Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON SOME PARADOXES IN VOTING THEORY

KRZYSZTOF CIESIELSKI

Abstract. We present some paradoxes concerning voting theory – both apportionment methods and elections of a winner. All the examples in the paper were constructed by the author. Mathematical background of some voting methods is explained. Also, the fundamental results on voting theory concerning the nonexistence of fair methods with some historical remarks are described. Finally, a new result on a weak method of n votes is presented.

Keywords: voting method, apportionment method, proportional representation, method, method of n votes, Arrow's Theorem.

MSC2010: 91B12, 91B14, 00A05

Received May 14, 2021; accepted 28 July 2021

1. On some paradoxical phenomena in apportionment methods

In many countries members of the parliament are elected with the use of *apportionment methods*. Number of seats given to a party is determined by the number of votes it receives. It is frequently said that thanks to such a method we get a proportional representation of parties or political groups in the parliament. For example, one may suppose that a party receiving 20% of the national vote would receive approximately 20% of the seats in the parliament. Anyway, although those methods are called "proportional", very frequently an obtained result is far from proportional and those methods lead to many paradoxes.

In such a method, usually a country is divided into wards, i.e. electoral subdivisions. In each ward some members of the parliament are elected. In voting, each voter indicates a preference for one party and for one candidate on the list of this party. As the result, numbers of seats that are given for those parties in the ward are determined.

https://doi.org/10.2478/9788366675360-005.

Let us explain in the example the method currently used in several countries, including Poland (the method is known as the *Jefferson method* or the *D'Hondt method*). The example is shown in Table 1.

	votes (:1)	:2	:3	:4	:5	Representatives
Fees	$\star$ 24000	$\star$ 12000	$\star$ 8000	$\star$ 6000	4800	4
Bees	$\star$ 21600	$\star$ 10800	$\star$ 7200	5400	4320	3
Jees	$\star$ 12000	6000	4000			1
Kees	$\star$ 8000	4000				1
Nees	$\star$ 7000	3500				1
Pees	4200					–
Rees	4000					–

Table 1.

The rule is as follows (its mathematical background will be explained in Section 2). We divide numbers of votes by consecutive natural numbers, i.e. by $1, 2, 3, \ldots$ If in the ward there are n seats to be taken, then n greatest quotients give seats to respective parties. In Table 1 we assume that there are 10 representatives to be elected. The quotients that give seats to the parties are indicated by stars.

In this example we already see that the distribution may be far from proportional. In the case of *Bees*:*Jees* we have the votes with the proportion $1.8 : 1$ and the seats with $3 : 1$. If we consider *Fees*:*Jees*, we have $2 : 1$ against $4 : 1$. It is not difficult to construct more such examples. In fact, if a number of seats in a ward is not too large (say: about 10 or smaller), very many natural models lead to some disproportions.

In many countries, the *election threshold* is introduced. This means that to be considered in the distribution of seats, a party must receive a specified minimum percentage of votes (in Poland, it is 5%). What is important, this percentage must be obtained in the whole country, not in a considered ward. Let us come back to the previous example, now with the final result shown in Table 2.

Assume that although in this ward *Bees* obtained a very good score, in the rest of the country the voters did not support *Bees* so strongly and as the result this party did not obtain 5% votes in the whole country. Thus the party is not considered in the distribution of seats. Nevertheless, in the ward 10 seats are to be taken, so three seats that were previously given to *Bees* must be given to other parties. The quotients that give seats are again indicated by stars.

Now, the result is even more strange. We see that *Pees* got a seat and *Bees* not, although the proportion is $5.14 : 1$. Note that *Pees* got a seat in this ward just because of the lack of a suitable number of votes for *Bees* in other wards, perhaps in a region of the country which is very far from this one and where *Pees* do not apply for any seat. We may also see some disproportions if we compare seats for parties which gained seats in this ward.

	votes (:1)	:2	:3	:4	:5	Representatives
Fees	⋆ 24000	⋆ 12000	⋆ 8000	⋆ 6000	⋆ 4800	5
Bees	~~21600~~	~~10800~~	~~7200~~	~~5400~~	~~4320~~	–
Jees	⋆ 12000	⋆ 6000	4000			2
Kees	⋆ 8000	4000				1
Nees	⋆ 7000	3500				1
Pees	⋆ 4200					1
Rees	4000					–

Table 2.

Suppose that in a certain ward there is a perfect candidate and all the voters in this ward vote for this candidate. However, the number of voters in this ward is smaller than 5% of the population in the country, so this candidate will not be elected to the parliament. He (or she) will not get 5% votes in the whole country, although all the voters in this region wants him (or her) as their representative.

We may see other paradoxes. Consider the example presented in Table 3.

In a certain ward, where 5 seats are to be distributed, three parties apply for those seats. *Bees* will get 6000 votes, *Fees* 5700 votes, *Kees* 1950 votes (see upper rows of Table 3). The distribution 3 : 2 : 0 is reasonable. However, assume that just before the day of election a group of 600 voters (possible because anti-*Fees* agitation organized by *Bees*) gave up voting for *Fees*. Then 400 of them changed their opinion and voted for *Bees*, but 200 decided to give their votes to *Kees*.

