Friedrich Pukelsheim

Proportional Representation

Apportionment Methods and Their Applications

Springer

Proportional Representation

Friedrich Pukelsheim

Proportional Representation

Apportionment Methods and Their Applications

With a Foreword by Andrew Duff MEP

Springer

Friedrich Pukelsheim
Universität Augsburg
Institut für Mathematik
Augsburg, Germany

ISBN 978-3-319-03855-1 ISBN 978-3-319-03856-8 (eBook)
DOI 10.1007/978-3-319-03856-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013956638

Mathematics Subject Classification (2010): 91B12, 62P25, 62H17

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

The just, then, is a species of the proportionate .... For proportion is equality of ratios, and involves four terms at least ...; and the just, too, involves at least four terms, and the ratio between one pair is the same as that between the other pair; for there is a similar distinction between the persons and between the things. As the term A, then, is to B, so will C be to D, and therefore, *alternando*, as A is to C, B will be to D. ...

This, then, is what the just is—the proportional; the unjust is what violates the proportion. Hence one term becomes too great, the other too small, as indeed happens in practice; for the man who acts unjustly has too much, and the man who is unjustly treated too little, of what is good. ...

This, then, is one species of the just.

Aristotle, Nichomachean Ethics, Book V, Chapter 3.
Translated and Introduced by Sir *David Ross*, 1953.
The World's Classics 546, Oxford University Press.

Foreword

The virtue of parliamentary democracy rests on the representative capability of its institutions. Even mature democratic states cannot take the strength of its representative institutions for granted. Newer democracies seek practicable ways and means on which to build lasting structures of governance which will command the affinity of the people they are set up to serve. The debate about the structural reform of parliamentary democracies is never far away. Nor should it be. The powers and composition of parliamentary chambers, their rules and working methods, the organisation and direction of the political parties which compete for votes and seats, the electoral systems (who to register, how to vote, how to count), the size and shape of constituencies—all these and more are rightly subject to continual appraisal and are liable to be reformed.

Electoral reform is a delicate business: handled well, it can be the basis on which new liberal democracies spread their wings; it can refresh the old, tired democracies. Handled badly, electoral reform can distort the people's will, entrench the abuse of power and sow the seeds of destruction of liberty. Electoral systems are central to the debate in the emerging democracies, and the relatively new practice of election observation by third parties highlights the need for elections to be run not only fairly but also transparently. Voting and counting should be simple, comprehensible and open to scrutiny—qualities which are too often lacking even in old established democracies.

Electoral reform is also very difficult to achieve. Those who must legislate for it are those very same people who have a vested interest in the status quo. That *Turkeys don't vote for Christmas* is amply demonstrated in the United Kingdom, where reform of the House of Lords has been a lost cause for over a century. Advocates of reform need to stack up their arguments well, be persistent and enjoy long lives.

Friedrich Pukelsheim has written a definitive work on electoral reform. He takes as his starting point the simple premise that seats won in a parliamentary chamber must represent as closely as possible the balance of the votes cast in the ballot box. Rigorous in his methodology, the author knows that there is no single perfect electoral system: indeed, in their quirky details every system affects the exact outcome of an election. We are fortunate indeed that this Professor of Mathematics is a profound democrat. He ably brings to the service of politicians the science of the mathematician.

Dr Pukelsheim was an indispensable participant at the meeting in Cambridge in 2011, chaired by Geoffrey Grimmett, which devised "CamCom"—the best consensual solution to the problem of how to apportion seats in the European Parliament. As the Parliament's rapporteur for electoral procedure, I am happy that our ideas are now taken forward in this publication.

THE EUROPEAN PARLIAMENT

The European Parliament presents unusual challenges both to the scientist and practitioner. It is one chamber of the legislature of the European Union with a lot of power but little recognition. It reflects a giant historical compromise between the international law principle of the equality of states and the democratic motto of "One person, one vote".

Proportional representation at the EU level needs to bear in mind not only party but also nationality. The European Parliament is the forum of the political single market where the different political cultures and constitutional practices of the 28 member states meet up. MEPs are constitutionally *representatives of the Union's citizens* but they are elected not by a uniform electoral procedure but by different procedures under which separate national political parties and candidates fight it out, largely untroubled by their formal affiliation to European political parties.[1] Efforts to make more uniform the election of the world's first multi-national parliament to be directly elected by universal suffrage have been frustrated.

Voter turnout, as we know, has declined at each election to the European Parliament from 62 percent in 1979 to 43 percent in 2009, although these overall figures disguise sharp contrasts among the states and between elections. The long financial and economic crisis since 2008 has brought to a head a crisis of legitimacy for the European Parliament. If the euro is to be salvaged, and the EU as a whole is to emerge strengthened from its time of trial, transnational democracy needs to work better. Banking union and fiscal union need the installation of federal government. That federal government must be fully accountable to a parliament which connects directly to the citizen and with which the citizen identifies. That parliament must be composed in a fair and logical way best achieved in accordance with a settled arithmetical formula and not as a result of unseemly political bartering which borders on gerrymandering and sparks controversy.

It is probable that in spring 2015 there will be a new round of EU constitutional change. This will take the form of a Convention in which heads of government and the European Commission will talk things through with members of the European and national parliaments. Part of the complex negotiations must include the electoral reform of the European Parliament. This will be the chance to progress CamCom for the apportionment of state seats alongside an ambitious proposal for the creation of a

[1] Article 14(2), Treaty on European Union.

pan-European constituency for which a certain number of MEPs will be elected from transnational party lists.[2]

There is no reason to doubt that the notion of *degressive proportionality*, which strikes mathematicians as odd, will survive these negotiations because it expresses quite well the broadly understood belief that in a federal polity the smaller need to be protected from subordination to the larger. CamCom copes logically with degressive proportionality in a way which should satisfy even the austere requirements of the Bundesverfassungsgericht at Karlsruhe.

Nevertheless, as Friedrich Pukelsheim recognises, fully-fledged CamCom means radical adjustments to the number of MEPs elected in several states. It is important, therefore, that changes to the electoral system for one chamber of the legislature are balanced by changes to the electoral system in the other. Here the Jagiellonian Compromise, which uses the square root as the basis for weighing the votes of the member states in the Council, deserves a good hearing.

In June 2013 the Council and European Parliament eventually agreed that the new Member State of Croatia should have 11 MEPs in the Parliament which is to be elected in May 2014. We worked hard to ensure that the re-apportionment of seats would not contradict the logic of CamCom. There is a first, albeit clumsy, legal definition of degressive proportionality. More importantly, the European Union has now formally decided to pursue the objective of a formulaic approach to the future distribution of seats in the Parliament, coupled with a commitment to revisit the matter of qualified majority voting (QMV) in the Council.

The decision of the European Council, now agreed by the European Parliament, lays down that a new system will be agreed in good time before the 2019 elections which *in future will make it possible, before each fresh election to the European Parliament, to allocate the seats between Member States in an objective, fair, durable and transparent way, translating the principle of degressive proportionality as laid down in Article 1, taking account of any change in their number and demographic trends in their population, as duly ascertained thus respecting the overall balance of the institutional system as laid down in the Treaties.*

So perhaps CamCom and JagCom are destined to surface together in the next EU treaty. Legislators who care to understand the maths should start with this book.

Andrew Duff MEP

Cambridge, United Kingdom
September 2013

[2] For a full exposition of this proposal see Spinelli Group, *A Fundamental Law of the European Union,* Bertelsmann Stiftung 2013.

Preface

Proportional representation systems determine how the political views of individual citizens, who are many, mandate the Members of Parliament, who are but a few. The same techniques apply when in Parliament the political groups are to be represented in a committee of a size much smaller than Parliament itself. There are many similar examples all showing that proportional representation inevitably culminates in the task of translating numbers into numbers—large numbers of those to be represented into small numbers of those serving as representatives. The task is solved by procedures called apportionment methods. Apportionment methods and their applications are the theme of this work. A more detailed *Outline of the Book* follows the *Table of Contents*.

By profession a mathematician rather than a politician, I have had the privilege of getting involved in several proportional representation reform projects in recent years. These include the introduction of a double-proportional electoral system in several Swiss cantons since 2006, the amendment of the German Federal Election Law during 2008–2013, and the discussion of the future composition of the European Parliament. The practical challenges and the teaching experience of many lectures and seminars on the subject of proportional representation and apportionment methods have shaped my view and provided the basis for this book.

Apportionment methods may become quite complex. However, these complexities are no ends in themselves. They are reflections of the historical past of a society, its constitutional framework, its political culture, its identity. On occasion the complexities are due to partisan interests of the legislators responsible. This mélange turns the topic into a truly interdisciplinary project. It draws on such fields as constitutional law, European law, political sciences, medieval history, modern history, discrete mathematics, stochastics, computational algorithms, to name but a few. I became increasingly fascinated by the interaction of so many disciplines. My fascination grew when I had the pleasure of conducting student seminars jointly with colleagues from the humanities on topics of common interest. These experiences made me realize that proportional representation and apportionment methods are a wonderful example to illustrate the *universitas litterarum*, the unity of arts and sciences.

In retrospect I find it much easier to conduct an interdisciplinary seminar than to author an interdisciplinary textbook. Nevertheless I hope that the present book may prove a useful reference work for apportionment methods, for scholars of constitutional law and political sciences as well as for other electoral system designers. The many apportionment methods studied span a wide range of alternatives in Germany, the European Union, and elsewhere. The book not only describes the mechanics of each method, but also lists the method's properties: biasedness in favor of stronger parties at the expense of weaker parties, preferential treatments of groups of stronger parties at the expense of groups of weaker parties, optimality with respect to goodness-of-fit or stability criteria, reasonable dependence on such variables as house size, vote ratios, size of the party system, and so on. These properties are rigorously proved and, whenever possible, substantiated by appropriate formulae.

Since the text developed from notes that I compiled for lectures and seminars, I am rather confident that it can be utilized for these purposes. The material certainly suffices for a lecture course or a student seminar in a curriculum of mathematics, quantitative economics, computational social choice, or electoral system design in the political sciences. I have used parts of the text with particular success in classes for students who are going to be high-school teachers. The chapters presuppose readers with an appreciation for rigorous derivations, and with a readiness to accept arguments from scientific fields other than their own. Most chapters can then be mastered with a minimum knowledge of basic arithmetic. Three chapters involve more technically advanced approaches. Chapters 6 and 7 use some stochastic reasoning, and Chapter 14 discrete optimization and computer algorithms.

The subject of the book is restricted to the quantitative and procedural rules that must be employed when a proportional representation system is implemented; as a consequence the book does *not* explicate the qualitative and normative foundations that would be called for when developing a comprehensive theory of proportional representation. As in all sciences, the classification of quantitative procedures starts with basic methods that later get modified to allow for more ambitious settings. The basic issue is to calculate seat numbers proportionately to vote counts. This task is resolved by divisor methods or by quota methods. Later, geographical subdivisions of the electoral region come into play, as do guarantees for small units to obtain representation no matter how small they are, as do restrictions for stronger groups to limit their representation lest they unduly dominate their weaker partners. In order to respond to these requirements the basic methods are modified into variants that may achieve an impressive degree of complexity.

When teaching the topic I soon became convinced that its intricacies can be appreciated only by contemplating real data. That is, data from actual elections in the real world, rather than imaginary data from contrived elections in the academic ivory tower. My Augsburg students responded enthusiastically and set out to devise an appropriate piece of software, BAZI. BAZI has grown considerably since 2000, and has proved an indispensable tool for carrying out practical calculations and theoretical investigations. I would like to encourage readers of this book to use the program to retrace the examples and to form their own judgment. BAZI is freely available from the website www.uni-augsburg.de/bazi.

ACKNOWLEDGMENTS

My introduction to the proportional representation problem was the monograph of *Michel Balinski / Peyton Young* (1982). In their book the authors recount the apportionment history in the House of Representatives of the United States of America, and then proceed to establish a Theory of Apportionment. This seminal source was soon complemented by the treatise of *Klaus Kopfermann* (1991) who adds the European dimension to the proportional representation heritage. *Svante Janson's* (2012) typescript proved invaluable for specific mathematical questions. These books provide the foundations on which the results of the present work are based.

Several colleagues and friends read parts or all of initial drafts of this book and proposed improvements. I have benefited tremendously from the critical comments and helpful suggestions of *Paul Campbell*, *Rudy Fara*, *Martin Fehndrich*, *Dan Felsenthal*, *Svante Janson*, *Jan Lanke*, and *Daniel Lübbert*.

Throughout the project I had the privilege to rely on the advice and inspiration of my colleagues *Karl Heinz Borgwardt*, *Lothar Heinrich*, and *Antony Unwin* in the Augsburg University Institute for Mathematics. A special word of thanks is due to my non-mathematical Augsburg colleagues who helped me mastering the interdisciplinary aspects of the topic. I wish to thank *Günter Hägele* (Medieval History, University Library), *Thomas Krüger* (Medieval History), *Johannes Masing* (Constitutional Law, now with the University of Freiburg im Breisgau), *Matthias Rossi* (Constitutional Law), and *Rainer-Olaf Schultze* (Political Sciences).

The largest debt of gratitude is due to the current and former members of my workgroup at Augsburg University. Many of them contributed substantially to this work through their research work and PhD theses. Moreover they helped in organizing lectures and seminars, in sorting the material, in optimizing the terminology, in polishing the presentation. For their cooperation I am extremely grateful to *Olga Birkmeier*, *Johanna Fleckenstein*, *Christoph Gietl*, *Max Happacher*, *Thomas Klein*, *Sebastian Maier*, *Kai-Friederike Oelbermann*, *Fabian Reffel*, and *Gerlinde Wolsleben*.

I would like to thank *Andrew Duff* MEP for graciously consenting to contribute the foreword to this book. Finally I wish to acknowledge support of the Deutsche Forschungsgemeinschaft.

Friedrich Pukelsheim

Augsburg, Germany
October 2013

Table of Contents

Outline of the Book

CHAPTERS 1, 2: APPORTIONMENT METHODS IN PRACTICE

The initial chapters present an abundance of apportionment methods used in practice. Chapter 1 reviews the European Parliament elections 2009. They provide a rich source of empirical examples. Chapter 2 deals with the German Bundestag election 2009. Emphasis is on the interplay between procedural steps and constitutional requirements. The presentation introduces vital concepts of proportional representation systems beyond their use in European or German elections. Concepts and terminology introduced in these chapters set the scene for the methodological approach that follows.

CHAPTERS 3, 4, 5: DIVISOR METHODS AND QUOTA METHODS

A rigorous apportionment methodology needs to appeal to rounding functions and rounding rules. They are introduced in Chapter 3. Chapters 4 and 5 discuss the two dominant classes of apportionment methods: divisor methods, and quota methods. Usually the sum of all vote counts is much larger than the number of seats available in a parliamentary body. Therefore, a first step converts vote counts into interim quotients of an appropriate order of magnitude. A second step rounds interim quotients to integers. Divisor methods use a flexible divisor for the first step, and a preordained rounding rule for the second. Quota methods employ a formulaic divisor—the quota—for the first step, and a flexible rounding rule for the second.

CHAPTERS 6, 7, 8: DEVIATIONS FROM PROPORTIONALITY

Many apportionment methods deviate from perfect proportionality in a systematic fashion. Chapters 6 and 7 investigate seat biases, that is, averages of the deviations between actual seat numbers and the ideal share of seats assuming that the vote shares follow the uniform distribution or some absolutely continuous distribution. Chapter 8 offers a deterministic comparison of two apportionment methods under the assumption that the vote counts are fixed. The majorization relation is a partial order among apportionment methods indicating whether one method is more beneficial to groups of stronger parties—and hence more disadvantageous to the complementary group of weaker parties—than the other.

CHAPTERS 9, 10, 11: COHERENCE, OPTIMALITY, SPECIFICS

Proportional representation aims at fairly representing voters in terms of their party preferences. Chapter 9 explores the idea that a fair division should be such that every part of it is fair, too. This requirement is captured by the notion of coherence. Divisor methods are coherent, quota methods are not. Chapter 10 evaluates the deviations between actual seat numbers and ideal shares of seats by means of goodness-of-fit criteria. The optimization of particular criteria is shown to lead to particular apportionment methods. Chapter 11 reverses the role of input and output. Given the seat numbers, the range of vote shares is determined that leads to the prespecified number of seats. As a matter of fact it may happen that a straight majority of votes fails to lead to a straight majority of seats. For this reason many electoral laws include an extra majority preservation clause. Three majority clauses are discussed, and their practical usage is illustrated by example.

CHAPTERS 12, 13: PRACTICAL IMPLEMENTATIONS

Many proportional representation systems go beyond abstract proportionality by imposing concrete restrictions. Chapter 12 shows how to handle minimum-maximum restrictions, and empirical examples illustrate their relevance. The most prominent example is the composition of the European Parliament, that is, the allocation of the seats of the European Parliament between the Member States of the Union. Chapter 13 describes the 2013 amendment of the German Federal Election Law. The law achieves impeccable proportionality by adjusting the Bundestag size beyond the nominal level of 598 seats. The system realizes practical equality of the success values of all voters' votes in the whole country. Mild deviations from proportionality may occur when apportioning the seats of a party to its lists of nominees.

CHAPTER 14: DOUBLE PROPORTIONALITY

Chapter 14 treats double-proportional divisor methods. Double-proportionality aims at a fair representation of the geographical division of the electorate as well as of the political division of the voters. The methods achieve this two-way fairness by apportioning seats to districts proportionately to population figures, and seats to parties proportionately to vote counts. The core is the sub-apportionment of seats to each party in each district in such a way that for every district the seats are summing to the given district magnitude, and for every party the seats are summing to their overall proportionate due. To this end two sets of electoral keys are required, district divisors and party divisors. While it is laborious to determine the electoral keys, their publication makes it rather easy to verify the double-proportional seat apportionment.

Notation

$\lfloor t \rfloor, \llbracket t \rrbracket$ (floor)	floor function, 45, rule of downward rounding, 46
$\lceil t \rceil, \llbracket t \rrbracket$ (ceiling)	ceiling function, 47, rule of upward rounding, 47
$\langle t \rangle, \langle\!\langle t \rangle\!\rangle$	commercial rounding, 47, rule of standard rounding, 48
$[t], \llbracket t \rrbracket$	general rounding function, 49, general rounding rule, 49
$\mathbb{N} = \{0, 1, 2, \ldots\}$	set of natural numbers, 44
$s(n), n \in \mathbb{N}$	signpost sequence $\big($always $s(0) = 0\big)$, 51
$s_r(n) = n - 1 + r$	stationary signpost ($n \geq 1$) with split parameter $r \in [0; 1]$, 52
$\widetilde{s}_p(n)$	power-mean signposts with power parameter $p \in [-\infty; \infty]$, 53
$\ell \in \{2, 3, \ldots\}$	number of parties entering the apportionment calculations, 55
$v = (v_1, \ldots, v_\ell)$	vector of vote weights $v_j \in (0; \infty)$ for parties $j \leq \ell$, 55
$v_+ = v_1 + \cdots + v_\ell$	component sum of the vector $v = (v_1, \ldots, v_\ell)$, 55
$h \in \mathbb{N}$	house size, 55
$\mathbb{N}^\ell(h)$	set of seat vectors $x \in \mathbb{N}^\ell$ with component sum $x_+ = h$, 55
$A(h; v)$	set of seat vectors for house size h and vote vector v, 56
A	apportionment rule, 56, apportionment method, 58
v_+/h	votes-per-seats ratio, 41, also known as Hare-quota, 72
$v_j/(v_+/h) = (v_j/v_+)\, h$	ideal share of seats for party $j \leq \ell$, 42
$n+ = \{n, n+1\}$	upward tie, increment option, 64
$n- = \{n-1, n\}$	downward tie, decrement option, 64
$x \preceq y$	majorization of vectors, 111
$A(h, v) \preceq B(h; v)$	majorization of sets of vectors, 112
$A \prec B$	majorization of apportionment methods, 112
$:=$	definitional equality
$\square$	end-of-proof mark
-ward	suffix of adjectives: the downward rounding etc.
-wards	suffix of adverbs: to round downwards etc.
I'	ambiguous prime: complement of the set I
x'	ambiguous prime: transposed vector (or matrix) x
$\bullet$	multi-purpose eye-catcher in tables

CHAPTER 1

Exposing Methods: The 2009 European Parliament Elections

The multitude of apportionment methods that is available for the translation of vote counts into seat numbers is exemplified. The examples are taken from the 2009 European Parliament elections. For each of the 27 Member States the vote counts that enter the apportionment calculations are given, and the computational steps to convert them into seat numbers are described. The exposition is interspersed with conceptual remarks and technical comments concerning proportional representation systems at large.

1.1. THE 27 MEMBER STATES OF THE 2009 UNION

The 2009–2014 European Parliament (EP) was elected in 2009. In this chapter we review the apportionment methods applied by the 27 Member States. Their 27 electoral laws are so distinct that we henceforth refer to the event in the plural, "elections". The domestic differences aptly demonstrate the many apportionment methods used to translate vote counts into seat numbers. Although the elections took place in the past, the properties of the apportionment methods do not depend on the particular instance of their application, and so we use the present tense throughout.

The Union's Interinstitutional Style Guide decrees a protocol order for the Member States that includes three identifiers for each state: an official name, a short name, and a two-letter code. The official names and the short names depend on which of the Union's official languages is used. We choose to list the Member States in the alphabetical order of their two-letter codes, shown in Table 1.1.

The table also includes the seat allocations assigned to the Member States for the 2009 election. The seat allocations have been modified since then, and will be modified again as further states accede. The exposition is restricted to the 27 Member States present in 2009. We group them into batches of three, in the sequence of Table 1.1. Each batch is complemented—somewhat arbitrarily—with general comments. The comments pertain to the procedures used in the Union, the presentation chosen in this book, and the problems addressed in the sequel.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_1,

Two-Letter Code	Short Name	Official Name	Seats
AT	Austria	Republic of Austria	17
BE	Belgium	Kingdom of Belgium	22
BG	Bulgaria	Republic of Bulgaria	17
CY	Cyprus	Republic of Cyprus	6
CZ	Czech Republic	Czech Republic	22
DE	Germany	Federal Republic of Germany	99
DK	Denmark	Kingdom of Denmark	13
EE	Estonia	Republic of Estonia	6
EL	Greece	Hellenic Republic	22
ES	Spain	Kingdom of Spain	50
FI	Finland	Republic of Finland	13
FR	France	French Republic	72
HU	Hungary	Hungary	22
IE	Ireland	Ireland	12
IT	Italy	Italian Republic	72
LT	Lithuania	Republic of Lithuania	12
LU	Luxembourg	Grand Duchy of Luxembourg	6
LV	Latvia	Republic of Latvia	8
MT	Malta	Republic of Malta	5
NL	Netherlands	Kingdom of the Netherlands	25
PL	Poland	Republic of Poland	50
PT	Portugal	Portuguese Republic	22
RO	Romania	Romania	33
SE	Sweden	Kingdom of Sweden	18
SI	Slovenia	Republic of Slovenia	7
SK	Slovakia	Slovak Republic	13
UK	United Kingdom	United Kingdom of Great Britain and Northern Ireland	72
Sum			736

TABLE 1.1 *The 2009 European Union.* The 27 Member States are listed alphabetically by their two-letter codes. The seat allocations are those pertaining to the 2009 elections.

