A Geometric Look at Manipulation

Jan van Eijck

CWI, Amsterdam

Abstract. We take a fresh look at voting theory, in particular at the notion of manipulation, by employing the geometry of the Saari triangle. This yields a geometric proof of the Gibbard/Satterthwaite theorem, and new insight into what it means to manipulate the vote. Next, we propose two possible strengthenings of the notion of manipulability (or weakenings of the notion of non-manipulability), and analyze how these affect the impossibility proof for non-manipulable voting rules.

1 Introduction

We start with fixing some terminology, mostly following the conventions of [11].

Let A be a finite set of goods, with $|A| > 2$. An A-*ballot* is a linear ordering of A. Let $\{1, \ldots, n\}$ be a set of voters. An (A, n)-*profile* is an n-tuple of A-ballots. If $\mathbf{P}$ is an (A, n)-profile, then $\mathbf{P}$ can be written as $(>_1, \ldots, >_n)$. $>_i$, the i-th component of profile $(>_1, \ldots, >_n)$, is the ballot of voter i. $>_i$ expresses "what voter i wants."

Let $\mathbf{P}(A)$ be the set of all (A, n)-profiles, for given $n \in \mathbb{N}$. A function $V : \mathbf{P}(A) \to A$ is a *resolute voting rule* for A. A function $V : \mathbf{P}(A) \to \mathcal{P}^+(A)$ is a *voting rule* for A. A function $V : \mathbf{P}(A) \to \mathcal{P}^+(A) \to \mathcal{P}^+(A)$ with the property that

$$V(\mathbf{P})(v) \subseteq v$$

is a *social choice function* for A. Let $\mathbf{ord}(A)$ be the set of all linear orderings of A. A function

$$V : \mathbf{P}(A) \to \mathbf{ord}(A)$$

is a *social welfare function* for A. A social welfare function transforms a sequence of ballots into a single ballot.

When a linear preference order $>$ on A is mentioned, we will use $<$ for $\{(x, y) \mid y > x\}$, $\geq$ for $\{(x, y) \mid x > y \vee x = y\}$, and $\leq$ for $\{(x, y) \mid x < y \vee x = y\}$.

Definition 1. *Let* $\mathbf{P} \sim_i \mathbf{P}'$ *express that* $\mathbf{P}$ *and* $\mathbf{P}'$ *differ only in the ballot of voter* i.

A voting rule satisfies *Pareto* if no x is winning if there is some y that every voter prefers to x:

$$\forall \mathbf{P} \forall x \in V(\mathbf{P}) \forall y \in A(y \neq x \longrightarrow \exists i \in N : \; x \geq_i y).$$

A resolute voting rule V is *non-manipulable (NM)* (or: *strategy-proof*) if $\mathbf{P} \sim_i \mathbf{P}'$ implies $V(\mathbf{P}) \geq_i V(\mathbf{P}')$.

Note that in $\mathbf{P}$, "what voter i wants" is expressed by $>_i$, and in $\mathbf{P}'$, "what voter i wants" is expressed by $>_i'$. If changing the ballot from $>_i$ to $>_i'$ gives a better outcome (better, given $>_i$) than sticking to $>_i$, then the voting rule invites strategic voting.

A voting rule V is *non-imposed* if any candidate can be a winner: $\forall a \in A \exists \mathbf{P} : a \in V(\mathbf{P})$. We will use a slightly weaker property. A resolute voting rule V is *weakly non-imposed (NI)* if at least three outcomes are possible: $|\{x \mid \exists \mathbf{P} : V(\mathbf{P}) = x\}| \geq 3$.

A resolute voting rule V is a *dictatorship* if there is some k such that $V : \mathbf{P}(A) \to A$ maps any $\mathbf{P}$ to the top ranking item in $>_k$.

A voter i is *effective* (or: *pivotal*) for V and $\mathbf{P}$ if there is some $\mathbf{P}'$ with $\mathbf{P} \sim_i \mathbf{P}'$ and $V(\mathbf{P}) \neq V(\mathbf{P}')$.

Here is what the famous 'Gibbard-Satterthwaite Theorem [5, 10] states:

GS. Any resolute voting rule that is NM and that is NI is a dictatorship.

In this paper we will reflect on the theorem by starting out with giving a new easy proof, and then analyzing what makes the proof so easy.

> The more important a theorem is, the harder we should look for explanations why it is true. And the Gibbard-Satterthwaite theorem is extremely important. [11]

The Gibbard/Satterthwaite theorem is closely related to Arrow's Theorem [1], although the GS theorem is about (resolute) voting rules, while Arrow's theorem is about social welfare functions. A proof in dialogue form of Arrow's theorem, with comments on the connection with Gibbard/Satterthwaite, is in [4]. Geometric proofs of Arrow's theorem exist [9, 7]. The geometric approach to paradoxes of preference aggregation from [9] is extended in [3] to paradoxes of judgement aggregation. In jugdgement aggregation not all outcomes are possible because the judgements are logically interconnected.

