A NEW WAY TO ANALYZE PAIRED COMPARISONS

DONALD G. SAARI

Abstract. The importance of "paired comparisons," whether ordinal or cardinal, has motivated the creation of several methodological approaches; missing, however, is a common way to compare the various techniques. To do so, an approach developed here creates a coordinate system for the "data space." The coordinates capture which data aspects satisfy a strong cardinal transitivity condition and which ones represent "noise;" i.e., this second component consists of highly cyclic aspects that are formed by the data. The data coordinate system is illustrated by analyzing a new paired comparison rule, by comparing behaviors of selective rules, and by answering concerns about the Analytic Hierarchy Process (AHP) (such as how to select an "appropriate" consistent data term for a non-consistent one). An elementary (and quick) way to obtain AHP weights is introduced.

1. Introduction

The ranking of alternatives, whether to obtain an ordinal ranking to indicate "which one is better" or a cardinal approach to display distinctions and priorities in terms of weights, is a need that cuts across all disciplines. Realities such as costs, complexities, and other pragmatic concerns may suggest using paired comparison approaches to analyze differences among the alternatives. These are favored approaches, for instance, for aspects of statistics, psychology, engineering, economics, individual decisions, voting, and on and on.

Many methods exist to determine the paired comparison outcomes, which introduces the concern of determining which rule should be used based on, perhaps, the criterion that it leads to more reliable outcomes. To analyze this question, and to address mysteries such as the emergence of cycles, the approach adopted here is to move the discussion a step above analyzing specific rules to create a tool — a coordinate system — for data spaces. The purpose of this coordinate system is to facilitate the analysis and comparisons of ranking rules and to explain why different methodologies can yield different answers.

This paper has three main themes. The primarily one (Section 4) is to create a coordinate system that holds for quite general, generic choices of "data spaces." By this I mean that the structure holds whether the data comes from subjective choices of how much one

The theme of this paper was formulated and the basic results developed during an August 16-20, 2010, American Institute of Mathematics "Mathematics of ranking" workshop organized by Shivani Agarwal and Lek-Heng Lim. During this workshop, T. Saaty introduced me to AHP. My thanks for discussions of these and related topics with the workshop participants, in particular K. Arrow and T. Saaty. Most of this paper was written during a September 2010 visit to the Shanghai University of Finance and Economics; my thanks to my hostess Lingfang (Ivy) Li. This research was supported by NSF DMS-0631362 and CMMI-1016785.

alternative is better than another, perhaps it comes from a statistical analysis, maybe it reflects experimental conclusions or engineering measurements, perhaps it reflects voting tallies, etc.

A problem with paired comparisons is that the data normally fails to satisfy even a crude sense of transitivity. Yet, for the outcomes to be useful, whether expressed as ordinal rankings or cardinal values where the weights indicate levels of priority, the outcomes must satisfy some form of transitivity. This reality suggests finding ways to separate those data components that support this aim from those components that can compromise the objective and even create paradoxical conclusions. This structure, after developed, will be used to analyze different decision rules. As an example, by use of this system, it becomes possible to determine whether or not a specified rule is influenced by those data components that cause paradoxical behaviors.

A second objective (Section 2.2) is to discuss a particular way to make paired comparisons. While this approach is not new (e.g., a version designed for engineering decisions is described in Saari-Sieberg [7]), the description given here is the first discussion about how to use this approach with a wide assortment of different kinds of data. The analysis of this approach illustrates features of the data coordinate system.

A third objective is to use the developed coordinate structures to analyze the well-known Analytic Hierarchy Process (AHP) (Section 2.1). While AHP (developed primarily by T. Saaty) appears to have delightful properties, there remain mysteries. The data coordinate system provides some answers.

2. Two different methods

To make the discussion concrete, two seemingly different rules are described. One (AHP) is based on the eigenvector structure of a matrix defined by the data. For reasons that will become clear, AHP can be thought of as a multiplicative rule. In contrast, the Borda assignment rule (BAR) introduced in Section 2.2 is additive.

2.1. Analytic Hierarchy Process.

The Analytic Hierarchy Process (AHP) ranks N alternatives based on how an individual assigns weights to each $\{A_i, A_j\}$ pair of the alternatives. When comparing A_i with A_j, the assigned $a_{i,j}$ value is intended to indicate, in some manner, the multiple of how much A_i is preferred to A_j. Thus the natural choice for $a_{j,i}$ is the reciprocal

$$a_{j,i} = \frac{1}{a_{i,j}}. \tag{1}$$

To ensure consistency with Equation 1, let $a_{j,j} = 1$.

These terms define a $N \times N$ matrix $((a_{i,j}))$ of positive entries. According to the Perron-Frobenius Theorem (e.g., Horn and Johnson [3]), matrix $((a_{i,j}))$ has an unique eigenvector $\mathbf{w} = (w_1, \dots, w_N)$, $w_j > 0$, $j = 1, \dots, N$, associated with the matrix's sole positive eigenvalue. For each j, the normalized value of w_j (e.g., let $\sum_{j=1}^{N} w_j = 1$) defines the weight, "intensity", or "priority" associated with alternative A_j.

A "consistency" setting for this $((a_{i,j}))$ matrix is where each triplet i, j, k, satisfies the expression

$$a_{i,j}a_{j,k} = a_{i,k}, \quad i,j,k = 1,\dots,N. \tag{2}$$

The power gained from Equation 2 is that (Saaty [8, 9])

$$a_{i,j} = \frac{w_i}{w_j}, \quad i,j = 1,\dots,N. \tag{3}$$

A delightful advantage gained from consistency, then, is that the $a_{i,j}$ values equal the natural quotients of the w_j weights. Multiplying the eigenvector with the matrix shows that should Equation 2 be satisfied, the eigenvalue equals N.

So, the Equation 2 consistency condition creates the intuitive interpretation for the $a_{i,j}$ and w_j values captured by Equation 3. But there remain several natural questions; some follow:

(1) While Equation 2 simplifies computations, does this condition have other interpretations that would appeal to modeling concerns?

(2) The $a_{i,j}$ terms, $i < j$, define the vector $\mathbf{a} = (a_{1,2}, a_{1,3}, \dots, a_{N-1,N}) \in \mathbb{R}_+^{\binom{N}{2}}$, which lists the entries in the upper triangular portion of $((a_{i,j}))$ according to rows. (So $\mathbb{R}_+^{\binom{N}{2}}$ is the positive orthant of $\mathbb{R}^{\binom{N}{2}}$; i.e., $\mathbb{R}_+^{\binom{N}{2}} = \{\mathbf{a} \in \mathbb{R}^{\binom{N}{2}} \mid a_{i,j} > 0\}$. When using $\mathbf{a}$ in an expression requiring an $a_{i,j}$ term where $i > j$, replace $a_{i,j}$ with $\frac{1}{a_{j,i}}$.)

Equation 2 defines a $(N-1)$-dimensional submanifold of $\mathbb{R}_+^{\binom{N}{2}}$ that I denote by $\mathbb{SC}_N$ (for "Saaty consistency"). Is there a natural interpretation for $\mathbb{SC}_N$?

(3) As $\mathbb{SC}_N$ is lower dimensional, it is unlikely for the selected $a_{i,j}$ values to define a point $\mathbf{a} \in \mathbb{SC}_N$. What does such an $\mathbf{a}$ mean? Is there a natural interpretation for this general setting where these $a_{i,j}$ terms fail to satisfy Equation 2?

(4) There are many ways to define an associated $\mathbf{a}' \in \mathbb{SC}_N$ for $\mathbf{a} \notin \mathbb{SC}_N$. One could, for example, let $\mathbf{a}'$ be the $\mathbb{SC}_N$ point closest (with some metric) to $\mathbf{a}$. But the infinite number of choices of metrics permits a continuum of possible $\mathbf{a}'$ choices. Is there a natural choice for $\mathbf{a}'$ that can be justified in terms of the AHP structure rather than with the arbitrary selection of a metric?

(5) Is there a simple way (other than computing eigenvectors) to find the w_j values, at least in certain settings?

These and other questions will be answered in this paper.

2.2. **Borda assignment rule.** An alternative way to compare a pair of alternatives $\{A_i, A_j\}$ also assigns a numeric value to each alternative in a pair; e.g., a typical choice comes from some interval, say $[-m_1, m_2]$, which I normalize to $[0,1]$. Let the intensity assigned to A_i over A_j in this pair be $b_{i,j} \in [0,1]$ where $b_{i,j}$ represents the share of the $[0,1]$ interval that is assigned to A_i; e.g., in a gambling setting, $b_{i,j}$ could represent the probability that A_i beats A_j. With this interpretation, the $b_{j,i}$ value assigned to A_j is

$$b_{j,i} = 1 - b_{i,j}. \tag{4}$$

As Equation 4 requires $b_{i,j} + b_{j,1} = 1$, it is natural to require $b_{j,j} = \frac{1}{2}$, $j = 1, \dots, N$.

To obtain more general conclusions, the differences between weights (rather than the actual weights) are emphasized. To convert to this setting, let

$$d_{i,j} = b_{i,j} - b_{j,i}, \text{ so } d_{i,j} = -d_{j,i}. \tag{5}$$

This expression requires $d_{j,j} = 0$, so there are only $\binom{N}{2}$ distinct $d_{i,j}$ values where $i < j$. These terms define $\mathbf{d} = (d_{1,2}, d_{1,3}, \dots, d_{N-1,N}) \in \mathbb{R}^{\binom{N}{2}}$. (No restrictions are imposed on the sign or magnitude of each $d_{i,j}$.)