Acronym	Political Group in the 2009 EP
EPP	European People's Party
S & D	Progressive Alliance of Socialists and Democrats
ALDE	Alliance of Liberals and Democrats for Europa
GREENS/EFA	European Greens / European Free Alliance
ECR	European Conservatives and Reformists
EFD	Europe of Freedom and Democracy
GUE/NGL	Gauche unitaire européenne / Nordic Green Left
NA	Non-Attached Members of the EP

TABLE 1.2 *Political Groups in the 2009 EP.* In the 2009 EP seven Political Groups were formed. We treat the non-attached Members of the EP as an eighth pseudo-group, NA.

In the 2009 EP elections political parties campaigned under their domestic names. We inject a European dimension by correlating a domestic party with the Political Group it joined after the election. This is illegal, strictly speaking. Political Groups carry the parliamentary business, and are barred from campaigning in elections. Although invisible in the 2009 EP elections, European Parties do exist. They were invented for funneling money from the Union's budget into political channels. But this book is about apportionment methodology, not money. Neglecting mundane subtleties, we audaciously replace the invisible European Parties by the visible Political Groups.

EP2009AT	Political Group	Votes	Quotient	DivDwn
ÖVP	EPP	858 921	6.1	6
SPÖ	S & D	680 041	4.9	4
Martin	NA	506 092	3.6	3
FPÖ	NA	364 207	2.6	2
GRÜNE	GREENS/EFA	284 505	2.03	2
BZÖ	NA	131 261	0.9	0
Sum (Divisor)		2 825 027	(140 000)	17

TABLE 1.3 *Austria, 2009 EP election.* Of the valid votes, 2 825 027 become effective and let six parties participate in the apportionment calculations. Every 140 000 votes justify roughly one seat out of the Austrian allocation of 17 seats, up to downward rounding of the interim quotients.

The seven Political Groups that formed in the 2009 EP are listed in Table 1.2, together with their acronyms. The last line is devoted to those Members of the EP not joining a group, coded "NA". The labeling "NA" is fortunate since it is standard statistical jargon for items that are Not Available, Not Applicable, Not Active or, for the present purpose, Not Attached to one of the existing Political Groups.

1.2. AUSTRIA–BELGIUM–BULGARIA: ELECTORAL KEYS

Austria has an allocation of 17 seats to fill, of the overall 736 EP seats. The Austrian voters cast 2 864 621 valid votes. However, not all of these become effective to participate in the apportionment calculations. The Austrian law sets an electoral threshold, at four percent of the valid votes. This means that a valid vote becomes effective to enter into the apportionment calculations only when cast for a party drawing at least four percent of all valid votes. Four percent of the Austrian valid votes is 114 584.8 vote fractions. Hence a valid vote becomes effective provided it is cast for a party that draws at least 114 585 votes. Of the eight parties campaigning, two fail the threshold, and the 39 594 votes for them turn ineffective. The apportionment of the 17 seats among the remaining six parties is shown in Table 1.3.

The right-most column in the table contains the final seat numbers and could be labeled "Seats". But the focus is on the procedure how to calculate these seat numbers, whence we label the column with the acronym of the apportionment method used. Since Austria employs the divisor method with downward rounding, the column is labeled "DivDwn". The apportionment methods form the main theme of this book to be treated in great detail, later. The present chapter provides only rudimentary method descriptions to set the scene.

Belgium subdivides the country into what we call electoral *districts*. Of its allocation of 22 seats, the Belgian electoral provisions assign 13 to District 1: *Nederlands kiescollege*, 8 to District 2: *Collège électoral français*, and 1 to District 3: *Deutschsprachiges Wahlkollegium*. Since Belgium does not use an electoral threshold, all valid votes become effective. Seats are apportioned using the divisor method with downward rounding except that now the method is applied three times, separately in each of the three districts. The apportionment is exhibited in Table 1.4.

EP2009BE	Political Group	Votes	Quotient	DivDwn
District 1: Nederlands kiescollege				
CD&V	EPP	948 123	3.8	3
Open VlD	ALDE	837 884	3.4	3
Vl. Belang	NA	647 170	2.6	2
sp.a	S & D	539 393	2.2	2
N-VA	GREENS/EFA	402 545	1.6	1
Groen!	GREENS/EFA	322 149	1.3	1
LDD	ECR	296 699	1.2	1
SLP	GREENS/EFA	26 541	0.1	0
3 Others	NA	55 440	—	0
Sum (Divisor)		4 075 944	(250 000)	13
District 2: Collège électoral français				
PS	S & D	714 947	3.1	3
MR	ALDE	640 092	2.8	2
ECOLO	GREENS/EFA	562 081	2.4	2
cdH	EPP	327 824	1.4	1
8 Others	NA	212 234	—	0
Sum (Divisor)		2 457 178	(230 000)	8
District 3: Deutschsprachiges Wahlkollegium				
CSP	EPP	12 475	1.2	1
PFF	ALDE	7 878	0.8	0
ECOLO	GREENS/EFA	6 025	0.6	0
PS	S & D	5 658	0.6	0
3 Others	NA	6 644	—	0
Sum (Divisor)		38 680	(10 000)	1

TABLE 1.4 *Belgium, 2009 EP election.* Belgium establishes three electoral districts, each evaluated by the divisor method with downward rounding. No threshold is imposed and all valid votes become effective. Of course, parties drawing too few votes fail to obtain representation.

The establishment of separate districts serves a double purpose. It accounts for the existence of distinct language groups, and it secures minority representation. Note that the three divisors vary significantly. To justify roughly one Belgian seat, 250 000 votes are needed in District 1 and 230 000 in District 2, while 10 000 suffice in District 3.

Bulgaria tells a story of its own. The threshold demands that a party secures at least as many votes as the average of valid votes per seat. With 2 576 434 valid votes and an allocation of 17 seats, the average amounts to 151 554.9 vote fractions per seat. Six parties miss the threshold, and their 389 911 votes are discarded as ineffective. The threshold of 151 555 votes amounts to 5.8 percent of votes cast. This percentage share violates the norm set by the European Union. The norm stipulates that the threshold may not exceed five percent of votes cast. Relative to the 2 601 677 votes cast in Bulgaria, five percent is but 130 084 votes. As a consequence the LIDER party, whose 146 984 votes would have justified a seat, is denied representation. Remarkably, nobody complained. No complaint, no redress.

The apportionment method used is the Hare-quota method with residual fit by greatest remainders, HaQgrR. The method relies on a quantity called the *Hare-quota* that is defined to be the votes-per-seats ratio. In the present instance the Hare-quota happens to be an integer, $2\,186\,523/17 = 128\,619$. Division of the Hare-quota 128 619 into the parties' vote counts produces the interim quotients shown in the "Quotient" column of Table 1.5.

EP2009BG	Political Group	Votes	Quotient	HaQgrR
GERB	EPP	627 693	4.880	5
BSP	S & D	476 618	3.706	4
DPS	ALDE	364 197	2.832	3
ATAKA	NA	308 052	2.395	2
NDSV	ALDE	205 146	1.595	2
SDS-DSB	EPP	204 817	1.592	1
Sum (Split)		2 186 523	(.594)	17

TABLE 1.5 *Bulgaria, 2009 EP election.* With the Hare-quota method with residual fit by greatest remainders, every 2 186 523/17 = 128 619 votes justify one seat, thus allocating 13 seats. The four residual seats are given to the parties with the greatest remainders, that is, with remainders above .594.

Each quotient is split into its integral part and its fractional part. The integral parts are instrumental to carry out the first stage of the apportionment, called *main apportionment*. Every full satisfaction of the quota is taken to justify one seat. The strongest party is apportioned 4 seats, while the weaker parties get 3, 2, 2, 1 and 1 seats. Thus the main apportionment allocates 13 seats representing $4 \times 128\,619 + 3 \times 128\,619 + 2 \times 128\,619 + 2 \times 128\,619 + 1 \times 128\,619 + 1 \times 128\,619 = 13 \times 128\,619 = 1\,672\,047$ voters. The second stage of the apportionment, the *residual apportionment*, deals with the $2\,186\,523 - 1\,672\,047 = 514\,476$ remaining votes and the four remaining seats. For the strongest party there are 113 217 ($= 627\,693 - 4 \times 128\,619$) remaining votes, for the others, 90 761, 106 959, 50 814, 76 527, and 76 198. In terms of the votes-per-seats ratio, the remaining votes correspond to the interim quotients' fractional parts .880, .706, .832, .395, .595, and .592. The four greatest claims are 113 217, 106 959, 90 761, and 76 527 votes, or equivalently .880, .832, .706, and .595 quota fractions. Each of these parties is awarded a residual seat. The two smallest remainders are left empty-handed.

Electoral keys: Divisors. An electoral key is a numerical quantity enabling a quick double-check of a published apportionment. We believe that the option to quickly check a result enhances its acceptance with the electorate, and so we take pains to always quote an electoral key.

The electoral key for a divisor method is a *divisor*. Austria provides an instructive example. A divisor of 140 000 means that every 140 000 votes justify roughly one seat. The qualification *roughly* is needed because, literally, the measure of 140 00 votes would justify 6.1 seat fractions for the strongest party, 4.9 seat fractions for the second-strongest party, and so on. The seat fractions, labeled "Quotient" in Table 1.3, are interim quantities deserving only passing attention. Members of Parliament are human beings, whence fractional quotients must always be rounded to whole numbers. Different ways of rounding induce different divisor methods. The method used in Austria is the divisor method with downward rounding. Downward rounding means that all quotients get rounded downwards to the integer below. In other words fractions are simply neglected. Thus the "Quotient" column documents the interim quantities that, after being rounded downwards, yield the seat numbers in the "DivDwn" column.

Caution must be exercised when utilizing divisors as indicators for representative equality. Not meant to serve this purpose, they provide no more than a meek measure of representativeness. In fact, there is some leeway which divisor to pick. Sometimes this is emphasized by speaking of a *flexible divisor* or, alternatively, of a *sliding divisor*.

EP2009CY	Political Group	Votes	Quotient	HQ3grR
DISY	EPP	109 209	2.139	2
AKEL	GUE/NGL	106 922	2.094	2
DI.KO	S & D	37 625	0.737	1
EDEK	S & D	30 169	0.591	1
EVROKO	ALDE	12 630	0.247	0
Ineffective votes		9 770	—	—
Sum (Split)		306 325	(.5)	6

TABLE 1.6 *Cyprus, 2009 EP election.* The Hare-quota variant-3 relies on the votes-per-seats ratio by rounding it downwards, HQ3 = $\lfloor 306\,325/6 \rfloor = \lfloor 51\,054.2 \rfloor = 51\,054$. The two residual seats go to the parties that have a quotient with a remainder above the quoted split .5.

For instance, in Belgium the apportionments in Districts 1 and 2 use divisors 250 000 and 230 000 (Table 1.4). A common divisor 238 000 would be equally feasible. To see this, consider District 1. If the divisor were smaller than $948\,123/4 = 237\,030.75$, the strongest party's quotient would grow greater than four, such as $948\,123/237\,000 = 4.0005$. Hence the party would be awarded a fourth seat and the seats would sum to 14 or more, even though there are only 13. If the divisor were larger than $539\,393/2 = 269\,696.5$, the fourth-strongest party would lose a seat, and the number of seats would fall to 12 or fewer. The given allotment of 13 seats is exhausted if and only if the divisor belongs to the *divisor interval* [237 030.75; 269 696.5] that is delimited by the two critical divisor values given above. Similarly District 2 hands out exactly eight seats if and only if the divisor lies in the interval pertaining to this district, [213 364; 238 315.7]. Feasibility of 238 000 follows since it is included in both the intervals.

We exploit the flexible nature of divisors by picking from the divisor interval a value that eases communication. The divisors quoted in our tables are calculated by starting from the midpoint of the divisor interval and reducing it to as few digits as the interval permits. For example, in District 1 the midpoint 253 500.1 of the divisor interval [237 030.75; 269 696.5] is reduced to 250 000. For District 2 the midpoint 225 839.8 of the interval [213 364; 238 315.7] leads to 230 000.

Electoral keys: Splits. The electoral key for a quota method with residual fit by greatest remainders is called a *split*, that is, a remainder value separating the parties that are awarded one of the residual seats from those that are not. Generally the term "quota" is indicative of a fixed divisor, a well-defined and unique quantity that serves to scale down (large) vote counts into (smaller) interim quotients of the size of the seat numbers. Hare-quota methods obtain their interim quotients by dividing the vote counts by the Hare-quota, the votes-per-seats ratio. We use the term "divisor" when the divisor is flexible, as opposed to "quota" when it is fixed; the distinction is technical jargon, but useful. Divisor methods employ a flexible divisor and a fixed rounding rule. Quota methods combine a fixed divisor with a flexible rounding rule.

The prime example of how quota methods round interim quotients to whole numbers is the residual fit by greatest remainders. In essence this means that the method uses a flexible rounding rule. This fact becomes evident when looking at numbers. In Bulgaria the quotients with the fourth- and fifth-greatest remainders are $205\,146/128\,619 = 1.594\,990$ and $204\,817/128\,619 = 1.592\,432$. Hence every remainder splitting point—*split*, for short—in the *split interval* [.592 432; .594 990] splits all parties into those that receive one of the residual seats, and those that do not. The strategy

EP2009CZ	Political Group	Votes	Quotient	DivDwn
ODS	ECR	741 946	9.9	9
ČSSD	S & D	528 132	7.04	7
KSČM	GUE/NGL	334 577	4.5	4
KDU-ČSL	EPP	180 451	2.4	2
Sum (Divisor)		1 785 106	(75 000)	22

TABLE 1.7 *Czech Republic, 2009 EP election.* Twenty-nine parties fail the threshold, and their 573 828 votes turn ineffective. The Czech allocation of 22 seats is apportioned among the remaining four parties by means of the divisor method with downward rounding.

which split to quote is the same as with divisors. Here we reduce the midpoint of the split interval, .593 711, to .594. This is the value quoted in Table 1.5. The publication of a split facilitates checking whether a particular quotient is rounded downwards or upwards. All that is needed is to compare the quotient's remainder to the split quoted. The party receives one of the residual seats if only if its remainder exceeds the split. Quoting a split circumvents the labor of ranking all parties by the size of their interim quotients' remainders when double-checking the seat numbers.

1.3. CYPRUS–CZECH REPUBLIC–GERMANY: TABLE DESIGN

Cyprus, referring its electoral threshold to valid votes, requires a peculiar level of 1.8 percent. Thus the 306 325 valid votes entail a threshold of 5 514 votes. Eight parties fail the threshold, and 9 770 valid votes become ineffective. The 296 555 effective votes are evaluated by means of a variant of the Hare-quota method with residual fit by greatest remainders. The variant is that the quota is based on the sum of all valid votes, not on the sum of all effective votes. Moreover, it is not precisely equal to the votes-per-seats ratio, but to the integer obtained from rounding the average downwards. In our classification we refer to this quota as the Hare-quota variant-3, HQ3. Here it amounts to $\lfloor 306\,325/6 \rfloor = \lfloor 51\,054.2 \rfloor = 51\,054$. The residual apportionment is a fit by greatest remainders. Since the rounding of the quotients comes out to be the same as with standard rounding, we prefer to quote the split .5. See Table 1.6.

The **Czech Republic** has an electoral threshold of five percent of valid votes. With 2 358 934 valid votes the threshold amounts to 117 947 votes. Twenty-nine parties miss it, and the 573 828 votes cast for them turn ineffective. The 22 Czech seats are apportioned according to the remaining 1 785 106 effective votes. The apportionment is carried out using the divisor method with downward rounding. See Table 1.7.

Germany employs an electoral threshold of five percent of valid votes. There are 26 333 444 valid votes, and so the threshold is 1 316 673 votes. Twenty-six parties fail it, and 2 840 893 votes are discarded as ineffective. Remarkably, somebody complained. In November 2011 the German Federal Constitutional Court ruled that the five percent threshold is unconstitutional. Of course, the court's opinion sets no binding precedent for other judiciaries nor for the Court of Justice of the European Union. The opinion does not mean that the Union's permission for domestic provisions to include a five percent threshold violates the Union's primary law. It only says is that it is unconstitutional to apply the Union's permission to the EP election in Germany.

EP2009DE	Political Group	Votes	Quotient	DivStd
CDU	EPP	8 071 391	34.3	34
SPD	S & D	5 472 566	23.3	23
GRÜNE	GREENS/EFA	3 194 509	13.6	14
FDP	ALDE	2 888 084	12.3	12
LINKE	GUE/NGL	1 969 239	8.4	8
CSU	EPP	1 896 762	8.1	8
Sum (Divisor)		23 492 551	(235 000)	99

District	Votes	Quotient	DivStd
Sub-apportionment to districts: CDU			
Schleswig-Holstein	308 368	1.3	1
Mecklenburg-Vorpommern	201 447	0.8	1
Hamburg	128 443	0.54	1
Niedersachsen	962 510	4.0	4
Bremen	45 886	0.2	0
Brandenburg	140 616	0.6	1
Sachsen-Anhalt	213 731	0.9	1
Berlin	208 395	0.9	1
Nordrhein-Westfalen	2 091 945	8.7	9
Sachsen	567 231	2.4	2
Hessen	596 878	2.49	2
Thüringen	304 858	1.3	1
Rheinland-Pfalz	660 252	2.8	3
Baden-Württemberg	1 478 135	6.2	6
Saarland	162 696	0.7	1
Sum (Divisor)	8 071 391	(240 000)	34

TABLE 1.8 *Germany, 2009 EP election.* The divisor method with standard rounding is used, DivStd, wherein quotients are rounded downwards when their fractional part is below one half, and upwards when above. Only the CDU with its fifteen district lists calls for a sub-apportionment.

Germany has 99 seats to apportion. All parties present their candidates on a single federal list, except for the CDU. The CDU submits fifteen state-lists, one for each state of the Federation where the party campaigns. As a consequence there are two calculatory stages. The first stage, the *super-apportionment*, evaluates the effective votes across all of Germany. The second stage is the CDU *sub-apportionment*. It apportions the 34 CDU seats to the fifteen CDU state-lists. Both stages use the divisor method with standard rounding, DivStd. A quotient is rounded upwards when its fractional part is larger than one-half, and downwards when it is smaller; if the fractional part were exactly equal to one-half then lots would be drawn. The divisor is adjusted so that all of the available seats are handed out. See Table 1.8.

Table design. Some general remarks on the design of tables may be in order. Usually, the column “Votes” lists the effective votes. In a rare case when it is valid votes that determine the electoral key, as in Cyprus, a line with the aggregate number of ineffective votes is adjoined. They do not otherwise participate in the calculations, and so dashes are entered into the “Quotient” and seats columns.

In the “Quotient” column, the number of decimal digits shown is contingent on the apportionment method used. For divisor methods, mostly a single digit suffices to see whether the rounding operation goes upwards or downwards. If not, then as many digits are exhibited as are needed for a clear decision. For quota methods with a residual fit by greatest remainders, the idea is to compare quotients’ fractional parts. Three digits usually suffice. The electoral key, whether a divisor or a split, is parenthetically recorded in the bottom line of a table.

EP2009DK	Political Group	Votes	Quotient	DivDwn
Alliance 1		975 136	6.1	6
Alliance 2		785 036	4.9	4
O	EFD	357 942	2.2	2
Alliance 3		224 014	1.4	1
Sum (Divisor)		2 342 128	(160 000)	13

Party	Political Group	Votes	Quotient	DivDwn
Alliance 1: Sub-apportionment				
A	S & D	503 439	4.03	4
F	GREENS/EFA	371 603	2.97	2
B	NA	100 094	0.8	0
Sum (Divisor)		975 136	(125 000)	6
Alliance 2: Sub-apportionment				
V	ALDE	474 041	3.2	3
C	EPP	297 199	1.98	1
I	NA	13 796	0.1	0
Sum (Divisor)		785 036	(150 000)	4
Alliance 3: Sub-apportionment				
N	GUE/NGL	168 555	1.7	1
J	NA	55 459	0.6	0
Sum (Divisor)		224 014	(100 000)	1

TABLE 1.9 *Denmark, 2009 EP election.* Three alliances are registered. The state-wide super-apportionment treats each of them as a virtual entity. It is only in the sub-apportionment that the share of seats of an alliance is apportioned among its partners.

The tables rank parties by decreasing voter support in the current election. Electoral bureaus sometimes rely on the party ranking of the previous legislative period. Formerly this helped to prepare the record sheets for the upcoming election. Nowadays modern computing equipment makes it easy to sort parties by their current vote counts. Moreover, we always include a final "Sum" line showing column sums. The sums are informative by themselves, such as the number of seats to be apportioned. Column sums provide a helpful and simple check on whether column entries got corrupted by copying, pasting, or any other editorial operation.

Multiple boxes in the tables point to multiple calculations. For example, the two boxes in Table 1.8 show the super-apportionment in the whole country, and the districtwise sub-apportionment for the CDU. The three boxes in Table 1.4 exhibit the apportionments in the three Belgian districts. The nature of a box is context-dependent. In the German example, the super-apportionment needs to be completed before its results can be handed down to the sub-apportionment. In the Belgian example, the three calculations are independent of one another.

1.4. DENMARK–ESTONIA–GREECE: ALLIANCES AND INDEPS

Denmark introduces a new twist into the exposition: *alliances*, also known as electoral cartels, or as list apparentements. Alliances 1 and 2 comprise three parties each, and Alliance 3 has two. Every alliance induces multiple stages into the apportionment calculations. In the first stage, that again is called the super-apportionment, an alliance is treated as a virtual unity that is allocated its due share of seats. The second stage consists of the sub-apportionment of the alliance's seats among its partners. In the

EP2009EE	Political Group	Votes	Quotient	DivDwn
KE	ALDE	103 506	2.9997	2
Indep Indrek Tarand	GREENS/EFA	102 460	2.97•	1
ER	ALDE	60 877	1.8	1
IRL	EPP	48 492	1.4	1
SDE	S & D	34 508	1.0001	1
ERR	GREENS/EFA	10 851	0.3	0
ERL	ECR	8 860	0.3	0
EÜUP	GREENS/EFA	3 519	0.1	0
LEE	EFD	2 206	0.1	0
8 Others	NA	21 703	—	0
Sum (Divisor)		396 982	(34 505)	6

TABLE 1.10 *Estonia, 2009 EP election.* The independent candidate—Indep, for short—Indrek Tarand draws so many votes that the quotient 2.97 would have justified two seats. However, an indep can occupy at most one seat (•).

Danish election three sub-apportionment calculations are called for, one for each of the three alliances. See Table 1.9.

