

A Mathematical Analysis of an Election System Proposed by Gottlob Frege

Paul Harrenstein* Marie-Louise Lackner† Martin Lackner‡

We provide a mathematical analysis of an election system proposed by the eminent logician Gottlob Frege (1848–1925). His proposal was written presumably in 1918, was (re)discovered around the turn of the millennium, and published for the first time in the original German in 2000. A remarkable feature of Frege’s proposal is its concern for the representation of minorities and its sensitivity to past election results. Frege’s proposal is based on some highly original and relevant ideas; his core idea is that the votes of unelected candidates are carried over to the next election. All candidates thus accumulate votes over time and eventually each candidate is elected at some point. We provide a mathematical formulation of Frege’s election system and investigate how well it achieves its aim of a fair representation of all political opinions in a community over time. We can prove that this goal is fulfilled remarkably well. However, we also show that, in other aspects, it falls short of Frege’s high ambition that no voter’s vote be lost. We propose a slight modification of his voting rule, the modified Frege method, that remedies these shortcomings. We analyse both methods from the perspective of modern social choice and apportionment theory, and can show that they are novel contributions with noteworthy proportionality properties over time.

1. Introduction

In the summer of 1998 a surprising discovery was made in the Thüringer Universitäts- und Landesbibliothek (ThULB) in Jena. Hidden among the legacy of the German politician Clemens von Delbrück (1856–1921), Uwe Dathe, curator at ThULB, found a typescript titled *Vorschläge für ein Wahlgesetz*, which translates to ‘Proposals for a Voting Law’.

*Department of Computer Science, University of Oxford, UK

†Christian Doppler Laboratory for Artificial Intelligence and Optimization for Planning and Scheduling, TU Wien, Vienna, Austria

‡Institute of Logic and Computation, TU Wien, Vienna, Austria

The author of the typescript turned out to be no one less than Gottlob Frege (1848–1925), the illustrious logician and recipient of the letter in which Bertrand Russell expounds his famous paradox. Although the typescript is undated, circumstantial evidence points to 1918 as the almost certain year of composition. The manuscript was finally published in the original German in the year 2000 (Gabriel and Dathe, 2000), accompanied by an extensive and excellent introduction by Dathe and Kienzler (2000). The original typescript has been made digitally available by the Thüringer Universitäts- und Landesbibliothek in Jena (Frege, 1918).

The discovery of Frege’s proposal was surprising and remarkable. Regarding it as a mere historical curiosity, however, would not do justice to its originality and perspective. Rather, we feel that some of its underlying ideas shed a fresh light on modern discussions on how to elect representatives to political assemblies. In particular, Frege takes a highly original temporal point of view, where the votes cast for unelected candidates in an election are carried over to the next election. He furthermore proposes a rudimentary system of how votes can be delegated from candidate to candidate. In this paper, we conduct a mathematical analysis of Frege’s proposal and can prove that, even from the standpoint of modern social choice and apportionment theory, it fares remarkably well. In the following, we present Frege’s proposal in detail and outline our findings.

1.1. Structure and content of Frege’s proposal

Frege’s typescript consists of four main parts and, in its original form, is presented on 24 numbered printed pages and is appended by two fold-out tables, *Tafel I* and *Tafel II*. In the following, the page numbering refers to Frege’s proposal as published in the year 2000 (Frege, 2000), whereas the bracketed numbers refer to the sheets of the original typescript from 1918 (Frege, 1918).

In a *preliminary note* (*Vorbemerkung*) (page 297, [1]–[3]), Frege sketches the general constitutional provisions of his proposal. He presumes a division of the electorate in constituencies, with the voters in each constituency delegating an elected representative to an electoral body like the German *Reichstag* (197, [1]). This is very much in line with constitutional law in Imperial Germany (1871–1918), where the representatives of the constituencies were elected by absolute majority and where, in case none of the candidates succeeded in securing an absolute majority, a run-off took place between the two candidates with the highest number of votes.¹ In this part, he moreover specifies restrictions on active and passive suffrage as he considered them appropriate.

The most original aspects of Frege’s election law are to be found in the second part (pages 297–299, [2]–[4]) of the manuscript. In eleven articles numbered §5 through §15, Frege presents the *voting method* by means of which the representative of a constituency is chosen.

Finally, Frege makes a short concluding remark (*Schlussbemerkung*, page 299, [4]–[5]) concerning the use of referenda to settle political disputes, which is followed by an

¹The succeeding Weimar Republic (1918–1933) adopted a voting system based on proportional representation, where multiple representatives were elected in each of a smaller number of constituencies of a considerably larger size than was usual in Imperial Germany.

extensive discussion elucidating the various aspects of his voting law (*Erläuterungen*, pages 299–311, [5]–[24]). The two tables appended by Frege illustrate his voting method by means of two examples he constructed for the purpose (pages 312–313).

1.2. Political views underlying Frege’s proposal

Frege’s proposal for a voting law displays a remarkable juxtaposition of highly conservative and nationalistic views with classical liberal ideals. Frege’s conservative views are most clearly manifested in his proposal to restrict the right to vote to married men without criminal convictions (“*unbescholten*”), who have performed military service, and who received no state support “*Almosen*” in the previous year (197, [1]–[2]). He explicitly declined women’s suffrage on grounds that the husband be the head of the family, the alleged political unit of German society as he saw it (311, [23]–[24]). To put this into perspective, it should be noted that women’s suffrage was made law on 30 November of the same year that Frege presumably wrote his proposal, with the first German women casting their votes for the *Wahl zur Deutschen Nationalversammlung* on 19 January 1919.

The deep conservatism inherent in these considerations² is in quite some contrast with the more liberal ideals underlying his voting method. According to Frege himself, the voting method constitutes the fundamental thought of his voting law and can largely be considered independently of his ideas about suffrage (197, [1]). In this paper, we will concentrate on the mathematical aspects of Frege’s voting method, and leave the other parts of his proposal largely uncommented.

Frege’s voting method is based on the classical liberal notion that the *voters* in a constituency should be represented by an elected representative in a national assembly or parliament while upholding the ideal of one-man-one-vote. Frege’s main concern is to guarantee that *no voters’ votes are lost in the election process* (197, [1]; 308–9, [19]–[20]) by ensuring that even political minorities should send the constituency’s representative from their midst at some point in time. Failing to do so is a problem that is notoriously inherent in electoral systems where representatives are elected in single-member constituencies by majoritarian or *first-past-the-post* methods, that is, systems where the voters indicate a single candidate on their ballot, and the candidate occurring on most ballots is elected as representative of the constituency. Arguably, any vote cast on a minority candidate serves the voter, and the candidate, just as well as had the voter not participated in the election, and can consequently be considered lost (*Erläuterungen*, 303, [10]).

1.3. Frege’s core idea: Elections as process over time

Frege’s solution to the problem of votes for non-winning candidates being lost during elections is to acknowledge the fact that elections take place repeatedly over time. He sees the election of representatives not so much as a one-shot event, but rather as a series

²Later in his life, Frege’s conservatism and nationalism would give way to even more extreme views on the political far right, as witnessed by his diary entries from 1924. The latter were published in 1994 (Gabriel and Kienzler, 1994; Mendelsohn, 1996) and have tainted Frege’s reputation ever since.

of connected and interdependent events proceeding in rounds. In his proposal, elections are held every five years (§6, 298, [2]), where voters submit a (plurality) ballot, indicating a single candidate, just as it was law in Germany between 1871 and 1918.

He proposes, however, that the representative of a constituency should not be elected on the basis of the votes received in the current election alone, *but also on those cast in previous elections*. More precisely, every candidate has a *voting score* (“*Stimmenzahl*”), which, after initially being set to zero (§5, §7, 297–8, [2]), is increased by the number of votes received in each election. Frege makes a crucial proviso for the incumbent representative of the constituency: on the Friday before the election, the incumbent’s voting score is decreased by the integer part of the *average voting score* of all candidates (§14, 299, [4]). The votes thus subtracted, Frege argues, have served their purpose and cannot be considered lost (*Vorbemerkung*, 197, [2]). We suggest that they can be viewed as the *cost of winning* that the elected representative incurs. The candidate with the highest voting score after the election is then elected as representative for the next five years. Possible ties are broken on the basis of age and, in case of an equal number of days lived, by lot (§13, 299, [4]).

With Frege’s voting method, candidates that only attract minor support among the electorate keep accumulating votes over time, and at some point will be elected as representative. Thus, also minority opinions are guaranteed to be represented in elected political assemblies, which complies with Frege’s guiding principle that no votes should be lost. Should his voting law be adopted, Frege anticipates the lively participation of all voters in the elections (*Erläuterungen*, 303–4, [11]), a concern that has also attracted attention in social choice theory (Fishburn and Brams, 1983; Moulin, 1988).

To the same end, Frege arranges for the installation of a maximum of twenty-five *choice candidates* (“*Erlesene*”) as those candidates in a constituency with maximal voting score (§8, 298, [2]). The representative is chosen among these choice candidates and non-choice candidates need to transfer their votes to one of the choice candidates. This is to ensure that the votes cast on non-choice candidates are not lost or too much scattered to be of any effect. Frege claims that the number of twenty-five choice candidates should suffice for all non-choice candidates to find a politically like-minded choice candidate to transfer their votes to (*Erläuterungen*, 303, [10]). Frege also provisions for the transfer of votes from a choice candidate to a deputy should the former die or otherwise lose his status as a choice candidate (§11, 298, [3]). In what follows, however, we will not make this distinction between choice and non-choice candidates.

Given the objectives of Frege’s voting method, a key issue is how often a candidate can be expected to be elected as representative given his support in the constituency. Even though Frege does not seem to pursue *proportional representation* (“*Proportionswahl*”) as a political aim in itself

he states without formal proof that: “If the strengths of political directions in a constituency remain the same over a prolonged period of time, then the number of times during which each of these directions is represented by the representative of the constituency will behave approximately proportionally to these strengths” (Frege, 2000, 302, [8]).³ The

³The original German reads: “Wenn die Richtungen in einem Wahlkreise längere Zeit hindurch dieselben

main goal of our paper is to investigate in a mathematically rigorous manner whether Frege’s claim can be vindicated.

1.4. Contributions of this paper

Frege’s intended audience presumably being politicians rather than academics, Frege neither provided a formal definition nor a formal analysis of his voting method.

