The whole and its parts: On the coherence theorem of Balinski and Young

Antonio Palomares [a], Friedrich Pukelsheim [b,*], Victoriano Ramírez [a]

[a] Departamento de Matemática Aplicada, Universidad de Granada, Spain
[b] Institut für Mathematik, Universität Augsburg, Germany

1. Introduction

The Coherence Theorem of Balinski and Young is a principal result of the axiomatic approach in the theory of apportionment. Apportionment methods serve to calculate seat numbers in parliament, for political parties proportionately to vote counts arising from a popular election, or for geographical regions proportionately to population figures emerging from a periodic census. Inevitably, then, apportionment methods deal with numbers. However, numerical peculiarities do not constitute the principal aspects of apportionment methods. Rather, we 'should argue the merits of the properties of methods, and let their conclusions in principle determine the methods and thus the numbers', as emphasized by Balinski and Young (1978, page 849). This is what the axiomatic approach is meant to achieve.

Over the past decades, the structural properties relevant for this paper have varied in content and nomenclature. Such variations may be found in the publications of the same author(s) and, of course, in those of different authors. For the purpose of this Introduction, we glance over many of the subtleties and unify the terminology by resorting to the notions as used in Pukelsheim (2014).

Apportionment methods are seat allocation rules satisfying five organizing principles: anonymity, balancedness, concordance, decency, and exactness. Coherence is yet another axiomatic property. It demands that every overall apportionment solution – a solution embracing all claimants present – is such that its embedded partial results are also viable solutions, whatever partial set of claimants is singled out. Coherence is the concrete specification of the abstract goal that the whole and its parts fit together in a fair manner. In a review of the origins of the coherence principle, Young (1994, page 173) points out that coherence and its ramifications play a vital role for many and diverse decision procedures. Balinski (2003), too, places coherence into a rather broad perspective.

In the theory of apportionment, the gist of a Coherence Theorem is to show that together with all or some of the organizing principles coherence forces the apportionment method under investigation to be a divisor method. Versions of a Coherence

* Corresponding author.
E-mail addresses: anpalom@ugr.es (A. Palomares), pukelsheim@math.uni-augsburg.de (F. Pukelsheim), vramirez@ugr.es (V. Ramírez).

Theorem are due to Balinski and Young (1978) and, independently, to Hylland (1978), see Balinski and Young (1980, page 2).

The Coherence Theorem of Hylland (1978, Theorems 3 and 10) says as follows:

> If a coherent apportionment rule is balanced, concordant, and decent, then it is compatible with a divisor method.

Anonymity is missing, its lack is outweighed by an appropriate definition of the other notions. Nor does Hylland mention exactness since his notion of divisor methods differs slightly from what has become the standard definition later on.

Balinski and coauthors present two versions of the Coherence Theorem.

1. If a coherent apportionment rule is anonymous, concordant, decent, and weakly exact, then it is a divisor method.

This is Theorem 8.4 in Balinski and Young (1982, page 147). The result is quoted as Theorem 8 in Young (1994, page 187), and as Theorem 1 in Balinski and Ramírez (1999, page 114). Balancedness is left out because it is implied by the other properties. The text preceding Theorem 8.4 sounds as if the theorem holds in the absence of completeness. The other version reads as follows:

2. If a coherent apportionment rule is anonymous, order-preserving, decent, weakly exact, and complete, then it is a divisor method.

This is Theorem 2.2 in Balinski and Rachev (1997, page 8). The authors announce the result as 'a modification' of the first version, without commenting on why they modified the previous version and to what avail.

The proof of version 1 goes a long way for passing from a rank-index method via its accompanying priority ordering to the signpost sequence sought. As moaned by Hylland (1978, page 108), the approach is 'implicit and indirect'. In contrast, we classify the proof of version 2 to be explicit and direct, albeit terse and compact. While version 2 has the nicer proof, version 1 has the nicer assumptions.

Our motivation for looking into the problem was the desire to merge the nice features of version 1 with those of version 2. The result is a version of the Coherence Theorem not assuming order preservation, but instead exploiting concordance. This is achieved by the VRM-Lemma in Section 4 showing that concordance (together with other principles) implies vote ratio monotonicity (and hence order-preservation). A minor challenge was to replace completeness by strong exactness.

In our work, we felt the need to amend the definitions of two properties, exactness and house size monotonicity. In Section 2, we distinguish between weak exactness, the traditional notion of exactness, and strong exactness, the new notion that carries with it a residual dose of completeness. Strong exactness is needed in Part IV of the proof of the Coherence Theorem in Section 5. In Section 4, we explicate the amendment of house size monotonicity. Two-party systems require a different definition than systems with three or more parties. The amended definition enters the proofs of the VRM-Lemma and of the Coherence Theorem.

The paper is organized as follows. Section 2 defines apportionment rules and discusses the five organizing principles that turn a rule into a method. Section 3 introduces divisor methods. Every divisor method originates from a suitable signpost sequence. We comment on the inverse problem of retrieving the signpost sequence when the divisor method is given.

Section 4 introduces the notion of coherence. The HSM-Lemma shows that coherent apportionment methods are house size monotone. In part, the lemma mimics Lemma 2.2 of Balinski and Rachev (1997), and Lemma 4 of Balinski and Ramírez (1999). The VRM-Lemma establishes that coherent apportionment methods are vote ratio monotone. This lemma is new. It links concordance and vote ratio monotonicity in a way missing so far.

With these preparations, the Coherence Theorem in Section 5 assumes a catchy form for which the present paper provides an explicit and direct proof:

> If a coherent apportionment rule is anonymous, balanced, concordant, decent, and strongly exact, then it is compatible with a divisor method.

Strong exactness supersedes weak exactness. Balancedness would be redundant since implied by the other properties; we prefer to list it explicitly.

When adjoining completeness, it is an immediate corollary that the rule actually is equal to a divisor method, rather than only being compatible with it. Thus, the Coherence Theorem entails a characterization of divisor methods: An apportionment method is coherent and complete if and only if it is a divisor method. Section 6 concludes with a discussion.

2. Apportionment methods

The problem formulation, in our standard setting, assumes that there is a parliamentary house whose h seats are to be apportioned among ℓ party lists proportionately to their vote weights $v_1, \ldots, v_\ell$. When the weights are vote counts they are integers, when they are shares of votes they are real numbers between zero and unity. Generally they are assumed to be nonnegative numbers, $v_j \geq 0$. The same apportionment problem arises when allocating h seats among s states proportionately to population figures $p_1, \ldots, p_s$, as in Balinski and Young (1982).

A solution to the problem is an integer vector $x = (x_1, \ldots, x_\ell)$ with vanishing components for vanishing weights, $v_j = 0 \Rightarrow x_j = 0$, and with all components summing to the "house size" h. The integer x_j is construed to be the "seat number of party j", for $j \leq \ell$. We call ℓ the "system size", and x a "seat vector".

Definition. An "apportionment rule" A is a set-valued mapping associating with a given house size h, a given system size ℓ, and a given "weight vector" $v = (v_1, \ldots, v_\ell)$ a nonempty set $A(h; v) = A\big(h; (v_1, \ldots, v_\ell)\big)$ of seat vectors.

Hylland (1978, page 5) calls A an "allotment method".

In order to be elevated to a viable method of practical interest, an apportionment rule must satisfy five organizing principles.

Definition. An "apportionment method" is defined to be an apportionment rule that is (A) anonymous, (B) balanced, (C) concordant, (D) decent, and (sE) strongly exact.

Balinski and Young (1982, pages 97, 144, 147) use the same concepts, but with different names. Anonymity is referred to as symmetry, balancedness has the same name, concordance is called weak population monotonicity, decency is homogeneity, and exactness is termed weak proportionality. Subsequent papers by Balinski and coauthors, and by other authors use further terminological alternatives. Our notions follow Pukelsheim (2014).