If the method was logical, there would be only two possibilities. Either nothing would change (as the modification would not be essential enough), or *Fees* would lose a seat (or seats), first for the benefit of *Bees*.

	votes (:1)	:2	:3	:4	Repr.
Bees	⋆ 6000	⋆ 3000	⋆ 2000	1500	3
	⋆ 6400	⋆ 3200	2133	1600	2
Fees	⋆ 5700	⋆ 2850	1900		2
	⋆ 5100	⋆ 2550	1700		2
Kees	1950	975			–
	⋆ 2150	1075			1

Table 3.

However, the result is shown in lower rows of Table 3. Although *Fees* lost votes, the party did not lose any seat. *Bees* got extra votes, moreover – they got more votes than anyone else, but they lost a seat. In this example we see a paradox which shows that this method is far from proportionality and logical rules.

One may presume that when the voting is organized in many wards, then the disproportions would disappear. Indeed, this may happen. However, the story may be completely different, as is shown in the following example.

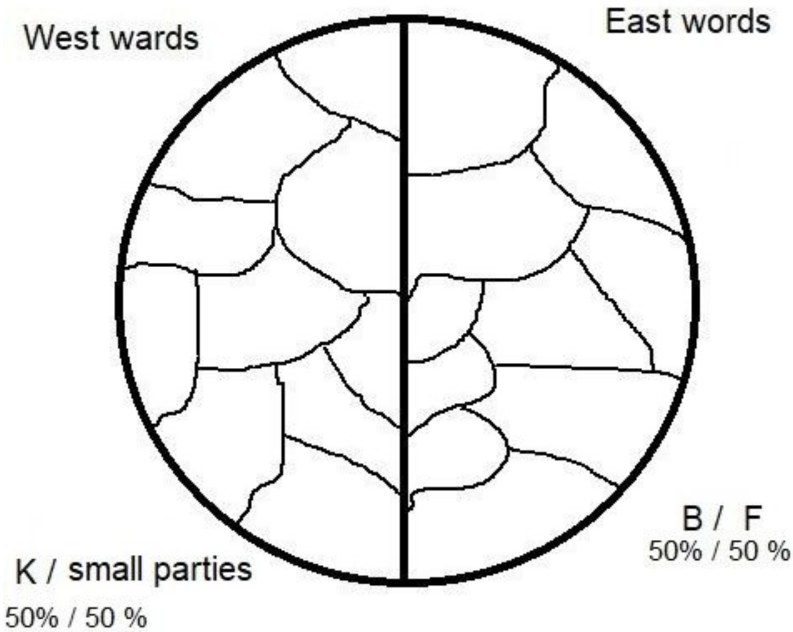

FIGURE 1.

Suppose that in a certain country the opinions of voters substantially differ in the East and in the West. In each East ward 50% voters vote for *Bees* and 50% voters vote for *Fees*. In each West ward 50% voters vote for *Kees* and other voters vote for many small parties which are finally eliminated because of the election threshold (see Figure 1). As the result, *Kees* get 50% seats in the parliament, *Bees* get 25%, and *Fees* get 25%, although all three parties have the same support in the whole country.

This example is artificial, but theoretically (as we consider mathematical models) possible. Nevertheless, generally parties that have representatives in the parliament get (in percent) much smaller support of voters than the number of seats in parliament.

Anyway, as a result of the election we get not anonymous seats, but concrete members of the parliament. Here the rule is also simple. Assume that *Bees* won n seats in a ward. As was mentioned above, each voter indicates one candidate on the list of the chosen party. Now, n candidates with the highest scores from this list are elected.

This also leads to strange phenomena. Consider the following example. Assume that in a ward 6 seats are to be taken and two parties apply for them: Mathematicians and Politicians. The result of voting is presented in Table 4.

	votes (:1)	:2	:3	:4	:5	:6
Mathematicians	$\star$1200	$\star$ 600	$\star$ 400	$\star$ 300	$\star$ 240	200
Politicians	$\star$ 201	101				

TABLE 4.

According to the rule in this method, 5 seats are won by Mathematicians and 1 seat by Politicians. Assume that the votes for the object in Mathematicians' list were as follows: Square 201, Triangle 200, Circle 200, Ball 200, Cube 200, Product 199. In the Politicians' list the greatest score (35) was due to Bureaucrat. Other politicians obtained the following results: one of them 34 and each of the remaining four candidates 33. In this model, Product gained 199 votes against 35 votes for Bureaucrat, but Product will not be elected, although Bureaucrat will become a member of the parliament. Note that Product itself got almost the same number of votes as Politicians altogether.

