 + 5s)$ plus one of the three possible permutations of $(30, 60s, 30)$, i.e.:

- $(60 + 10s, 85s, 40 + 5s)$

- $(60 + 10s, 30 + 25s, 10 + 65s)$

- $(30 + 70s, 30 + 25s, 40 + 5s)$

Again, the analysis reduces to elementary algebra. For instance, with the first vertex, $\{A, B\}, \{A, C\}, \{B, C\}$ ties require, respectively, $60 + 10s = 85s, 60 + 10s = 40 + 5s, 85s =$

Table 2.2: Possible rankings by value of s

s	Rankings	s	Rankings
$s=0,$	$A\succ C\succ B, A\succ B\succ C, C\succ A\sim B$	$0<s<\frac{2}{13},$	$A\succ C\succ B, A\succ B\succ C, C\succ A\succ B$
$s=\frac{2}{13},$	$A\succ C\succ B, A\succ B\sim C, C\sim A\succ B$	$\frac{2}{13}<s<\frac{1}{2},$	$A\succ C\succ B, A\succ B\succ C$
$s=\frac{1}{2},$	$A\succ B\sim C$	$\frac{1}{2}<s<\frac{4}{5},$	$A\succ B\succ C, A\succ C\succ B,$
$s=\frac{4}{5},$	$A\sim B\succ C, A\succ C\succ B, A\succ B\succ C$	$\frac{4}{5}<s<\frac{10}{11},$	$B\succ A\succ C, A\succ C\succ B, A\succ B\succ C$
$s=\frac{10}{11},$	$B\succ A\succ C, A\sim C\succ B, A\succ B\succ C$	$\frac{10}{11}<s\leq 1,$	$B\succ A\succ C, C\succ A\succ B, A\succ B\succ C$

(2.40)

$40+5s$), which have respective solutions $s=\frac{4}{5}, -4, \frac{1}{2}$. As only the first and third lie in the $0\leq s\leq 1$ range, the vertex has the $A\succ C\succ B$ ranking for $0\leq s<\frac{1}{2}$, the $A\succ B\succ C$ ranking for $\frac{1}{2}<s<\frac{4}{5}$, and the $B\succ A\succ C$ ranking for $\frac{4}{5}<s\leq 1$, with two obvious transition rankings involving a tie.

The second vertex defines $A\succ B\succ C$ for $0\leq s<\frac{1}{2}$, $A\succ C\succ B$ for $\frac{1}{2}<s<\frac{10}{11}$, and $C\succ A\succ B$ for $\frac{10}{11}<s\leq 1$ with transition rankings with ties at $s=\frac{1}{2},\frac{10}{11}$, while the third vertex has the plurality $C\succ A\sim B$ that becomes $C\succ A\succ B$ for $0<s<\frac{2}{13}$, continues as $A\succ C\succ B$ for $\frac{2}{13}<s<\frac{1}{2}$, and ends as $A\succ B\succ C$ for $\frac{1}{2}<s\leq 1$.

A complete analysis follows from the above with the transition values of $s=0,\frac{2}{13},\frac{1}{2},\frac{4}{5},\frac{10}{11}$. Leaving out the obvious transition rankings for each s, a complete analysis of the initial example is given by the following list of vertex rankings: Thus, for $0\leq s<\frac{2}{13}$, either A or C could be the $\mathbf{w}_s$ winner, for $\frac{2}{13}<s<\frac{4}{5}$, A must be both the Condorcet and $\mathbf{w}_s$ winner, for $\frac{4}{5}<s<\frac{10}{11}$, either A or B could be the $\mathbf{w}_s$ winner, and for $\frac{10}{11}<s\leq 1$, anyone could be the positional winner.

Chapter 3

A Geometric Approach to Voting Probability Problems

3.1 Introduction

In Chapter 2, we introduced a method of discerning the complete relationship between pairwise comparisons and possible election results for positional methods with $n = 3$ candidates. This allowed us to establish necessary and sufficient conditions for a profile to exist that meets a given set of pairwise and positional criteria - for example, when the plurality winner and the Condorcet winner can differ. These necessary and sufficient conditions tell us which pairwise outcomes are associated with which positonal results.

In this chapter, we will extend these techniques to determine how *probable* these combinations are, using natural extensions of the techniques of Chapter 2. We will compute the probabilities that certain combinations of pairwise and/or positional criteria are met; in particular, how likely it is that Condorcet winner fails to win a Borda or plurality vote, how likely it is that a Condorcet winner exists at all, and how likely it is that a Condorcet winner or loser will, under a plurality or antiplurality ("vote for two") rule, win an election.

Importantly, these new techniques are both intuitive and transparent, and they give robust, meaningful results across a wide range of probabilistic assumptions. In particular, we can compute results for large classes of probability distributions of voter preferences, rather than being restricted to one or two particular probability distributions, as is the case in the analysis of Gehrlein and Lepelley [2011], which explores the probability of a Condorcet paradox under the Impartial Culture (IC) and Impartial Anonymous Culture (IAC) assumptions.

Since the actual probability distribution of profiles (collections of preferences of voters) for $n \geq 3$ candidate elections is not well-known (and, for that matter, may differ substantially among different populations), computing results for an entire class of probability distributions at once sharply increases the scope of our ability to predict electoral anomalies. The transparent and intuitive nature of the techniques means that we can qualitatively understand how all these probabilities change as the probability distributions associated with voters' voting behavior change.

3.2 Contributions to the literature

There is interest in computing how likely voting paradoxes are in principle and in practice, though the techniques used previously have limitations. Some prior techniques also only apply to a single assumed probability distribution. For example, Gehrlein and Fishburn [1976] showed that under the assumption of Impartial Anonymous Culture (IAC), the probability of a Condorcet paradox approaches 6.25%, or $\frac{1}{16}$, as the number of voters, becomes large. However, the method of calculation used in Gehrlein and Fishburn [1976] is specific to IAC; and it is not intuitive. This pattern continues in subsequent literature, e.g., Gehrlein and Lepelley [2011], which investigates the rate at which Condorcet paradoxes occurs under

Impartial Culture (IC) and IAC, and compares these rates to a selection of empirical data.

Using our techniques, it is easy to show that large classes of probability distributions lead to the exact same conclusions with regard to the calculated probability of a Condorcet paradox. In particular, a large class of distributions that includes the distribution associated with IAC produces a 1 in 16 (6.25%) chance of a Condorcet paradox; and a large class which includes good approximations of the distribution associated with IC (for large numbers of voters), produces a probability of a Condorcet paradox equal to twice the solid angle of a tetrahedron out of a sphere (about 8.8%). It is already known (see again Gehrlein and Lepelley [2011]) for IAC and IC specifically that the limiting chance of a Condorcet paradox as $N \to \infty$ approaches these two values; the extension to a larger class of probability distributions is new.

This will be accomplished by use of the representation cube. Some of the analysis will be done within on the tally space in which the representation cube was originally introduced. We will also introduce essential octohedral coordinates for the representation cube, and show how this coordinate change is a useful tool for furthering the understanding of probability distributions over profile space and their effects on the relationships between different voting rules, and between those different voting rules and pairwise voting criteria.

Both of these classes of probability distribution in question are large (containing related families of distributions superficially very unlike those associated with IC and IAC) but are tractable using our approach. Further, within this framework, it is easy to see that nearly any probability distribution we are likely to assume *a priori* for large N will produce a theoretical estimate between the values associated with these classes, i.e., 6.25% and 8.8%. This is because those two classes of probability distribution represent natural extreme cases. Moreover, that the percentage of profiles that are *strongly* non-cyclic (as defined in Chapter 2) holds steady at 75% for both classes of distributions, suggesting this is likely to remain constant for most *a priori* choices of probability distribution.

Although of importance, the rate at which a Condorcet winner can be expected to win a Borda Count election has not been carefully investigated. In this chapter, we provide an exact calculation for the probability with 3 candidates and two different classes of probability distributions, showing that a Condorcet winner is 91.2-90.2% likely to win for a large class of probability distributions that either includes or closely approximates most distributions of theoretical interest.

We will address, though significantly less completely, the question of how likely a Condorcet winner is to win a plurality or antiplurality vote. The answer to this question is very sensitive to the choice of probability distribution even within the classes of tractable symmetric distributions, but usually not more than 89% or 64%, respectively. In the case of an antiplurality vote, 77.5% represents an upper limit for a very large class of probabilities. We will also address the issue of agreement between a plurality vote and a Borda count, noting that it is, in cases where a Condorcet winner exists, strictly less likely than either the plurality vote or Borda count agreeing with a Condorcet winner. Finally, we will also address the subject of the Condorcet loser, which is generally unlikely to win a plurality or antiplurality vote (and never wins a Borda count.)

3.3 Constructing the spaces of interest

There are three spaces used in Chapter 2 to compute all possible positional tallies for each combination of pairwise majority voting outcomes. The first space is the most fundamental and best known: Profile space, an $n!$-dimensional space where each dimension represents a separate voter preference. In this case, profile space is a subset of $\mathbb{R}^6$. This space is not utilized directly, but is instead broken up into two orthogonal component subspaces. The first is the representation cube, an object in $\mathbb{R}^3$ which represents all possible pairwise majority votes, introduced in Saari [1995]; while the second is the supporting space, lying in another

$\mathbb{R}^3$, which underlies each point in the representation cube. The supporting space is also known as the reversal component. We will mostly be concerned with the representation cube itself (a subset of $\mathbb{R}^3$), within two coordinate systems: The tally coordinates in which it was originally introduced, and also essential octohedral coordinates, in which the representation cube takes the shape of a regular octohedron.

3.3.1 The representation cube

The representation cube is an object lying in $\mathbb{R}^3$. In Saari [1995], the representation cube is given in tally space where the three axes are the pairwise majority votes margins, or tallies. (This is exhibited in Figure 3.1.)

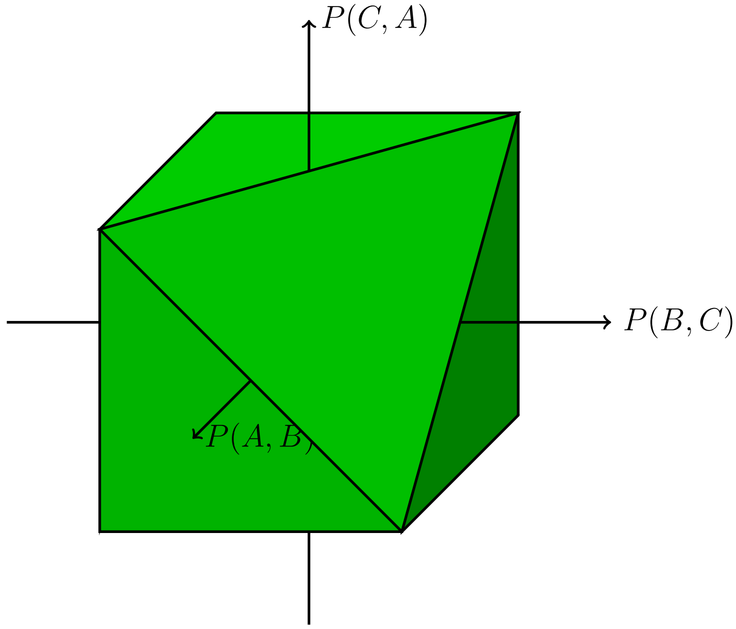
- P(A,B) - the majority vote margin of A over B.
- P(B,C) - the majority vote margin of B over C.
- P(C,A) - the majority vote margin of C over A.

To translate from $\mathbb{R}^6$ of profile space to the $\mathbb{R}^3$ of the representation cube, that is to say to go from $(p_{ABC}, p_{ACB}, p_{CAB}, p_{CBA}, p_{BCA}, p_{BAC})$ to $(P(A,B), P(B,C), P(C,A))$, we use the following linear transformation:

$$L = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & -1 \end{bmatrix} \tag{3.1}$$

For our purposes in this chapter, it is most convenient to use a *normalized* representation cube in $\mathbb{R}^3$, based on a *normalized* profile space in $\mathbb{R}^6$. A normalized profile has all compo-

Figure 3.1: Representation cube



nents adding up to one, meaning that *normalized* profile space corresponds to Δ^5, the unit 5-simplex within $\mathbb{R}^6$. We will use the natural inscription of the representation cube into a proper subset of $[-1, 1]^3$, and assume that N, the number of voters, is large. By "large," we mean large enough that error in approximating the representation cube as a continuous object are small enough to be neglected. If N itself is not entirely constant in addition to being large (e.g., overall population is fixed but turnout is variable), this makes the continuous approximation more accurate and more useful, but even if N is fixed, the approxmation tends to be close when $N \geq 100$, and is generally very good when $N \geq 1000$.

The representation cube is a cube that has had two of its vertices truncated. This is because it is the convex hull of unanimous points, and there are only six combinations of pairwise preferences which describe a transitive ordering of three candidates. Each of the six possible unanimous preference orders corresponds to a vertex, and a full cube has eight vertices. If $(1, 1, 1)$ were included, we would need all voters to hold the preference $A \succ B$, $B \succ C$, and also $C \succ A$. (One might consider relaxing the assumption of individually transitive preferences. The resulting analysis, wherein each $P(X, Y)$ ranges freely from -1 to 1 independently, is relatively straightforward, but outside of the scope of this chapter.)

3.3.2 Essential octohedral coordinates

The essential profile is introduced in Chapter 2 as the minimum sub-profile with the same pairwise majority votes ($P(X,Y)$ values). While essential profiles enjoy a bijective correspondence with points in the representation cube, tally space is not always the most convenient space in which to display essential profiles. The essential profile can be easily defined directly in terms of its three non-zero components; assigning signs to each component and assigning each component along with its reversal to the same axis (only one of which will be non-zero for the essential profile) gives us a space in which the essential profile's components can be immediately identified. This is a more natural space for understanding the essential profile, and in the normalized case, we can see that all profiles are mapped to a regular octohedron. We will demonstrate that this is, in fact, isomorphic to the standard representation cube.

Formally, the following linear transformation takes essential profiles in $\mathbb{R}^6$ (profile space) to unique points in $\mathbb{R}^3$ (the essential coordinates):

$$E = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} \tag{3.2}$$

That is to say, the three component axes are:

- $p_{ABC} - p_{CBA}$
- $p_{ACB} - p_{BCA}$
- $p_{BAC} - p_{CAB}$

As with the representation cube, we find it convenient to normalize by defining the above quantities not in terms of the number of voters (which would give us a lattice of integers from $-N$ to N) but in terms of a fraction of the population, starting with a normalized profile that is the unit simplex $\Delta^5 \subset \mathbb{R}^6$. In the essential octohedral coordinates, the representation cube is then the set of ordered triplets in the convex hull of $\{(1,0,0),(0,0,-1),(0,1,0),(0,-1,0),(0,0,1),(0,0,-1)\}$. (Only ordered triples of *rational* numbers correspond to specific voting situations with a finite number of possible voters.)

3.3.3 Relating tally space and essential octohedral coordinates

To relate these two coordinate systems, we state the following lemma:

Lemma 1. *The essential octohedral coordinates and tally coordinates describe the same underlying space, and in particular, there exists a natural isomorphism between them, T, such that with L from Eq. 3.1 and E from Eq. 3.2:*

$$TL = E^{-1}T \tag{3.3}$$

From this lemma we may infer that the representation cube can be displayed equivalently both in tally space and in essential octohedral coordinates. A simple proof of this lemma can be had by exhibiting such an isomorphism; in particular, the following linear transformation:

$$T := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \tag{3.4}$$

We can note that each vertex of the representation cube is transformed from its coordinate in tally space by $T : \mathbb{R}^3 \to \mathbb{R}^3$ to a unanimous point, i.e., standard unit vector showing a value of 1 for one profile coordinate and 0 for all other profile coordinates, in essential octohedral coordinates:

- $T(1, 1, -1) = (1, 0, 0)$
- $T(1, -1, -1) = (0, 1, 0)$
- $T(-1, 1, -1) = (0, 0, 1)$
- $T(-1, -1, 1) = (-1, 0, 0)$
- $T(-1, 1, 1) = (0, -1, 0)$
- $T(1, -1, 1) = (0, 0, -1)$

This linear transformation T and its inverse T^{-1} allow us to freely move between tally space and essential octohedral coordinates. The bijective correspondence tells us that the information contained in the representation cube is available within both coordinate systems. Any object (point, boundary, subset, or even density function) in either space has a unique corresponding object in the other; since T is linear, this object's structure is largely preserved. This also means that the maps from the full profile space to essential octohedral coordinates and to tally space share the same kernel within profile space.

The kernel of the maps from profile space to tally space or essential octohedral coordinates (the kernels of these two maps are the same) contains what we will call the *supporting space* for each essential profile. The space of all supporting spaces is, for $n = 3$ candidates and normalized profiles, adequately termed the *reversal component*, as contrasted with the *essential component* displayed in the representation cube.

3.3.4 Supporting space

All profiles leading to the same essential profile we call $binary - equivalent$. The space of all profiles which are binary-equivalent to some essential profile will be called the *supporting space* of a particular essential profile; that is to say that for each point in the representiation cube, there is a slice of profile space which corresponds to that particular point. For $n = 3$ candidates (the scope of this chapter), the supporting space is most naturally represented as a two-dimensional object lying inside of a three-dimensional space.

Consider the space spanned by a set of three linearly independent vectors:

$$\{(1, -0.5, -0.5, 1, -0.5, -0.5), (-0.5, 1, -0.5, -0.5, 1, -0.5), (1, 1, 1, 1, 1, 1)\} \tag{3.5}$$

Each of these is in the kernel of L from Eq. 3.1 and E from Eq. 3.2, and each is linearly independent from the other two. In particular, this set of vectors spans the *neutral* component of the profile and the $reversal$ component of the profile. The neutral component - the component spanned by (1,1,1,1,1,1) - is not of interest in this chapter. In particular, normalization trivializes the neutral component - for a normalized profile P, we have that:

$$P \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 1 \tag{3.6}$$

This, the meaningful components of the supporting space is the reversal component of the profile. However, while the reversal component *can* be described completely using two coordinates, it is meaningful to talk about three different reversal components (R_A, R_B, and R_C), and this is the natural space in which to describe the supporting space. We then define a map from profile space to the supporting space, ignoring the trivial component and mapping from Δ^5 in $\mathbb{R}^6$ to Δ^2 in $\mathbb{R}^3$:

$$R = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix} \tag{3.7}$$

Note that R is injective if the domain is restricted to a set of binary-equivalent profiles. It is also worth noting that in each set of binary-equivalent profiles, there is a unique profile which maximizes each of the three coordinates within the reversal component, so the set of all RP_i in a set $\{P_i\}_{i \in I}$ of binary-equivalent profiles make up the convex hull of $\{k_A + (1 - r_P)R_A, k_B + (1 - r_P)R_B, k_C + (1 - r_P)R_C\}$ for some $(1 - r_P) \in (0, 1)$ and $(k_A, k_B, k_C) \in \mathbb{R}^3$. In this case, r_P turns out to b a very useful radius-like quantity within the space of the essential octohedon or representation cube. We will define it, and then show that $(1 - r_P)$ happens to be the relevant parameter determining the size of the supporting space associated with a given essential profile.

3.3.5 Defining r_P and r_S

To fully understand the relationship between an essential profile and its supporting space, it is necessary to define a useful radius-like quantity r_P. This variable is also the key to understanding a large class of convenient probability distributions. We will also define a similar quantity r_S. Within tally space, we can define r_P in terms of pairwise comparisons:

$$r_P := \max_{X,Y,Z} \{|P(X, Y)|, |P(X, Y) + P(Y, Z) + P(Z, X)|\} \tag{3.8}$$

While in essential octohedral coordinates, we define r_P in terms of the profile components:

$$r_P := |p_{ABC} - p_{CBA}| + |p_{ACB} - p_{BCA}| + |p_{BAC} - p_{CAB}| \tag{3.9}$$

Finally, if we plot the essential sub-profile in profile space, we can write that:

$$r_P = |e_{ABC}| + |e_{ACB}| + |e_{CAB}| + |e_{CBA}| + |e_{BCA}| + |e_{BAC}| \tag{3.10}$$

The following theorem gives three useful properties of r_P:

Theorem 3.1. *1. Eqs. 3.8, 3.9, and 3.10 are equivalent definitions of* r_P*.*

2. The natural distance function defined by r_P *induces a metric on tally space and essential octohedral space, and a pseudometric on profile space, with the maximum radius from the origin being* N *(non-normalized cas) or* 1 *(normalized case).*

3. r_P *gives the quantity of voters in the essential profile.*

4. $1 - r_P$ *(normalized) or* $N - r_P$ *(non-normalized) gives the quantity of voters in the supporting space.*

That is to say, r_P is a radius-like quantity that can be defined within any of the spaces we are interested in.

We will define a second useful radius-like quantitity. To define r_S on the representation cube, the following formula suffices:

$$r_S := \frac{1}{2}\sqrt{(P(A,B) + P(B,C))^2 + (P(A,C) + P(C,B))^2 + (P(B,A) + P(A,C))^2} \tag{3.11}$$

Or, equivalently, in essential octohedral coordinates:

$$r_S := \sqrt{(p_{ABC} - p_{CBA})^2 + (p_{ACB} - p_{BCA})^2 + (p_{BAC} - p_{CAB})^2} \tag{3.12}$$

Within the original profile space, a third equivalent definition comes from the essential octohedral coordinates:

$$r_S := \sqrt{e_{ABC}^2 + e_{CBA}^2 + e_{ACB}^2 + e_{BCA})^2 + e_{BAC}^2 + e_{CAB}^2} \tag{3.13}$$

The following theorem gives some useful properties of r_S:

Theorem 3.2. *1. Eqs. 3.11, 3.12, and 3.13 are equivalent definitions of r_S.*

2. At the origin of each space, $r_S = 0$, and r_S is at most N in a non-normalized space, or 1 in the normalized version of the space.

The proof of part (1) relies on the equivalences established in the proof of the earlier theorem, namely, that:

$$e_{XYZ} = |p_{XYZ} - p_{ZYX}| = \max\left\{\frac{P(X,Y) + P(Y,Z)}{2}, 0\right\} \tag{3.14}$$

To show (2), note that r_S is exactly the Cartesian radius in essential octohedral coordinates, which indeed produces a metric. The fact that r_P and r_S are familiar metrics within essential octohedral coordinates is one of the things which makes essential octohedral coordinates a natural choice for carrying out related calculations.

3.4 Measure and probability

In general, to calculate the probability that something happens, we take a measure of the set of ways in which that can happen; the measure of the set of all things that could possibly happen; and divide the former by the latter. For any particular number of voters N and any particular number of candidates n, the natural measure to turn to would be the counting measure; that is, counting the number of points; and in the general case, probability distributions over that space are equivalent to an assignment of a weight to each particular point.

Counting measures are not practical for many real-world cases (where N can be approximately guessed before the election, but not precisely known) or the general case (where N is left undefined). Nor is it computationally tractable for specific N when N is large (particularly as n, the number of candidates, increases, though we will only consider $n = 3$ here), because there are $\binom{N}{n!-1}$ points in the set. (Note that if voters are not anonymous, a complete coding of a general N voter n candidate profile requires $N^{n!}$ points.)

For our purposes, it is adequate to approximate the set of interest - the representation cube and its supporting sets of profiles - as continuous spaces; with density functions on those spaces. The integrals of those density functions then become measures on the space of possible results; and by appropriate choice of normalization, become probability measures.

3.4.1 Density functions

We now turn to the problem of constructing an appropriate density distribution. In principle, common assumptions about density distributions are defined in terms of profile space. Two common assumptions within social choice literature are:

Impartial culture (IC) is the assumption that each *preference order* is equally likely for each voter, independently. That is to say, each one of the N voters' preferences is an independent random variable, distributed uniformly over the $n!$ possible preference orders between the n candidates. To draw a profile from a profile distributed per the IC assumption is equivalent to drawing N independent draws from a uniform distribution. This defines a very specific multinomial distribution in the $n!$-dimensional profile space. This means that a unanimous profile is exceptionally unlikely under the assumption of IC.

Impartial anonymous culture (IAC) is the assumption that every possible *profile* of the population is equally likely. This means that voters do not act independently, and that profiles are distributed uniformly in profile space. For example, in an IAC probability distribution over a population of 120 voters, we are as likely to have a unanimous profile with 120 voters preferring $A \succ B \succ C$ as 60 voters preferring $A \succ B \succ C$ and 60 preferring $C \succ B \succ A$. (This assumption was introduced in Gehrlein and Fishburn [1976]).