In this paper we give a geometric proof of the Gibbard-Satterthwaite Theorem. Our first aim is to make (still) clearer *why the GS Theorem is true*. It will turn out that our proof is easy because the notion of manipulability is strong. Next, we will analyse what the proof tells us about the notion of manipulability, and study what some slight modifications of this notion would do to the proof.

2 Geometry of Voting: the Saari Triangle

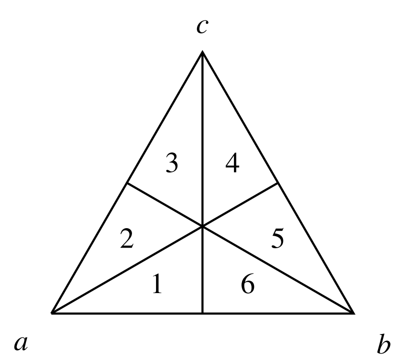
If a voter i changes his ballot from $>_i$ to $>'_i$, then this change can be decomposed into a sequence of adjacent transpositions. E.g., the change from *abcd* to *cbad* can be decomposed into

$$abcd \to bacd \to bcad \to cbad.$$

First a and b are swapped, then c and a, and finally b and c.

We call such an adjacent transposition where x and y are swapped from a situation where x is preferred over y to a situation where y is preferred over x an $x : y$ crossing. Geometrically, an $x : y$ crossing crosses the line between the set of ballots where x is preferred over y to the set of ballot where y is preferred over x.

For the case of three alternatives, this is made clear in the geometry of profiles given by the Saari Triangle from [9]. Following [9], we call call the voters that have ballot $a > b > c$ voters of type 1. The voters of type 2 are those that have ballot $a > c > b$. The voters of type 3 have ballot $c > a > b$. The voters of type 4 have ballot $c > b > a$. The voters of type 5 have ballot $b > c > a$. The voters of type 6 have ballot $b > a > c$.



Here is how to read this. The closer a region is to a vertex, the more preferred the vertex. Now the six regions represent the six types of voters. The subtriangle marked 1 is closest to the a vertex and farthest from the c vertex, so this area represents the $a > b > c$ voters. Note that every time a boundary between regions in the triangle gets crossed, one binary preference gets swapped. E.g., in crossing from the 1 into the 2 region, $b < c$ gets swapped to $c < b$.

Geometrically, $\mathbf{P} \sim_i \mathbf{P}'$ gets represented as a single voter i changing type by moving from one region in the triangle to another.

To allow draws in the ballots one can allow voters positioned on the lines. It is easy to extend all that follows to cover such cases as well; it is left to the reader to check that the argument is not affected by this.

3 A 'Geometric' Proof of the Gibbard Satterthwaite Theorem

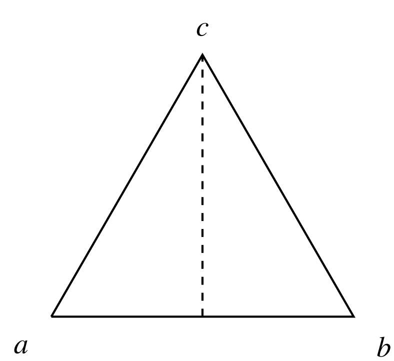
We will now chart the possible effects of minimal type changes, on the assumption that the voting rule satisfies NM. It will turn out that the constraints on vote changing that follow from NM are very strong.

The Crossing Lemma describes the possible effects of $x : y$ crossings on the outcome of the vote, given that the voting rule satisfies NM. Note that the lemma holds for any number of alternatives.

Lemma 1 (Crossing Lemma). *Let V be NM, and let $\mathbf{P} \sim_i \mathbf{P}'$ be such that $>_i$ and $>'_i$ are related by an adjacent transposition that exchanges x and y (with $x >_i y$). Then $V(\mathbf{P}) \neq V(\mathbf{P}')$ implies $V(\mathbf{P}) = x$ and $V(\mathbf{P}') = y$.*

Proof. Suppose V, $\mathbf{P}$, $\mathbf{P}'$, i, x, y are as in the statement of the Lemma. Assume $V(\mathbf{P}) \neq V(\mathbf{P}')$. By NM we have that $V(\mathbf{P}) \geq_i V(\mathbf{P}')$ and $V(\mathbf{P}) \leq'_i V(\mathbf{P}')$. Since the two orderings differ only in the positions of x and y, and x and y are adjacent, it follows that $V(\mathbf{P}) = x$ and $V(\mathbf{P}') = y$. (This continues to hold if ballots are allowed to have ties.) □

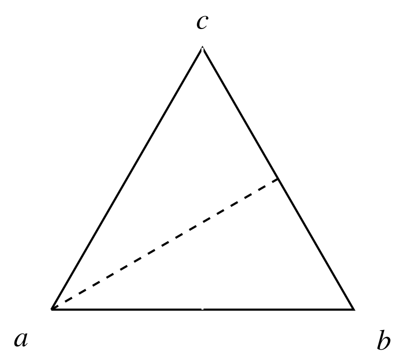
For the Saari triangle, where the regions are arranged by adjacent transpositions, we can visualize the constraints on crossing the $a \sim b$ divide as follows:



$a \sim b$: dividing line between a and b regions

The only vote change that can take place when this line is crossed is from a to b if the line is crossed from the a region into the b region, and from b to a if it is crossed in the other direction.