Estonia apportions its six seats on the basis of 396 982 valid votes. There is no electoral threshold. The apportionment uses the divisor method with downward rounding. It turns out that 34 505 votes justify roughly one seat. The Estonian election features an *indep*, an independent candidate who stands in the election with no party affiliation nor party support. See Table 1.10.

Greece employs a unique quota method, HQ3-EL, unrivaled for its complexity. An electoral threshold of three percent of valid votes applies. With 5 127 537 valid votes altogether, the threshold amounts to 153 827 votes. Twenty-one parties fail it, and the 377 997 votes for them are discarded as ineffective. Six parties pass the threshold, and share the Greek allocation of 22 seats. The calculations are carried out in three stages: the main apportionment, the initial residual apportionment, and the final residual apportionment. The main apportionment is based on the Hare-quota variant-3, HQ3, as in Cyprus. In the Greek election, HQ3 equals $\lfloor 5\,127\,537/22 \rfloor = 233\,069$ votes per seat. Table 1.11 exhibits the resulting quotients and the ensuing main apportionment of 18 seats.

The remaining four seats are taken care of in the two residual apportionments. They refer to *Unused Voting Power*, UVP. It is determined as follows. In the main apportionment the strongest party is allocated eight seats and thus uses up $8 \times 233\,069 = 1\,864\,552$ of its power of 1 878 982 votes, leaving an UVP of 14 430 votes. The unused voting powers of the other parties are found similarly. The aggregate UVP is taken to be the sum of the parties' unused voting powers plus the number of ineffective votes, 932 295. Now a new quota is introduced with reference to the aggregate UVP. Applied to 4 seats the quota is $\text{DQ5} = \lfloor 932\,295/(4+1) \rfloor = 186\,459$, called Droop-quota variant-5 in Section 1.6. The resulting quotients of UVP per DQ5 are shown in the penultimate column of Table 1.11. They are used to apportion the four residual seats. In the initial residual apportionment, every party with a quota exceeding unity gets a seat and drops out from further consideration (here just K.K.E.). Three seats and five parties make it into the final stage. The seats left are allocated to the parties left according to the greatest remainders of their quotients.

EP2009EL	Political Group	Votes	Quotient-1	Main	UVP	Quotient-2	HQ3-EL
Pa.So.K	S & D	1 878 982	8.062	8	14 430	0.077	8
N.D.	EPP	1 655 722	7.104	7	24 239	0.130	8
K.K.E.	GUE/NGL	428 282	1.838	1	195 213	1.–	2
LA.O.S	EFD	366 637	1.573	1	133 568	0.716	2
SY.RIZ.A	GUE/NGL	240 930	1.034	1	7 861	0.042	1
OP	GREENS/EFA	178 987	0.768	0	178 987	0.960	1
Ineffective votes		377 997	—	—	377 997	—	—
Sum (Quotas HQ3, DQ5)		5 127 537	(233 069)	18	932 295	(186 459)	22

TABLE 1.11 *Greece, 2009 EP election.* Greece applied the method HQ3-EL, unrivaled for its complexity. The main apportionment leaves Unused Voting Power as listed in column "UVP". It is assessed using the Droop-quota variant-5, DQ5 = 186 459.

EP2009ES	Political Group	Votes	Quotient	DivDwn
PP	EPP	6 670 377	23.8	23
PSOE	S & D	6 141 784	21.9	21
CpE	ALDE	808 246	2.9	2
IU-ICV-EU/IA-BA	GUE/NGL, GREENS/EFA	588 248	2.1	2
UPyD	NA	451 866	1.6	1
EdP-V	GREENS/EFA	394 938	1.4	1
II	NA	178 121	0.6	0
28 Others, each below 90 000 votes		381 716	—	0
Sum (Divisor)		15 615 296	(280 000)	50

TABLE 1.12 *Spain, 2009 EP election.* The divisor method with downward rounding is used. The two representatives of the fourth-strongest party join two different Political Groups, GUE/NGL and GREENS/EFA.

Alliances and indeps. Alliances are peculiar dispositions that came into being as a consequence of the divisor method with downward rounding. The method is notorious for awarding stronger parties an overproportional share of seats at the expense of weaker parties. Hence it is desirable for weaker parties to become stronger by joining together. This is what alliances are meant to achieve. However, once an electoral law allows formation of alliances, stronger parties may also join into an alliance and thus grow stronger yet. We shall see that alliances fail to serve their purpose of neutralizing the bias that marks the divisor method with downward rounding.

The term "indep" for an independent, individual candidate is the sole neologism in this book, a kind of antipode to an alliance. An alliance gathers a large ensemble of nominees who make do with a small common denominator politically. An indep boasts a maximum of individuality. Since an individual cannot fill more than one seat, the apportionment calculation imposes a maximum restriction of one seat per indep. In case the restriction is active we earmark the relegated interim quotient by a trailing dot •. The indep Indrek Tarand in Table 1.10 provides an example.

1.5. SPAIN–FINLAND–FRANCE: VOTE CATEGORIES

Spain aggregates the votes across all of its fifty provinces. No threshold applies, and the divisor method with downward rounding is used to apportion 50 seats. Every 280 000 votes justify roughly one seat. The votes for the fourth-strongest party justify two Members of the EP who, however, join distinct Political Groups. See Table 1.12.

EP2009FI	Political Group	Votes	Quotient	DivDwn
KOK	EPP	386 416	3.9	3
KESK	ALDE	316 798	3.2	3
SDP	S & D	292 051	2.9	2
Alliance 1		232 388	2.3	2
VIHR	GREENS/EFA	206 439	2.1	2
SFP(RKP)	ALDE	101 453	1.01	1
VAS	GUE/NGL	98 690	0.99	0
6 Others	NA	30 596	—	0
Sum (Divisor)		1 664 831	(100 000)	13

Party	Political Group	Votes	Plurality
Alliance 1: Sub-apportionment			
PS	EFD	162 930	1
KD	EPP	69 458	1
Sum		232 388	2

TABLE 1.13 *Finland, 2009 EP election.* The super-apportionment uses the divisor method with downward rounding. In the sub-apportionment to Alliance 1 the two seats go to the candidates with the most personal votes. One of them joined the EFD group, the other, EPP.

Finland features an alliance of two parties. There is no electoral threshold. The super-apportionment uses the divisor method with downward rounding. In the sub-apportionment for the alliance, the two seats are allotted by plurality. The candidates with the most votes receive the seats. The PS top-runner draws 130 715 votes, followed by others with fewer than 10 000 votes each. The strongest KD candidate has 53 803 votes. Hence the first seat falls to PS, the second to KD. See Table 1.13.

France subdivides the country into eight districts. The 72 French seats are assigned to the districts well ahead of the election, with district magnitudes 10, 9, 9, 10, 13, 5, 13, and 3. The electoral threshold is set at five percent of valid votes, *voix exprimées*, and pertains to each district separately. In all districts the seat apportionment is carried out using the divisor method with downward rounding, DivDwn. The results are displayed in Table 1.14.

The apportionment of the three seats of District 8: *Outre-Mer* has to follow certain rules to secure a fair geographical representation. The district is subdivided into three sections, Atlantique, Océan Indien, and Pacifique. Each section is guaranteed representation in the EP. To this end the candidate lists of the parties are obliged to include at least one nominee from each section. The seats allocated to the strongest party are filled with the nominees from the sections where the strongest party performs best. The seat of the second-strongest party goes to that section among the remaining sections where the second-strongest party scores best. In case the third-strongest party gets a seat it is allocated to the section remaining.

Vote categories. An electoral system generally operates with various categories of voters, and of votes. The all-embracing reference set is the entire *citizenry*. The citizens who have the franchise to vote form the *electorate*. Those of the electorate who go to the polls are the *voters*. As the term "electorate" is a singular, many experts also refer to a singular, "the voter", when they really mean many people who vote. We find the typifying singular misleading. If there were only one voter we would not have to deal with numbers.

EP2009FR Party	Political Group	Votes	Quotient	Div-Dwn
District 1: Nord-Ouest				
UMP	EPP	601 556	4.002	4
PS	S & D	449 533	2.99	2
EuÉco	GREENS/EFA	300 579	1.9999	1
FN	NA	253 009	1.7	1
MoDem	ALDE	215 482	1.4	1
FG	GUE/NGL	169 813	1.1	1
NPA	NA	143 967	0.96	0
Sum (Divisor)		2 133 939	(150 300)	10
District 2: Ouest				
UMP	EPP	680 829	3.4	3
PS	S & D	433 309	2.2	2
EuÉco	GREENS/EFA	417 449	2.1	2
Lib.	EFD	257 437	1.3	1
MoDem	ALDE	212 524	1.1	1
NPA	NA	128 641	0.6	0
Sum (Divisor)		2 130 189	(200 000)	9
District 3: Est				
UMP	EPP	635 016	4.04	4
PS	S & D	374 971	2.4	2
EuÉco	GREENS/EFA	310 620	1.98	1
MoDem	ALDE	205 256	1.3	1
FN	NA	164 672	1.05	1
NPA	NA	122 767	0.8	0
Sum (Divisor)		1 813 302	(157 000)	9
District 4: Sud-Ouest				
UMP	EPP	705 900	4.2	4
PS	S & D	465 076	2.7	2
EuÉco	GREENS/EFA	415 457	2.4	2
MoDem	ALDE	225 917	1.3	1
FG	GUE/NGL	214 079	1.3	1
FN	NA	155 806	0.9	0
NPA	NA	147 422	0.9	0
Sum (Divisor)		2 329 657	(170 000)	10
District 5: Sud-Est				
UMP	EPP	862 556	5.4	5
EuÉco	GREENS/EFA	537 151	3.4	3
PS	S & D	426 043	2.7	2
FN	NA	249 695	1.6	1
MoDem	ALDE	216 630	1.4	1
FG	GUE/NGL	173 576	1.1	1
Sum (Divisor)		2 465 651	(160 000)	13
District 6: Massif-Central / Centre				
UMP	EPP	382 632	3.2	3
PS	S & D	238 806	1.99	1
EuÉco	GREENS/EFA	182 311	1.5	1
MoDem	ALDE	109 369	0.9	0
FG	GUE/NGL	108 194	0.9	0
NPA	NA	73 162	0.6	0
FN	NA	68 665	0.6	0
Sum (Divisor)		1 163 139	(120 000)	5
District 7: Ile-de-France				
UMP	EPP	828 172	5.9	5
EuÉco	GREENS/EFA	583 690	4.2	4
PS	S & D	379 908	2.7	2
MoDem	ALDE	238 341	1.7	1
FG	GUE/NGL	176 862	1.3	1
Sum (Divisor)		2 206 973	(140 000)	13
District 8: Outre-Mer				
UMP	EPP	103 247	1.7	1
AOM	GUE/NGL	73 110	1.2	1
PS	S & D	70 514	1.2	1
EuÉco	GREENS/EFA	56 502	0.9	0
MoDem	ALDE	32 322	0.5	0
Sum (Divisor)		335 695	(60 000)	3

TABLE 1.14 *France, 2009 EP election.* France establishes eight districts. Each district uses the divisor method with downward rounding, with geographical restrictions on District 8: *Outre-Mer.*

The *votes cast* are categorized into *valid votes* or good votes, versus *invalid votes* or rejected votes. The valid votes subdivide into the *effective votes* that enter into the apportionment calculations, and the *ineffective votes* that though valid are nevertheless discarded. For the purpose of analyzing the diversity of apportionment methods we mostly rely on effective votes, and refer to valid votes or votes cast only occasionally. Weak parties not getting a seat nor affiliated with a Political Group are aggregated into the category "Others".

Ballot design in the 27 Member States is by no means uniform. In most states voters cast their votes for party lists. Before the election every party publicizes the list of their nominees. Later, any party seats are filled in the sequence of this list. *Closed lists*, with a definite sequence of nominees, are practiced in Germany. *Open lists*, where the terminal ranking relies on the personal votes for the nominees, are employed in Finland. Passing over these subtleties we assume that the vote counts as reported by the domestic electoral bureaus are all of a comparable quality.

EP2009HU	Political Group	Votes	Quotient	DivDwn
FIDESZ-KDNP	EPP	1 632 309	14.8	14
MSZP	S & D	503 140	4.6	4
JOBBIK	NA	427 773	3.9	3
MDF	ECR	153 660	1.4	1
Sum (Divisor)		2 716 882	(110 000)	22

TABLE 1.15 *Hungary, 2009 EP election.* The threshold of five percent of valid votes shuts out four parties and turns 179 297 votes ineffective. For the four remaining parties, every 110 000 votes justify roughly one seat.

1.6. HUNGARY–IRELAND–ITALY: QUOTAS

Hungary is assigned 22 seats. There is an electoral threshold of five percent of valid votes. Four parties miss the threshold, and 179 297 votes become ineffective. The seat apportionment is carried out by means of the divisor method with downward rounding. Four parties take part in the apportionment, with every 110 000 votes justifying roughly one seat. See Table 1.15.

Ireland is assigned 12 seats. There is no electoral threshold. The country is divided into four separately evaluated districts with three seats each: Dublin, East, North-West, and South. In each district the Droop-quota DrQ $= \lfloor v_+/(3+1) \rfloor + 1$ is calculated, where v_+ designates the valid vote total in that district. A single transferable vote (STV) scheme is used. In brief it works as follows. On the ballot sheet voters mark the candidates as first preference, second preference, and so on downwards. When the first-preference tally for a candidate reaches the Droop-quota, the candidate is declared elected. The ballots in excess of the Droop-quota are transferred to the candidate who is next according to the voter's preferences. The decision which ballots are in excess is random, as they happen to be filed in their pile. We label the system with the acronym STVran. See Table 1.16.

An effect of the accumulation of lower-order preferences is seen in District 1: *Dublin*. Fourth-ranked socialist Joe Higgins has fewer first preferences than third ranked Fianna Fáil candidate Eoin Ryan, but is elected by means of the additional votes transferred to him during the evaluation process. Another effect emerges in District 4: *South*. The data given in Table 1.16 are *not* sufficient to reconstruct these vote transfers and to double-check the final seat apportionment. The vote transfer process is time-consuming. Evaluation of an STV scheme —extends over several days, even in a smaller state such as Ireland. The evaluation with an apportionment method is finished within a few hours.

Generally, STV schemes are considered to be viable proportional representation systems. At first glance an STV scheme links those voting and those being elected in a more direct way. Mediating parties seem to be circumvented. At a second glance it transpires that the differences are minute. Consider an alternative evaluation whereby the divisor method with standard rounding apportions the district's seats in proportion to the parties' aggregated first-preference scores, and seats are filled with the parties' candidates who did best. With the data in Table 1.16 voters in Districts 2 and 3 elect the same candidates. In Districts 1 and 4 only the third seat is allocated differently. In District 1 the pooled first preferences for the two Fianna Fáil nominees would secure a

EP2009IE	Party	Political Group	1st Pref	STVran
District 1: Dublin				
Gay Mitchell	FG	EPP	96 715	1
Proinsias de Rossa	Lab.	S & D	83 471	1
Eoin Ryan Jnr	FF	ALDE	55 346	0
Joe Higgins	SP	GUE/NGL	50 510	1
Mary Lou McDonald	SF	GUE/NGL	47 928	0
Deirdre de Burca	Green/Comhaontas Glas	GREENS/EFA	19 086	0
Eibhlin Byrne	FF	ALDE	18 956	0
Caroline Simons	Libertas	EFD	13 514	0
2 Others	Indeps	NA	21 104	0
Sum (Droop-quota)		(101 658)	406 630	3
District 2: East				
Mairead McGuinness	FG	EPP	110 366	1
Nessa Childers	Lab.	S & D	78 338	1
Liam Aylward	FF	ALDE	74 666	1
John Paul Phelan	FG	EPP	61 851	0
Thomas Byrne	FF	ALDE	31 112	0
Kathleen Funchion	SF	GUE/NGL	26 567	0
Tomas Sharkey	SF	GUE/NGL	20 932	0
Ray O'Malley	Libertas	EFD	18 557	0
3 Others	Indeps	NA	6 860	0
Sum (Droop-quota)		(107 313)	429 249	3
District 3: North-West				
Marian Harkin	Indeps	ALDE	84 813	1
Pat Gallagher	FF	ALDE	82 643	1
Jim Higgins	FG	EPP	80 093	1
Declan Ganley	Libertas	EFD	67 638	0
Padraig MacLochlainn	SF	GUE/NGL	45 515	0
Paschal Mooney	FF	ALDE	42 985	0
Joe O'Reilly	FG	EPP	37 564	0
Susan O'Keeffe	Lab.	S & D	28 708	0
5 Others	Indeps	NA	25 348	0
Sum (Droop-quota)		(123 827)	495 307	3
District 4: South				
Brian Crowley	FF	ALDE	118 258	1
Sean Kelly	FG	EPP	92 579	1
Toireasa Ferris	SF	GUE/NGL	64 671	0
Alan Kelly	Lab.	S & D	64 152	1
Colm Burke	FG	EPP	53 721	0
Ned O'Keeffe	FF	ALDE	16 596	0
Dan Boyle	Green/Comhaontas Glas	GREENS/EFA	15 499	0
3 Others	Indeps	NA	72 651	0
Sum (Droop-quota)		(124 532)	498 127	3

TABLE 1.16 *Ireland, 2009 EP election.* Ireland uses the Single Transferable Vote scheme. The decision which votes for a candidate exceed the Droop-quota and are transferred to lower preferences is random (STVran). By and large party seats are in proportion to party votes.

seat for Eoin Ryan. In District 4 the 64 671 first preferences for Toireasa Ferris would win over the 64 152 votes for Alan Kelly.

Italy is assigned 72 seats. A threshold of four percent of valid votes applies, except for a minority protection clause. A minority party gets a seat if it registers an alliance with a party standing in all five districts, and if one of its candidates draws at least 50 000 votes. In 2009 the Südtiroler Volkspartei (SVP) is allied with Partito democratico, the Vallee d'Aoste party with Il Popolo della libertà, and the party Autonomie liberté et démocratie with Di Pietro Italia dei Valori. The protection clause secures a seat for Herbert Dorfmann (SVP) who received 84 361 votes. The

EP2009IT	Political Group	Votes	Quotient	HQ1grR
PdL+VA	EPP	10 828 525	29.348	29
PD+SVP	S & D	8 140 766	22.063	22
LN	EFD	3 125 418	8.471	9
IdV+Ald	ALDE	2 476 695	6.712	7
UDC	EPP	1 994 813	5.406	5
Sum (Split)		26 566 217	(.44)	72

District	Votes	Quotient	HQ1grR
Sub-apportionment to districts: PdL+VA			
Nord-Occidentale+VA	2 935 126	7.861	8
Nord-Orientale	1 777 869	4.761	5
Italia Centrale	2 344 306	6.278	6
Italia Meridionale	2 869 765	7.686	8
Italia Insulare	901 459	2.414	2
Sum (Split)	10 828 525	(.5)	29
Sub-apportionment to districts: PD+SVP			
Nord-Occidentale	2 002 790	5.412	5
Nord-Orientale+SVP	1 915 846	5.177	5
Italia Centrale	2 030 062	5.486	6
Italia Meridionale	1 575 928	4.259	4
Italia Insulare	616 140	1.665	2
Sum (Split)	8 140 766	(.45)	22
Sub-apportionment to districts: LN			
Nord-Occidentale	1 684 842	4.852	5
Nord-Orientale	1 204 785	3.469	3
Italia Centrale	186 988	0.538	1
Italia Meridionale	39 521	0.114	0
Italia Insulare	9 282	0.027	0
Sum (Split)	3 125 418	(.5)	9
Sub-apportionment to districts: IdV+Ald			
Nord-Occidentale+Ald	663 495	1.875	2
Nord-Orientale	454 801	1.285	1
Italia Centrale	483 471	1.366	1
Italia Meridionale	688 368	1.946	2
Italia Insulare	186 560	0.527	1
Sum (Split)	2 476 695	(.5)	7
Sub-apportionment to districts: UDC			
Nord-Occidentale	460 487	1.154	1
Nord-Orientale	353 714	0.887	1
Italia Centrale	341 612	0.856	1
Italia Meridionale	582 421	1.460	1
Italia Insulare	256 579	0.643	1
Sum (Split)	1 994 813	(.5)	5

TABLE 1.17 *Italy, 2009 EP election.* The state-wide super-apportionment is followed by a sub-apportionment for each party. The resulting sums of seats per district fail to meet the district magnitudes that are guaranteed in the electoral law.

overall threshold amounts to 1 224 615 votes. Eight parties fail the threshold, discarding 4 049 147 votes as ineffective. Five parties with three allied minority parties pass it.

Italy forms five districts. The apportionment takes place in two stages. The first stage is the super-apportionment of all 72 seats among the five parties and their alliance partners. The second stage consists of five sub-apportionments to allocate the state-wide seats of a party to the five districts. All calculations use the Hare-quota variant-1 method with residual fit by greatest remainders, HQ1grR. See Table 1.17.

EP2009LT	Political Group	Votes	Quotient	HQ2gR1
TS-LKO	EPP	147 756	3.918	4
LSDP	S & D	102 347	2.714	3
TT	EFD	67 237	1.783	2
DP	ALDE	48 368	1.283	1
LLRA(AWPL)	ECR	46 293	1.228	1
ULRls	ALDE	40 502	1.074	1
Sum (Split)		452 503	(.5)	12

TABLE 1.18 *Lithuania, 2009 EP election.* The quota method HQ2gR1 is used. Participation in the residual apportionment gR1 requires HQ2 = 37 709 votes, that is, close to seven percent of the 564 803 votes cast.

The Italian electoral system does not agree with the Italian electoral law. The law guarantees every district a certain number of seats based on census figures. The 2009 EP election misses the guaranteed district magnitudes in each case. Nord-Occidentale is guaranteed 19 seats but gets 21; Nord-Orientale is guaranteed 13 but gets 15; Italia Centrale is guaranteed 14 but gets 15; Italia Meridionale is guaranteed 18 but gets 15; and Italia Insulare is guaranteed 8 but gets 6. The deficiencies were pointed out to the public by academics, yet nobody filed a complaint in court. No complaint, no redress.

Quotas. The number of seats to be apportioned is denoted by h, the house size. Here is an overview over various quota definitions:

$$\mathrm{HaQ} = \frac{\text{effective votes}}{h}, \qquad \mathrm{DrQ} = \left\lfloor \frac{\text{effective votes}}{h+1} \right\rfloor + 1,$$

$$\mathrm{HQ1} = \left\lfloor \frac{\text{effective votes}}{h} \right\rfloor \vee 1, \qquad \mathrm{DQ1} = \left\lfloor \frac{\text{effective votes}}{h+1} \right\rfloor \vee 1,$$

$$\mathrm{HQ2} = \left\lceil \frac{\text{effective votes}}{h} \right\rceil, \qquad \mathrm{DQ2} = \left\lceil \frac{\text{effective votes}}{h+1} \right\rceil,$$

$$\mathrm{HQ3} = \left\lfloor \frac{\text{valid votes}}{h} \right\rfloor \vee 1, \qquad \mathrm{DQ3} = \left\langle \frac{\text{effective votes}}{h+1} \right\rangle \vee 1,$$

$$\mathrm{HQ4} = \left\lceil \frac{\text{valid votes}}{h} \right\rceil, \qquad \mathrm{DQ5} = \left\lfloor \frac{\text{unused voting power}}{r+1} \right\rfloor \vee 1.$$

The missing link, the unrounded Droop-quota DQ4, is adjoined in Section 5.8. Quotas, being employed as fixed divisors, must not be zero. When the rounding may yield zero the definition forces the quota to stay positive, by setting $t \vee 1 := \max\{t, 1\} \geq 1$; these cases are of no practical interest. Variants 1 and 2 of the Droop-quota were applied in the Swiss Canton Solothurn in 1896–1977 and 1981–1993, but are absent in the 2009 EP elections.