However, generally such situations rather do not occur in practice. Frequently the majority of votes goes to one candidate on the list (usually the one put in "pole position"). So, assume that Mathematicians gained 100000 votes and because of that have 4 representatives in the parliament. Personally, Triangle got 99995 votes, Tangent Bundle 2 votes, Fibre Bundle 2 votes and Frame Bundle 1 vote. However, in the parliament all the representatives are equal and three Bundles may force the act on the superiority of differential geometry over classical geometry. Not speaking about the case where a member of the parliament changes the party that put him on a list to another one, which happens in some countries. Although in fact three Bundles became members of the parliament thanks to votes for Triangle, being the representatives they may change their party and become members of the party of Physicists.

There are many apportionment methods. Another popular method is the *Webster method* (known also as the *Sainte-Laguë method*). Here the numbers of votes are divided by consecutive odd numbers, i.e. by 1, 3, 5, 7,. . . .

Let's come back to the example concerning Mathematicians and Polititians. If the seats were distributed according to the Sainte-Laguë method, even the paradox worse than previously may occur, see Table 5.

Now Mathematicians' candidates gained the following numbers of votes: Square 316, Triangle 315, Circle 315, Ball 315, Cube 315, Product 314. Product obtained much bigger support than all the politicians together (200), but Product will not become a member of the parliament although one politician will get a seat.

The second paradox will be shown in the example presented in Table 6.

	votes (:1)	:3	:5	:7	:9	:11
Mathematicians	$\star$1890	$\star$ 630	$\star$ 378	$\star$ 270	$\star$ 210	171.8
Politicians	$\star$ 200	66.6				

Table 5.

	votes (:1)	:3	:5	:7	:9	Representatives
Waiters	$\star$ 1050	$\star$ 350	$\star$ 210	$\star$ 150	117	4
Sportsmen	$\star$ 1008	$\star$ 336	$\star$ 202	144	112	3
Waiters	$\star$ 1050	$\star$ 350	$\star$ 210	150	117	3
Footballists	$\star$ 504	$\star$ 168	101	56		2
Basketballists	$\star$ 504	$\star$ 168	101	56		2

Table 6.

In this example, Waiters gained 1050 votes and got 4 seats of 7 (see upper rows of Table 6), whereas Sportsmen gained 1008 votes and got 3 seats of 7. This seems reasonable. However, if instead of Sportsmen's party two ,,sport parties" take part in the election and the votes initially intended for Sportsmen are equally distributed among two parties: Footballists and Basketballists, then the distribution of seats will be as 3 : 2 : 2 (see lower rows of Table 6). Note that Waiters get more than 50% votes, but do not have majority in the parliament! If Sportsmen predicted a possible result (for example on the base of a precise poll) they could artificially split their party into two parties to get majority in the parliament.

Let us move on to a practical surprising application. In 1997 the author of this paper published in a Polish popular monthly *Wiedza i Życie* (*Knowledge and Life*) a popular article [6] explaining the rules of elections. Then in Poland the so-called the *modified Sainte-Laguë method* was in use – here the first seat for the party was obtained for the division the number of votes by 1.4, the rest was the same as in Sainte-Laguë method. This method also shared paradoxes described above. The paradox on the possible splitting of the party to get more votes was described in this article. The article was noticed by a physicist, Jerzy Przystawa (1939–2012), who actively worked to introduce single-winner voting in Poland. Przystawa spread this property and called it *Ciesielski law* (*prawo Ciesielskiego*, see [9]). Then it was applied in real life. In 2002, in the elections of the City Council in Nysa, 23 seats were to be taken. The Mayor of Nysa, Janusz Sanocki (1954–2020) and other members of his group realized that the group would not get majority in the election and, consequently, would not have majority in the City Council. Thus they split their association into two: *Liga Nyska* and *Komitet Obywatelski Ziemi Nyskiej*. The first one won 9 seats in the City Council, the second won 3 seats. Together they had 12 seats of 23 in the City Council, so they had majority. If they had run in this election as one association, they would not have obtained a suitable results. The story is described in [13] as “a practical application of Ciesielski law”.

2. The mathematical background of apportionment methods

The apportionment methods (which nowadays in public opinion are associated mainly with elections) have a source completely different from party lists. In the House of Representatives in the United States the number of seats given to each state is determined by the population of this state. According to the Constitution of the US, the representatives should be apportioned to the states according to their respective numbers of persons. A natural question arises: how to apportion those seats? The problem appeared at the end of the 18th century. The number of states was changing, so rules must have been analyzed and possibly modified. Then, different methods were used and those methods were changing, as new paradoxes were being discovered and noticed. For the description of the history of those changes, see [10].

Let us introduce the mathematical background of these methods. The following description will be based not on elections to the parliament, but on the distribution of seats in the House of Representatives.

First, fix the notation. Assume that there are k states that should have their representatives. Denote the populations in these states by $p_1, \ldots, p_k$ and assume that $p_1 + \ldots + p_k = p$, which is total population in the country. The numbers of seats given to states are $s_1, \ldots, s_k$ and $s_1 + \ldots + s_k = s$, where s is the number of seats in the House of Representatives. Mathematically, in an apportionment method we have to find a function

$$(k, s, (p_1, \ldots, p_k)) \;\mapsto\; (s_1, \ldots, s_k)$$

satisfying the condition $s_1 + \ldots + s_k = s$. Of course, the number of seats for a state must be an integer.