We aim to fill this gap and, in Section 2, start by formulating his proposal in the precise mathematical framework of modern social choice theory. We first observe that Frege’s method fails to be *proportional* in the strict sense even under the assumption that the voters’ number and preferences remain constant over time: Examples are easily found in which at some point a candidate is elected more often (and another less often) than would be justified by the number of votes he received as a proportion of the total of votes cast (Example 2). We show, however, that this failure of proportionality can be attributed to the particularity of Frege’s method that the cost of being chosen representative varies over time. Even more, the variation of this cost creates a massive, unproportional advantage for strong candidates who pay a lower cost than minority candidates.

In Section 3, we can nevertheless prove that the cost of winning will converge and stabilise at the number of voters after a finite number of elections—provided that the size of the constituency remains constant (Lemma 1). Frege appears to have been aware of this phenomenon, as he now and then speaks of a stable state (“*Dauerzustand*”) in this context (302, [8]). This convergence of the cost of winning can take a very long time, as is also indicated by the two examples Frege himself constructed to illustrate his procedure (*Erläuterungen*, 300–1, [5]–[7]; *Tafel I* and *Tafel II*). In both of the examples, the cost of winning can be shown to stabilise only after 184 elections. As Frege proposes that elections are held every five years, this process will take 920 years.

As soon as the cost of winning has stabilised, however, the behaviour of Frege’s voting method becomes more favorable.

We can show that, as time proceeds, the proportion of times a candidate is chosen will converge towards the proportion of the candidate’s support in the electorate. In other words, Frege’s method achieves proportionality in the limit (Theorem 1). Again, this result requires that the size of the electorate remains constant and that the average number of votes each candidate receives converges. In this context, it is interesting that Frege explicitly expresses the desirability of constituencies being of about equal size and their composition changing as little as possible (*Vorbemerkung*, 197, [1]). If instead the size of the electorate is not assumed to be constant but can grow over time, we give an example showing that this convergence to proportionality does no longer hold (Example 3).

The problems that come with the long initialisation phase before the cost of winning stabilises suggest a modified version of Frege’s voting method which we present in Section 4. For this modified version, the number of votes cast for each candidate is normalised to lie between 0 and 1, and the cost of winning is invariably 1. The latter

Stärken behalten, werden sich die Zeiten, während deren die einzelnen durch den Abgeordneten des Wahlkreises vertreten werden, annähernd wie diese Stärken verhalten”. Translation by the authors.

stipulation intuitively corresponds to the cost of winning being equal to the number of voters. The modified version of Frege’s voting method is well-behaved immediately and has stronger proportionality guarantees over time than Frege’s original method. In particular, we prove that, at any point in time, the number of times each candidate has been chosen lies within a bounded margin from his or her proportional share of votes aggregated up to that time (Theorems 3 and 4).

Frege’s original method was conceived for the election of a single representative in a constituency and rests on the liberal principle of voting for individual candidates (“*Persönlichkeitswahl*”) instead of a party-list system (“*Listenwahl*”).

That is, the method should guarantee fair representation of citizens’ opinions in parliament rather than reflect the strength of political parties.

Accordingly, it would be inappropriate to represent Frege as making a case for proportional representation as such. Nevertheless, the modified Frege method, as proposed in this paper, can naturally be interpreted as an *apportionment method* as is commonly used for assigning seats to parties in a political assembly in systems of proportional representation.

This apportionment method, which we introduce in Section 5 and refer to as the *Frege’s apportionment method*, assigns to each political party as many seats as the number of times it would be elected as representative if the modified Frege method were run for so many times as there are seats in the assembly (parliament) while keeping the electorate fixed.

We can show that Frege’s apportionment method is not mathematically equivalent to any of the methods that are common in the literature. In particular, we demonstrate that it differs from the Adams method, the D’Hondt (or Jefferson) method, the quota method, the Sainte-Laguë (or Webster) method, the largest remainder method, and the Huntington-Hill method. Analysing its compliance with the customary axioms for apportionment, we prove that Frege’s apportionment method satisfies house monotonicity and upper quota but fails population monotonicity. Lower quota is only satisfied if the number of candidates is at most three (Theorem 5). As an apportionment method, it therefore behaves surprisingly well, especially given the fact that it was not designed as one. Only the quota method by Balinski and Young (1975) satisfies all of the axioms mentioned above that are also satisfied by Frege’s method. The quota method, however, is notoriously biased against small parties. Frege’s method fares considerably better in this respect, as suggested by a numerical experiment in which we compare the number of votes per representative of the largest party and the number of votes per representative of the smallest party (Section 5.3).

The conclusion seems to be fair that Frege’s apportionment method is an interesting and novel addition to the apportionment literature.

Further discussions of Frege’s proposals follow at the end of the paper in Section 6. So as not to interrupt the flow of the argument, mathematical proofs are deferred to the appendix, Sections A and B.

2. Mathematical Formulation of Frege's Voting Method

Frege couched his voting law in legal terms, and also his subsequent discussions are mathematically informal.

In this section we provide a mathematical formulation of Frege's voting law, where we concentrate on the voting mechanism Frege proposed. In particular we focus on the way candidates accumulate votes over the course of multiple elections and whether this leads to a fair (proportional) representation of opinions over time. We will make the simplifying assumption that candidates remain the same over time, but we do allow voters to change their number, their identity, and opinion as to their most preferred candidate, unless stated otherwise. In particular, we will disregard Frege's distinction between choice candidates ("*Erlesene*") and non-choice candidates, and the delegation mechanism it enables.

Thus, in our analysis, every candidate accumulates votes.

Let C be a set of $m \geq 2$ candidates. Voting proceeds in rounds over time, meaning that at every point in time $t \geq 1$ an election takes place and a new *representative* $\text{repr}(t)$ is chosen. We assume the set of candidates to remain the same over time. We make no such assumption for the voters and we have n_t denote the *number of voters* participating in the election at time t .

In every election, the voters submit their preferences by means of *plurality ballots*, that is, for every $t \geq 1$, the voters specify their most preferred candidate only. We denote by π_j^t the *plurality score* of candidate j at time t , that is, the number of voters that put j on their ballot at time t . Since every and only participating voters cast votes on candidates, we thus have $n_t = \sum_{j \in C} \pi_j^t$.

We speak of a *fixed electorate* if the number of voters and the candidates' plurality scores remain constant over time, that is, if $n_t = n$ and $\pi_j^t = \pi_j$ for all $t \geq 1$ and $j \in C$.

In each round t of the election process, an *aggregate score* σ_j^t is calculated for every candidate j on the basis of the scores obtained at the time of the election (π_j^t) and the scores obtained in past elections. The candidate which obtains the maximal aggregate score at time t is chosen as representative for round t , that is, $\text{repr}(t) = \text{argmax}_{j \in C} \sigma_j^t$. In case of a tie, the candidate that is lexicographically first is chosen. This is equivalent to assuming a fixed tie-breaking order, for instance, by breaking ties in favour of the oldest candidate, as suggested by Frege himself. Formally, we define the *aggregate score* σ_j^t of a candidate j at time t inductively such that, for every $t \geq 1$,

$$\sigma_j^1 = \pi_j^1$$

$$\sigma_j^{t+1} = \begin{cases} \sigma_j^t + \pi_j^{t+1} - \lfloor \frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t \rfloor & \text{if } \text{repr}(t) = j, \\ \sigma_j^t + \pi_j^{t+1} & \text{otherwise.} \end{cases}$$

The term $\lfloor \frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t \rfloor$ can be seen as the *cost of winning* the election at time t , as it is later subtracted from the aggregate score of the winning candidate. Note that this number is chosen in such a way that the aggregate scores of all candidates are guaranteed to remain non-negative at all times. Furthermore, observe that the aggregate scores at

time t are used to elect the representative at time t and consequently only include the costs of winning of previous rounds and not of time t .

Example 1. Consider a fixed electorate with three candidates and ten voters, and let the corresponding plurality scores of candidates a , b , and c be 5, 3, and 2, respectively. The table below depicts the values of σ_j^t for $t = 1, \dots, 10$. Maximum aggregate scores of each round are printed in bold.

time t	σ_a^t	σ_b^t	σ_c^t	$\text{repr}(t)$	$\lfloor \frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t \rfloor$
1	5	3	2	a	3
2	7	6	4	a	5
3	7	9	6	b	7
4	12	5	8	a	8
5	9	8	10	c	9
6	14	11	3	a	9
7	10	14	5	b	9

8	15	8	7	a	10
9	10	11	9	b	10
10	15	4	11	a	10

At time 1, candidate a is chosen, because a has a maximum aggregate score of 5. At time 2, each candidate keeps the votes he had obtained at time 1 plus the votes obtained at time 2, which we assumed to be the same as at time 1. Thus for candidates b and c the aggregate scores at time 2 are $3 + 3 = 6$ and $2 + 2 = 4$, respectively. The cost of winning incurred by candidate a at time 1 amounts to 3, which has to be subtracted from the number of votes candidate a received at time 1 and time 2. Accordingly, candidate a 's aggregate score σ_a^2 at time 2 is calculated as $5 + 5 - 3 = 7$. Hence, at time 2, candidate a again has the highest aggregate score and is elected representative another time. The cost of winning at time 2, however, has increased to 5, and at time 3, it is candidate b who has the maximum aggregate score and is elected representative, this time at a cost of 7. And so on. Note that the cost of winning increases over time, starting with 3 and increasing to 10, the number of voters. Once the cost of winning has reached 10—indicated in the table by the dashed line—it stabilises and remains constant at all subsequent times.

The increasing cost $\lfloor \frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t \rfloor$ of winning, as we saw in Example 1, suggests an unfairness inherent in Frege's voting method. Early winners, that is, those candidates with the highest plurality scores, incur lower costs for being elected than those candidates that win later. This makes it advantageous to win early in the election process, creating a positive bias towards strong candidates and accordingly constitutes a disadvantage for minority candidates. This is also reflected in these candidates being elected more often than would seem to be justified by the proportion of the electorate that supports them.

This phenomenon is all the more remarkable as it was Frege's intention to also strengthen minority opinions (Frege 2000, *Erläuterungen*, 306, [15]; Gabriel and Dathe 2000, page 292). How extreme this distortion can be is illustrated by the following example.