As an example we consider District 4: *South* in Table 1.16. With 498 127 effective votes the default Hare-quota amounts to 166 042.3 vote fractions. Its variant-1 rounds the number downwards, 166 042, variant-2 upwards, 166 043. The default Droop-quota divides $3 + 1 = 4$ into the vote count, then rounds the quotient 124 531.75 downwards, and finally adds 1, resulting in 124 532 votes. Variant-1 of the Droop-quota equals 124 531 votes, variant-3 uses commercial rounding and yields 124 532. *Commercial rounding* rounds downwards when the first digit after the decimal point is 0, 1, 2, 3, or 4; it rounds upwards when the first digit is 5, 6, 7, 8, or 9.

EP2009LU	Political Group	Votes	Quotient	DivDwn
CSV	EPP	353 094	3.2	3
LSAP	S & D	219 349	1.99	1
DP	ALDE	210 107	1.9	1
déi gréng	GREENS/EFA	189 523	1.7	1
4 Others	NA	153 959	—	0
Sum (Divisor)		1 126 032	(110 000)	6

TABLE 1.19 *Luxembourg, 2009 EP election.* Every voter has up to six votes. On average, a ballot sheet contains 5.7 votes. The numbers shown ought to be scaled by 5.7 in order to appreciate the size of the Luxembourg electorate.

1.7. LITHUANIA–LUXEMBOURG–LATVIA: RESIDUAL FITS

Lithuania implements three rules for the electoral threshold. First, an explicit threshold is set at five percent of the 564 803 votes cast, 28 241 votes. Nine parties fail the threshold, and the 97 514 votes for them turn ineffective. Second, a salvation clause says that, if need be, the percentage is lowered so that at least sixty percent of the valid votes become effective. In the 2009 election 452 503 of the 550 017 valid votes are effective, more than eighty percent. Hence the second rule is dormant. The 12 Lithuanian seats are apportioned using the method HQ2gR1. It combines the Hare-quota variant-2, HQ2 $= \lceil 452\,503/12 \rceil = 37\,709$, with a *full-seat restricted* residual fit by greatest remainders, gR1. That is, parties with a quotient smaller than a full quota of votes are excluded from the residual apportionment. In Lithuania, gR1 excludes parties with fewer than 37 709 votes, thus embodying a third threshold. The third threshold exceeds the five percent threshold. Hence one may wonder whether it is permissible under the Union's norm that the threshold may not exceed five percent of the votes cast. But nobody complained. No complaint, no redress. See Table 1.18.

Luxembourg gives every voter six votes. The six votes may be distributed across party lines. Two of the six votes may be accumulated on a single candidate. There are 1 126 032 votes marked on the 198 364 valid ballot sheets. On average, there are 5.7 marks per ballot. The vote numbers would need to be scaled by the marks-per-ballot average in order to mirror the number of people who back the parties. There is no electoral threshold. The seat apportionment is carried out by means of the divisor method with downward rounding, DivDwn. See Table 1.19.

Latvia has a threshold of five percent of the 791 597 votes cast. Eleven parties fail the threshold of 39 580 votes, whence of the 777 084 valid votes 182 144 are discarded as ineffective. The eight Latvian seats are apportioned among the six parties that pass the threshold using the divisor method with standard rounding, DivStd. One hundred thousand votes constitute a feasible electoral key. See Table 1.20.

Residual fits. The Lithuanian full-seat restricted residual fit by greatest remainders, gR1, must be seen in connection with the rationale that underlies quota methods. The seats allocated in the main apportionment are understood to be fully justified, because each of these seats is backed by a full quota of citizens who voted in their favor. In contrast, the residual seats are no longer justified by a full quota of voters, and are awarded only to parties deserving them. The argument is felt to provide a sufficient reasoning to exclude parties supported by so few voters that the one and only seat of the party would be a remainder seat.

EP2009LV	Political Group	Votes	Quotient	DivStd
PS	EPP	192 537	1.9	2
SC	S & D, GUE/NGL	154 894	1.55	2
PCTVL	GREENS/EFA	76 436	0.8	1
LPP/LC	ALDE	59 326	0.6	1
TB/LNNK	ECR	58 991	0.6	1
JL	EPP	52 751	0.53	1
Sum (Divisor)		594 935	(100 000)	8

TABLE 1.20 *Latvia, 2009 EP election.* Latvia uses the divisor method with standard rounding, DivStd. It so happens that every one hundred thousand votes justify roughly one seat. The apportionment of the eight seats becomes as transparent as could possibly be.

EP2009MT	Party	Political Group	1st Pref	STVran
Simon Busuttil	PN	EPP	68 782	1
Louis Grech	PL(MLP)	S & D	27 753	1
Edward Scicluna	PL(MLP)	S & D	24 574	1
Joseph Cuschieri	PL(MLP)	S & D	19 672	0
Marlene Mizzi	PL(MLP)	S & D	17 724	0
John Montalto Attard	PL(MLP)	S & D	12 880	1
Baldacchino Abela	PL(MLP)	S & D	12 309	0
David Casa	PN	EPP	6 539	1
26 further nominees	—	NA	57 936	—
Sum (Droop-quota)		(41 362)	248 169	5

TABLE 1.21 *Malta, 2009 EP election.* Malta uses the Single Transferable Vote scheme. The decision which votes exceed a candidate's Droop-quota and are transferred to lower preferences is random (STVran). By and large party seats are in proportion to party votes.

Yet another residual apportionment procedure was in operation in former times. During 1896–1917 the Swiss Canton Solothurn allocated all remaining seats to the strongest party. We abbreviate this *winner-take-all* directive by WTA. In summary, here is an overview of the variety of residual apportionment procedures used in the 2009 EP elections and elsewhere:

- grR The remaining seats are allocated, one by one via greatest remainders, among all eligible parties.
- gR1 The remaining seats are allocated, one by one via greatest remainders, among those parties drawing at least one full quota of votes.
- WTA Winner-take-all: All remaining seats go to the strongest party.
- -EL The remaining seats are allocated as in the Greek EP election 2009.

1.8. MALTA–NETHERLANDS–POLAND: NESTED STAGES

Malta bases its election on the same single transferable vote scheme that is used in Ireland, STVran. If first preferences are pooled per party, hypothetically as in Section 1.6 for Ireland, the Partit Laburista would have been apportioned three seats, the Partit Nazzjonalista two. If these seats are filled with the most successful candidates then just a single seat would have been allocated differently, to Joseph Cuschieri rather than to John Montalto Attard. As in Ireland, the STV result in Malta can be regarded as coming close to what a divisor or quota method would produce. See Table 1.21.

EP2009NL	Political Group	Votes	Quotient	DivDwn
Alliance 1		1 223 773	7.7	7
Alliance 2		1 034 065	6.6	6
Alliance 3		952 711	6.03	6
PV	NA	772 746	4.9	4
SP	GUE/NGL	323 269	2.05	2
PD	NA	157 735	0.998	0
Libertas	EFD	14 612	0.1	0
De Groenen	GREENS/EFA	8 517	0.1	0
6 Others	NA	66 436	—	0
Sum (Divisor)		4 553 864	(158 000)	25

Party	Political Group	Votes	Quotient	HaQgrR
Alliance 1: Sub-apportionment				
CDA	EPP	913 233	5.224	5
CU-SGP	ECR,EFD	310 540	1.776	2
Sum (Split)		1 223 773	(.5)	7
Alliance 2: Sub-apportionment				
VVD	ALDE	518 643	3.009	3
D66	ALDE	515 422	2.991	3
Sum (Split)		1 034 065	(.5)	6
Alliance 3: Sub-apportionment				
PvdA	S & D	548 691	3.456	3
GL	GREENS/EFA	404 020	2.544	3
Sum (Split)		952 711	(.5)	6

TABLE 1.22 *Netherlands, 2009 EP election.* The super-apportionment uses the divisor method with downward rounding, the sub-apportionments for the three alliances the Hare-quota method with residual fit by greatest remainders.

The Netherlands has no electoral threshold. Three alliances are registered, of two parties each. This calls for a super-apportionment, followed by three sub-apportionments. The super-apportionment uses the divisor method with downward rounding, DivDwn. The sub-apportionments apply the Hare-quota method with residual fit by greatest remainders, HaQgrR. See Table 1.22.

Poland also evaluates the election in two stages, but for the reason of securing regional representation. The electoral threshold amounts to 368 239 votes, five percent of the 7 364 763 valid votes. Eight parties miss it, and 650 393 votes turn ineffective. Four parties pass the threshold and become eligible to participate in the apportionment calculation. In the first stage all 50 Polish seats are apportioned proportionately to the parties' state-wide vote counts. The calculation uses the divisor method with downward rounding, DivDwn. The second stage consists of the per-party sub-apportionments, and allocates a party's seats to the district lists in proportion to the votes in that district. All sub-apportionments use the Hare-quota method with residual fit by greatest remainders, HaQgrR. See Table 1.23.

Nested stages. The selective use of apportionment procedures is due to the fact that some methods are notorious for being biased in favor of stronger parties at the expense of weaker parties, such as DivDwn and DrQgrR. Other methods are known to be unbiased, such as DivStd and HaQgrR. Weaker parties might be reluctant to join an alliance that would expose them to a disadvantageous bias. Hence an alliance becomes more attractive when it promises its partners an unbiased procedure.

EP2009PL	Political Group	Votes	Quotient	DivDwn
PO	EPP	3 271 852	25.3	25
PiS	ECR	2 017 607	15.6	15
SLD-UP	S & D	908 765	7.02	7
PSL	EPP	516 146	3.99	3
Sum (Divisor)		6 714 370	(129 400)	50

District	Votes	Quotient	HaQgrR	Votes	Quotient	HaQgrR
Sub-apportionment to districts:		*PO*			*PiS*	
Gdańsk	285 268	2.180	2	105 946	0.788	1
Bydgoszcz	162 556	1.242	1	73 183	0.544	1
Olsztyn	159 943	1.222	1	121 921	0.906	1
Warszawa 1	434 421	3.319	3	196 720	1.463	1
Warszawa 2	114 000	0.871	1	129 165	0.960	1
Łódź	204 798	1.565	2	134 947	1.003	1
Poznań	289 442	2.212	2	121 216	0.901	1
Lublin	112 221	0.857	1	136 986	1.018	1
Rzeszów	107 092	0.818	1	153 661	1.142	1
Kraków	327 854	2.505	2	383 631	2.852	3
Katowice	523 602	4.001	4	207 429	1.542	1
Wrocław	347 617	2.656	3	163 197	1.213	1
Gorzów Wielkopolski	203 038	1.551	2	89 605	0.666	1
Sum (Split)	3 271 852	(.53)	25	2 017 607	(.543)	15
Sub-apportionment to districts:		*SLD-UP*			*PSL*	
Gdańsk	50 427	0.388	0	13 170	0.077	0
Bydgoszcz	79 400	0.612	1	38 092	0.221	0
Olsztyn	59 194	0.456	0	38 012	0.221	0
Warszawa 1	84 740	0.653	1	22 899	0.133	0
Warszawa 2	30 225	0.233	0	72 551	0.422	1
Łódź	62 923	0.485	0	32 390	0.188	0
Poznań	94 180	0.725	1	52 716	0.306	1
Lublin	24 725	0.190	0	51 954	0.302	0
Rzeszów	27 147	0.209	0	45 685	0.266	0
Kraków	95 277	0.734	1	60 846	0.354	1
Katowice	117 884	0.908	1	23 566	0.137	0
Wrocław	93 172	0.718	1	41 975	0.244	0
Gorzów Wielkopolski	89 471	0.689	1	22 290	0.130	0
Sum (Split)	908 765	(.5)	7	516 146	(.304)	3

TABLE 1.23 *Poland, 2009 EP election.* The super-apportionment uses the divisor method with downward rounding, the sub-apportionments the Hare-quota method with residual fit by greatest remainders.

Alliances lose their appeal, though, when seen with the eyes of the voters. In a one-stage calculation, without alliances, people know that a vote for party A directly helps party A, rather than being re-interpreted as an indirect support for party B. With alliances, the dedication of the votes is determined not directly by the citizens, but also indirectly by "the system". In the Netherlands, a vote for Alliance 1 may be felt to be supportive of any of three Political Groups: EPP, ECR, or EFD.

The Netherlands and Poland implement two-stage systems consisting of a super-apportionment at the top level, followed by several sub-apportionments on a lower level. The two states entertain two-stage systems for quite different reasons. In the Netherlands, political parties join into an alliance in order to increase their weights in the calculations. In Poland, geographical divisions of the same party are aggregated so as to combine local representativeness with state-wide uniformity.

EP2009PT	Political Group	Votes	Quotient	DivDwn
PPD/PSD	EPP	1 129 243	8.9	8
PS	S & D	946 475	7.5	7
BE	GUE/NGL	382 011	3.01	3
CDU(PCP-PEV)	GUE/NGL	379 707	2.99	2
CDS-PP	EPP	298 057	2.3	2
8 Others	NA	189 934	—	0
Sum (Divisor)		3 325 427	(127 000)	22

TABLE 1.24 *Portugal, 2009 EP election.* There is no electoral threshold. The divisor method with downward rounding is used. Every 127 000 votes justify roughly one seat, out of the Portuguese allocation of 22 seats.

EP2009RO	Political Group	Votes	Quotient	DivDwn
PSD+PC	S & D	1 504 218	11.2	11
PD-L	EPP	1 438 000	10.7	10
PNL	ALDE	702 974	5.2	5
UDMR	EPP	431 739	3.2	3
PRM	NA	419 094	3.1	3
Indep Elena Băsescu	EPP	204 280	1.5	1
Sum (Divisor)		4 700 305	(134 000)	33

TABLE 1.25 *Romania, 2009 EP election.* The threshold that applies to parties is five percent of valid votes, 242 002 votes. Another threshold is for indeps, the Hare-quota variant-4 HQ4 = 146 668. Five parties and one indep pass the thresholds.

1.9. PORTUGAL–ROMANIA–SWEDEN: METHOD OVERVIEW

Portugal is assigned 22 seats. There is no electoral threshold. The divisor method with downward rounding is used. Every 127 000 votes justify roughly one seat. Eight parties are too weak to obtain representation. See Table 1.24.

Romania uses two electoral thresholds. The first threshold applies to parties, and is five percent of valid votes. The second threshold applies to indeps, and is the Hare-quota variant-4. Since there are 4 840 033 valid votes and 33 seats, the quota is $\text{HQ4} = \lceil 4\,840\,033/33 \rceil = 146\,668$. Four parties and five indeps fail their thresholds, and the 139 728 votes for them are discarded. Five parties and one indep enter into the apportionment calculation. The apportionment method used is the divisor method with downward rounding. See Table 1.25.

Sweden uses a threshold of four percent of valid votes. With 3 168 546 valid votes it amounts to 126 742 votes. Six parties fail the threshold, and their 292 172 votes become ineffective. The apportionment of the 18 seats among the eight eligible parties is carried out using the *Swedish modification* of the divisor method with standard rounding. The modification applies a special prescription only for quotients between zero and unity, by replacing the standard rounding point 0.5 by the Swedish rounding point 0.7. That is, if a quotient stays below 0.7 it is rounded downwards to zero, if it is larger than 0.7 it is rounded upwards to unity. Our acronym for this method is Div0.7. The special prescription does not take effect for the 2009 EP data. See Table 1.26.

Method overview. Before turning to the last batch of states we classify the apportionments methods met so far. Divisor methods follow the motto "Divide and round". Many electoral system designers favor the divisor method with downward rounding, DivDwn. It is the procedure of choice in 16 of the 27 Member States. We

EP2009SE	Political Group	Votes	Quotient	Div0.7
S	S & D	773 513	4.8	5
M	EPP	596 710	3.7	4
FP	ALDE	430 385	2.7	3
MP	GREENS/EFA	349 114	2.2	2
PP	GREENS/EFA	225 915	1.4	1
V	GUE/NGL	179 182	1.1	1
C	ALDE	173 414	1.1	1
KD	EPP	148 141	0.9	1
Sum (Divisor)		2 876 374	(160 000)	18

TABLE 1.26 *Sweden, 2009 EP election.* The divisor method with standard rounding is used, except that a quotient below unity is rounded upwards to unity if it lies above 0.7, and downwards to zero otherwise. Because of this modification we use the acronym Div0.7.

may put forward at least three reasons for its dominant position. First, the divisor method with downward rounding originates from the early days of the proportional representation movement in the late nineteenth century. Endorsed by the leading protagonists of the movement, the Belgian *Victor D'Hondt* and the Swiss *Eduard Hagenbach-Bischoff*, it is distinguished by history. Second, its technical instructions are elementary. Whoever calculates the quotients of votes and divisor may drop their pen when reaching the decimal point. The missed fractional parts do not matter since they are rounded downwards to zero anyway. Superficially, downward rounding conveys the impression that no rounding operation is going on at all. Presumably the third reason is most tempting for parliamentary actors and their helpers. The divisor method with downward rounding produces seat numbers that are biased in favor of stronger parties at the expense of weaker parties. Usually a parliamentary majority involves stronger parties. When they have to bow to a proportional representation system, the divisor method with downward rounding is a favorite choice to secure an advantage. Advantages promise to accumulate as soon as the electoral regions are subdivided into districts that are separately evaluated.

The unbiased alternative is the divisor method with standard rounding, DivStd. It implements the rounding rule that has stood the test of time. To round downwards or upwards according as fractions are smaller or larger than one-half has proved to treat partners fairly, and this is true for commerce and business as much as it is true for electoral systems.

The differences between various divisor methods point towards the general theory. To see this, just contemplate interim quotients in the first interval $[0; 1]$. Standard rounding compares them with the decision point—generally termed *signpost*—one-half. The Swedish modification moves the signpost to 0.7. Downward rounding moves the signpost yet further up, to unity. Generally the theory equips every integer interval $[n-1; n]$ with a signpost $s(n)$, and these signposts become the decision points of an induced rounding rule. The location of the signposts within their intervals determines the properties of the corresponding divisor method.

Quota methods with residual fit by greatest remainders are captured by the motto “Divide and rank”. It is tempting to believe that there ought to be a fixed quota of votes justifying a seat, a kind of certified measure how many voters are represented by every Member of Parliament. However, the EP elections demonstrate that there is a bewildering variety of quotas to choose from. There is the Hare-quota, and its

EP2009SI	Political Group	Votes	Quotient	DivDwn
SDS	EPP	123 563	2.9	2
SD	S & D	85 407	2.03	2
N.Si.	EPP	76 866	1.8	1
LDS	ALDE	53 212	1.3	1
Zares	ALDE	45 238	1.1	1
DeSUS	NA	33 292	0.8	0
Sum (Divisor)		417 578	(42 000)	7

TABLE 1.27 *Slovenia, 2009 EP election.* There is a four percent threshold, but we are unable to ascertain whether it refers to votes cast or to valid votes. Every 42 000 votes justify roughly one seat.

many variants, as well as the Droop-quota, with yet more variants. Furthermore, the main apportionment leaves some residual seats to be looked after. Most residual apportionments rank the parties by decreasing magnitude of the fractional parts of their quotients. Lithuania imposes an additional full-seat restriction, Greece uses a peculiar approach of its own. Quota methods address the seat apportionment issue in a somewhat eclectic manner.

The third category of apportionment methods is the STV schemes. In Ireland and Malta surplus votes are transferred to lower-order preferences using a random procedure, STVran. In the Northern Ireland district of the United Kingdom, yet to be discussed, surplus votes are redistributed in a deterministic fashion according to the fractions of votes in favor of lower preferences, STVfra. We do not know of any initiative to export STV schemes to the EP elections at large, whence we forgo a theoretical study of their computational aspects. We find the available empirical data encouraging that, for the purposes of the EP elections, the results from STV schemes can be virtually duplicated with the divisor method with standard rounding.

1.10. SLOVENIA–SLOVAKIA–UNITED KINGDOM: LOCAL REPRESENTATION

Slovenia is assigned 7 seats. The electoral threshold is set at four percent, but we do not know whether it pertains to votes cast or valid votes. Either way the threshold removes six parties, and discards 45 894 votes as ineffective. The divisor method with downward rounding is used. Every 42 000 votes justify roughly one seat. See Table 1.27.

Slovakia uses a threshold of five percent of valid votes. With 826 782 valid votes, the threshold amounts to 41 340 votes. Eleven parties fail it, and 117 778 votes are discarded as ineffective. The allocation of the 13 Slovakian seats uses DQ3grR, the Droop-quota variant-3 method with residual fit by greatest remainders. Hence the quota is $\mathrm{DQ3} = \langle 709\,004/(13+1)\rangle = \langle 50\,643.1\rangle = 50\,643$. See Table 1.28.

The United Kingdom subdivides its area into 12 districts. The 72 seats are assigned to the districts well ahead of the election. Each district is evaluated separately. Eleven districts apply the divisor method with downward rounding, DivDwn. The twelfth district, Northern Ireland, uses the STV scheme with fractional vote transfer, STVfra, based on the Droop-quota, DrQ. There is no electoral threshold. Table 1.29 shows Districts 1–8, Table 1.30 continues with Districts 9–12.

EP2009SK	Political Group	Votes	Quotient	DQ3grR
SMER	S & D	264 722	5.227	5
SDKÚ-DS	EPP	140 426	2.773	2
SMK-MKP	EPP	93 750	1.851	2
KDH	EPP	89 905	1.775	2
LS-HZDS	ALDE	74 241	1.466	1
SNS	EFD	45 960	0.908	1
Sum (Split)		709 004	(.774)	13

TABLE 1.28 *Slovakia, 2009 EP election.* The Droop-quota variant-3 is used, DQ3 = $\langle 709\,004/(13+1)\rangle = \langle 50\,643.1\rangle = 50\,643$. Ten seats are allocated in the main apportionment, the three remaining seats by greatest remainders.

Local representation. Electoral districts are meant to strengthen the local ties between the electorate and those elected. To this end the electoral area is divided into two or more districts, to which the seats of a Member State are allocated prior to election day. The evaluation of districts varies. Belgium, France, Ireland, and the United Kingdom choose to evaluate them separately, each in its own right. In Germany and Poland districts are incorporated with a kind of nested evaluation. Yet another approach is to employ a double-proportional apportionment method (Section 14).