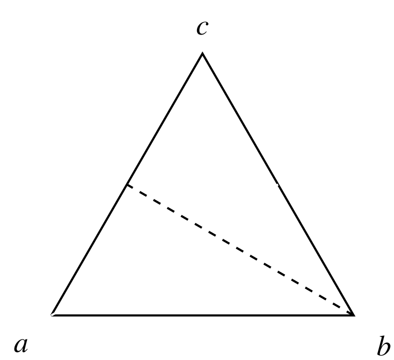
Similarly for the consequences of NM for crossing the $b \sim c$ divide:



$b \sim c$: dividing line between b and c regions

The only shift in the vote (given NM) that can take place when this line is crossed from the b to the c region is from b to c, and vice versa.

Finally, the consequences of NM for crossing the $a \sim c$ divide:



$a \sim c$: dividing line between a and c regions

Crossing from a to c by a single voter can only cause a vote shift from a to c, and crossing in the other direction can only cause a vote shift from c to a.

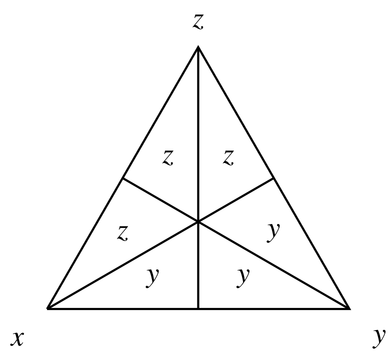
Summing up, we get the following **crossing rule** for crossings in the Saari triangle:

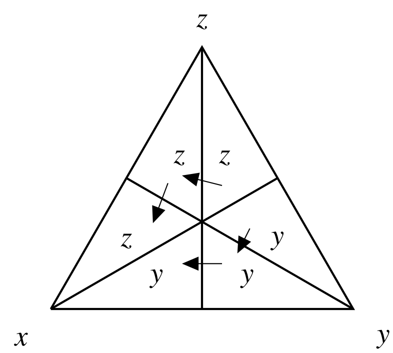
Fact 1. *In crossing the $x \sim y$ divide, from the x region into the y region (call this an $x : y$ crossing) the only value change that is allowed is from x to y.*

The Crossing Lemma is our main tool to prove the following Effectiveness Lemma:

Lemma 2 (Effectiveness Lemma). *If V is NM and NI, and i is effective for* $\mathbf{P}$*, then $V(\mathbf{P})$ is equal to the top of the i-ballot in* $\mathbf{P}$.

Proof. Assume V is NM and i is effective for $\mathbf{P}$, and suppose for a contradiction that $V(\mathbf{P}) <_i x$, where x is the top element of the i-ballot in $\mathbf{P}$. Since i is effective, there is a profile $\mathbf{P}'$ with $\mathbf{P} \sim_i \mathbf{P}'$ and $V(\mathbf{P}) \neq V(\mathbf{P}')$. By NM, $V(\mathbf{P}) >_i V(\mathbf{P}')$. Let $y = V(\mathbf{P})$ and $z = V(\mathbf{P}')$. Then the i-ballot in $\mathbf{P}$ has the pattern $x \cdots \dot{y} \cdots z$, with $\dot{y}$ indicating that y is the outcome of the vote. Now change all ballots in $\mathbf{P}$ by pushing values different from x, y, z below the third position, while keeping the order of x, y, z. Let the result be $\mathbf{Q}$. Then by the Crossing Lemma, $V(\mathbf{Q}) = V(\mathbf{P}) = y$. So i-ballot and result of the vote in $\mathbf{Q}$ are given by $x\dot{y}z$. All ballots different from those with one of the six permutations of x, y, z on top are irrelevant for the argument that follows, and we can disregard them. By the Crossing Lemma, the only consistent configuration for how the vote can change as i moves through the relevant ballots is given by: $x\dot{y}z \sim_i x\dot{z}y \sim_i \dot{z}xy \sim_i \dot{z}yx \sim_i \dot{y}zx \sim_i \dot{y}xz$ $(\sim_i x\dot{y}z)$. See the lefthand side picture below.