An important concept in the theory is *quota*. A quota of a state i is defined as $q_i = \frac{p_i}{p} \cdot s$. Roughly speaking, quota represents the appropriate number of seats which should be assigned to a state, if the distribution is done fairly. Unfortunately, usually quota is not an integer. We define a standard divisor $d = \frac{p}{s}$ (roughly speaking, this is ,,the value of a seat”). Note that $q_i = \frac{p_i}{d}$, i.e a result of the division of the population of state i by the standard divisor.

We define the lower quota as the floor $\lfloor q_i \rfloor$ of the quota q_i, and the upper quota as the ceiling $\lceil q_i \rceil$ of the quota q_i.

It is natural that each stage should get at least its lower quota seats. When we give to each state its lower quota, it is almost certain then some seats will not be distributed. It seems also natural that one may give remaining seats to the states with the highest differences $q_i - \lfloor q_i \rfloor$. This method, called the *Hamilton method* or the *largest remainder method* has been used several times. In fact, it was the first method suggested to be used in the apportionment in the United States, but then it was vetoed. Nevertheless, later it was used from time to time. This method admits some original paradoxes, in particular the Alabama Paradox and the Oklahoma Paradox. In short, the Alabama Paradox is connected with the situation where the number of all seats is enlarged and as a consequence one state loses a seat. The Oklahoma Paradox deals with the situation where an extra state joins the country, so some seats (according to the proportion in populations), say n, have to be added to the House of Representatives, and then this extra state really gets n seats, but the numbers of seats for some other states change. The precise description of those paradoxes may be easily found in the literature, for example in [10].

Let's move to the *Jefferson method*, proposed at the end of the 18th century by Thomas Jefferson (then a Secretary of State, later the third President of the US) and, independently, about a century later, by Belgian lawyer Victor D'Hondt. Assume that we want to give to each state its lower quota $\lfloor q_i \rfloor = \lfloor \frac{p_i}{d} \rfloor$; recall that $d = \frac{p}{s}$ is a real value of a seat. Then the situation in which not all seats will be taken is an event with probality close to one. Let us modify a standard divisor d and take a number $\widetilde{d}$ (in fact, $\widetilde{d} < d$). This gives us modified quotas $\widetilde{q_i} = \frac{p_i}{\widetilde{d}}$ and consequently modified lower quotas $\lfloor \widetilde{q_i} \rfloor$. We change the divisor up to the moment when the numbers of seats obtained by modified lower quotas sum up to s. Almost always seats obtained by lower quotas with standard divisor d result with a number smaller than s, so an appropriate modified divisor $\widetilde{d}$ is slightly smaller than d. It is proved that the distribution of seats by this method gives the same result as the algorithm of dividing p_i by $1, 2, 3, \ldots$ and taking s greatest quotients.

If in this procedure we will consider upper quotas instead of lower quotas, we get the *Adams method* (named by John Quincy Adams, the sixth President of the US). Then an appropriate modified divisor $\widetilde{d}$ is slightly greater than d. We may round the modified quota $\widetilde{q_i} = \frac{p_i}{\widetilde{d}}$ to the nearest integer (round up if $\widetilde{q_i} = k.5$ for some k) – this gives the *Webster method*. This method was later independently introduced in Europe by André Sainte-Laguë and here the distribution of seats boils down to dividing p_i by consecutive odd numbers. Note that in this method a standard divisor may be appropriate.

In the Webster method rounding to the nearest integer may be presented as a comparison of $\widetilde{q_i}$ to the arithmetic mean $\widetilde{A_i} = \frac{\lfloor \widetilde{q_i} \rfloor + \lceil \widetilde{q_i} \rceil}{2}$: if $\widetilde{q_i} \geq \widetilde{A_i}$, we round it up, and if $\widetilde{q_i} < \widetilde{A_i}$, we round it down. However, we may consider here also other means and they are also taken into account in certain methods. In the case of the geometric mean we have the *Hill–Huntington method* (since 1930s used in the distribution of seats in the House of Representatives in the US) and in the case of the harmonic mean we have the *Dean* method.

A good apportionment method does not exist. Below are stated three natural conditions. If in a certain method of apportionment any of them is not fulfilled, then a method cannot be regarded as fair. We use the notation introduced above. The conditions are:

- *quota condition*: $\lfloor q_i \rfloor \leq s_i \leq \lceil q_i \rceil$ for any state i
- *monotonicity property*: $p_i > p_j \;\Rightarrow\; s_i \geq s_j$ for any states i, j
- *population property*: assume that k and s are given, but populations and seats of states change (we denote it by $a \mapsto \overline{a}$); then there are no i, j with $\overline{p_i} > p_i$, $\overline{s_i} < s_i$ and $\overline{p_j} < p_j$, $\overline{s_j} > s_j$.