In what I call the *Borda assignment rule* (BAR), assign the value

$$b_i = \sum_{j=1, j\neq i}^{N} b_{i,j} \tag{6}$$

to A_i, $i = 1, \dots, N$. Because

$$b_{i,j} = (b_{i,j} - \frac{1}{2}) + \frac{1}{2} = \frac{1}{2}(b_{i,j} - b_{j,i}) + \frac{1}{2} = \frac{1}{2}d_{i,j} + \frac{1}{2} \tag{7}$$

and $b_{i,i} = \frac{1}{2}$, equivalent expressions for BAR are $b_i = \sum_{j=1}^{N}[b_{i,j} - \frac{1}{2}] + \frac{N-1}{2}$, $i = 1, \dots, N$, and the one used here,

$$b_i = \overline{b}_i + \frac{N-1}{2} \text{ where } \overline{b}_i = \frac{1}{2}\sum_{j=1}^{N} d_{i,j}, \quad i = 1, \dots, N. \tag{8}$$

The simpler $\overline{b}_i$ is used more often than b_i.

A special, but important case of BAR (which provides its name) comes the part of voting theory that analyzes majority votes over pairs. Here, with m voters, $b_{i,j}$ is the fraction of all voters who prefer candidate A_i over A_j. The BAR values for this election (Equation 6) turn out to be equivalent to what is known as the "Borda Count" tallies (Saari [6]). As a $m = 100$ voter example, if the A_1:A_2 tally is 60:40, the A_1:A_3 tally is 45:55, and the A_2:A_3 tally is 70:30, then

$$\begin{array}{l} b_{1,2} = 0.60,\ b_{2,1} = 0.40,\ d_{1,2} = 0.20; \quad b_{1,3} = 0.45,\ b_{3,1} = 0.55,\ d_{1,3} = -0.10; \\ b_{2,3} = 0.70,\ b_{3,2} = 0.30,\ d_{2,3} = 0.40. \end{array} \tag{9}$$

In turn, this means that

$$\begin{array}{rl} b_1 = & b_{1,2} + b_{1,3} = 0.60 + 0.45 = 1.05 \text{ or } 1 + \overline{b}_1 = 1 + \frac{1}{2}(0.20 - 0.10) \\ b_2 = & b_{2,1} + b_{2,3} = 0.40 + 0.70 = 1.10 \text{ or } 1 + \overline{b}_2 = 1 + \frac{1}{2}(-0.20 + 0.40) \\ b_3 = & b_{3,1} + b_{3,2} = 0.55 + 0.30 = 0.85 \text{ or } 1 + \overline{b}_3 = 1 + \frac{1}{2}(0.10 - 0.40) \end{array} \tag{10}$$

While BAR is agnostic about the origin of the $b_{i,j}$ values, voting theory traditionally requires voters with complete transitive preferences over the alternatives. This assumption leads to the actual definition of the Borda Count, which is to tally a voter's ballot by assigning $N - j$ points to the j^{th} positioned candidate; e.g., if $N = 4$, then 3, 2, 1, 0 points are assigned, respectively, to the top-, second-, third-, and bottom-ranked candidates. With $N = 3$, the tallying weights are 2, 1, 0. As these tallies are based on the number of voters, they must be scaled by the multiple of $\frac{1}{m}$ to obtain the Equation 6 value.

To illustrate, the Equation 9 values also arise with transitive voter preferences

(11)

Number	Ranking	Number	Ranking
30	$A_1 \succ A_2 \succ A_3$	15	$A_2 \succ A_1 \succ A_3$
30	$A_3 \succ A_1 \succ A_2$	25	$A_2 \succ A_3 \succ A_1$

Because A_1 is top-ranked 30 times and second-ranked 45 times, her Borda tally is

$$(2 \times 30) + (1 \times 45) = 105.$$

When scaled by $\frac{1}{m}$, or $\frac{1}{100}$, the resulting value equals that of b_1 in Equation 10. Similar computations hold for A_2 and A_3.

3. Creating connections; a data coordinate system

The promised coordinate system (which will separate data portions that satisfy a particularly strong version of transitivity from data portions that manifest cycles) will be used to discover properties of BAR (Section 5). Of interest, the same coordinate system can be used to extract and compare AHP properties with those of BAR (Section 5).

To do so, an isomorphic relationship is established between $\mathbb{R}_+^{\binom{N}{2}}$, the domain of AHP, and $\mathbb{R}^{\binom{N}{2}}$, the domain of BAR. This relationship premits AHP concerns to be transferred to the simpler setting that will be endowed with the data coordinate system. If the conveyed problem can be solved within this coordinate system, the answer can be transferred back to the AHP setting. In this manner, answers are found for the questions raised about AHP.

To define the isomorphism $F : \mathbb{R}_+^{\binom{N}{2}} \to \mathbb{R}^{\binom{N}{2}}$, let the $\mathbb{R}^{\binom{N}{2}}$ coordinates be given by $d_{i,j} \in \mathbb{R}$. The $F(\mathbf{a}) = \mathbf{d}$ definition is to let

$$d_{i,j} = \ln(a_{i,j}),\ 1 \leq i < j \leq N, \ \text{ so } F^{-1}(\mathbf{d}) = \mathbf{a} \text{ is given by } a_{i,j} = e^{d_{i,j}}. \tag{12}$$

Clearly, F is an isomorphism. Using these spaces in a way to handle the extra $a_{i,j}$ and $d_{i,j}$ terms with $i > j$ is easy; indeed, F^{-1} essentially converts the $d_{i,j} = -d_{j,i}$ condition (Equation 5) into a constraint that is equivalent to Equation 1.

For another transferred relationship, when Equation 2 is expressed in $F(\mathbf{a})$ terms as $\ln(a_{i,j}) + \ln(a_{j,k}) = \ln(a_{i,k})$, it is equivalent to

$$d_{i,j} + d_{j,k} = d_{i,k}, \qquad \text{for all } i, j, k = 1, \ldots, N. \tag{13}$$

Equation 13 plays a central role in what follows, so it is worth providing an interpretation in terms of "transitivity of rankings."

Strong transitivity: The requirement for transitivity of ordinal rankings is that if $A_i \succ A_j$ and $A_j \succ A_k$, then it must be that $A_i \succ A_k$. Re-expressed in terms of $d_{i,j}$ values, the definition for transitivity becomes

$$\text{if } d_{i,j} > 0 \text{ and } d_{j,k} > 0, \text{ then it must be that } d_{i,k} > 0; \quad i, j, k = 1, \ldots, N. \tag{14}$$

The Equation 13 cardinal expression is much stronger than Equation 14; it has the flavor of measurements along a line where the signed distance from point i to point j plus the signed distance from j to point k equals the signed distance from i to k. In fact, Equation 13 restricts $\mathbf{d}$ to a $(N-1)$-dimensional linear subspace $\mathcal{ST}_N$ that I call the *strong transitivity plane:*

$$\mathcal{ST}_N = \{\mathbf{d} \in \mathbb{R}^{\binom{N}{2}} \mid d_{i,j} + d_{j,k} = d_{i,k} \text{ for all } i, j, k = 1, \dots, N\}. \tag{15}$$

This connection provides an interpretation for Saaty's Equation 2 consistency condition.

Theorem 1. $F(\mathcal{SC}_N) = \mathcal{ST}_N$. *Stated in words,* $\mathcal{SC}_N$ *and* $\mathcal{ST}_N$ *are isomorphic.*

Thus the Equation 2 consistency condition is equivalent to a "strong transitivity" constraint, which is an attractive modeling requirement.

4. Coordinate system for $\mathbb{R}^{\binom{N}{2}}$

Not only does F convert the nonlinear $\mathcal{SC}_N$ into the linear $\mathcal{ST}_N$, it reduces complexities by replacing multiplication with addition. The natural next step is to determine how the data coming from the different $\mathbb{R}^{\binom{N}{2}}$ subspaces affect BAR and AHP values.

By definition, the vectors (i.e., data) in the $\mathcal{ST}_N$ plane in $\mathbb{R}^{\binom{N}{2}}$ satisfy a strong transitivity condition. This connection suggests developing a coordinate system for $\mathbb{R}^{\binom{N}{2}}$ in terms of a basis for $\mathcal{ST}_N$ and its orthogonal complement that is denoted by $\mathcal{C}_N$. The approach developed here is motivated by results from voting theory developed in (Saari [5, 6]).

The basis for $\mathcal{ST}_N$ comes from the following:

Definition 1. *For each* $i = 1, \dots, N$, *let the vector* $\mathbf{B}_i \in \mathbb{R}^{\binom{N}{2}}$ *be where each* $d_{i,j} = 1$ *for* $j = 1, \dots, N$, *and each* $d_{k,j} = 0$ *if* $k, j \neq i$.

Before proving that these vectors define a basis for $\mathcal{ST}_N$ (Theorem 2), it is worth previewing the results with examples.

Example ($N = 3$): For $N = 3$, we have $\mathbf{d} = (d_{1,2}, d_{1,3}, d_{2,3})$, so $\mathbf{B}_1 = (1, 1, 0)$, $\mathbf{B}_2 = (-1, 0, 1)$, $\mathbf{B}_3 = (0, -1, -1)$. Because $\mathbf{B}_1 + \mathbf{B}_2 + \mathbf{B}_3 = \mathbf{0}$, the system is linearly dependent. Any two vectors, however, are independent, so they span a two-dimensional space.