In both cases, the probability measure of any particular set S of possible events is defined via the integral over S of the probability density function ρ within $\mathbb{R}^6$:

$$|S| = \int_S dV = \int_S \rho dp_{ABC} dp_{ACB} dp_{CAB} dp_{CBA} dp_{BCA} dp_{BAC} \tag{3.15}$$

That is to say that a differential element of measure is given on $\mathbb{R}^6$ by:

$$\rho dp_{ABC} \wedge dp_{ACB} \wedge dp_{CAB} \wedge dp_{CBA} \wedge dp_{BCA} \wedge dp_{BAC} \tag{3.16}$$

Profile space, however, is not the natural space for solving problems of the type this chapter is concerned with. Six-dimensional structures can be dealt with mathematically, but it is less convenient, less intuitive, and does not display the same symmetries as are visible on the representation cube.

It is convenient to define density functions directly and explicitly in essential octohedral coordinates (or, equivalently, representation cube) - and in particular, to do so systematically and correctly.

Lemma 2. *Given a probability density function $\rho_{profile}$ defined on profile space and a point $x \in \mathbb{R}^3$ within essential octohedral coordinates, E defined in Eq. 3.2, we can construct the density function $\rho_{essential}$ in essential octohedral coordinates as:*

$$\rho_{essential}(x) = \int_{RE^{-1}(x)} \frac{1}{8} \rho_{profile} dV_R \tag{3.17}$$

Or, equivalently, with S_x, the supporting space of x in the space of reversal components, write that:

$$\rho_{essential}(x) = \int_{S_x} \frac{1}{8} \rho_{profile} dV_R \tag{3.18}$$

To convert from a measure (density function) on $\mathbb{R}^6$ to a measure on $\mathbb{R}^3 \times \mathbb{R}^3$ (the product space of the representation cube with the supporting spaces), we note that by differentiating Eq. 3.2, we get, for each triplet (X, Y, Z):

$$dp_{XYZ-ZYX} = dp_{XYZ} - dp_{ZYX} \tag{3.19}$$

Differentiating Eq. 3.7 gives:

$$dp_{XYZ+ZYX} = dp_{XYZ} + dp_{ZYX} \tag{3.20}$$

Combining Eqs. 3.19 & 3.20 , we have, by the antisymmetry of the wedge product, that:

$$\frac{1}{2} dp_{XYZ-ZYX} \wedge dp_{XYZ+ZYX} = dp_{XYZ} \wedge dp_{ZYX} \tag{3.21}$$

Instantiating and applying Eq. 3.21 several times tells us that:

$$\rho dp_{ABC} \wedge dp_{BAC} \wedge dp_{BCA} \wedge dp_{CBA} \wedge dp_{CAB} \wedge dp_{ACB} = \frac{1}{8} \rho (dp_{ABC+CBA} \wedge dp_{CAB+BAC} \wedge dp_{ACB+BCA}) \wedge (dp_{ABC-CBA} \wedge dp_{CAB-BAC} \wedge dp_{ACB-BCA}) \tag{3.22}$$

I.e., that the differential volume element of profile space is equal to $\frac{1}{8}$ times the wedge product of the differential volume element of essential octohedral space. This tells us, conveniently, that:

$$\int_S \rho dV = \int_{ES} \left(\int_{RE^{-1}(x) \cap RS} \frac{1}{8} \rho dV_R \right) dV_E \tag{3.23}$$

That is, $\rho_{essential}(x)$ is defined as the integral over the region $RE^{-1}(x)$ of the natural projection of the original density function $\rho_{profile}$ on $E^{-1}(x) \subset \mathbb{R}^6$ - adjusted by an important factor of $\frac{1}{8}$.

We will define three large classes of density functions in essential octohedral coordinates ($\mathbb{R}3$); these classes will have equivalent classes of density functions on the representation cube. One includes IAC, one includes density functions which correspond closely to IC (for large N); the third class is constructed from the first two classes, a considerably larger superset that includes both.

Definition 3.1 (ρ_P-type density function)**.** *Suppose that ρ is a density function on the representation cube. We say that ρ is a ρ_P-type distribution if and only if we may write:*

$$\rho(\vec{x}) = \alpha(r_P(\vec{x})) \tag{3.24}$$

Where r_P is the radius-like quantity defined in Eq. 3.8, and $\vec{x}$ is the position vector on the representation cube.

This class includes the distribution induced by the IAC assumption. It also includes a large number of other distributions, some of which are superficially unlike IAC. Any function of r_P that is integrable on the representation cube falls within this class, and there are a very large number of integrable functions of one variable on a bounded domain.

Definition 3.2 (ρ_S-type density function)**.** *Suppose that ρ is a density function on the representation cube. We say that ρ is a ρ_S-type distribution if and only if we may write:*

$$\rho(\vec{x}) = \alpha(r_S(\vec{x})) \tag{3.25}$$

Where r_S is the radius-like quantity defined in Eq. 3.11.

Definition 3.3 (ρ_{P+S}-type density function)**.** *Suppose that ρ is a density function on the representation cube. We say that ρ is a ρ_{P+S}-type distribution if and only if we may write:*

$$\rho(\vec{x}) = \alpha\rho_S + (1-\alpha)\rho_P \tag{3.26}$$

Where $\alpha \in [0,1]$.

3.4.2 The assumption of IAC

We begin with a simple statement:

Lemma 3. *The Impartial Anonymous Culture assumption induces an r_P-type density function on the representation cube.*

Proving this is straightforward. Since fixing the three majority vote comparisons fixes the essential profile, we can recall that each point in essential octohedral coordinates corresponds one to one with an essential profile. Per Chapter 3.1, the total fraction of the population in the essential profile component *plus* the total fraction in the supporting profile is constant at a normalized value of 1. Recall that we can compute the density function in essential octohedral coordinates as:

$$\rho_{essential}(x) = \int_{RE^{-1}(x)} = \frac{1}{8}\rho_{profile} \tag{3.27}$$

We know exactly what $\rho_{profile}$ is: Some constant number k. We also know exactly what the measure of $RE^{-1}(x)$ is: The surface area of S_x. So, for IAC:

$$\rho_{essential}(x) = \|S_x\| \tag{3.28}$$

How large is S_x? The natural way to describe the supporting space of profiles behind each point of the representation cube is as a triangle in $\mathbb{R}^3$. This means that our integral is the integral of some constant function k over an equilateral triangle with side length $\sqrt{2}(1-r_P)$, i.e., $\frac{\sqrt{3}}{2}k(1-r_P)^2$. In particular, it is useful to work with one of the following three density functions in $\mathbb{R}^3$:

$$\rho(r_P) = (1 - r_P)^2 \tag{3.29}$$

This density function makes some calculations simple. It is not a probability distribution; but it is the simplest density distribution that has the correct form.

$$\rho(r_P) = \frac{15}{8}(1 - r_P)^2 \tag{3.30}$$

This density function integrates to 1 over the entire cube, and so is a probability density function over the representation cube.

$$\rho(r_P) = 12(1 - r_P)^2 \tag{3.31}$$

For this choice of leading constant, the integral over a single transitive octant is conveniently 1, making this an appropriate choice when dealing with conditional probabilities contingent on the existence of a Condorcet winner and loser.

3.4.3 The assumption of IC

The Impartial Culture (IC) assumption induces a very specific probability distribution for each choice of N, the number of voters. However, we may state that:

Lemma 4. *The IC (Impartial Culture) assumption is, for large values of N, closely approximated by an r_S type distribution.*

We will prove this lemma in the context of essential octohedral coordinates. Lemma 3.2 immediately extends this to all other spaces of interest. The assumption of impartial culture is that each voter decides with equal probability between each possible preference, independently If N voters express preferences in accordance with an Impartial Culture distribution, the resulting profile can be considered to have been constructed as a sum of N unit vectors:

$$P = \sum_{i=1}^{N} e_i \tag{3.32}$$

Where each e_i is one of the unit basis vectors:

- (1,0,0,0,0,0)
- (0,1,0,0,0,0)
- (0,0,1,0,0,0)
- (0,0,0,1,0,0)
- (0,0,0,0,1,0)
- (0,0,0,0,0,1)

Applying the transformation of Eq. 3.2 then gives us:

$$EP = \sum_{i=1}^{N} Ee_i \tag{3.33}$$

For each possible e_i, Ee_i is then, with equal probability:

- $\pm(1, 0, 0)$
- $\pm(0, 1, 0)$
- $\pm(0, 0, 1)$

To draw a series of N independent random voter vectors is, in essential octohedral coordinates, equivalent to taking N steps of unit size in a random walk along a three-dimensional lattice. This is a very well known problem, the solution to which has been explored very extensively. As might be expected from the Central Limit Theorem, the resulting distribution closely approximates a particular multinormal distribution; as might be expected from the symmetry of the vectors, it happens to be spherically symmetric:

$$\rho = e^{-\frac{r^2}{9\pi N}} \tag{3.34}$$

Where r is the ordinary Cartesian radius, and thus:

$$\rho = e^{-\frac{r_S^2}{9\pi N}} \tag{3.35}$$

Which is to say that the density distribution induced by the IC distribution is, for large N, approximated closely (via the Central Limit Theorem) by a ρ_S-type distribution. This means that while the density distribution induced by the IC distribution is not a ρ_S-type distribution, it is (for large N) close enough to a ρ_S-type distribution that the ρ_S class of distributions meaningfully approximates it. It is also worth noting that a variety of *similar* statistical assumptions lead to similarly symmetric bell-curve type distributions; so this class includes many distributions that are of potential interest.

3.5 Simplifying symmetries

There are several useful symmetries present in the general model of voting behavior. The first thing to take note of is that names are generally considered to be arbitrarily chosen. Because names are arbitrarily chosen, we can know everything that is going on by examining simply two octants of the entire space. First, the octant in which:

$$P(A, B) > 0, P(B, C) > 0, P(C, A) < 0 \tag{3.36}$$

I.e., the transitive octant in which $A \succ B \succ C$ transitively by pairwise majority votes. Second, the octant in which:

$$P(A, B) > 0, P(B, C) > 0, P(C, A) > 0 \tag{3.37}$$

I.e., the cyclic octant in which we have a Condorcet cycle where $A \succ B \succ C \succ A$. There are six transitive octants and two cyclic octants in total. The operation of candidate permutation can be used to map each transitive octant onto each other transitive octant; and likewise with the two cyclic octants. Conceptually, "candidate permutation" refers to changing the map between candidates; unless we are concerned with ordering effects (or incumbency, or other asymmetric effects) it does not matter which candidate is identified with which variable. This means we only need to examine at most one transitve octant (which gives us the result for $\frac{15}{16}$ of the space) and one cyclic octant (which gives us the result for $\frac{1}{16}$ of the space).

In cases where we assume the existence or nonexistence of a Condorcet winner, we need only examine only a single representative octant. If we assume a Condorcet winner, such as when considering the probability of a Condorcet winner winning under a particular $(1, s, 0)$ rule, we only need a single transitive octant; and if we assume that no Condorcet winner exists, we need only examine a single non-transitive octant.

Some calculations are easier to perform in essential octohedral coordinates, where each octant contains only one face; and each octant contains $\frac{1}{8}$ of the total space. In this case, the octants may be divided into six strongly non-cyclic octants (accounting for 75% of the total space), and two octants which are not strongly non-cyclic (weakly cyclic octants accounting for 25% of the total space).

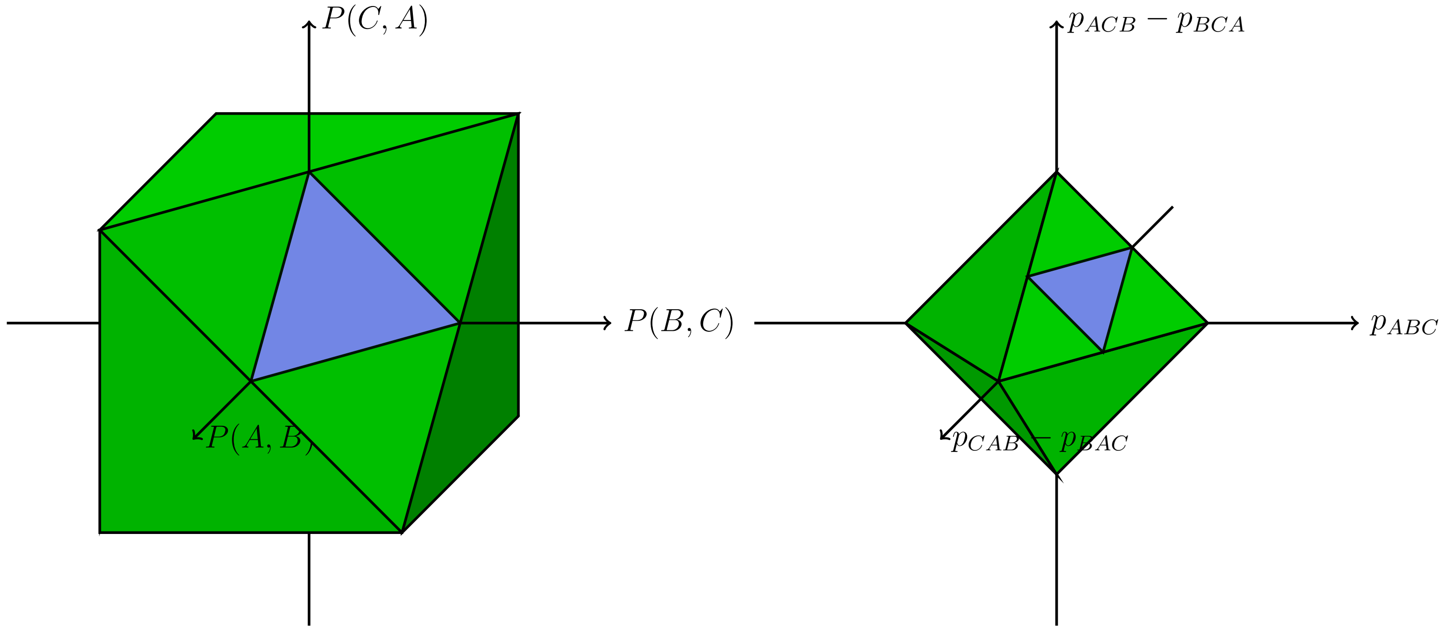
3.6 Probability of a Condorcet paradox

We have at this point the tools needed to calculate the probability of a Condorcet paradox, where $P(A, B)$, $P(B, C)$, and $P(C, A)$ all share the same sign on several probability distributions. We will start with a uniform distribution on the representation cube, and extend this to a large class of probability distributions including the one induced by IAC.

3.6.1 Uniform distribution, IAC, and related distributions

The first and simplest case to consider is that of a uniform distribution. The Condorcet paradox occurs in the two cyclic octants of the representation cube. When this is translated to the more symmetric essential octohedral coordinates, the faces become symmetric, and the cyclic octant now occupies the central quarter of one face of the regular octohedron (Figure 3.2.) The geometry of the situation is simple. Each octant of the representation cube occupies equal and identical spaces in essential octohedral coordinates ($\mathbb{R}^3$); the two weakly transitive regions each take up $\frac{1}{8}$ of the total octohedron. The Condorcet paradox occurs in the blue region - a pyramid with a base equal to one quarter of the whole face, going to a singular point at the origin. Each region in which the Condorcet paradox occurs takes up $\frac{1}{4}$ of a weakly transitive region, or $\frac{1}{32}$ of the entire reperesentation cube. That there are two such regions immediately brings us to the figure of $\frac{1}{16}$, which (not coincidentally)

Figure 3.2: Representation cube in tally space and essential octohedral coordinates (Condorcet paradox region in blue)



is equal to the probability of a Condorcet paradox under the IAC assumption and a large number of voters.

That is volume in $\mathbb{R}^3$ - equal to probability only if we assume a uniform probability distribution over $\mathbb{R}^3$ (the representation cube) - and the probability density function implied by the IAC assumption is *not* a uniform distribution. However, it *is* an integrable ρ_P-type distribution, which means that we can construct arbitrarily good approximations of the distribution induced by IAC with a stepwise function of r_P. In this case, the following version of the classic Riemann sum will suffice:

$$\rho_k(r_P) = \sum_{i=1}^{k} a_i H\left(\frac{i}{k} - r_P\right) \tag{3.38}$$

Where H is the Heaviside function (0 if < 0, 1 if ≥ 0). The same geometry that holds for the uniform distribution, moreover, holds for each $a_i H(\frac{i}{n} - r_P)$ in that summation, as each is simply no more than a scaled copy of the original representation cube; and in holding for each component function, holds for the entire sum. More to the point, this holds, in effect, for every integrable function of r_P:

Theorem 3.3. *If the probability distribution of pairwise majority votes follows a probability distribution ρ which may be written as an integrable function $\rho_P(r_P)$ of the single variable r_P, the probability of a Condorcet paradox is $\frac{1}{16}$.*

Since IAC leads to a probability distribution that can be written as a function of r_p, namely:

$$\rho_P(r_p) = k(1 - r_P)^2 \tag{3.39}$$

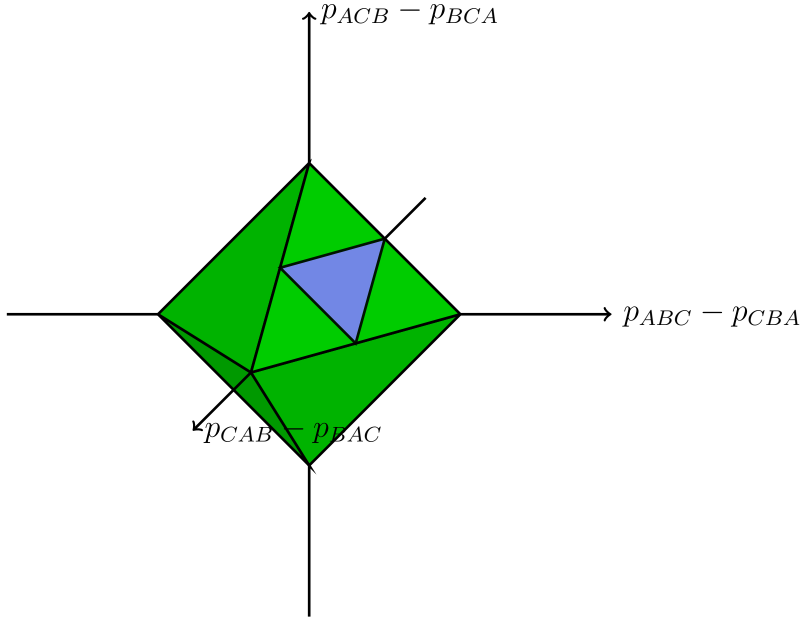
The known fact that assuming IAC leads to a 6.25% chance of a Condorcet paradox for large N is a special case of the more general theorem above. The extension to any integrable function of r_P describes a much larger class of probability distributions. This makes a much stronger case for why we might expect to continue to see figures close to 6.25% in empirical data; but also means that the case for IAC being an accurate description of the structure of voter preferences is not greatly helped by empirical data matching this prediction.

3.6.2 Approximating IC and related distributions

In essential octohedral coordinates, the class of ρ_S-type distributions become the class of spherically symmetric integrable functions. Of particular note with respect to the calculation of the probability of a Condorcet paradox, it takes the convex hull of $\{(0,0,0),(1,0,0),(0,1,0),(0,0,1)\}$ to the regular tetrahedron whose corners are $\{(0,0,0),(\frac{1}{2},\frac{1}{2},0),(\frac{1}{2},0,\frac{1}{2}),(0,\frac{1}{2},\frac{1}{2})\}$.

While integrals over ρ_S-type distributions can be computed readily within the original representation cube, it is in essential octohedral coordnates that results about those distributions become most intuitive. It is the case that r_S, like r_P, is independent of the boundaries deter-

Figure 3.3: Condorcet paradox region



mining whether or not a Condorcet winner exists, so we can start with a sphere of arbitrary radius inside essential octohedral coordinates, and compute the fraction of the sphere taken up by the region in which no Condorcet winner exists. Just as in the case of a ρ_P-type distribution, we can approximate integrable functions of r_S with summations:

$$\rho_k(r_P) = \sum_{i=1}^{k} a_i H\left(\frac{i}{k} - r_P\right) \tag{3.40}$$

Meaning they all give the same result for the likelihood of a Condorcet paradox. At this point, the only thing that remains to calculate the probability of a Condorcet paradox under *any* assumption leading to a ρ_S-type distribution (which in turn closely approximates the multinormal distribution produced by the Impartial Culture assumption, for large numbers of voters) is computing the solid angle of the blue region in Figure 3.3.

That region is a regular tetrahedron. The solid angle occupied by a regular tetrahedron in $\mathbb{R}^3$ as viewed from one of its vertices (the origin, in this case) is well known - $\frac{cos^{-1}\left(\frac{23}{27}\right)}{4\pi}$. There are two such regions, and so we may state:

Theorem 3.4. *If the probability distribution of pairwise majority votes follows a probability* *distribution ρ which may be written as an integrable function $\rho_P(r_P)$ of the single variable r_P, the probability of a Condorcet paradox is $\frac{cos^{-1}\left(\frac{23}{27}\right)}{2\pi}$.*

This is, in other words, about an 8.8% chance; meaning that, since the IC assumption may be closely approximated by such distributions, that IC gives us an approximately 8.8% chance of a Condorcet paradox. This is higher, although not greatly distinct from, the probability of a Condorcet paradox under IAC and similar distributions.

By the intermediate value theorem, we have as an immediate consequence of our two theorems the following corrolary:

Corollary 5. *If $\rho(r_S, r_P) = t\rho_S(r_S) + (1-t)\rho_P(r_P)$ is a probability distribution over pairwise majority votes that can be written as the linear combination of a ρ_S type and ρ_P type distribution, then the probability of a Condorcet paradox given the distribution ρ is between $\frac{1}{16}$ and $\frac{cos^{-1}\left(\frac{23}{27}\right)}{2\pi}$.*

In other words, every function that can be approximated by summing together ρ_S and ρ_P type distributions will predict Condorcet paradoxes at a rate that falls somewhere within a relatively narrow range. This class of probability density functions is considerably larger than either the class of ρ_P-type distributions or the class of ρ_S-type distributions, and includes almost every distribution that is likely to be considered as a candidate for describing voter behavior *a priori*.

3.7 Probability that a Borda count selects a Condorcet winner or loser

It is known that a Borda count cannot select a Condorcet loser. A slightly more complex example is the probability that a Borda count selects a Condorcet winner. The first problem we face is that we must exclude the probability of a Condorcet paradox from our set of possible results; the answers to that problem are in Sec. 3.6, however, and are in particular $6.25\% - 8.8\%$ under our favored assumptions.

This is not the most complex example we will demonstrate. The Borda count is constant over the supporting space at any point in the representation cube; all binary-equivalent profiles result in the same Borda count totals. We may therefore still read results directly off the representation cube, or any reasonable fascimile thereof, and the techniques used largely mirror the above techniques used to find the probability of a Condorcet paradox.