Example 2. Let us consider a scenario that highlights the unfairness introduced by increasing costs. We have a fixed electorate with 6 candidates and 10 voters. The corresponding plurality scores π_j of the candidates a through f are 1, 1, 1, 1, 1, and 5, respectively. The table below depicts the values of σ_j^t for $t = 1, \dots, 10$. Ties occur in rows with more than one element printed in bold. We assume that ties are broken in alphabetical order and thus always to the disadvantage of f .

time	a	b	c	d	e	f	representative	$\lfloor \frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t \rfloor$
1	1	1	1	1	1	5	f	1
2	2	2	2	2	2	9	f	3
3	3	3	3	3	3	11	f	4
4	4	4	4	4	4	12	f	5
5	5	5	5	5	5	12	f	6
6	6	6	6	6	6	11	f	6
7	7	7	7	7	7	10	f	7
8	8	8	8	8	8	8	a	8
9	1	9	9	9	9	13	f	8
10	2	10	10	10	10	10	b	8

After ten rounds, candidate f has been chosen eight times, candidates a and b have been chosen once, and candidates c , d and e not at all. This shows that Frege’s voting method does not select representatives in a proportional fashion: in ten rounds it would be possible to perfectly reflect the distribution of votes, that is, by choosing candidate f five times and all other candidates once.

At this point, we would like to make a minor remark concerning Frege’s use of the floor function in the definition of the cost of winning as $\lfloor \frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t \rfloor$ in his voting method (cf. σ_j^{t+1} , as defined above).⁴ This was most likely motivated by considerations of numerical simplicity, but, for our purposes, does not have a significant mathematical effect. Replacing the term $\lfloor \frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t \rfloor$ with $\frac{1}{m} \cdot \sum_{k \in C} \sigma_k^t$ does not help to resolve any of the issues identified in Example 2.

In Example 1, we also saw that after time 7, the cost of winning stabilises at 10. After this time, therefore, candidates no longer gain an advantage by being elected earlier rather than later, at least not with respect to the cost of winning. This phenomenon is not specific to Example 1; also in Example 2 the cost of winning will stabilise at time 16. Rather, as long as the *size* of the electorate remains constant, the convergence of the cost of winning to the number of voters will hold generally (Lemma 1, below). Under further conditions on how the opinions in the electorate evolve over time, moreover, the unfairness caused by the varying costs of winning in the initial phase will taper off and result in a proportional representation in the long run. We formally prove this in the next section.

⁴Frege writes: “A remainder that is smaller than the number of choice candidates of a constituency, is left disregarded” (“*Ein Rest, der kleiner als die Anzahl der Erlesenen des Wahlkreises ist, bleibt dabei unberücksichtigt*”) (§14, 299, [4], English translation by the authors).

The unfairness towards minority candidates caused by the increasing costs of winning suggests a variation of Frege’s method where the cost of winning is stipulated to be constant from the outset. We introduce the *modified Frege method* in Section 4 and show that it not only guarantees proportionality in the long run, but also has stronger proportionality properties, which we will make formally precise.

3. Proportionality Guarantees for Frege’s Method

In the previous section we saw how the cost of winning converges as long as the number of voters is fixed. We now make this observation mathematically precise. The proof of this lemma as well as all further proofs can be found in the appendix.

Lemma 1. *Assume that the number of voters is fixed, i.e., $n_t = n$ for all $t \geq 1$. Then the function $a(t) = \sum_{k \in C} \sigma_k^t$ is monotonically increasing. Moreover, there exists a positive number t_0 such that $a(t) = n \cdot m$ for all $t \geq t_0$.*

This convergence process of the cost of winning (which is $\lfloor a(t)/m \rfloor$) to the number of voters can take quite a long time: In Frege’s proposal, he provided two explanatory examples with $n = 1000$ and $m = 25$. With these parameters, Frege’s voting method reaches a constant cost of winning at time $t_0 = 184$. As Frege proposes that elections are held every five years, this would amount to 920 years.

Let us continue by specifying in which sense Frege’s voting method violates even a most basic form of proportionality, as noted in Example 2. In what follows, let $\rho_j(t)$ denote the number of times candidate j is chosen as representative up until time t , that is, $\rho_j(t) = |\{s \leq t : repr(s) = j\}|$. In t rounds, each candidate j should ideally win $t \cdot \frac{\pi_j}{n}$ times, that is, the number of rounds multiplied by the proportion of the electorate that supports candidate j . If $t \cdot \frac{\pi_j}{n}$ is an integer for all j , a perfectly proportional outcome is possible. This observation gives rise to the following definition.

Definition 1. A fixed electorate with plurality scores $(\pi_j)_{j \in C}$ has *integral quotas at time t* if $t \cdot \frac{\pi_j}{n}$ is integral for all candidates j . We say that an (infinite) sequence of chosen representatives $(repr(1), repr(2), \dots)$ satisfies *variable integral quota* if for any time $t \geq 1$ at which the electorate has integral quotas, it holds that

$$\rho_j(t) = \tau \cdot \frac{\pi_j}{n} \quad \text{for every } j \in C.$$

Example 2 shows that Frege’s method cannot guarantee this property: a sequence of representatives chosen by Frege’s method may violate variable integral quota. Note that variable integral quota applies to rather few electorates, namely only fixed electorates considered at time points when a perfectly proportional outcomes is possible. Hence, we consider variable integral quota as a weak and very basic form of proportionality.

However, the following theorem shows under which conditions $\rho_j(t)/t$ will converge to π_j/n , thus Frege’s method provides a form of proportionality *in the long run*.

Theorem 1. *If we assume a fixed electorate, i.e., $n_t = n$ and $\pi_j^t = \pi_j$ for all $t \geq 1$, the following holds for Frege's voting method:*

$$\lim_{t \rightarrow \infty} \frac{\rho_j(t)}{t} = \frac{\pi_j}{n}.$$

If $n_t = n$ and for all $t \geq 1$ and for all $j \in C$ there is some $\pi_j^ \in [0, 1]$ such that*

$$\lim_{t \rightarrow \infty} \frac{\sum_{s=1}^t \pi_j^s}{t} = \pi_j^*,$$

then the following holds for all candidates $j \in C$:

$$\lim_{t \rightarrow \infty} \frac{\rho_j(t)}{t} = \frac{\pi_j^*}{n}.$$

The following proposition shows that, for a fixed electorate, proportionality is not only guaranteed in the limit but eventually also within (finite) intervals.

Proposition 1. *If we assume a fixed electorate with n voters, there exists a time $t^* \geq 1$ and a period length $P \in \mathbb{N}$ such that, for all $t \geq t^*$ and all $j \in C$,*

$$\frac{\rho_j(t+P) - \rho_j(t)}{P} = \frac{\pi_j}{n}.$$

If we do not assume a fixed number of voters, proportionality cannot be guaranteed even as $t \rightarrow \infty$. To see this, consider the following example.

Example 3. *We consider a scenario with two candidates, a and b . For $t = 1$, there are three voters and $\pi_a^1 = 2$, $\pi_b^1 = 1$. In every following round the number of voters is doubled and the ratio π_a^t/π_b^t remains the same, namely 2. We thus have $\pi_a^t = 2^{t+1}$ and $\pi_b^t = 2^t$. As we will see, candidate b never wins despite receiving one third of the votes.*

In order to see this, let us prove that $\sigma_a^t > \sigma_b^t$ for all $t \geq 1$ by induction over t . The basis is clearly fulfilled since $\sigma_a^1 = \pi_a^1 = 2 > 1 = \pi_b^1 = \sigma_b^1$. For the induction step, we assume that $\sigma_a^s > \sigma_b^s$ for all $s \leq t$, that is, candidate a has always won so far. We thus have:

$$\sigma_a^{t+1} = \sigma_a^t + 2^{t+2} - \left\lfloor \frac{\sigma_a^t + \sigma_b^t}{2} \right\rfloor \quad \text{and} \quad \sigma_b^{t+1} = \sigma_b^t + 2^{t+1}.$$

Since b has not been chosen so far, $\sigma_b^t = \sum_{s=1}^t \pi_b^s = \sum_{s=1}^t 2^s = 2^{t+1} - 1$. Moreover, since $\sigma_a^t > \sigma_b^t$, it follows:

$$\left\lfloor \frac{\sigma_a^t + \sigma_b^t}{2} \right\rfloor \leq \frac{\sigma_a^t + \sigma_b^t}{2} < \sigma_b^t = 2^{t+1} - 1.$$

We thus have:

$$\sigma_a^{t+1} > \sigma_b^t + 2^{t+2} - \left\lfloor \frac{\sigma_a^t + \sigma_b^t}{2} \right\rfloor > \sigma_b^t + 2^{t+2} - 2^{t+1} + 1 > \sigma_b^t + 2^{t+1} = \sigma_b^{t+1}.$$

Thus, candidate b will never be chosen as representative.

In conclusion, Frege's method is not proportional for arbitrary time intervals, but converges to proportional outcomes if the size of the electorate is fixed.

4. The Modified Frege Method

The examples in Section 2 and the results of Section 3 point to an increasing cost of winning as the reason for Frege's original method failing a reasonable form of proportionality in the initial phase before the cost of winning has stabilised. This observation suggests that a natural variation of Frege's original method, for which the cost of winning is stipulated to be constant, might do better. We thus introduce the following modification of Frege's method, which we will refer to as the *modified Frege method*. In the formal definition of the modified Frege method, we abstract from the size of the electorate and accordingly use *normalised plurality scores* p_j^t for candidates j and times $t \geq 1$:

$$p_j^t = \frac{\pi_j^t}{n_t},$$

where n_t denotes the total number of voters at time t . The aggregate scores for the modified Frege method are defined as follows, where we use Latin letters to denote variables instead of Greek ones as in the definition of Frege's original method:

$$s_j^1 = p_j^1$$

$$s_j^{t+1} = \begin{cases} s_j^t + p_j^{t+1} - 1 & \text{if } \text{repr}(t) = j, \\ s_j^t + p_j^{t+1} & \text{otherwise.} \end{cases}$$

One may wonder why the number 1 is subtracted from the winning candidate. The reason is that this choice ensures that the sum $\sum_{j \in C} s_j^t$ of aggregated scores is invariably 1 for all times t . Hence, this stipulation intuitively corresponds to the cost of winning being equal to the number of voters, as it is eventually the case for Frege's original method (cf. Lemma 1). A potential disadvantage of the modified Frege method is that aggregated scores can become negative; we will further discuss this issue at the end of the paper in Section 6.