Suppose some other agent j is able to influence the vote. Then all j can do is make the vote switch to x. The only way of doing that is by moving x up in his ballot. This ballot change can be decomposed into adjacent transposition steps. According to the Crossing Lemma, this is what j could do: move from zxy to xzy to make the vote switch from z to x, move from zyx to zxy to make the vote switch from z to x, move from yxz to xyz to make the vote switch from y to x, or move from yzx to yxz to make the vote switch from y to x. See the righthand side picture above. Here are two of the cases:

$$\begin{tikzcd}
\begin{array}{l} i : x\dot{y}z \\ j : \dot{y}xz \end{array} \arrow[r, phantom, "\sim_i"] \arrow[d, "j"'] & \begin{array}{l} i : x\dot{z}y \\ j : yx\dot{z} \end{array} \arrow[d, "j"] \\
\begin{array}{l} i : \dot{x}yz \\ j : \dot{x}yz \end{array} \arrow[r, phantom, "\sim_i"] & \begin{array}{l} i : x\dot{z}y \\ j : xy\dot{z} \end{array}
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
\begin{array}{l} i : x\dot{y}z \\ j : z\dot{y}x \end{array} \arrow[r, phantom, "\sim_i"] \arrow[d, "j"'] & \begin{array}{l} i : x\dot{z}y \\ j : \dot{z}yx \end{array} \arrow[d, "j"] \\
\begin{array}{l} i : \dot{x}yz \\ j : z\dot{x}y \end{array} \arrow[r, phantom, "\sim_i"] & \begin{array}{l} i : x\dot{z}y \\ j : \dot{z}xy \end{array}
\end{tikzcd}$$

In both cases, $\dot{x}yz \sim_i x\dot{z}y$ in the bottom line contradicts the Crossing Lemma. The other two cases are similar. So, x cannot be forced, and contradiction with NI. This proves the Lemma. □

Theorem 1 (Gibbard-Satterthwaite). *Any resolute voting rule that is NM and NI is a dictatorship.*

Proof. Let V be a resolute voting rule that is NM and NI. Then by NI, there has to be a profile $\mathbf{P}$ with at least one effective voter i.

Suppose $\mathbf{P}$ has another effective voter j. By the Effectiveness Lemma, i determines the vote for every $\mathbf{P}'$ with $\mathbf{P} \sim_i \mathbf{P}'$, and j determines the vote for every $\mathbf{P}''$ with $\mathbf{P} \sim_j \mathbf{P}''$. By the Effectiveness Lemma, $V(\mathbf{P}) = x$ is the favourite of both i and j in $\mathbf{P}$. By the Effectiveness Lemma, $V(\mathbf{P}') = y \neq x$ is the favourite of i in $\mathbf{P}'$, while the favourite of j in $\mathbf{P}'$ is still x. By the Effectiveness Lemma, $V(\mathbf{P}'') = z \neq x$ is the favourite of j in $\mathbf{P}''$, while the favourite of i in $\mathbf{P}''$ is still x. We may assume that $z \neq y$, for by NI there is some z different from both x and y that can be the outcome of the vote, and if j moves z to the top of her ballot, z will be the outcome of the vote by the Effectiveness Lemma. Let $\mathbf{Q}$ be the result of both i and j changing their ballots from those in $\mathbf{P}$, i to his $\mathbf{P}'$ ballot and j to her $\mathbf{P}''$ ballot. Then $\mathbf{P}' \sim_j \mathbf{Q} \sim_i \mathbf{P}''$. Suppose $V(\mathbf{Q}) \neq V(\mathbf{P}')$. Then j is effective in $\mathbf{P}'$, so by the Effectiveness Lemma, $V(\mathbf{P}')$ should equal the favourite of j in $\mathbf{P}'$, which is not the case. So $V(\mathbf{Q}) = V(\mathbf{P}')$. This means that i is effective in $\mathbf{P}''$. By the Effectiveness Lemma, $V(\mathbf{P}'')$ should equal the favourite of i in $\mathbf{P}''$, which is not the case. Contradiction.

So i is the one and only effective voter for $\mathbf{P}$. But then i must be the one and only effective voter for *any* $\mathbf{P}$, and therefore i is the dictator. □

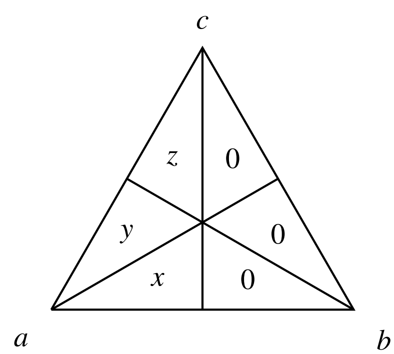
4 Some Other Properties of Resolute Voting Rules

The following theorem provides another example of use of the Saari triangle for proving simple properties of voting rules.

Theorem 2. *For any resolute voting rule V that satisfies NM it holds that V satisfies Pareto iff V satisfies NI.*

Proof. $\Rightarrow$: We will show that any resolute voting rule V that satisfies NM but not Pareto is imposed.

Suppose V satisfies NM, but not Pareto. Then (wlog) there is a profile $\mathbf{P}$ that has everywhere a above b, but $V(\mathbf{P}) = b$.