3.7.1 ρ_P-type distributions

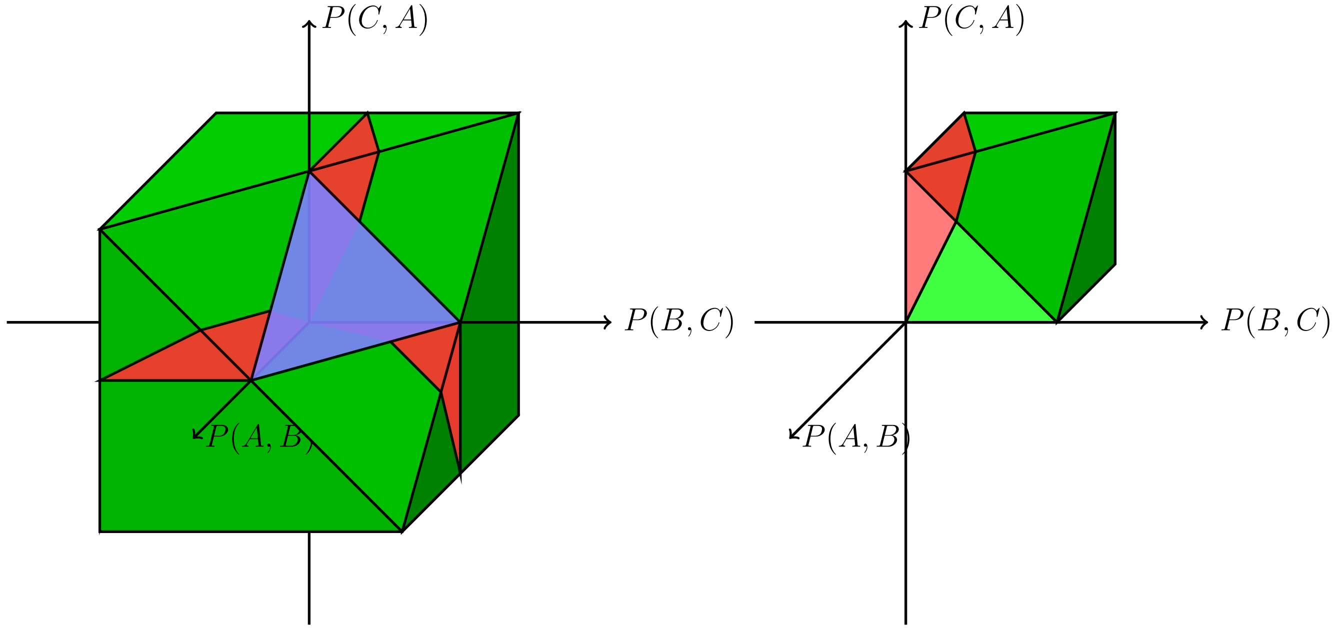
If A is the Condorcet winner, and C is the Condorcet loser, the necessary and sufficient condition for A to be the Borda winner is given by the following condition from Chapter 2:

$$2P(A,B) + P(A,C) > P(B,C) \tag{3.41}$$

Where A is the Condorcet winner and C is the necessary Condorcet loser. As we are restricting our attention to situations in which there is a Condorcet winner, we can (per the argument of Sec. 3.5) restrict our attention to a single octant where:

- $P(A,B) \geq 0$
- $P(A,C) \geq 0$
- $P(B,C) \geq 0$

Figure 3.4: Borda-Condorcet agreement
(Full cube, left; transitive octant, right)



With our final boundary equation given by the original requirement that individual voters have transitive preferences (and thus, by the plane that truncates the representation cube):

$$P(A,B) + P(B,C) \leq 1 + P(A,C) \tag{3.42}$$

The figure below shows both the Condorcet paradox region and the regions where the Borda count disagrees with the Condorcet winner on the whole of the representation cube, and the region where the Borda Count disagrees with the Condorcet winner within a single transitive octant (Figure 3.4).

Note that Eq. 3.41 describes a plane through the origin, a scale-invariant surface. Consequentially, as in Sec. 3.6.1 any density distribution function ρ_P that is a function of r_P will yield an equal probability of the Borda winner and Condorcet winner coinciding as in the constant case $\rho_P \equiv k$. All that is needed, therefore, is to calculate the volume of the regions where Eq. 3.41 does and does not hold.

In this case, the smaller and simpler region is that which fails to meet Eq. 3.41. The portion within a given transitive octant has volume $\frac{13}{162}$. The total volume of the entire transitive octant of the normalized representation cube is $\frac{5}{6}$; which gives us $\frac{13}{135}$ as the probability of a Condorcet winner losing given that our election result falls within that transitive octant in $\mathbb{R}^3$. Symmetry dictates that the same figure holds in each of the other five transitive octants. It is worth emphasizing that this calculation holds accurate for any ρ_P-type distribution.

Theorem 3.5. *If the probability distribution of pairwise majority votes, ρ, may be written as an integrable function $\rho_P(r_P)$ of the single variable r_P, and there exists a Condorcet winner, the probability that the Condorcet winner wins a Borda count is $\frac{122}{135}$.*

That is to say, if we have a large number of voters whose voting behavior is distributed in the way we would expect from the assumption of Impartial Anonymous Culture (IAC) and sincere accurate voting, Condorcet winners should win Borda counts $\frac{122}{135}$ of the time; and this rate will hold for a wide variety of similar distributions.

3.7.2 ρ_S-type distributions

If we want to move to an approximation of IC, or a ρ_S-type density distribution, it is convenient to transition to essential octohedral coordinates; at which point the boundaries of the region where a Condorcet winner A with Condorcet loser C is also a Borda winner are given by:

$$p_{ABC-CBA} + p_{CAB-BAC} > 2p_{BCA-ACB} \tag{3.43}$$

$$p_{ABC-CBA} + p_{CAB-BAC} > p_{BCA-ACB} \tag{3.44}$$

$$p_{ABC-CBA} + p_{BCA-ACB} > p_{CAB-BAC} \tag{3.45}$$

$$p_{ABC-CBA} > p_{CAB-BAC} + p_{BCA-ACB} \tag{3.46}$$

This lends itself naturally to a very direct representation in terms of spherical coordinates on $\mathbb{R}^3$; letting $p_{ABC-CBA}$ become the azimuthal axis with azimuthal angle ϕ and lateral angle θ, we transform that set of equations into a set of constraints on ϕ as functions of θ:

$$\phi < M_\phi = \min \left\{ \begin{array}{c} \cot^{-1}\left(2\sin(\theta) - \cos(\theta)\right) \\ \cot^{-1}\left(\sin(\theta) - \cos(\theta)\right) \\ \cot^{-1}\left(-\sin(\theta) + \cos(\theta)\right) \\ \cot^{-1}\left(\sin(\theta) + \cos(\theta)\right) \end{array} \right\} \tag{3.47}$$

Which gives us a simple integral to find the solid angle:

$$\int_0^{2\pi} \int_0^{M_\phi} \sin(\phi) d\phi d\theta \tag{3.48}$$

Integrating over ϕ and θ gives us a solid angle of about 1.72 steradians; after factoring in the fact that we have six different possible pairs of Condorcet winner and loser, and the regions in which there is no Condorcet winner, we come up with the following conditional probability:

Theorem 3.6. *If the probability distribution of pairwise majority votes follows a probability distribution ρ which may be written as an integrable function $\rho_S(r_S)$ of the single variable r_S, and there is a Condorcet winner, the probability of a Condorcet winner being the winner* *of a Borda count is* $\approx 90.1\%$.

It is worth noting that this is very close to the probabliity cited in Thm. 3.5, the probability of affirming the Condorcet winner under a very different class of probability distributions. Significant changes in how likely a Condorcet winner is to win a Borda count require very unusual underlying probability distributions that are, by and large, not being considered in the current literature.

3.7.3 Generalizing

As in Sec. 3.6, we have an immediate corollary giving us a range of figures for a much larger class of probability functions; a class large enough to include or closely approximate most probability functions that people are likelyto use. Combining Thms. 3.5 and 3.6, we have as an immediate corollary a useful statement about a wide range of probability distributions:

Corollary 6. *If* $\rho(r_S, r_P) = t\rho_S(r_S) + (1-t)\rho_P(r_P)$ *is a probability distribution over pairwise majority votes that can be written as the linear combination of a* ρ_S *type and* ρ_P *type distribution, then if pairwise majority votes follow the distribution* ρ*, the conditional probability of a Condorcet winner losing a Borda count (given the existence of such a Condorcet winner) is* $\approx 9.6 - 9.9\%$.

Note that this happens slightly more often than a Condorcet paradox; but is higher for the IC-related distributions than the IAC-related distributions, just as with the Condorcet paradox. These are not the only probability distributions which produce a probability of a Condorcet winner losing a Borda count between 9-10%; however, it contains most probability distributions of interest or close approximations thereof. Any empirical study of electoral behavior would be hard-pressed to distinguish between most behavioral assumptions based on the frequency with which Condorcet winners lose Borda counts.

An empirical measurement of this frequency falling outside of this range by a significant margin would suggest that the underlying distribution of voter preferences, or at least behavior, does not tend towards any of the common assumptions made in the current literature, including both Impartial Culture and Impartial Anonymous Culture.

3.8 Probability that a plurality vote selects a Condorcet winner or loser

To examine the cases where a plurality vote agrees or disagrees with the Condorcet criterion, it helps to start first with the case of the Condorcet winner. In calculating the probability of a plurality vote selecting a Condorcet winner, we cannot consider all ρ_S-type distributions simultaneously, nor all ρ_P-type (IAC-type) distributions simultaneously. We will, in fact, only consider five distributions, three of which are ρ_P-type (IC-type) distributions, and infer a few things about the class of distributions.

3.8.1 Limiting cases

We will first consider some extreme cases. First, if all possible votes are concentrated tightly in a tiny structure with a maximum radius of ϵ, but distributed uniformly across the supporting space, symmetry dictates that the Condorcet winner has a chance of winning that is within an $\mathcal{O}(\epsilon)$ error term of $\frac{1}{3}$. This is because the effect of the supporting space dominates. A second limiting case: If all possible votes are concentrated on the external boundary, the Condorcet winner always wins the plurality vote. This leads to the following summarization of results:

Theorem 3.7. *1. Given a probability distribution that is a linear combination of ρ_P and*

ρ_S type distributions, the probability of a Condorcet winner winning a plurality vote falls between $\frac{1}{3}$ and 1.

2. *Conversely, given a probability $p \in \left(\frac{1}{3}, 1\right)$, a probability distribution that is a ρ_P-type distribution can be constructed that gives a probability p of a Condorcet winner winning a plurality vote.*

The proof of this follows immediately from the boundary cases. It can be illustrated very easily by picture. As shown in Figure 3.5, the $r_P \equiv 1$ surface contains no cases in which the Condorcet winner is defeated by another candidate in a plurality vote. Figure illustrates what happens at the origin.

3.8.2 Condorcet winner under IAC

Thms. 1-2 in Chapter 2 give the following five conditions determining the winner of a plurality vote:

- If $P(A, C)$ is the largest pairwise victory, A must win if $2P(A, B) + P(A, C) > 1$.
- If $P(B, C)$ is the largest pairwise victory, A must win if $2P(A, B) + P(B, C) > 1$.
- If $P(A, B) > P(A, C) > P(B, C)$, A must win if $P(A, B) + 2P(A, C) > 1$.
- If $P(A, B) > P(B, C) > P(A, C)$ A must win if $P(A, B) + 2P(B, C) > 1$.
- It is possible for any candidate to win if and only if $P(X, Y) < \frac{1}{3}$ for all X, Y.

These conditions give boundaries, which in turn describe two surfaces enclosing two regions. These are pictured in Figure 3.5.

The region U (left) gives the boundaries where A must win; the region V (right) gives the

Figure 3.5: Regions where the plurality vote only sometimes agrees with the Condorcet winner (left) and where all winners are possible (right)

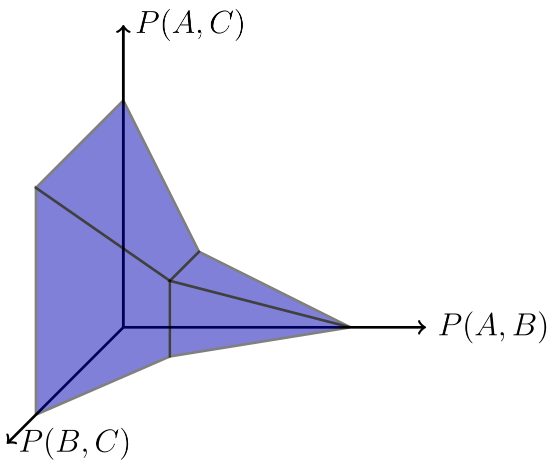
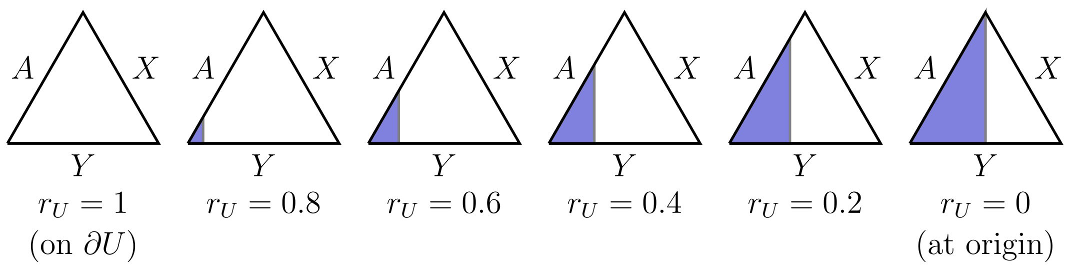


Figure 3.6: Plurality-Condorcet agreement in the supporting space, $X \succ A$ cases

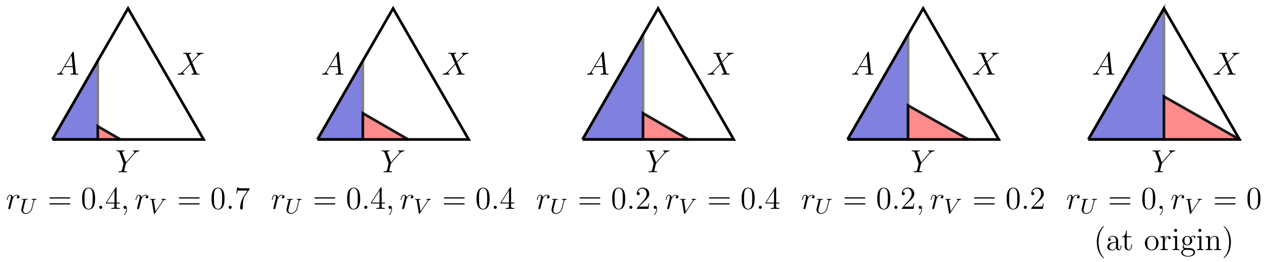


cases where both B and C can win. This region U has a volume $V = \frac{5}{27}$. It is worth noting that any scaled copy of ∂U has a uniform density of the Condorcet winner A defeating the next-most likely candidate to win. (Over the larger part of the surface, this is B; over the smaller part of this surface, this is C.) Defining $r_U \equiv 1$ on this surface, dropping linearly to 0 at the center, the associated differential volume of is $\frac{5}{9} r_U^2 dr_U$.

It is at this point necessary to consider the supporting space. In Figure 3.6 , we see a series of boundaries between A winning a plurality vote and X winning a plurality vote, as we move from a point on the boundary of U to the origin. The shaded region in Figure 3.6 marks areas in the supporting space where X defeats A.

In particular, X is C if $P(B,C)$ is larger than both $P(A,C)$ and $P(A,B)$, and B otherwise.

Figure 3.7: Plurality-Condorcet disagreement in the supporting space



The scale of the supporting space varies; the dimension of the shaded triangle is linearly proportionate to $1 - r_U$. The total density of B defeating A and goes from a density of 0 at the boundary of U to a density of $\frac{1}{2}$ of the total density at the center. Using the probability density function of $12(1 - r_P)^2$ as a basis, the probability density of A being defeated by the second place Borda candidate in a plurality vote is then exactly $6(1 - r_U)^2$.

$$\int_U \rho_2 dV = \int_0^1 6(1 - r_U)^2 \frac{5}{9} r_U^2 dr_U^2 = \frac{1}{9} \tag{3.49}$$

This gives a probability of $\frac{1}{9}$ of the total space. This is already a good estimate of the cases in which a Condorcet winner loses; but not quite an exact calculation. It is not only possible that we have $X \succ A$ by the plurality vote it is also possible, for some cases, to have $Y \succ A$ by the plurality vote. To add cases where $Y \succ A$ (but not $X \succ A$) is fairly simple. In particular, this happens inside the region V. Figure 3.7 shows the supporting space for a selection of values of r_U and r_V.

The red region represents the region where the plurality vote gives the result of $Y \succ A \succ X$. The size of this region depends on r_V but not r_U, and has a limiting value of $\frac{1}{6}$ as we go towards the origin. It increases simply by a square factor, in particular:

$$\int_V \rho_3 dV = \int_0^1 2(1 - r_V)^2 \frac{1}{9} r_V^2 dr_V = \frac{1}{135} \tag{3.50}$$

Note that if ρ^* is the total density of A (the Condorcet winner) losing a plurality vote, we have that:

$$\rho_2 + \rho_3 = \rho^* \tag{3.51}$$

Combining Eqs. 3.51, 3.49 & 3.50 allows us to compute the integral of ρ^* over a sample transitive octant T, giving:

$$p_{CWP} = \int_T \rho^* dV = \frac{1}{9} + \frac{1}{135} = \frac{16}{135} \tag{3.52}$$

This is fairly unlikely, in other words; but it is sensitive to distribution. We will consider another ρ_P-type distribution next.

3.8.3 Condorcet winner under another ρ_P-type distribution

In particular, we will use the example of a piecewise function, such that:

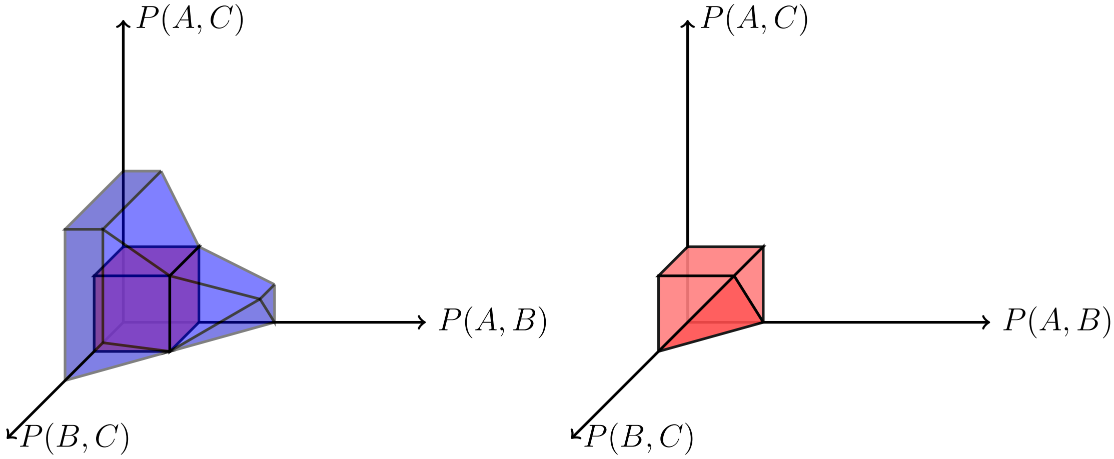
$$\rho_P(r_P) = \begin{array}{l} k_1(1-r_P)^2, r_P \in [0, \frac{1}{3}] \\ k_2(1-r_P)^2, r_p \in (\frac{1}{3}, \frac{2}{3}] \\ k_3(1-r_P)^2, r_p \in (\frac{2}{3}, 1] \end{array} \tag{3.53}$$

With the distribution within the supporting space of any point x being held to some constant c_x. To compute this case, we consider two additional geometric objects, Q and W, which we will associate to radius-like quantities r_Q and r_W, illustrated in Figure 3.8.

Note that:

$$W \subsetneq V \subsetneq Q \subsetneq U \tag{3.54}$$

Figure 3.8: Plurality-Condorcet disagreement.
From left to right, $r_P \leq \frac{2}{3}$, and $r_P \leq \frac{1}{3}$



If we restrict ourselves to the domain of a single transitive octant and the following pair of conditions is met:

$$k_1 + k_2 + k_3 = 36 \tag{3.55}$$

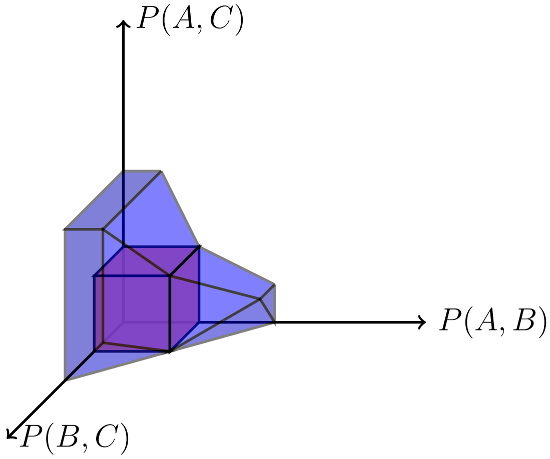
$$k_1 \geq k_2 \geq k_3 \geq 0 \tag{3.56}$$

Then the density function ρ_P is, in fact, a probability distribution, and we have that the probability of a Condorcet winner losing is given by:

$$p = \frac{k_1 - k_2}{12} \int_0^1 f_W dr_W + \frac{k_2 - k_3}{12} \int_0^1 f_Q dr_Q + \frac{k_3}{12} \int_0^1 f_U dr_U + \frac{k_2}{12} \int_0^1 f_V dr_V \tag{3.57}$$

Given, in each case, that f_I an appropriately function of r_I. This is already done for f_U and f_V in, respectively, Eq. 3.49 and Eq. 3.50. Note that on the surface of Q, if we define a parameter s on ∂Q, the density function of the Condorcet winner being defeated by the

Figure 3.9: Density of plurality-Condorcet disagreement



candidate who would place second is given by:

$$\rho^* = 6\left(1 + \frac{r_Q s}{3} - r_Q\right)^2 \tag{3.58}$$

This parameter s is uniformly zero on the lightly shaded portions of Q, and attains a value of 1 on the $P(B,C) = 0$ plane, as well as the point where the dark regions intersect the $P(B,C)$ axis in Figure 3.9.

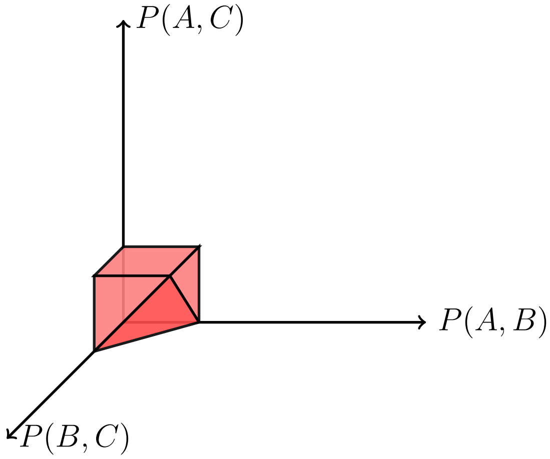
This treatment leads to breaking the integral into three separate parts. The two darkly shaded triangular regions at the base each have:

$$\int_0^1 \int_0^1 \frac{\sqrt{3}}{9}(1-s)r_Q^2 \left(1 + \frac{r_Q s}{3} - r_Q\right)^2 ds dr_Q = \frac{79\sqrt{3}}{48600} \tag{3.59}$$

The dark rectangular panel on top, which is equivalent to the lower panel plus the smaller triangular region, gives:

$$\int_0^1 \int_0^1 \frac{2}{3} r_Q^2 \left(1 + \frac{r_Q s}{3} - r_Q\right)^2 ds dr_Q = \frac{31}{810} \tag{3.60}$$

Figure 3.10: Plurality-Condorcet agreement, central region



The lightly shaded regions can be combined in the following integral:

$$\int_0^1 \int_0^1 \frac{23}{36} r_Q^2 (1 - r_Q)^2 ds dr_Q = \frac{23}{1080} \tag{3.61}$$

Combining Eq. 3.59, Eq. 3.60, and Eq. 3.61, we get:

$$\int_Q f_Q dr_Q = 2\frac{79\sqrt{6}}{48600} + 2\frac{31}{810} + \frac{23}{1080} = \frac{4755 + 158\sqrt{3}}{48600} \approx 0.103 \tag{3.62}$$

Note this is almost the total given by Eq. 3.49 - this is not coincidental, as most cases where the Condorcet winner loses a plurality vote occur with $r_P \leq \frac{2}{3}$. We may now consider W, as illustrated in Figure 3.10.