The precise definitions of the five organizing principles are as follows.

(A) An apportionment rule A is called "anonymous" when every rearrangement of the vote weights goes along with the same rearrangement of the seat numbers.

(B) An apportionment rule A is called "balanced" when the seat numbers of equally strong parties differ by at most one seat. That is, all seat vectors $(x_1, \ldots, x_\ell) \in A\big(h; (v_1, \ldots, v_\ell)\big)$ and all parties $i, j \leq \ell$ satisfy $v_i = v_j \Rightarrow |x_i - x_j| \leq 1$.

(C) An apportionment rule A is called "concordant" when a stronger party gets at least as many seats as a weaker party. That is, all seat vectors $(x_1, \ldots, x_\ell) \in A\big(h; (v_1, \ldots, v_\ell)\big)$ and all parties $i, j \le \ell$ satisfy $v_i > v_j \Rightarrow x_i \ge x_j$.

(D) An apportionment rule A is called "decent", or homogeneous of degree zero, when a rescaling of the weight vector does not change the solution sets. That is, for all positive scalars $a > 0$ the rule satisfies $A(h; v) = A(h; av)$.

(E) Exactness is the only organizing principle amended by the present paper. We distinguish between "weak exactness", the common notion so far in use, and "strong exactness", the novel notion needed to prove the Coherence Theorem in Section 5.

(wE) An apportionment rule A is called "weakly exact" when every integer weight vector $x = (x_1, \ldots, x_\ell)$ with component sum h reproduces itself as the unique solution. That is, we have $x \in A(h; x)$, and if $y \in A(h; x)$ then $y = x$.

(sE) An apportionment rule A is called "strongly exact" when every integer weight vector $x = (x_1, \ldots, x_\ell)$ with component sum h reproduces itself as the unique solution for all weight vectors in neighborhood of the weight vector x. That is, we have $x \in A(h; x)$, and

$$y \in A\big(h; v(k)\big) \quad \text{for all } k \ge 1 \Longrightarrow y = x$$

for all sequences of weight vectors $v(k)$, $k \ge 1$, that converge to x, where $v_j(k) = 0$ when $x_j = 0$.

Strong exactness implies weak exactness. This is evident from inserting the constant sequence $v(k) = x$.

Weak exactness does not imply strong exactness. As an example, consider the apportionment method Q that starts out by allocating to every party $j \le \ell$ its lower quota, $x_j = \lfloor w_j h \rfloor$, where $w_j = v_j/(v_1 + \cdots + v_\ell)$. The seats remaining are taken care of by instead assigning the upper quota $x_j = \lceil w_j h \rceil$ to as many stronger parties as need be. Evidently Q is weakly exact. However, when the vote weights are $(4 + 1/k, 2 - 1/k)$, the apportionment of six seats results in the seat vector $y = (5, 1)$ for all $k \ge 1$. For the limiting weights $(4, 2)$, the method produces the seat vector $x = (4, 2)$. Since $y \ne x$, the method Q is not strongly exact.

Another notion met in the study of coherent apportionment methods is completeness. An apportionment method A is called "complete" when all seat vectors that are obtainable in a neighborhood around a weight vector v belong to its solution set,

$$y \in A\big(h; v(k)\big) \text{ for all } k \ge 1 \Longrightarrow y \in A(h; v)$$

for all sequences of weight vectors $v(k)$, $k \ge 1$, that converge to v, where $v_j(k) = 0$ when $v_j = 0$.

In the presence of completeness, weak exactness implies strong exactness.

Completeness is a notion of disputed status as it is not in line with electoral practice. It forces an apportionment method to include all tied seat vectors into its solutions sets. But there are electoral laws which, while aiming at an apportionment method A, modify it into a method B by adopting a tie resolution rule such as favoring stronger parties at the expense of weaker parties. That is, in the presence of ties the seat numbers of stronger parties are rounded upwards and the seat numbers of weaker parties are rounded downwards. Thus, B features ties only between parties whose vote counts are literally the same. This modification was already pointed out by Hylland (1978, page 120, note 50), and is actually practiced in Spain and in other countries. If the original method A is complete and weakly exact (and hence strongly exact), then the modified method B, though no longer complete, continues to be strongly exact.

The class of all apportionment methods comprises two important subclasses: the family of divisor methods, and the family of quota methods. In essence, every apportionment method involves two steps: scaling and rounding. Divisor methods employ a flexible scaling step and a fixed rounding step, while quota methods rely on a fixed scaling step and a flexible rounding step, see Pukelsheim (2014, Chapters 4 and 5).

A divisor method scales all weights v_j by some quantity d traditionally called "divisor"; hence the name divisor method. Then, it applies a preordained rounding rule to pass from the quotients v_j/d to integers x_j. The flexibility of the divisor is instrumental to secure the desired component sum, $x_1 + \cdots + x_\ell = h$.

A quota method scales all weights v_j by a formulaic quantity q commonly called "quota"; hence the name quota method. Then, it applies some rounding procedure to turn the quotients v_j/q into integers x_j. The flexibility of the rounding step is employed to achieve the desired component sum, $x_1 + \cdots + x_\ell = h$.

At first glance, the two approaches may seem to differ only insignificantly. That this is not so is one of the main messages of the monograph of Balinski and Young (1982). Divisor methods are procedures that are superior to quota methods from almost every practical and theoretical viewpoint. The Coherence Theorem is but one instance. For this reason, the next section takes a closer look at divisor methods.

3. Divisor methods

Every divisor method D is induced by a specific rounding rule $[\![\cdot]\!]$. The method D maps a house size h and a weight vector $v = (v_1, \ldots, v_\ell)$ into $D(h; v)$, the seat vector set given by

$$D(h; v) = \left\{ (x_1, \ldots, x_\ell) \,\middle|\, x_1 \in \left[\!\!\left[\frac{v_1}{d} \right]\!\!\right], \ldots, x_\ell \in \left[\!\!\left[\frac{v_\ell}{d} \right]\!\!\right] \right. \\ \left. \text{for some } d > 0,\ x_1 + \cdots + x_\ell = h \right\}.$$

That is, the seat number x_j is obtained by applying the rounding rule $[\![\cdot]\!]$ to the quotient v_j/d of the weight v_j and the divisor d. The divisor is determined so that the sum of all seat numbers becomes equal to the house size, $x_1 + \cdots + x_\ell = h$.

The rounding rule $[\![\cdot]\!]$ is a set-valued mapping, as follows. A quotient $q = v_j/d$ lying in the interval $[n-1; n]$ is rounded to the lower or upper integer endpoint, or to both. The value $s(n) \in [n-1; n]$ that determines the split into the region where q is rounded downwards, and the other region where q is rounded upwards, is called the "nth signpost", for $n \ge 1$. It is convenient to prepend the initial signpost $s(0) = 0$. Thus, the interval $[s(n); s(n+1)]$ is the domain of attraction for rounding to the integer n, for all $n \ge 0$, and the union of these intervals covers the half-axis $[0; \infty)$.

The notation in Balinski and Young (1982, page 99) differs by an index shift of one unit. Those authors place their dividing points $d(a)$ into the interval $[a; a+1]$, while our signposts $s(n)$ are located in the interval $[n-1; n]$. Hence the notations are related through $d(a) = s(a+1)$, and $s(n) = d(n-1)$. As a consequence, the sequence of dividing points is initialized by $d(-1) = 0$, see Balinski and Young (1982, page 120).