Example 4. *Let us now reconsider Example 2 for the modified Frege method, as depicted in the table below. The normalised plurality scores for candidates a through f are 0.1,*

0.1, 0.1, 0.1, 0.1, and 0.5, respectively.

time	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	representative
1	0.1	0.1	0.1	0.1	0.1	0.5	<i>f</i>
2	0.2	0.2	0.2	0.2	0.2	0.0	<i>a</i>
3	−0.7	0.3	0.3	0.3	0.3	0.5	<i>f</i>
4	−0.6	0.4	0.4	0.4	0.4	0.0	<i>b</i>
5	−0.5	−0.5	0.5	0.5	0.5	0.5	<i>c</i>
6	−0.4	−0.4	−0.4	0.6	0.6	1.0	<i>f</i>
7	−0.3	−0.3	−0.3	0.7	0.7	0.5	<i>d</i>
8	−0.2	−0.2	−0.2	−0.2	0.8	1.0	<i>f</i>
9	−0.1	−0.1	−0.1	−0.1	0.9	0.5	<i>e</i>
10	0.0	0.0	0.0	0.0	0.0	1.0	<i>f</i>

We now see that this method produces a proportional outcome: candidate *f* wins five times and all other candidates once, which is exactly in accordance with the candidates' proportional share of the votes.

The modified Frege method enjoys a number of proportionality properties that are stronger than the ones that can be proven for Frege's original method. In the following, let $r_j(t)$ denote the number of times candidate *j* is chosen as representative up to time *t*, that is $r_j(t) = |\{s \leq t : \text{repr}(s) = j\}|$. First, we show that a similar statement to Theorem 1 holds for the modified Frege method:

Theorem 2. *If we assume that $p_j^t = p_j$ for all $t \geq 1$, the following holds for the modified Frege method:*

$$\lim_{t \rightarrow \infty} \frac{r_j(t)}{t} = p_j.$$

If the the normalised plurality scores p_j^t are not fixed but for all $j \in C$ there is some $p_j^ \in [0, 1]$ such that*

$$\lim_{t \rightarrow \infty} \frac{\sum_{s=1}^t p_j^s}{t} = p_j^*,$$

then the following holds for all candidates $j \in C$:

$$\lim_{t \rightarrow \infty} \frac{r_j(t)}{t} = \lim_{t \rightarrow \infty} \frac{\sum_{s=1}^t \pi_j^s}{t} = p_j^*.$$

Theorem 2 provides a proportionality guarantee for the modified Frege method in the long run. Note that Theorem 2 does not require the number of voters to be fixed, in contrast to the analogous result for Frege's original method where this assumption is necessary (cf. Theorem 1 and Example 3). We will now aim for much stronger guarantees, namely guarantees that hold for arbitrary time intervals.

We strengthen the definition of variable integral quota (Definition 1) to hold for arbitrary electorates, inspired by the lower and upper quota axioms in the apportionment setting (cf. Section 5).

Definition 2. For all candidates $j \in C$, let p_j^1, p_j^2, \dots be an infinite sequence of normalized plurality scores. We say that an (infinite) sequence of chosen representatives $\text{repr}(1), \text{repr}(2), \dots$ satisfies *variable upper quota* if for any time $t \geq 1$ it holds that

$$r_j(t) \leq \left\lceil \sum_{s=1}^t p_j^s \right\rceil \quad \text{for every } j \in C,$$

and it satisfies *variable lower quota* if for any time $t \geq 1$ it holds that

$$r_j(t) \geq \left\lfloor \sum_{s=1}^t p_j^s \right\rfloor \quad \text{for every } j \in C.$$

Note that both variable upper and lower quota imply variable integral quota: in case of integral quotas any deviation from a proportional distribution would also violate variable upper and lower quota. In the following we say that the modified Frege method satisfies variable lower or upper quota if any sequence of winners produced by this method satisfies the corresponding axiom.

Theorem 3. *The modified Frege method satisfies variable upper quota.*

As a consequence, the modified Frege method also satisfies variable integral quota. By contrast, it violates variable lower quota, as the following example illustrates.

Example 5. *Let us consider a fixed electorate with 6 candidates and 2750 voters. The plurality scores are 1001, 1000, 206, 182, 181, and 180, respectively; the corresponding normalised plurality scores are obtained by dividing these numbers by 2750. For increased readability, the aggregate scores s_j^t are multiplied by 2750 in the table below.*

time	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	representative
1	1001	1000	206	182	181	180	<i>a</i>
2	−748	2000	412	364	362	360	<i>b</i>
3	253	250	618	546	543	540	<i>c</i>
4	1254	1250	−1926	728	724	720	<i>a</i>
5	−495	2250	−1720	910	905	900	<i>b</i>
6	506	500	−1514	1092	1086	1080	<i>d</i>
7	1507	1500	−1308	−1476	1267	1260	<i>a</i>
8	−242	2500	−1102	−1294	1448	1440	<i>b</i>
9	759	750	−896	−1112	1629	1620	<i>e</i>
10	1760	1750	−690	−930	−940	1800	<i>f</i>
11	2761	2750	−484	−748	−759	−770	<i>a</i>

The variable lower and upper quota of candidate *b* at round 11 is $\frac{11 \cdot 1000}{2750} = 4$, but candidate *b* has been chosen only 3 times, that is, $r_b(11) = 3$.

Similar examples can be found for $m = 4$ (for instance, with plurality scores of 1001, 1000, 115, and 26, and for $t = 30$) and $m = 5$ (for instance, with plurality scores of 1001, 1000, 300, 107, and 92 and for $t = 15$).

We can nevertheless show that the violations of variable lower quota by the modified Frege method are not too severe. This is in particular the case for electorates with few candidates, as the following theorem shows.

Theorem 4. *For $m \in \{2, 3\}$, the modified Frege method satisfies variable lower quota. For $m \geq 4$, we have $r_j(t) \geq \left\lfloor \sum_{s=1}^t p_j^s \right\rfloor - \left\lceil \frac{m-3}{2} \right\rceil$ for every candidate j and time $t \geq 1$.*

The following example shows, however, that variable lower quota violations can still be arbitrarily large. Yet, this construction requires a number of candidates that is exponential in the size of the violation. Thus it may still be possible to strengthen the bounds of Theorem 4 for the cases in which $m \geq 6$. (Example 5 shows that Theorem 4 is optimal for $m = 4$ and $m = 5$.)

Example 6. *We define a variable electorate with candidates $C = \{1, \dots, m\}$ and $n_t = m - t + 1$ voters for $t \in \{1, \dots, m\}$. At time t , we have plurality scores of*

$$\begin{aligned} \pi_j^t &= 0 & \text{for } j \in \{1, \dots, t-1\}, \\ \pi_j^t &= 1 & \text{for } j \in \{t, \dots, m\}. \end{aligned}$$

The corresponding normalized plurality scores are

$$\begin{aligned} p_j^t &= 0 & \text{for } j \in \{1, \dots, t-1\}, \\ p_j^t &= \frac{1}{m-t+1} & \text{for } j \in \{t, \dots, m\}. \end{aligned}$$

Furthermore, we assume that if candidate i and j are tied, then the tie is broken in favor of $\min(i, j)$. Due to this tie-breaking assumption, the modified Frege method selects candidate 1 in the first round, candidate 2 in the second, candidate t in round t . Let us consider round $m-1$, in which candidate $m-1$ wins. The variable lower quota of candidate m is $\left\lfloor \sum_{s=1}^{m-1} p_m^s \right\rfloor = \left\lfloor \frac{1}{m} + \frac{1}{m-1} + \dots + \frac{1}{2} \right\rfloor = \left\lfloor \sum_{i=2}^m \frac{1}{i} \right\rfloor > \sum_{i=1}^m \frac{1}{i} - 2$. We can bound this sum by

$$-2 + \sum_{i=1}^m \frac{1}{i} > -2 + \int_2^{m+1} \frac{1}{i} di = -2 + \log(m+1),$$

which is unbounded for a growing number of candidates m . Candidate m has not won yet and hence we have the following lower bound on the violation of m 's variable lower quota:

$$\left\lfloor \sum_{s=1}^{m-1} p_m^s \right\rfloor - \underbrace{r_m(m-1)}_{=0} > -2 + \log(m+1).$$

Thus, if m tends to infinity, so does the violation of candidate m 's variable lower quota at time $t = m-1$.

5. The Apportionment Setting

In this section, we want to analyse Frege’s methods from the viewpoint of apportionment. Let us first review the apportionment problem and well-known methods that provide apportionment solutions.

5.1. Apportionment methods

An *apportionment problem* for m parties is given by a vote distribution $\mathbf{p} = (p_1, \dots, p_m)$ with $\sum_{i=1}^m p_i = 1$ and a desired house size k . A *solution* to the apportionment problem (\mathbf{p}, k) is an m -sequence of non-negative integers (a_1, \dots, a_m) with $\sum_{i=1}^m a_i = k$. An *apportionment method* is a function that returns for every apportionment problem a valid solution⁵. Apportionment has two main applications: to assign a fixed number of parliamentary seats to parties (proportionally to their vote count), and to assign representatives in a senate to states (proportionally to their population count). From a mathematical point of view, these two applications are indistinguishable, and, in particular, this distinction is not relevant for our study. For the sake of clarity, we speak in the following of parties and seats. We are now going to introduce important apportionment methods (cf. Balinski and Young, 1982; Pukelsheim, 2014).

Largest remainder method. The earliest proposal for an apportionment method is the *largest remainder method* (also: Hamilton method). The largest remainder method assigns in a first step $\lfloor kp_i \rfloor$ seats to each party. In a second step, all remaining seats are distributed so that each party receives at most one seat. Priority is given to parties with the largest remainder, i.e., those with largest $kp_i - \lfloor kp_i \rfloor$.

Divisor methods. Divisor methods are the most commonly used apportionment methods. Their definition is based on *divisor criteria*: A divisor criterion is a monotonically increasing function $d : \mathbb{N} \rightarrow \mathbb{R}$ that satisfies $i \leq d(i) \leq i + 1$ for all $i \geq 0$. A divisor criterion d induces a d -rounding, defined as follows:

$$[x]_d = \{a \in \mathbb{N} : d(a-1) \leq x \leq d(a)\}.$$

If $x = d(a)$ for some a , then $[x]_d$ contains two integers, otherwise only one. For example, the divisor criterion $d(a) = a + 1$ corresponds to rounding down, with the slight difference that rounding down an integer a yields here both a and $a - 1$.

Given a divisor criterion d we define a corresponding divisor method: the set of d -admissible solutions is defined as

$$\left\{ (a_1, \dots, a_m) \in \mathbb{N}^m : \sum_{i=1}^m a_i = k \text{ and } a_i \in \left[\frac{p_i}{x} \right]_d \text{ for some } x \in \mathbb{R} \right\}.$$

⁵To simplify the presentation, we assume that ties are broken in some fashion and thus apportionment methods always return a single solution.