To prove that this two-dimensional space is the desired $\mathcal{ST}_3$, it must be shown that if $\mathbf{d}^* = \alpha\mathbf{B}_1 + \beta\mathbf{B}_2 = (\alpha - \beta, \alpha, \beta)$, then $d^*_{1,2} + d^*_{2,3} = d^*_{1,3}$; that is, $\mathbf{d}^*$ satisfies the defining condition for $\mathcal{ST}_3$ (Equation 13). The proof reduces to showing that $d^*_{1,2} + d^*_{2,3} = (\alpha - \beta) + (\beta)$ always equals $d^*_{1,3} = \alpha$, which is true.

The normal space to $\mathcal{ST}_3$ is spanned by $(1, -1, 1)$. While this vector does *not* satisfy Equation 13, it provides a convenient way to capture the differences between transitivity (Equation 14) and strong transitivity (Equation 13). Namely, if $\mathbf{d} \notin \mathcal{ST}_3$ satisfies ordinal transitivity of Equation 14 but not the strong cardinal transitivity condition of Equation 13, then $\mathbf{d}$ must have a component in the $(1, -1, 1)$ normal direction. It also follows from this comment that an open region of ordinal transitive rankings in $\mathbb{R}^{\binom{3}{2}}$ surrounds $\mathcal{ST}_3$; $\mathcal{ST}_3$ is a lower dimensional subspace. (This holds for all N.)

Example ($N = 4$): For $N = 4$ where $\mathbf{d} = (d_{1,2}, d_{1,3}, d_{1,4}, d_{2,3}, d_{2,4}, d_{3,4})$, we have that

$$\begin{array}{lll} \mathbf{B}_1 = (1, 1, 1, 0, 0, 0), & \mathbf{B}_2 = (-1, 0, 0, 1, 1, 0), & \mathbf{B}_3 = (0, -1, 0, -1, 0, 1), \\ \mathbf{B}_4 = (0, 0, -1, 0, -1, -1). & & \end{array} \tag{16}$$

The sum of these four vectors equals zero, so at most three of them are linearly independent; any three are. Condition Equation 13 is easily verified in the same manner as above.

The next theorem asserts that these conditions hold for all values of $N \geq 3$.

Theorem 2. *For $N \geq 3$, any subset of $(N-1)$ vectors from $\{\mathbf{B}_i\}_{i=1}^N$ spans $\mathcal{ST}_N$.*

Proof. To prove that $\sum_{j=1}^N \mathbf{B}_j = \mathbf{0}$, notice that each $d_{i,j}$ coordinate in this summation has only two non-zero terms; one comes from $\mathbf{B}_i$ and the other from $\mathbf{B}_j$. Because $d_{i,j} = -d_{j,i}$, one term is $+1$ and the other is -1, which proves the assertion.

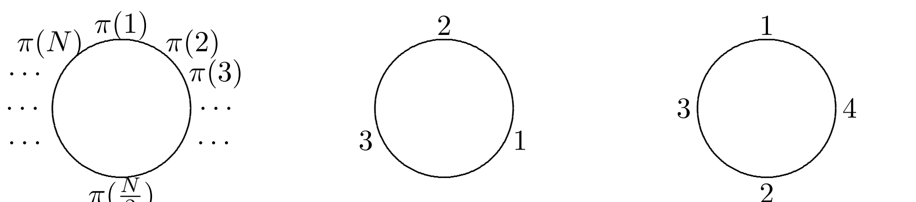
At most $(N-1)$ of the $\mathbf{B}_j$ vectors are linearly independent; by symmetry, it suffices to show that the first $(N-1)$ are. That is, if $\sum_{j=1}^{N-1} \alpha_j \mathbf{B}_j = \mathbf{0}$, it must be that $\alpha_j = 0$ for all choices of j. But if $\alpha_j \neq 0$, then $\alpha_j \mathbf{B}_j$ has a non-zero $d_{j,N}$ component. This $d_{j,N}$ component, however, equals zero for all remaining $\mathbf{B}_i$ vectors, which proves the assertion.

It remains to show that any vector $\mathbf{d}^*$ in the span of $\{\mathbf{B}_j\}_{j=1}^{N-1}$ satisfies the strong transitivity condition of Equation 13. To simplify the proof (by avoiding special cases for the index N), let

$$\mathbf{d}^* = \sum_{j=1}^{N} \beta_j \mathbf{B}_j.$$

It must be shown that $d^*_{i,j} + d^*_{j,k} = d^*_{i,k}$ for all triplets $i, j, k = 1, \ldots, N$. In this coordinate system, $d^*_{i,j} = [\beta_i - \beta_j]$ (To see this, if $i < j$, then $d^*_{i,j} = [\beta_i - \beta_j]$, so $d^*_{j,i} = -d^*_{i,j} = [\beta_j - \beta_i]$.) Thus $d^*_{i,j} + d^*_{j,k} = [\beta_i - \beta_j] + [\beta_j - \beta_k] = [\beta_i - \beta_k]$ which is the desired $d^*_{i,k} = [\beta_i - \beta_j]$ value.

As $\{\mathbf{B}_j\}_{j=1}^N$ spans a $(N-1)$-dimensional subspace of vectors that satisfy Equation 13, this set defines an $(N-1)$-dimensional subspace of $\mathcal{ST}_N$. But $\mathcal{ST}_N$ (defined by Equation 13) is a $(N-1)$-dimensional subspace, so $\{\mathbf{B}_j\}_{j=1}^N$ spans $\mathcal{ST}_N$. □



a. General case **b.** Special $N = 3$ case **c.** Special $N = 4$ case

Figure 1. Cyclic arrangements of data

The $\mathcal{C}_N$ subspace. The next step is to find the normal space for $\mathcal{ST}_N$. As part of the analysis, it will turn out that these data portions always define cycles.

To define the basis (based on developments in [5, 6]) for this orthogonal, or *cyclic space* $\mathcal{C}_N$, list the N indices, in any specified order, along the edges of a circle as indicated in Figure 1. In Figure 1a, $\pi(j)$ represents the integer listed in the j^{th} slot around the circle; Figures 1a, b,c illustrate special cases. (So in Figure 1b, $\pi(1) = 2, \pi(2) = 1, \pi(3) = 3$.) With respect to a clockwise motion about the circle, each integer is preceded by an integer and followed by a different integer.

Definition 2. *Let π be a specific listing, or permutation, of the indices $1, 2, \dots, N$ where the indices are listed in the cyclic fashion $(\pi(1), \pi(2), \dots, \pi(N))$ around a circle. Define $\mathbf{C}_\pi \in \mathbb{R}^{\binom{N}{2}}$ as follows: If j immediately follows i in a clockwise direction, then $d_{i,j} = 1$; if j immediately precedes i, then $d_{i,j} = -1$. Otherwise $d_{i,j} = 0$. Vector $\mathbf{C}_\pi$ is the "cyclic direction defined by π".*

Example ($N = 3$): Let $\pi = (2, 1, 3)$ as represented in Figure 1b. To compute $\mathbf{C}_\pi = (d_{1,2}, d_{1,3}, d_{2,3})$, because 2 precedes 1 on the circle, $d_{1,2} = -1$. Likewise, 3 immediately follows 1, so $d_{1,3} = 1$. For the remaining term of $d_{2,3}$, as 3 immediately precedes 2, we have that $d_{2,3} = -1$ Thus $\mathbf{C}_\pi = (-1, 1, -1)$.

Vector $\mathbf{C}_\pi$ significantly violates the strict transitivity of Equation 13. In sharp contrast to the $d_{1,2} + d_{2,3} = d_{1,3}$ requirement, the $\mathbf{C}_\pi$ entries are $d_{1,2} = d_{2,3} = -1$, where rather than the required $d_{1,3} = -2$, we have $d_{1,3} = 1$. When expressed in terms of ordinal rankings, these entries represent the cycle $A_2 \succ A_1, A_1 \succ A_3$, and $A_3 \succ A_2$ each by the same $d_{i,j}$ difference. Indeeed, a direct computation proves that cyclic data direction $\mathbf{C}_\pi$ is orthogonal to both $\mathbf{B}_1 = (1, 1, 0)$ and $\mathbf{B}_2 = (-1, 0, 1)$, so $\mathbf{C}_\pi$ defines a normal direction for $\mathcal{ST}_3$.

The indices are listed on a circle, so any rotation of these numbers does not affect which integers follow and precede a specified value; they all define the same $\mathbf{C}_\pi$ vector. Thus, each of the $(2, 1, 3)$, $(1, 3, 2)$ and $(3, 2, 1)$ orderings define the same $\mathbf{C}_{(2,1,3)} = (-1, 1, -1)$.

Three remaining orderings come from $(1, 2, 3)$ and its rotations. The cyclic direction representing these choices is the earlier $\mathbf{C}_{(1,2,3)} = (1, -1, 1) = -\mathbf{C}_{(2,1,3)}$. Thus, $\mathbf{C}_\pi$ spans the normal space $\mathcal{C}_3$ to $\mathcal{ST}_3$.