Check that type changes to the empty regions, moving clockwise, can never produce a V-value different from b. So V is imposed. (In fact, no outcome other than b is possible!)

$\Leftarrow$: Let V satisfy NM and Pareto. Suppose for a contradiction that V is imposed, i.e., for any profile $\mathbf{P}$, either $V(\mathbf{P}) = x$ or $V(\mathbf{P}) = y$. Since we assume that $|A| > 2$ there is some z different from both x and y. Consider some $\mathbf{P}$ with $V(\mathbf{P}) = x$. Move any alternative different from x, y, z down below any of these three alternatives in all ballots of $\mathbf{P}$. Call the resulting profile $\mathbf{Q}$. By Pareto, $V(\mathbf{Q}) = x$. Let $\mathbf{Q}'$ be the result of all voters changing their ballots in $\mathbf{Q}$ by moving z to the top position of their ballot. Then by Pareto, $V(\mathbf{Q}') = z$ and contradiction with the assumption that V is imposed. □

A resolute voting rule V is *monotonic* if whenever $V(\mathbf{P}) = a$ and $\mathbf{P}'$ is the result of changing each $>_i$ to $>'_i$ in such a way that for all $b \in A$, $a >_i b$ implies $a >'_i b$, then $V(\mathbf{P}') = a$.

Theorem 3. *If V is NM then V is monotonic.*

Proof. It follows immediately from the crossing lemma that any adjacent transposition that does not move a b up past a winning a will not make the vote change from a to b. □

Theorem 4. *If V is mononotic then V is NM.*

Proof. Assume V is monotonic. Suppose for a contradiction that V can be manipulated. Then there has to be a pair $\mathbf{P} \sim_i \mathbf{P}'$ such that $V(\mathbf{P}') >_i V(\mathbf{P})$. Since any individual vote change can be decomposed into adjacent transpositions, there has to be a pair $\mathbf{P} \sim_i \mathbf{P}_i$ with $V(\mathbf{P}') >_i V(\mathbf{P})$ and such that an adjacent transposition relates $>_i$ to $>'_i$. Let $V(\mathbf{P}') = a$ and $V(\mathbf{P}) = b$. Then $>_i = \alpha ab\beta$ and $>'_i = \alpha ba\beta$. By monotonicity, changing $>'_i$ to $>_i$ by means of moving a up over b should not change the vote, and contradiction with $V(\mathbf{P}) = b$. □

Theorem 5 (Muller-Satterthwaite [6]). *Any resolute voting rule V that is* monotonic *and satisfies Pareto is a dictatorship.*

Proof. Let V be a resolute voting rule that is *monotonic* and satisfies Pareto. Then by Theorem 4, V is NM. By Theorem 2, V also is NI. It follows from Theorem 1 that V is a dictatorship. □

This ends our discussion of the proof of the Gibbard-Satterthwaite theorem and related results.

5 Reflections on Manipulation

In this section we will argue that the notion of manipulation that was used for proving the Gibbard-Satterthwaite theorem is too general to serve as a useful classifier of voting rules. If we define our concept of sin in such manner that every human (including saints and law-abiding citizens) is a sinner, then we can take this as a condemnation of humanity, but we can also conclude that there may be something wrong with our definition of sin. To escape from the uniform condemnation of humanity, we could ask ourselves if some sins are perhaps worse than others.

So let us ask ourselves some questions about possible manipulations. Which of the following is worse?

champion rearrangement My preferences are $a > b > c$, and the V-value is c. I switch my preferences to $b > a > c$ and the value becomes b.

champion bashing My preferences are $a > b > c$, and the V-value is c. I switch my preferences to $b > c > a$ and the value becomes b.

In the first case, the only thing that happens is that a does not beat c, but I have another candidate b for beating c. So I push b, and it turns out b is better for the job.

One might think about vote manipulation from the perspective of reasoning about other minds, as follows. Taking an epistemic perspective on manipulation, we see:

- I *know* that the outcome of the voting process if I stick to ordering $a > b > c$ is c. This in information about *what the others think*.
- I *know* that the outcome if I change the order of my two most favoured candidates is b. This is information about *how others would react if I readjust*.
- Who, in his right mind, would not readjust? Not adjusting would be worse than a crime: it would be a stupidity.

– I can even explain it to a: "Sorry, in the circumstances you are not the right choice. If I insist on you, I will not be able to beat c. But I will not deny that you are way better than that abject c."

These considerations show that 'manipulation' is simply too coarse for making useful distinctions. To illustrate that it is possible to do better, we propose a couple of notions that are stronger than manipulation.

Definition 2. *The* knights *of a voter i, given profile* $\mathbf{P}$ *and resolute voting rule V, are the goods that are above* $V(\mathbf{P})$ *on the i-ballot. The* knaves *of a voter i, given profile* $\mathbf{P}$ *and resolute voting rule V, are the goods that are below* $V(\mathbf{P})$ *on the i-ballot.*

For example, if i has ballot $a > b > c > d$ in $\mathbf{P}$, and the outcome of the vote is c, then a and b are knights of i in $\mathbf{P}$, and d is a knave of i in $\mathbf{P}$.