In other words, quota condition says that each state should not get less seats than its lower quota and more seats than its upper quota (note that this implies that if quota is an integer, then a state gets its quota). Monotonicity property means that one state is allowed to get more seats than another only if it has a greater (or equal) population. Population property says that it is impossible that the population of one state increases and this state loses a seat, but simultaneously the population of another state decreases and this second state gets an extra seat.

In 1982 Michel Balinski and H. Peyton Young proved in [4] a fundamental theorem.

Theorem 2.1 (Balinski-Young Theorem)**.** *If $k \geq 4, s \geq 7$, then there is no apportionment method satisfying all three conditions: quota condition, monotonicity property, and population property.*

To end this chapter note that although apportionment methods are frequently paired off with the elections, they are applicable to many situations – in fact, frequently in much reasonable sense. We need to use apportionment methods in many situations in real life. Say, there are some shareholders that contributed in some venture and their contributions to this venture are different. As the result of their contributions they should get some goods, let's call them shares. Contributions and shares are expressed in integers and shares are expressed in much smaller integers. It is easy to give several examples from real life. The question is: how to distribute shares fairly?

Note that one share is indistinguishable from another one. In the distribution of seats to the House of Representatives in the United States the seats are given to states and there is not indicated who personally would get a seat. Some paradoxes in the application of apportionment methods applied to voting appear because the concrete people are elected simultaneously with determining of numbers seats; in other words, shares (seats) are distinguishable.

3. Single winner voting methods.

Now we turn to methods that lead to the selection of a winner. Of course, such a selection need not be connected with politics, parliament etc. First fix the terminology.

Assume that two sets are given: $\mathbf{V}$ – a *set of voters* and $\mathbf{C}$ – a *set of candidates*. Each voter ranks all candidates (an order in $\mathbf{C}$ fixed by a particular voter is called a *voter's preference order*). A collection of all voters' preference orders is called a profile. A *winner method* is a method that on the base of a profile outputs a set of winners (a subset of $\mathbf{C}$, possibly empty) – of course, we prefer getting a single winner. Formally, if we denote a set of all profiles by $\mathbf{\Sigma}$, then a winner method is a function $f\colon \mathbf{\Sigma} \to 2^{\mathbf{C}}$. If by $Or(\mathbf{C})$ we denote the set of all (linear) orders in $\mathbf{C}$, then $\mathbf{\Sigma} = \{M\colon \mathbf{V} \to Or(\mathbf{C})\}$.

Now consider the following example. Assume that five candidates: Pooh, Tigger, Rabbit, Kanga and Eeyore apply for a title of Milne's Star and it will be decided by 55 voters. Of course, it is possible that two different voters will rank candidates with different orders (as $5! > 55$). We consider much simpler profile: there are only six possible ordered preferences. They are shown in Table 7 (in the first row, numbers of voters that rank candidates this way is written).

	14	11	10	9	6	5
1.	P	T	K	R	E	P
2.	E	K	R	E	K	R
3.	T	R	E	K	R	E
4.	R	E	T	T	T	T
5.	K	P	P	P	P	K

Table 7.

Now consider five different voting methods.

In so-called *plurality method* the winner is the candidate who is ranked as the first choice of the most voters. Here Pooh wins with 19 votes.

Now take so-called *president method*. Then the winner is the candidate who gets more than 50% first position. If such a candidate does not exist, we take two candidates with the greatest number of first-place votes and compare them. In our case, the winner is Tigger, who in second round conquers Pooh.

Another method is the *Hare method*. We find the candidate (or candidates) who has the fewer first-place votes and eliminate this candidate. Then voters rearrange their lists (removing that candidate and moving up all the candidates that were ranked lower), and vote again. We proceed up to the moment when there is a candidate with a majority of first-place votes. This method was for many years used in the elections of rector in many Polish universities. In our profile, Kanga wins (the first eliminated candidate is Eeyore, then Rabbit, Tigger and Pooh).

We may use points. If there are n candidates, we may assign points for positions of voters' lists: $n-1$ points for the first position, $n-2$ points for the second position, $\ldots$, 0 points for the last position. A sum of points obtained by candidates in all lists in the profiles determine the winner. This is called the *Borda Count* *method.* In the case of the profile presented in Table 7 Eeyore is the winner with 134 points (R:129, K:109, T:102, P:76).

Now we need to find a profile where Rabbit wins. We may use pairwise comparison between candidates. For two candidates, say A and B, let's count how many voters rank A above B and conversely. If the number of voters that rank A above B is greater than the number of voters that rank B above A, then A conquers B. If there is a candidate who conquers anyone else, then that candidate seems to be an obvious winner. In our case, Rabbit is such a candidate (we have: R:P – 36:19, R:T – 30:25, R:M – 35:20, R:K – 28:27).

The disadvantage of this metod is that in several cases it will not determine a winner (there is a famous *Condorcet paradox*: A conquers B, B conquers C, C conquers A). But it is easy to repair this method; a modified method is called the *Copeland method.* If A conquers B in pairwise comparison, A gets one point. If the comparison results in a tie, a method gives half a point to A and half a point to B. Finally, a candidate (candidates in the case of a tie) with the greatest number of points is a winner. This method is used, for example, in the Handball World Championship or the Handball European Championship, in preliminary round (with points doubled by 2, i.e. 2, 1, 0). For many years it was used in football championship, now 2 point for a winner are replaced by 3 points.