Within W, all three candidates can win; we have, as with the original computation, two different cases to consider, one in which A loses to the candidate who would place second with the essential profile alone - first, on the sides and top:

$$\int_0^1 \int_0^1 \frac{4}{9} r_W^2 \left(1 + \frac{2}{3} s r_W - r_W\right)^2 ds dr_W = \frac{52}{1215} \tag{3.63}$$

Then the triangular region, which must be divided into two components:

$$\int_0^1 \int_0^{\frac{3}{4}} \frac{2\sqrt{6}}{9} s r_W^2 \left(1 + \frac{2}{3} s r_W - r_W\right)^2 ds dr_W = \frac{11\sqrt{6}}{1920} \tag{3.64}$$

$$\int_0^1 \int_{3/4}^1 \left(\frac{7\sqrt{6}}{24} - \frac{4\sqrt{6}s}{6}\right) r_W^2 \left(1 + \frac{2}{3} s r_W - r_W\right)^2 ds dr_W = \frac{163\sqrt{6}}{51840} \tag{3.65}$$

Or, in total:

$$\int_W \rho_2 = \frac{52}{1215} + \frac{11\sqrt{6}}{1920} + \frac{163\sqrt{6}}{51840} \tag{3.66}$$

Then there is the case in which the Condorcet winner places second - and is defeated by the candidate who would place last with the essential profile alone:

$$\int_0^1 \int_0^1 \left(\frac{4}{27} r_W^2 (1 - r_W)^2 + \frac{\sqrt{3}}{108}(1 - s)\left(1 + \frac{1}{2} s r_W - r_W\right)^2\right) ds dr_W = \frac{2}{405} + \frac{29\sqrt{3}}{15552} \tag{3.67}$$

(Note that the first term is simply 2/3 of Eq. 3.50.) Combining Eqs. 3.66 and 3.67 then gives:

$$\int_W \rho^* = \frac{52}{1215} + \frac{11\sqrt{6}}{1920} + \frac{163\sqrt{6}}{51840} + \frac{2}{405} + \frac{29\sqrt{3}}{15552} \approx 0.073 \tag{3.68}$$

Taking each of Eqs. 3.49, 3.50, 3.62, and 3.68, and plugging these values back into Eq. 3.57, we can compute direct values for any particular choice of (k_1, k_2, k_3). The following table gives several examples:

Of note for the third line is that it roughly mirrors a normal distribution. Simply working within the center third of representation, instead of the entire cube, comes close to doubling the probability that a Condorcet winner loses; note that this still includes a large number of

Table 3.1: Condorcet-plurality agreement in a piecewise ρ_P distribution

k_1	k_2	k_3	Probability
12	12	12	0.1185
18	12	6	0.1510
24	10	2	0.1785
36	0	0	0.2181

highly decisive elections where a single candidate, out of a field of three, achieves as much as two thirds of the vote.

3.8.4 Condorcet loser under a plurality vote

In a subset of the cases where the Condorcet winner does not win a plurality vote, the Condorcet loser wins instead. The conditions relevant to this are:

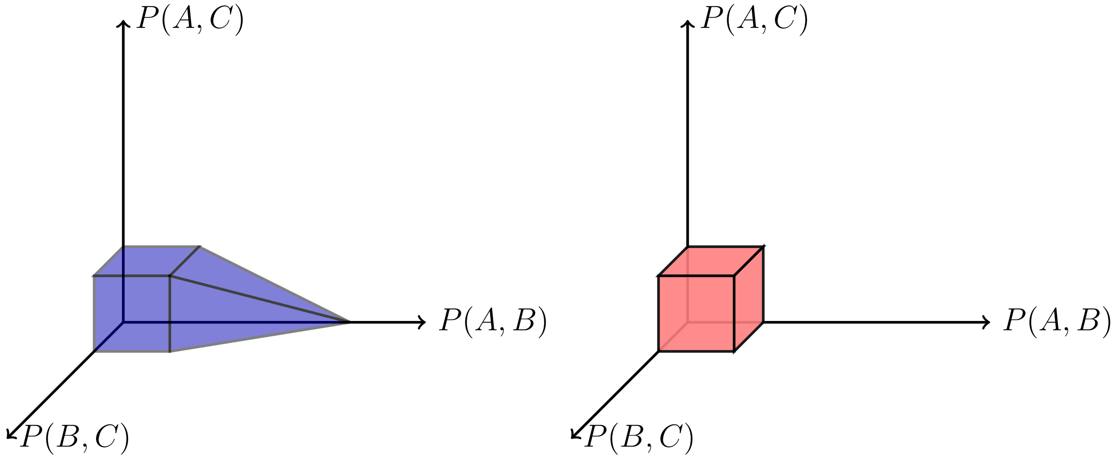
- If $P(A,B) > P(A,C) > P(B,C)$, A must win if $P(A,B) + 2P(A,C) > 1$.
- If $P(A,B) > P(B,C) > P(A,C)$ A must win if $P(A,B) + 2P(B,C) > 1$.
- It is possible for any candidate to win if and only if $P(X,Y) < \frac{1}{3}$ for all X, Y.

This gives us the regions U and V, illustrated in Figure 3.11. We will calculate, as a specific example, the probability of a Condorcet loser winning a plurality vote under the IAC assumption; this can be compared to the probability under the IAC assumption that the Condorcet winner loses a plurality vote.

The method of calculating this again to construct a density function ρ^* where our case of interest happens - here, ρ^* is the density of C winning.

$$p_{CLP} = \frac{\displaystyle\int \rho^* dV}{\displaystyle\int \rho dV} = \frac{\int \rho_2 dV - \int \rho_3 dV}{\int \rho dV} \tag{3.69}$$

Figure 3.11: Condorcet loser conditions, plurality vote



Here ρ_2 is defined as the density of cases where C (the Condorcet loser) defeats A (the Condorcet winner), and ρ_3 is the density of cases where C defeats A but loses to B. The support for ρ_2 is U.

As before, ρ_2 is zero on ∂U, and it is convenient to construct a radius-like quantity r_U that is 1 on ∂U and 0 at the origin to produce the explicit form of ρ_2:

$$\rho_2 = 6(1 - r_U)^2 \tag{3.70}$$

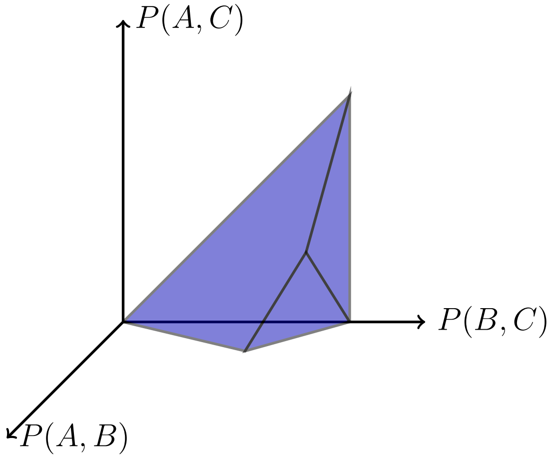
Noting that the volume of U is $\frac{5}{81}$, we can construct a differential volume element for integrating over r_U:

$$dV = \frac{5}{27} r_U^2 dr_U \tag{3.71}$$

Note that ρ_3 and V are just the same as in the case of the Condorcet winner and plurality vote, so we can refer back to Equation 3.50 and compute the probability of a Condorcet loser winning as:

$$\int_0^1 6(1 - r_U)^2 \frac{5}{27} r_U^2 dr_U - \int_0^1 2(1 - r_V)^2 \frac{1}{9} r_V^2 dr_V \qquad = \frac{1}{27} - \frac{1}{135} = \frac{4}{135} \tag{3.72}$$

Figure 3.12: Region where Borda Count disagrees with a Condorcet winner:



So the probability of a Condorcet loser winning under the IAC assumption for $n = 3$ candidates and a large number of voters is $\frac{4}{135}$. As an intermediate result, we have also that the probability that a Condorcet loser defeats a Condorcet winner in those circumstances with probability $\frac{1}{27}$ (and the ranking $B \succ C \succ A$ happens only with probability $\frac{1}{135}$).

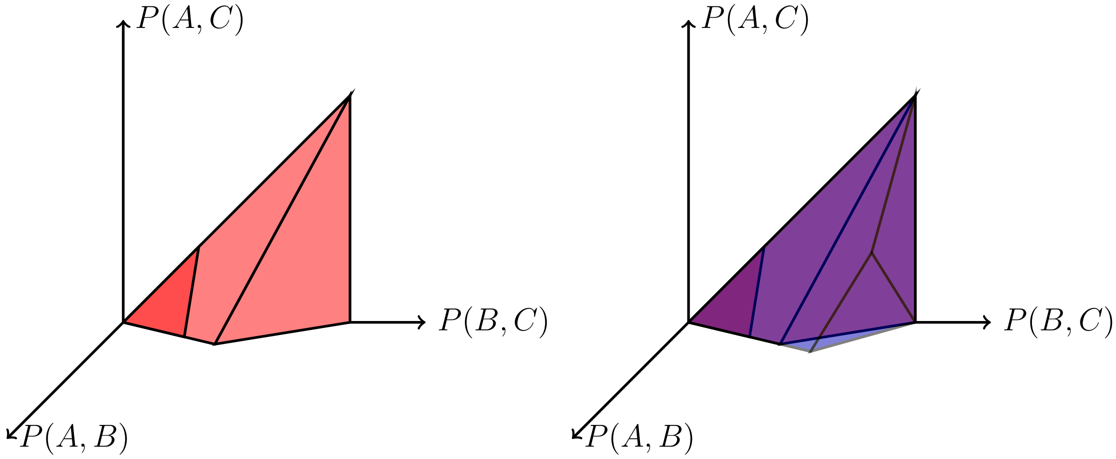
3.8.5 Borda-plurality agreement under IAC

A further question that might be asked is this: How frequently does a plurality vote agree with a Borda count? For $n = 3$ and ρ_P-type density distributions, this turns out to be a relatively simple question. It is simplified by the fact that, per Thm. 3.5, any ρ_P-type density function leads to a hard and fast $\frac{122}{135}$ probability that a Borda count gives victory to a Condorcet winner. First, we note that the region in which the Borda and Condorcet winners disagree is given by a theorem from Chapter 2, which is to say Eq. 3.41:

$$2P(A,B) + P(A,C) > P(B,C) \tag{3.73}$$

Figure 3.12 displays this region, which we will call M. Note that M makes up $\frac{13}{135}$ of the transitive octant under any ρ_P-type distribution. Also note that we have reflected coordi-

Figure 3.13: Regions in which a plurality vote sometimes, but not always, disagrees with a Borda count within M
(To the left, the subregion by itself; to the right, the subregion inside M.)



nates, to better display the object in question.

$M \cap U$, in Figure 3.13, shows the region where the plurality vote sometimes, but not always, disagrees with the Borda count. Note that the volume of $M \cap U$ is $\frac{1}{12}$. Producing a differential element gives:

$$dV = \frac{1}{4} r_{M\cap U}^2 dr_{M\cap U} \tag{3.74}$$

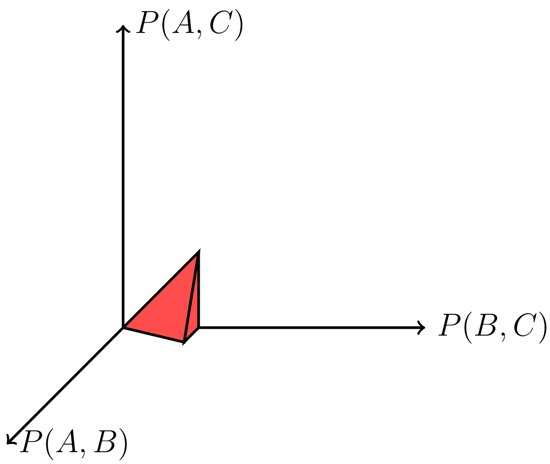
Which leads to the computation that B defeats A by plurality vote in this region, under the IAC assumption:

$$\int_0^1 6(1 - r_{M\cap U})^2 r_{M\cap U}^2 \frac{1}{4} dr_{M\cap U} = \frac{1}{20} \tag{3.75}$$

Then we should, of course, consider $M \cap V$, which gives us the cases where C wins.

$$dV = \frac{1}{324} r_{M\cap V}^2 dr_{M\cap V} \tag{3.76}$$

Figure 3.14: Cases within M where any candidate can win a plurality vote:



Which leads to the computation that $C \succ B \succ A$ in this region under the IAC assumption:

$$\int_0^1 2(1 - r_{M\cap V})^2 r_{M\cap V}^2 \frac{1}{324} dr_{M\cap V} = \frac{1}{4860} \tag{3.77}$$

How often do the plurality vote and Borda count agree and disagree? Well, within the transitive octant, we can:

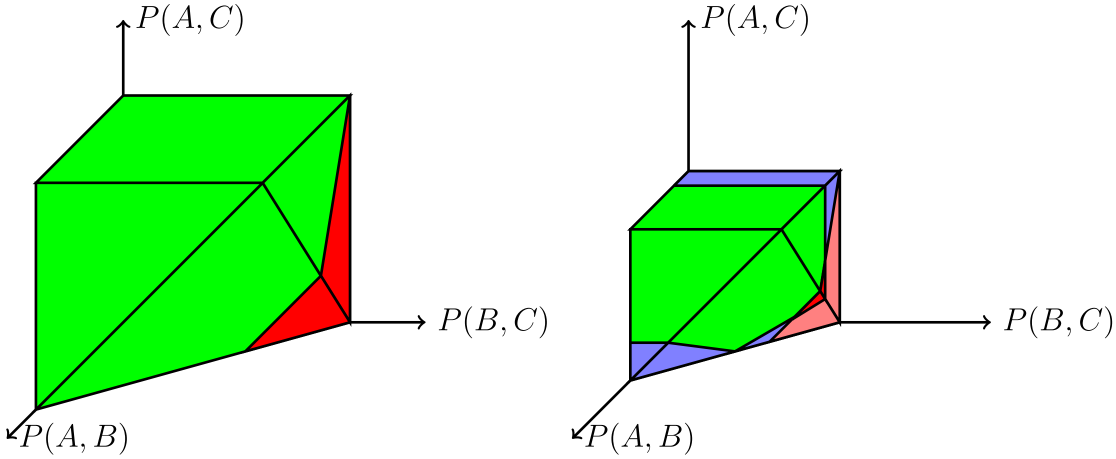
- Take the probability that a plurality and Condorcet winners are not the same. (For IAC, $\frac{16}{135}$.)
- Add all cases in M, where the Borda and Condorcet winner are not the same. (For any ρ_P distribution, $\frac{13}{135}$).
- Subtract the cases where the plurality and Condorcet winner agree within M (For IAC, $\frac{1}{20} - \frac{1}{4860}$.)

This gives us a probability (for the IAC assumption) of:

$$p = \frac{16}{135} + \frac{13}{135} - \frac{1}{20} + \frac{1}{4860} = \frac{1703}{10800} \approx 0.165 \tag{3.78}$$

It is worth noting this is higher than the probability that the plurality vote fails to elect

Figure 3.15: Level sets, Borda-Plurality interactions, for $r_P \equiv 1$ and $r_P \equiv \frac{2}{3}$



the Condorcet winner under IAC. We might ask if this holds qualitatively for all ρ_P type distributions. The answer is yes:

Theorem 3.8. *For $n = 3$ candidates and a probability distribution leading to a ρ_P-type distribution on the representation cube and a locally uniform distribution on each supporting space S_x, if a Condorcet winner exists, a plurality vote is at least as likely to select the Condorcet winner than it is to select the Borda winner. This becomes a strict inequality if the distribution assigning all weight to the $r_P \equiv 1$ shell is excluded.*

The proof of this is simple geometry: Take any shell of $r_P = k$, and examine the level of agreement or disagreement on that shell. For $r_P = 1$, we have a disagreement of $\frac{13}{135}$. Figure 3.15 illustrates $k = \frac{2}{3}$.

The bright green and red regions show no change in agreement between the plurality vote and Borda count from the case of $r_P = 1$. The light red and blue regions show increased disagreement with the Condorcet winner; which means, for the blue regions, increased disagreement with the Borda winner as well; the peak densities occurring on the $P(A,C) - P(A,B)$ plane and $P(B,C)$ axis (where r_U is minimized; note that r_U takes the same values at those two peaks). Note that eventually, these regions engulf the entire shell.

On any shell of $r_P \equiv k < 1$, we have that the relative density of the plurality vote agreeing with the Borda count, as opposed to the Condorcet winner, is lower. To complete the proof, one need simply refer to the contrapositive case of the intermediate value theorem and note that any ρ_P-type distribution is non-negative in the density it assigns to each value of r_P.

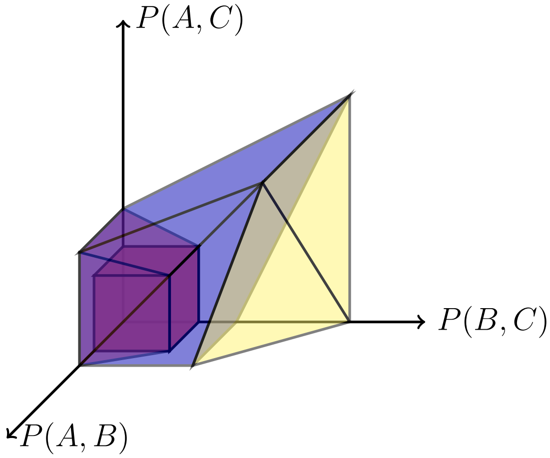
3.9 Condorcet winners and losers in an antiplurality vote

3.9.1 Condorcet winner

The following conditions (per the techniques of Chapter 2) determine the winner of an antiplurality vote:

- If $P(A, C)$ is the largest pairwise victory, A must win if $2P(A, C) - P(B, C) > 1$.
- If $P(A, C)$ is the smallest pairwise victory, A must win if $2P(A, B) - P(A, C) > 1$.
- If $P(A, B) > P(A, C) > P(B, C)$, A must win if $2P(A, B) - P(B, C) > 1$.
- If $P(B, C)$ is the largest pairwise victory, B must win if $2P(B, C) - P(A, C) > 1$.
- If $P(A, C)$ is the largest pairwise victory, C can win if $2P(A, C) + P(B, C) > 1$.
- If $P(A, B)$ is the largest pairwise victory, C can win if $2P(A, B) + P(B, C) > 1$.
- If $P(B, C) > P(A, C) > P(A, B)$, C can win if $2P(B, C) + P(A, C) > 1$.
- If $P(B, C) > P(A, B) > P(A, C)$, C can win if $2P(B, C) + P(A, B) > 1$.
- It is possible for any rank ordering of candidates to occur if and only if $P(X, Y) < \frac{1}{3}$ for all (X, Y).

Figure 3.16: Regions of interest under an antiplurality vote



It is worth taking special note of the fourth of these conditions. There is a portion of the representation cube where B is at an advantage in an antiplurality vote, and accordingly, this sector contributes a significant probability that a Condorcet winner loses. The sixth and seventh conditions are therefore not relevant to the question of how likely a Condorcet winner is to win an antiplurality vote, as when $P(B, C)$ is largest and $> \frac{1}{3}$, the ranking $C \succ B \succ A$ can occur, but not the ranking $C \succ A \succ B$.

The region with multiple possible winners, which we will call Q, is shown in in blue in Figure 3.16. It is useful to defining a radius-like quantity r_Q on the surface of this region Q, and divide Q into piecewise parts Q_1 and Q_2. The yellow region adjacent to Q is the region where B must win *not* favored to win on the entire surface of the representation cube. The region on the right we will refer to as R with associated radius-like quantity r_R. (We have also re-oriented the axes, as this perspective shows the shape a little better.)

Here it is useful to define ρ_2 as the density of B defeating A and ρ_3 as the density of C winning while A defeats B. Thus, we divide up the calculation of ρ^* as follows:

$$\frac{\int_T \rho^* dV}{\int_T \rho dV} = 12\left(\frac{1}{12} - \int_{Q_1} \rho_2 - \int_{Q_2} \rho_2 - \int_R \rho_3\right) \tag{3.79}$$

Note that the sub-region where $P(B, C)$ is the largest has a volume of $\frac{1}{4}$. Integrating over r_P using the density function $12(1-r_P)^2$ gives:

$$\int \rho dV = \int_0^1 12(1-r_P)^2 \frac{3}{4} r_P^2 dr_P = \frac{3}{10} \tag{3.80}$$

Which is a $\frac{3}{10}$ probability. This region can be divided in halves with volume $\frac{1}{8}$ each by the boundary where A defeats B with zero density. The density with which A defeats B increases quadratically to 6 at the origin (half the peak probability density value of 12 at the origin).

$$\int_{Q_1} \rho - \rho_2 dV = \int_0^1 6(1-r_Q)^2 \frac{3}{8} r_Q^2 dr_Q = \frac{3}{40} \tag{3.81}$$

Combining these gives a total probability of $\frac{3}{40}$ for the case where B defeats A while $P(B, C)$ is the largest pairwise victory. We have a $\frac{7}{10}$ probability of landing elsewhere in the octant. We can then calculate the total proability that B defeats A (but $P(B, C)$ is not the largest pairwise victory):

$$\int_{Q_2} \rho_2 dV = \int_0^1 6(1-r_Q)^2 \frac{5}{8} r_Q^2 dr_Q = \frac{1}{8} \tag{3.82}$$

We should now consider the case where A defeats B but loses to C. The region R in which this occurs has volume $\frac{19}{324}$, so in terms of a tailored radius r_R we have a differential volume element $\frac{19}{108}$:

$$\int_R \rho_3 dV = \int_0^1 2(1-r_R)^2 \frac{19}{108} r_V^2 dr_V = \frac{19}{1620} \tag{3.83}$$

This gives us everything we need to compute how frequently A wins:

$$\frac{\int_T \rho^* dV}{\int_T \rho dV} = 1 - \frac{9}{40} - \frac{1}{8} - \frac{19}{1620} = \frac{517}{810} \tag{3.84}$$

Which is to say about 63.8%. The majority of the exceptions happen when $P(B, C)$ is the largest pairwise comparison.

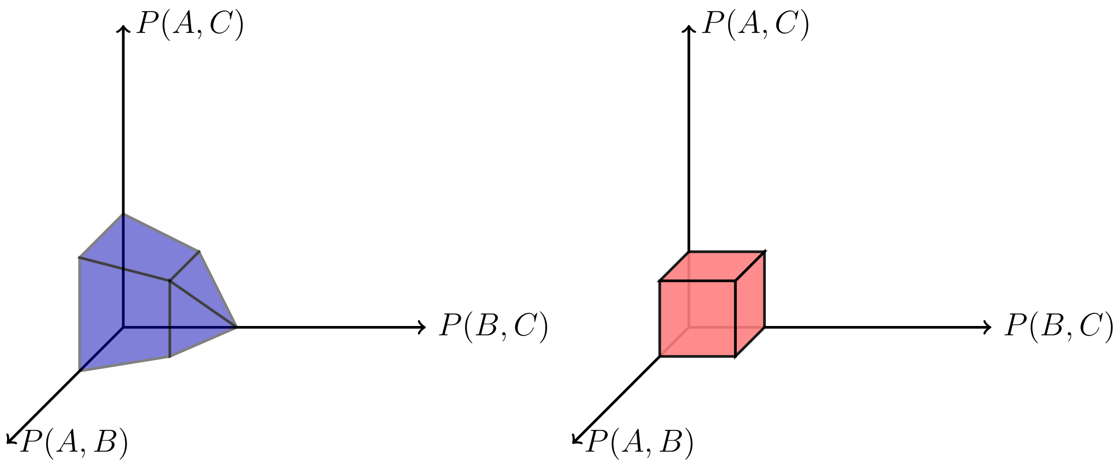
3.9.2 Condorcet loser

The conditions which apply to a Condorcet loser winning an antiplurality vote are:

- If $P(A, C)$ is the largest pairwise victory, C can win if $2P(A, C) + P(B, C) > 1$.
- If $P(A, B)$ is the largest pairwise victory, C can win if $2P(A, B) + P(B, C) > 1$.
- If $P(B, C) > P(A, C) > P(A, B)$, C can win if $2P(B, C) + P(A, C) > 1$.
- If $P(B, C) > P(A, B) > P(A, C)$, C can win if $2P(B, C) + P(A, B) > 1$.
- It is possible for any rank ordering of candidates to occur if and only if $P(X, Y) < \frac{1}{3}$ for all (X, Y).

These conditions give the boundaries for two regions: An outer region Q in which C can place first by surpassing the candidate who would place first if the reversal component were zero, and an inner region R where C can defeat that candidate but still lose. As in the previous sections, this leads directly to the definition of two surfaces, which we will label Q and R, respectively.

