

A MISSING PROOF OF THE GIBBARD-SATTERTHWAITE THEOREM*

UUGANBAATAR NINJBAT

*Department of Mathematics, The National University of Mongolia
Ulaanbaatar P.O.Box 46-145, Mongolia
uugnaa.ninibat@gmail.com.*

Received April 2017; Accepted November 2017

Abstract

A short and direct proof of the Gibbard-Satterthwaite theorem à la Amartya Sen's proof of Arrow's impossibility theorem is given.

Keywords: the Gibbard-Satterthwaite theorem, field expansion lemma, group contraction lemma
JEL Classification Codes: D70, D71

I. *Introduction*

Several different approaches are used in proving social choice impossibility theorems, such as

- By induction (see e.g. Satterthwaite (1975), Sen (2001), Ninjbat (2012) and Svensson and Reffgen (2014)),
- By contradiction (see e.g. Fishburn (1970) and Suzumura (1988)),
- By the pivotal voter approach (for recent modifications, see Yu (2012), Fey (2014)),
- By other known impossibility theorems (see e.g. Gibbard (1973)), and
- By using computer assistance (see e.g. Takekuma (1997), Tang and Ling (2009)).

Among these, proofs in the style of Arrow (1963) and Sen (1986) have a distinct feature to treat the impossibility result under consideration as a self contained mathematical structure and deduce the result from its setting without referring to an external mathematical device; see Sen's discussions in Maskin and Sen (2014). In Sect. 3, we give a short and direct proof for the Gibbard-Satterthwaite theorem (Gibbard (1973), Satterthwaite (1975)) using this approach.

II. *The Preliminaries*

A denotes the set of alternatives with $|A| \geq 3$ elements, and X denotes the set of strict linear orders (i.e. complete, transitive and asymmetric binary relations) on A . Let there be N

* Financial support from the National University of Mongolia (P2016-1225) is acknowledged. I am also thankful to an anonymous referee for the remarks.

individuals in the group $I=\{1, 2, \dots, N\}$. A function $f:X^N \rightarrow A$ is called as a social choice function. A member $x=(x_1, \dots, x_N)$ of X^N is called as a profile, and its i' th component, $x_{i'}$, is called the individual i' 's ranking. For any $x \in X^N$ and $i \in I$, let $(x'_i, x_{-i}) \in X^N$ denote the profile that has $x'_i \in X$ in its i' th component instead of $x_i \in X$, and otherwise the same as $x \in X^N$. When $a \in A$ is ranked above $b \in A$ according to x_i we write $a \succ_{x_i} b$.

A group of individuals $G \subseteq I$ is *decisive over* $a \in A$ if $f(x) = a$ for all $x \in X^N$ such that a is on the top of x_i for all $i \in G$. $G \subseteq I$ is *decisive* if it is decisive over all $a \in A$. We say that $f:X^N \rightarrow A$ is *unanimous* (UNM) if I is decisive. It is *manipulable* (MNP) at $x \in X^N$ by $i \in I$ via $x'_i \in X$ if $f(x'_i, x_{-i}) \succ_{x_i} f(x)$. It is *strategy proof* (STP) if it is not manipulable. Finally, it is *dictatorial* (DT) if there is a decisive group consisting of a single individual. The following result is known as the Gibbard-Satterthwaite theorem:

Theorem 1. (Gibbard (1973), Satterthwaite (1975)) $f:X^N \rightarrow A$ is UNM and STP if and only if it is DT.

III. The Proof

From now on we assume $f:X^N \rightarrow A$ is UNM and STP, and we shall prove three subsequent lemmas.

Lemma 1. (Tops only) Let $x \in X^N$ and $a, b \in A$ be such that the top ranked alternative in x_i is in $\{a, b\}$ for all $i \in I$. Then, $f(x) \in \{a, b\}$.