As we require that apportionment methods return only one solution, a tie-breaking mechanism may be necessary to choose one solution in this set.

We can now define the most common divisor methods: The D'Hondt method (or Jefferson method) is defined by $d(a) = a + 1$, i.e., rounding down. The Adams method is defined by $d(a) = a$ (rounding up). The Sainte-Laguë method (or Webster method) is defined by $d(a) = a + 0.5$, which corresponds to rounding to the nearest integer. Finally, the Huntington-Hill method uses the $d(a) = \sqrt{a(a+1)}$ criterion.

Quota method. The quota method (Balinski and Young, 1975) is the most recent addition to this list of apportionment methods. It is defined iteratively, starting with the empty solution $(0, \dots, 0)$. In round $k \geq 1$, if (a_1, \dots, a_m) is the current solution, we consider all parties that would not violate upper quota (cf. Definition 4) if they received an additional seat, i.e., all parties i with $a_i + 1 < \lceil p_i k \rceil$ or, equivalently, $a_i < p_i k$. Then we choose among these parties the one party i with maximum $p_i/(a_i + 1)$ (subject to a tie-breaking, if necessary); party i receives another seat.

Frege's apportionment method. Both Frege's original method and the modified Frege method can easily be transformed into apportionment methods. However, since Frege's original method violates even a very basic proportionality property (weak proportionality, see below), it is not a sensible method in this context and we omit it here from further study. To apply the modified Frege method, we view the vote distribution (p_1, \dots, p_m) as a fixed electorate and apply the method for k rounds, thus obtaining an apportionment solution $(r_1(k), \dots, r_m(k))$. Let us refer to this method as *Frege's apportionment method*.

Given this interpretation, it is natural to ask how Frege's apportionment method compares to other apportionment methods, and, in particular, whether it is equivalent to an already established method in the apportionment setting. We answer these questions in the following section.

5.2. Apportionment axioms

For an overview of apportionment methods and their respective properties, we refer the reader to Table 1; the corresponding analysis can be found, e.g., in the book by Balinski and Young (1982). In the following, we discuss axiomatic properties of Frege's apportionment method. The column "bias" is discussed in Section 5.3.

As a first step, we want to discuss a basic requirement of apportionment methods, called weak proportionality.

Definition 3. An apportionment method satisfies *weak proportionality* if, given an apportionment problem $((p_1, \dots, p_m), k)$ with $k \cdot p_i$ being integer for every $i \in \{1, \dots, m\}$, the method returns (kp_1, \dots, kp_m) .

It is easy to see that our concept of variable integral quota is closely related to weak proportionality. Thus, Frege's original method, seen as an apportionment method, violates this property. This is the reason why we focus on Frege's apportionment method

	house monot.	popul. monot.	lower quota	upper quota	quota for $m = 3$	bias
Largest Remainder	—	—	+	+	+	48.6%
D'Hondt (Jefferson)	+	+	+	—	—	11.9%
Adams	+	+	—	+	—	87.6%
Sainte-Laguë (Webster)	+	+	—	—	+	48.6%
Huntington-Hill	+	+	—	—	—	55.8%
Quota Method	+	—	+	+	+	12.7%
Frege	+	—	—	+	+	54.6%

Table 1: An overview of apportionment methods and their respective properties. The column bias is based on numerical simulations and describes the percentage of instances where the smallest party has been favored over the largest party; a value of 50% would correspond to “no bias” for the used set of apportionment problems (see details in Section 5.3).

(which is based on the modified Frege method). Let us now consider two stronger proportionality requirements:

Definition 4. An apportionment method satisfies *upper quota* if, for any apportionment problem $((p_1, \dots, p_m), k)$, the method returns a solution (a_1, \dots, a_m) satisfying $a_i \leq \lceil kp_i \rceil$ for all $i \in \{1, \dots, m\}$. An apportionment method satisfies *lower quota* if, for any apportionment problem $((p_1, \dots, p_m), k)$, the method returns a solution (a_1, \dots, a_m) satisfying $a_i \geq \lfloor kp_i \rfloor$ for all $i \in \{1, \dots, m\}$. An apportionment method satisfies *quota* if it satisfies both lower and upper quota.

Note that upper quota implies weak proportionality since any deviation from the proportional solution (kp_1, \dots, kp_m) would violate upper quota for some voter. The same holds for lower quota.

Frege’s apportionment method satisfies upper quota (and thus weak proportionality) but fails lower quota. This is an immediate consequence of Theorem 3 and Example 5, respectively. Note that Theorem 4 also holds in the apportionment setting and thus Frege’s apportionment method satisfies quota for $m \in \{2, 3\}$, and for $m \geq 4$ violates lower quota by at most $\lceil \frac{m-3}{2} \rceil$.

Let us now turn to two monotonicity axioms, viz., house and population monotonicity.

Definition 5. An apportionment method satisfies *house monotonicity* if the following holds: for any vote distribution \mathbf{p} and positive integer k , if this method returns the solution (a_1, \dots, a_m) for the problem (\mathbf{p}, k) and the solution (b_1, \dots, b_m) for the problem $(\mathbf{p}, k + 1)$, then there exists $1 \leq i \leq m$ such that (i) $a_i + 1 = b_i$ and (ii) $a_j = b_j$ for all $j \neq i$.

In other words, if the house size increases by one, then the apportionment solution can change only by an increase of 1 for one party. The largest remainder method is notable

in that it actually violates this basic criterium. As Frege's apportionment method is calculated iteratively, it is easy to see that it satisfies house monotonicity.

Definition 6. An apportionment method satisfies *population monotonicity* if the following holds: for vote distributions \mathbf{p} , \mathbf{p}' and a positive integer k , if this method returns the solution (a_1, \dots, a_m) for the problem (\mathbf{p}, k) and the solution (b_1, \dots, b_m) for the problem (\mathbf{p}', k) , then for any $i, j \in \{1, \dots, m\}$:

$$\frac{p'_i}{p'_j} \geq \frac{p_i}{p_j} \text{ implies that either } a'_i \geq a_i \text{ or } a'_j \leq a_j.$$

In other words, if party i increases its vote count relative to party j , then either i does not lose seats or j does not gain seats. We speak of a population paradox if this property is violated: a gain for party i relative to party j grants extra seats for j while i loses seats. The following example shows that Frege's apportionment method suffers from the population paradox and thus violates population monotonicity.

Example 7. Consider the following two scenarios with three parties (a , b , and c) and three seats ($k = 3$): In the first, we have $\mathbf{p} = (2/5, 3/20, 9/20)$, for which Frege's apportionment method yields:

time	a	b	c	representative
1	2/5	3/20	9/20	c
2	4/5	3/10	-1/10	a
3	1/5	9/20	7/20	b

In the second scenario, we have $\mathbf{p}' = (1/4, 1/5, 11/20)$.

time	a	b	c	representative
1	1/4	1/5	11/20	c
2	1/2	2/5	1/10	a
3	-1/4	3/5	13/20	c

Now consider parties b and c . We have

$$\frac{p_b}{p_c} = \frac{1}{3} < \frac{4}{11} = \frac{p'_b}{p'_c},$$

that is, the relative strength of b over c increases, but b loses a seat while c gains one.

To summarize:

Theorem 5. Frege's apportionment method satisfies house monotonicity, upper quota, and quota for $m \in \{2, 3\}$, but fails lower quota for $m \geq 4$ and population monotonicity.

This theorem provides a negative answer to whether Frege's apportionment method corresponds to one of the major apportionment methods. It is not a divisor method, as divisor methods satisfy—and can even be uniquely characterized by—population monotonicity (Balinski and Young, 1982). Furthermore, it is different from the largest remainder method (as the largest remainder method fails house monotonicity) and different from the quota method (as the quota method satisfies lower quota).

5.3. Bias

As a final aspect of apportionment methods, we consider “bias”: Does a method favor small over large parties—or vice versa? Bias is generally more of a concern when using apportionment methods for assigning representatives to states and less so for parties. In parliamentary elections, a bias for larger parties can support the formation of governments and disincentivize schisms of parties (Rae, 1967; Balinski and Young, 1982). However, even when assigning seats to parties, a certain amount of fairness between large and small parties is desirable.

To formalize “bias” as an axiom is difficult, as it is best described as a tendency. Balinski and Young (1982) formalize what it means for a divisor method to be unbiased, but this definition does not extend to arbitrary apportionment methods. Pukelsheim (2014) provides a more general, probabilistic analysis assuming that vote distributions are distributed uniformly at random and that the house size converges to infinity. A third approach is to compute bias in given data sets. This has been done by Balinski and Young (1982) and Birkhoff (1976) based on Congressional apportionment in the USA. It is noteworthy that all these analyses yield similar results, which we summarize below together with our own findings.

Our approach is to determine bias via numerical simulations⁶. We employ the following simple test: we assume five parties, each having a vote count between 1 and 1000, drawn uniformly at random. Furthermore, we assume a house size of 100 seats. For each apportionment method, we computed the number of votes per representative of the strongest and the weakest party. If p_s and p_w are the vote counts for the strongest and weakest party, respectively, and a_s and a_w are the number of seats of the strongest and weakest party, we say that the given apportionment method favors the smaller party if

$$\frac{p_w}{a_w} < \frac{p_s}{a_s},$$

that is, if the smaller party requires fewer votes per seat. We computed the fraction of apportionment problems where the smaller party had this advantage based on 1.000.000 instances. A value of 50% would correspond to being perfectly unbiased, as small and large parties are favored equally often.

The results are shown in Table 1, column “bias”, and can be summarized as follows. Adams favors small parties; D’Hondt and the quota method favor large parties. Sainte-Laguë and the Largest remainder method are well-balanced, as is Huntington-Hill, but to a lesser degree. All these findings are in alignment with previous work (Balinski and Young, 1982; Birkhoff, 1976; Pukelsheim, 2014). Frege’s apportionment method achieves a ratio of 54.5%, in between Sainte-Laguë and Huntington-Hill, and thus can be seen as a rather unbiased apportionment method. Thus, in settings where a bias towards smaller or larger parties is undesirable, Frege’s apportionment method may be more favorable than the quota method.

⁶The source code can be found at <https://github.com/martinlackner/frege>.