Example ($N = 4$): The permutation $\pi = (1, 4, 2, 3)$, as depicted in Figure 1c, can be used to underscore the simplicity of finding the corresponding $\mathbf{C}_\pi$. Start at the top and move in a clockwise direction: 4 follows 1 so $d_{1,4} = 1$; 2 follows 4 so $d_{4,2} = 1$ or $d_{2,4} = -1$; 3 follows 2 so $d_{2,3} = 1$, and 1 follows 3 so $d_{3,1} = 1$ or $d_{1,3} = -1$. All other $d_{j,k}$ entries equal zero. Thus, the corresponding

$$\mathbf{C}_{(1,4,2,3)} = (d_{1,2}, d_{1,3}, d_{1,4}, d_{2,3}, d_{2,4}, d_{3,4}) = (0, -1, 1, 1, -1, 0). \tag{17}$$

The cyclic arrangement represented by $\mathbf{C}_\pi$ can also be read from Figure 1c; it is

$$A_1 \succ A_4, \quad A_4 \succ A_2, \quad A_2 \succ A_3, \quad A_3 \succ A_1$$

where each difference is captured by a common $d_{i,j}$ value.

The cyclic nature of the data represented by $\mathbf{C}_\pi$ suggests that $\mathbf{C}_\pi$ is orthogonal to $\mathcal{ST}_4$. It is, and a simple proof is to show that $\mathbf{C}_\pi$ is orthogonal to each $\mathbf{B}_j$ basis vector. The only non-zero entries of $\mathbf{B}_1$, for instance, are the $d_{1,j}$, $j > 1$, terms, which all equal unity. In the permutation π, "1" is immediately preceded and immediately followed by different integers; these $d_{1,j}$ terms will be, respectively, -1 and 1 in $\mathbf{C}_\pi$, thus they cancel in the scalar product with $\mathbf{B}_1$. As all other $d_{i,j}$ terms are zero in $\mathbf{C}_\pi$, orthogonality is verified.

Figure 1c represents one of $4! = 24$ permutations of the 4 indices; rotations of this ordering preserve which integers precede and follow others, so they all define the same Equation 17 vector. Four other orderings are obtained by flipping the ordering (which is equivalent to using the ordering in a counter-clockwise manner). These orderings define a cyclic vector that is a (-1) multiple of the Equation 17 choice.

Three mutually orthogonal $\mathbf{C}_\pi$ vectors are orthogonal to $\mathcal{ST}_4$; they are given by

$$\mathbf{C}_{(1,2,3,4)} = (1, 0, -1, 1, 0, 1) = -\mathbf{C}_{(1,4,3,2)}, \quad \mathbf{C}_{(1,3,2,4)} = (0, 1, -1, -1, 1, 0) = -\mathbf{C}_{(1,4,2,3)}, \quad \mathbf{C}_{(1,3,4,2)} = (-1, 1, 0, 0, -1, 1) = -\mathbf{C}_{(1,2,4,3)}$$

These vectors span the normal space – the cyclic space, $\mathcal{C}_4$. For a dimension count, $\mathbb{R}^{\binom{4}{2}}$ has dimension six and the strongly transitive subspace $\mathcal{ST}_4$ has dimension three, so the subspace $\mathcal{C}_4$ (spanned by the cyclic directions) accounts for the remaining three dimensions. This establishes a coordinate system for $\mathbb{R}^{\binom{4}{2}}$.

The next two statements extend this coordinate system to all choices of N.

Theorem 3. *For any $N \geq 3$, a $\mathbf{C}_\pi$ vector represents a particular cyclic ranking of the N alternatives $\{A_j\}$.*

For any triplet of distinct indices $\{i, j, k\}$, there exists a linear combination of the $\mathbf{C}_\pi$ vectors so that, rather than fulfilling Equation 13, the data satisfies

$$d_{i,j} = d_{j,k} = 1,\ d_{i,k} = -1,\ \textit{all other } d_{u,v} = 0. \tag{18}$$

Equation 18 corresponds to the cycle of the triplet

$$A_i \succ A_j, \quad A_j \succ A_k, \quad A_k \succ A_i,$$

each by the same $d_{i,j}$ difference; all other pairs are ties.

Proof. The $\mathbf{C}_\pi$ entries defined by permutation $(\pi(1), \pi(2), \dots, \pi(N))$ represents the cycle

$$A_{\pi(1)} \succ A_{\pi(2)}, \quad A_{\pi(2)} \succ A_{\pi(3)}, \quad \dots, \quad A_{\pi(N-1)} \succ A_{\pi(N)}, \quad A_{\pi(N)} \succ A_{\pi(1)}$$

each with the same $d_{i,j}$ difference. All remaining $d_{u,v}$ values equal zero, so all remaining pairs of alternatives are tied.

The symmetry among indices means that to prove Equation 18, it suffices to show that there exist arrangements of the cyclic directions so that $d_{1,2} = d_{2,3} = d_{3,1} > 0$ where all remaining $d_{i,j}$ values equal zero. For $N = 3$, the conclusion follows by using $\mathbf{C}_{(1,2,3)}$. For $N = 4$, the conclusion follows by using $\mathbf{C}_{(1,2,4,3)}$, $\mathbf{C}_{(1,2,3,4)}$, $\mathbf{C}_{(1,4,2,3)}$. In two of these arrangements, 1 is followed by 2, 2 by 3, and 3 by 1, so $d_{1,2} = d_{2,3} = d_{3,1} = 2$. For each of the three remaining pairs, $\{1, 4\}$, $\{2, 4\}$, $\{3, 4\}$, each index is immediately preceded in one permutation and immediately followed in another by the other index, so each $d_{u,v} = 0$.

For $N \geq 5$, consider the four permutations

$$\pi_1 = (1, 2, 3, r), \quad \pi_2 = (1, 2, 3, r'), \quad \pi_3 = (3, 1, s), \quad \pi_4 = (3, 1, s'),$$

where r, s are permutations of the remaining indices and the prime indicates the reversal of this listing. (So, if $r = (4, 6, 5)$, then $r' = (5, 6, 4)$.) The only restriction imposed on the r, s listings is that they must begin with 4 and end with 5.

To compute the $d_{i,j}$ values, each of $\{1, 2\}$, $\{2, 3\}$, and $\{3, 1\}$ appears, in this immediate order, in precisely two of the four listings; they never appear in a reversed order. Thus $d_{1,2} = d_{2,3} = d_{3,1} = 2$. To show that all remaining $d_{u,v} = 0$, the critical terms are $d_{4,j}$ and $d_{5,j}$, $j = 1, 3$. (This is because these indices are adjacent in certain permutations.) To compute $d_{1,4}$, notice that 1 immediately precedes 4 in π_3 and immediately follows 4 in π_2; in all other listings the two indices are separated, so $d_{1,4} = 0$. A similar argument shows that $d_{1,5} = d_{3,4} = d_{3,5} = 0$. For any other pair of indices, if i and j are adjacent in r or s, then they are adjacent in the opposite order in the respective r', s', so $d_{i,j} = 0$. This completes the proof. □

Theorem 4. *For any $N \geq 3$ and any permutation π of the indices, $\mathbf{C}_\pi$ is orthogonal to $\mathcal{ST}_N$. These $\mathbf{C}_\pi$ vectors span $\mathcal{C}_N$, which is the $[\binom{N}{2} - (N-1)] = \binom{N-1}{2}$-dimensional subspace of vectors normal to $\mathcal{ST}_N$.*

Proof. To show that $\mathbf{C}_\pi$ is orthogonal to $\mathcal{ST}_N$, it suffices to show that $\mathbf{C}_\pi$ is orthogonal to each basis vector $\mathbf{B}_i$. The only non-zero entries in $\mathbf{B}_i$ are the $d_{i,j}$, $j \neq i$, terms, which all equal $+1$. But $\mathbf{C}_\pi$ has only two non-zero $d_{i,j}$ terms; one is for the j value immediately following i, and the other is for the j value immediately preceding i in the permutation. Because one entry is $+1$ and the other is -1, the scalar product is zero.

To show that the set $\{\mathbf{C}_\pi\}$ over all permutations of the indices π spans $\mathcal{C}_N$, it suffices to show that this set spans a linear subspace of dimension $\binom{N-1}{2}$. This is done by using the triplet cycles of Theorem 3 to show that the $d_{i,j}$ values of these triplets define $\binom{N-1}{2}$ linearly independent vectors. Key to the proof are the $d_{j,k}$ coordinates where $2 \leq j < k \leq N$, there are precisely $\binom{N-1}{2}$ of them. Define $\mathbf{V}_{j,k}$, $2 \leq j < k$, by the triplet $d_{1,j}, d_{j,k}, d_{1,k}$. Because the index "1" occurs in two of these terms, it follows that $\mathbf{V}_{j,k}$ has unity in the $d_{j,k}$ component, and zeros for all other $d_{u,v}$, $u, v \geq 2$. From this fact, it follows that $\{\mathbf{V}_{j,k}\}_{2\leq j<k\leq N}$ spans a $\binom{N-1}{2}$ dimensional subspace. This completes the proof. □

5. Examples and comparing decision rules

With its cyclic symmetry, no natural transitive (ordinal or cardinal) ranking can be assigned to $\mathcal{C}_N$ terms other than complete indifference. Combining this fact with the cyclic nature of the $\mathcal{C}_N$ data, it is reasonable to anticipate that these terms are responsible for paired comparison paradoxes and cause different methods to have, with the same data, different answers. This is the case. A desired property of a paired comparison rule, then, is if it can filter out all $\mathcal{C}_N$ content. As shown next, BAR does this.