A vanishing quotient, $q = 0$, is uniquely rounded to zero, $[\![0]\!] = \{0\}$. A quotient $q \in [n-1; n]$ to the right of the signpost, $q \in \big(s(n); n\big]$, is rounded upwards, $[\![q]\!] = \{n\}$. A quotient $q \in [n; n+1]$ to the left of the signpost, $q \in \big[n; s(n+1)\big)$, is rounded downwards, $[\![q]\!] = \{n\}$. A quotient that is equal to a positive signpost, $q = s(n) > 0$, is rounded ambiguously to either endpoint, $[\![s(n)]\!] = \{n-1; n\}$. It is this ambiguity which necessitates a rounding rule to be a set-valued mapping. Altogether we have, for all $q \ge 0$ and $n \ge 0$,

$$[\![q]\!] = \begin{cases} \{0\} & \text{in case } q = 0, \\ \{n\} & \text{in case } q \in \big(s(n), s(n+1)\big), \\ \{n-1, n\} & \text{in case } q = s(n) > 0. \end{cases}$$

Besides initialization and localization, $s(0) = 0$ and $s(n) \in [n-1; n]$ for $n \geq 1$, a "signpost sequence" $s(0), s(1), s(2), \ldots$ must fulfill yet another property so that the induced rounding rule leads to a proper divisor method. The third property, called the "left–right disjunction", demands the following. If there is a signpost hitting the right endpoint of its localization interval then all signposts lie above their left endpoints, and if there is a signpost hitting the left endpoint of its localization interval then all signposts stay below their right endpoints:

$$s(n) = n \text{ for some } n \geq 1 \Longrightarrow s(m) > m-1 \text{ for all } m \geq 2,$$
$$s(m) = m-1 \text{ for some } m \geq 2 \Longrightarrow s(n) < n \text{ for all } n \geq 1.$$

The left–right disjunction is missing from the monograph of Balinski and Young (1982). Its first mentioning is in Balinski and Rachev (1997, page 6) who show that it is indispensable for proving that divisor methods are weakly exact.

In fact, every divisor method is an apportionment rule satisfying the five organizing principles of being anonymous, balanced, concordant, decent, and strongly exact, as is easily verified. Hence a divisor method is a proper apportionment method. This is reflected by speaking of a divisor *method* (and not just of a divisor *rule*).

Normally, a divisor method D is given by the underlying signpost sequence $s(0), s(1), s(2), \ldots$ and its accompanying rounding rule $[\![\cdot]\!]$. Then any solution set $D(h; v)$ calls for a suitable divisor d to verify the definition. As a preparation for the Coherence Theorem in Section 5, it is instructive to study the inverse problem. Given all solution sets $D(h; v)$ of a divisor method D, how do we go about to reconstruct the underlying signpost sequence $s(0), s(1), s(2), \ldots$?

The clue is to start out from a sequence that is proportional to the signpost sequence, but that has one of its members scaled to be equal to unity. Scaling by the first signpost $s(1) \in [0; 1]$ is problematic since $s(1)$ may be zero. The second signpost is certainly positive, $s(2) \geq 1$. Therefore, we introduce the scaled sequence

$$t(1) = \frac{s(1)}{s(2)}, \quad t(2) = 1, \quad t(3) = \frac{s(3)}{s(2)}, \ldots, \quad t(n) = \frac{s(n)}{s(2)}, \ldots.$$

Two-party systems suffice to rebuild the signpost sequence, as we shall see in our coherence discussion. Fix $n \geq 1$ and consider the apportionment of $n+1$ seats among a first party with weight $t(n)$ and a second party with weight unity. Decency permits to multiply the weights by $s(2)$, whence we get $D\big(n+1; (t(n), 1)\big) = D\big(n+1; \big(s(n), s(2)\big)\big) = \big\{(n, 1), (n-1, 2)\big\}$. This is so because (with divisor $d = 1$) the signpost $s(n)$ may be rounded upwards to n or downwards to $n-1$, while $s(2)$ may be rounded upwards to 2 or downwards to 1. Since the house size is $n+1$, one weight must be rounded upwards and the other weight must be rounded downwards. This gives rise to the two solution vectors $(n, 1)$ and $(n-1, 2)$.

Now let the first party lose weight, $t < t(n)$, while the second party maintains its weight unity. The seat vector $(n, 1)$ ceases to be valid. Thus, the value $t(n)$ is the infimum weight for a party to win n seats in a house of size $n+1$ when competing against another party whose weight is unity,

$$t(n) = \inf\left\{t > 0 \,\middle|\, (n, 1) \in D\big(n+1; (t, 1)\big)\right\}.$$

The infimum is taken over a set containing the value $t = n$, by exactness. This entails $t(n) \leq n$ and $t(n)/n \leq 1$. The quotients $t(n)/n$ converge to a limit L,

$$L = \lim_{n\to\infty} \frac{t(n)}{n} = \frac{1}{s(2)} \lim_{n\to\infty} \frac{s(n)}{n} = \frac{1}{s(2)},$$

since the localization property $s(n) \in [n-1; n]$ implies $\lim_{n\to\infty} s(n)/n = 1$. A division by L reproduces the original signpost sequence, $t(n)/L = s(n)$.

In summary, the reconstruction of the signpost sequence underlying a given divisor method D would proceed in three steps. The first step uses the relations $(n, 1) \in D\big(n+1; (t, 1)\big)$ to define the sequence $t(n)$. The second step calculates the limit L of the quotients $t(n)/n$. The third step scales the sequence $t(n)$ to retrieve the signposts, $t(n)/L = s(n)$.

These steps reappear in the proof of the Coherence Theorem in Section 5. However, the Coherence Theorem is more demanding since it does not presuppose the existence of the signpost sequence, but needs to construct it from scratch.

4. Coherence and monotonicity

An established strategy to analyze a large problem with many variables is to dissect it into partial problems with fewer variables. A solution for the whole problem should comprise viable solutions for the partial problems. Balinski and Young (1982, page 141) put it this way: 'An inherent principle of any fair division is that *every part of a fair division should be fair*.' The whole and its parts must fit together in a coherent way.

The definition below follows Balinski and Young (1982, page 141) where the concept is called "uniformity". Young (1994, page 171) speaks of "consistency". The term "coherence" is popularized by Balinski (2003). In the following, we continue to use the latter term since uniformity and consistency are notions to be found also in other contexts of mathematics and statistics.

An apportionment method appears to be coherent when its overall seat vectors are such that their partial seat vectors are valid solutions for the partial problems. In other words, if the overall seat vector $(x_1, \ldots, x_\ell)$ is a solution for the apportionment of h seats in a large system of ℓ parties, then a party subsystem $I \subset \{1, \ldots, \ell\}$ admits the subvector $(x_i)_{i\in I}$ as a solution for the apportionment of $\sum_{i\in I} x_i$ seats among the parties in I. Moreover, if the subproblem features another solution $(y_i)_{i\in I}$ besides $(x_i)_{i\in I}$, then substituting one for the other yields another overall solution.

While the abstract motivation is persuasive, the concrete definition of coherent apportionment methods is notationally cumbersome. Let $x_+ = x_1 + \cdots + x_\ell$ be an abbreviation for the overall component sum of the seat vector x. We denote partial sums of components by $x_I = \sum_{i\in I} x_i$, and complementary sets by $I' = \{1, \ldots, \ell\} \setminus I$.

Definition. An apportionment method A is called "coherent" when, for all system sizes $\ell \geq 2$ and all weight vectors $(v_1, \ldots, v_\ell)$, every seat vector $x \in A(x_+; v)$ fulfills the following two properties for all party subsets $I \subset \{1, \ldots, \ell\}$:

1. Coherence of partial problems: A partial vector of x solves the associated partial apportionment problem, that is, $(x_i)_{i\in I} \in A\big(x_I; (v_i)_{i\in I}\big)$.
2. Coherence of substituted solutions: Whenever tied partial solutions are substituted for partial vectors of x, then the resulting vector is an overall solution as is x, that is, all seat vectors $(y_i)_{i\in I} \in A\big(x_I; (v_i)_{i\in I}\big)$ and $(z_k)_{k\in I'} \in A\big(x_{I'}; (v_k)_{k\in I'}\big)$ satisfy $\big((y_i)_{i\in I}, (z_k)_{k\in I'}\big) \in A\big(x_+; \big((v_i)_{i\in I}, (v_k)_{k\in I'}\big)\big)$.