Proof. By UNM, we may assume that $a, b \in A$ are distinct. Let $G(a, x) = \{i \in I: x_i \text{ ranks } a \text{ as the top}\}$ and $G(b, x) = \{j \in I: x_j \text{ ranks } b \text{ as the top}\}$. Without losing generality, we may assume that $G(a, x) = \{1, \dots, k\}$ and $G(b, x) = \{k+1, \dots, N\}$ for some $k \leq N$.¹ Consider $x' \in X^N$ such that $x'_i = (a \succ b \succ \dots)$ for all $i \in G(a, x)$, and $x'_j = (b \succ a \succ \dots)$ for all $j \in G(b, x)$. We claim that $f(x') \in \{a, b\}$. To see this, suppose $f(x') \notin \{a, b\}$. Transform $x' \in X^N$ by reversing the positions of $a, b \in A$ in x'_i for all $i \in G(a, x)$, one at a time. Let x^i be the resulting profile after x'_1, \dots, x'_i are changed, and we set $x^0 = x'$. Then, $f(x^i) \notin \{a, b\}$ for all $i \in G(a, x)$, since

- $f(x^0) \notin \{a, b\}$, and
- if $f(x^i) \notin \{a, b\}$, but $f(x^{i+1}) \in \{a, b\}$ for some $i \in \{0, \dots, k-1\}$, then f is MNP at $x^i \in X^N$ by individual $i+1$ via x'_{i+1} .

In particular, $f(x^k) \notin \{a, b\}$, which then contradicts to UNM, and this proves our claim.

Transform $x' \in X^N$ by replacing x'_i with x_i for all $i \in G(a, x)$, one at a time. Let $y^i \in X^N$ be the profile obtained after x'_1, \dots, x'_i are replaced, and we set $y^0 = x'$. Then, $f(y^i) \in \{a, b\}$ for all $i \in G(a, x)$, since

- $f(y^0) \in \{a, b\}$, and
- if $f(y^i) \in \{a, b\}$ then $f(y^{i+1}) \in \{a, b\}$ for all $i \in \{0, \dots, k-1\}$. To see this, suppose for some $i \in \{0, \dots, k-1\}$, $f(y^i) \in \{a, b\}$ but $f(y^{i+1}) \notin \{a, b\}$. Start with $y^{i+1} \in X^N$ and transform everyone's preferences across $\{k+1, \dots, N\}$ by bringing $a \in A$ to the top. Then STP

¹ The reader should check that our argument works if we instead had $G(a, x) = \{i_1, \dots, i_k\}$ and $G(b, x) = \{i_{k+1}, \dots, i_N\}$.

ensures that social choice is never in $\{a,b\}$, which then eventually contradicts to UNM after transforming the individual N 's preferences.

In particular, $f(y^k) \in \{a,b\}$. Now transform $x' \in X^N$ by replacing x'_j with x_j for all $j \in G(b,x)$, one at a time. Let $z^j \in X^N$ be the profile obtained after x'_{k+1}, \dots, x'_j are replaced, and we set $z^k = x'$. Then, $f(z^j) \in \{a,b\}$ for all $j \in G(b,x)$, since

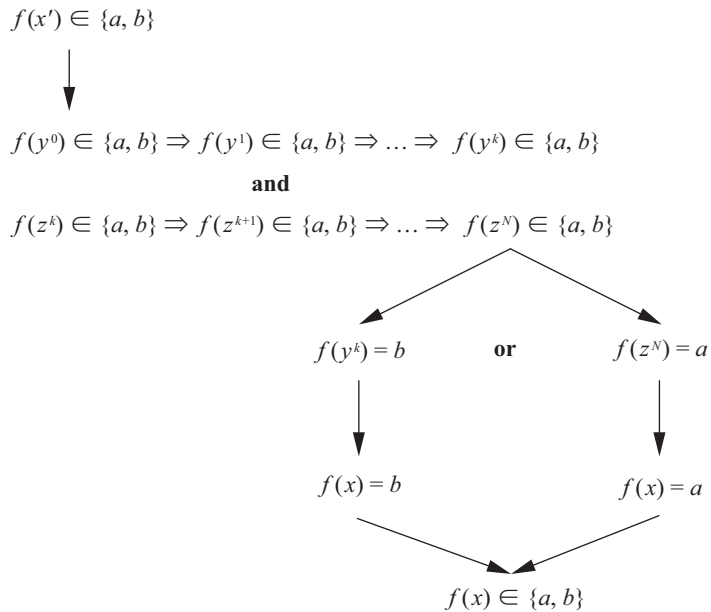
- $f(z^k) \in \{a,b\}$, and
- if $f(z^j) \in \{a,b\}$ then $f(z^{j+1}) \in \{a,b\}$ for all $j \in \{k, \dots, N-1\}$. To see this suppose for some $j \in \{k, \dots, N-1\}$, $f(z^j) \in \{a,b\}$ but $f(z^{j+1}) \notin \{a,b\}$. Start with $z^{j+1} \in X^N$ and transform everyone's preferences across $\{1, \dots, k\}$ by bringing $b \in A$ to the top. Then STP ensures that social choice is never in $\{a,b\}$, which then eventually contradicts to UNM after transforming the individual k 's preferences.