Figure 3.17: Condorcet loser winning an antiplurality vote (left)
Any ranking possible (right)



As before, we construct ρ^*, the density function of interest, from a pair of easier-to-compute functions:

$$\rho^* = \rho_2 - \rho_3 \tag{3.85}$$

Here ρ_2 represents the density of C defeating the candidate who would place first without a reversal component (B if $P(B, C)$ is the largest pairwise comparison, A otherwise), and ρ_3 represents the probability of them doing so but losing. Q is the support of ρ_2 and R the support of ρ_3, and r_Q and r_R can be conveniently used to show their density functions. Note that the volume of Q is $\frac{59}{972}$, so this gives a differential volume element of $\frac{59}{324} r_Q^2 dr_Q$, while the volume of R is $\frac{1}{27}$. We can then integrate ρ_2 and ρ_3:

$$\int_Q \rho_2 dV = \int_0^1 6(1 - r_Q)^2 \frac{59}{324} r_Q^2 dr_Q = \frac{59}{1620} \tag{3.86}$$

$$\int_Q \rho_2 dV = \int_0^1 2(1 - r_Q)^2 \frac{1}{9} r_Q^2 dr_Q = \frac{1}{135} \tag{3.87}$$

So our probability is:

$$\frac{\int_T \rho^* dV}{\int_T \rho dV} = \int_Q \rho_2 dV - \int_R \rho_3 = \frac{47}{1620} \tag{3.88}$$

This is about 2.9% - surprising as that may be, it is only slightly less likely than a Condorcet loser winning a plurality vote, in spite of the significant difference in probabilities of a Condorcet winner winning under the IAC assumption. This will vary slightly with different ρ_P-type distributions.

3.10 Conclusions

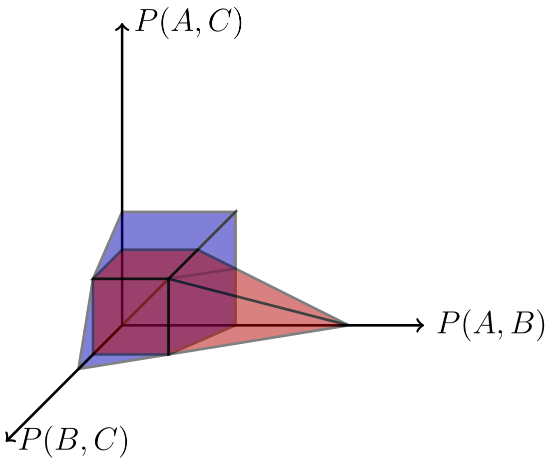
3.10.1 Discussion: r_P and r_S distributions, [anti]plurality, and Condorcet rankings

IAC is a special case of r_P-type distribution, in particular one in which the distribution in supporting space is uniform and can easily be reduced to a function of a single variable. However, our method of calculating the probabilities for IAC has clear extension to other r_P-type distributions, and moreover, we can tell which portions of the space contribute to the probability of a Condorcet winner or loser winning or losing an election.

For example, we can see readily that while the probability of a Condorcet loser winning under IAC is similar for both plurality and antiplurality are quite close (2.96% and 2.90%, respectively), the two cases rely on a different underlying geometry.

Both share in common the $(0, \frac{1}{3})^3$ inner region; but a Condorcet loser can win a plurality

Figure 3.18: Comparison of Condorcet loser under plurality with Condorcet loser under antiplurality



vote even when r_P is very close to 1; a Condorcet loser cannot win an antiplurality vote when $r_P > \frac{2}{3}$. Lower the variance of an r_P-type distribution in the representation cube, and the advantage a Condorcet loser has in a plurality vote over an antiplurality vote will be reduced or even reversed. Reduce the variance in the supporting space, and this will generally disadvantage the Condorcet loser, who is under no circumstances advantaged without a reversal component. Reversal components *must* align in favor af a Condorcet loser in order for a Condorcet loser to win; basic profile components disfavor Condorcet losers.

For Condorcet winners, the story is different, and in particular different for different rules. For a plurality vote, a Condorcet winner benefits by having a high variance in the distribution on the representation cube and also from low variance in the supporting space. The less of a role the supporting space plays in a plurality vote, the more likely a Condorcet winner is to win, as the basic profile components dominate. If the supporting space is irrevelant, the Condorcet winner wins 100% of the time. For an antiplurality vote, however, the effect of basic components is more ambiguous. The limiting value of a Condorcet winner as the supporting space plays a smaller and smaller role is in fact 77.5%, and while this is higher than with the IAC distribution, the Condorcet winner receives a much smaller benefit from reducing the variance of the supporting space or increasing the variance of the density distribution in the representation cube.

The antiplurality vote is simply less friendly to a Condorcet winner. The differences become smaller as distributions concentrate nearer the origin, and in particular as we move away from the outer edges and corners of the representation cube - which is also one consequence of shifting to an r_S-type distribution. Closer to the origin, the behavior of plurality and antiplurality become more similar.

3.10.2 The larger picture

There are several significant results that should be emphasized in addition to the computational results. The most important development within this chapter is the demonstration of an intuitive method of calculating the probability of a Condorcet paradox under a given probability distribution. While the specific results of 6.25% for IAC itself and 8.8% for large numbers of voters and IC are known, we can extend these figures to large numbers of other related distributions, and also provide a method of swiftly computing the results for most possible probability distributions.

Another key result is that there are certain figures which are highly insensitive to choice of probability distribution; that is to say, nearly every "natural" choice of probability distribution gives us an identical probability that a profile is *strongly* transitive (75%), and an incredibly wide variation in probability distributions all give us a similar small chance that a Condorcet winner loses a Borda count (9-10%). What this tells us is that if we observe rates outside this range, some very basic assumptions common in the literature are wrong; and it also tells us that we cannot meaningfully distinguish between probability distributions by observing those particular rates.

Predictions of Condorcet upset probabilities with plurality and antiplurality votes are, on the other hand, sensitive to choice of distribution. While a large class of distributions underlies any particular observation of (for example) the probability with which a Condorcet winner wins, It is also worth noting that the antiplurality vote has significantly more difficulty in meeting the Condorcet winner criterion, and that this rate is *not* especially sensitive to the distribution of reversal components though antiplurality votes violating the Condorcet loser criterion *is* highly dependent on the presence of reversal components.

This is not true of all combinations, and some, such as the combination of a Condorcet winner and a plurality vote, have a predicted frequency that varies wildly based on different probability distributions. In spite of this, the techniques of this chapter illustrate a potential hazard in trying to guess the probability distribution of profiles based on plurality vote outcomes: For intermediate figures, there are a wide range of very different probability distributions that give the same probability of a Condorcet winner losing a plurality vote.

Chapter 4

A Geometric Model of Manipulation and Error in Single and Multiple Stage Elections

4.1 Introduction

In the study of voting systems, a common concern is whether or not particular systems are vulnerable to strategic action. From the Gibbard-Satterthwaite theorem (Gibbard [1973] and Satterthwaite [1975]), it is known that nearly all voting systems[1] become vulnerable to strategic action when there are three or more candidates - in theory. It has been suggested that *in practice*, the Borda count is less vulnerable (see saa and Forsythe et al. [1996]) to strategic action than other positional methods, but others have also suggested that in practice, it encourages insincere voting (in the sense used in Bassi [2014], insincere voting is a subset of strategic voting). It is also well known that multiple stage systems (runoffs) have

[1] The principal exceptions being a dictatorship, and a lottery system where a random voter is selected to be the single decisive voter.

difficulties with non-monotonicity, which provides an incentive for strategic voting. Some have even suggested that runoff elections are less likely (e.g., Niou [2001]) to inspire strategic behavior from voters. The element of strategic voting is one of the chief reasons for advocacy both for and against runoff elections.

This chapter uses a geometric model to explore the interaction between positional voting rules used in a given stage, whether that voting stage is final or intermediate (to be followed by another election), and the depth of vulnerability to a certain type of strategic action. By *certain type of strategic action*, we are referring to a comparison between the results with a given electorate, and the results given the addition of a fixed share ϵ of votes controlled by an agent who has full information about the other votes cast. In addition to providing an estimate of the power of good information about other voters' votes, this model can also be applied to the case of accidental error, random variations in turnout, small shifts in the electorate, or the fraudulent addition of a small number of ballots during the counting stage. this chapter does not address large-scale strategic manipulation on a platform or ticket level.

Under the most basic assumptions, in which results of the voting rule are uniformly distributed, the model in question suggests parity between particular single and multiple stage voting rules. It also exhibits a set of symmetries which show that a single stage election using a plurality vote is most closely related to a multiple stage election using an antiplurality vote to eliminate one candidate at a time. Conversely, an antiplurality vote carried out in a single stage is most closely related to a multiple stage election using a plurality vote to eliminate one candidate at a time. More sophisticated treatment bringing in a class of more plausible probability distributions will suggest that having more stages generally makes a system less vulnerable to manipulation or error; this breaks the symmetries described above.

The model also shows the Borda count to be considerably less vulnerable to small-scale manipulations, particularly as the number of candidates n increases, in line with saa. There is little difference between a single and multiple stage Borda count under the assumption of a uniform distribution of results; however, as with all other positional methods, more plausible probability distributions suggest a significant difference between a single stage election using a Borda count and a multiple stage election using a Borda count.

4.2 Contributions to the literature

Studies of actual voters voting in actual elections can and have been conducted which ask voters how they would have voted with a different electoral system (for a professional society election, see Saari [2001] & Brams and Fishburn [2001]; and for a political election, see Laslier and Van der Straeten [2008]). There is a long list of difficulties faced by this empirical literature:

- Voters may vote strategically more often, less often, or in a qualitatitively different way if the voting rule is different.
- Voters may express preferences based *post hoc* on their strategic vote in the election at hand, rather than write down what is effectively an admission of strategic voting on their ballot.
- Voters may collectively learn how a voting rule works for a given electorate over multiple elections; thus, the general behavior of a given body of voters under an alternate electoral system may not emerge immediately even after actual elections, making voters' forecasts of their own hypothetical behavior less likely to be accurate.
- Different voting rules may alter the dynamics of the *campaign* leading up to the election.
- In some cases, voters may fail to understand a voting system that is not currently being used, and not be willing to invest time into thinking seriously about how to vote in that system when there is little reason to.

- It is very difficult to obtain the information necessary to begin an analysis of a 3+ candidate election under alternate voting rules; thus, most studies examine a relatively small number of elections. In such an environment, it is difficult to understand what appears to be a general case and what is particular to a given election; and extremely difficult to test hypotheses about probability distributions of profiles of voters.

Some of these difficulties are amenable to study in the laboratory.Strategic behavior has been noted in voters in a laboratory setting (see Forsythe et al. [1996] Bassi [2014] Blais et al. [2007] Blais et al. [2010] directly, and Palfrey [2009] for an overview of multiple experiments). Past experiments have generally been restricted to three candidates, limiting the ability to measure the strategic response, especially for elections with multiple stages, and have often used a small number of very specific profiles of voter preferences (as in Forsythe et al. [1996]).

The model developed in this chapter is useful for several reasons. First, it provides a theoretical framework under which choice of probability distribution is relatively easy to manipulate; and under which it can be seen that some results emerge not simply for a couple of particular probability distributions, but for a large class of probability distributions meeting a simple set of criteria. Second, it will offer some perspective in systematically examining the dynamic of strategic behavior. The model used in this chapter shows that a Duverger's Law type dynamic emerges naturally in a plurality system as the product of a feedback loop between information and intention to engage in strategic voting; but that other voting rules will show a different dynamic.

4.3 Result space

To begin with, we consider the space of possible outcomes of a given voting rule over a given number of candidates. This could be viewed as a discrete set, containing the set of possible rankings, but doing so is trivial. We will take result space to be a subset of $\mathbb{R}^n$, and in particular normalized to be a subset of $[0,1]^n$; each axis gives each of the n's candidates share of points, relative to the maximum number of points they could receive. (If votes are assigned a rational weight and the voting body is finite, the actual result space is a subset of $(\mathbb{Q} \cap [0,1])^n$, but this distinction is not useful in the present work.)

This produces two general classes of voting rules. For convenience, when referring to the result space of a particular *rule* over n candidates, we will refer to the normalized result space for the general case for that voting rule. Every specific election has a particular result space associated with a particular election and electorate; it has N voters, each with a weight w_i, leading to a grid of possible results. By normalizing, we restrict our analysis to a familiar space; and by examining the general case, we avoid dealing with specific integer effects tied to some particular value of N or combination of weights. (For example, even-odd effects related to ties.) Note that for some theoretical scenarios involving infinite voters, the space cannot be normalized; we are excluding such cases *a priori* from our analysis. This defines a continuous subspace of $[0,1]^n$, R_{rule}, for each voting rule.

Each particular positional voting rule (plurality, anti-plurality, "vote for 2," etc) has a result space contained within a minimal continuous object R_{rule} of dimension $n-1$ within $\mathbb{R}^n$. Some multiple positional rules (such as an approval vote) have a result space R_{rule} corresponding to a non-planar object of dimension n. To give a number of specific examples:

For an **approval vote**,in which a voter may choose to give each candidate either 0 or 1 point, , $R_{n,approval}$ is $[0,1]^n$. For any point $(x_1...x_n) \in [0,1]^n$ with rational points and common denominator dividing a number of voters N, we can construct that point by placing N voters in order, having the first x_iN of the voters approve of candidate i, and the next $(1-x_i)N$ disapprove. Since $[0,1]^n$ is the maximum possible result space, a proof is complete.

For a **range vote**, in which a voter may choose to give each candidate any number of points from 0 to 1, $R_{n,range}$ is also $[0,1]^n$. Again, for any $(x_1..x_n)$, we may construct this outcome by having a unanimous election where each voter gives candidate i points equal to x_i; summed over all voters, this averages out to x_i of the maximum possible vote.

For a **plurality vote**, where each voter votes for a single candidate (giving them 1 point and other candidates 0 points), $R_{n,plurality}$ is the standard unit simplex Δ^{n-1} of dimension $n-1$; or equivalently, the intersection of $[0,1]^n$ with the plane of points whose coordinates sum to 1. We can see this by noting that every outcome must lie inside the convex hull of unanimous outcomes; and each unanimous outcome is one of the standard unit vectors.

For a **cumulative vote**, where each voter divides up a single point between as many candidates as they like, $R_{n,cumulative}$ is the same as $R_{n,plurality}$. For any point in the standard unit simplex, $(x_1..x_n)$, we have that $\sum x_i = 1$, which means that if a voter assigns x_i points to each candidate i, the voter has cast a complete cumulative vote; so we can construct the result $(x_1..x_n)$ with a unanimous electorate by having every voter assign x_i points to the ith candidate.

For an **antiplurality vote**, where each voter votes against one candidate, giving them 0 points and all other candidates 1 point, $R_{n,antiplurality}$ is a simplex that is the convex hull of $\mathbf{1} - e_i$, where e_i is one of the stardand unit vectors and $\mathbf{1}$ is the sum of all standard unit vectors. This is, as with the plurality vote, the convex hull of the n unanimous outcomes where voters unanimously vote against candidate n. Note that we may write, as a vector equation:

$$R_{n,antiplurality} = \{\mathbf{1} - x \mid x \in R_{n,plurality}\} \tag{4.1}$$

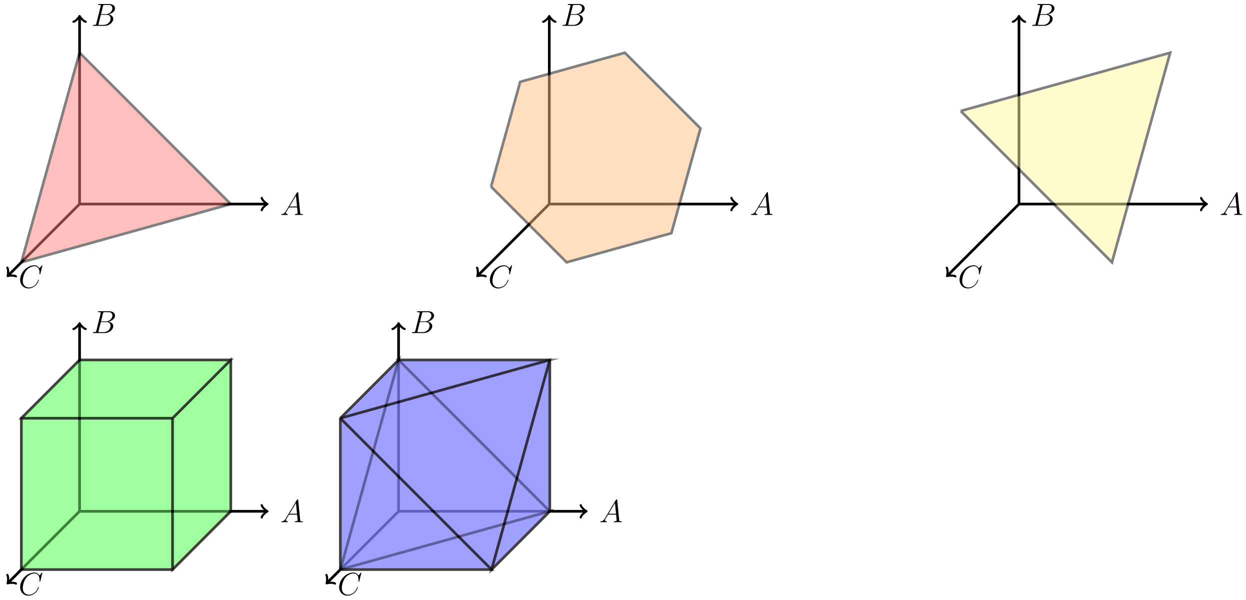
This tells us that $R_{n,antiplurality}$ is an inverted image of $R_{n,plurality}$. In particular, it is reflected through the point $\frac{1}{2}\mathbf{1}$.

For a voting rule that assigns 1 point to the candidates ranked above average, and 0 points to candidates below average, with a median candidate (for odd n) getting $\frac{1}{2}$ point, $R_{n,approvehalf}$ is equal to $\left\{x|\sum_{i=1}^{n} x_i = \frac{n}{2}\right\} \cap [0,1]^n$. Note that any result within this object (up to the rational restriction) can be constructed using a range vote. Moreover, it can be constructed with a unanimous electorate, whose average level of approval of candidates is $\frac{1}{2}$. Assuming the index is ordered such that $x_i \geq x_j$ when $i < j$, we begin at the center with the indices i, j closest to but not equal to $n/2$, and add $k_0 = \min\{x_i, 1 - x_j\}$ ballots which assign 1 point to all candidates with index $\leq i$, assign 0 points to candidates with index $\geq j$, assigning a $\frac{1}{2}$ 1/2 point to a candidate with index $n/2$ if such exists. Then we define $x_{i,1}$ to be the ith largest value of $\{x_j - k \mid j \geq n/2\} \cup \{x_j \mid j \leq n/2\}$, re-indexing our remaining needed total, and repeat this process. This algorithm completes to a legal "approve half" set of ballots after at most n steps.

For a **Borda count**, $R_{n,borda}$ is a subset of $R_{n,approvehalf}$. Note that there are $n!$ unique permutations, and thus $n!$ unique unanimous electorates; $R_{n,Borda}$ is thus the convex hull of the set of points of form $(\frac{m_1}{n-1}...\frac{m_2}{n-1})$, where $\{m_i\}_{i=1..n} = \{0,..,n-1\}$. For $n = 3$, this convex hull is the whole of $R_{3,approvehalf}$; for $n \geq 4$, this is a strict subset of $R_{n,approvehalf}$. (For a particular point of reference, consider $(0, 0, 1, 1)$; it is not in $R_{4,Borda}$, as at most one candidate may have a normalized score of 1 (being then top-ranked on all ballots unanimously), and at most one candidate may have a normalized score of 0 (being then bottom-ranked on all ballots unanimously).

When we talk about an undefinined number of voters N and make $R_{n,rule}$ a continuous object, as in the above treatment, results tend to apply very accurately when the number of voters is large, and especially when it is both large and not fixed; though it is less accurate with a fixed small number of voters. The analysis produced by the continuous approximation remains *qualitatively* useful, however, even in those cases. As exhibited above, the result space is not necessarily unique to a given rule. In scenarios where the number of voters is

Figure 4.1: Result spaces for plurality, Borda count, antiplurality, approval, and restricted approval

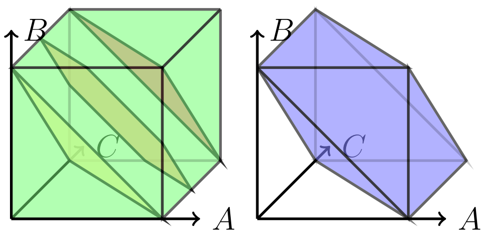


large enough that a continuous result space is a reasonable approximation, the difference between a cumulative vote and plurality vote is not very different when it comes to manipulating the results. (Such differences that exist in this model will exist solely in terms of probability distributions on result space.) We have listed seven distinct rules, but only four unique result spaces for $n = 3$. A fifth space emerges with the "non-trivial" or restricted approval vote - an approval vote where voters must approve at least 1 and at most $n - 1$ candidates. Those five $n = 3$ result spaces are shown in Figure 4.1

Reflecting the C axis gives a useful perspective on the three planar result spaces relative to the approval cube (left in Figure 4.2), and also gives a better look at the restricted approval vote result space (right in Figure 4.2).

Note that the restricted approval vote is the convex hull of the plurality and antiplurality result spaces. It is also worth noting that positional and multiple positional voting rules,

Figure 4.2: Result spaces compared



once normalized appropriately, all have the same characteristic *type* of result space; that result space a subset of the intersection of $[0,1]^n$ with a band of possible point totals. In the case of a plurality vote, that total is 1, meaning that the result space is the intersection of the level set of 1 total vote with the $[0,1]^n$ cube; in the case of an antiplurality vote, it is $n-1$; in the case of a Borda count, it is $\frac{n}{2}$.

This relationship goes both ways. Each level set corresponds to a class of positional and multiple positional rules where a voter has that many points to cast (once first place is normalized to 1 and last place is normalized to 0); while each band from one level set to another level set corresponds to a class of multiple positional rules where a voter can cast a number of points ranging from the first level set's value to the second level set's value. These classes can be further distinguished in a discrete treatment; but from within a continuous approximation (and therefore when the number of voters is fairly large), these classes are treated as equivalence classes of voting rules.

4.4 Regions of influence

The next key tool is the *region of influence.* This is a subset of the result space, tied to a value ϵ representing the proportion of the total vote in play.

Definition 4.1 (Region of influence for n candidates)**.** *Given a positional or multiple positional rule, a number* $\epsilon \in (0,1)$*, and a number of candidates* n*, the region of influence* $I_{\epsilon,n,rule}$

is the subset of $R_{n,rule}$ in which the results could be altered by the addition of ϵ votes.

We restrict our definition to $\epsilon \in (0,1)$ because cases outside of $(-1,0) \cup (0,1)$ are trivial, and for $\epsilon < 0$, we are dealing with subtraction of votes rather than addition of votes. We will also define the region of influence over a particular set S as being:

Definition 4.2 (Region of influence over a set S of candidates)**.** *Given a positional or multiple positional rule, a number $\epsilon \in (0,1)$, and a set S of candidates the region of influence $I_{\epsilon,S,rule}$ is the subset of $R_{n,rule}$ in which ϵ votes can be added to produce an outcome where the members of S are ranked in any order.*

The *indifference set* over a set S of candidates, a related concept, is:

Definition 4.3 (Indifference set)**.** *Given a particular voting rule, a space of possible outcomes, and a set of candidates S, the indifference set $I_{S,rule}$ is the set of possible voting outcomes where all candidates in S have the same ranking. These may be familiar as a generalization of the indifference planes described in Section 2.4.*

We use I to represent both sets, because we have that necessarily:

$$I_{S,rule} \subsetneq I_{\epsilon,S,rule} \tag{4.2}$$

Moreover, for any set X such that for each $\epsilon > 0$:

$$X \subsetneq I_{\epsilon,S,rule} \tag{4.3}$$

We have that:

$$I_{S,rule} = X \tag{4.4}$$

That is, $I_{S,rule}$ is the unique subset common to $I_{\epsilon,S,rule}$ for each $\epsilon > 0$. Subset operations give us relations common to $I_{S,rule}$, and over $I_{\epsilon,S,rule}$:

Lemma 5. *If $S \subseteq T$, then $I_{S,rule} \supseteq I_{T,rule}$.*

Likewise:

If $S \subseteq T$, then $I_{\epsilon,S,rule} \supseteq I_{\epsilon,T,rule}$.

To prove this, note that if $S \subseteq T$, and all members of T are tied, as they are within $I_{T,rule}$, then all members of S, being members of T, are also tied, so we are within $I_{S,rule}$. Then, to address the region of influence, suppose that P is an ordering of of S, a set of m candidates and a subset of T. Then the ordering P^* such that P^*'s first m terms are equal to P, followed by all members of $T - S$, is an ordering of T, and the ordering P^* produces the order of P on the set S. So any point within $I_{\epsilon,T,rule}$, an agent with ϵ votes can produce P^* on T, and therefore P on S.