Coherence of partial problems means that every partial vector that is extracted from an overall seat vector is a valid apportionment solution of the associated partial problem. Coherence of substituted solutions says that tied solutions of a partial problem, when substituted into an overall solution, yield tied overall solutions. That is, if an overall seat vector $x = \big((x_i)_{i\in I}, (x_k)_{k\in I'}\big) \in$

$A\big(x_+; \big((v_i)_{i\in I}, (v_k)_{k\in I'}\big)\big)$ is modified by replacing the partial vector $(x_i)_{i\in I}$ by a tied seat vector $(y_i)_{i\in I} \in A\big(x_I; (v_i)_{i\in I}\big)$, the resulting seat vector yields an overall tie, $\big((y_i)_{i\in I}, (x_k)_{k\in I'}\big) \in A\big(x_+; \big((v_i)_{i\in I}, (v_k)_{k\in I'}\big)\big)$. The same reasoning applies to the complementary subsystem I'.

The definition makes sense only in the presence of anonymity. Without anonymity we would have to observe the order in which the parties are presented, which would force us to consider subsequences. With anonymity the order is negligible, and consideration of subsets suffices.

As soon as coherence is adjoined to the five organizing principles of anonymity, balancedness, concordance, decency, and exactness, it is well known that balancedness becomes dispensable. Balancedness is implied by anonymity, exactness, and coherence, see, for example, Lemma 2.1 in Balinski and Rachev (1997, page 7).

The mutual dependences among these properties is of lesser interest than their joint consequences. Consequences that prove crucial for the Coherence Theorem concern two notions of monotonicity, house size monotonicity and vote ratio monotonicity. Their proper meanings are detailed below. In a nutshell, house size monotonicity deals with varying house sizes in the presence of a constant weight vector. Vote ratio monotonicity addresses varying weight vectors in the presence of a constant house size. The present section shows that coherence implies monotonicity, in either sense.

House size monotonicity is the only other notion besides exactness which we need to amend. Let $x \le y$ denote the componentwise ordering of two seat vectors x and y, that is, $x_j \le y_j$ for all $j \le \ell$. Balinski and Young (1982, page 117) define an apportionment method A to be house size monotone when, given a house size h and a weight vector v, for every seat vector $x \in A(h; v)$ there exists a seat vector $y \in A(h+1; v)$ satisfying $x \le y$. This is the definition adopted by most of the current literature.

The fact that there is at least some vector $y \in A(h+1; v)$ which satisfies the componentwise increase $x \le y$ accounts for the occurrence of ties, or more precisely, of too many ties. Among the tied solutions there may be others for which the componentwise increase does not hold true. For an example, see Table 4.2 in Pukelsheim (2014, page 64). For this reason, it is asking too much to have *all* seat vectors $y \in A(h+1; v)$ satisfy $x \le y$. The definition demands only that at least one of the tied seat vectors works out fine. Yet the current definition is deficient, for two reasons.

Firstly, ties are unproblematic in two-party systems. In tied situations, there are just two equally justified seat vectors according to whether the last seat goes to one party or the other. Here, the notion of house size monotonicity is less sophisticated, by requiring all seat vectors $x \in A(h; v)$ and all seat vectors $y \in A(h+1; v)$ to satisfy $x \le y$. This is equivalent to saying that all seat vectors $y \in A(h+1; v)$ and all seat vectors $x \in A(h; v)$ satisfy $y \ge x$. Hence, the requirement covers house sizes that are increasing, from h to $h+1$, as well as house sizes that are decreasing, from $h+1$ to h.

Secondly, systems with three or more parties may comprise situations where the number of tied solutions is much larger. For systems of size $\ell \ge 3$ the existential quantifier "there exists a seat vector y" is indispensable and cannot be relaxed to the universal quantifier "for all seat vectors y", see Pukelsheim (2014, page 120). However, since the particular succession of quantifiers demands that "for every seat vector $x \in A(h; v)$ there exists a seat vector $y \in A(h+1; v)$", the current definition applies only to house sizes that are increasing, from h to $h+1$. It is silent on what happens when stepping down from a larger house size $h+1$ to the smaller house size h, as observed already by Hylland (1978, page 19).

The amended definition resolves the two deficiencies, as follows.

An apportionment method A is called "house size monotone" when, given house sizes $h < k$ and a two-party weight vector (v_1, v_2), all seat vectors $(x_1, x_2) \in A\big(h; (v_1, v_2)\big)$ and all seat vectors $(y_1, y_2) \in A\big(k; (v_1, v_2)\big)$ satisfy $(x_1, x_2) \le (y_1, y_2)$, and when, given house sizes $h \ne k$ and a weight vector v in a system with three or more parties, for every seat vector $x \in A(h; v)$ there exists a seat vector $y \in A(k; v)$ satisfying $x \le y$ in case $h < k$, and $x \ge y$ in case $h > k$.

HSM-Lemma. *Every coherent apportionment method is house size monotone.*

Proof. The proof is subdivided into two parts. Part I treats systems with two parties, Part II systems with an arbitrary number of parties. Let A denote a coherent apportionment method.

I. Part I is restricted to two-party systems. It suffices to establish the assertion for house size $k = h + 1$. Given a vote vector $v = (v_1, v_2)$ we set $t = v_1/v_2$ and pass to the scaled weights $v/v_2 = (t, 1)$, which is allowed by decency. For two arbitrary seat vectors, $(y_1, y_2) \in A\big(h; (t, 1)\big)$ and $(z_1, z_2) \in A\big(h+1; (t, 1)\big)$ we need to show that $(y_1, y_2) \le (z_1, z_2)$.

Consider the apportionment of $2h + 1$ seats among four parties with respective vote weights t, 1, t, and 1. We claim that in every solution (x_1, x_2, x_3, x_4) the sum of the first two components equals h or $h + 1$,

$$(x_1, x_2, x_3, x_4) \in A\big(2h+1; (t, 1, t, 1)\big) \Longrightarrow x_1 + x_2 \in \{h, h+1\}.$$

The claim is proved as follows. Since x_1 and x_3 are tied via the weight t, and x_2 and x_4 are tied via the weight 1, balancedness restricts the ranges of x_3 and x_4, namely $x_1 - 1 \le x_3 \le x_1 + 1$ and $x_2 - 1 \le x_4 \le x_2 + 1$. If the sum of the first two components is $h - 1$ seats or less, or if it is $h + 2$ seats or more, then the sum of all four components cannot match the house size $2h + 1$,

$$\begin{aligned} x_1 + x_2 \le h - 1 \Longrightarrow x_+ &\le (h-1) + (x_1 + 1) + (x_2 + 1) \\ &\le 2(h-1) + 2 = 2h, \\ x_1 + x_2 \ge h + 2 \Longrightarrow x_+ &\ge (h+2) + (x_1 - 1) + (x_2 - 1) \\ &\ge 2(h+2) - 2 = 2h + 2. \end{aligned}$$

The only possibilities left are $x_1 + x_2 = h$ or $x_1 + x_2 = h + 1$. The claim is proved.

The set $A\big(2h+1; (t, 1, t, 1)\big)$ also contains the solution (x_3, x_4, x_1, x_2) where the first two components are interchanged with the last two components, because of anonymity. If $x_1 + x_2 = h$ then $x_3 + x_4 = h + 1$, while if $x_1 + x_2 = h + 1$ then $x_3 + x_4 = h$. We continue with the first case, $x_1 + x_2 = h$. The second case is handled similarly.