Note that the profile was not artificial. Moreover, all methods are quite natural and logical. In fact, all of them have been used somewhere. However, each of them gives another winner.

Anyway, note that in each of above methods the winner had a significant support. This is one of advantages of such a method in comparison to the election of representatives on the base of the list of a party in an apportionment method. Here the case where a winner gets a very small number of votes is impossible.

Also in the case of voting methods that select the winner, there is the theorem on the nonexistence of a good method. Before presenting a fundamental theorem we introduce some notation and formulate some conditions.

Consider more detailed voting method. As a result of voting, we do not pick up only winners, but a weak order in the set of candidates. A weak order is defined as follows. Take an equivalence relation $\sim$ in $\mathbf{C}$ and consider a linear order in the quotient space $\mathbf{C}/\sim$. Then for $x, y \in \mathbf{C}$ we say that $x < y$ if and only if $[x]_\sim < [y]_\sim$. In other words, we rank elements of $\mathbf{C}$ and admit ties. If as a result of voting we consider a weak order in $\mathbf{C}$, we call a method an *order voting method.* Of course, a winner method is a particular case of an order voting method – we have two classes in the set $\mathbf{C}$: winners and losers.

Formally, if we denote by $We(\mathbf{C})$ the set of all weak orders in $\mathbf{C}$, the order voting method is a function $f\colon \mathbf{\Sigma} \to We(\mathbf{C})$; recall that $\mathbf{\Sigma} = \{M\colon \mathbf{V} \to Or(\mathbf{C})\}$.

The notation $A \overset{v,M}{<} B$ means that in profile M voter v prefers B to A. The notation $A \underset{M}{<} B$ means that as a result in profile M candidate B is ranked higher than A.

An order voting method

- is *anonymous* if
$$\forall M \in \mathbf{\Sigma}\ \forall w, v \in \mathbf{V}$$
$$w \text{ and } v \text{ exchange their votes} \Rightarrow \text{result does not change};$$
- satisfies *Pareto rule* if
$$\forall M \in \mathbf{\Sigma}\ \forall A, B \in \mathbf{C}$$
$$(\forall v \in \mathbf{V}:\ A \overset{v,M}{<} B) \ \Rightarrow\ A \underset{M}{<} B;$$
- satisfies *Independence of Irrelevant Alternatives* (IIA) if
$$\forall M, N \in \mathbf{\Sigma} \text{ and } \forall A, B \in \mathbf{C}$$
$$(\forall v \in \mathbf{V}:\ A \overset{v,M}{<} B \Leftrightarrow A \overset{v,N}{<} B) \Rightarrow (A \underset{M}{<} B \Leftrightarrow A \underset{N}{<} B)$$

In other words, a method is anonymous if the votes of all voters are equal. The Pareto rule means that if all voters prefer B to A, then as a result B is ranked above A. The condition IIA means that if voters change their votes but none of them changes the relation between A and B, then the final relation between A and B will not change (that is, the opinion about other candidates should not impact on the final relation of A and B; we may also say that the final preference between two candidates depends only on the individual voter's preferences between those two candidates).

Fundamental Arrow's Impossibility Theorem proved in 1950 by Kenneth Arrow (Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel winner in 1972), see [1] and [2], says:

Theorem 3.1 (Arrow's Impossibility Theorem)**.** *If* $\mathbf{C}$ *contains at least 3 elements and* $\mathbf{V}$ *contains at least 2 elements, then there is no anonymous order voting method satisfying both the Pareto rule and IIA.*

The theorem may be presented in slightly more general way. For this purpose, introduce the next definition. By the *dictatorship* we mean a method where the final results is identical with a vote of one particular voter. Of course, this method is not anonymous. Arrow's Theorem says:

Theorem 3.2 (Arrow's Impossibility Theorem)**.** *If* $\mathbf{C}$ *contains at least 3 elements and* $\mathbf{V}$ *contains at least 2 elements, then the only order voting method satisfying both the Pareto rule and IIA is dictatorship.*

Arrow's Theorem may be formulated in more general mathematical form, without referring to voting. Note that the Pareto rule and IIA may be formulated just for ordered sets. Let us keep the notation introduced above and present a next definition.

A set $\mathbf{T} \subset \mathbf{V}$ is called a *decision set* if

$$\forall M \in \mathbf{\Sigma}\ \forall A, B \in \mathbf{C}\ \ (\forall v \in \mathbf{T}:\ A \overset{v,M}{<} B) \ \Rightarrow\ A \underset{M}{<} B.$$

In the language of voting this means that if all voters in $\mathbf{T}$ prefer B to A, then finally B is ranked above A.

In the following theorem we do not assume that $\mathbf{V}$ and $\mathbf{C}$ are finite. For the basic information of filters, see for example [3].