6. Conclusions and Discussion

In our mathematical study of Frege’s voting method we focussed on the extent to which it guarantees various forms of proportionality. Accordingly, we ignored a number of its other features that are still worth discussing.

Practical applicability in political elections. It should be noted that the proportionality guarantees of Frege’s method only apply to single constituencies (when observed over time) but not to the political assemblies formed by the chosen representatives. It is thus possible that the political assembly does not at all reflect the entire electorate’s current political opinion. This issue becomes even more dramatic if one considers the actual political decision power within such an assembly (cf. the work on power indices, e.g., Rae, 1969; Dubey and Shapley, 1979). In particular, it may be beneficial for a group of candidates (e.g., a party) to receive few additional votes at time t , so that all of them are elected at time $t + 1$ and thus potentially achieve a majority in the assembly. This paradoxical behaviour leads us to the conclusion that Frege’s method and the modified Frege method are only sensible for single decisions and not so much in the broader sense for electing assemblies. In addition, Frege’s idea is only desirable if the main concern is fairness towards candidates (in the sense that no votes are lost) and only in the absence of harmful extremist opinions (as also extremist candidates would win eventually). Frege’s apportionment method, as introduced in Section 5, is not affected by these considerations and can be recommended in situations where its axiomatic properties appear desirable.

Gerrymandering. Frege claimed in his proposal (*Erläuterungen*, 302, [9]) that his method provides a safeguard against gerrymandering (“*Wahlkreisgeometrie*”), that is, the strategic districting by a political party for electoral gain, an iniquity that infamously pervades representative systems based on first-past-the-post methods for electing representatives (see, e.g., Ricca et al. (2013)). The validity of Frege’s claim, however, much depends on the exact assumptions that are being made and accordingly warrants a careful analysis.

The effectiveness of gerrymandering obviously rests upon the possibility of affecting the proportional support of a party in constituencies. However, if the constituencies are of equal size and the combined electorate of all constituencies is fixed, redistricting will not affect the *sum* of these proportions, no matter how clever the gerrymander.⁷ Thus, if redistricting can only be performed once, Theorem 1 shows that, as time goes to infinity, the number of times each candidate for a party is elected in her respective constituency will be in accordance with this proportion. It would thus follow that Frege’s method does indeed contravene the designs of gerrymanderers. The argument can be generalised to

⁷To make this precise, let there be k constituencies of n voters each and let the distribution of a party’s support be given by a vector $\mathbf{x} = (x_1, \dots, x_k)$, where x_j ($1 \leq j \leq k$) denotes the number of supporters of a given party in constituency j . The proportion of a constituency j supporting a party’s candidate is thus given by x_j/n . Now, redistricting may result in an alternate distribution $\mathbf{y} = (y_1, \dots, y_k)$. Obviously, it holds that $\sum_{1 \leq i \leq k} x_i = \sum_{1 \leq i \leq k} y_i$. Hence, $\sum_{1 \leq j \leq k} \frac{x_j}{n} = \frac{1}{n} \sum_{1 \leq j \leq k} x_j = \frac{1}{n} \sum_{1 \leq j \leq k} y_j = \sum_{1 \leq j \leq k} \frac{y_j}{n}$.

electorates that are not necessarily fixed but still comply with the convergence conditions as in the second part of Theorem 1.

A number of caveats, however, are in place regarding the sweep of this argument. First, as Frege’s voting method is based on the plurality rule, and as such it is still susceptible to gerrymandering for gain in the short run if an individual election in a constituency is seen as a singular event (ignoring past and future elections) or as particularly important. Second, even when considering the temporal nature of Frege’s method, incidental gerrymandering may even be successful in achieving short-term benefits without harming the candidates’ long-term chances. This point also relates to the question of power distribution within an assembly, as discussed above. Third, Frege argues that constituencies should be kept of a similar size and largely remain unchanged over time (*Vorbemerkung*, 197, [1]). This demand by itself excludes some forms of gerrymandering but cannot be seen as a (mathematical) guarantee of his voting method. Therefore, Frege’s method certainly prevents or hinders certain forms of gerrymandering, but to which extent and under which assumptions is a question we leave for future research. Furthermore, it would be interesting to investigate whether the modified Frege method guarantees a better protection against gerrymandering.

Choice candidates and delegation. Among all candidates in a constituency, Frege proposes to distinguish so-called *choice candidates* (“*Erlesene*”), the twenty-five candidates in a constituency with maximal electoral backing, among which the representative will be chosen. So as to ensure that no votes are lost, Frege also provisioned for a delegation mechanism, in which non-choice candidates or deceased choice candidates can transfer the votes cast on them to one of the (living) choice candidates. What exactly the social choice theoretic ramifications of this delegation mechanism are is left as a topic for future research. Furthermore, this mechanism could be compared with modern proposals for vote delegation (Alger, 2006; Green-Armytage, 2015; Blum and Zuber, 2016).

Negative aggregate scores and strategic voting. We have seen how an increasing cost of winning undermines the proportionality guarantees of Frege’s method until this cost stabilizes (after a potentially very long time). The modified Frege method has a constant cost of winning and, consequently, also stronger proportionality guarantees. However, the aggregate scores of the candidates can be negative, which could be considered a serious defect.

The possibility of negative scores renders the modified Frege method vulnerable to the following type of manipulation. Once a candidate has a negative score, it is advantageous to this candidate to retract his or her candidacy in favor of a like-minded person (in social choice terminology a so-called clone), who then starts with an aggregate score of 0 and by that increases their score. A complete study of the manipulability of Frege’s voting method and the modified Frege method is subject to future research.

Transition from plurality to proportionality. The plurality rule performs very badly when it comes to proportionality in the long run. If the electorate is assumed to be fixed,

it will always elect the same candidate! Frege’s method has much better proportionality properties, and the modified Frege method even better ones still. In an important and interesting sense, Frege’s method can be seen as a gradual transition between a system based on plurality towards a system based on the modified Frege method, as the cost of winning is increased until it stabilises at the size of the electorate and henceforward behaves like the modified Frege method.

Outlook and research directions. The temporal or dynamic aspect of the Frege methods distinguishes them from most other voting methods that have been considered in the literature. As such they can also take into account changing electorates and changing opinions among the electorate. Yet, the dynamic Frege methods still rely on the plurality rule in that plurality ballots are used and, for any election at any one single time, they simply select the candidate with the highest aggregate score. Seen this way, the Frege methods could easily be varied upon by considering other ‘static’ social choice rules instead of the plurality rule, thus defining a new class of dynamic voting rules that can be studied from a social choice perspective in their own right. An obvious variation, for instance, would be to assume that the voters’ ballots specify complete preference orders over the candidates. This would allow the computation of Borda scores. The candidates then aggregate their respective Borda scores over time in a similar way as they aggregate plurality scores for the Frege methods. At each election the candidate with the highest Borda score could then be chosen as representative and subsequently incur a certain cost of winning yet to be defined. In order to investigate this class of dynamic voting rules in a systematic and principled fashion, one may want to define axioms that are specific to the temporal setting, like variable quota axioms. In particular it would be interesting to see if there is a dynamic voting rule that satisfies both variable upper and lower quota.

Acknowledgements

Paul Harrenstein was supported by the European Research Council via ERC Advanced Investigator Grant 291528 (“RACE”) and ERC Grant 639945 (“ACCORD”) at Oxford, as well as by the Alan Turing Institute in London. The financial support of Marie-Louise Lackner by the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged. Martin Lackner was supported by the Austrian Science Foundation FWF, grant P31890. We thank Edith Elkind, Piotr Faliszewski, and the participants of the research colloquium on economics and computation at TU Berlin for their helpful comments and suggestions.

A. Proofs of Section 3

Lemma 1. Assume that the number of voters is fixed, i.e., $n_t = n$ for all $t \geq 1$. Then the function $a(t) = \sum_{k \in C} \sigma_k^t$ is monotonically increasing. Moreover, there exists a positive number t_0 such that $a(t) = n \cdot m$ for all $t \geq t_0$.

Proof. Let us establish a recursive definition for $a(t)$, starting with

$$a(1) = \sum_{k \in C} \sigma_k^1 = \sum_{k \in C} \pi_k^1 = n.$$

Let $j^* = \text{repr}(t)$. Then the following equalities hold:

$$\begin{aligned} a(t+1) = \sum_{k \in C} \sigma_k^{t+1} &= \sigma_{j^*}^t + \pi_{j^*}^{t+1} - \left\lfloor \frac{1}{m} \sum_{k \in C} \sigma_k^t \right\rfloor + \sum_{k \neq j^*} (\sigma_k^t + \pi_k^{t+1}) \\ &= \sum_k \sigma_k^t + \sum_k \pi_k^{t+1} - \left\lfloor \frac{1}{m} \sum_{k \in C} \sigma_k^t \right\rfloor \\ &= a(t) + n - \left\lfloor \frac{1}{m} a(t) \right\rfloor \text{ for } t \geq 1. \end{aligned}$$

Now, let us start by proving that there exists a $t_0 \in \mathbb{N}$ such that $a(t) \geq n \cdot m$ for all $t \geq t_0$. For this purpose, let us consider the simpler recursion $b(1) = n$ and $b(t+1) = b(t) + n - \frac{1}{m} \cdot b(t)$. This recursion has the solution $b(t) = nm \left(1 - \left(\frac{m-1}{m}\right)^t\right)$. Note that $b(t)$ converges to nm for $t \rightarrow \infty$. Furthermore, it holds that $b(t) \leq a(t)$ (this can easily be shown by induction). Since $b(t)$ converges to nm and $a(t)$ is integer-valued, there has to be a point in time t_1 such that $a(t) \geq nm$ for all $t \geq t_1$.

Let $t_0 \geq 1$ be the smallest possible choice for t_1 , i.e., t_0 is chosen such that $a(t_0) \geq nm$ and $a(t_0 - 1) < nm$. We want to show that $a(t_0) = nm$. Let $i \in \mathbb{N}$ be such that $a(t_0 - 1) = nm - i$. Then, using the recursion for $a(t)$, we obtain the following:

$$\begin{aligned} a(t_0) &= a(t_0 - 1) + n - \left\lfloor \frac{a(t_0 - 1)}{m} \right\rfloor \\ &= nm - i + n - \left\lfloor \frac{nm - i}{m} \right\rfloor \\ &= nm - i + n - n - \left\lfloor -\frac{i}{m} \right\rfloor \\ &= nm - i + \left\lceil \frac{i}{m} \right\rceil \\ &\leq nm, \end{aligned}$$

since $-i + \left\lceil \frac{i}{m} \right\rceil \leq 0$. It follows that $a(t_0) = nm$.