On the one hand coherence of partial problems secures that (x_1, x_2) lies in $A\big(h; (t, 1)\big)$, as does (y_1, y_2). Coherence of substituted solutions allows to replace in the seat vector (x_1, x_2, x_3, x_4) the first two components by (y_1, y_2),

$$(y_1, y_2, x_3, x_4) \in A\big(2h+1; (t, 1, t, 1)\big).$$

On the other hand, coherence of partial problems ascertains that (x_3, x_4) lies in $A\big(h+1; (t, 1)\big)$, as does (z_1, z_2). Coherence of substituted solutions allows to replace in the seat vector (y_1, y_2, x_3, x_4) the last two components by (z_1, z_2),

$$(y_1, y_2, z_1, z_2) \in A\big(2h+1; (t, 1, t, 1)\big).$$

Here, y_1 and z_1 are tied via the weight t, and y_2 and z_2 are tied via the weight 1. Balancedness restricts the ranges of z_1 and z_2 to $y_1 - 1 \le z_1 \le y_1 + 1$ and $y_2 - 1 \le z_2 \le y_2 + 1$. The option $z_1 = y_1 - 1$ fails since it implies $z_1 + z_2 \le (y_1 - 1) + (y_2 + 1) \le h$ which contradicts $z_1 + z_2 = h + 1$. The option $z_2 = y_2 - 1$ fails for the same reason. The remaining options entail $y_1 \le z_1$ and $y_2 \le z_2$. The proof of Part I is complete.

II. Part II turns to larger party systems, $\ell \geq 3$. The case $k = h+1$ is presented in detail. We assume that the apportionment method A is house size monotone for all systems up to size $\ell - 1$, but violates house size monotonicity for an ℓ-party system with some weight vector $v = (v_1, \ldots, v_\ell)$. That is, there exists a seat vector $(x_1, \ldots, x_\ell) \in A(h; v)$ such that all seat vectors $z = (z_1, \ldots, z_\ell) \in A(h+1; v)$ feature some party i that loses seats, $x_i > z_i$. We fix such a vector z. Since the house size grows there must be another party $j \neq i$ that does better than before, $x_j < z_j$. Anonymity permits to move the two parties up front, $i = 1$ and $j = 2$. Hence, we have $x_1 > z_1$ and $x_2 < z_2$.

Coherence of partial problems yields $(x_1, x_2) \in A\big(x_1 + x_2; (v_1, v_2)\big)$ and $(z_1, z_2) \in A\big(z_1 + z_2; (v_1, v_2)\big)$. In case $x_1 + x_2 < z_1 + z_2$, Part I implies $x_1 \leq z_1$ which contradicts $x_1 > z_1$. In case $x_1 + x_2 > z_1 + z_2$, Part I implies $x_2 \geq z_2$ which contradicts $x_2 < z_2$. The case left is $x_1 + x_2 = z_1 + z_2$. Setting $k = h - (x_1 + x_2) = h - (z_1 + z_2)$, coherence of partial problems secures

$$(x_3, \ldots, x_\ell) \in A\big(k; (v_3, \ldots, v_\ell)\big),$$
$$(z_3, \ldots, z_\ell) \in A\big(k+1; (v_3, \ldots, v_\ell)\big).$$

By assumption there is a seat vector $(y_3, \ldots, y_\ell) \in A\big(k+1; (v_3, \ldots, v_\ell)\big)$ satisfying $(x_3, \ldots, x_\ell) \leq (y_3, \ldots, y_\ell)$. Coherence of substituted solutions allows to replace in z the first two components by (x_1, x_2), and the last $\ell - 2$ components by $(y_3, \ldots, y_\ell)$. The two substitutions yield a vector

$$y = (x_1, x_2, y_3, \ldots, y_\ell) \in A(h+1; v)$$

satisfying $x \leq y$. This contradicts the assumption that the apportionment method A violates house size monotonicity for the ℓ-party system with weight vector v.

The case $k = h - 1$ is handled similarly. The proof of Part II is complete. □

Part I of our proof is similar to the proofs of Lemma 2.2 in Balinski and Rachev (1997), and of Lemma 4 in Balinski and Ramírez (1999). A closely related result is Theorem 3 in Hylland (1978, page 25) though the technical assumptions differ slightly.

The other monotonicity notion is vote ratio monotonicity, as it is called in Pukelsheim (2014, page 121). Balinski and Young (1982, page 108) speak of population monotonicity. The issue is to compare a seat vector x originating from a weight vector $v = (v_1, \ldots, v_\ell)$ with a seat vector y belonging to a weight vector $w = (w_1, \ldots, w_\ell)$, while keeping the house size the same, $x \in A(h; v)$ and $y \in A(h; w)$. When the first two parties are to be compared, the input becomes more transparent by standardizing the weight of the second party to be unity. Since this kind of standardization is permitted by decency we switch to the weight vectors

$$\frac{1}{v_2} v = \left(\frac{v_1}{v_2}, 1, \frac{v_3}{v_2}, \ldots, \frac{v_\ell}{v_2}\right),$$
$$\frac{1}{w_2} w = \left(\frac{w_1}{w_2}, 1, \frac{w_3}{w_2}, \ldots, \frac{w_\ell}{w_2}\right).$$

Which seat numbers appear adequate when the weight of the first party increases, $v_1/v_2 < w_1/w_2$, while the second party continues to have weight unity? Common sense would frown upon an outcome where the first party loses seats, $x_1 > y_1$, while simultaneously the second party gains seats, $x_2 < y_2$. The opposite should happen, the first party, doing better, may gain seats, $x_1 \leq y_1$, or the second party, not doing better, may lose seats, $x_2 \geq y_2$. This view leads to the notion of vote ratio monotonicity as in Pukelsheim (2014, page 121).

An apportionment method A is called "vote ratio monotone" when, for every house size h, system size ℓ, and weight vectors $v = (v_1, \ldots, v_\ell)$ and $w = (w_1, \ldots, w_\ell)$, all seat vectors $x \in A(h; v)$ and $y \in A(h; w)$ and all parties $i, j \leq \ell$ satisfy

$$\frac{v_i}{v_j} < \frac{w_i}{w_j} \Longrightarrow x_i \leq y_i \text{ or } x_j \geq y_j.$$

VRM-Lemma. *Every coherent apportionment method is vote ratio monotone.*

Proof. The proof comes in two parts, first treating systems with two parties and then passing to systems with an arbitrary number of parties. Let A denote a coherent apportionment method.

I. Part I handles two-party systems. For every pair of seat vectors $(z_1, z_2) \in A\big(h; (v_1, v_2)\big)$ and $(y_1, y_2) \in A\big(h; (w_1, w_2)\big)$, we need to show that $v_1/v_2 < w_1/w_2$ implies $z_1 \leq y_1$ or $z_2 \geq y_2$. Upon setting $s = v_1/v_2$ and $t = w_1/w_2$ decency allows the more comfortable notation $(y_1, y_2) \in A\big(h; (t, 1)\big)$ and $(z_1, z_2) \in A\big(h; (s, 1)\big)$.

Consider the apportionment of $2h$ seats among four parties with respective vote weights $t, 1, s$, and 1, where $t > s$. We claim that in every solution (x_1, x_2, x_3, x_4) the sum of the first two components is greater than or equal to h,

$$(x_1, x_2, x_3, x_4) \in A\big(2h; (t, 1, s, 1)\big) \Longrightarrow x_1 + x_2 \geq h.$$

The claim is proved by establishing the inequality $x_1 + x_2 \geq x_3 + x_4$. The inequality instantly gives $2(x_1 + x_2) \geq x_+ = 2h$, whence $x_1 + x_2 \geq h$. In view of $t > s$ concordance yields $x_1 \geq x_3$. In case $x_2 \geq x_4$, the inequality $x_1 + x_2 \geq x_3 + x_4$ is immediate. In case $x_2 < x_4$, balancedness entails $x_4 = x_2 + 1$, whence $x_2 + x_4 = 2x_2 + 1$ is odd. Since $x_+ = 2h$ is even, $x_1 + x_3$ is odd, too. It follows that $x_1 \neq x_3$, whence $x_1 > x_3$, that is, $x_1 \geq x_3 + 1$. Now $x_1 + x_2 \geq (x_3 + 1) + (x_4 - 1) = x_3 + x_4$ establishes the inequality in case $x_2 < x_4$. The claim is proved.