5.1. **BAR properties.** Basic properties of BAR follow almost immediately from the data coordinate system. To develop intuition about what to expect, consider the following general representation of what can occur for $N = 3$:

$$\mathbf{d} = \beta_1\mathbf{B}_2 + \beta_2\mathbf{B}_2 + c\mathbf{C}_{(1,2,3)} = (\beta_1 - \beta_2 + c, \beta_1 - c, \beta_2 + c) \tag{19}$$

where $\mathbf{C}_{(1,2,3)} = (1, -1, 1) \in \mathcal{C}_3$. The BAR values are

$$\begin{aligned} \bar{b}_1 &= \tfrac{1}{2}[(\beta_1 - \beta_2 + c) + (\beta_1 - c)] = \tfrac{1}{2}[2\beta_1 - \beta_2], \\ \bar{b}_2 &= \tfrac{1}{2}[(\beta_2 - \beta_1 - c) + (\beta_2 + c)] = \tfrac{1}{2}[2\beta_2 - \beta_1], \\ \bar{b}_3 &= \tfrac{1}{2}[(-\beta_1 + c) + (-\beta_2 - c)] = -\tfrac{1}{2}(\beta_1 + \beta_2) \end{aligned} \tag{20}$$

Notice; according to Equation 20, the $\bar{b}_j$ values are not affected, in any manner, by the $\mathcal{C}_3$ cyclic data components! As these terms cancel, BAR filters out this noise component.

Another feature to observe is how the BAR values faithfully represent the multiples of strongly transitive $\mathcal{ST}_3$ components of the data. If, for example, $\beta_1 > \beta_2$, (so the strongly transitive *data* components reflect a preference for A_1 over A_2), then (Equation 20) the BAR values rank $A_1 \succ A_2$ with the cardinal values $\overline{b}_1 = \beta_1 - \frac{1}{2}\beta_2 > \overline{b}_2 = \beta_2 - \frac{1}{2}\beta_1$.

(As an aside concerning notation, the $\{\mathbf{B}_j\}_{j=1}^3$ vectors are not independent because $\mathbf{B}_3 = -(\mathbf{B}_1 + \mathbf{B}_2)$. By substitution, it follows that if $\mathbf{d} = \beta_1\mathbf{B}_1 + \beta_2\mathbf{B}_2 = \tilde{\beta}_1\mathbf{B}_1 + \tilde{\beta}_3\mathbf{B}_3$, then $\tilde{\beta}_1 = \beta_1 - \beta_2$, $\tilde{\beta}_3 = -\beta_2$. Substituting into Equation 20 leads to $\overline{b}_1 = \frac{1}{2}[2\tilde{\beta}_1 - \tilde{\beta}_3]$, $\overline{b}_2 = -\frac{1}{2}(\tilde{\beta}_1 + \tilde{\beta}_3)$, and $\overline{b}_3 = \frac{1}{2}[2\tilde{\beta}_3 - \overline{\beta}_1]$; a representation that is consistent with Equation 20.)

As shown in the next theorem, these BAR properties extend to all values of N.

Theorem 5. *For any $N \geq 3$, a general representation for $\mathbf{d} \in \mathcal{R}^{\binom{N}{2}}$ is*

$$\mathbf{d} = \sum_{j=1}^{N-1} \beta_j \mathbf{B}_j + \sum_{k=1}^{\binom{N-1}{2}} c_j \mathbf{C}_{\pi_k}, \tag{21}$$

where $\{\mathbf{C}_{\pi_k}\}_{k=1}^{\binom{N-1}{2}}$ spans $\mathcal{C}_N$.

For each $j = 1, \dots, N-1$,

$$2\overline{b}_j = (N-1)\beta_j - \sum_{k=1, k\neq j}^{N-1} \beta_k = \sum_{k=1,k\neq j} (\beta_j - \beta_k) \tag{22}$$

and $2\overline{b}_N = -\sum_{j=1}^{N-1} \beta_j$.

Basic BAR properties follow directly from Theorem 5. As catalogued next, for instance, $\overline{b}_j$ values depend only upon the data's strongly transitive components, so the difference between any two $\overline{b}_j$ and $\overline{b}_i$ is not affected, in any manner, by $\mathcal{C}_N$ data cyclic components, nor, in fact, by how any other alternative is ranked. Indeed, BAR weights respect how the strongly transitive data components rank specific alternatives in that if $\beta_i > \beta_j$, then $\overline{b}_i > \overline{b}_j$ and $A_i \succ A_j$.

Corollary 1. *For any $N \geq 3$, the $\overline{b}_j$ values are not influenced in any manner by the $\mathcal{C}_N$ cyclic data components; they are strictly determined by $\mathcal{ST}_N$ data components. All differences in the BAR values of alternatives faithfully reflect differences in the strongly transitive data components in that*

$$\overline{b}_j - \overline{b}_k = \frac{N}{2}(\beta_j - \beta_k). \tag{23}$$

Thus $A_j \succ A_k$ iff $\beta_j > \beta_k$.

Proof. Equation 21 (Theorem 5) follows from the coordinate representation of $\mathbb{R}^{\binom{N}{2}}$.

The definition for $\overline{b}_j$ is linear (Equation 8), so it suffices to establish Equation 22 by computing $\overline{b}_j$ for each term in Equation 21. In computing the $\overline{b}_j$ value for a $\mathbf{C}_{\pi_k}$ term, notice that $\mathbf{C}_{\pi_k}$ has precisely two non-zero entries with the index j; one is accompanied with the index preceding j in the permutation, and the other is the index following j. Thus these two terms cancel in the computation of $\bar{b}_j$.

When computing $\bar{b}_j$ for $\beta_k \mathbf{B}_k$, the two cases are where $j = k$ and $j \neq k$. In the first setting, all $d_{j,i}$ terms equal unity, so these $N-1$ terms have the total contribution of $(N-1)\beta_j$. In the second case, $d_{j,k} = -1$ and all other terms involving j are zero; the contribution from this term is $-\beta_k$. (Recall, for $\mathbf{B}_k$, $d_{k,j} = 1$, so $d_{j,k} = -1$.) After including the $\frac{1}{2}$ multiple of Equation 8, Equation 22 follows. Equation 23 follows from Equation 21. □

5.2. **Other rules.** Rather than selecting widely used rules, the following choices are chosen as a means to illustrate the $\mathbb{R}^{\binom{N}{2}}$ coordinate system. As it will become clear, a purpose of these rules (and many others) is to convert setting where the $d_{i,j}$ values define cycles into some form of transitive outcomes.

(1) The *Condorcet winner* (after Condorcet [2]) rule selects the alternative A_i for which $d_{i,j} > 0$ for all $j \neq i$. In other words, A_i beats all other alternatives.
(2) A method called here the *Kemeny approach* (as it resembles the method in Kemeny [4]) finds the "closest" transitive ranking to $\mathbf{d}$. This can be done by replacing certain $d_{i,j}$ values with $-d_{i,j}$ in a manner that satisfies a specified criterion; e.g., minimize the number of $d_{i,j}$ values that need to be reversed, minimize the sum of magnitudes of reversed numbers, etc.
(3) Whatever version of the Kemeny method is adopted, it can be extended to find a point $\mathbf{d}^* \in \mathcal{ST}_N$. $\mathbf{d}^*$ could be, for instance, the closest $\mathcal{ST}_N$ point to the data point obtained with the Kemeny method. Call this the *cardinal Kemeny method.*

Many other rules can, and have been, developed, but the above suffice to demonstrate what happens. A way to demonstrate the role of $\mathcal{C}_N$ in affecting the outcomes of these rules is to show that even if $\mathbf{d}$ defines a transitive ranking, this ranking need not agree with the BAR values.

Example. With $N = 3$, let $\mathbf{d} = 3\mathbf{B}_1 + 2\mathbf{B}_2 = (1, 3, 2)$. Both $d_{1,2}$, $d_{1,3} > 0$, so $\mathbf{d}$ has A_1 ranked above A_2 and A_3. Now create

$$\mathbf{d}^* = \mathbf{d} + x(1, -1, 1) = (1 + x,\ 3 - x,\ 2 + x)$$

where $\mathbf{C}_{(1,2,3)} = (1, -1, 1) \in \mathcal{C}_3$ is a cyclic data component and the value of x is to be determined. For any x satisfying $-2 < x < -1$, the data has $d_{2,1}, d_{2,3} > 0$, so $\mathbf{d}^*$ has A_2, not A_1, ranked above both other alternatives. Thus all of the above rules favor A_2 over A_1 in their outcomes, But as shown above, BAR ignores the distorting cyclic terms, so it retains the original ranking with A_1 ranked above A_2. Cyclic $\mathcal{C}_N$ terms, in other words, are the components that can force different paired comparison rules to have different outcomes.

Extending the x values to $x < -2$ creates the cycle associated with $d_{1,2}, d_{2,,3} < 0$, and $d_{1,3} > 0$. In other words, the above change in the transitive ranking was only a preliminary stage in the continuum moving from the noise free (i.e., free of $\mathcal{C}_N$ terms) setting to cycles. (For $x > 3$, the opposite cycle (where $d_{1,2}, d_{2,3} > 0$, and $d_{1,3} < 0$) is created.)

Theorem 6. *If* $\mathbf{d} \in \mathcal{ST}_N$*, then the outcomes for the Condorcet, all versions of the Kemeny, and cardinal Kemeny methods agree with the BAR outcomes. Thus all differences these methods have in their outcomes are completely due to components of* $\mathbf{d} \notin \mathcal{ST}_N$ *in the cyclic* $\mathcal{C}_n$ *direction. For each method, there exist* $\mathbf{d}^*$ *where its outcome disagrees with BAR.*

Proof. it is clear that all of these rules agree with $\mathbf{d} \in \mathcal{ST}_N$. Thus it must be shown that there exist $\mathbf{d}^* \notin \mathcal{ST}_N$ where the rules disagree with BAR.