Now, let us prove that $a(t) = nm$ for all $t \geq t_0$. Using the recursion for $a(t)$, we have $a(t_0 + 1) = nm + n - \left\lfloor \frac{nm}{m} \right\rfloor = nm$. By induction $a(t) = nm$ for all $t \geq t_0$.

Let us finally show that the function $a(t)$ is monotonically increasing. From what was proven so far, we now that $0 \leq a(t) \leq nm$ and thus $0 \leq \left\lfloor \frac{a(t)}{m} \right\rfloor \leq n$ for all $t \geq 1$. Thus, we have for all $t \geq 1$:

$$a(t+1) = a(t) + n - \left\lfloor \frac{1}{m}a(t) \right\rfloor \geq a(t) + n - n = a(t).$$

□

Theorem 1. If we assume a fixed electorate, i.e., $n_t = n$ and $\pi_j^t = \pi_j$ for all $t \geq 1$, the following holds for Frege's voting method:

$$\lim_{t \rightarrow \infty} \frac{\rho_j(t)}{t} = \frac{\pi_j}{n}.$$

If $n_t = n$ and for all $t \geq 1$ and for all $j \in C$ there is some $\pi_j^* \in [0, 1]$ such that

$$\lim_{t \rightarrow \infty} \frac{\sum_{s=1}^t \pi_j^s}{t} = \pi_j^*,$$

then the following holds for all candidates $j \in C$:

$$\lim_{t \rightarrow \infty} \frac{\rho_j(t)}{t} = \frac{\pi_j^*}{n}.$$

Let us first prove a technical lemma, which is required in the proof of Theorem 1:

Lemma A.1. Assume that the number of voters n is fixed. For every candidate $j \in C$ there exists a positive integer c_j such that for all $t \geq t_0$

$$\sigma_j^{t+1} = \sum_{s=1}^{t+1} \pi_j^s - c_j - n \cdot (\rho_j(t) - \rho_j(t_0)), \quad (1)$$

where $t_0 \geq 1$ is such that $\sum_{k \in C} \sigma_k^{t_0} = nm$.

Proof. Let $t_0 \geq 1$ be such that $\sum_{k \in C} \sigma_k^{t_0} = nm$ (the existence of such a t_0 was proven in Lemma 1), and let $c_j \geq 0$ such that

$$\sigma_j^{t_0+1} = \sum_{s=1}^{t_0+1} \pi_j^s - c_j.$$

We shall prove equation (1) by induction over $t \geq t_0$.

For the induction start, let us consider time t_0 :

$$\sigma_j^{t_0+1} = \sum_{s=1}^{t_0+1} \pi_j^s - c_j = \sum_{s=1}^{t_0+1} \pi_j^s - c_j - n \cdot (\rho_j(t_0) - \rho_j(t_0)).$$

For the induction step, we distinguish whether $\text{repr}(t+1) = j$ or not. We may assume that $\sigma_j^{t+1} = \sum_{s=1}^{t+1} \pi_j^s - c_j - (\rho_j(t) - \rho_j(t_0)) \cdot n$. In case $\text{repr}(t+1) = j$, we have:

$$\begin{aligned} \sigma_j^{t+2} &= \sigma_j^{t+1} + \pi_j^{t+2} - \left\lfloor \frac{1}{m} \sum_{k \in C} \sigma_k^{t+1} \right\rfloor \\ &= \sum_{s=1}^{t+1} \pi_j^s - c_j - (\rho_j(t) - \rho_j(t_0)) \cdot n + \pi_j^{t+2} - n \\ &= \sum_{s=1}^{t+2} \pi_j^s - c_j - (\rho_j(t) + 1 - \rho_j(t_0)) \cdot n \\ &= \sum_{s=1}^{t+2} \pi_j^s - c_j - (\rho_j(t+1) - \rho_j(t_0)) \cdot n. \end{aligned}$$

In case $\text{repr}(t+1) \neq j$, we have:

$$\begin{aligned} \sigma_j^{t+2} &= \sigma_j^{t+1} + \pi_j^{t+2} \\ &= \sum_{s=1}^{t+1} \pi_j^s - c_j - (\rho_j(t) - \rho_j(t_0)) \cdot n + \pi_j^{t+2} \\ &= \sum_{s=1}^{t+2} \pi_j^s - c_j - (\rho_j(t) - \rho_j(t_0)) \cdot n \\ &= \sum_{s=1}^{t+1} \pi_j^s - c_j - (\rho_j(t+1) - \rho_j(t_0)) \cdot n, \end{aligned}$$

which concludes the induction step. \square

Proof of Theorem 1. Assuming a fixed electorate, equation (1) becomes:

$$\sigma_j^{t+1} = (t+1) \cdot \pi_j - c_j - (\rho_j(t) - \rho_j(t_0)) \cdot n \quad (2)$$

or, equivalently,

$$\frac{\rho_j(t)}{t} = \frac{\pi_j}{n} + \frac{1}{t} \left(\frac{\pi_j}{n} + \rho_j(t_0) - \frac{c_j}{n} - \frac{\sigma_j^{t+1}}{n} \right)$$

To see that the expression $\frac{1}{t}(\dots)$ converges to 0, note the following: First, $0 \leq \frac{\pi_j}{n} \leq 1$ and $0 \leq \rho_j(t_0) \leq t_0$. Moreover, $0 \leq c_j \leq \rho_j(t_0) \cdot n \leq t_0 \cdot n$ and $0 \leq \sigma_j^{t+1} \leq n \cdot m$. Thus, the term in brackets is bounded from below and above. We obtain:

$$\lim_{t \rightarrow \infty} \frac{\rho_j(t)}{t} = \frac{\pi_j}{n}.$$

Similarly, if the electorate is not fixed but the mean plurality scores converge, i.e., $\lim_{t \rightarrow \infty} 1/t \cdot \sum_{j \in C} \pi_j^t = \pi_j^*$ for all $j \in C$, we have:

$$\frac{\rho_j(t)}{t} = \frac{\sum_{j \in C} \pi_j^t}{t \cdot n} + \frac{1}{t} \left(\frac{\pi_j^{t+1}}{n} + \rho_j(t_0) - \frac{c_j}{n} - \frac{\sigma_j^{t+1}}{n} \right)$$

and thus

$$\lim_{t \rightarrow \infty} \frac{\rho_j(t)}{t} = \lim_{t \rightarrow \infty} \frac{\sum_{j \in C} \pi_j^t}{t \cdot n} = \frac{\pi_j^*}{n}.$$

□

The following proposition shows that, for a fixed electorate, proportionality is not only guaranteed in the limit but eventually also within (finite) intervals.

Proposition 1. If we assume a fixed electorate with n voters, there exists a time $t^* \geq 1$ and a period length $P \in \mathbb{N}$ such that, for all $t \geq t^*$ and all $j \in C$,

$$\frac{\rho_j(t+P) - \rho_j(t)}{P} = \frac{\pi_j}{n}.$$

Proof. From Lemma 1 we know that $0 \leq \sigma_k^t \leq nm$.

Thus the tuple $(\sigma_1^t, \sigma_2^t, \dots, \sigma_m^t)$ can only take finitely many values. Therefore there must be a time $t^* \geq 1$ and an integer P such that

$$(\sigma_1^{t^*}, \sigma_2^{t^*}, \dots, \sigma_m^{t^*}) = (\sigma_1^{t^*+P}, \sigma_2^{t^*+P}, \dots, \sigma_m^{t^*+P}).$$

Given these two values t^* and P , it clearly also holds that

$$(\sigma_1^{t^*+k}, \sigma_2^{t^*+k}, \dots, \sigma_m^{t^*+k}) = (\sigma_1^{t^*+P+k}, \sigma_2^{t^*+P+k}, \dots, \sigma_m^{t^*+P+k}) \text{ for } k \in \mathbb{N} \text{ and thus}$$

$$(\sigma_1^t, \sigma_2^t, \dots, \sigma_m^t) = (\sigma_1^{t+P}, \sigma_2^{t+P}, \dots, \sigma_m^{t+P}) \text{ for all } t \geq t^*.$$

From equation (1) in Lemma A.1 it follows that the following holds for $t \geq t^* - 1$:

$$\rho_j(t+P) - \rho_j(t) = \frac{1}{n} \cdot (\pi_j \cdot (t+P+1) - c_j - \sigma_j^{t+P+1})$$

and

$$\rho_j(t) - \rho_j(t-1) = \frac{1}{n} \cdot (\pi_j \cdot t - c_j - \sigma_j^t).$$

Thus

$$\rho_j(t+P) - \rho_j(t) = \frac{1}{n} \cdot (\pi_j \cdot P - (\sigma_j^{t+P+1} - \sigma_j^{t+1})) = \frac{\pi_j \cdot P}{n},$$

which concludes the proof. □

B. Proofs of Section 4

Theorem 2. If we assume that $p_j^t = p_j$ for all $t \geq 1$, the following holds for the modified Frege method:

$$\lim_{t \rightarrow \infty} \frac{r_j(t)}{t} = p_j.$$

If the the normalised plurality scores p_j^t are not fixed but for all $j \in C$ there is some $p_j^* \in [0, 1]$ such that

$$\lim_{t \rightarrow \infty} \frac{\sum_{s=1}^t p_j^s}{t} = p_j^*,$$

then the following holds for all candidates $j \in C$:

$$\lim_{t \rightarrow \infty} \frac{r_j(t)}{t} = \lim_{t \rightarrow \infty} \frac{\sum_{s=1}^t \pi_j^s}{t} = p_j^*.$$

Let us prove two technical lemmas, which will yield the desired proportionality results.