Definition 3. *A resolute voting rule V is* demotion pervertible *(DP) if there exists an i-minimal pair of profiles* $\mathbf{P}$, $\mathbf{P}'$ *such that*

– $V(\mathbf{P}) <_i V(\mathbf{P}')$*, and*
– $\exists x : \; V(\mathbf{P}) <_i x <_i' V(\mathbf{P})$.

A resolute voting rule V is NDPe *(non-demotion-pervertible) if V is not DP.*

Note: the demotion of knight x from above $V(\mathbf{P})$ to a new position below $V(\mathbf{P})$ is the perversion.

For example, suppose i has ballot *abcd* in $\mathbf{P}$, and the outcome of the vote is c, and $\mathbf{P} \sim_i \mathbf{P}'$ where i has ballot *bcad* in $\mathbf{P}'$, and the outcome of the vote in $\mathbf{P}'$ is b. Then V is demotion pervertible, for we have that $V(\mathbf{P}) = c <_i b = V(\mathbf{P}')$, and $V(\mathbf{P}) = c <_i a <_i' c = V(\mathbf{P})$, that is to say, a was demoted from a knight to a knave position (from the perspective of $\mathbf{P}$).

Definition 4. *A resolute voting rule V is* promotion pervertible *(PP) if there exists an i-minimal pair of profiles* $\mathbf{P}$, $\mathbf{P}'$ *such that*

– $V(\mathbf{P}) <_i V(\mathbf{P}')$*, and*
– $\exists x : \; V(\mathbf{P}) <_i' x <_i V(\mathbf{P})$.

A resolute voting rule V is NPPe *(non-promotion-pervertible) if V is not PP.*

Note: the promotion of knave x from below $V(\mathbf{P})$ to a new position above $V(\mathbf{P})$ is the perversion.

For example, suppose i has ballot *abcd* in $\mathbf{P}$, and the outcome of the vote is c, and $\mathbf{P} \sim_i \mathbf{P}'$ where i has ballot *abdc* in $\mathbf{P}'$, and the outcome of the vote in $\mathbf{P}'$ is b. Then V is promotion pervertible, for we have that $V(\mathbf{P}) = c <_i b = V(\mathbf{P}')$, and $V(\mathbf{P}) = c <_i' d <_i c = V(\mathbf{P})$, that is to say, d was promoted from a knave to a knight position (from the perspective of $\mathbf{P}$).

Definition 5. *An i-minimal pair of profiles* $\mathbf{P}$, $\mathbf{P}'$ invites decency towards knights *if the following hold:*

- $\exists x : V(\mathbf{P}) <_i x <'_i V(\mathbf{P})$ *implies* $V(\mathbf{P}) \geq_i V(\mathbf{P}')$*,*
- $\exists x : V(\mathbf{P}') <'_i x <_i V(\mathbf{P}')$ *implies* $V(\mathbf{P}) \leq'_i V(\mathbf{P}')$*.*

Conversely:

- $V(\mathbf{P}) <_i V(\mathbf{P}')$ *implies* $\forall x : V(\mathbf{P}) <_i x \Rightarrow V(\mathbf{P}) \leq'_i x$*,*
- $V(\mathbf{P}) >'_i V(\mathbf{P}')$ *implies* $\forall x : V(\mathbf{P}') <'_i x \Rightarrow V(\mathbf{P}') \leq_i x$*.*

"If the shift from $\geq_i$ to $\geq'_i$ is an improvement, then no knight was demoted", and similarly in the other direction.

Definition 6. *An i-minimal pair of profiles* $\mathbf{P}$, $\mathbf{P}'$ invites decency towards knaves *if the following hold:*

- $\exists x : V(\mathbf{P}) <'_i x <_i V(\mathbf{P})$ *implies* $V(\mathbf{P}) \geq_i V(\mathbf{P}')$*,*
- $\exists x : V(\mathbf{P}') <_i x <'_i V(\mathbf{P}')$ *implies* $V(\mathbf{P}) \leq'_i V(\mathbf{P}')$*.*

Conversely:

- $V(\mathbf{P}) <_i V(\mathbf{P}')$ *implies* $\forall x : V(\mathbf{P}) <'_i x \Rightarrow V(\mathbf{P}) \leq_i x$*,*
- $V(\mathbf{P}) >'_i V(\mathbf{P}')$ *implies* $\forall x : V(\mathbf{P}') <_i x \Rightarrow V(\mathbf{P}') \leq'_i x$*.*

Lemma 3 (Non-Perversion Lemma: the Meaning of Perversion). *A resolute voting rule V is NDPe iff V invites decency towards knights for every i-minimal pair of profiles. A resolute voting rule V is NPPe iff V invites decency towards knaves for every i-minimal pair of profiles.*