To construct an example, let $\mathbf{d}$ in Equation 21 have no cyclic terms and the β_j values satisfy $\beta_1 > \beta_2 > \cdots > \beta_{N-1}$ where at least $\beta_2 > 0$, and $\beta_1 - \beta_2 < \beta_j - \beta_{j+1}$ for $j = 2, \dots, N-2$. Next let $\mathbf{d}^* = \mathbf{d} + x\mathbf{C}_{(2,1,\dots)}$ where x is to be determined; the only important part of the permutation is that 1 immediately follows 2. Thus, in the pairwise vote of $\mathbf{C}_{(2,1,\dots)}$, the A_2:A_1 tally is $(N-1)$:1, with the difference $d_{2,1} = (N-2)$. By selecting x so that $\beta_1 - \beta_2 < x(N-2) < \beta_j - \beta_{j+1}, j = 2, \dots, N-2$, it follows that $d_{1,2} > 0$ in $\mathbf{d}$ and negative for $\mathbf{d}^*$. All remaining $d_{i,j}$ coordinates have the same sign in $\mathbf{d}$ and $\mathbf{d}^*$. In terms of ordinal rankings, $\mathbf{d}$ is accompanied by $A_1 \succ A_2 \succ A_3 \succ \cdots \succ A_N$, while $\mathbf{d}^*$ by $A_2 \succ A_1 \succ A_3 \succ \cdots \succ A_N$.

All of the specified methods, whether ordinal or cardinal, will have the ranking of $\mathbf{d}^*$, while the BAR outcome (because BAR cancels these cyclic terms) is that of $\mathbf{d}$. □

The problem, of course, is that even with the basic objective of these other methods to eliminate cyclic aspects from the outcome, problems arise because the "starting point" for these efforts remain strongly influenced by $\mathcal{C}_N$ terms.

5.3. **AHP outcomes; AHP and BAR connections.** It is reasonable to wonder whether the isomorphism F relates AHP and BAR weights. Because F coverts $a_{i,j}$ values into $d_{i,j} = \ln(a_{i,j})$ terms, should a connection exist, $\bar{b}_j$ would be related to $\ln(w_j)$ values.

Indeed, as the weights assigned by BAR to each of the N-alternatives is determined by Equation 8, we have that

$$\bar{b}_i = \frac{1}{2}\sum_{j=1}^{N} d_{i,j} = \frac{1}{2}\sum_{j=1}^{N}\ln(a_{i,j}) = \frac{1}{2}\ln(\Pi_{j=1}^{N} a_{i,j}). \tag{24}$$

This expression nicely connects BAR and AHP outcomes when Equation 2 is satisfied, which demonstrates at least a partial commonality between the approaches.

Theorem 7. *When the consistency Equation 2 is satisfied, then*

$$\bar{b}_i = \frac{N}{2}\ln(w_i) - B, \quad i = 1, \dots, N, \tag{25}$$

where B *is a constant.*

Proof. If Equation 2 is satisfied, then $a_{i,j} = \frac{w_i}{w_j}$. It follows from Equation 24 that

$$\bar{b}_i = \frac{N}{2}\ln(w_i) - B, \quad B = \frac{1}{2}\ln(w_1 w_2 \dots w_N) \text{ for } i = 1, \dots, N,$$

which is Equation 25. Re-expressing this equation in terms of the w_i weights,

$$w_i = D \times exp(\bar{b}_i)^{\frac{2}{N}}, \quad i = 1, \dots, N, \tag{26}$$

where D is some constant. Equation 26 will useful when describing AHP settings where Equation 2 is not satisfied. □

An advantage of relating the AHP and BAR weights is that the fairly extensive literature (e.g., see [6] and its references) establishing advantages of using the Borda approach become available to AHP. In voting theory, for example, the Borda Count is the unique positional method (tally ballots by assigning specified weights to candidates based on how their position on a ballot) that minimizes the numbers and kinds of consistency paradoxical outcomes that can occur. Many of these positive properties transfer, via the isomorphism, to provide new types of support, or maybe criticism, for AHP.

The next AHP issue is to understand what happens when Equation 2 is not satisfied. As simple examples prove, the useful Equation 3, which relates $a_{i,j}$ values to w_i and w_j, no longer applies. In fact, it is not entirely clear how to interpret the $a_{i,j}$ terms. The AHP literature (e.g., see [9]), on the other hand, proves that AHP proponents strive to attain Equation 2, perhaps by encouraging "users" to refine their $a_{i,j}$ values, or by finding an appropriate $\mathbf{a}' \in \mathcal{SC}_N$ to represent the actual $\mathbf{a} \in \mathbb{R}_+^{\binom{N}{2}}$. This concern is addressed below in three ways:

(1) an explanation is given for the source of $\mathbf{a} \neq \mathcal{SC}_N$,
(2) a nonlinear coordinate system is developed for $\mathbb{R}_+^{\binom{N}{2}}$ that helps to analyze AHP,
(3) and two natural ways to find an $\mathbf{a}' \in \mathcal{SC}_N$ are compared and discussed; one is recommended.

5.4. **Violating AHP consistency; $N = 3$.** A way to preview what happens in general is to develop some of the consequences of the $N = 3$ nonlinear coordinate system. Start by expressing Equation 21 as

$$\mathbf{d} = \sum_{j=1}^{2} \ln(\beta_j)\mathbf{B}_j + \ln(c)\mathbf{C}_{(1,2,3)}, \tag{27}$$

where $\beta_1, \beta_2, c > 0$ and $\mathbf{C}_{(1,2,3)} = (1, -1, 1)$. The $d_{1,2}$ term is $\ln(\beta_1) - \ln(\beta_2) + \ln(c) = \ln(\frac{c\beta_1}{\beta_2})$, and in the same manner, the general form for $\mathbf{d}$ becomes:

$$\mathbf{d} = (\ln(\frac{c\beta_1}{\beta_2}), \ln(\frac{\beta_1}{c}), \ln(c\beta_2)),$$

which defines the representation for $\mathbf{a} \in \mathbb{R}_+^{\binom{3}{2}}$ of

$$\mathbf{a} = (\frac{c\beta 1}{\beta_2}, \frac{\beta 1}{c}, c\beta_2), \quad \beta_1, \beta_2, c \in (0, \infty). \tag{28}$$

As each $\mathbb{R}^3$ point has a unique Equation 27 representation, it follows from the properties of F that each $\mathbb{R}^3_+$ point has a unique Equation 28 representation.

Notice how Equation 28 reflects AHP features; it describes a natural progression from where Equation 2 is satisfied ($c = 1$) to all various levels of inconsistent choices defined by specific $c \neq 1$ values. Each leaf of this natural foliation of $\mathbb{R}^3_+$, which describes all $\mathbf{a}$ values associated with different but specified c levels, is given by

$$\mathcal{L}_c = \{\, (\frac{c\beta 1}{\beta_2}, \frac{\beta_1}{c}, c\beta_2) \,|\, \beta_1, \beta_2 \in (0, \infty) \text{ for a fixed } c\}) \tag{29}$$

These leafs nicely mimic the structure of $\mathcal{SC}_3$, which arises when $c = 1$.

Re-expressing Equation 28 in terms of an $\mathbf{a}' = (a'_{1,2}, a'_{1,3}, a'_{2,3}) \in \mathcal{SC}_3$ (so, for appropriate β_j values, $a'_{1,2} = \frac{\beta_1}{\beta_2}$, $a'_{1,3} = \beta_1$, $a'_{2,3} = \beta_2$), the following statement is obtained.

Theorem 8. *For any* $\mathbf{a} \in \mathbb{R}^{\binom{3}{2}}_+$, *there is a unique* $\mathbf{a}' \in \mathcal{SC}_3$ *and* c *defining*

$$\mathbf{a} = (a'_{1,2}c, \frac{a'_{1,3}}{c}, a'_{2,3}c). \tag{30}$$

Stated in words, an $\mathbf{a} \notin \mathcal{SC}_3$ can be viewed as being a distorted version of a consistent $\mathbf{a}' \in \mathcal{SC}_3$; the distortion is created by the cyclic noise component given by c. (A similar interpretation extends to all $\mathbf{a} \notin \mathcal{SC}_N$ for all $N \geq 3$.) Indeed, the noise, the c terms, twist an $\mathbf{a}' \in \mathcal{SC}_3$ to generate the associated $\mathbf{a}$ terms throughout $\mathbb{R}^{\binom{3}{2}}_+$. The consistent $\mathbf{a}' = (6, 2, \frac{1}{3})$, for instance, defines and is the base for the curve $\{(6t, \frac{2}{t}, \frac{t}{3}) \,|\, 0 < t < \infty\}$; all points on this infinitely long curve are distorted versions of $\mathbf{a}'$.

Finding an appropriate, consistent representative. The next question is to find an $\mathbf{a}' \in \mathcal{SC}_3$ that appropriately represents an $\mathbf{a} \notin \mathcal{SC}_3$. *With no other available information*, a natural choice would be to find an $\mathbf{a}' \in \mathcal{SC}_3$ closest to $\mathbf{a}$. Letting $x = a_{1,2}$, $y = a_{1,3}$, $z = a_{2,3}$, this reduces to the Lagrange multiplier problem of finding such an $\mathbf{a}'$ subject to the constraint that $xz - y = 0$. The answer need not be unique; with $\mathbf{a} = (5, 5, 5)$, for instance, there are two choices: $\mathbf{a}'_1 = (3, 6, 2)$ and $\mathbf{a}'_2 = (2, 6, 3)$ where each is distance $\sqrt{14}$ from $\mathbf{a}$. (The weights associated with the first choice are $w_1 = 2$, $w_2 = 1, w_3 = \frac{1}{3}$.) This approach tacitly assumes that no information exists about the structure of "error behavior;" that is, it is not known how an $\mathbf{a}'$ is converted into an $\mathbf{a}$.