As a matter of fact local representativeness is the historical origin of parliamentary elections. In former days the country was divided into single-seat *constituencies*, and in each constituency the electorate voted a candidate into Parliament to represent them and their constituency. The person-to-person relationship was the natural ideal at a time when a constituency's electorate embraced fewer people than it does today.

In modern democracies voters may count into the millions. Political parties intervene and provide the institutional link that mediates between the many voters and the few parliamentarians. A candidate who is nominated by a party communicates to the electorate in both ways, through personal standing as well as through party affiliation. Current proportional representation systems shift the focus towards the electorate's division along party lines, and aim at fairly mapping the voter support of a party into this party's parliamentary seats.

Yet electoral systems provide various means to maintain the original intention of local representativeness. To this end we need a hierarchy of geographical notions for use in proportional representation systems. Since no standardized set of terms is available we use the following classification. The *electoral region* is the largest possible territorial extension where the election takes place. The electoral region may be composed of various *electoral areas*. An electoral area may be further subdivided into *electoral districts*. On occasion, an electoral district is split into *electoral sections*, the finest level we consider for proportional representation systems.

For EP elections the electoral region is the aggregation of the territories of the 27 Member States. Each Member State constitutes an electoral area. Some Member States subdivide their area into electoral districts: Belgium, France, Germany, Ireland, Poland, and the United Kingdom. The French Outre-Mer district features electoral sections: Atlantique, Océan Indien, and Pacifique. When the electoral region is smaller than that of the European Union, often two levels suffice: the electoral region and, if applicable, its subdivision into electoral districts.

EP2009UK Party	Political Group	Votes	Quotient	Div-Dwn	Party	Political Group	Votes	Quotient	Div-Dwn
District 1: East					*District 2: East Midland*				
Cons.	ECR	500 331	3.3	3	Cons.	ECR	370 275	2.6	2
UKIP	EFD	313 921	2.1	2	Lab.	S & D	206 945	1.5	1
LD	ALDE	221 235	1.5	1	UKIP	EFD	201 984	1.4	1
Lab.	S & D	167 833	1.1	1	LD	ALDE	151 428	1.1	1
Greens	GREENS/EFA	141 016	0.9	0	BNP	NA	106 319	0.8	0
BNP	NA	97 013	0.6	0	Greens	GREENS/EFA	83 939	0.6	0
9 Others	NA	161 991	—	0	7 Others	NA	107 175	—	0
Sum (Divisor)		1 603 340	(150 000)	7	Sum (Divisor)		1 228 065	(140 000)	5
District 3: London					*District 4: North East*				
Cons.	ECR	479 037	3.4	3	Lab.	S & D	147 338	1.5	1
Lab.	S & D	372 590	2.7	2	Cons.	ECR	116 911	1.2	1
LD	ALDE	240 156	1.7	1	LD	ALDE	103 644	1.04	1
Greens	GREENS/EFA	190 589	1.4	1	UKIP	EFD	90 700	0.9	0
UKIP	EFD	188 440	1.3	1	BNP	NA	52 700	0.5	0
BNP	NA	86 420	0.6	0	Greens	GREENS/EFA	34 081	0.3	0
13 Others	NA	193 794	—	0	6 Others	NA	44 488	—	0
Sum (Divisor)		1 751 026	(140 000)	8	Sum (Divisor)		589 862	(100 000)	3
District 5: North West					*District 6: South East*				
Cons.	ECR	423 174	3.2	3	Cons.	ECR	812 288	4.95	4
Lab.	S & D	336 831	2.6	2	UKIP	EFD	440 002	2.7	2
UKIP	EFD	261 740	1.998	1	LD	ALDE	330 340	2.01	2
LD	ALDE	235 639	1.8	1	Greens	GREENS/EFA	271 506	1.7	1
BNP	NA	132 094	1.01	1	Lab.	S & D	192 592	1.2	1
Greens	GREENS/EFA	127 133	0.97	0	BNP	NA	101 769	0.6	0
7 Others	NA	135 214	—	0	9 Others	NA	186 361	—	0
Sum (Divisor)		1 651 825	(131 000)	8	Sum (Divisor)		2 334 858	(164 000)	10
District 7: South West					*District 8: West Midlands*				
Cons.	ECR	468 742	3.1	3	Cons.	ECR	396 847	2.8	2
UKIP	EFD	341 845	2.3	2	UKIP	EFD	300 471	2.1	2
LD	ALDE	266 253	1.8	1	Lab.	S & D	240 201	1.7	1
Greens	GREENS/EFA	144 179	0.96	0	LD	ALDE	170 246	1.2	1
Lab.	S & D	118 716	0.8	0	BNP	NA	121 967	0.9	0
BNP	NA	60 889	0.4	0	Greens	GREENS/EFA	88 244	0.6	0
11 Others	NA	149 084	—	0	6 Others	NA	95 060	—	0
Sum (Divisor)		1 549 708	(150 000)	6	Sum (Divisor)		1 413 036	(140 000)	6

TABLE 1.29 *Districts 1–8, United Kingdom, 2009 EP election.* The United Kingdom subdivides its area into twelve districts. Its 72 seats are allocated to the district prior to the election. Districts are evaluated separately, eleven of them use the divisor method with downward rounding, DivDwn.

1.11. DIVERSITY VERSUS UNIFORMITY

Ever since its inception the EP has confirmed its intention to standardize the procedures that the Member States employ for EP elections. Electoral systems comprise more than just counting votes. They determine who stands in the election, how they register, if they are given access to the media, whether they are reimbursed for their expenses, which ballot design is submitted to the voters and much more. But even when the view is narrowed down to what happens with the resulting vote counts, the multitude of procedures in the 27 Member States is perplexing.

Many States do not subdivide their area into electoral districts. Some do and evaluate the districts separately. Others do too, but handle districts through nested calculations of a super-apportionment followed by several sub-apportionments. Some states admit alliances, others do not. Some states forgo an electoral threshold or even

EP2009UK (*continued*) Party	Political Group	Votes	Quotient	Div-Dwn
District 9: Yorkshire and Humber				
Cons.	ECR, NA	299 802	2.5	2
Lab.	S & D	230 009	1.9	1
UKIP	EFD	213 750	1.8	1
BNP	NA	120 139	1.001	1
LD	ALDE	161 552	1.3	1
Greens	GREENS/EFA	104 456	0.9	0
6 Others	NA	96 472	—	0
Sum (Divisor)		1 226 180	(120 000)	6

Party	Political Group	Votes	Quotient	Div-Dwn
District 10: Scotland				
SNP	GREENS/EFA	321 007	2.9	2
Lab.	S & D	229 853	2.1	2
Cons.	ECR	185 794	1.7	1
LD	ALDE	127 038	1.2	1
Greens	GREENS/EFA	80 442	0.7	0
UKIP	EFD	57 788	0.5	0
BNP	NA	27 174	0.2	0
6 Others	NA	75 416	—	0
Sum (Divisor)		1 104 512	(110 000)	6

Party	Political Group	Votes	Quotient	DivDwn
District 11: Wales				
Cons.	ECR	145 193	1.8	1
Lab.	S & D	138 852	1.7	1
PC	GREENS/EFA	126 702	1.6	1
UKIP	EFD	87 585	1.1	1
LD	ALDE	73 082	0.9	0
Greens	GREENS/EFA	38 160	0.5	0
BNP	NA	37 114	0.5	0
4 Others	NA	37 832	—	0
Sum (Divisor)		684 520	(80 000)	4

Candidate	Party	Political Group	1st Pref	STVfra
District 12: Northern Ireland				
Bairbre de Brún	SF	GUE/NGL	126 184	1
Diane Dodds	DUP	NA	88 346	1
Jim Nicholson	UUP	ECR	82 893	1
Alban Maginness	SDLP	S & D	78 489	0
Steven Agnew	Greens (NI)	GREENS/EFA	15 764	0
2 Others	—	NA	92 896	0
Sum (Droop-quota)		(121 144)	484 572	3

TABLE 1.30 *Districts 9–12, United Kingdom, 2009 EP election.* In District 12: *Northern Ireland*, the STVfra system is used. Surplus votes are redistributed in fractions as given by the distribution of lower preferences.

declare it unconstitutional. Others install a threshold but refer to distinct ensembles, votes cast or valid votes or something else, relative to the whole electoral area or relative to separate electoral districts. The seat apportionment methods vary to an extent that it is even hard to tell how many of them are in use. Ten? Twelve? The states' electoral systems excel in diversity, not in uniformity. See Table 1.31.

Does it matter? After all, the Union may be viewed as a timely political construction allowing the citizenries of its Member States to preserve their domestic identities and idiosyncrasies in a diverse world. On the other hand, all parliaments in this world derive their political legitimization from the way how they get elected, and uniformity is always part of the underlying electoral principles. We will have more to say about electoral principles in the next chapter. As far as the EP is concerned, the Union's electoral principles are enshrined in its primary law and, to some extent, promise that all citizens of the Union are treated equally.

Electoral equality can be reliably gauged only when the European Parties start functioning on the Union level and give rise to a political system in which the many domestic parties agree to find their place. Certainly this scenario does not apply to the 2009 elections. We opt for replacing the invisible European Parties in the

Member State	Separate Apportionments	Nested Evaluations	Electoral Threshold	Apportionment Method(s)
AT	1	—	4% of valid votes	DivDwn
BE	3	—	—	DivDwn
BG	1	—	(See Sect. 1.2)	HaQgrR
CY	1	—	1.8% of valid votes	HQ3grR
CZ	1	—	5% of valid votes	DivDwn
DE	1	16 districts	5% of valid votes	DivStd
DK	1	3 alliances	—	DivDwn
EE	1	—	—	DivDwn
EL	1	—	3% of valid votes	HQ3-EL
ES	1	—	—	DivDwn
FI	1	1 alliance	—	DivDwn+plurality
FR	8	—	5% of valid votes	DivDwn
HU	1	—	5% of valid votes	DivDwn
IE	4	—	—	STVran
IT	1	5 districts	4% of valid votes	HQ1grR
LT	1	—	(See Sect. 1.7)	HQ2gR1
LU	1	—	—	DivDwn
LV	1	—	5% of votes cast	DivStd
MT	1	—	—	STVran
NL	1	3 alliances	—	DivDwn+HaQgrR
PL	1	13 districts	5% of valid votes	DivDwn+HaQgrR
PT	1	—	—	DivDwn
RO	1	—	(See Sect. 1.9)	DivDwn
SE	1	—	4% of valid votes	Div0.7
SI	1	—	(See Sect. 1.10)	DivDwn
SK	1	—	5% of valid votes	DQ3grR
UK	12	—	—	DivDwn+STVfra

TABLE 1.31 *Electoral indices of the 27 Member States.* Belgium, France, Ireland, and the United Kingdom subdivide their area into districts that are evaluated separately. Germany, Italy and Poland handle their districts by means of a super-apportionment and the induced sub-apportionments.

Union by the visible Political Groups in the EP. To this end Tables 1.3–1.30 mention for every domestic party the Political Group it joined. In Spain, Italy, Latvia, the Netherlands, and the United Kingdom some parties split their seats between several Political Groups; we split their votes accordingly. The Luxembourg counts are scaled by the marks-per-ballot average to obtain figures signifying human beings. In Ireland, Malta, and the Northern Ireland district of the United Kingdom, where STV schemes are used, the vote aggregation process is restricted to the first preferences shown in the tables. Domestic parties not affiliated to a Political Group nor obtaining a seat are neglected. The resulting union-wide aggregation of vote counts by Political Groups is exhibited in Table 1.32.

The table confronts the actual seat allocation with a hypothetical union-wide solution. The hypothetical apportionment yields a mirror image of the division of the Union's citizens along the political dimension. Using the divisor method with standard rounding every 196 000 votes would justify roughly one of the 736 seats. The actual allocation deviates from the hypothetical apportionment by twenty seat transfers, ranging from fifteen seats more than are indicated by proportionality up to seven seats fewer. The actual result exhibits a discordant seat assignment, in that the EFD group features more votes but fewer seats than the GUE/NGL group. The hypothetical allocation is concordant. However, due to the bold and artificial vote

Political Group	Actual Seats	Votes	Quotient	DivStd	Difference
EPP	265	52 324 413	267.0	267	−2
S & D	184	36 776 044	187.6	188	−4
ALDE	84	16 058 094	81.9	82	2
GREENS/EFA	55	12 070 029	61.6	62	−7
ECR	54	7 610 712	38.8	39	15
EFD	32	7 153 584	36.498	36	−4
GUE/NGL	35	6 280 876	32.0	32	3
NA	27	5 970 692	30.46	30	−3
Sum (Divisor)	736	144 244 444	(196 000)	736	20 − 20

TABLE 1.32 *Actual seats by Political Groups versus hypothetical seat apportionment, 2009 EP election.* The hypothetical apportionment is based on the groups' aggregate vote counts. The two seat vectors differ by a transfer of twenty seats.

aggregation rule the hypothetical solution is by no means authoritative. Yet Table 1.32 is indicative of a comprehensive view that would aid uniformity.

While a single union-wide apportionment faithfully reflects the political division of the electorate, it misses out on the geographical dimension of how citizens relate to the 27 Member States of the Union. Divisor methods can be adapted to honor both dimensions: the geographical distribution of the Union's citizens across Member States, and the political division as expressed through their party votes. These methods come under the heading of double proportionality, and are treated in Chapter 14.

Double proportionality merges the two approaches how to handle districts. The first approach is separate district evaluations. It relies on prespecified district magnitudes and, in each district, apportions the district magnitude proportionally among parties. The second approach is nested district evaluations. It relies on a super-apportionment to determine overall party-seats and, for each party, sub-apportions its seats locally in the districts. An iterative merger of the two approaches preserves the merits of both. The result is a double-proportional seat apportionment achieving proportionality relative to party votes as well as relative to district populations.

All seat apportionment methods have to be calibrated against what is required by a state's constitution or, in the case of the EP, by the Union's primary law. As for the European Union, its legal ramifications are still in flux. The Court of the European Union has yet to specify to what extent the treaty articles bind the electoral provisions. As a substitute, we exemplify the embedding of the electoral procedures into a constitutional framework with one of the Union's Member States, Germany. Chapter 2 describes the 2009 election to the German Parliament. The description includes an outline how the Federal Election Law responds to the five electoral principles that are set forth in the German Basic Law.

CHAPTER 2

Imposing Constitutionality: The 2009 Bundestag Election

Electoral systems veer between constitutional demands and political desires on the one hand side, and procedural rules and practical manageability on the other. The diverse requirements are exemplified by the 2009 election of the German Bundestag and its underlying Federal Election Law. The law provides citizens with two votes: a first vote to elect a constituency representative by plurality, and a second vote to mirror the electorate's division along party lines by proportionality. As Germany is a federation of sixteen states, federal components are also included. The Bundestag electoral system exemplifies the five electoral principles that underly Europe's electoral heritage: to elect the Members of the Bundestag by direct and universal suffrage in a free, equal, and secret ballot.

2.1. THE GERMAN FEDERAL ELECTION LAW

The Federal Election Law (Bundeswahlgesetz, BWG) defines the electoral system for the Bundestag (Federal Diet), Germany's parliament. The Bundestag is one of the five constitutional organs of the country. The other four are the Federal President (Bundespräsident), the Federal Government (Bundesregierung), the Federal Council (Bundesrat), and the Federal Constitutional Court (Bundesverfassungsgericht). The Federal Election Law is designed to implement *a proportional representation system that is combined with the election of persons* (eine mit der Personenwahl verbundene Verhältniswahl). The law's aims and instructions grow out of the country's history.

Parliamentary representation in Germany began with the North-German Confederation 1867–1871. It was established through the impetus and under the leadership of Prussian Prime Minister Otto von Bismarck. Impressed by the effectiveness of the franchise in Napoléon III's Second Republic, Bismarck had the members of the Confederate Reichstag (Diet) elected in single-seat constituencies by straight majority, with a second-round runoff if the straight majority was missed in the first round. Like many electoral system designers after him, Bismarck wanted to secure a safe majority for the ruling government. Trusting that the masses would support monarchic needs more enthusiastically than the upper class, Bismarck extended the franchise to universal male suffrage, practically a novum in Europe.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_2,

With the formation of the Imperial Reich in 1871, in the aftermath of waging a war against France, the confederate electoral system was copied into the new Constitution, with a noticeable exception. In order to admonish the members of the Imperial Reichstag to serve the Reich's interests and not just those of their local constituencies, and also to shield them from the evolving influence of political parties that Bismarck distrusted wholeheartedly, the Constitution's Article 29 obliged them to represent *the whole people* (das ganze Volk). The constitutional obligation that every Member of Parliament represents the whole people has since been upheld, in the Weimar Republic 1919–1933 as well as in the present Federal Republic.

In the Confederate Reichstag and in the Imperial Reichstag, candidates were elected on the basis of their personal standing. As parliamentary politics grew in its importance, the members of the Reichstag informally assembled in groups, and later formally aligned in political parties. In addition, the sheer number of voters under a universal male franchise necessitated an elaborate coordination of party politics. Eventually candidates were perceived as party nominees as much as they stood as individuals. They were presented to the electorate as one name among many on lists of nominees, submitted by a particular party and identified with this party.

The developing party system met with the disapproval of those who feared that it would corrupt the purpose of an election to select persons who would act as representatives of the people. The deficiency was strongly felt and widely discussed during the Weimar Republic. By then a proportional representation system had been adopted with every 60 000 votes justifying one seat, called the *automatic system*. With a large electoral turnout the Reichstag would have many seats. When people stayed away from the polls, the number of seats would decrease. The house size of the Republican Reichstag varied between a low of 459 seats in the elections 1920–1922, and a high of 647 seats in 1933. Variability of the house size was deemed inefficient when the Reichstag of one legislative period handed business over to its successor.

Yet the bigger issue was how parties filled the seats they won. Parties strictly followed the lists of nominees they had submitted prior to the election. Voters could neither delete nor add names. A party's list of nominees was definitive for the sequence how seats were filled. Critics maintained that the rigid lists put voters at the mercy of party bosses rather than encouraging the election of dedicated individuals. Strong candidates with a leadership personality, unwilling to submit to the intervention of a party's nomination assembly, would turn away from the political life of the Republic.

Such problems had been foreseen much earlier. Already *Siegfried Geyerhahn* (1902) had proposed a system to combine proportional representation with the election of persons. Essentially his system was built on two ingredients. First, the electoral region was to be subdivided into equal-sized constituencies, half as many as there were seats to fill. In every constituency voters would elect a representative by plurality. Second, the proposed ballot design was what might be called a *double-evaluated single vote*. On their ballot sheets voters would mark a candidate of a party. These marks would be evaluated twice. One count would be country-wide and by parties, to determine the number of seats a party would deserve proportionally.

The other count would be per constituency and by candidates. Locally, in each constituency, the candidate with the most votes would be declared elected. Overall, across the whole electoral region, the non-elected candidates of a party would be ranked by the number of votes they drew in their constituencies. This would generate *ex post* candidate lists based on the voters' say, as an alternative to lists of nominees submitted by the parties. The overall seat allotment of a party would be reduced by its number of constituency seats. The remaining seats would be filled in the sequence given by the party's *ex post* candidate list.

Geyerhahn also addressed the possibility that a party would win more constituency seats than entitled to by proportionality. Nowadays a surplus seat is called an *overhang seat* (Überhangmandat). Dismissed by *Geyerhahn* as a rare eventuality, the occurrence of overhang seats has become a common event. They have emerged in all Bundestag elections over the last thirty years, and in ever increasing numbers. Overhang seats are at the core of the difficulties arising with the current Federal Election Law. *Geyerhahn's* pamphlet was published in a prestigious series edited by prominent law scholars. The author did not enter into an academic career, though, and his name is absent from the debate during the Weimar Republic. A system similar to his was proposed in a 1925 newspaper article by *Richard Thoma*, a renowned law professor. *Thoma* opted for a ballot with two votes, a first vote for the election of a constituency representative, and a second vote for the election of a party list. A year later *Thoma*'s contribution triggered a response article by *Wilhelm Heile*.

After the Second World War the Allied Powers installed a Parliamentary Council with the mission to draft a constitution for a new, democratic Germany. The Council decided that the constitution was to include just the electoral principles that were deemed fundamental, but no procedural particulars. Indeed Article 38 of the Basic Law stipulates that the Members of the Bundestag are elected by direct and universal suffrage in a free, equal, and secret ballot. The Members of the Bundestag are representatives of the whole people. They are not bound by orders or directives, and shall submit solely to their conscience.

The task of designing an electoral system for the Bundestag was relegated to a subcommittee of the Parliamentary Council, the Committee on Electoral Procedure. The committee invited the expert witness *Richard Thoma* to review the electoral systems past and present. *Wilhelm Heile*, who had commented on *Thoma*'s 1925 newspaper article, was among the deputy committee members. After extensive deliberations the committee proposed an electoral system with a double-evaluated single vote, with rigid party lists to be registered before the election, and with a separate apportionment in every state of the Federal Republic. The common ground with *Geyerhahn's* pamphlet is striking, but regrettably we lack evidence on how the proposal came about. It eventually found its way into the Federal Election Law, and was used in the first two Bundestag elections in 1949 and 1953. In 1956 the law was amended substantially by introducing the two-votes ballot design. Furthermore, the seat apportionment calculations were arranged in a way persisting until 2009. The first stage is the superapportionment, that is, the apportionment of all Bundestag seats among the eligible parties proportionate to their country-wide vote counts. The second stage consists of the per-party sub-apportionments, that is, the apportionment of the country-wide seats of a party among its state-lists.

Current Bundestag ballot sheets, printed and issued by the election authorities, consist of one piece of paper displaying two columns. The left column, in black print, is the voter's *first vote* (Erststimme). This is the vote to elect a constituency representative. The right column, in blue print, is the voter's *second vote* (Zweitstimme). The second vote serves to elect one of the candidate lists that parties register in each of the sixteen states. The header of the right column includes a small-print hint, awkwardly worded, reminding voters that the second vote is the *decisive vote for the distribution of the seats altogether among the distinct parties* (maßgebende Stimme für die Verteilung der Sitze insgesamt auf die einzelnen Parteien).

The two-votes electoral system has grown into an export hit of democratic Germany. For instance New Zealand converted to it from first-past-the-post plurality in 1993, and the Scottish Parliament has been using it since its establishment in 1998. When adapted, the name may change, of course. New Zealand's acronym MMP is indicative of a *mixed member proportional* parliament. The Scottish AMP is short for an *additional member proportional* parliament. In Germany, the name *two-votes system* (Zweistimmensystem) is input-oriented. It places all the emphasis on offering voters a dual choice, in fulfillment of the law's aim to combine proportional representation with the election of persons.

2.2. COUNTRY-WIDE SUPER-APPORTIONMENT 2009

The Federal Election Law introduces an electoral threshold consisting of three components. A valid second vote becomes effective provided it is cast for a party (1) drawing at least five percent of the country-wide valid second votes, or (2) winning at least three constituencies, or (3) representing a national minority. The three components honor second votes (1), first votes (2), and minority representation (3).