Proof. By an easy check on the definitions. □

We can use the notion of decency-inviting pairs to work out what the decent V-value switches are for adjacent transpositions (in the case of three alternatives: crossings in the Saari triangle). It then turns out that the constraints on the crossings change. Here is the new crossing lemma for decent behaviour towards both knights and knaves:

Lemma 4 (Lemma For Decent Crossings). *Let V be NDP and NPP, and let* $\mathbf{P} \sim_i \mathbf{P}'$ *be such that* $>_i$ *and* $>'_i$ *are related by an adjacent transposition that exchanges x and y (with* $x >_i y$*). Then* $V(\mathbf{P}) \neq V(\mathbf{P}')$ *implies:*

- *if* $V(\mathbf{P}) = x$ *then* $V(\mathbf{P}') <_i x$*,*
- *if* $V(\mathbf{P}) = y$ *then* $V(\mathbf{P}') <'_i y$*.*

Proof. Assume V is NDP and NPP, let $\mathbf{P} \sim_i \mathbf{P}'$ be such that $>_i$ and $>'_i$ are related by an adjacent transposition that exchanges x and y (with $x >_i y$). Assume $V(\mathbf{P}) \neq V(\mathbf{P}')$.

Suppose $V(\mathbf{P}) = x$. Then $V(\mathbf{P}') \geq_i x$ and $V(\mathbf{P}) \neq V(\mathbf{P}')$ imply $V(\mathbf{P}) <_i V(\mathbf{P}')$. Moreover, $V(\mathbf{P}) = x <_i y <'_i x$. Contradiction with the given that V is NDP. Therefore $V(\mathbf{P}') <_i x$.

Suppose $V(\mathbf{P}) = y$. Then $V(\mathbf{P}') \geq'_i y$ and $V(\mathbf{P}) \neq V(\mathbf{P}')$ imply $V(\mathbf{P}) <_i V(\mathbf{P}')$. Moreover $V(\mathbf{P}) = y <'_i x <_i y$. Contradiction with the given that V is NPP. Therefore $V(\mathbf{P}') <'_i y$. □

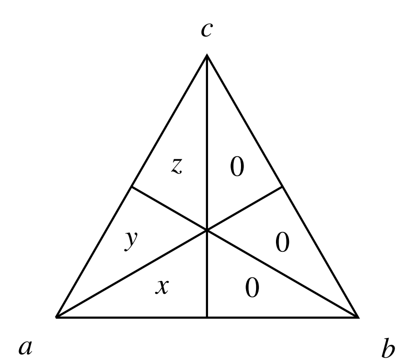
What the new crossing lemma says is that in decent $x : y$ crossing a shift in the vote from x has to be to a position that is above x on the original ballot, and a shift of the vote from y has to be to a position that is above y on the new ballot (in particular, a shift to x is forbidden).

Working this out we find what the decent V-value switches are for walking through the Saari triangle, moving in clockwise direction.

- From 1 to 2: all except (c, b).
- From 2 to 3: all except (a, c).
- From 3 to 4: all except (b, a).
- From 4 to 5: all except (b, c).
- From 5 to 6: all except (a, c).
- From 6 to 1: all except (a, b).

Armed with this, we can turn back to the manipulability proof, to see where it breaks down. It is clear, then, that the Effectiveness Lemma cannot be proved anymore with the new, much weaker version of the Crossing Lemma.

Other theorems also break down. The equivalence of the Pareto condition and non-imposition no longer holds for resolute voting rules that satisfy NPP and NDP. This can be seen by analyzing the picture again:



It remains to show that the notion of pervertibility allows for useful distinctions. For that, notice that these notions can also be applied to voting rules (functions from profiles to non-empty subsets of alternatives), as follows. A voting rule V is single winner demotion pervertibible if there is a pair of profies $\mathbf{P} \sim_i \mathbf{P}'$ with $V(\mathbf{P}) = \{x\}$ and $V(\mathbf{P}') = \{y\}$, and $x <_i y$, and moreover $>'_i$ is the result of demoting at least one knight in $>_i$. Similarly for promotion pervertibility.

The following example shows that the Borda Count rule (the set of alternatives with the highest Borda counts tie for a win, where the Borda count of an alternative a in ballot $>_i$ is given by the number of alternatives strictly below a according to $>_i$) is promotion pervertible in this sense. Let $\mathbf{P} \sim_1 \mathbf{P}_i$ be given by:

$$\begin{array}{|c|c|c|c|} 1 & 2 & 3 & 4 \\ a & b & d & c \\ b & d & c & a \\ c & c & a & b \\ d & a & b & d \end{array} \qquad \sim_1 \qquad \begin{array}{|c|c|c|c|} 1 & 2 & 3 & 4 \\ b & b & d & c \\ a & d & c & a \\ d & c & a & b \\ c & a & b & d \end{array}$$

The BC scores in $\mathbf{P}$ are given by $a : 6$, $b : 6$, $c : 7$, $d : 5$, so the outcome of the Borda vote in $\mathbf{P}$ is $\{c\}$. The BC scores in $\mathbf{P}'$ are given by $a : 5$, $b : 7$, $c : 6$, $d : 6$, so the outcome of the Borda vote in $\mathbf{P}'$ is $\{b\}$. The vote change involves a promotion of the 1-knave d in $\mathbf{P}$; moreover, this promotion is necessary to get b to win. (This example is used in [11] to show that the Borda Count rule is single winner manipulable.)