In particular, $f(z^N) \in \{a,b\}$. We then claim that exactly one of the following two statements holds true:

- (a) $f(y^k) = b$, or
- (b) $f(z^N) = a$.

To see this, assume none of them holds true. Then, since both $f(y^k)$ and $f(z^N)$ are in $\{a,b\}$, we have $f(y^k) = a$ and $f(z^N) = b$. Transform $y^k \in X^N$ back to $x' \in X^N$ by reversing the above procedure. Then STP ensures that social choice remains at $a \in A$ throughout this transformation, in particular $f(x') = a$. Similarly, transform $z^N \in X^N$ back to $x' \in X^N$. Again, STP ensures that social choice remains at $b \in A$ throughout this transformation, in particular $f(x') = b$, which is a

FIGURE 1. LOGIC OF THE PROOF OF LEMMA 1



contradiction as we already concluded that $f(x')=a$ and $a,b \in A$ are different alternatives. Thus, at least one of the two statements must be true. However, with a very similar argument one can also show that the two statements in our claim can not be true at the same time. This proves our claim.

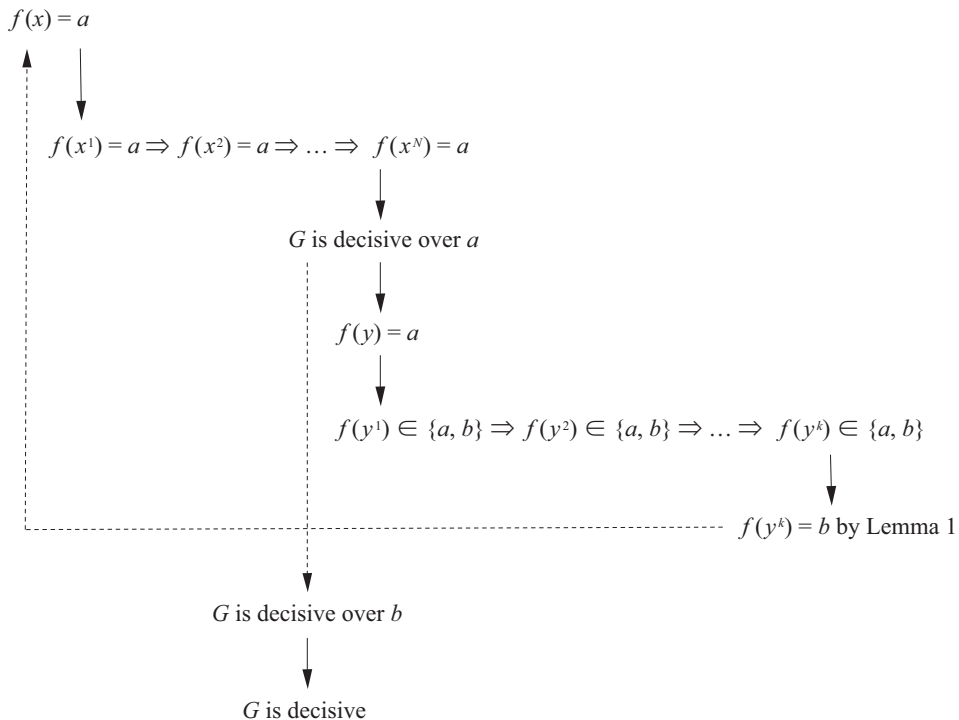
To complete the proof of Lemma 1, assume $f(y^k)=b$ and transform $y^k \in X^N$ into $x \in X^N$ by changing preferences of the individuals in $G(b,x)$ into their preferences in $x \in X^N$, one at a time. Then, STP ensures that social choice remains at $b \in A$ throughout this transformation. In particular, $f(x)=b$. If instead we had $f(z^N)=a$, then we can show that $f(x)=a$ with a similar argument. Thus, in either case, $f(x) \in \{a,b\}$. \square

Lemma 2. (Expansion) *Let $G \subseteq I$ and $x \in X^N$ be such that $a \in A$ is ranked at the top of x_i for all $i \in G$, and at the bottom of x_j for all $j \in I \setminus G$. If $f(x)=a$, then G is decisive.*