Lemma 6. $I_{\cup_i S_i,rule} \subseteq \cap_i I_{S_i,rule}$

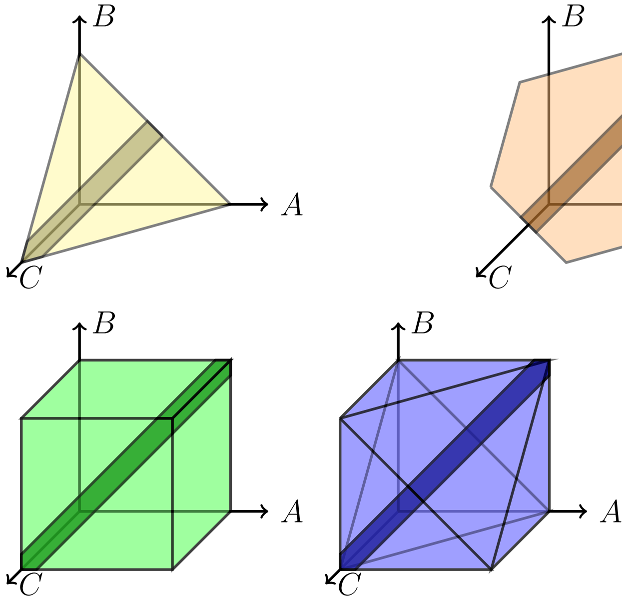
Likewise:

$I_{\epsilon,\cup_i S_i,rule} \subseteq \cap_i I_{\epsilon S_i,rule}$

This follows from the previous lemma by finite induction; we need only recall the basic set-theoretic fact that $A \subseteq B$ and $A \subseteq C$ together imply that $A \subseteq B \cap C$.

The basic building blocks are regions of influence over pairs of candidates. These are the maximally sized regions of influence; the intersections of those sets bound regions of influence over triplets and larger sets. Regions of influence over pairs of candidates are also the only regions of influence of scale ϵ, as opposed to ϵ^2 and higher order terms; thus, if ϵ is small, accurate approximations of the vulnerability of a system to manipulations of order ϵ need only consider regions of influence over pairs.

Figure 4.3: Regions of influence over $\{A, B\}$ in, respectively, plurality, Borda count, antiplurality, approval, and restricted approval result spaces



4.4.1 Motivating examples

To illustrate the regions of influence over pairs, we go to $n = 3$ candidates and each of the five distinct result spaces we have constructed for $n = 3$. Figure 4.3 shows $I_{0.1,\{A,B\},plurality}$, $I_{0.1,\{A,B\},borda}$, $I_{0.1,\{A,B\},antiplurality}$, $I_{0.1,\{A,B\},approval}$, and $I_{0.1,\{A,B\},restrictedapproval}$

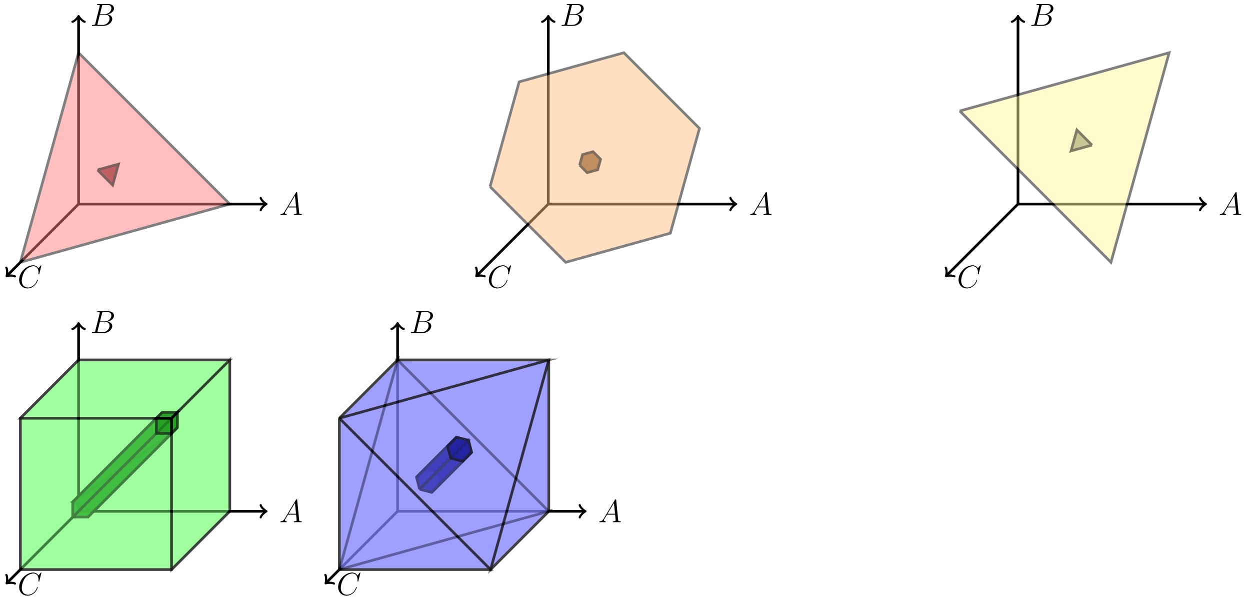
Compare these with the equivalent regions of influence over $\{A, B, C\}$ in Figure 4.4

For sets of 2 candidates, $\{X, Y\}$, we may state that:

$$I_{\epsilon,\{X,Y\},rule} = I_{\epsilon,\{X,Y\},approval} \cap R_{rule} \tag{4.5}$$

Essentially, this is a consequence of normalization; after normalization, each rule offers an equal opportunity for an added set of ϵ votes to give ϵ points to one candidate in the set while giving 0 points to the other candidate in that set. It is, however, worth using the case of a region of influence over two candidates to demonstrate the systematic construction of a region of influence.

Figure 4.4: Regions of influence over $\{A, B, C\}$ in, respectively, plurality, Borda count, antiplurality, approval, and restricted approval result spaces



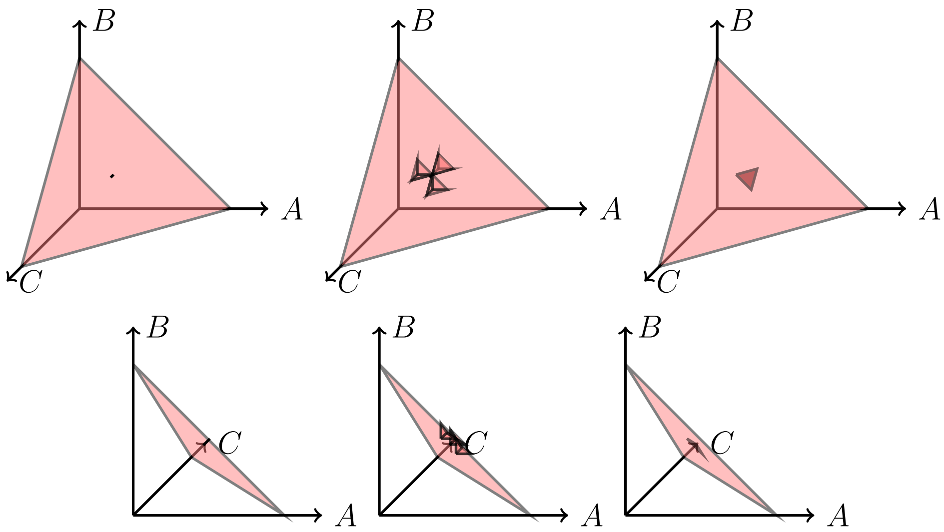
The construction is as follows:

$$I_{\epsilon,S,rule} = \bigcup_{x \in I_{S,rule}} \left\{ y \in R_{rule} \mid \exists z \in R_{rule} \exists k \in \mathbb{R} \left(y + \epsilon z = x + \frac{\epsilon}{n}(k, .., k) \right) \right\} \tag{4.6}$$

Specifically, k in the above equation will be a number in the set of possible total number of points given by a ballot under the rule in question. So for a plurality vote, $k = 1$; for an antiplurality vote, $k = n - 1$. For convenience, we will refer to the maximal such k as k_{rule}. To explain this process as a step by step algorithm:

- Consider a target point within R_{rule}, $x = (x_1, .., x_n)$.
- An equivalent outcome is given by $x' = (x_1 + \frac{\epsilon k_{rule}}{n}, .., x_n + \frac{\epsilon k_{rule}}{n})$ for any particular choice of k.
- For each point $y \in R_{rule}$, let Y be the set given by the possible addition of ϵ votes. (Y is then the ϵ-scale image of R_{rule}, translated to $(1 + k\epsilon)R_{rule}$).
- Define E_x as follows: $y \in E_x$ iff $x' \in Y$. X is then the set of results within R_{rule} that

Figure 4.5: Constructing a region of influence



can become equivalent to x with the addition of ϵ votes.

- The union of all such E_x over every choice of x in the indifference set over S is the region of influence.

An important fact about E_x that allows for better conceptual understanding is that it is a reflected, ϵ-scaled, and translated image of R_{rule}. k in the above process will be a number of points that can be legally assigned by the rule

A visual illustration of this process is shown in Figure 4.5

If the result space R_{rule} is an $n-1$ dimensional object within the n dimensional space of candidate outcomes, the indifference set over the set S containing all n out of n candidates is necessarily a point, which in turn means that the region of influence $I_{\epsilon,S,rule}$ is a reflected and $\epsilon-$scaled image of the original result space. We may also state that:

$$I_{\epsilon,\{X,Y,Z\},rule} \subseteq I_{\epsilon,\{X,Y,Z\},approval} \cap R_{rule} \tag{4.7}$$

This follows directly from Lemma 6 and Eq. 4.5. That this subset relation is not always equality is illustrated directly by the plurality and antiplurality examples.

4.5 A simplistic analysis of the approval vote

As exhibited in Eq. 4.7, we know that analysis of the approval vote (or, equivalently in this treatment, the range vote) gives us very interesting information about all other positional and multiple positional methods. Exploring this further, we may very directly compute the exact size of $I_{\epsilon,S,approval}$ for any S and an approval vote over n candidates. Letting s be the number of elements of S, we have that:

$$\|I_{\epsilon,S,approval}\| = s\epsilon^{s-1} - (s-1)\epsilon^{s} \tag{4.8}$$

Similarly, we can count the number of such distinct regions of influence as $\binom{n}{s}$; so the total volume of all regions of influence over sets of size s in an approval vote is approximately:

$$\binom{n}{s} s\epsilon^{s-1} \tag{4.9}$$

If we have that, for some particular s:

$$\epsilon << \frac{s+1}{sn - s^{-2}} \tag{4.10}$$

Then:

$$\binom{n}{s+1}(s+1)\epsilon^{s} << \binom{n}{s}(s)\epsilon^{s-1} \tag{4.11}$$

Implying that the sum total area of regions of influence over sets of 3 candidates is considerably smaller than the sum total area of regions of influence over sets of 2 candidates. Note that the right half of Eq. 4.10 is increasing in s. So if Eq. 4.10 holds for any particular s, it holds for all larger s. Inductively, then, if we have the simple base case of:

$$\epsilon << \frac{3}{2}\frac{1}{n-2} \tag{4.12}$$

Then we have that the regions of influence over 2 candidates are significantly larger than all other regions of influence. So in the case that Eq. 4.12 holds, it is appropriate to focus almost entirely on regions of influence over 2 candidates; and further, the total volume within non-trivial regions of influence is approximately:

$$2\binom{n}{2}\epsilon^1 = n(n-1)\epsilon \tag{4.13}$$

This is an overestimate, with higher order (e.g., ϵ^2) terms reducing this volume. On the other hand, the areas where multiple regions of influence overlap are also regions in which that influence can be exercised in more directions, and hence it is far more likely that this influence can be exercised to produce a more *desirable* outcome; so if this is a good approximation of how likely someone controlling an ϵ-share of the vote is *able* to influence the outcome, then:

$$\frac{n(n-1)\epsilon}{2} \tag{4.14}$$

is an excellent (that is to say, *better*) approximation of the probability that such an agent would be able to influence the outcome in favor of their own preferences; or that a randomly acting agent *would* cause the result of an election to change. This total volume increases roughly linearly with ϵ, and roughly quadratically in n. The total volume of the result space is 1; if we assume a uniform distribution over possible results (not a particularly realistic assumption) the total volume of the regions of influence is also a measure of total probability.

4.5.1 Approval vote, runoff, uniform distribution

We may now compare what happens if we conduct, instead of a single approval vote over n candidates, a sequence of approval votes. We will assume that the distribution of results is independently uniform over the candidates in each stage. (This is an exceptionally unrealistic assumption for typical scenarios, but useful for the purpose of creating a baseline case.) During each stage, one candidate is eliminated. During the kth stage, we have eliminated $k-1$ candidates, and there are $n-k+1$ candidates remaining. Thus, the total volume of the regions of influence is approximately:

$$(n-k+1)(n-k)\epsilon \tag{4.15}$$

Within each region of influence over a pair of candidates $\{X, Y\}$, outside of the subset that overlaps other regions of influence, we may divide the region of influence into $n-k$ distinct sub-regions: One where X and Y are in first and second place, one with X and Y in second and third place, all the way out to $n-k+1$th place and $n-k$th place. For an approval vote, each of these sub-regions is of equal size (this is not true of a plurality or antiplurality vote). Moreover, since each stage decides the elimination of a single candidate alone, rather than deciding a complete rank order of candidates, *only* the sub-region where X and Y are competing for last, i.e., $n-k+1$th, place is relevant to the final outcome. This is approximately $\frac{1}{n-k}$ of each region of influence over a pair of candidates; so the total volume in which the final results can be influenced is given by:

$$(n-k+1)\epsilon \tag{4.16}$$

Summing over $k=1$ to $k=n-1$ for all stages gives a simple finite summation:

$$\sum_{k=1}^{n-1}(n-k+1)\epsilon = \sum_{k=2}^{n} k = \frac{n(n+1)}{2} - 1 \tag{4.17}$$

This is strictly less than the quantity seen in Eq. 4.13. As n increases, Eq. 4.17 rapidly approaches one half of Eq. 4.13. It would appear as though introducing runoff stages makes an approval vote less easily manipulated. This is true, provided that one of two conditions holds. The first of those two conditions is that we are equally concerned with all place rankings. The alternate condition is that the probability distribution of results is more favorable (meaning more decisive more often) for elections involving smaller numbers of candidates. We will examine the first condition, as it is particularly relevant to elections, and also serves to further illuminate distinctions between plurality and antiplurality votes.

4.5.2 The special case of a unique winner

For many elections (or competitions), there is very little difference between a second-place finish and a third-place finish - or last-place finish - but a great deal of difference between a first-place finish or a last place finish. In the extremal (but not at all uncommon) case that a competition or election is designed to select a single winner, we may optimistically ignore most possible manipulations, concentrating our intentions strictly on those that chance the first-place winner. By the symmetry of the approval vote, we have that in a single stage election, approximately $\frac{1}{n-1}$ of each region of influence over pairs involves a difference between first and second place, giving us a total volume of:

$$\frac{n(n-1)\epsilon}{n-1} = n\epsilon \tag{4.18}$$

For a multi-stage vote, the reasoning becomes more complex. Note first that the size of $\mathcal{O}(\epsilon^2)$ terms suggests that in most cases where alteration of the results is possible, it occurs only in one key stage. If that key stage is the final ($k = n - 1$) stage, *any* alteration of the rank-ordering is decisive. However, for all previous stages, causing candidate X to be eliminated in place of candidates Y alters the results under a more narrow set of circumstances. In particular, the direct methods of doing so:

- Candidate X, who wins without the addition of ϵ votes, would otherwise survive all subsequent rounds, including the final round, winning out of the remaining $n-k$ candidates.

- Candidate Y who loses without the addition of ϵ votes, would subsequently survives all subsequent rounds, including the final round, winning out of the remaining $n-k$ candidates.

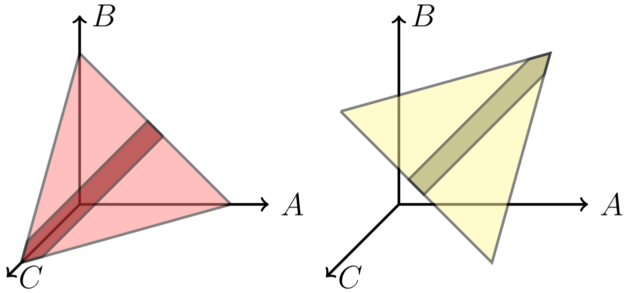
Considering the direct methods of influence, we divide Eq. 4.16 by $n-k$, which in turn implies a summation of:

$$\sum_{k=1}^{n-1} \frac{n-k+1}{n-k}\epsilon = n\epsilon + \epsilon \sum_{k=2}^{n-1} \frac{1}{k} \tag{4.19}$$

This is close to Eq. 4.18; it is higher, although ratio between the two is no higher than 7:6. The basic *possibilities* for direct manipulation are comparable; the question is balancing the expected positive correlation in the electorate from stage to stage (which decreases the likelihood of earlier eliminations affecting the eventual first-place winner) against indirect manipulation for $n \geq 4$. This is, in some sense, a negative result: This analysis shows a significant difference in manipulability only in the case where the complete ranking matters, but not in the case where only a first place winner matters.

This is important as a base case. We will next consider what happens with systems other than an approval vote.

Figure 4.6: Regions of influence over $\{A, B\}$ in plurality (left) and antiplurality (right) result spaces



4.6 Plurality and antiplurality, uniform distribution

For plurality and antiplurality votes, the indifference sets over pairs of candidates are simplexes of dimension $n-2$, within the result space that is a simplex of dimension $n-1$. To recall the regions of influence, see Figure 4.6.

The ratio between the original result space and $I_{\epsilon,S,plurality}$ over a pair is then approximated, to within $\mathcal{O}(\epsilon^2)$, by the ratio between simplex volumes:

$$\frac{\|I_{\epsilon,S,plurality}\|}{\|R_{plurality}\|} = (n-1)\epsilon + \mathcal{O}(\epsilon^2) \tag{4.20}$$

Which means that the total region of influence is approximately given by:

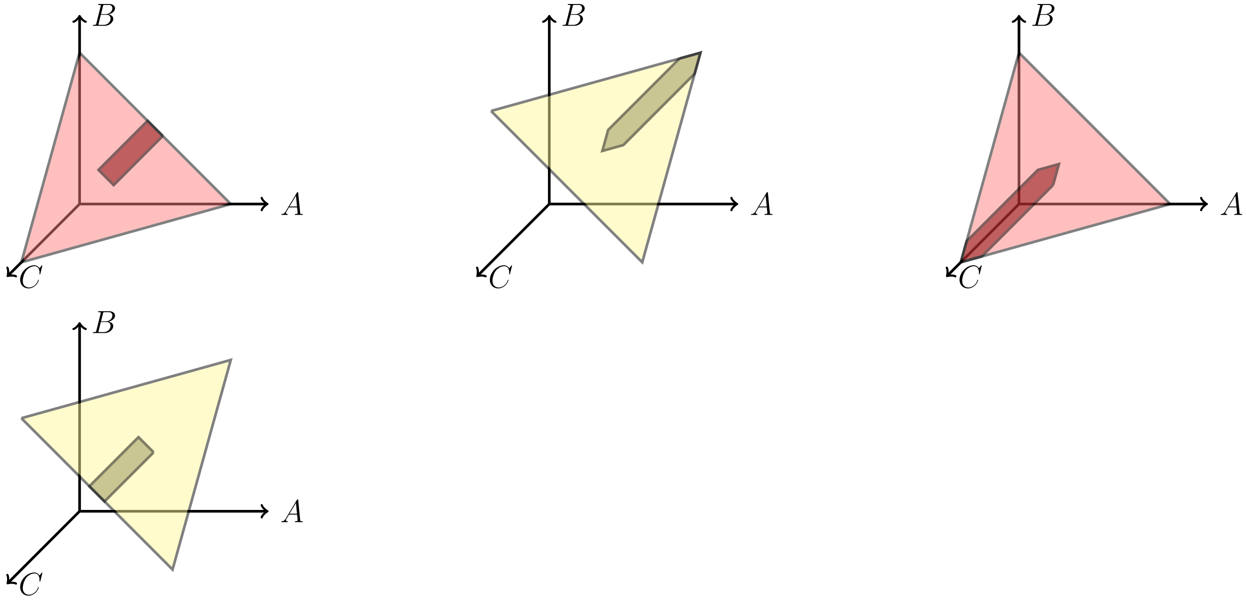
$$\frac{\|\bigcup_S I_{\epsilon,S,plurality}\|}{\|R_{plurality}\|} = n(n-1)^2\epsilon \tag{4.21}$$

These two figures apply, identically, to an antiplurality vote. However, the symmetry present in the regions of influence for an approval vote is absent. Combining with Eq. 4.13, we have that:

Lemma 7. *The region of influence over S with an approval vote is approximately $\frac{1}{n-1}$ of the region of influence over S in a plurality or antiplurality vote.*

This suggests, but does not prove, that as n increases, plurality and antiplurality votes gain

Figure 4.7: First-place and last-place subregions of influence over $\{A, B\}$



a signficant amount of resistance to small-scale manipulation and small errors. In order to examine the case for a multi-stage election, we must specifically examine the regions of influence over different place ranks. In particular, we will focus on the region of influence over first place (important to the case of a unique winner) and over last place (important for a maximized series of elimination stages).

From left to right in Figure 4.7: First place in a plurality vote, first place in an antiplurality vote, last place in a plurality vote, last place in an antiplurality vote. Note that these all include, as a subregion, the region of influence over $\{A, B, C\}$. This is an antisymmetric relationship; the last place region in an antiplurality vote corresponds to the first place regions in a plurality vote, and *vice versa*. In the case of $n = 3$, these subregions are approximately 1/3 and 2/3 of the region of influence over a pair of candidates.

The indifference set over $\{A, B\}$ for a plurality vote over n candidates is an $n-2$-dimensional object (for $n = 3$, a line segment)

For general n, in a plurality vote, we have that the indifference set over $\{A, B\}$ in last place is given by the convex hull of the universal indifference point (where all candidates receive $\frac{1}{n}$ votes) and the following set of $n-2$ points:

$$\{(0, 0, x_1, .., x_{n-2}) \mid \exists k(x_i = 1 \leftrightarrow i = k \& x_i = 0 \leftrightarrow i \neq k)\} \tag{4.22}$$

The set in 4.22 plus the indifference point forms an $n-2$-simplex. The volume of an $n-2$-simplex is given by:

$$\frac{\sqrt{det(W^T W)}}{(n-2)!} \tag{4.23}$$

Where W is the matrix of column vectors of form $v_i - v_0$ for $i = 1..n-2$. In this particular case, taking the indifference point as v_0 gives:

$$W = \begin{pmatrix} -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & \cdots & -\frac{1}{n} \\ \frac{n-1}{n} & \ddots & \vdots \\ -\frac{1}{n} & \ddots & -\frac{1}{n} \\ \vdots & \ddots & \frac{n-1}{n} \end{pmatrix} \tag{4.24}$$

Which means that:

$$W^T W = \begin{pmatrix} \frac{n-1}{n} & \frac{-1}{n} & \cdots & \frac{-1}{n} \\ \frac{-1}{n} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{-1}{n} \\ \frac{-1}{n} & \cdots & \frac{-1}{n} & \frac{n-1}{n} \end{pmatrix} \tag{4.25}$$

Conveniently:

$$det(W^T W) = \frac{1}{n} \tag{4.26}$$

Which leads to the conclusion that:

$$\|I_{plurality,\{X,Y\}}\| = \frac{\sqrt{det(W^T W)}}{(n-2)!} = \frac{1}{(n-2)!\sqrt{n}} \tag{4.27}$$

The first order approximation of the volume of the subregion of influence over last place is then:

$$\|I_{\epsilon,\{X,Y\},plurality,last}\| = \frac{\sqrt{n-1}\epsilon}{(n-2)!\sqrt{n}} \tag{4.28}$$

Note that the volume of the result space is given, similarly, by:

$$\|R_{plurality,n}\| = \frac{\sqrt{n}}{(n-1)!} \tag{4.29}$$

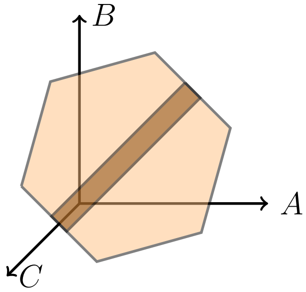
The ratio of this subregion to the whole space is then:

$$\frac{\|I_{\epsilon,\{X,Y\},plurality,last}\|}{\|R_{plurality,n}\|} = \frac{(n-1)^{3/2}\epsilon}{n} \tag{4.30}$$

Compare this with Eq. 4.20, which is linear in n: What this tells us is that the share of the result space taken up by the subregion of influence over last place is persistently large, taking up an increasingly disproportionate share of the total region of influence. If the subregions were equal, the ratio between 4.30 and 4.20 would be $n-1$. Correspondingly, the subregion of influence over *first* place shrinks significantly as n increases, taking up a similarly disproportionately small share of the region of influence. These leads immediately to a large set of results:

Theorem 4.1. *1. In the case where an election leads to a unique winner, rather than a ranking outcome, a plurality vote, conducted in a single stage, offers fewer possibilities for manipulation than an antiplurality vote, conducted in a single stage.*

Figure 4.8: Region of influence over $\{A, B\}$ in a Borda count



2. *In the case where an election is carried out in multiple stages, an antiplurality voting rule in each stage offers fewer possibilities for manipulation than a plurality voting rule.*

3. *An antiplurality voting rule $n-1$ stage election over n candidates, eliminating one candidate per stage, offers fewer possibilities for manipulation than a single stage election using a plurality voting rule.*

4. *A plurality voting rule $n-1$ stage election over n candidates, eliminating one candidate per stage, offers fewer possibilities for manipulation than a single stage election using an antiplurality voting rule.*

This theorem extends from possibility to probability if we add certain assumptions relating the probability distribution of results for different voting rules.