The law is exemplified with the election of the Members of the 17th Bundestag, on 27 September 2009. We begin with the proportionality part of the system, the evaluation of the decisive second votes. The 40 764 288 effective second votes are cast for six parties. Of the six parties, five stand in two or more states. The SPD, FDP, LINKE, and GRÜNE present candidate lists in all sixteen states. The CDU stands in fifteen states, but not in Bavaria. The CSU campaigns in Bavaria, only. For every party, the second votes for their state-lists are aggregated into a country-wide count of second votes. The Federal Election Law decrees a notional Bundestag size of 598 seats, with the proviso that it may be modified by subsequent sections of the law. Indeed, the 2009 election generates 24 overhang seats and leads to 622 seats eventually. The last time when the notional house size persisted was the election to the 8th Bundestag in 1976. In any case the apportionment process begins with the allocation of 598 seats proportionate to second votes. This stage is called the *super-apportionment*. The divisor method with standard rounding, DivStd, is used to translate second votes into seats. See Table 2.1.

The 2009 numbers easily reveal the pertinent divisor interval. A divisor much below 68 196 would lead to more seats than 598. More precisely, the SPD is closest to acquire the next seat and determines the critical lower limit for the divisor, $9\,990\,488/146.5 = 68\,194.46$. On the other extreme, if the divisor were much larger than 68 196 then fewer seats would result. Clearly the party securing the last seat

17BT2009	Second Votes	Quotient	DivStd
Super-apportionment			
CDU	11 828 277	173.4	173
SPD	9 990 488	146.497	146
FDP	6 316 080	92.6	93
LINKE	5 155 933	75.6	76
GRÜNE	4 643 272	68.1	68
CSU	2 830 238	41.502	42
Sum (Divisor)	40 764 288	(68 196)	598

TABLE 2.1 *Super-apportionment of 598 seats by second votes, election to the 17th Bundestag 2009.* Every 68 196 votes justify roughly one seat. The seat apportionment is carried out using the divisor method with standard rounding, DivStd. The notional house size of 598 seats is subsequently modified.

is the CSU. Hence the upper limit for the divisor is $2\,830\,238/41.5 = 68\,198.50$. Any number in the interval $[68\,194.46; 68\,198.50]$ may serve as a viable divisor. For ease of reference we pick the midpoint, 68 196.48, and reduced it to as few digits as the interval permits, 68 196. This is the divisor quoted in Table 2.1.

As the CSU campaigns only in one state, Bavaria, its result may be finalized right away. The super-apportionment awards the CSU 42 *proportionality seats*. On the other hand the CSU wins 45 direct seats, because all of the 45 Bavarian constituencies are won by the CSU candidate. The law stipulates that in such a case the notional house size of 598 seats is enlarged by three seats, referred to as overhang seats. Each of the other parties calls for a sub-apportionment so that a party's super-apportionment seats are handed down to that party's state-lists.

2.3. PER-PARTY SUB-APPORTIONMENTS 2009

Parties present their nominees on *state-lists* (Landeslisten) when they campaign in more than a single state. Thus a state functions as a lower-level electoral district. The names for such districts vary greatly. It is *Land* in Germany, *kiescollege* and *circonscription électorale* in Belgium, *circonscription* in France, *constituency* in Ireland, *circoscrizione* in Italy, *okręgach* in Poland, and *electoral region* in the United Kingdom. To evade language barriers we uniformly refer to these units as *districts*. Districts are usually listed in some standard order. In Germany, the Federal Election Officer sorts the sixteen states from North to South by their northern-most latitude. The states' names are abbreviated by two-letter codes, see the top box of Table 2.2.

Every party with multiple state-lists calls for a *sub-apportionment*. Hence the 2009 election features five sub-apportionments, as shown in Table 2.2. Column "Dir." contains the number of direct seats a party wins in the state, column "Second Votes" lists the per-district counts of second votes. The overall party-seats from the super-apportionment are apportioned on the basis of the per-district second votes, again by means of the divisor method with standard rounding. The interim quotients for the CDU are included in the table. For lack of space they are omitted for SPD, FDP, LINKE and GRÜNE. The proportionality seats are exhibited in column "DivStd".

The final step relates direct seats to proportionality seats. If in a district a party's number of direct seats exceeds its number of proportionality seats, then the direct seats persist and the proportionality seats become moot, as marked by a trailing dot •

17BT2009 (*continued*)		Dir.	Second Votes	Quotient	DivStd	Overhang
Sub-apportionment to districts: CDU						
SH	Schleswig-Holstein	9	518 457	7.51	8•	1
MV	Mecklenburg-Vorpommern	6	287 481	4.2	4•	2
HH	Hamburg	3	246 667	3.6	4	0
NI	Niedersachsen	16	1 471 530	21.3	21	0
HB	Bremen	0	80 964	1.2	1	0
BB	Brandenburg	1	327 454	4.7	5	0
SA	Sachsen-Anhalt	4	362 311	5.3	5	0
BE	Berlin	5	393 180	5.7	6	0
NW	Nordrhein-Westfalen	37	3 111 478	45.1	45	0
SN	Sachsen	16	800 898	11.6	12•	4
HE	Hessen	15	1 022 822	14.8	15	0
TH	Thüringen	7	383 778	5.6	6•	1
RP	Rheinland-Pfalz	13	767 487	11.1	11•	2
BY	Bayern	—	—	—	—	—
BW	Baden-Württemberg	37	1 874 481	27.2	27•	10
SL	Saarland	4	179 289	2.6	3•	1
Sum (Divisor)		173	11 828 277	(69 000)	173	21

District	Dir.	Second Votes	DivStd	Overhang	Dir.	Second Votes	DivStd	Overhang
	Sub-apportionment to districts: SPD				*Sub-apportionment to districts: FDP*			
SH	2	430 739	6	0	0	261 767	4	0
MV	0	143 607	2	0	0	85 203	1	0
HH	3	242 942	4	0	0	117 143	2	0
NI	14	1 297 940	19	0	0	588 401	9	0
HB	2	102 419	2	0	0	35 968	1	0
BB	5	348 216	5	0	0	129 642	2	0
SA	0	202 850	3	0	0	124 247	2	0
BE	2	348 082	5	0	0	198 516	3	0
NW	27	2 678 956	39	0	0	1 394 554	20	0
SN	0	328 753	5	0	0	299 135	4	0
HE	6	812 721	12	0	0	527 432	8	0
TH	0	216 593	3	0	0	120 635	2	0
RP	2	520 990	8	0	0	364 673	5	0
BY	0	1 120 018	16	0	0	976 379	14	0
BW	1	1 051 198	15	0	0	1 022 958	15	0
SL	0	144 464	2	0	0	69 427	1	0
Sum	64	9 990 488	146	0	0	6 316 080	93	0

District	Dir.	Second Votes	DivStd	Overhang	Dir.	Second Votes	DivStd	Overhang
	Sub-apportionment to districts: LINKE				*Sub-apportionment to districts: GRÜNE*			
SH	0	127 203	2	0	0	203 782	3	0
MV	1	251 536	4	0	0	47 841	1	0
HH	0	99 096	1	0	0	138 454	2	0
NI	0	380 373	6	0	0	475 742	7	0
HB	0	48 369	1	0	0	52 283	1	0
BB	4	395 566	6	0	0	84 567	1	0
SA	5	389 456	6	0	0	61 734	1	0
BE	4	348 661	5	0	1	299 535	4	0
NW	0	789 814	11	0	0	945 831	14	0
SN	0	551 461	8	0	0	151 283	2	0
HE	0	271 455	4	0	0	381 948	6	0
TH	2	354 875	5	0	0	73 838	1	0
RP	0	205 180	3	0	0	211 971	3	0
BY	0	429 371	6	0	0	719 265	10	0
BW	0	389 637	6	0	0	755 648	11	0
SL	0	123 880	2	0	0	39 550	1	0
Sum	16	5 155 933	76	0	1	4 643 272	68	0

TABLE 2.2 *Sub-apportionments of party-seats to districts, election to 17th Bundestag 2009.* For the CDU, the direct seats (column "Dir.") overrule the proportionality seats (column "DivStd") in seven states (marked •), thus giving rise to 21 overhang seats. For the other four parties, all direct seats in all states can be incorporated into the corresponding proportionality seats.

in Table 2.2. In all other cases the proportionality seats persist. As an example we examine the CDU results. Its candidates win nine Schleswig-Holstein constituencies while Schleswig-Holstein's share of the 173 country-wide CDU seats is only eight seats. Thus one overhang seat is created, as recorded in column "Overhang". More overhang seats are brought into being in Mecklenburg-West Pommerania (2), Saxony (4), Thuringia (1), Rheinland-Palatinate (2), Baden-Württemberg (10), and Saarland (1). Thus the notional house size of the Bundestag increases by 21 CDU overhang seats.

For the other four parties all direct seats are carried by the corresponding proportionality seats, if only barely so for the SPD in the Free Hanseatic City of Bremen (HB: 2 direct seats and 2 proportionality seats) and in the State of Brandenburg (BB: 5 seats in either category). Nevertheless, the proportionality seats suffice to seat the constituency winners. Any remaining seats are filled from the parties' state-lists. In summary, the 17th Bundestag 2009 started out with 622 seats, 598 notional seats plus an additional 24 overhang seats, 3 for the CSU and 21 for the CDU. The seats are shared between the six parties CDU : SPD : FDP : LINKE : GRÜNE : CSU in the relation 194 : 146 : 93 : 76 : 68 : 45.

2.4. NEGATIVE VOTING WEIGHTS

Does the Bundestag seat apportionment fairly represent the whole people on the basis of the voters' ballots on election day? As a matter of fact on 3 July 2008 the Federal Constitutional Court ruled that the seat apportionment procedure, as described in the previous sections, violated the electoral principles of an equal and direct suffrage. Acknowledging that any amendment of the Federal Election Law poses a major challenge to the Bundestag, the court allowed the subsequent 2009 election to be conducted according to the prevalent law despite its unconstitutionality. However, the court ordered the Bundestag to restore the law's constitutionality by 30 June 2011. The government majority tabled a proposal on 28 June 2011, and later voted it into law against a vehemently dissenting opposition. The amended law entered into force on 3 December 2011. The opposition minority and a group of dedicated citizens, maintaining that the amended law still violated the electoral principles of an equal and direct suffrage, took the amended law to the Federal Constitutional Court. The court ruled that the amended law was indeed incompatible with the electoral principles in the Basic Law, and declared it null and void. As of ten o'clock on 25 July 2012, when the court read its opinion, Germany had no valid Federal Election Law.

The bone of contention is a bizarre effect called *negative (or inverse) voting weights*. The name indicates a discordant behavior of votes and seats. A party may profit from fewer votes by getting more seats, other things being equal. In other words voters may support the party of their choice by *not* casting their votes for this party. Conversely, more votes may be detrimental because of entailing fewer seats. For example, in 2009 the CDU might have profited from losing 18 000 second votes in Saxony (782 898 instead of 800 898), thereby winning an additional Bundestag seat (195 instead of 194). How? The loss of votes releases a proportionality seat of the CDU in Saxony, that instantly resurfaces for the CDU in Lower Saxony. But CDU seats in Saxony are sealed as direct seats. Hence the CDU in Saxony stays put, and the CDU in Lower Saxony increases their seats by one. The final tally is fewer votes, more seats.

What seems a mere hypothetical possibility turned into reality during the election to the 16th Bundestag in 2005. The main election took place on 18 September 2005. In the constituency Dresden I in Saxony, however, a candidate had died too suddenly for the party to nominate a substitute candidate for the main election. Thus a by-election had to be called that took place two weeks later, on 2 October 2005. The scenario then was of the same type as that in the previous paragraph. The Dresden I by-election leaves no doubt that about ten thousand CDU supporters withheld their second votes from the CDU in order not to harm the party of their choice.

This incident proved to the Federal Constitutional Court that the Federal Election Law makes voters speculate whether casting their votes for the party of their choice helps its cause, or hinders it. The fact that casting a vote may prove detrimental undermines the legitimizing function of democratic elections, and fools voters to a degree that is unacceptable. The court ruled the law unconstitutional in as far as it leads to negative vote weights, and called upon the Bundestag to amend these provisions. The story continues in Chapter 13.

2.5. DIRECT AND UNIVERSAL SUFFRAGE

The reasoning in the Constitutional Court's decision builds on the constitution of course. Article 38, Section 1, of the German Basic Law specifies five electoral principles:

> Die Abgeordneten des Deutschen Bundestages werden in allgemeiner, unmittelbarer, freier, gleicher und geheimer Wahl gewählt. Sie sind Vertreter des ganzen Volkes, an Aufträge und Weisungen nicht gebunden und nur ihrem Gewissen unterworfen.
>
> The Members of the German Bundestag are elected in a universal, direct, free, equal, and secret election. They are representatives of the whole people, not bound to orders nor instructions, and accountable solely to their conscience.

The five principles of a direct and universal suffrage in a free, equal, and secret ballot also constitute Europe's electoral heritage, according to the 2002 *Code of Good Practice in Electoral Matters* of the European Commission for Democracy Through Law (Venice Commission) of the Council of Europe. It is beyond the scope of this book to analyze the principles' meaning in full depth. We restrict ourselves to some brief comments to elucidate the principles' meanings.

The principle of direct suffrage demands that the translation of votes into seats is not hampered by any intervention. A prototype *indirect* election is the election of a President of the United States. The electorate is called to the polls to elect the President, but their votes are filtered through the actions of the Electoral College. The President is elected by the electors in the Electoral College, not directly by the people.

Indirect, stratified electoral systems were common in the multi-layered stratified societies of the Middle Ages. In medieval times Augsburg, a financial center of Europe, was a Free Imperial City in the Holy Roman Empire of the German Nation. The proceedings for the election of the Augsburg Mayor were somewhat circumstantial. Every male Augsburg citizen was a member of a guild, and his franchise was bound to the guild. In other words the electorate was subdivided into districts on the basis of social, rather than geographical, membership. As a guild member a man voted to elect his guild's Council of Twelve (Zwölferrat). The Councils of Twelve of all Augsburg guilds elected the Great Council (Großer Rat). The Great Council elected the Governing Council (Kleiner Rat). The Governing Council elected the Mayor. Part of the layered structure of the society was the claim of the upper level to be the *sounder part* (sanior pars), the political board that knew better than the lower level. When in 1502 the Augsburg carpenter Marx Neumüller was elected to the Great Council and he came to take his seat, he was sent *home again* (wider haim), because the council's majority did not approve of his election. In contrast, the principle of a direct election leaves no leeway for any political body to claim to be the sounder part, nor to interfere otherwise in the translation of votes into seats.

The principle of universal suffrage grants the franchise to every German. The meaning of "every German" has expanded over time. In imperial Germany it meant every male German of age twenty-five years and older, excluding the military, people living on welfare payments, and a few other groups of society. Nowadays the notion of every German embraces all male and female Germans who are at least eighteen years old. Exclusions from the franchise are kept to a minimum.

In the Middle Ages the concept of universal elections was unknown. All elections were limited to small electoral colleges. The archetype has always been the College of Cardinals to elect the Pope of the Roman Catholic Church. In addition many low level clerical institutions conducted elections to select their leaders. So did the secular world, again always limiting the electorate to an ensemble of privileged individuals. For example the King of the Holy Roman Empire of the German Nation was elected by the Electoral College of the seven Prince Electors. The inception and enforcement of direct and universal elections are an achievement of modern and contemporary history, growing out of the Era of Enlightenment, the independence of the United States, the French Revolution, and the times thereafter.

2.6. FREE, EQUAL, AND SECRET BALLOTS

The Middle Ages pondered already how electors could cast a free, equal, and secret ballot. In those days the term "electors" meant the members of a small, distinguished, and well-defined electoral college. In 1299, the Catalan philosopher *Ramon Llull* (1232–1316) elaborated on the pros and cons of open versus secret votes. In 1433, the German clergyman *Nicolaus Cusanus* (1401–1464), later promoted Cardinal of the Roman Curia, argued forcefully in favor of secret balloting to secure a free vote. It would keep electors from offering their votes for sale to the candidates, and it would prevent candidates from frightening the electors and exerting undue pressure on them. Moreover *Cusanus* held that a secret ballot was a necessity for all votes to acquire an equal impact on the final outcome. The historical sources prove that a free, equal, and secret ballot always has been considered a prerequisite for the validity of an election.

The principle of a free ballot means that voters are not subjected to any pressures about whom to vote for, that individuals may stand as candidates at their own discretion without anybody hindering them to do so, and that parties may participate with a minimum of bureaucratic requirements.

The principle of a secret ballot serves the same purpose today as it did in *Cusanus*'s time. It shields voters from being frightened or pressured by those who stand in the election, and it keeps them from selling their votes to candidates or parties.

The principle of an equal ballot constitutes the essence of contemporary democratic elections. The German Federal Constitutional Court has developed a comprehensive jurisdiction concerning the issue of electoral equality in proportional representation systems. Of course, the court's rulings are binding only within Germany. We feel that they may radiate beyond, in view of their inner consistency of what parliamentary elections are supposed to accomplish.

Equality is a relation among many subjects, not a property of a single item. It depends on the reference set within which it applies. In Sections 2.7–2.9 we explicate three distinct sets: voters, Members of Parliament, and political parties. In each set, equality may be captured by a precise numerical quantity. These quantities would be in a one-to-one relationship if seats were continuously divisible and could be fractionated. This is not so, seats are discrete entities and come in whole numbers. For this reason it makes a difference whether equality is referred to the voters, to the Members of Parliament, or to the parties, as emphasized in Section 2.10.

The equality principle is missing from the parts in the 2010 Treaty of Lisbon (Treaty on European Union) that deal with the election of the EP. Article 14, Section 3, reads:

> The members of the European Parliament shall be elected for a term of five years by direct universal suffrage in a free and secret ballot.

Is the inclusion, or omission, of electoral equality at the discretion of the legislator? Every Member State of the European Union is among the 47 members of the Council of Europe and, as such, endorses the Venice Commission's *Code of Good Practice in Electoral Matters*. However, it is a challenge to fill the abstract principle of electoral equality with a concrete meaning that would limit the margin of discretion of the legislator. Hence varied interpretations, not disagreement on its importance, may have prevented inclusion of the equality principle into the Treaty of Lisbon.

2.7. EQUALITY OF THE VOTERS' SUCCESS VALUES

The German Federal Constitutional Court uses the notion of *success value equality* (Erfolgswertgleichheit) of the voters' votes to asses electoral equality in proportional representation systems. The court coined the notion right at the beginning of its functioning, in a decision of 5 April 1952 that refers to a similar decision of the Bavarian State Constitutional Court a month earlier. The subjects on whom this notion of electoral equality focuses are the voters:

> ...; alle Wähler sollen mit der Stimme, die sie abgeben, den gleichen Einfluss auf das Wahlergebnis haben.

> ...; all voters shall have the same influence on the result of the election with the vote they cast.

Ever since, success value equality for all voters has been the central reference point for constitutional jurisdiction on proportional representation in Germany.

Since the voters are the main protagonists on election day it appears utterly adequate to focus on them, and not on the candidates elected, nor on the parties mediating between those voting and those elected. In fact, the goal was formulated decades earlier in 1910 by *André Sainte-Laguë* (1910a):

> Pour que l'égalité des bulletins de vote soit aussi complète que possible, chacun des électeurs doit avoir la même part d'influence.
>
> For the equality of the ballots to be as complete as possible, each voter must have an equal part in influence.

Similarly *George Pólya* put voters ahead of political parties rather explicitly in his 1919 writings on proportional representation systems:

> Das Prinzip des gleichen Wahlrechts fordert die möglichst gleichmäßige Berücksichtigung der Wünsche aller Wähler, aber nicht der Parteien, als solcher.
>
> The principle of equal suffrage demands observing as evenly as possible the wishes of all voters, but not of the parties, as such.

The reference to Members of Parliament, or to parties, justifies alternative manifestations of electoral equality, as explicated in subsequent sections.

Pursuant to its initial definition of success value equality of the voters' votes, the German Federal Constitutional Court has developed a consistent body of jurisdiction on the principle of electoral equality. The jurisdiction is in text form, of course, but it is precise to a degree that it lends itself to a unique and compelling form of quantification. The electoral success manifests itself through the number of seats apportioned to the party of the voter's choice, party P. Any voter of party P shares the individual success equally with the other voters who cast their vote for the same party. Hence the *success share* of a voter of party P is given by the ratio of seats relative to votes,

$$\frac{\text{seat number of party } P}{\text{vote count for party } P}.$$

For example, in Table 2.1 the success share of a CDU voter amounts to 173/11 828 277 = 0.000 014 626. Since the vote counts may range into the millions, a success share is a minute quantity. Besides being of an awkward order of magnitude numerically, success shares do not tell the full story. The impact of 173 seats is contingent on how many seats are available altogether. For instance, in a house of size 300 they constitute a straight majority, while in a house of size 598 they account for a bit more than a quarter of the seats. Similarly, the vote count of a party exhibits its true weight only when referred to the overall vote total. Thus the *success value* of a voter's vote for party P is defined to be the ratio of seat shares relative to vote shares,

$$\frac{\text{seat number of party } P \text{ / house size}}{\text{vote count for party } P \text{ / vote total}}.$$

The definition turns the success values into manageable quantities. Theoretically, if all votes would enjoy the same success value, seat shares would coincide with vote shares. All voters would enjoy the success value unity, a one-hundred percent success.

Practically, deviations from theoretical equality are unavoidable. For example, in Table 2.1 the success value of a CDU voter turns out to be 0.997, while an SPD voter has success value 0.996. In other words, a CDU voter realizes a 99.7 percent success, an SPD voter a 99.6 percent success. The complete picture of the success values in the 2009 Bundestag election is as follows:

Voter	Success value [in%]	Deviation from 100%
CDU voter	99.7	−0.3
SPD voter	99.6	−0.4
FDP voter	100.4	0.4
LINKE voter	100.5	0.5
GRÜNE voter	99.8	−0.2
CSU voter	101.2	1.2

Some voters stay short of a one-hundred percent success, others reach beyond. Immediately the question comes to mind whether the observed deviations from ideal equality can be reduced any further, or not. We return to this question later, in Chapter 10. At present the point is more modest, to enable us to ask such questions. We may do so because the success value of a voter's vote, while originally introduced as a normative, qualitative standard of constitutional law, may be identified with a quantitative, procedural concept that is amenable to a conceptual analysis.

2.8. EQUALITY OF REPRESENTATIVE WEIGHTS

Another group of electoral protagonists are those elected, the Members of Parliament. They enjoy the constitutional right of *statutory equality.* From the viewpoint of Members of Parliament the election is equal provided every Member of Parliament represents the same number of voters. To this end we define the *representative weight* of a Member of Parliament of party P to be the average number of voters per seat,

$$\frac{\text{vote count for party } P}{\text{seat number of party } P}.$$

Representative weights are measured in vote fractions. They indicate the number of people represented by a party's Member of Parliament. Representative weights are interpretable without further standardization. Ideally, if they are all equal to each other, then they would coincide with the votes-per-seats ratio.