Now that added information about AHP is available, a more natural approach is to use the AHP structure of Equation 30. The objective is to use this equation to drop the noise introduced by c values and to find the associated $\mathbf{a}' \in \mathcal{SC}_3$. Finding $\mathbf{a}'$ is simple; the answer is unique. Illustrating with the example problem of $\mathbf{a} = (5, 5, 5)$, it follows that

$$a'_{1,2} = \frac{5}{c}, \quad a'_{1,3} = 5c, \quad a'_{2,3} = \frac{5}{c},$$

which, with Equation 2, means that $\frac{25}{c^2} = 5c$, or $c = 5^{\frac{1}{3}}$. As such, the unique

$$\mathbf{a}' = (5^{\frac{2}{3}}, 5^{\frac{4}{3}}, 5^{\frac{2}{3}}), \qquad \text{with weights } w_1 = 5^{\frac{2}{3}},\ w_2 = 1,\ w_3 = 5^{-\frac{2}{3}} \tag{31}$$

Although $\mathbf{a}'$ is farther from $(5, 5, 5)$ than $(2, 6, 3)$ or $(3, 6, 2)$, it more accurately reflects the *structure* of AHP by eliminating only the noise caused by the cyclic effects of the data. Moreover, this method faithfully retains the consistent $\mathbf{a}'$ part of the data; the approach of finding the "closest point" to $\mathbf{a}$ does not.

Adding support to this approach is the following somewhat unexpected result.

Theorem 9. *For $N = 3$, the AHP (eigenvalue) weights assigned to $\mathbf{a} \in \mathbb{R}_+^{\binom{3}{2}}$ agree with the AHP weights assigned to the unique $\mathbf{a}' \in \mathcal{SC}_3$ from Equation 30.*

For $N = 3$, then, the $\mathbf{a}'$ found by removing the noise c is compatible with the AHP structure.[1] This assertion introduces another argument for directly removing the noise in an $\mathbf{a}$ data element as compared with using some other approach.

Proof. Because $\mathbf{a}'$ satisfies Equation 2, the $\mathbf{a}$ matrix with noise c becomes

$$A = \begin{pmatrix} 1 & \frac{w_1}{w_2}c & \frac{w_1}{c\,w_3} \\ \frac{w_2}{c\,w_1} & 1 & \frac{w_2}{w_3}c \\ \frac{w_3}{w_1}c & \frac{w_3}{c\,w_2} & 1 \end{pmatrix}$$

where $a'_{i,j} = \frac{w_i}{w_j}$ and the w_j values come from the eigenvector in the $c = 1$ consistency setting. A direct computation shows for this $\mathbf{w} = (w_1, w_2, w_3)$ that

$$A\,\mathbf{w}^T = (1 + c + \frac{1}{c})\mathbf{w}^T,$$

where the superscript "T" represents the transpose. As this expression proves, the main affect of the cyclic "c" value is to change the positive eigenvalue from 3 to $(1 + c + \frac{1}{c})$; the eigenvector that determines AHP weights remains unchanged. □

Another, quicker way to find the AHP weights associated with $\mathbf{a} \in \mathbb{R}^3$ is to use Theorem 7 and Equation 25. Illustrating with the $\mathbf{a} = (5, 5, 5)$ example, the associated choice under the image of F is $\mathbf{d} = (\ln(5), \ln(5), \ln(5))$. Thus

$$\bar{b}_1 = \frac{1}{2}[\ln(5) + \ln(5)] = \ln(5), \bar{b}_2 = \frac{1}{2}[-\ln(5) + \ln(5)] = 0, \bar{b}_3 = \frac{1}{2}[-\ln(5) - \ln(5)] = -\ln(5).$$

According to Equations 25 and 26 with $N = 3$, all AHP weights are common multiples of

$$w_1 = (exp(\ln(5)))^{\frac{2}{3}} = 5^{\frac{2}{3}},\ w_2 = 1,\ w_3 = 5^{-\frac{2}{3}}.$$

These values are precisely those in Equation 31. But these weights differ from what would emerge by using the "closest" $(3, 6, 2)$ or $(2, 6, 3)$. Again, the AHP structure supports reducing the induced noise rather than finding "closest points."

Much of what is developed above extends to $N \geq 3$. This is the topic of the next subsection.

[1]This result extends to $N > 3$ only for noise caused by a single $\mathbf{C}_\pi$ of the kind introduced in Section 4. Readers interested in carrying out details will notice that a multiplier x and its reciprocal $\frac{1}{x}$ appear once in each column of $A\,\mathbf{w}^T$. That is, the new eigenvalue changes from N to $N - 2 + x + \frac{1}{x}$, while the eigenvector remains fixed. But the eigenvectors change should there be two or more $\mathbf{C}_{pi}$ components. Thus, the conclusion does not hold if the $N > 3$ triplets described Section 4 are used to capture the noise because these triplets are linear combinations of the $\mathbf{C}_\pi$ vectors.

5.5. **AHP for $N \geq 3$.** The following theorem describes what happens for any $N \geq 3$. A word of caution; there are many convenient ways to represent a basis for $\mathcal{C}_N$, which means that there are many ways to represent the nonlinear structure of $\mathbb{R}_+^{\binom{N}{2}}$; each of which reflects the AHP structure. As all choices are equivalent, a choice that is convenient for a particular problem should be selected. The selection in Theorem 10 emphasizes the structure of triplets (as described in Theorem 3 and used in the proof of Theorem 4) was chosen only because it was easier to use with certain examples; e.g., if certain triplets satisfy Equation 2, then the associated $c_{i,j}$ value can be set equal to unity.

Theorem 10. *For $N \geq 3$, a nonlinear coordinate system of $\mathbb{R}_+^{\binom{N}{2}}$ that reflects the structure of AHP is given by the the $\binom{N}{2}$ positive parameters $\beta_1, \dots, \beta_{N-1}$ and $\{c_{i,j}\}_{2 \leq i < j \leq N}$. They are combined to create the coordinates in the following manner:*

- *For $2 \leq i < j < N$, let $a_{i,j} = \frac{\beta_i}{\beta_j} c_{i,j}$.*
- *For $2 \leq i < N$, let $a_{i,N} = \beta_i c_{i,N}$.*
- *For $1 < j \leq N$ let $a_{1,j} = \frac{\beta_1}{\beta_j} C_{1,j}$ where $C_{1,j} = \frac{\Pi_{k>j} c_{j,k}}{\Pi_{k<j} c_{k,j}}$.*

These coordinates satisfy the AHP consistency equation, Equation 2, iff all $c_{j,k} = 1$.

This result immediately leads to the following statement of Corollary 2. This result asserts that associated with each $\mathbf{a} \in \mathbb{R}_+^{\binom{N}{2}}$ is a unique $\mathbf{a}' \in \mathcal{SC}_N$ and (once the form of the noise is selected; the triplet form is used here) a unique choice and structure for the associated noise. The contributions of this corollary are similar to that discussed above for $N = 3$; it provides a natural interpretation for a non-consistent $\mathbf{a}$ (in terms of the cyclic noise distorting the associated base $\mathbf{a}'$ values), it identifies a way to find an appropriate, consistent $\mathbf{a}'$ that is associated with $\mathbf{a}$, it provides a foliation of the full space that can be used to understand the levels and kinds of inconsistencies (generated by different $c_{i,j}$ combinations and values), and so forth.

Corollary 2. *For any $\mathbf{a} \in \mathbb{R}_+^{\binom{N}{2}}$, there exists a unique $\mathbf{a}' \in \mathcal{SC}_N$ and $\binom{N-1}{2}$ values of $c_{i,j} > 0$ so that for $2 \leq i < j < N$, $a_{i,j} = a'_{i,j} c_{i,j}$; for $2 \leq i < N$, $a_{i,N} = a'_{i,N} c_{i,N}$, and for $i = 1 < j \leq N$ $a_{1,j} = a'_{i,j} C_{1,j}$ where $C_{1,j} = \frac{\Pi_{k>j} c_{j,k}}{\Pi_{k<j} c_{k,j}}$.*

Proof. The coordinate representation for $\mathbb{R}_+^{\binom{N}{2}}$ follows directly from the coordinate representation of $\mathbb{R}^{\binom{N}{2}}$

$$\mathbf{d} = \sum_{i=1}^{N-1} \ln(\beta_i) \mathbf{B}_i + \sum_{1<i<j} \ln(c_{i,j}) \mathbf{C}_{i,j} \tag{32}$$

where scalars b_i, $c_{i,j} > 0$, and the vectors $\mathbf{C}_{i,j}$ represent the cyclic terms given by triplets of the $\{1, i, j\}$ form. Thus, vector $\mathbf{C}_{i,j}$ has only three non-zero terms; two are where the $\{1, i\}$ and $\{i, j\}$ terms equal unity while the third, the $\{1, j\}$ term, equals -1. According to the structure of these basis vectors, the $1 < i < j < N$ term becomes $d_{i,j} = \ln(\frac{\beta_i c_{i,j}}{\beta_j})$, which leads to the $a_{i,j}$ representation. The only difference for $1 < i < N$ is that there are no β_N values, so $d_{1,N} = \ln(\beta_i c_{i,N}))$.