4.7 Borda Count and the complete $n = 3$ comparison

For $n = 3$, $I_{\epsilon,\{X,Y\},borda}$ is shown in Figure 4.8.

This region has an area of:

$$\|I_{\epsilon,\{A,B\},borda}\| = \sqrt{3}\epsilon \tag{4.31}$$

Note, incidentally, that this is only an $\mathcal{O}(\epsilon^2)$ term away from the area of the region of influence over the same pair for a plurality or antiplurality vote. The result space, however, is larger for the Borda count:

$$\|R_{3,borda}\| = \frac{3\sqrt{3}}{4}\epsilon \tag{4.32}$$

Which leads to a smaller ratio:

$$\frac{I_{\epsilon,\{X,Y\},borda}}{\|R_{3,borda}\|} = \frac{4}{3}\epsilon \tag{4.33}$$

Recall that the corresponding figure for a plurality or antiplurality vote is given by:

$$\frac{I_{\epsilon,\{X,Y\},plurality}}{\|R_{3,plurality}\|} = 2\epsilon - \mathcal{O}(\epsilon^2) \tag{4.34}$$

This is, in other words, quite substantially less. As a first-order approximation of the total region of influence, we have that:

$$\frac{\bigcup_{\{X,Y\}} I_{\epsilon,\{X,Y\},borda}}{\|R_{3,borda}\|} = 4\epsilon - \mathcal{O}(\epsilon^2) \tag{4.35}$$

The Borda count is symmetric with regard to first and last place; so in the event that we are only concerned with first place manipulations, we have that:

$$\frac{\bigcup_{\{X,Y\}} I_{\epsilon,\{X,Y\},borda,first}}{\|R_{3,borda}\|} = 2\epsilon - \mathcal{O}(\epsilon^2) \tag{4.36}$$

Up to $\mathcal{O}(\epsilon)$ terms, this is comparable to the subregions of influence over first place in a plurality vote, which represent roughly one third of the total region of influence for a plurality vote at $n = 3$:

$$\frac{\bigcup_{\{X,Y\}} I_{\epsilon,\{X,Y\},plurality,first}}{\|R_{3,plurality}\|} = 2\epsilon - \mathcal{O}(\epsilon^2) \tag{4.37}$$

This suggests the following result:

Theorem 4.2. *For $n = 3$ candidates, if ϵ is small, the following holds:*

1. *Under a uniform probability distribution over possible results, a Borda count with majority runoff has regions of influence that differ only by a small amount from a single stage election using a Borda count*

2. *A Borda count result space has first-place subregions of influence that differ only by a small amount from the first-place subregions of influence for a positional or multiple positional method that distributes between 1 and $\frac{3}{2}$ normalized points per ballot; and by a larger amount for positional or multiple positional methods that distribute $\geq \frac{3}{2}$ normalized points per ballot.*

3. *A Borda count result space has smaller regions of influence, relative to its result space, than any other positional or multiple positional method.*

Part 1 is proven by the above calculations. The above calculations also demonstrate the specific case of plurality and antiplurality for parts 2-3. To extend this to the whole continuum of possible results is fairly simple.

To prove part 2, we note first that for $n = 3$ all positional and multiple positional rules that are normalized (i.e., a maximum value of 1 for a first place candidate, and minimum value of 0 for a last place candidate) produce a result space that is simply a cross-section of the approval result space at the level set of k points. From $k = 1$ to $k = \frac{3}{2}$, the area of these cross sections increases linearly; from $k = \frac{3}{2}$ to $k = 2$ it decreases linearly; and from $k = 2$ to $k = 3$ quadratically.

At the same time, the size of the indifference subsets presenting ties between first and second place increases linearly from $k = 1$ to $k = 2$. The base case of a plurality vote shows that the ratio of the size of the indifference sets to the size of the result space is equal at $k = 1$

and $k = \frac{3}{2}$; since this is a ratio of two linearly increasing components, it must be monotone, which gives us equality throughout this span. For $k > \frac{3}{2}$, we have a decreasing area of the result space as the indifference subsets increase in size, which necessarily means that the ratio between the indifference subset affecting first place and the result space increases. This extends, up to $\mathcal{O}(\epsilon^2)$ terms, to the regions of influence.

To prove part 3, we use a similar argument: The indifference sets remain of constant size from $k = 1$ to $k = 2$. Moving away from this band with $k < 1$ ($k > 2$), the indifference sets decrease in size linearly, while the cross-sectional area decreases quadratically as k decreases (increases, for the $k > 2$ case). The ratio of cross-sectional size to indifference set size is therefore maximized at $k = \frac{3}{2}$, where the cross-sectional area is maximized.

4.8 Dynamics

One of the key results of this chapter is taking note of the strong symmetry between an $n-1$-stage plurality vote over n candidates, and an antiplurality vote. In the above, the reasoning is quantitative and static. We considered the result space as it is, with a single strategic agent who has a single opportunity to attempt to change the results.

Elections are frequently preceded by a period during which voters form opinions about candidates, and also consider predictions about how other voters are likely to vote (a *campaign*, in other words). This is a dynamic process, and strategic voting in reality occurs within this dynamic process. Different voters have, at different points in a campaign, different amounts of information. Voters support candidates based on a combinaiton of preferences, viability, et cetera.

Does the choice of voting rule affect this dynamic? It seems that it *ought* to. There can be a feedback cycle between information about the likely outcomes of an election, and voters' intentions. In particular, Duverger's Law (see Riker [1982]) can be viewed as a dynamic result: Voters perceive a candidate as non-viable, and make plans to vote for a more viable candidate. This, in turn, makes the candidate look less viable, leading eventually to removal of that candidate from the set of considered candidates entirely.

In this section, we will move from a single stage of manipulation to an iterative process, which can be viewed as simulating either a dynamic exchange between voting intentions and information over time, or increasing level-k information about the rest of the electorate.

4.8.1 A response heuristic

Formally speaking, rational evaluation of information requires careful evaluation of error structures. In practice, the typical error structure will be closely associated with a normal or lognormal distribution. When information is of good quality, there will generally only be one region of influence over one pair of candidates which contains a non-trivial weight of probability. If we have an agent with ϵ share of the vote (with ϵ small), who views all candidates as having non-trivial differences, the rational utility-maximizing behavior of that agent closely approximates a very simple heuristic:

- Identify the nearest indifference set over a pair of candidates.
- Cast an ϵ share of the vote, maximizing the component of vote orthogonal to that indifference set.

Figure 4.9 illustrates this process.

From left to right in Figure 4.9, we have a projected election result; the nearest indifference set; and then the two vectors which maximize moving towards or away from that indifference

Figure 4.9: Process of voting strategically, plurality vote

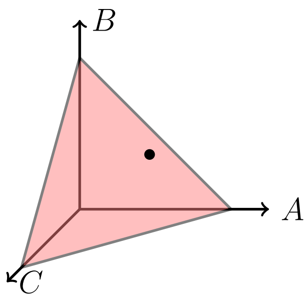
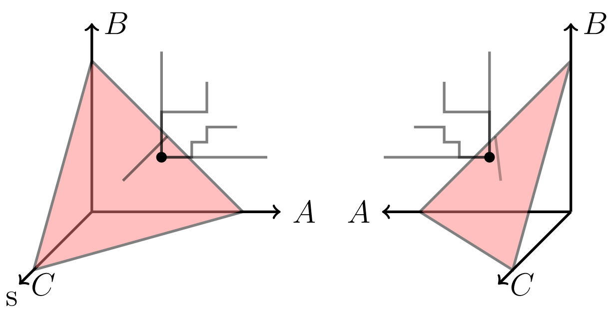


Figure 4.10: Iterated strategic votes



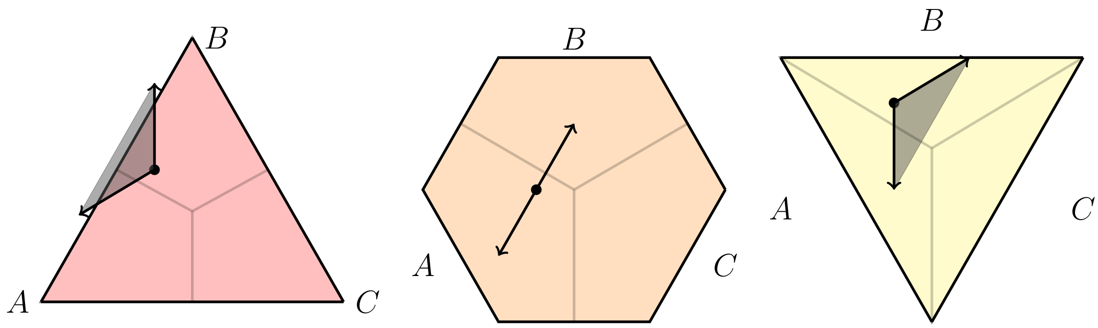
s

set. Note that we have illustrated a plurality vote. If we iterate this process, we construct a path; this can be viewed as a sequential response to increasing information, changing information, or k-level responses.. There are several interesting features of the paths in question. (Figure 4.10 shows several sample paths, viewed from two different perspectives.)

The salient boundary (nearest indifference set) remains the same regardless of how many voters are added. This, in turn, implies that the strategic responses remain the same - in the case of a plurality election, to vote for one of the two most popular candidates. It is also worth noting that the strategic response is to vote in alignment with existing voters, rather than against it. This behavior holds for higher n, and is more striking in that environment. This dynamic reproduces Duverger's Law.

Note that the same dynamic - with the same illustrations - applies to any given stage of a multi-stage election using an antiplurality vote and eliminating one candidate at a time.

Figure 4.11: Process of voting strategically
Plurality, antiplurality, and Borda Count



Next, we consider the case of a single-stage election using an antiplurality voting rule, displayed in Figure 4.11.

In the case of a plurality vote, the strategic responses move inward. Agents acting to make B defeat A, or vice versa, add support to C along the way. This may eventually bring the projected result closer to the $\{A, C\}$ or $\{B, C\}$ indifference set, at which point the appropriate strategic response shifts, as well. Strategic voters are encouraged to vote differently from how the projected population votes. Anti-coordination leads to an unstable dynamic, diverging to different possible results depending on the order in which voters respond to information about the projected outcome of the election. Again, the geometric symmetry of the situation dictates that this instability also applies to multi-stage elections run with a plurality vote, eliminating a single candidate at a time.

Instability, interestingly, leads to strong motive for deception if strategic voters are numerous.

We now consider a Borda count, also illustrated in Figure 4.11.

The dynamic of the Borda count is unstable, but marginally so. Rather than heading towards the origin, it tends towards orbiting at a fixed radius from the axis of universal indifference. A large shift in one or the other direction can bring the projected result closer to another indifference set; albeit considerably more slowly than with an antiplurality vote.

4.9 Conclusions

The models used in this chapter has implications for comparing the various positional and multiple positional voting systems. In particular, they tell us about two different forms of *stability*. In a static sense, we can think of stability as resistance to manipulation by a small number of well-informed voters, and resistance to random errors. In a dynamic sense, we can think of stability in terms of whether a dynamic feedback cycle between information and voting intentions involves significant changes; and also, whether or not it tends to lead to a stable equilibrium.

This model shows a number of intriguing relationships between systems, and between the two types of stability (dynamic and static). A plurality election has a strong kinship to a Coombs rule. An antiplurality election has a strong kinship to a plurality vote with successive elimination stages. A Borda count is somewhat akin to an approval vote if voters approve and disapprove symmetrically, but especially close kin to the positional rule where the top half of candidates get points. Plurality votes and Coombs elections show coordination between voters to vote similarly to one another in a dynamic setting (leading to a Duvergian-type dynamic), and also have less vulnerability to manipulation of first place rankings in a static setting; while antiplurality votes, and plurality-based runoff systems, show anti-coordination in a dynamic setting and a strong chance of manipulation of first-place winners.

The Borda count performs relatively well in all theoretical settings, both in a single stage and as a rule for multi-stage elimination elections; it does not stabilize dynamically in a Duvergian fashion. For $n = 3$, it and other positional and multiple positional methods that share in common with a Borda count assigning $\frac{n}{2}$ all minimize the static vulnerability to small-scale manipulations; and it can be strongly suggested that somewhere in this class of rules is the rule that minimizes static vulnerability to small-scale manipulations.

In general, elections involving fewer candidates are better behaved. Classically, this is exhibited in the difference between May's Theorem, for $n = 2$ candidates, and the various impossibility theorems for $n \geq 3$ candidates. A multiple stage rule takes advantage of this fact by dividing the regions of influence up between elections with a large number of candidates and elections with a small number of candidates; in spite of this, though, eventual first-place candidates are in theory vulnerable to elimination in an early stage, especially if the electorate itself shows more independent behavior from stage to stage.

Chapter 5

Concluding remarks

Ultimately, understanding of electoral dynamics requires both empirical understanding and a good theoretical framework. The preceding work addresses, in general terms, questions of probability and possibility regarding positional and multiple positional voting methods, and iterative voting over subsets of candidates. This began in Chapter 2, which provides a complete classification, for $n = 3$, of all possibilities for all positional methods compared to the collected votes over subsets of size $n = 2$ (pairwise votes). Pairwise votes are of particular interest, in that they are the simplest form of election. Condorcet winners, Condorcet losers, and the Condorcet paradox are determined in terms of pairwise votes. May's theorem applies to majority votes, and the question of which rule to use for a pairwise vote (majority vote) is a simple one.

Chapter 3 extends the work in Chapter 2 to probabilities, and also introduces probabilistic comparisons between different positional methods via the intermediary of pairwise criteria. One key feature of Chapter 3 is that it answers the question of what sort of probability distributions over profiles lead to what probability of various voting paradoxes; and in particular, this work tells us that it is very difficult to distinguish empirically between a very large class of probability distributions using the probability of a Condorcet paradox alone, or of any part of the relationship between pairwise votes and Borda counts, as these are highly insensitive to otherwise dramatic changes in probability distribution of profiles. Another key feature of Chapter 3 is that it shows that plurality and antiplurality do not lead to simply symmetric outcomes; although the methods are symmetric pairs in the most obvious way, pairwise criteria behave qualitatively differently, and in particular are much more strongly consistent with a plurality vote.

Chapter 3 also opens up probabilistic comparisons of results directly between different positional methods, not merely between positional methods and pairwise votes. It answers the question of how likely a plurality vote and Borda count are to agree, again contingent upon choice of probability distribution. We can again see from this comparison as well that the plurality and antiplurality votes are not simply mirror images of one another, but behave qualitatively differently when it comes to agreement with Borda votes as well as pairwise votes.

With Chapter 4, the $n = 3$ barrier of analysis is broken in looking at questions related to *stability* - stability with respect to small-scale manipulation by informed voters, small-scale errors, and also stability in terms of an informational dynamic between voters and polling. Chapter 4, in looking at this relatively narrow problem, does so for all n and *most* voting systems that have arisen in the literature, comparing positional methods and multiple positional methods over n candidates with successive stages of elimination rounds, each of which uses a positional or multiple positional method. Chapter 4 directly addresses *possibilities* and does not directly address the problem of variable probability distributions, although the question of likelihood can be expected to qualitatively follow the measures given for the spaces of possibilities. (Some discussion of the influence of probability distributions is found in an appendix.)

A key result of Chapter 4 is that plurality and antiplurality have a counter-intuitive antisymmetry when reflected across multiple rounds; a series of plurality vote elimination rounds is, with respect to strategic action, manipulation, and random error, more akin to an antiplurality vote carried out in a single stage than with a plurality vote carried out in a single stage. Chapter 4 does not answer whether or not this relationship will be visible in actual voting behavior, that is, that voters will recognize the symmetry; an outline for an experiment to test that hypothesis in a laboratory setting is contained within an appendix.

Bibliography

A. Bassi. Voting systems and strategic manipulation: An experimental study. *Journal of Theoretical Politics*, page 0951629813514300, 2014.

A. Blais, J.-F. Laslier, A. Laurent, N. Sauger, and K. Van der Straeten. One-round vs two-round elections: An experimental study. *French Politics*, 5(3):278–286, 2007.

A. Blais, S. Labbé-St-Vincent, J.-F. Laslier, N. Sauger, and K. Van der Straeten. Strategic vote choice in one-round and two-round elections: An experimental study. *Political Research Quarterly*, 2010.

S. J. Brams and P. C. Fishburn. A nail-biting election. *Social Choice and Welfare*, 18(3): 409–414, 2001.

R. Forsythe, T. Rietz, R. Myerson, and R. Weber. An experimental study of voting rules and polls in three-candidate elections. *International Journal of Game Theory*, 25(3):355–383, 1996.

W. V. Gehrlein and P. C. Fishburn. Condorcet's paradox and anonymous preference profiles. *Public Choice*, 26(1):1–18, 1976.

W. V. Gehrlein and D. Lepelley. Voting paradoxes and their probabilities. In *Voting Paradoxes and Group Coherence*, pages 1–47. Springer, 2011.

A. Gibbard. Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society*, pages 587–601, 1973.

J.-F. Laslier and K. Van der Straeten. A live experiment on approval voting. *Experimental Economics*, 11(1):97–105, 2008.

E. M. Niou. Strategic voting under plurality and runoff rules. *Journal of Theoretical Politics*, 13(2):209–227, 2001.

H. Nurmi. *Voting procedures under uncertainty.* Springer, 2002.

T. R. Palfrey. Laboratory experiments in political economy. *Annual Review of Political Science*, 12:379–388, 2009.

W. H. Riker. The two-party system and duverger's law: An essay on the history of political science. *American Political Science Review*, 76(04):753–766, 1982.

D. Saari. *Basic geometry of voting*, volume 12. Springer, 1995.

D. Saari. Disposing dictators. *Demystifying Voting Paradoxes Cambridge University, New York*, 2008.

D. G. Saari. A dictionary for voting paradoxes. *Journal of Economic Theory*, 48(2):443–475, 1989.

D. G. Saari. Explaining all three-alternative voting outcomes. *Journal of Economic Theory*, 87(2):313–355, 1999.

D. G. Saari. Mathematical structure of voting paradoxes. *Economic Theory*, 15(1):1–53, 2000.

D. G. Saari. Analyzing a nail-biting election. *Social Choice and Welfare*, 18(3):415–430, 2001.

D. G. Saari. Systematic analysis of multiple voting rules. *Social Choice and Welfare*, 34(2): 217–247, 2010.

D. G. e. a. Saari. Chapter twenty-seven-geometry of voting. *Handbook of Social Choice and Welfare*, 2:897–945, 2011.

M. A. Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, 10(2):187–217, 1975.

K. Sieberg and M. D. McDonald. Probability and plausibility of cycles in three-party systems: A mathematical formulation and application. *British Journal of Political Science*, 41(03): 681–692, 2011.

Appendix A

A.1 Supplemental proofs for Chapter 2

Proof of Thms. 2.1 – 4: Some assertions are proved in Sect. 2.3, others are special cases ($s = 0, 1$) of Thms. 2.6, 2.7. The exceptions are proved here. Assume Thm. 2.5 is correct.

To prove Thm. 1, parts 1 and 2, add α, β, γ values to each of the first three Fig. 2.2 essential profiles and determine what it takes to ensure the assertion. With Fig. 2.2a, because $P(A, C) \geq 0$, we have that $e_1 + e_2 \geq e_3$, or that the plurality tally for the essential profile has A at least tied for first. If the ranking is strict (i.e., $P(A, C) > 0$), then A is the sole plurality winner. Adding β terms improves A's tally at the expense of C. (B cannot gain enough advantage.) The same argument holds for Fig. 2.2b by interchanging B and C. A similar argument holds for Fig. 2.2c, where the $P(A, B) \geq 0$ assumption ensures A can be a plurality winner. If $P(A, B) > 0$, A is the sole plurality winner for the essential profile. Adding α terms helps A at B's expense.

The argument for Thm. 3 and the antiplurality rule involves minor differences. To prove Eq. 2.8, start with the AB essential profile where the A, B, C antiplurality tallies are, respectively,

$$e_1 + e_2 + e_3 + \alpha + \beta + 2\gamma, \quad e_1 + 2\alpha + \beta + \gamma, \quad e_2 + e_3 + +\alpha + 2\beta + \gamma. \tag{A.1}$$

By comparing tallies, for A to always beat B, it must always be that $e_2 + e_3 > \alpha - \gamma$, for A to always beat C, it must always be that $e_1 > \beta - \gamma$. So, for A to beat B for all supporting profiles, it must be that $e_2+e_3 = \frac{1}{2}[P(A,C)+P(C,B)]+\frac{1}{2}[P(C,A)+P(A,B)] = \frac{1}{2}[P(A,B)-P(B,C)]$ is greater than $q = \frac{1}{2}[n - P(A,B)]$. This requires $2P(A,B) > n + P(B,C)$, which is Eq. 2.8. The other constraint of $e_1 = \frac{1}{2}[P(A,B)+P(B,C)] > q = \frac{1}{2}[n-P(A,B)]$ provides the more relaxed $2P(A,B) + P(B,C) > n$.

The analysis for the AC essential profile is the essentially the same; the A, B, C antiplurality tallies are, respectively, $e_1 + e_2 + e_6 + \alpha + \beta + 2\gamma$, $e_1 + e_6 + 2\alpha + \beta + \gamma$, and $e_2 + \alpha + 2\beta + \gamma$. The sharpest constraint again involves ensuring that A always beats B; this requires $e_2 = \frac{1}{2}[P(A,C)+P(C,B)] > q = \frac{1}{2}[n-P(A,C)]$, which becomes the Eq. 2.8 condition $2P(A,C) > n+P(B,C)$. The condition ensuring that A beats C is the more relaxed $2P(A,C)+P(B,C) > n$.

To prove the Thm. 3 assertion that a BC essential profile makes it impossible for A to be the sole antiplurality winner, the A, B, C antiplurality tallies with the BC essential profile are, respectively, $e_1 + e_6 + \alpha + \beta + 2\gamma$, $e_1 + e_5 + e_6 + 2\alpha + \beta + \gamma$, and $e_5 + \alpha + 2\beta + \gamma$. By comparing the A and B tallies, for A to always be the sole antiplurality winner, it must always be that $\gamma > e_5 + \alpha$. Because $e_5 \geq 0$ and γ can be set equal to zero, this inequality cannot always be satisfied. But if $\gamma < e_5 + \alpha$ always is true, then (as asserted in Thm. 3), A *never* is an antiplurality winner. This holds if and only if $e_5 > q \geq \gamma$, or $\frac{1}{2}[P(B,C) + P(C,A)] > q = \frac{1}{2}[n - P(B,C)]$, or $2P(B,C) > n + P(A,C)$.