Proof. As above, we may assume that $G = \{1, \dots, k\}$ with $k \leq N$. We first show that G is decisive over $a \in A$. Let $x' \in X^N$ be such that $a \in A$ is at the top of x'_i for $i \in G$. Transform $x \in X^N$ into $x' \in X^N$ by replacing x_i with x'_i , for $i=1, 2, \dots, N$, one at a time. Let $x^i \in X^N$ be the profile obtained after changing x_1, \dots, x_i , and we set $x^0 = x$. Notice that $f(x^i) = a$ for $i=1, \dots, k$, since

- $f(x^0) = a$, and
- if $f(x^i) = a$, but $f(x^{i+1}) \neq a$ for some $i \in \{0, \dots, k-1\}$, then f is MNP by individual $i+1$ at x^{i+1} via x_{i+1} .

FIGURE 2. LOGIC OF THE PROOF OF LEMMA 2



Notice also that $f(x^j) = a$ for $j = k+1, \dots, N$, since

- $f(x^k) = a$, and
- if $f(x^j) = a$, but $f(x^{j+1}) \neq a$ for some $j \in \{k, \dots, N-1\}$, then f is MNP by individual $j+1$ at x^j via x_{j+1}^j .

We showed that $f(x') = f(x^N) = a$ and hence, G is decisive over a .

We next show that G is decisive over any $b \in A \setminus \{a\}$. Let $y \in X^N$ be a profile where $y_i = (a > b > \dots)$ for all $i \in G$, and $y_j = (c > \dots > b)$ for all $j \in I \setminus G$, for some $c \in A \setminus \{a, b\}$. Then, $f(y) = a$ as G is decisive over $a \in A$. Change the positions of $a, b \in A$ in y_i for $i = 1, 2, \dots, k$, one at a time. Let $y^i \in X^N$ be the profile obtained after changing y_1, \dots, y_i , and we set $y^0 = y$. Notice that $f(y^i) \in \{a, b\}$ for $i = 1, \dots, k$, since

- $f(y^0) = a \in \{a, b\}$, and
- if $f(y^i) \in \{a, b\}$, but $f(y^{i+1}) \notin \{a, b\}$ for some $i \in \{0, \dots, k-1\}$, then f is MNP by individual $i+1$ at $y^{i+1} \in X^N$ via y_{i+1} .

Then, $f(y^k) \in \{a, b\}$, but by Lemma 1, $f(y^k) \in \{b, c\}$. Thus, $f(y^k) = b$ and we can repeat the argument above to show that G is decisive over b . Thus, G is decisive. \square

Lemma 3. (Contraction) *If a group $G \subseteq I$ with $|G| \geq 2$ is decisive, then it has a proper subset which is decisive.*

Proof. We may again assume that $G = \{1, \dots, k\}$ with $k \leq N$. Let $x \in X^N$ be a profile with $x_1 = (a > \dots > b)$, and for $2 \leq i \leq k$, $x_i = (b > \dots > a)$, and for all $k+1 \leq j \leq N$, $x_j = (a > \dots > b)$. Then, by Lemma 1, $f(x) \in \{a, b\}$. If $f(x) = b$ then we found $\{2, \dots, k\} \subsetneq G$ which is decisive by Lemma 2.