The Borda Count rule is not single winner demotion pervertible, because of the following fact:

Theorem 6. *The Borda Count Rule has the following property: Although it is* possible *to pervert the Borda Count rule by knight demotion, this is never* profitable. *For any knight demoting manipulation there is a rearrangement alternative that does not demote knights and that works just as well.*

Proof. Let $\mathbf{P} \sim_i \mathbf{P}'$ be an i-minimal pair of profiles with $V(\mathbf{P}) = \{x\}$, and $V(\mathbf{P}') = \{y\}$, with $x <_i y$. To make y the winner in $\mathbf{P}'$, either the Borda count of y must have gone *up*, or the Borda count of x must have gone *down*, or both. Moving y up does not involve knight demotion, and moving x down does not involve knight demotion either. □

6 Conclusion

Don Saari analyzes Arrow's impossibility theorem using geometry, and argues that the principle of IIA (Indepence of Irrelevant Alternatives) is to blame (a very brief summary is in [8]). Saari shows how modifying IIA can turn the impossibility theorem into a useful possibility result.

The same seems possible (and necessary) for the notion of manipulability. Our analysis of proof of the Gibbard/Satterthwaite theorem highlights the severity of the constraints that NM imposes on how the value of the vote can change.

Our notions of *pervertibility* are just examples of possible ways out. Hopefully, the distinction between pervertible and non-pervertible voting rules will turn out more useful than that between manipulable and non-manipulable voting rules. Classifying the pervertible voting rules is future work.

An earlier proposal for modifying the notion of manipulability is in [2]. The criticism in that paper of the notion of manipulability is two-fold: the authors argue that manipulations can be sincere, and they argue that the non-transparancy that results from manipulability can be a boon. Let's ignore the second criticism, and focus on the first.

> We do not think that someone has revealed a preference for beer over champagne when they buy beer rather than champagne, when we know their finances will not stretch to a bottle of the bubbly. [2]

The cited paper calls a manipulation of V in $\mathbf{P}$ given by $\mathbf{P}'$ *sincere* if $\mathbf{P}'$ is the result of a subset S of the voters moving some y that they all prefer to $V(\mathbf{P})$ to the top of their ballots, while leaving the rest of the ballot unchanged. (Actually, the definition is stated in game-theoretic terms; this is my paraphrase.) Clearly, this is a special case of our proposal: y is among the knights of all voters in S, so no knave is promoted and no knight demoted in the switch from $\mathbf{P}$ to $\mathbf{P}'$. But the proposal is less general than ours, for a ballot change from $xyz\dot{w}$ to $x\dot{z}yw$ (with the dots indicating the outcome of the vote) is not sincere in the sense of [2], but it is decent in our sense.

Acknowledgement Thanks to Krzysztof Apt, Vince Conitzer, Ulle Endriss, Floor Sietsma and Sunil Simon for enlightening discussions about the topic of this paper. Three anonymous CLIMA XII reviewers also gave useful feedback.

References

1. K. Arrow. *Social Choice and Individual Values*. Wiley, New York, 1951, second edition: 1963.
2. Keith Dowding and Martin van Hees. In praise of manipulation. *British Journal of Political Science*, 38:1–15, 2008. `http://dx.doi.org/10.1017/S000712340800001X`.
3. D. Eckert and C. Klamler. A geometric approach to paradoxes of majority voting: From Anscombe's paradox to the discursive dilemma with Saari and Nurmi. *Homo Oeconomicus*, 26(3/4):471–488, 2009.

4. Jan van Eijck. *Discourses on Social Software*, chapter On Social Choice Theory, pages 71–85. Amsterdam University Press, 2009. `www.cwi.nl/~jve/books/pdfs/just0SCT.pdf`.
5. A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41:587–601, 1973.
6. E. Muller and M. Satterhwaite. The equivalence of strong positive association and strategy proofness. *Journal of Economic Theory*, 14:412–418, 1977.
7. J. Perote-Peña and A. Piggins. Geometry and impossibility. *Economic Theory*, 20:831–836, 2002.
8. D. Saari. Arrow impossibility theorem. In *Encyclopaedia of Mathematics*, Springer Online Reference Works. Springer, 2001. `http://eom.springer.de/a/a110710.htm`.
9. D.G. Saari. *Basic Geometry of Voting*. Springer, 1995.
10. M.A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.
11. Alan D. Taylor. *Social Choice and the Mathematics of Manipulation*. Mathematical Association of America and Cambridge University Press, 2005.