To prove Thm. 4 and Eq. 2.9, it is well known that a Condorcet winner (A) always is Borda ranked over a Condorcet loser (C). Thus, the only three strict Borda rankings are

$A \succ B \succ C, A \succ C \succ B$, and $B \succ A \succ C$. To show when the last choice cannot occur, compute A's and B's Borda tallies. Because a candidate's Borda tally is the sum of her tallies from all paired comparison elections, they are, respectively, $(P(A,B) + \frac{n}{2}) + (P(A,C) + \frac{n}{2})$ and $(P(B,A) + \frac{n}{2}) + (P(B,C) + \frac{n}{2})$. By setting A's tally over B's, canceling terms and using $P(B,A) = -P(A,B)$, Eq. 2.9 follows. □

Proof of Cor. 2: All settings satisfying Def. 2.4 are proved as described in Sect.2.3.1. Only the case not satisfying Def. 2.4 remains. From Fig. 2.2d, the plurality tallies for A, B, and C are, respectively,

$$e_1 + \alpha + \beta, \quad e_5 + \beta + \gamma, \quad e_3 + \alpha + \gamma.$$

The proof reduces to simple algebra; e.g., for A to always beat B, it must be for all α, β, γ that $e_1 + \alpha + \beta > e_5 + \beta + \gamma$, or $e_1 - e_5 > \gamma$. As the largest choice of γ is q, the condition becomes

$$e_1 - e_5 = \frac{1}{2}[P(A,B) + P(B,C) - P(B,C) - P(C,A)] > q = \frac{1}{2}[n - P(A,B) - P(B,C) - P(C,A)],$$

which is $2P(A,B) + P(B,C) > n$. All other assertions are similarly proved. □

Proof of Thm. 2.5: Results in Saari [1999, 2008] prove that only reversal pairs have no effect upon $P(X,Y)$ values. Start with any profile in the Fig. 2.1b form and remove reversal pairs; these are pairs where one entry is in region j, $j = 1, 2, 3$, and the other is in the diametrically opposite region $j+3$. Removing entries from each $\{j, j+3\}$ pair until one entry is zero has no effect on $P(X,Y)$ values; as the remaining entry is non-negative there are, at most, three positive entries.

This reduction defines eight possible arrangements; either all entries on one side of one of the perpendicular dividers are zero, or an arrangement arises where the zeros occur in an alternating fashion to define either $(n_1, 0, n_3, 0, n_5, 0)$ or $(0, n_2, 0, n_4, 0, n_6)$. All eight settings can be catalogued with Fig. 2.2; each case has a parallel arrangement where the e_j's in the figure are replaced by zeros and the zeros with e_j terms. That three of the sought after arrangements aregiven by Figs. 2.2abc follows from the $P(A,B), P(B,C), P(A,C) \geq 0$ assumptions; a direct computation proves that the parallel choices violate these conditions.

For the last two settings, the $(n_1, 0, n_3, 0, n_5, 0)$ choice is represented by Fig. 2.2d with appropriate n_j values; Fig. 2.2d clearly is compatible with the $P(A,B), P(B,C) \geq 0$ conditions. To show that the remaining $(0, n_2, 0, n_4, 0, n_6)$ choice is not compatible, the $P(A,B) \geq 0$ constraint requires $n_2 \geq n_4 + n_6$. But unless $n_4 = 0$ and $n_2 = n_6$ (which defines a Fig. 2.2b setting), this is incompatible with the $P(B,C) \geq 0$ condition that requires $n_6 \geq n_2 + n_4$.

To find the e_j values, each setting defines a linear algebra problem; e.g., Fig. 2.2a defines the set of three equations and three unknowns

$$n_1 + n_2 + n_3 = P(A,B), \quad n_1 - n_2 - n_3 = P(B,C), \quad n_1 + n_2 - n_3 = P(A,C), \quad \text{(A.2)}$$

with similar equations for each of the three remaining settings. As each set of equations is linearly independent, each set admits a unique solution. Thus, to show that these solutions are giving by Eq. 2.13, it suffices to show that the e_j values satisfy the equations. With the first of Eq. A.2, for instance, this becomes

$$e_1 + e_2 + e_3 = \frac{1}{2}[P(A,B) + P(B,C)] + \frac{1}{2}[P(A,C) + P(C,B)] + \frac{1}{2}[P(C,A) + P(A,B)].$$

Using the relationships $P(C,B) = -P(B,C)$, $P(C,A) = -P(A,C)$, it follows that the sum is $P(A,B)$ as required. All other sets of equations are verified in the same manner.

All remaining assertions in the theorem are immediate. □

Proof of Thms. 2.6, 2.7: This elementary algebra exercise follows the outline leading to Eq. 2.23. □

Proof of Cor. 4: These conditions are equivalent to determining the s values for which a linear equation is positive. To illustrate with $0 \leq s < \frac{1}{2}$, Eq. 2.26 can be expressed as determining where the linear equation $(1-s)[2P(A,B)+P(A,C)]-(1-2s)n-sP(B,C)$ is positive. But if a linear equation is positive at two points, it is positive for all values between these points. The conclusion follows because, with the possible exception of Eq. 2.30, all conditions are satisfied for the Borda Count $s = \frac{1}{2}$. Thus, if the conditions hold for $s_1 < \frac{1}{2}$ and $s_2 > \frac{1}{2}$, they hold for $s_1 \leq s \leq \frac{1}{2}$ and $\frac{1}{2} \leq s \leq s_2$, or for all s in $s_1 \leq s \leq s_2$. When Eq. 2.9 is satisfied (so the Borda and Condorcet winners agree), the assertion extends to Eq. 2.30. (If Eq. 8 is not satisfied, then a similar "clustering" assertion holds for s values that violate, rather than agree with, Eq. 2.30.) □

Proof of Thm. 2.8: Again, this is a direct algebraic computation following the lead of the example developed before the statement of the theorem. □

Proof of Thm. 2.9: This is a direct computation. □

A.2 Proof of Theorem 3.1

(3) follows immediately from Eq. 3.10, which defines r_P as the sum of each profile component of the essential component, and (4) follows directly from (3), the fact that in a normalized profile space:

$$p_{ABC} + p_{ACB} + p_{BAC} + p_{BCA} + p_{CBA} + p_{CAB} = 1 \tag{A.3}$$

Or, in a non-normalized space:

$$p_{ABC} + p_{ACB} + p_{BAC} + p_{BCA} + p_{CBA} + p_{CAB} = N \tag{A.4}$$

And that the profile equals the sum of its essential and reversal components, that is:

$$p_{XYZ} = e_{XYZ} + r_{XYZ} \tag{A.5}$$

To prove (1), we turn to Thm. 5 in Chapter 2, which gives a complete classification of essential profiles. In particular, each essential component has a collection of at most three nonzero profile components, which in no case includes both p_{XYZ} and p_{ZYX} simultaneously. If it did, it would be possible to remove an equal fraction of voters preferring $X \succ Y \succ Z$ and $Z \succ Y \succ X$, creating a smaller sub-profile with identical pairwise majority vote margins. The reversal component consists entirely of such pairs of voters, such that:

$$r_{XYZ} = r_{ZYX} \tag{A.6}$$

Applying Eq. A.6 to Eq. 3.9 gives us that:

$$|p_{XYZ} - p_{ZYX}| = |e_{XYZ} + r_{XYZ} - e_{ZYX} - r_{ZYX}| = |e_{XYZ} - e_{ZYX}| \tag{A.7}$$

Applying the fact that $(e_{XYZ}$ or $e_{ZYX}\}$ is zero, with the other being non-negative (possibly also zero), we may write that:

$$|e_{XYZ} - e_{ZYX}| = |e_{XYZ}| + e_{ZYX}| \tag{A.8}$$

This shows that Eq. 3.9 and Eq. 3.10 are equivalent. Conveniently, Thm. 5 in Chapter 2 also defines explicitly the formula for each component of the essential profile; we have, for each nonzero profile component p_{XYZ}:

$$|e_{XYZ}| = e_{XYZ} = \frac{P(X,Y) + P(Y,Z)}{2} = \left|\frac{P(X,Y) + P(Y,Z)}{2}\right| \tag{A.9}$$

Which three components are non-zero are determined (Thm. 5 from Chapter 2 again) by the pairwise majority vote values. If the essential profile is weakly cyclic or cyclic, with $P(A,B) > 0$ and $P(B,C) > 0$:

$$e_{ABC} + e_{ACB} + e_{CAB} + e_{CBA} + e_{BCA} + e_{BAC} + e_{ACB} = e_{ABC} + e_{CAB} + e_{BCA} \tag{A.10}$$

And for any X, Y, we have that :

$$P(A,B) + P(B,C) + P(C,A) \geq P(X,Y) \tag{A.11}$$

Applying Eq. A.9 to the classification of a cyclic or weakly cyclic profile in Thm. 5 of Chapter 2:

$$e_{ABC} + e_{CAB} + e_{BCA} = P(A,B) + P(B,C) + P(C,A) \tag{A.12}$$

This shows that Eqs. 3.9 and 3.10 are equivalent for the weakly cyclic and cyclic cases. In the strongly non-cyclic case, we can write that if $P(X,Y)$ is the largest pairwise value,

$$e_{ABC} + e_{ACB} + e_{CAB} + e_{CBA} + e_{BCA} + e_{BAC} + e_{ACB} = e_{XYZ} + e_{XZY} + e_{ZXY} \tag{A.13}$$

And that:

$$e_{XYZ} + e_{XZY} + e_{ZXY} = P(X,Y) \tag{A.14}$$

This then implies that Eqs. 3.9 and 3.10 are equivalent for strongly non-cyclic cases. We may also state another useful lemma, which applies in particular to the normalized version of the representation cube. To prove (2), first note that all of the above formulae are 0 when each coordinate variable is set to 0; and at the origin of both spaces, each coordinate variable is set to 0. To see that it is 1 on the surface of the representation cube, it suffices to note first that for any unanimous profile where $p_{XYZ} = 1$, $P(X, Y)$ is the (non-unique) maximum pairwise vote margin, so in a non-normalized space, we have at each unanimous vertex:

$$r_P = P(X, Y) = N \tag{A.15}$$

Or in a normalized space:

$$r_P = P(X, Y) = 1 \tag{A.16}$$

To show (2): Each unanimous profile occupies a vertex of the representation cube; each other point on the surface of the representation cube is the linear weighted average of three such points; and, within essential octohedral coordinates, lies within the same octant as every other point sharing the same three generating points. Since Eqs. 3.8 & 3.9 are linear if restricted to a single octant (with no sign changes in $p_{XYZ} - p_{ZYX}$), the values of r_P on each face must lie between the values r_P takes on the unanimous vertices surrounding that face, which is to say $r_P \equiv 1$ on the entire surface. This shows the bounds on r_P; to note that r_P is a metric, it suffices to note, first, that r_P can be recognized immediately as a Minkowski distance from the origin (with $p = 1$) in essential octohedral coordinates; and second, that the map from essential octohedral coordinates to tally space is a homeomorphism, so this property is preserved across coordinates. The map from profile space down to the representation cube is not one-to-one. In particular, has a three-dimensional kernel, or two-dimensional in the normalized case, but remains locally linear, which gives us a pseudometric.

A.3 An outline of an experiment testing the informational dynamics of Chapter 4

In Chapter 4, we outlined a mathematical model that explores the degree of potential vulnerability to manipulation; and creates qualitative distinctions between different types of strategic action. There are, in particular, some symmetries and antisymmetries in strategic action that are predicted across different positional rules and runoff systems. This leaves open the question of whether or not these symmetries in structure would be reflected in actual human behavior.

While the model addresses the whole class of positional and multiple positional methods, there are several particular results that are of some interest, in particular, the kinship between a plurality vote and an antiplurality runoff. These methods also should show the most dramatic differences in terms of convergent and divergent behavior.

A.3.1 Experiment design

The simple way to test this is using a 2×2 between-subjects factorial design. The two variables are:

1. Whether a (simulated) election is carried out in a single round of voting, or multiple rounds of voting (with one candidate eliminated per round).

2. Whether a (simulated) election is carried out using a *plurality* counting rule ("voting for") or an *antiplurality* counting rule ("voting against").

A between-subjects design is recommended, because otherwise, subjects may learn strategies for one voting system and apply them to another haphazardly; or become confused. Preliminary work suggests that subjects require some time to learn how voting systems other than a plurality vote work.

A.3.2 Hypotheses

Although a relatively simple experiment in design, this could test a large range of results from Chapter 4.

1. A plurality election, conducted in a single stage will show convergent strategic behavior, where voters converge on (usually two) candidates as they gain information about others' preferences; that is, the individual voter tends to try to coordinate with their peers. (This prediction is closely related to Duverger's Law.)

2. An antiplurality election, conducted in a single stage, will show divergent strategic behavior, where voters diverge in behavior, trying to vote against strong canidates; that is, the individual voter tends to try to anti-coordinate with their peers.

3. The effect of system type on the type of strategic behavior is reversed if candidates are eliminated one at a time; more plurality-like systems will show divergent behavior, while more antiplurality-like systems will show convergent behavior.

4. In multiple-round systems, the effects of system type on the type of strategic behavior is weaker, and less strategic behavior will be observed.

5. Divergent strategic behavior is associated with increased levels of attempts at deception and decreased credibility of polls, while convergent strategic behavior is associated with sincere elicitation of preferences.

6. A single-stage plurality vote, decided in a single stage, will show less strategic action than a single-stage antiplurality vote.

In particular, we expect these effects to be reinforced by learning over time, as subjects acclimate to the rule and dynamics thereof.

A.4 Building the relationship between stages

The type of analysis of Chapter 4 could be readily applied to any sequence of positional or multiple positional rules, with any sequence of probability distributions. We could at this point construct an example demonstrating nearly anything; which would prove nothing. In order to get meaningful results, we must find a way to insure that we compare apples to apples instead of apples to oranges. The main difficulty in this arises when we are comparing elections over different numbers of candidates.

We can very easily imagine mixing different voting rules with different stages; we can perform similar analyses with this. If I introduce a plurality stage after a Borda count, what does it tell me about the Borda count's interaction with runoffs? Formally speaking, a Borda count over four candidates is not the same rule as a Borda count over three candidates, just as it is not the same rule as a plurality count over three candidates. But the two Borda counts are related in some fashion; and we wish to exploit this relationship to describe what happens when we use the "same" rule from stage to stage. Both a $(1, 1, 1/2, 0, 0)$ rule and a

$(1, 3/4, 1/2, 1/4/, 0)$ rule are generalizations of a $(1, 1/2, 0)$ rule; and nearly every positional or multiple positional rule is a generalization of a majority vote over $n = 2$ candidates.

There is another formal difficulty we face: What about the underlying probability distribution on the space?

For example, we might assume subtotals are distributed uniformly when $n = 5$ and $n = 4$, but at $n = 3$ suddenly become concentrated around the edges, and at $n = 2$ tend to be concentrated in the middle. In such a case, minimizing sensitivity to small manipulations tells us we wish to actually, we want to avoid $n = 2$ stages (in spite of the fact that any such manipulation is unlikely to be classed as *strategic* in nature), and get great returns out of an $n = 3$ stage; but this is due to our choice of probability distributions, rather than something intrinsic to the space.

The analysis within Chapter 4 does not directly address the role of probability distributions over different numbers of candidates; it implicitly considers a uniform distribution over result space (which is unrealistic), but can be extended to cope with different probabilistic assumptions without much alteration.

There are several caveats that should be explicitly included. To this end, we will define the following four things:

- Related voting rules over m and n candidates.
- Corresponding voting rules over m and n candidates.
- Corresponding probability distributions over m and n candidates' results.
- Related probability distributions over m and n candidates' results.
- "Expanding" and "compressing" related probability distributions.

Our main results relate to the combination of *corresponding* voting rules and *expanding related* probability distributions.

A.4.1 Corresponding & related rules

It is intuitively clear that a plurality vote over five candidates is related to a plurality vote over six candidates; and that a plurality vote over five candidates is somehow the uniquely "best" way to capture the same spirit of a plurality vote over six candidates. (Moreover, we cannot simply assert *vica versa*; more on that later.) To formalize this, we

To formulate this more precisely, we first define a general class of related rules:

Definition A.1 (Related voting rules)**.** *For a voting rule α over n candidates and a voting rule β over $n+k$ candidates, we say that α and β are related rules if the set of possible results R_α for α over n candidates is the natural projection of a level set of k candidates in the set of possible results R_β for β down to n dimensional result space.*

In other words, α is related to β if α looks like what happens if we hold k candidates constant and keep using the rule β. This is not very restrictive; the trivial voting rule $(0, 0)$ is related to the three candidate plurality rule $(1, 0, 0)$. However, if we look at the class of voting rules related to a six-candidate plurality rule, we find that the Borda count is conspicuously absent; as is anything but the trivial voting rule and $(s, 0...0)$ rules, scale-equivalent to one or another plurality vote.

We are almost there. We want, however, to use a method that will get us a *unique* rule; and avoid trivial rules and the like. Refining our definition gets us this:

Definition A.2 (Corresponding voting rules, I)**.** *A voting rule α over n candidates corresponds to a voting rule β over $n+k$ candidates if α and β are related, and there is no other* *rule γ related to β such that for the corresponding result spaces, R_γ is larger than the result space R_α.*

That is to say, α preserves as much discrimination between candidates as possible from β. Since the result space is a convex hull of the possible unanimous voting vectors, there is a level set that is unique up to translation; so our corresponding rule is unique up to its result space. Note that some rules may have the same result space in the continuous approximation, e.g., cumulative voting and plurality voting; the lower dimensional corresponding rule is *nearly* unique. Existence of a corresponding rule is guaranteed.

The higher dimensional one is not; as most non-trivial voting rules correspond, at $n = 2$, to the majority vote. We actually want the majority vote to correspond to all rules; so we will add an additional rule to allow us to collapse a three-candidate approval vote to a two-candidate majority vote instead of a two-candidate approval vote; this comes at the cost of making the corresponding rule less unique:

Definition A.3 (Corresponding voting rules, II)**.** *Suppose a voting rule α over n candidates with result space R corresponds to a voting rule β over $n+k$ candidates. Suppose that α' is a rule which has a result space $R' = R \cap L$, where L is a level set where $\sum x_i = k$ for some constant k; and L is a level set which maximizes the area of $R \cap L$. Then we will say α' also corresponds to β.*

What is worth noting is that α' is constructed from β in the same way as α; we just have one extra projection step, which is carried out in much the same manner. A very simple theorem follows from fact that R_β is normalized and we are constructing related voting rules from maximal level sets of the result space. That is to say:

Theorem A.1. *If β is a rule over $n+k$ candidates whose result space is a q-dimensional subset of $[0,1]^{n+k}$ (usually $q = n+k$ or $q = n+k-1$), and α is a rule over n candidates* *corresponding to β, then the q-dimensional volume of R_β is no larger than the $q-k$ or $q-k-1$-dimensional volume of R_α.*

This follows immediately from the fact that R_α is the largest possible $q-k$ or $q-k-1$ dimensional level subset and the intermediate value theorem. What is less immediately obvious is that, for any positional or multiple positional rule, we have that:

$$\frac{\|R_{A,\epsilon,\alpha}\|}{\|R_\alpha\|} \leq \frac{\|R_{A,\epsilon,\beta}\|}{R_\beta} \tag{A.17}$$

When dealing with positional and multiple positional rules, we are not only dealing with a convex hull; but a convex hull with substantial symmetries, and with a fixed number of possible faces for each cross-section. If we take a level set fixing k candidates' vote totals, and shift those vote totals by some small vector v, one of three things happens.

1. All possible faces of the level set expand together. Some may "appear" out of a lower-dimensional edge.
2. It contracts symmetrically, with possibly some faces collapsing into vertices or lower dimension edges.
3. Distorts, with every other possible face moving outward and the others moving inward; since there is an even number of faces, this means that pairs of opposite faces will have one face moving inward and one outward.

When the length (area, 3-volume, etc) of the intersection between the indifference set and increases, the area (resp. 3-volume, 4-volume, etc) of the level set it is contained increases more quickly relative to its current value. This means that when we maximize the area, it happens at a minimum of the ratio of length to area (resp. area to 3-volume, 3-volume to 4-volume) between the indifference set intersected with the level set and the level set itself.

A.4.2 Corresponding & related probability distributions

We will define the relationship between probability distributions over n and $n+k$ candidates constructively, by creating a function $f : R_\beta \to R_\alpha$.

Definition A.4 (Corresponding probability distributions)**.** *Let α be a voting rule over n candidates and β be a voting rule over $n+k$ candidates.*
Let R_α and R_β be their respective result spaces.
Let ρ_β be a probability distribution over R_β.
Fix k particular candidates. Take each L that is a level set of R_β, such that L_x fixes those k candidates' vote totals, and the cumulative total if and only if R_α does so. L therefore consists of a slice of R_β parallel (in a sense) to an image of R_α, and with the same dimension.
Take $f_L : \partial L \to \partial R_\alpha$ such that if $x(t)$ is a constant speed loop on ∂L, $f_L(x(t))$ is a constant speed loop on ∂R_α.
Take ∂L_r for each such L to be a scalar image of ∂L multiplied by $r \in [0, 1]$, positioned such that ∂L_r has the same center as ∂L; and likewise ∂R_{α_r} and $x_r \in \partial R_{\alpha_r}$
Define $f(\partial L_r) = f(\partial L)_r$.
Define $\rho_\alpha(x) = \int_{f^{-1}(x)} \rho_\beta$.
Then ρ_α corresponds to ρ_β.

Given this constructive definition of a corresponding lower-dimensional probability distribution, we can naturally extend this to a looser definition of merely *related* lower-dimensional probability distribution. This modifies the above definition in several ways:

- We relax the requirement that the inner versions of the boundaries ∂L_r and ∂R_{α_r} are necessarily the same shape as the original boundary. We require only that it be a continuous family that takes each point $x \in \partial L$ on a straight line from the edge to the center as r goes from 1 to 0.

- We relax the radial requirement to $f(\partial L_r) = f(\partial L)_{g(r)}$ for a strictly increasing function $g(r)$.
- In particular, if $|f(x) - m|$, the distance between $f(x)$ and the center of the R_α, is concave viewed as a function along any straight line from the center of L to ∂L, the resulting ρ_g it is an *expanding* related distribution.
- Similarly, if $|f(x) - m|$ is convex, the resulting ρ_g is a *compressing* related distribution.

It is worth noting that a uniform distribution corresponds to a uniform distribution. It may be more useful use ∂L_r which rapidly become circular in terms of spherical shells rather than the actual shape of the result space, especially when (say) dealing with normal distributions.

A.4.3 Expanding and compressive distributions

What should be readily apparent is that the more concentrated results are near the center of the result space (nearer to an n-way tie), the more likely the result is to lie near a boundary, and thus the more easily the result may be altered. So a compressive chain of related probability distributions increases the chance of manipulation in later rounds; while an expansive chain of related probability distributions decreases the chance of manipulation in later rounds.

Whether having a runoff increases or decreases the amount of room for manipulation depends some, in other words, on the true probability distribution; which is not clear. If the chain of related probability distributions is expanding by any noticeable margin, then adding runoff stages nearly always reduces the opportunity (with the mainly the exception of plurality). If the chain of related probability distributions is compressive, however, there are no conclusive results available (except that adding a plurality runoff at $n = 3$ would increase vulnerability).

Arguments can be made in favor of both expanding and compressive distributions. It is my considered opinion that the argument for distributions that are more concentrated at higher n, i.e., leading to expansive chains of related probability distributions as n decreases in subsequent rounds, is the case with the stronger argument. It may be the case that different populations have different tendencies, and the ultimate answer to this question must be empirical, though this answer may not be easily available.

We can answer the question directly for Impartial Culture. In this case, the variance of the distribution in non-normalized result space is proportionate to $\frac{N(1-\frac{1}{n})}{n}$, where n is the number of candidates and N is the number of voters; normalization gives $\frac{(1-\frac{1}{n})}{n\sqrt{N}}$. Note that this is strictly decreasing with n, which means that if we reduce n, the variance increases. The assumption of Impartial Culture therefore gives us (approximately; to the degree that IC is approximated well by a normal bell curve) an expanding sequence of distributions.

We obtain a qualitatively similar result if we introduce a condition of dependence from round to round, and assume that all candidates keep their existing pool of support; with the former supporters of other candidates introducing noise (particularly suitable for discussing the plurality vote). For this reason, it is very natural to expect normalized variance to increase as n, the number of candidates, decreases, given the same population. It is ultimately this relationship over probability distributions (and over informational dynamics) that can drive a significant difference in the probability of the manipulation of first place rankings.