For example, it is straightforward to evaluate the representative weight of a *Mitglied des Bundestages* (MdB, Member of Parliament) in Table 2.1. The weight of a CDU MdB amounts to $11\,828\,277/173 = 68\,371.5$ vote fractions. An SPD MdB carries a weight that is a bit heavier, $9\,990\,488/146 = 68428$ votes. The ideal representative weight, the votes-per-seats ratio, is $40\,764\,288/598 = 68\,167.7$ vote fractions. Altogether the numbers come out as follows:

MdB	Representative weight	Deviation from 68 167.7
CDU MdB	68 371.5	203.8
SPD MdB	68 428.0	260.3
FDP MdB	67 914.8	−252.9
LINKE MdB	67 841.2	−326.5
GRÜNE MdB	68 283.4	115.7
CSU MdB	67 386.6	−781.1

When looking for numerical evidence the German Federal Constitutional Court calculates representative weights and rounds them to whole numbers. In this way the court evades the reference to vote fractions, and replaces them by numbers of voters. While the court's practice is convenient for communication purposes, it distracts from the technical difficulties of having to deal with interim quotients that are not whole numbers, in a context where only whole numbers make sense. In this book we stick to the original definition. Representative weights are votes-per-seats quotients, and hence generally are fractional numbers.

2.9. SATISFACTION OF THE PARTIES' IDEAL SHARES OF SEATS

Usually a parliament features just a handful or so of political parties, in contrast to hundreds of its members, and even millions of its voters. In view of the size of the numbers it may seem that parties are simplest to deal with. They are not the most important group in an electoral system, though. From a constitutional viewpoint parties rank only third in importance, behind voters, and behind Members of Parliament.

In a proportional representation system a party may claim its *ideal share of seats*, that is, the share of seats going along with its share of votes,

$$\frac{\text{vote count for party } P}{\text{vote total}} \times \text{house size.}$$

For example, if a party gets 12.3 percent of the votes then it would claim 12.3 percent of the seats. Since an arbitrary percentage of the house size usually yields a fractional number of seats, ideal shares are measured in seat fractions.

For the 2009 Bundestag election in Table 2.1, the actual seats of a party, its ideal share of seats, and their difference, its seat excess, are as follows:

Party	Actual seats	Ideal share	Seat excess
CDU	173	173.5173	−0.5173
SPD	146	146.5575	−0.5575
FDP	93	92.6550	0.3450
LINKE	76	75.6360	0.3640
GRÜNE	68	68.1154	−0.1154
CSU	42	41.5188	0.4812

Within each of the three groups—voters, candidates, and parties—the qualitative message is the same. For example, a CSU voter enjoys a success value that is larger than the ideal success value, by 1.2 percentage points. A CSU MdB is better off with a representative weight that is lighter than the ideal representative weight, by 781.1 voter fractions. The CSU as a party is allocated a seat number that exceeds the ideal share of seats, by 0.4812 seat fractions. What about these quantitative indices?

2.10. CONTINUOUS FITS VERSUS DISCRETE APPORTIONMENTS

Let us hypothetically imagine an ideal world where Members of Parliament are divisible into continuous fractions. Then the ideal share of seats of a party P would constitute the solution. In terms of a formula, party P would be allocated an amount of seats given by the formula

$$y_P = \frac{v_P}{v_+}\, h,$$

where v_P denotes the vote count for party P, v_+ designates the total of all effective votes, and h signifies the house size. The outcome y_P on the left-hand side is the amount of seat fractions sought.

This assignment would comply perfectly well with all equality standards mentioned. The success values of all voters turn out to be equal to unity and signal a uniform one-hundred percent success,

$$\frac{y_P/h}{v_P/v_+} = 1.$$

The representative weights of all Members of Parliament would become equal,

$$\frac{v_P}{y_P} = \frac{v_+}{h}.$$

The ideal share of seats of all parties coincides with the seat allocation, by the very definition of y_P. The ideal world has no problems in coping with ideal equality.

But the world is real, not ideal. Members of Parliament are human beings who are entitled to be treated discretely, each in her or his own right. The problem is not to calculate continuous seat fractions. Rather, the discrete charm of each seat must be honored. The task is to determine a discrete apportionment, a procedure respecting the discrete character of the seats. Whatever continuous quantities are calculated in-between, in the end they must be rounded to whole numbers.

It seems easy enough to round fractional quotients to whole numbers. Rounding unfolds an enigmatic complexity, however, when it concerns parliamentary seats. This is no different from every-day life where we round in various ways to meet various sentiments. When asked for age we round downwards until the very last minute when on our birthday we grow a year older. When paying a bill in a restaurant the sole accepted rounding method is upward rounding, if only because of its implied expression of appreciation. When business partners negotiate contracts or convert currencies they use commercial rounding because experience has shown that it treats both partners in a fair and symmetric fashion. To cope with these issues in a systematic way, Chapter 3 develops a theory of rounding functions, and rounding rules.

CHAPTER 3

From Reals to Integers: Rounding Functions, Rounding Rules

A rounding function is a function mapping positive quantities into integers. Prominent examples are the floor function, the ceiling function, the commercial rounding function, and the even-number rounding function. Every rounding function induces a sequence of jumppoints, called signposts, where it advances from one integer to the next. In contrast, a rounding rule is more liberal, in that it maps into subsets of integers. A non-signpost is always mapped into a singleton. However, a signpost is mapped into a two-element subset comprising two neighboring integers. Thus rounding rules are elusive whether a signpost is rounded to one integer or to the other. Prominent examples are the rules of downward rounding, of standard rounding, and of upward rounding. Of particular interest are the rounding rules that belong to the one-parameter family of stationary signposts, or to the one-parameter family of power-mean signposts.

3.1. ROUNDING FUNCTIONS

All seat apportionment methods need to map interim quotients of some sort into whole numbers. This is achieved by rounding functions, and by rounding rules. We begin by introducing rounding functions.

Definition. *A* rounding function f *is an increasing function mapping the non-negative half-axis* $[0;\infty)$ *onto the set of natural numbers* $\mathbb{N} := \{0, 1, 2, 3, \ldots\}$,

$$f : [0;\infty) \to \mathbb{N}, \qquad f \text{ increasing and onto.}$$

The qualifiers "increasing" and "onto" mean that the function starts at $f(0) = 0$ and, as the arguments t grow from zero towards infinity, the values $f(t) = n$ grow from an integer n to its successor $n + 1$ unboundedly. The specific requirement that the rounded value $f(t)$ is an integer near t is postponed until Section 3.10. Four rounding functions are of particular interest: the floor function, the ceiling function, and the commercial and even-number rounding functions, as detailed in the following sections.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*, DOI 10.1007/978-3-319-03856-8_3, © Springer International Publishing Switzerland 2014

3.2. FLOOR FUNCTION

The mother of all rounding functions is the *floor function*, defined for all $t \geq 0$ through

$$\lfloor t \rfloor := \max \left\{ n \in \mathbb{N} \mid n \leq t \right\}.$$

The value $\lfloor t \rfloor$ is the natural number just below t, called the *integral part* of t. From $\lfloor t \rfloor \leq t < \lfloor t \rfloor + 1$, that is $t - 1 < \lfloor t \rfloor \leq t$, we see that $\lfloor t \rfloor$ is the unique integer in the half-open interval $(t-1; t]$. The remainder $t - \lfloor t \rfloor \in [0; 1)$ is the *fractional part* of t. The floor function is instrumental to decompose a non-negative number into its integral part and its fractional part,

$$t = \lfloor t \rfloor + \big(t - \lfloor t \rfloor \big).$$

Thus 12.34 has integral part $\lfloor 12.34 \rfloor = 12$, and fractional part $12.34 - \lfloor 12.34 \rfloor = .34$.

The floor function acts as the truncation operator that ignores all digits after the decimal point; this is why it is particularly suited for computing machinery. The function dates back to *Gauss* (1808) who denoted it by $[t]$; some authors refer to this notation as *Gauss brackets*. In this book the brackets $[\cdot]$ signify a general rounding function which we do not want to specify to a particular one.

3.3. TIES AND THE NEED FOR ROUNDING RULES

Seat apportionment problems involve the side condition that the sum of all seats must be equal to the preordained house size. The handling of these problems calls for a modification of rounding functions into rounding rules. We illustrate the insufficiency of rounding functions by means of the floor function. To this end we manipulate the data in the EP 2009 election in Austria (Table 1.3) so that they exhibit the cause of irritation, *ties*. Table 3.1 lowers party A to 840 000 votes, and raises party B to 700 000 votes. Thus both vote counts become integral multiples of the divisor 140 000.

In a first attempt we stick to the divisor 140 000, and obtain the interim quotients in the fifth column of Table 3.1. The floor function rounds party A's quotient 6.0 to 6 seats, and B's quotient 5.0 to 5 seats (not shown as a separate column in Table 3.1). The other parties get 3 : 2 : 2 : 0 seats. A total of 18 seats are handed out, one seat too many compared to the Austrian allocation of 17 seats.

In a second attempt we increase the divisor to 140 001, so as to decrease the number of seats apportioned. Naturally, all interim quotients decrease. Party A's quotient becomes genuinely smaller than 6.0 and is rounded downwards to 5 seats. Party B's quotient falls strictly below 5.0 and is rounded downwards to 4 seats. Adjoining the 3 : 2 : 2 : 0 seats of the others only 16 seats are alloted, one seat too few.

There are two options how to proceed. Either we declare the problem of apportioning 17 seats to be unsolvable and turn away, with weird implications for the Austrian EP allocation. Or we surmise that the solution lies in-between the first attempt that divides eleven seats between parties A and B according to 6 : 5, and the second attempt that divides nine seats according to 5 : 4. The solution sought ought to handle ten seats, and divide them into 6 : 4 or 5 : 5. The handling of ten seats, in either way, is the pragmatic conclusion to be adopted. It is common sense that a competition may terminate with some of the contestants tied, if only rarely so.

(EP2009AT)	Votes	Quotient	DivDwn	Quotient	DivDwn	Quotient	DivDwn
A	840 000	5.999 96	5	6.0	6−	6.000 04	6
B	700 000	4.999 96	4	5.0	4+	5.000 04	5
C	506 092	3.6	3	3.6	3	3.6	3
D	364 207	2.6	2	2.6	2	2.6	2
E	284 505	2.03	2	2.03	2	2.03	2
F	131 261	0.9	0	0.9	0	0.9	0
Sum (Divisor)	2 826 065	(140 001)	16	(140 000)	17	(139 999)	18

TABLE 3.1 *Occurrence of ties.* The vote counts of the two strongest parties in Table 1.3 are manipulated until they are tied. For house size 17 it is equally justified to allocate either 6 and 4 seats, or 5 seats each. The two options are indicated by means of the trailing plus- and minus-signs.

To save space we code ties by trailing plus-signs and trailing minus-signs. In Table 3.1, the notation 6− and 4+ indicates two-element sets of feasible seat numbers, $6- := \{5,6\}$ and $4+ := \{4,5\}$. Every choice of seat numbers from these sets is legitimate provided the sum of all seats exhausts the preordained house size. With two parties tied there are two equally justified apportionments, 6 : 4 and 5 : 5. Generally the number of tied apportionments is given by a binomial coefficient, see Section 4.7.

3.4. RULE OF DOWNWARD ROUNDING

Rounding rules pave the way to handle ties in an efficient manner. They are set-valued mappings, admitting at their jumppoints the two-element set comprising the integer before the jump, and the integer after the jump. We denote rounding rules by double brackets, as a reminder that they embrace two feasible rounded values. The *rule of downward rounding* is defined for all $t \geq 0$ through

$$\llbracket t \rrbracket := \begin{cases} \{\lfloor t \rfloor\} & \text{in case } t \neq 1,2,3,\ldots, \\ \{t-1, t\} & \text{in case } t = 1,2,3,\ldots. \end{cases}$$

The expression $n \in \llbracket t \rrbracket$ reads as "n results from t via downward rounding", or "t is rounded downwards to n". For instance, in Table 3.1 we get for the first three parties

$$\left\llbracket \frac{840\,000}{140\,000} \right\rrbracket = \llbracket 6 \rrbracket = \{5,6\}, \quad \left\llbracket \frac{700\,000}{140\,000} \right\rrbracket = \llbracket 5 \rrbracket = \{4,5\}, \quad \left\llbracket \frac{506\,092}{140\,000} \right\rrbracket = \llbracket 3.6 \rrbracket = \{3\}.$$

That is, downward rounding of the quotient of the first party yields 6 or 5, of the second party 5 or 4, of the third party 3.

The information inherent in the rule of downward rounding is the same as that provided by the floor function in cases when t is a non-integer. In these cases the floor function yields $\lfloor t \rfloor$, while downward rounding is more circumstantial by packaging the same answer into a one-element set, $\{\lfloor t \rfloor\}$. However, in cases when $t = 1,2,3,\ldots$ is an integer the floor function results in a unique value, $\lfloor t \rfloor = t$, whereas downward rounding yields a two-element set, $\llbracket t \rrbracket = \{t-1, t\}$. Rounding rules are more laborious to deal with, but the added labor is worth the gain. The exposition provides persuasive evidence that rounding rules capture the occurrence of ties in a rather practical manner.

3.5. CEILING FUNCTION AND RULE OF UPWARD ROUNDING

The counterpart of the floor function is the *ceiling function*, defined for all $t \geq 0$ through

$$\lceil t \rceil := \min \left\{ n \in \mathbb{N} \;\middle|\; n \geq t \right\}.$$

The value $n = \lceil t \rceil$ is the natural number just above t. From $\lceil t \rceil - 1 < t \leq \lceil t \rceil$, that is $t \leq \lceil t \rceil < t + 1$, we see that $\lceil t \rceil$ is the unique integer in the half-open interval $[t; t+1)$. Hence positive reals are mapped to positive integers, $t > 0 \Rightarrow \lceil t \rceil \geq 1$. Or, the other way around, only zero is mapped into zero, $\lceil t \rceil = 0 \Rightarrow t = 0$.

The *rule of upward rounding* that goes along with the ceiling function is given by

$$[\![t]\!] := \begin{cases} \left\{ \lceil t \rceil \right\} & \text{in case } t \neq 1, 2, 3, \ldots, \\ \left\{ t, t+1 \right\} & \text{in case } t = 1, 2, 3, \ldots. \end{cases}$$

The distinct orientation of downward rounding versus upward rounding becomes evident at the jumppoints $t = 1, 2, 3, \ldots$ where $[\![t]\!] = \{t-1, t\}$, but $[\![t]\!] = \{t, t+1\}$.

3.6. COMMERCIAL ROUNDING FUNCTION

A rounding function with a neutral orientation is the *commercial rounding function.* Fractional parts are rounded downwards when strictly smaller than one-half, and upwards otherwise. The commercial rounding function is defined for all $t \geq 0$ through

$$\langle t \rangle := \begin{cases} \lceil t \rceil & \text{in case } t - \lfloor t \rfloor \geq \frac{1}{2}, \\ \lfloor t \rfloor & \text{in case } t - \lfloor t \rfloor < \frac{1}{2}. \end{cases}$$

A mechanical prescription refers to the decimal representation of t. If its first digit after the decimal point is 0, 1, 2, 3 or 4, then the fractional part is considered to be a *minor fraction* and t gets rounded downwards. If the first decimal place is 5, 6, 7, 8 or 9, then the fractional part is taken to be a *major fraction* and t is rounded upwards. In particular a fractional part exactly equal to one-half is rounded upwards; this prescription leaves a residue of an upward drift that lacks symmetry.

The missing symmetry is restored by the *even-number rounding function.* By definition it coincides with the commercial rounding function except that numbers with a fractional part of one-half are rounded to the nearest even integer:

$$\langle t \rangle^* := \begin{cases} \lceil t \rceil & \text{in case } t - \lfloor t \rfloor > \frac{1}{2}, \text{ or } t - \lfloor t \rfloor = \frac{1}{2} \text{ and } \lceil t \rceil \text{ even}, \\ \lfloor t \rfloor & \text{in case } t - \lfloor t \rfloor < \frac{1}{2}, \text{ or } t - \lfloor t \rfloor = \frac{1}{2} \text{ and } \lfloor t \rfloor \text{ even}. \end{cases}$$

Hence the jumppoints 0.5, 1.5, 2.5, 3.5, 4.5, 5.5 etc. are rounded to 0, 2, 2, 4, 4, 6 etc. The even-number rounding function is the default `round` command implemented in the popular statistics software R. Here the reader may well ask Why even? Why not odd? A good question with no good answer.

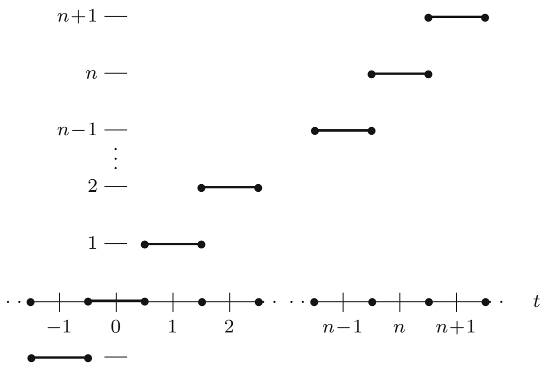

EXHIBIT 3.1 *The rule of standard rounding.* If the fractional part of t is equal to one-half, standard rounding is undecided and may round either way, downwards and upwards (•). Otherwise t is rounded downwards if its fractional part is smaller than one-half, and upwards if larger.

3.7. RULE OF STANDARD ROUNDING

The commercial and even-number rounding functions both jump at the midpoint $n - 1/2$ of the integer interval $[n-1; n]$. For this reason they induce the same rounding rule, the *rule of standard rounding*. It is defined for all $t \geq 0$ through

$$\langle\!\langle t \rangle\!\rangle := \begin{cases} \{\lceil t \rceil\} & \text{in case } t - \lfloor t \rfloor > \frac{1}{2}, \\ \{\lfloor t \rfloor, \lceil t \rceil\} & \text{in case } t - \lfloor t \rfloor = \frac{1}{2}, \\ \{\lfloor t \rfloor\} & \text{in case } t - \lfloor t \rfloor < \frac{1}{2}. \end{cases}$$

That is, if the fractional part of t is smaller than one-half, then standard rounding rounds downwards. If the fractional part is larger than one-half, then standard rounding rounds upwards. If the fractional part is exactly equal to one-half then standard rounding is undecided and may turn either way, downwards or upwards.

It is unambiguous to extend the rule of standard rounding from the half-axis $[0; \infty)$ to the whole real line $\mathbb{R}$. To this end multiplication and translation of real subsets is indicated by $-\{a, b, \ldots\} := \{-a, -b, \ldots\}$, and $\{a, b, \ldots\} + z := \{a + z, b + z, \ldots\}$. The definition of standard rounding is extended to the nonpositive half-axis via $\langle\!\langle -t \rangle\!\rangle := -\langle\!\langle t \rangle\!\rangle$, for all $t \geq 0$. With this extension all real numbers $t \in \mathbb{R}$ and all integers $z \in \mathbb{Z}$ obey the identity

$$\langle\!\langle t + z \rangle\!\rangle = \langle\!\langle t \rangle\!\rangle + z.$$

Standard rounding is the only rounding rule that is equivariant under reflections and equivariant under translations. This is a pleasing feature of standard rounding; see Exhibit 3.1. In contrast, a general rounding rule has an extension by reflection that differs from the extension by translation. For this reason the domain of definition of general rounding rules remains restricted to the positive half-axis $[0; \infty)$.

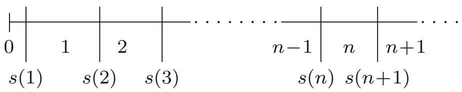

EXHIBIT 3.2 *Rounding rules.* Given a jumppoint sequence $s(0) = 0 \le s(1) < s(2) < \cdots$, an argument t in an open interval $\big(s(n); s(n+1)\big)$ is rounded to n, $[\![t]\!] = \{n\}$. The jumppoints $s(n)$ themselves are rounded to a two-element set, $[\![s(n)]\!] = \{n-1, n\}$, except for the initial value $[\![0]\!] = \{0\}$.

3.8. GENERAL ROUNDING RULES

A rounding rule allows a jumppoint to be rounded downwards or upwards and thus admits two values. A rounding function, being single-valued, cannot be so liberal and must turn either one way, or else the other way. Given a rounding function f, its nth jumppoint $s(n)$ is retrieved from the formula

$$s(n) := \inf \big\{ t \ge 0 \bigm| f(t) \ge n \big\} .$$

This formula does not depend on whether f rounds the jumppoint $s(n)$ downwards to $n-1$, or upwards to n. Since this is the type of freedom aimed at, the definition of rounding rules starts out from jumppoint sequences rather than rounding functions.

Definition. *A* jumppoint sequence $s(0), s(1), s(2), \ldots$ *is defined to be an unbounded sequence of nonnegative numbers satisfying* $s(0) = 0 \le s(1) < s(2) < \cdots$. *A jumppoint sequence defines a* rounding rule $[\![\cdot]\!]$ *by setting, for all* $t \ge 0$ *and* $n \in \mathbb{N}$,

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

A rounding rule is called pervious *when* $s(1) > 0$. *A rounding rule is called* impervious *when* $s(1) = 0$.

The definition ensures that zero is always and unambiguously rounded to zero, $[\![0]\!] = \{0\}$. In other words, all rounding rules obey the *no input–no output law.* In the empirical examples parties with vote count zero are not even mentioned as they certainly do not get a seat. For impervious rules these are the only have-nots. For pervious rules, it is possible that rounding annihilates very small quantities, $t \in \big[0; s(1)\big) \Rightarrow [\![t]\!] = \{0\}$. Its functioning may be compared to a sieve that loses quantities too small to get hold of. The attribute "pervious" (German: durchlässig) is meant to be indicative of annihilating a positive input that is too small. A pervious rounding rule is sketched in Exhibit 3.2.

In most instances a rounding rule returns a singleton, $[\![t]\!] = \{n\}$. These unambiguous instances apply whenever t lies strictly between the nth jumppoint and its successor, $s(n) < t < s(n+1)$. They may be paraphrased by saying that "t is rounded to n". All empirical examples in Chapters 1 and 2 fall into this category.

The remaining instances are two-way ties. A tie arises when t hits a positive jumppoint, $t = s(n) > 0$ for some $n \ge 1$. Then a rounding rule delivers the two-element set $\{n-1, n\}$. The rule considers it equally justified to round a tied input

$t = s(n) > 0$ downwards to $n-1$, or upwards to n. The ambivalence that a tie may be rounded either way, downwards or upwards, is indispensable for realistically modeling practical electoral systems, as pointed out in Section 3.3.