Now we have

Theorem 3.3 (Arrow's Theorem in form of ultrafilters)**.** *If $\mathbf{C}$ contains at least 3 elements and $\mathbf{V}$ contains at least 2 elements, $f : \mathbf{\Sigma} \to We(\mathbf{C})$, f satisfies both the Pareto rule and IIA, then the family of decision sets is an ultrafilter on $\mathbf{V}$.*

Arrow's Impossibility Theorem for voting theory is an immediate consequence of above theorem, as an ultrafilter on a finite set must be generated by one element, so one voter determines the final result.

Arrow's Theorem in form of ultrafilters was first published by Alan P. Kirman and Dieter Sondermann in [7]. The authors credit this idea to Don Brown and Peter S. Fishburn.

For more information on Arrow's Impossibility Theorem see [8], [10] and [14].

4. Methods of k votes

The topic of the last chapter are methods of k votes. Again, first introduce the terminology.

As previously, assume that each voter ranks candidates (there are n candidates). The method is called a *positional method* and denoted $P(a_1, a_2, \dots, a_n)$ if a candidate obtains a_1 points for each first-place vote, a_2 points for each second-place vote and so on. We assume that $a_1, \dots, a_n$ are integers and $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$.

We have already considered such methods. The plurality method is a positional method $P(1, 0, \dots, 0)$. The Borda Count method is a positional method $P(n-1, n-2, \dots, 1, 0)$.

Now we consider a method

$$P(\underbrace{1, 1, \dots, 1}_{k\ \text{times}}, \underbrace{0, 0, \dots, 0}_{n-k\ \text{times}})$$

and denote it by $P_n(k)$. In other words, each voter votes for k candidates, treating them equally, although the voter ranked these k candidates. If as the result we obtain one winner (or more in the case of ties) we call this method *the method of k votes*. If as result we obtain k winners (or more in the case of a tie in the last winning position) we call this method *the weak method of k votes*. Both methods are in use on several occasions. For example, the weak method of k votes is used in many Polish universities in the election of students' representatives or young academic staff representatives in the Faculty Council. Sometimes election threshold 50% for a candidate is required. Then if l candidates ($l < k$) obtain this threshold, they are elected and the voters vote again according to their lists, but now the list of candidates is reduced to $2(k-l)$ candidates with greatest scores in the first round (not elected yet) and so on.

First take into account a method of k votes. Consider profiles presented in Table 8 and Table 9.

	4	3	6
1.	A	A	C
2.	B	C	B
3.	C	B	A

TABLE 8.

In profile presented in Table 8 there are 13 voters and 3 candidates. A wins under the method of 1 vote (with 7 votes), B wins under the method of 2 votes (with 10 votes). Considering here the method of 3 votes is useless, as all the candidates would have equal results, but we may find the winner under the Borda Count method. Here it is C with 15 points (A obtains 14 points, B obtains 10 points).

	4	2	2	1
1.	D	A	A	A
2.	B	B	C	D
3.	C	C	D	C
4.	A	D	B	B

TABLE 9.

In profile presented in Table 9, A wins under the method of 1 vote (with 5 votes), B wins under the method of 2 votes (with 6 votes), C wins under the method of 3 votes (with 9 votes), D wins under the Borda Count method with 16 points (A: 15, B: 12, C: 11).

This paradoxical effect may be generalized for each number of candidates greater than 2. In 1992 Donald G. Saari proved the following theorem.

Theorem 4.1 (Saari's Theorem)**.** *For any $n \geq 3$ and the set of candidates $\{c_1, c_2, \ldots, c_n\}$ there exist a profile so that c_k wins the election under the method of k votes $P_n(k)$ for $k = 1, 2, \ldots, n-1$ and c_n wins under the Borda Count method.*

The theorem is presented and clearly explained (without proof) in [11]. The original proof was published in [12] and was based on advanced geometrical theory constructed in this purpose. However, the theorem may be proved elementary ([5]).

Now turn to weak method of k votes.

The weak method of two votes was used in the election to Polish Senate in 1990–2010. In all (except two) election wards two senators were elected. Consider the example presented in Table 10.

	30%	30%	20%	20%
1.	Pooh	Pooh	Rabbit	Tigger
2.	Rabbit	Tigger	Tigger	Rabbit
3.	Tigger	Rabbit	Pooh	Pooh

TABLE 10.

If only one candidate was elected in one-vote method, Pooh would be the winner. It is obvious that Pooh is the best candidate, as 60% of voters put Pooh in the first place. However, in the weak method of 2 votes Rabbit and Tigger win (with 70% votes each, and Pooh still keeps his 60% votes).

This paradoxical result may be generalized as follows.