Some bookkeeping of the indices is needed to handle the $1 < j$ setting. For each k, $1 < k < j$, the index pairs $\{1, k\}$ and $\{k, j\}$ have a positive entry in $\mathbf{C}_{k,j}$, while $\{1, j\}$ has the negative value. This means that this cyclic term adds $-\ln(c_{k,j})$ to the $\ln(\frac{\beta_1}{\beta_j})$ value. In the other direction, for each k, $1 < j < k$, the $\{1, j\}$ index pair as a positive entry, so the $\ln(\frac{\beta_1}{\beta_j})$ value is changed by adding the $\ln(c_{j,k})$ values. When these logarithmic terms are combined, and transferred via F, the representation given in Theorem 10 follows.

Each point in $\mathbb{R}_+^{\binom{N}{2}}$ is uniquely represented by choices of $\beta_j, c_{i,j}$ because Equation 32 is a coordinate system for $\mathbb{R}^{\binom{N}{2}}$ and the two spaces are isomorphic.

It is immediate to prove that these coordinates satisfy Equation 2 iff all $c_{j,k} = 1$. □

Example ($N = 4$): A nonlinear coordinate system for $N = 4$ is

$$\mathbf{a} = (a'_{1,2}c_{2,3}c_{2,4}, \quad a'_{1,3}\frac{c_{3,4}}{c_{2,3}}, \quad \frac{a'_{1,4}}{c_{2,4}c_{3,4}}, \quad a'_{2,3}c_{2,3}, \quad a'_{2,4}c_{2,4}, \quad a'_{3,4}c_{3,4}), \tag{33}$$

where

$$\mathbf{a}' = (a'_{1,2},\, a'_{1,3},\, a'_{1,4}, a'_{2,3},\, a'_{2,3},\, a'_{3,4}) = (\frac{\beta_1}{\beta_2},\, \frac{\beta_1}{\beta_3},\, \beta_1,\, \frac{\beta_2}{\beta_3},\, \beta_2,\, \beta_3) \in \mathcal{SC}_4$$

satisfies Equation 2; all $\beta_j, c_{j,k}$ values are positive.

As an illustration, the $\beta_1 = 2$, $\beta_2 = 3$, $\beta_3 = 4$ values define $\mathbf{a}' = (\frac{2}{3}, \frac{1}{2}, 2, \frac{3}{4}, 3, 4) \in \mathcal{SC}_4$. Thus the three-dimensional surface defined by the expression

$$\mathbf{a}_{x,y,z} = (\frac{2}{3}xy,\, \frac{1}{2}\frac{z}{x},\, \frac{2}{yz},\, \frac{3}{4}x,\, 3y,\, 4z), \quad 0 < x, y, z < \infty$$

represents all possible "noise" distortions of this original $\mathbf{a}'$. (Here, $x = c_{2,3}$, $y = c_{2,4}$, $z = c_{3,4}$.) Without help from the coordinate system (Equation 33), there is no way (at least that I can see) to realize that $\mathbf{a}_{2,3,3} = (4, \frac{3}{4}, \frac{2}{9}, \frac{3}{2}, 9, 12)$ is AHP related to $\mathbf{a}_{3,2,1} = (4, \frac{1}{6}, 1, \frac{9}{4}, 6, 4)$ in that both share the same base $\mathbf{a}' = (\frac{2}{3}, \frac{1}{2}, 2, \frac{3}{4}, 3, 4)$.

Now consider $\mathbf{a} = (4,\, 6,\, 1.5,\, 3,\, 3,\, 8) \notin \mathcal{SC}_4$ where the goal is to find an appropriate $\mathbf{a}'$. The Equation 33 coordinate system (with $x = c_{2,3}$, $y = c_{2,4}$, $z = c_{3,4}$) leads to the expressions

$$a'_{1,2} = \frac{4}{xy},\, a'_{1,3} = 6\frac{x}{z},\, a'_{1,4} = 1.5yz,\, a'_{2,3} = \frac{3}{x},\, a'_{2,4} = \frac{3}{y},\, a'_{3,4} = \frac{8}{z},$$

which are then substituted into the three consistency constraints to obtain the three equations with three unknowns: $a'_{1,2}a'_{2,3} = a'_{1,3}$ or $(\frac{4}{xy})(\frac{3}{x}) = 6\frac{x}{z}$; $a_{1,2}a_{2,4} = a_{1,4}$ or $(\frac{4}{xy})(\frac{3}{y}) = 1.5yz$, and $a_{1,3}a_{3,4} = a_{1,4}$, or $(6\frac{x}{z})(\frac{8}{z}) = 1.5yz$. (A preferred way to handle these systems is to use F to convert the system into easier solved linear equations.) The solution $x = 2$, $y = 1$, $z = 4$ leads to $\mathbf{a}' = (2, 3, 6, 1.5, 3, 2)$. The associated weights, then, are multiples of $\mathbf{w} = (2, 1, \frac{2}{3}, \frac{1}{3})$.

As advertised, a much simpler, more elementary approach to find these $\mathbf{a}'$ weights is to use Theorem 7 and Equation 25. The starting $\mathbf{a} = (4, 6, 1.5, 3, 3, 8)$ values define $d_{1,2} = \ln(4)$, $d_{1,3} = \ln(6)$, $d_{1,4} = \ln(1.5)$, $d_{2,3} = \ln(3)$, $d_{2,4} = \ln(3)$, $d_{3,4} = \ln(8)$. Thus

$$\begin{array}{llll} \overline{b}_1 & = \frac{1}{2}[\ln(4) + \ln(6) + \ln(1.5)] = \ln(6), & \overline{b}_2 & = \frac{1}{2}[-\ln(4) + \ln(3) + \ln(3)] = \ln(\frac{3}{2}) \\ \overline{b}_3 & = \frac{1}{2}[-\ln(6) - \ln(3) + \ln(8)] = \ln(\frac{2}{3}), & \overline{b}_4 & = \frac{1}{2}[-\ln(1.5) - \ln(3) - \ln(8)] = \ln(\frac{1}{6}) \end{array}$$

According to Equation 26 (with $N = 4$) the weights are a common multiple of $w_1 = \sqrt{6}$, $w_2 = \sqrt{\frac{3}{2}}$, $w_3 = \sqrt{\frac{2}{3}}$, $w_4 = \sqrt{\frac{1}{6}}$. Dividing each by $\sqrt{\frac{3}{2}}$ (so that $w_2 = 1$) yields the $w_1 = 2$, $w_2 = 1$, $w_3 = \frac{2}{3}, w_4 = \frac{1}{3}$ choice of the previous paragraph.

6. Summary

It is interesting how any paired comparison data space can be decomposed into those data aspects that satisfy a strong transitivity conditions and those data combinations that create cyclic effects. Adding to the interest is that nothing goes wrong on the $\mathcal{ST}_N$ subspace; on this data subspace, paired comparison rules can be expected to agree. An immediate consequence of this observation is that all disagreements among different rules, all paradoxical kinds of behavior that occur, must be caused by the way in which rules react to the cyclic data terms from $\mathcal{C}_N$.

Part of the utility of this decomposition is displayed by how it simplifies the analysis of different rules. The properties of the additive BAR rule, for instance, are quickly extracted; BAR outcomes strictly depend upon the $\mathcal{ST}_N$ components of data and they cancel the cyclic $\mathcal{C}_N$ terms. Moreover, when concentrating on the strongly transitive terms, the BAR difference for any specified pair strictly depends on the data components for these two alternatives; what happens with other alternatives is irrelevant. (A word of caution; this statement does not mean that BAR satisfies Independence of Irrelevant Alternatives from Arrow's seminal theorem [1]. This feature, which will be discussed elsewhere, is affected, in a subtle manner, by $\mathcal{C}_N$ terms.)

The value of this decomposition also is demonstrated by how it simplifies answering some natural questions about AHP. It shows that the multiplicative consistency condition (Equation 2) can be equated with the strong transitivity condition; it shows that terms not satisfying this condition can be uniquely represented in terms of an entry that satisfies Equation 2, which is then distorted by multiplicative cyclic effects. This nonlinear coordinate system provides the appropriate AHP structure to analyze the structure of these non-conforming data choices, and it simplifies the selection of an associated consistent data term for a given non-consistent choice. (For the second, just eliminate the noise.) Of particular surprise is how this characterization of the cyclic noise provides an elementary and quick way to compute the AHP weights for a "corrected" data structure.

References

[1] Arrow, K., 1951 (2nd ed. 1963), *Social Choice and Individual Values*, Wiley, New York.

[2] Condorcet, M. 1785, *Éssai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris.

[3] Horn, R., and C. Johnson, 1990, *Matrix Analysis,* Cambridge University Press.

[4] Kemeny, J., 1959, Mathematics without numbers. *Daedalus* **88**, 571-591.

[5] Saari, D. G., 2000, Mathematical structure of voting paradoxes 1; pairwise vote. *Economic Theory* **15**, 1-53.

[6] Saari, D. G., 2008, *Disposing Dictators, Demystifying Voting Paradoxes*, Cambridge University Press.

[7] Saari, D. G., and K. Sieberg, 2004. Are part wise comparisons reliable? *Research in Engineering Design* **15**: 62-71.

[8] Saaty, T., 1980, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, Mcgraw-Hill

[9] Saaty, T., 2009, *Principia Mathematica Decernendi: Mathematical Principles of Decision Making*, RWS Publications.