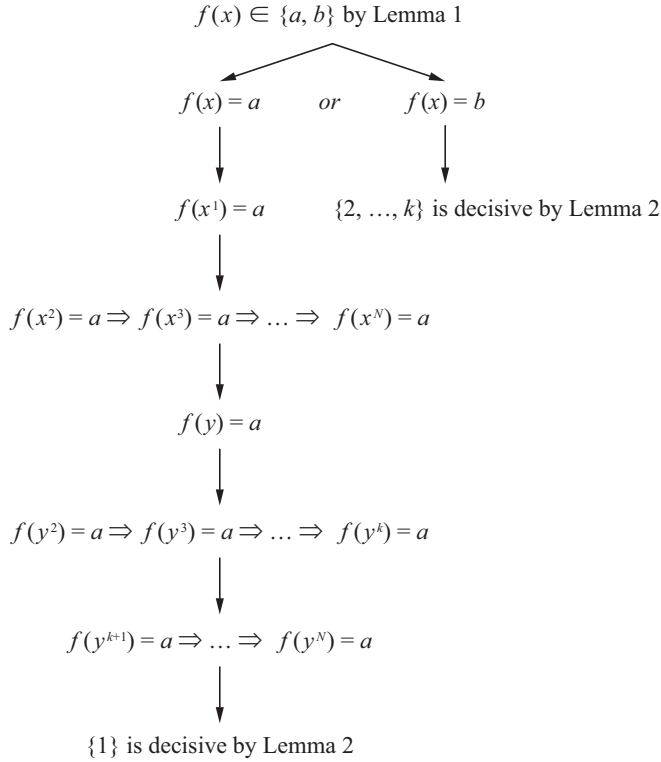
Assume $f(x) = a$ and we show that $\{1\}$ is decisive. Take $c \in A \setminus \{a, b\}$ and let $x^1 \in X^N$ be the profile we obtained from $x \in X^N$ by changing x_1 with $x_1' = (a > b > \dots > c)$. Then, $f(x^1) = a$ as otherwise f is MNP by individual 1 at $x^1 \in X^N$ via x_1 . Let us start with $x^1 \in X^N$ and change its j 'th coordinate as $x_j' = (c > \dots > b)$ for all $j \in \{k+1, \dots, N\}$, one at a time. Let $x^j \in X^N$ be resulting profile after x_{k+1}, \dots, x_j are changed, and we set $x^k = x^1$. Then, $f(x^j) = a$ for $j = k+1, \dots, N$, since

- $f(x^k) = a$,
- $f(x^j) \in \{a, b\}$ for all $j \in I \setminus G$, as otherwise individual 1 can manipulate f at $x^j \in X^N$ by reporting $b \in A$ at the top (recall that G is decisive), and
- if $f(x^j) = a$, but $f(x^{j+1}) \neq a$ for some $j \in \{k, \dots, N-1\}$, then $f(x^{j+1}) = b$, and f is MNP by individual $j+1$ at x^{j+1} via x_{j+1} .

In particular, $f(x^N) = a$. Let $y \in X^N$ be as $y = (x_1, x_2^N)$, i.e. the profile obtained from $x^N \in X^N$ after changing its first coordinate back to x_1 . Then, $f(y) = a$ as otherwise f is MNP at $y \in X^N$ by individual 1 via x_1' . Start with $y \in X^N$ and for $2 \leq i \leq k$ change $y_i = (b > \dots > a)$ as $y_i' = (c > \dots > a)$, one at a time. Let $y^i \in X^N$ be the profile obtained after changing y_2, \dots, y_i , and we set $y^1 = y$. Notice that $f(y^i) = a$ for $i = 2, \dots, k$, since

- $f(y^1) = a$, and
- if $f(y^i) = a$, but $f(y^{i+1}) \neq a$ for some $i \in \{1, \dots, k-1\}$, then f is MNP by individual $i+1$ at $y^i \in X^N$ via y_{i+1}' .

FIGURE 3. LOGIC OF THE PROOF OF LEMMA 3



In particular, $f(y^k) = a$. Finally, start with $y^k \in X^N$ and for $k+1 \leq j \leq N$ change its j 'th coordinate as $y_j' = (c > \dots > a)$, one at a time. Let $y^j \in X^N$ be the resulting profile after y_{k+1}, \dots, y_j are changed. Then, $f(y^j) = a$ for $j = k+1, \dots, N$, since

- $f(y^k) = a$,
- $f(y^j) \in \{a, c\}$ for all $j \in \{k+1, \dots, N\}$ by Lemma 1, and
- if $f(y^j) = a$, but $f(y^{j+1}) \neq a$ for some $j \in \{k, \dots, N-1\}$, then $f(y^{j+1}) = c$, and f is MNP by individual $j+1$ at $y^j \in X^N$ via y_{j+1}' .

Thus, $f(y^N) = a$, and $a \in A$ is ranked as the top by individual 1, and as the bottom by everybody else. Then Lemma 2 implies that $\{1\}$ is decisive. \square

By UNM we know that I is decisive. Then, repeated application of Lemma 3 gives the result in Theorem 1.

IV. *Final Remarks*

Let us make a few comparisons. Lemma 2 and 3 are counterparts of the field expansion and group contraction lemmas in Sen (1986), but with a small difference. Lemma 2 states that if a group is decisive over an alternative at a particular profile, then it is globally decisive, whereas Sen's field expansion lemma is not profile specific. Another difference between our proof and Sen's proof is Lemma 1, which is a non trivial result that the former needs.