The conclusion $x_1 + x_2 \geq h$ is split into two cases: $x_1 + x_2 = h$ and $x_1 + x_2 > h$. The case $x_1 + x_2 = h$ comes with $x_3 + x_4 = h$. Coherence of substituted solutions allows to substitute (y_1, y_2) for (x_1, x_2), and (z_1, z_2) for (x_3, x_4). In the seat vector $(y_1, y_2, z_1, z_2) \in A\big(2h; (t, 1, s, 1)\big)$ concordance ascertains $y_1 \geq z_1$, whence $y_2 \leq z_2$.

The case $x_1 + x_2 > h$ is more intricate. On the one hand, coherence of partial problems secures $(x_1, x_2) \in A\big(x_1 + x_2; (t, 1)\big)$. Compared with $(y_1, y_2) \in A\big(h; (t, 1)\big)$ house size monotonicity implies $x_2 \geq y_2$.

On the other hand, coherence of partial problems ascertains $(x_3, x_4) \in A\big(x_3 + x_4; (s, 1)\big)$, where $x_3 + x_4 = 2h - (x_1 + x_2) < h$. Compared with $(z_1, z_2) \in A\big(h; (s, 1)\big)$ house size monotonicity implies $x_4 \leq z_2$.

The set $A\big(2h; (t, 1, s, 1)\big)$ also contains the solution with second and fourth components interchanged, (x_1, x_4, x_3, x_2), by anonymity. By balancedness, one of the two solutions has its second component less than or equal to its fourth component; without loss of generality, we continue with $x_2 \leq x_4$. This allows to concatenate the preceding inequalities, $y_2 \leq x_2 \leq x_4 \leq z_2$, implying $y_1 = h - y_2 \geq h - z_2 = z_1$. The proof of Part I is complete.

II. Part II treats systems of size $\ell \geq 3$. The argument is indirect. Suppose that a coherent apportionment methods A fails to be vote ratio monotone. Then, it features seat vectors $z \in A(h; v)$ and $y \in A(h; w)$ such that two parties i and j—where without loss of generality we take $i = 1$ and $j = 2$—satisfy $v_1/v_2 < w_1/w_2$ as well as $z_1 > y_1$ and $z_2 < y_2$. Upon setting $s = v_1/v_2$ and $t = w_1/w_2$, coherence of partial problems yields $(z_1, z_2) \in A\big(z_1 + z_2; (s, 1)\big)$ and $(y_1, y_2) \in A\big(y_1 + y_2; (t, 1)\big)$, where $s < t$ as well as $z_1 > y_1$ and $z_2 < y_2$. Contradictions evolve, as follows.

In case $z_1 + z_2 = y_1 + y_2$, Part I yields $z_1 \leq y_1$, whence $z_2 \geq y_2$. This contradicts both, $z_1 > y_1$ as well as $z_2 < y_2$.

In case $z_1 + z_2 < y_1 + y_2$, we introduce a solution $(x_1, x_2) \in A\big(y_1 + y_2; (s, 1)\big)$. For the passage from (z_1, z_2) to (x_1, x_2), house size monotonicity yields $z_1 \leq x_1$ and $z_2 \leq x_2$. For the passage from

(x_1, x_2) to (y_1, y_2), Part I gives $x_1 \leq y_1$, whence $x_2 \geq y_2$. Altogether we get $z_1 \leq y_1$, which contradicts $z_1 > y_1$.

In case $z_1 + z_2 > y_1 + y_2$, we introduce a solution $(x_1, x_2) \in A\big(z_1 + z_2; (t, 1)\big)$. For the passage from (z_1, z_2) to (x_1, x_2), Part I gives $z_1 \leq x_1$, whence $z_2 \geq x_2$. For the passage from (x_1, x_2) to (y_1, y_2), house size monotonicity yields $x_1 \geq y_1$ and $x_2 \geq y_2$. Altogether, we get $z_2 \geq y_2$, which contradicts $z_2 < y_2$. The proof of Part II is complete. □

Balinski and Rachev (1997, page 6) resort to the notion of order-preservation. They call an apportionment method A "order-preserving" when all seat vectors $x \in A(h; v)$ and $y \in A(h; w)$ with $v_1 < w_1$ and $v_j = w_j$ for all $j \geq 2$ satisfy $x_1 \leq y_1$. We mention in passing that Balinski and Young (1982, page 106) address this notion and promptly brush it aside because 'it is not relevant to the problem in an applied sense, since such comparisons scarcely ever occur in practice'.

As a matter of fact, every apportionment method that is vote ratio monotone is order-preserving. The opposite ordering $x_1 > y_1$ is ruled out. Because of a common house size h there would have to be a party $j \geq 2$ with $v_j = w_j$ and $x_j < y_j$. But vote ratio monotonicity does not allow a situation where $v_1/v_j < w_1/w_j$ occurs together with $x_1 > y_1$ and $x_j < y_j$. Hence order-preservation comes for free once we know that vote ratio monotonicity holds true. There is no need to assume order-preservation as a premise when the VRM-Lemma delivers vote ratio monotonicity as a conclusion.

A breach of monotonicity is commonly referred to as a paradox. Quota methods provide ample examples. The most notorious example is the Alabama paradox which illustrates a breach of house size monotonicity. Moreover quota methods may fail vote ratio monotonicity, as is exemplified by Brams and Straffin (1982), Young (1994, page 60), and Pukelsheim (2014, page 122). Quota methods even violate order-preservation. Dančišin (2014) presents an example where a party wins more votes, but gets fewer seats. This is an instance of the no-show paradox, also known under the heading of negative vote weight.

5. The Coherence Theorem

The main result of the present paper states that coherent apportionment methods coincide with divisor methods, except possibly for the handling of ties. See Section 1 for a more detailed survey of the relevant literature.

An apportionment method A is called "compatible with the apportionment method D", when every solution set of A is included in the corresponding solution set of D, that is $A(h; v) \subseteq D(h; v)$ for all house sizes h and all weight vectors v.

Hylland (1978, pages 6, 32) calls a method A that is compatible with D a "submethod" of D and, when D is a divisor method, a "partial divisor method".

Compatibility means that every seat vector of A is a valid seat vector of D as well. Hence, the two methods actually yield identical solutions whenever the solution set $D(h; v)$ is a singleton. Only when $D(h; v)$ contains two or more seat vectors, not all of them need to make their appearance in the solution set $A(h; v)$.

Coherence Theorem. *Every coherent apportionment method is compatible with a divisor method.*

Proof. The proof is subdivided into Parts I–V. Let A denote a coherent apportionment method. Parts I–III investigate two-party solution sets $A(h; v)$. Part I considers equal weights $v = (1, 1)$, Part II special weights $v = (t, 1)$, and Part III arbitrary weights $v = (v_1, v_2)$. Part IV constructs the signpost sequence that induces the desired divisor method D. Part V verifies that A is compatible with D.