Theorem 4.2. *If the set of candidates* $\mathbf{C}$ *contains at least* $2n+1$ *candidates* ($c_1, \dots, c_{2n+1} \in \mathbf{C}$)*, then for each* $n \geq 1$ *there exists such a profile that in voting with weak method of* n *votes the winners are* $c_1, \dots, c_n$*, and in voting with weak method of* $n+1$ *votes the winners are* $c_{n+1}, \dots, c_{2n+1}$*. Moreover, in both cases each winner is supported by more than 50% voters.*

Proof. Consider the following profile (in Table 11 rankings of first $n+1$ positions are presented). Assume that there are v voters.

	k_1	k_2	...	k_n	k_{n+1}	k_{n+2}	...	k_{2n+1}
1.	c_1	c_2	...	c_n	c_{n+1}	c_{n+2}	...	c_{2n+1}
2.	c_2	c_3	...	c_1	c_{n+2}	c_{n+3}	...	c_{n+1}
3.	c_3	c_4	...	c_2	c_{n+3}	c_{n+4}	...	c_{n+2}
...	...				...			
$n-1.$	c_{n-1}	c_n	...	c_{n-2}	c_{2n-1}	c_{2n}	...	c_{2n-2}
$n.$	c_n	c_1	...	c_{n-1}	c_{2n}	c_{2n+1}	...	c_{2n-1}
$n+1.$	$\star$				c_{2n+1}	c_{n+1}	...	c_{2n}

TABLE 11.

The numbers of voters are:

$$k_1 = k_2 = \ldots = k_n = v\left(\frac{1}{2n} + \varepsilon\right),$$

$$k_{n+1} = k_{n+2} = \ldots = k_{2n+1} = v\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right)$$

In the $(n+1)$th positions in the lists of voters $k_1, \ldots, k_n$ (denoted by $\star$) candidates $c_{n+1}, c_{n+2}, \ldots c_{2n+1}$ are posed, each of them $v\frac{1+2n\varepsilon}{2(n+1)}$ times.

In the position $n+2$ we put each of candidates $c_1, c_2, \ldots c_n$ on $v\left(\frac{1}{2n} - \varepsilon\right)$ lists and each of candidates $c_{n+1}, c_{n+2}, \ldots c_{2n+1}$ on $v\frac{2n\varepsilon+1}{2(n+1)}$ lists, of course each candidate is placed on the list that this candidate was not ranked before. The same procedure is applied to the position $n+3$ and so on, up to the position $2n+1$.

If the number of candidates is greater than $2n+1$, then we put candidate c_{2n+2} in the position $2n+2$ in each list, and so on.

Now we need to show the following properties:

- the construction of the profile is correct
- $v\left(\frac{1}{2n} + \varepsilon\right) > v\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right)$
- $nv\left(\frac{1}{2n} + \varepsilon\right) < nv\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right) + \frac{v}{n+1}$
- $n\left(\frac{1}{2n} + \varepsilon\right) > \frac{1}{2}$
- $n\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right) + \frac{1}{n+1} > \frac{1}{2}$

This can be done be simple calculations. Then, for ε and n satisfying those conditions (it turns out that it is enough to assume that $\varepsilon < \frac{1}{2(2n^2+n)}$) and chosen in such a way that a suitable numbers are integers, we get the desired result. $\square$

References

[1] K.J. Arrow, *A difficulty in the concept of social welfare*, Journal of Political Economy **58** (1950), 328-346.

[2] K.J. Arrow, *Social choice and individual values*, John Wiley & Sons, New York 1951.

[3] N. Bourbaki, *Éléments de mathématique. Topologie générale. Chapitres 1 à 4*, Hermann, Paris 1971.

[4] M.L. Balinski, H.P. Young, *Fair Representation: Meeting the Ideal of One Man, One Vote*, Yale University Press, New Haven and London 1982.

[5] K. Ciesielski, *An elementary proof of Saari's Theorem*, in preparation.

[6] K. Ciesielski, *Paradoksy ordynacji, czyli matematyka wyborcza*, Wiedza i Życie 1997, No. 8, 18–23.

[7] A.P. Kirman, D. Sondermann, *Arrow's Theorem, many agents, and invisible dictators*, Journal of Economic Theory **5** (1972), 267–277.

[8] J. Lützen, *How mathematical impossibility changed welfare economics: a history of Arrow's impossibility theorem*, Historia Math. **46** (2019), 56–87.

[9] J. Przystawa, *Konieczność nieuświadomiona*, Wektory, Kąty Wrocławskie 2006.

[10] E.A. Robinson, D.H.Ullman, *A Mathematical Look at Politics*, CRC Press, New York 2011.

[11] D.G. Saari, *Chaotic elections! A Mathematical Looks at Voting*, AMS, Providence, RI 2001.

[12] D.G. Saari, *Millions of election rankings from a single profile*, Soc. Choice & Welfare **9** (1992), 277–306.
[13] J. Sanocki, *Wojownicy*, Nyskie Towarzystwo Społeczno-Kulturalne, Nysa 2005.
[14] A.D. Taylor, A.M. Pacelli, *Mathematics and Politics. Strategy, Voting, Power and Proof*, Springer, New York 2008.

Mathematics Institute, Jagiellonian University, Poland
`krzysztof.ciesielski@im.uj.edu.pl`