The former difference can be attributed to the fact that the key axiom in Theorem 1, strategy proofness, is a local (intra-profile) condition, while as already noted in Fishburn (1987) most of the other key axioms such as monotonicity and Arrow's IIA are more global (inter-profile). The latter difference can be explained by different set ups used in stating impossibility results, i.e. social welfare function vs. social choice function. However, one can remove these differences by

- Removing Lemma 1 and incorporating it in the proofs of the other lemmas wherever it is needed, and
- Replacing Lemma 2 with the following weaker statement, proof of which is already embedded in the proof of Lemma 2: *If a group is decisive over $a \in A$, then it is decisive.*

Such a change will make the two proofs parallel, but each of the resulting lemmas would have a longer proof.

Since the differences are syntactical and can be removed, we believe that the above proof is very close to Sen's proof. Moreover, it is one of the shortest among the existing proofs of Theorem 1, and it being missing is a surprise. One possible cause of this delay is the emergence of other approaches; in particular, the popular pivotal voter approach (see Sect. 1). On the other hand, Moulin (1988), Taylor and Pacelli (2008) and Wallis (2014) apply variants of this approach to prove Theorem 1. But the former two proofs use the Muller-Satterthwaite theorem (see Chap. 10.1 in Moulin (1988), Chap. 7.5 in Taylor and Pacelli (2008)), while the latter proof consists of seven lemmas, one of which is stated without a proof (see Chap. 5.5 in Wallis (2014)). As such, the current proof is more direct than any of these, and could well be 'the missing proof' for an advocate of this approach.

REFERENCES

- Arrow, K.J. (1963), *Social Choice and Individual Values*, 2nd edn., New York, John Wiley and Sons.
- Fey, M. (2014), "A Straightforward Proof of Arrow's Theorem," *Economics Bulletin* 34, pp.1792-1797.
- Fishburn, P.C. (1970), "Arrow's Impossibility Theorem: Concise Proof and Infinite Voters," *Journal of Economic Theory* 2, pp.103-106.
- Fishburn, P.C. (1987), *Interprofile Conditions and Impossibility*, Chur, Harwood Academic Publishers.
- Gibbard, A. (1973), "Manipulation of Voting Schemes: A General Result," *Econometrica* 41, pp.587-601.

- Maskin, E.S. and A.K. Sen (2014), *The Arrow Impossibility Theorem*, New York, Columbia University Press.
- Moulin, H. (1988), *Axioms of Cooperative Decision Making*, Cambridge University Press.
- Ninjabat, U. (2012), "Symmetry vs. Complexity in Proving The Muller-Satterthwaite Theorem," *Economics Bulletin* 32, pp.1434-1441.
- Satterthwaite, M. (1975), "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory* 10, pp.187-217.
- Sen, A.K. (1986), "Social Choice Theory," in K.J. Arrow and M.D. Intriligator eds., *Handbook of Mathematical Economics*, vol. 3, Amsterdam, North Holland, pp.1073-1181.
- Sen, A. (2001), "Another Direct Proof of The Gibbard-Satterthwaite Theorem," *Economics Letters* 70, pp.381-385.
- Suzumura, K. (1988), "Reduction of Social Choice Problems: A Simple Proof of Arrow's General Possibility Theorem," *Hitotsubashi Journal of Economics* 29, pp.219-222.
- Svensson, L-G. and A. Reffgen (2014), "The Proof of The Gibbard-Satterthwaite Theorem Revisited," *Journal of Mathematical Economics* 55, pp.11-14.
- Takekuma, S. (1997), "A Proof of Arrow's Impossibility Theorem by Mathematica," *Hitotsubashi Journal of Economics* 38, pp.139-148.
- Tang, P. and F. Ling (2009), "Computer Aided Proofs of Arrow's and Other Impossibility Theorems," *Artificial Intelligence* 173, pp.1041-1053.
- Taylor, A.D. and A.M. Pacelli (2008), *Mathematics and Politics: Strategy, Voting, Power and Proof*, 2nd edn., Springer.
- Wallis, W.D. (2014), *Mathematics of Elections and Voting*, Springer.
- Yu, N.N. (2012), "A One-shot Proof of Arrow's Impossibility Theorem," *Economic Theory* 50, pp.523-525.