I. Part I begins with two parties with equal vote weights, $v_1 = v_2$. Decency permits a passage to the weight vector $(1, 1)$. Here, the solution sets are fully determined through the organizing principles. If the house size is even, balancedness forces the two parties to split the seats evenly. If the house size is odd, balancedness permits a party to be one seat ahead of the other party. Anonymity lets the solution set also comprise the permuted solution, with the other party ahead of the first. In summary,

$$A\big(n+1; (1,1)\big) = \begin{cases} \left\{\left(\dfrac{n+1}{2}, \dfrac{n+1}{2}\right)\right\} & \text{in case } n+1 \text{ is even,} \\ \left\{\left(\dfrac{n}{2}+1, \dfrac{n}{2}\right), \left(\dfrac{n}{2}, \dfrac{n}{2}+1\right)\right\} & \text{in case } n+1 \text{ is odd.} \end{cases} \tag{1}$$

II. Part II discusses two-party systems where the second party has weight unity. That is, the weight vector is of the form $(t, 1)$. For $n \geq 1$ we introduce the set T_n consisting of all $t > 0$ so that in a house of size $n+1$ the first party with weight t gets n seats and the second party with weight unity gets one seat,

$$T_n = \left\{ t > 0 \,\middle|\, (n, 1) \in A\big(n+1; (t, 1)\big) \right\}.$$

These sets are nonempty since $n \in T_n$, by exactness. Vote ratio monotonicity entails that they are intervals, but this fact is not needed in the sequel. Central quantities for the proof are the infima of the sets T_n,

$$t(n) = \inf T_n \in [0; n]. \tag{2}$$

The infimum weights fulfill

$$t(2) = 1 \leq t(n) \quad \text{for all } n > 2. \tag{3}$$

Indeed (1) says $(2, 1) \in A\big(3; (1, 1)\big)$, whence $1 \in T_2$; thus $t(2) \leq 1$. For $t < 1$ concordance yields $(2, 1) \notin A\big(3; (t, 1)\big)$, whence $t \notin T_2$ and $t(2) \geq 1$. Thus, we get $t(2) = 1$. For $n > 2$ we have $t \notin T_n$ for $t < 1$, whence $t(n) \geq 1$.

For large weights $t > t(n)$ definition (2) implies a lower bound for the seats of the first party or an upper bound for the seats of the second party, for all house sizes h and for all seat vectors $(x_1, x_2) \in A\big(h; (t, 1)\big)$:

$$t > t(n) \Longrightarrow x_1 \geq n \text{ or } x_2 \leq 1. \tag{4}$$

To see this, we observe that vote ratio monotonicity and (2) let all solutions $(y_1, y_2) \in A\big(n+1; (t, 1)\big)$ satisfy $y_1 \geq n$, whence $y_2 \leq 1$. For house sizes $h \neq n+1$ and arbitrary solutions $(x_1, x_2) \in A\big(h; (t, 1)\big)$, we apply house size monotonicity: either $h > n+1$ and $x_1 \geq y_1 \geq n$, or $h < n+1$ and $x_2 \leq y_2 \leq 1$. This leads to (4).

Similarly, small weights $0 < s < t(m)$ imply an upper bound for the seats of the first party or a lower bound for the seats of the second party, for all house sizes h and for all seat vectors $(x_1, x_2) \in A\big(h; (s, 1)\big)$:

$$s < t(m) \Longrightarrow x_1 \leq m - 1 \text{ or } x_2 \geq 2. \tag{5}$$

Here, vote ratio monotonicity and (2) let all solutions $(y_1, y_2) \in A\big(m+1; (s, 1)\big)$ fulfill $y_1 \leq m-1$, whence $y_2 \geq 2$. For house sizes $h \neq m+1$ and arbitrary solutions $(x_1, x_2) \in A\big(h; (s, 1)\big)$ again house size monotonicity is applied: either $h < m+1$ and $x_1 \leq y_1 \leq m-1$, or $h > m+1$ and $x_2 \geq y_2 \geq 2$. This proves (5).

III. Part III turns to two-party systems with an arbitrary weight vector (v_1, v_2). We claim that, for all house sizes h and $n, m \geq 1$ with $t(m) > 0$, every seat vector $(y_1, y_2) \in A\big(h; (v_1, v_2)\big)$ satisfies

$$\frac{v_1}{v_2} > \frac{t(n)}{t(m)} \Longrightarrow y_1 \geq n \text{ or } y_2 \leq m - 1. \tag{6}$$

The proof of (6) is as follows. By assumption, we have $t(n)/v_1 < t(m)/v_2$. Using the average $a = \big(t(n)/v_1 + t(m)/v_2\big)/2$, we rescale the weights into $t = av_1$ and $s = av_2$. Since $t > t(n)$ and $s < t(m)$ the new weights allow to employ (4) and (5). We remark that (6) reduces to (4) when $m = 2$, and to (5) when $n = 2$.

Consider the apportionment of $n+m$ seats among three parties with respective weights t, s, and 1. We verify that in every solution (x_1, x_2, x_3) the component x_1 is greater than or equal to n, while x_2 is less than or equal to $m-1$,

$$(x_1, x_2, x_3) \in A\big(n+m; (t, s, 1)\big) \implies x_1 \geq n \text{ and } x_2 \leq m-1. \quad (7)$$

Verification of (7) is in two steps. The first step assumes $x_3 \geq 2$. Coherence of partial problems ascertains $(x_1, x_3) \in A\big(x_1 + x_3; (t, 1)\big)$. From (4), we get $x_1 \geq n$ or $x_3 \leq 1$. The latter inequality is ruled out by assumption, hence $x_1 \geq n$ holds true. We cannot have $x_2 > m-2$ since it implies $x_+ > n + (m-2) + 2 = n + m = x_+$, a contradiction. Hence $x_2 \leq m-1$ holds true, too, as demanded in (7).

The second step handles the remaining case, $x_3 \leq 1$. Coherence of partial problems ascertains $(x_2, x_3) \in A\big(x_2 + x_3; (s, 1)\big)$. From (5), we get $x_2 \leq m-1$ or $x_3 \geq 2$. The latter inequality is ruled out by assumption, hence $x_2 \leq m-1$ holds true. We cannot have $x_1 < n$ since it implies $x_+ < n + (m-1) + 1 = n + m = x_+$, a contradiction. Hence, $x_1 \geq n$ holds true, too, as called for in (7).

In (7), coherence of partial problems secures $(x_1, x_2) \in A\big(x_1 + x_2; (t, s)\big)$, with $x_1 \geq n$ and $x_2 \leq m-1$. Finally we consider arbitrary seat vectors $(y_1, y_2) \in A\big(h; (v_1, v_2)\big) = A\big(h; (t, s)\big)$, the identity following from decency. In case $h = x_1 + x_2$ coherence of substituted solutions yields $(y_1, y_2, x_3) \in A\big(n+m; (t, s, 1)\big)$, whence (7) implies (6). Otherwise (7) is supplemented by house size monotonicity: either $h > x_1 + x_2$ and $y_1 \geq x_1 \geq n$, or $h < x_1 + x_2$ and $y_2 \leq x_2 \leq m-1$. The proof of (6) is complete.

IV. Part IV establishes and explores the inequalities

$$\frac{n}{m-1} > \frac{t(n)}{t(m)} \quad \text{for all } n \geq 1 \text{ and for all } m \geq 2. \quad (8)$$

We remark that (3) secures $t(m) > 0$. The proof of (8) is indirect. Assume that there are integers $n \geq 1$ and $m \geq 2$ satisfying $n/(m-1) \leq t(n)/t(m)$. This inequality necessitates $t(n) > 0$. By passing to reciprocals, $(m-1)/n \geq t(m)/t(n)$, we move closer to (6) except that the premise in (6) features a strict inequality.