Lemma B.1. *For the modified Frege method, we have $\sum_{j \in C} s_j^t = 1$ and $-1 < s_j^t - p_j^t$ for all $j \in C$ and $t \geq 1$.*

Proof. The proof of the first statement is by induction on t . For the basis $t = 1$, we immediately have

$$\sum_{j \in C} s_j^1 = \sum_{j \in C} p_j^1 = \sum_{j \in C} \frac{\pi_j^1}{n_1} = 1.$$

For the induction step, we assume $\sum_{j \in C} s_j^t = 1$. Let $j^* = \text{repr}(t)$. Then,

$$\begin{aligned} \sum_{j \in C} s_j^{t+1} &= \sum_{k \neq j^*} (s_k^t + p_k^{t+1}) + s_{j^*}^t + p_{j^*}^{t+1} - 1 \\ &= \sum_{j \in C} s_j^t + \sum_{j \in C} p_j^{t+1} - 1 = 1 + 1 - 1 = 1. \end{aligned}$$

The proof of the second statement is by induction on t as well. For the basis $t = 1$, we immediately have

$$s_j^1 - p_j^1 = p_j^1 - p_j^1 = 0 > -1$$

For the induction step, we assume $-1 < s_j^t - p_j^t$. We distinguish two cases. First, if $j \neq \text{repr}(t)$, we have:

$$s_j^{t+1} - p_j^{t+1} = s_j^t + p_j^{t+1} - p_j^{t+1} = s_j^t \geq s_j^t - p_j^t > -1.$$

Second, if $j = \text{repr}(t)$, it has to hold that $s_j^t \geq \frac{\sum_{j \in C} s_j^t}{m} = \frac{1}{m}$ (in order for j to win). Thus we have:

$$s_j^{t+1} - p_j^{t+1} = s_j^t + p_j^{t+1} - 1 - p_j^{t+1} = s_j^t - 1 \geq \frac{1}{m} - 1 > -1.$$

□

Proof of Theorem 2. Note that the following holds for the modified Frege method:

$$s_j^{t+1} = \sum_{s=1}^{t+1} p_j^s - r_j(t) \text{ for all } t \geq 1 \text{ and } j \in C. \quad (3)$$

This corresponds to a normalized version of equation (1) in the proof of Lemma A.1 with $r_j(t_0) = 0$ and $c_j = 0$.

It follows that

$$\frac{r_j(t)}{t} = \frac{\sum_{s=1}^{t+1} p_j^s}{t} - \frac{s_j^{t+1}}{t}.$$

Since we know from Lemma B.1 that $-1 \leq s_j^{t+1} \leq 1$, both asymptotic results follow immediately. □

Lemma B.2. For every $t \geq 1$ and $j \in C$:

$$-1 < \sum_{s=1}^t p_j^s - r_j(t) \leq \frac{m-1}{2}.$$

The second inequality is strict for $m \geq 3$.

Proof. The lower bound of -1 follows from a combination of Lemma B.1 and Theorem 2: By virtue of equation (3) in the proof of Theorem 2, the lower bound is equivalent to $-1 < s_j^{t+1} - p_j^{t+1}$. Lemma B.1 ensures that this inequality holds for all $j \in C$ and all $t \geq 1$.

To prove the upper bound of $\frac{m-1}{2}$, by virtue of equation (3), it suffices to show by induction on t that $s_j^{t+1} - p_j^{t+1} \leq \frac{m-1}{2}$ for all candidates j and $t \geq 1$. For the basis, let $t = 1$. First assume that $j = \text{repr}(1)$. Then, observing that $0 < p_j^1 \leq 1$,

$$s_j^2 - p_j^2 = s_j^1 + p_j^2 - 1 - p_j^2 = p_j^1 - 1 \leq 0 < \frac{m-1}{2},$$

where the last inequality is strict as we always assume $m \geq 2$. Now assume that $j \neq \text{repr}(1)$. Then, $p_j^1 \leq \frac{1}{2}$, as otherwise j would have been chosen as representative. Accordingly,

$$s_j^2 - p_j^2 = s_j^1 + p_j^2 - p_j^2 = p_j^1 \leq \frac{1}{2} \leq \frac{m-1}{2},$$

where the last inequality is strict if $m \geq 3$.

For the induction step, we may assume $s_j^{t+1} - p_j^{t+1} \leq \frac{m-1}{2}$ to prove that $s_j^{t+2} - p_j^{t+2} \leq \frac{m-1}{2}$ (for $m \geq 3$ we assume strict inequalities). First assume that $j = \text{repr}(t+1)$. Now the following inequalities hold.

$$s_j^{t+2} - p_j^{t+2} = s_j^{t+1} + p_j^{t+2} - 1 - p_j^{t+2} = s_j^{t+1} - 1 \leq s_j^{t+1} - p_j^{t+1} \leq_{i.h.} \frac{m-1}{2}.$$

Finally, let $j^* = \text{repr}(t+1)$ and assume that $j \neq j^*$. Accordingly, $s_j^{t+1} \leq s_{j^*}^{t+1}$. By virtue of Lemma B.1, we have that

$$s_j^{t+1} + s_{j^*}^{t+1} + \sum_{k \in C \setminus \{j, j^*\}} s_k^{t+1} = 1.$$

As we saw above that $-1 < s_k^{t+1} - p_j^{t+1}$, we also have the following:

$$\begin{aligned} s_j^{t+1} + s_{j^*}^{t+1} &= 1 - \sum_{k \in C \setminus \{j, j^*\}} s_k^{t+1} \\ &\leq 1 - \sum_{k \in C \setminus \{j, j^*\}} (s_k^{t+1} - p_k^{t+1}) \\ &< 1 + m - 2 \\ &= (m-1). \end{aligned}$$

As $s_j^{t+1} \leq s_{j^*}^{t+1}$, it follows that $s_j^{t+1} < \frac{m-1}{2}$ and further $s_j^{t+2} - p_j^{t+2} = s_j^{t+1} < \frac{m-1}{2}$. \square

Theorem 3. The modified Frege method satisfies variable upper quota.

Proof. By Lemma B.2, it holds that

$$\sum_{s=1}^t p_j^s - r_j(t) > -1,$$

and consequently

$$r_j(t) < \sum_{s=1}^t p_j^s + 1.$$

This is equivalent to

$$r_j(t) \leq \left\lceil \sum_{s=1}^t p_j^s \right\rceil,$$

which is exactly the condition for variable upper quota. \square

Theorem 4. For $m \in \{2, 3\}$, the modified Frege method satisfies variable lower quota. For $m \geq 4$, we have $r_j(t) \geq \left\lfloor \sum_{s=1}^t p_j^s \right\rfloor - \left\lceil \frac{m-3}{2} \right\rceil$ for every candidate j and time $t \geq 1$.

Proof. By Lemma B.2, it holds that

$$\sum_{s=1}^t p_j^s - r_j(t) < \frac{m-1}{2} \text{ for } m \geq 3,$$

and consequently

$$r_j(t) > \sum_{s=1}^t p_j^s - \frac{m-1}{2}.$$

This is equivalent to

$$r_j(t) \geq \left\lfloor \sum_{s=1}^t p_j^s \right\rfloor - \left\lceil \frac{m-3}{2} \right\rceil.$$

For $m = 3$, the last inequality becomes

$$r_j(t) \geq \left\lfloor \sum_{s=1}^t p_j^s \right\rfloor$$

and thus lower quota is fulfilled.

For $m = 2$, variable lower quota follows from variable upper quota: Let $C = \{a, b\}$ and $\sum_{s=1}^t p_j^s = x_j$ for $j \in \{a, b\}$. By variable upper quota, $r_a(t) \leq \lceil x_a \rceil$ and $r_b(t) \leq \lceil x_b \rceil$. Since $x_a + x_b = t$ and $r_a(t) + r_b(t) = t$, either $r_a(t) = \lfloor x_a \rfloor$ and $r_b(t) = \lceil x_b \rceil$ or $r_a(t) = \lceil x_a \rceil$ and $r_b(t) = \lfloor x_b \rfloor$. In both cases, variable lower quota holds for both a and b . \square

References

- D. Alger. Voting by proxy. *Public Choice*, 126(1-2):1–26, 2006.
- M. Balinski and H. P. Young. *Fair Representation: Meeting the Ideal of One Man, One Vote*. Yale University Press, 1982.
- M. L. Balinski and H. P. Young. The quota method of apportionment. *The American Mathematical Monthly*, 82(7):701–730, 1975.
- G. Birkhoff. House monotone apportionment schemes. *Proceedings of the National Academy of Sciences*, 73(3):684–686, 1976.
- C. Blum and C. I. Zuber. Liquid democracy: Potentials, problems, and perspectives. *Journal of Political Philosophy*, 24(2):162–182, 2016.
- U. Dathe and W. Kienzler. Einleitung. In G. Gabriel and U. Dathe, editors, *Gottlob Frege: Werk und Wirkung. Mit den unveröffentlichten Vorschlägen für ein Wahlgesetz von Gottlob Frege*, pages 285–296. Mentis, 2000.
- P. Dubey and L. S. Shapley. Mathematical properties of the banzhaf power index. *Mathematics of Operations Research*, 4(2):99–131, 1979.
- P. C. Fishburn and S. J. Brams. Paradoxes of preferential voting. *Mathematics Magazine*, 56(4):207–214, 1983.
- G. Frege. Vorschläge für ein Wahlgesetz. Original typescript at the Thüringer Universitäts- und Landesbibliothek (ThULB), 1918. URL https://archive.thulb.uni-jena.de/collections/receive/HisBest_cbu_00022979.
- G. Frege. Vorschläge für ein Wahlgesetz. In G. Gabriel and U. Dathe, editors, *Gottlob Frege: Werk und Wirkung. Mit den unveröffentlichten Vorschlägen für ein Wahlgesetz von Gottlob Frege*, pages 297–313. Mentis, 2000.
- G. Gabriel and U. Dathe, editors. *Gottlob Frege: Werk und Wirkung. Mit den unveröffentlichten Vorschlägen für ein Wahlgesetz von Gottlob Frege*. Mentis, 2000.
- G. Gabriel and W. Kienzler. Gottlob Freges politisches Tagebuch. *Deutsche Zeitschrift für Philosophie*, 42(6):1057–1066, 1994.
- J. Green-Armytage. Direct voting and proxy voting. *Constitutional Political Economy*, 26(2):190–220, 2015.
- R. L. Mendelsohn. Diary: Written by professor Dr Gottlob Frege in the time from 10 March to 9 April 1924. *Inquiry*, 39(3-4):303–342, 1996.
- H. Moulin. Condorcet’s principle implies the no show paradox. *Journal of Economic Theory*, 45(1):53–64, 1988.

- F. Pukelsheim. *Proportional Representation: Apportionment Methods and Their Applications*. Springer, 2014.
- D. W. Rae. *The political consequences of electoral laws*. New Haven: Yale University Press, 1967.
- D. W. Rae. Decision-rules and individual values in constitutional choice. *American Political Science Review*, 63(1):40–56, 1969.
- F. Ricca, A. Scozzari, and B. Simeone. Political districting: from classical models to recent approaches. *Annals of Operations Research*, 204(1):271–299, 2013.