By exactness, the seat vector $x = (m-1, n)$ is contained in the solution set $A(m-1+n; x)$. For $k \geq 1$, we define the weights $v_1(k) = m-1+1/k > m-1$, and the constant weights $v_2(k) = n$. They fulfill

$$\frac{v_1(k)}{v_2(k)} > \frac{v_1(k+1)}{v_2(k+1)} > \frac{m-1}{n} \geq \frac{t(m)}{t(n)}.$$

Upon setting $v(k) = \big(v_1(k), v_2(k)\big)$, the vectors $\big(y_1(k), y_2(k)\big) \in A\big(m-1+n; v(k)\big)$ satisfy $y_1(k) \geq y_1(k+1)$, by vote ratio monotonicity. From (6) we get the lower bound $y_1(k) \geq m$. Therefore, the integer sequence $y_1(k), k \geq 1$, is eventually constant to some integer $y_1 \geq m$. Now the seat vector $y = (y_1, m-1+n-y_1)$ lies in $A\big(m-1+n; v(k)\big)$ for eventually all k. But $y \neq x$ contradicts strong exactness. Thus, (8) is established.

Exploration of (8) relies on the equivalent relations $t(n)/n < t(m)/(m-1)$. These relations entail a string of inequalities,

$$\limsup_{n\to\infty} \frac{t(n)}{n} \leq \sup_{n\geq 1} \frac{t(n)}{n} \leq \inf_{m\geq 2} \frac{t(m)}{m-1} \leq \liminf_{m\to\infty} \frac{t(m)}{m-1}$$
$$= \liminf_{n\to\infty} \frac{t(n)}{n}.$$

Since the right limit inferior is less than or equal to the left limit superior, equality holds throughout. Thus, the quotients $t(n)/n$ converge to a limit, $\lim_{n\to\infty} t(n)/n = L$ say. The limit equals the supremum of $t(n)/n$ and the infimum of $t(m)/(m-1)$, whence

$$\frac{t(n)}{n} \leq L \leq \frac{t(m)}{m-1} \quad \text{for all } n \geq 1 \text{ and for all } m \geq 2. \quad (9)$$

This puts us in a position to define the numbers $s(0) = 0$ and $s(n) = t(n)/L$ for $n \geq 1$. They form a signpost sequence. Indeed, inequality (9) with $n = m \geq 2$ yields $s(n)/n \leq 1 \leq s(n)/(n-1)$. This is just another way of expressing the localization property $s(n) \in [n-1; n]$. For $n = 1$, we get $s(1) \in [0; 1]$ from (2) and (9). Furthermore, (8) may be rearranged into $s(n)/n < s(m)/(m-1)$, for all $n \geq 1$ and $m \geq 2$. This is just another way of expressing the left–right disjunction.

V. Part V verifies that the apportionment method A is compatible with the divisor method D that is induced by the signpost sequence $s(n)$, $n \geq 0$. It suffices to show that the seat vectors $x \in A(h; v)$ satisfy the max–min inequality that belongs to D,

$$\max_{i\leq\ell} \frac{v_i}{s(x_i+1)} \leq \min_{j\leq\ell} \frac{v_j}{s(x_j)}. \quad (10)$$

The proof of (10) is as follows. For every two parties $i \neq j$ coherence of partial problems ascertains that $(x_i, x_j) \in A\big(x_i + x_j; (v_i, v_j)\big)$. Suppose that some such pair with $x_i \geq 0$ and $x_j \geq 1$ violates (10) by fulfilling

$$\frac{v_i}{v_j} > \frac{s(x_i+1)}{s(x_j)} = \frac{t(n)}{t(m)},$$

where we have set $n = x_i + 1$ and $m = x_j$. From (6) we get $x_i \geq n = x_i + 1$ or $x_j \leq m - 1 = x_j - 1$, a contradiction. Therefore, we have $v_i/s(x_i+1) \leq v_j/s(x_j)$, and obtain $\max_{i\leq\ell} v_i/s(x_i+1) \leq \min_{j:x_j\geq 1} v_j/s(x_j)$. Upon recalling the convention $v_j/0 = \infty$, no harm is done by adjoining to the right hand minimum the cases $x_j = 0$. The proof of (10) is complete and concludes the proof of the Coherence Theorem. □

Corollary. *An apportionment method is coherent and complete if and only if it is a divisor method.*

Proof. First consider the direct part of the proof. The Coherence Theorem tells us that a coherent apportionment method A has solution sets which are subsets of the solutions sets of some divisor method D. Since A is also assumed to be complete, its solution sets include all tied solutions. The solution sets of D do so, too, by the definition of divisor methods. Hence, the solutions sets coincide. Now $A = D$ shows that A is a divisor method.

For the converse part, we need to verify that a divisor method is coherent and complete. But this is well known, see for instance Balinski and Rachev (1997, page 8) or Pukelsheim (2014, page 118). □

6. Discussion

Balinski and Young (1982, page 98) declare 'the rock-bottom requirements that must be satisfied by any method that is worthy of consideration' to be anonymity, decency, exactness, and completeness. The present paper suggests to drop completeness from the list, and to add balancedness and concordance. This handful of properties constitute the organizing principles by which an apportionment rule is elevated to an apportionment method, in the terminology of Pukelsheim (2014, page 58).

Coherence reduces the vast class of all apportionment methods to the subclass of methods compatible with a divisor method. If additionally assuming that the divisor method transfers properly then it must be parametric, see Theorem 3 in Balinski and Ramírez

(2014, page 44). In the language of Pukelsheim (2014), these are the divisor methods with stationary rounding.

Moreover, Young (1994, pages 49–50, 190) proves that the divisor method with standard rounding is the unique apportionment method that is a coherent extension of the natural two-party apportionments. See also Section 9.3 in Pukelsheim (2014).

Altogether, the axiomatic approach points to a dominant family of apportionment methods, the parametric stationary divisor methods. As envisioned by Balinski and Young (1978, page 849), the Coherence Theorem characterizes these methods primarily by structural properties; mechanical peculiarities are secondary. Procedural transparency and numerical ease of these methods are an added bonus that comes for free.

References

Balinski, M.L., 2003. Quelle équité? Pour Sci., September, 82–87, German translation, Die Mathematik der Gerechtigkeit. Spektrum der Wissenschaft, March 2004, 90–97. English translation, 2005: What is just? American Mathematical Monthly 112, 502–511.

Balinski, M.L., Rachev, S.T., 1997. Rounding proportions: Methods of rounding. Math. Sci. 22, 1–26.

Balinski, M.L., Ramírez, V., 1999. Parametric methods of apportionment, rounding and production. Math. Social Sci. 37, 107–122.

Balinski, M.L., Ramírez, V., 2014. Parametric vs. divisor methods of apportionment. Ann. Oper. Res. 215, 39–48.

Balinski, M.L., Young, H.P., 1978. Stability, coalitions and schism in proportional representation systems. Am. Polit. Sci. Rev. 72, 848–858.

Balinski, M.L., Young, H.P., 1980. The Webster method of apportionment. Proc. Natl. Acad. Sci. USA 77, 1–4.

Balinski, M.L., Young, H.P., 1982. Fair Representation—Meeting the Ideal of One Man, One Vote, second ed. Yale University Press, Brookings Institution Press, New Haven, CT, Washington, DC, [with identical pagination], 2001.

Brams, S.H., Straffin, P.D., 1982. The apportionment problem. Science 217, 437–438.

Dančišin, V., 2014. Negative vote weight and the no-show paradox in party-list proportional system. Working Paper No. 1/2014, University of Prešov, Slovakia [www.researchgate.net/profile/Vladimir_Dancisin].

Hylland, A., 1978. Allotment Methods: Procedures for Proportional Distribution of Indivisible Entities. Mineographed Report, John F. Kennedy School of Government, Harvard University, 1978. Published as Working Paper 1990/11, Research Centre, Norwegian School of Management, Sandvika 1990.

Pukelsheim, F., 2014. Proportional Representation—Apportionment Methods and their Applications. With a Foreword by Andrew Duff MEP. Springer International Publishing, Cham (CH).

Young, H.P., 1994. Equity—In Theory and Practice. Princeton University Press, Princeton, NJ.