How do rounding rules relate to rounding functions? A rounding function $[\cdot]$ is said to be *compatible* with the rounding rule $[\![\cdot]\!]$ when the rounding function maps to values that are feasible for the given rounding rule,

$$[t] \in [\![t]\!] \qquad \text{for all } t \geq 0.$$

For example, the commercial rounding function is compatible with the rule of standard rounding, and so is the even-number rounding function. Evidently the relation of rounding rules to rounding functions is one-to-many. In contrast, the correspondence between rounding rules and jumppoint sequences is one-to-one.

The definition requires rounding functions to be increasing and onto (Section 3.1). Rounding rules share these properties in the sense of set-valued mappings. A rounding rule is *set-monotonic* in the sense of

$$t < T \qquad \Longrightarrow \qquad [\![t]\!] \leq [\![T]\!],$$

where the right-hand side means that all $n \in [\![t]\!]$ and all $N \in [\![T]\!]$ satisfy $n \leq N$. Moreover, a rounding rule maps onto the set $\mathbb{N}$ in the sense that the union of its image sets is equal to the set of all natural numbers, $\bigcup_{t \geq 0} [\![t]\!] = \mathbb{N}$.

In order not to overload the presentation we identify a rounding rule with its jumppoint sequence without any explicit reference. Hence we may start with a rounding rule $[\![\cdot]\!]$, and instantly refer to its underlying jumppoint sequence $s(n)$, $n \geq 0$. Or we name a specific jumppoint sequence, and then immediately apply the induced rounding rule. The identification often makes use of the *fundamental relation*

$$n \in [\![t]\!] \qquad \Longleftrightarrow \qquad s(n) \leq t \leq s(n+1)$$

that holds for all $t \geq 0$ and for all $n \in \mathbb{N}$. The left-hand side refers to the rounding rule, the right-hand side to its jumppoint sequence. The fundamental relation is a direct consequence of the definition, and is going to be called upon again and again.

3.9. GENERALIZED JUMPPOINT SEQUENCES

In the definition of a rounding rule $[\![\cdot]\!]$ we allow a jumppoint sequence to initially stay put, $0 = s(0) = s(1)$, but insist that it is strictly increasing thereafter, $s(n) < s(n+1)$ for all $n \geq 1$. This level of generality is sufficient for the inclusion of the special rounding rules mentioned earlier. For example upward rounding has the jumppoints $s(n) = n - 1$ for all $n \geq 1$, and hence satisfies $s(0) = s(1) = 0$.

More generality would be achieved by admitting unbounded sequences that are only weakly increasing, $s(0) = 0 \leq s(1) \leq s(2) \leq \cdots$. The definition of the induced rounding rule would have to be amended to admit multi-way ties,

$$[\![t]\!] := \begin{cases} \{n\} & \text{in case } t \in \big(s(n), s(n+1)\big), \\ \{m-1, \ldots, n\} & \text{in case } s(m-1) < t = s(m) = \cdots = s(n) < s(n+1), \end{cases}$$

so that the mapping remains onto $\mathbb{N}$. We are unaware of any empirical relevance of this level of generality. Most applications rely on jumppoint sequences that are more specific rather than more general, and have the structure of signpost sequences.

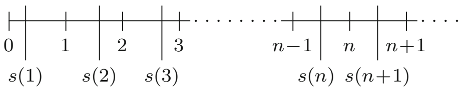

EXHIBIT 3.3 *Signpost sequences.* Signpost sequences are specific jumppoint sequences obeying the localization requirement $s(n) \in [n-1; n]$. The exhibit sketches a signpost sequence with signposts $s(1) = 0.3$, $s(2) = 1.6$, $s(3) = 2.7$, $s(n) = n - 1/2$, and $s(n+1) = n + 1/2$.

As to be defined next, signpost sequences localize the nth jumppoint within the nth integer interval, $[n-1; n]$. There are at least two electoral systems failing to satisfy this provision. The first is the rounding rule used for parliamentary elections in Estonia, with jumppoints $s(n) = n^{0.9}$ for all $n \geq 1$. Hence 7.5, lying between $s(9) = 9^{0.9} = 7.2$ and $s(10) = 10^{0.9} = 7.9$, is rounded to 9. The other electoral system is that for parliamentary elections in Macau. Its jumppoint sequence is $s(n) = 2^n$ for $n \geq 1$. Hence $t \in (0; 2)$ is rounded to 0, while $t \in (2; 4)$ is rounded to 1, and $t \in (4; 8)$ to 2, and so on. We note that the truncated rounding rules $[\![\cdot]\!]_a^b$ of Section 12.3 may be expressed by means of the generalized jumppoints

$$s_a^b(n) := \begin{cases} \infty & \text{in case } n \geq b, \\ s(n) & \text{in case } n = a+1, \ldots, b, \\ 0 & \text{in case } n \leq a. \end{cases}$$

Since not much seems to be gained by this type of generalization we restrict attention to specific jumppoint sequences that qualify as signpost sequences, in the following sense.

3.10. SIGNPOST SEQUENCES

Definition. *A* signpost sequence $s(0), s(1), s(2), \ldots$ *is characterized by the three properties* a, b, *and* c*:*

a. (Initialization) *The starting signpost is fixed at zero,* $s(0) = 0$.

b. (Localization) *All subsequent signposts belong to consecutive integer intervals,*

$$s(n) \in [n-1; n] \qquad \text{for all } n = 1, 2, \ldots.$$

c. (Left-right disjunction) *If there exists a signpost hitting a left limit of its localization interval then all signposts stay away from their right limits, and if there is a signpost hitting a right limit then all signposts stay away from their left limits,*

$$s(n) = n-1 \text{ for some } n \geq 1 \qquad \Longrightarrow \qquad s(n) < n \text{ for all } n \geq 1,$$

$$s(n) = n \text{ for some } n \geq 1 \qquad \Longrightarrow \qquad s(n) > n-1 \text{ for all } n \geq 1.$$

Every signpost sequence is strictly increasing, with the sole exception that impervious sequences start out with $s(0) = s(1) = 0$. Indeed, the localization property implies that non-strictness $s(n) = s(n+1)$ actually entails $s(n) = n = s(n+1)$. Then $s(n)$ hits the right limit of its interval and $s(n+1)$ its left limit. This constellation is excluded by the left-right disjunction. This proves strict monotonicity, and verifies that signpost sequences are special jumppoint sequences.

A signpost sequence decomposes the non-negative half-axis $[0;\infty)$ into successive intervals, $\big[0;s(1)\big)$, $\big[s(1);s(2)\big)$, $\big[s(2);s(3)\big)$, etc. For impervious sequences the initial interval $\big[0;s(1)\big) = [0;0)$ is empty and dispensable, yet is retained for the sake of notational uniformity. The intervals are localized according to $n \in \big[s(n);s(n{+}1)\big]$. That is, the interval $\big[s(n);s(n+1)\big]$ is the domain of attraction to round to the integer n. See Exhibit 3.3.

The name "signpost" is borrowed from *Balinski / Young* (1982 [62]). It nicely alliterates with *Sprungstellen* in German, and *seuils* in French. It is convenient to refer to $s(n)$ as the nth signpost. Aside from the starting signpost $s(0) = 0$, the labeling indicates that the first signpost $s(1)$ lies in the first integer interval $[0;1]$, the second signpost $s(2)$ in the second integer interval $[1;2]$, and so on. The signpost sequences of the rounding rules mentioned earlier are as follows:

Downward rounding, $\llbracket\cdot\rrbracket$:	0,	1,	2,	3,	$\dots$,
Standard rounding, $\langle\!\langle\cdot\rangle\!\rangle$:	0,	0.5,	1.5,	2.5,	$\dots$,
Upward rounding, $\lceil\!\lceil\cdot\rceil\!\rceil$:	0,	0,	1,	2,	$\dots$.

The multitude of signpost sequences comprises a pair of one-parameter families that deserve special attention because of their relevance for the subsequent development. The first is the family of stationary signposts with split parameter r, the second is the family of power-mean signposts with power parameter p.

3.11. STATIONARY SIGNPOSTS

Definition. *The sequence of* stationary signposts with split $r \in [0;1]$ *is defined through* $s_r(0) := 0$ *and, for all* $n \geq 1$, *through*

$$s_r(n) := n - 1 + r.$$

That is, the sequence with split r is $0,\ r,\ 1+r,\ 2+r,\ 3+r$ etc. Within the integer intervals $[n-1;n]$, the relative position of the signposts $s_r(n) = n-1+r$ stays the same. The distance to the lower limit is r, to the upper limit it is $1-r$. The attribute "stationary" alludes to the stability of the signposts under shifts from one integer interval to the next. For stationary signpost sequences all domains of attraction for rounding to a positive integer $n \geq 1$ have the same length unity, $s_r(n{+}1) - s_r(n) = (n{+}r) - (n-1+r) = 1$. However, the length of the domain of attraction for rounding to zero is $s_r(1) - s_r(0) = r$. It equals unity if and only if the split is unity, $r = 1$. All other splits $r < 1$ start with a first interval of length less than unity. This boundary effect has to be kept in mind.

The family of stationary signposts starts from upward rounding ($r = 0$), passes through standard rounding ($r = .5$), and finishes with downward rounding ($r = 1$). It thus affords a smooth transition and embeds the three rounding rules into a wider family. A similar embedding is provided by the family of power-mean signposts. While more elaborate notationally they embrace another two rules of traditional interests, harmonic rounding and geometric rounding.

3.12. POWER-MEAN SIGNPOSTS

Definition. *The sequence of* power-mean signposts with power parameter $p \in [-\infty; \infty]$ *is defined through* $\widetilde{s}_p(0) := 0$ *and, for all* $n \geq 1$, *through*

$$\widetilde{s}_p(n) := \left(\frac{(n-1)^p + n^p}{2} \right)^{1/p} \qquad \textit{in case } p \neq -\infty, 0, \infty,$$

$$\widetilde{s}_{-\infty}(n) = n - 1, \qquad \widetilde{s}_0(n) = \sqrt{(n-1)n}, \qquad \widetilde{s}_\infty(n) = n.$$

The expression $(n-1)^p$ is nonsensical for $n = 1$ and $p < 0$, and so we set $\widetilde{s}_p(1) := 0$ for $p \in (-\infty; 0)$. The power-mean sequences with powers $p \in [-\infty; 0]$ have $\widetilde{s}_p(1) = 0$ and hence are impervious. Those with $p \in (0; \infty]$ satisfy $\widetilde{s}_p(1) > 0$, they are pervious.

The signpost $\widetilde{s}_p(n)$ is the power-mean with power parameter p of the limits $n - 1$ and n of the integer interval $[n-1; n]$. Generally, power-means are an average of two positive quantities $a, b > 0$,

$$\left(\frac{a^p + b^p}{2} \right)^{1/p}.$$

The exponents $p \neq -\infty, 0, \infty$ are immediately applicable. Since the expression is convergent as p tends to $-\infty, 0, \infty$, the three exceptional cases fit in. Moreover, convergence takes place as a or b tend to zero. The limits give rise to the case distinctions in the definition of power-mean signposts. It is worth remembering that the parameterization is continuous, $\lim_{q \to p} \widetilde{s}_q(n) = \widetilde{s}_p(n)$, for all $n \in \mathbb{N}$ and all $p \in [-\infty; \infty]$.

The family of power-mean signposts retrieves the special rules of upward rounding ($p = -\infty$), of standard rounding ($p = 1$), and of downward rounding ($p = \infty$). For $p = -1, 0$ it adjoins another two rules of traditional interest. The signpost $\widetilde{s}_{-1}(n) = \big((n-1)^{-1}/2 + n^{-1}/2\big)^{-1}$ is the harmonic mean of $n-1$ and n, whence the rule with power $p = -1$ is called *harmonic rounding*. The signpost $\widetilde{s}_0(n) = \sqrt{(n-1)n}$ designates the geometric mean of $n - 1$ and n, whence the rule with power parameter $p = 0$ is called *geometric rounding*. While the stationary family picks up three special rules, the power-mean family includes each of the five traditional rounding rules:

	Split r	Power p
Downward rounding:	1	∞
Standard rounding:	1/2	1
Geometric rounding:	—	0
Harmonic rounding:	—	-1
Upward rounding:	0	$-\infty$

Finally note that l'Hôpital's rule yields $\lim_{n \to \infty} \big(\widetilde{s}_p(n) - (n - 1/2)\big) = 0$ whenever $p \in (-\infty; \infty)$. Thus all power-mean rules for which the power parameter is finite converge to standard rounding as n tends to infinity. The diversity inherent in the power-mean family evaporates and reduces to the special rules of upward, standard, and downward rounding. However, the transition to the limit $n \to \infty$ models a parliament with an infinite number of seats, whereas practical seat apportionment problems mostly deal with smallish seat numbers. Both families merit attention, stationary signposts as well as power-mean signposts.

1975	Population	Proportion	Percent
Asia	2 295 000 000	0.57289	57
Europe	734 000 000	0.18323	18
Americas	540 000 000	0.13480	13
Africa	417 000 000	0.10409	10
Australia and Oceania	20 000 000	0.00499	0
Sum	4 006 000 000	1.00000	98

TABLE 3.2 *Insufficiency of simple rounding.* The percentages, obtained via commercial rounding of the proportions, sum to 98 percent only. The discrepancy of −2 percent, of a world population of four billion, accounts for more than eighty million people, a country of the size of Germany.

3.13. SIMPLE ROUNDING DOES NOT SUFFICE!

While rounding rules succeed in mapping a single quotient into a whole number, they need not achieve a collective side condition that often forms an indispensable part of an apportionment problem. For example when quotients are rounded to percentages, the mere mention of the term "percentages" promises that the total is equal to 100. It is rather likely though that the promise does not come true, and that the sum misses the target value 100 by a discrepancy of some percentage points too few or too many. Chapter 6 investigates the discrepancy distribution in detail. For now we satisfy ourselves with an example. Table 3.2 exhibits the 1975 World Population, see *Kopfermann* (1991 [109]). The percentage total equals 98, not 100, and leaves a discrepancy of −2 percent. The missing two percents of a population of 4 006 000 000 does away with more than 80 million people, a country of the size of Germany.

The procedure applied may be called *simple rounding* or naïve rounding. Quantities are individually rounded, without regard to a collective side condition. Since simple rounding makes no provisions to fulfill the side condition, the resulting roundings generally fail to do so. For this reason many statistical publications contain a disclaimer, somewhere in the small print, that any percentages quoted may fail to sum to 100 "due to rounding effects". However, when the 100 units signify parliamentary seats rather than percentage points, nobody would dare to suggest that some seats disappear "due to rounding effects. The way-out of the dilemma is to introduce apportionment methods. Chapter 4 is devoted to a class of apportionment methods called divisor methods, Chapter 5 deals with another class called quota methods.

CHAPTER 4

Divisor Methods of Apportionment: Divide and Round

Apportionment methods are procedures to allocate a preordained number of seats proportionately to vote counts, census figures, or similar quantities. Apportionment methods must be anonymous, balanced, concordant, decent, and exact. Beyond these organizing principles the central issue is proportionality. The chapter focuses on the family of divisor methods; they follow the motto "Divide and round". The properties of general divisor methods are elaborated in detail. Five divisor methods are of particular traditional interest: the divisor methods with downward rounding, with standard rounding, with geometric rounding, with harmonic rounding, and with upward rounding.

4.1. APPORTIONMENT RULES

The standard setting for seat apportionment problems presumes that a preordained number of seats h, called *house size*, must be allocated. The house size is taken to be a natural number, $h \in \mathbb{N}$. The h seats are to be apportioned among ℓ political parties in proportion to the parties' *vote weights* $v_1, \dots, v_\ell$. Usually the weights are vote counts. At times, it is convenient to take them to be vote shares. Vote counts are integers, of course, while vote shares are fractions. To cover both cases we allow vote weights to be arbitrary positive values, $v_j \in (0; \infty)$. We assemble them into the *vote vector*

$$(v_1, \dots, v_\ell) \in (0; \infty)^\ell.$$

Because component sums of vectors keep occurring in the sequel we abbreviate them by a subscript plus-sign, $v_+ := v_1 + \dots + v_\ell$. The output is a *seat vector* $x = (x_1, \dots, x_\ell)$ for house size h, that is, a vector with integral components that are summing to h. A convenient shorthand notation for the set of all seat vectors for house size h is

$$\mathbb{N}^\ell(h) := \left\{ (x_1, \dots, x_\ell) \in \mathbb{N}^\ell \;\middle|\; x_+ = h \right\} \subseteq \{0, 1, \dots, h-1, h\}^\ell.$$

Since the last set is finite, so is its subset $\mathbb{N}^\ell(h)$. The component x_j of a seat vector $x \in \mathbb{N}^\ell(h)$ signifies the *seat number* of party $j \le \ell$.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*, DOI 10.1007/978-3-319-03856-8_4,

The problem is the same when seats are apportioned among several states in proportion to their census figures. Then j indicates a state, and the weight v_j is its apportionment population. In other applications we may be given a set of items whose measurements are rounded to integral percentages $0, 1/100, \ldots, 99/100, 1$. Then j is the name of the item, the weight v_j is its measured score, and the $h = 100$ apportionment units are percentage points. Since the task of apportioning seats to parties of a political body is a core issue, we continue to orient the terminology towards this type of application. It is obvious how to re-interpret the results for other applications.

A spontaneous definition of an apportionment function A would demand that A maps a house size $h \in \mathbb{N}$ and a vote vector $v \in (0;\infty)^\ell$ into a seat vector $x \in \mathbb{N}^\ell(h)$. However, a lasting definition must support the occurrence of ties, as pointed out in Section 3.3. This is achieved by switching from a single-valued function to a set-valued rule. Hence A is assumed to map the input $(h; v_1, \ldots, v_\ell)$ into a subset of seat vectors that is nonempty, $\emptyset \neq A(h;v) \subseteq \mathbb{N}^\ell(h)$. However, this notion defies reality because it is contingent on the size of the party system, ℓ. Nobody distinguishes between the divisor method with downward rounding for $\ell = 6$ parties as in the Austrian EP election (Table 1.3), and the divisor method with downward rounding for $\ell = 25$ parties as in the Spanish EP election (Table 1.12). It is common practice to quote the method without any reference to the size of the party system. Hence it remains to ensure that the definition does not involve the number of participating parties, ℓ.

To this end we denote by V the union of all vote vectors and by W the union of the power sets of $\mathbb{N}^\ell$, that is, $V := \bigcup_{\ell \geq 2} (0;\infty)^\ell$ and $W := \bigcup_{\ell \geq 2} 2^{\mathbb{N}^\ell}$. For a given input $(h; v)$, the set of seat vectors aimed at is an element in the range space W. For a vector $v \in V$ we designate the number of its components by $\ell(v)$. Now we are in a position to provide a definition of apportionment rules that conforms with their usage.

Definition. *An* apportionment rule A *maps house sizes* $h \in \mathbb{N}$ *and vote vectors* $v \in V$ *into nonempty subsets of seat vectors for house size* h*,*

$$A : \mathbb{N} \times V \to W \quad \textit{such that} \quad \emptyset \neq A(h;v) \subseteq \mathbb{N}^{\ell(v)}(h).$$

A solution set $A(h;v)$ is called *tie-free* when it is a singleton, $A(h;v) = \{x\}$. In this case it is unambiguous to identify the solution set $\{x\}$ with its element x. In the general case the notation $x \in A(h;v)$ says that the seat vector x is an apportionment of h seats according to the vote vector v. For an abstract apportionment rule to prove practically useful it must satisfy five general principles.

4.2. ORGANIZING PRINCIPLES

A reasonable apportionment rule must be anonymous, balanced, concordant, decent, and exact. We discuss the five principles one after the other.

Anonymity. An apportionment rule A is called *anonymous* when every rearrangement of the vote weights goes along with the same rearrangement of the seat numbers. Whether a party is listed first or last has no effect on its seat number. Chapters 1 and 2 make use of anonymity in that parties are ranked by decreasing vote counts. On the other hand districts usually follow a fix geographical order. Whichever order applies, the resulting seat numbers are carried along.

Balancedness. An apportionment rule A is called *balanced* when the seat numbers of equally strong parties differ by at most one seat. Formally, all seat vectors $(x_1, \ldots, x_\ell) \in A(h; v_1, \ldots, v_\ell)$ and all parties $j, k = 1, \ldots, \ell$ are required to satisfy

$$v_j = v_k \qquad \Longrightarrow \qquad \left| x_j - x_k \right| \leq 1.$$

It would be tempting to insist that equally strong parties get the same number of seats. But when two parties with the same vote counts must share a house size that is odd, a one-seat imbalance is unavoidable. Balancedness ascertains that in tied instances the spread does not grow beyond the inevitable minimum, one seat.

Concordance. An apportionment rule A is called *concordant* when of two parties the stronger party does not get fewer seats than the weaker party. Formally, all seat vectors $(x_1, \ldots, x_\ell) \in A(h; v_1, \ldots, v_\ell)$ and all parties $j, k = 1, \ldots, \ell$ satisfy

$$v_j > v_k \qquad \Longrightarrow \qquad x_j \geq x_k.$$

Concordance is easy to check visually provided parties are listed by decreasing vote counts, because then the corresponding seat numbers must be non-increasing too. When vote counts are ordered otherwise, as with districts, a visual check cannot be relied upon and needs to be supplemented by a computer program. One would tend to believe that a *discordant result*, that is a non-concordant results, is an academic artifact and does not occur in practice. This is not so. They emerge in electoral systems that use a succession of several computational stages. The formation of alliances gives rise to many examples of discordant seat apportionments, see Section 7.11.

Decency. An apportionment rule A is called *decent*, or positively homogeneous of degree zero, when its results for vote weights $(\alpha v_1, \ldots, \alpha v_\ell)$ stay the same for all scale factors $\alpha > 0$. In particular, a scaling with the vote total $v_+ = \sum_{j \leq \ell} v_j$ affords a passage from the raw vote counts v_j to the induced vote shares $w_j = v_j / v_+$. This transition gives the apportionment problem a probabilistic twist. Since vote shares sum to unity, $w_+ = 1$, they may be interpreted as probabilities. Because of the usually large size of the denominator v_+, the vote shares w_j are virtually continuous probability weights. In contrast, the output signifies probabilities $x_1/h, \ldots, x_\ell/h$ that are distinguished by the discrete character of the numerators, $x_j \in \{0, 1, \ldots, h-1, h\}$. Hence from a stochastic viewpoint a probability distribution with continuous weights is to be approximated by a probability distribution with discrete weights.

Exactness. An apportionment rule A is called *exact* when any integer input with the desired component sum reproduces itself as the unique output,

$$A(h; x_1, \ldots, x_\ell) = \{(x_1, \ldots, x_\ell)\}$$

for all $(x_1, \ldots, x_\ell) \in \mathbb{N}^\ell(h)$. Hence the results of an exact rule cannot be changed, let alone be improved, by repeated apportionment cycles. For rules that are decent and exact we may even insert scaled vote weights $(\alpha x_1, \ldots, \alpha x_\ell)$, with any scale factor $\alpha > 0$, without jeopardizing uniqueness of the solution $(x_1, \ldots, x_\ell)$.

4.3. APPORTIONMENT METHODS

Definition. *An* apportionment method *is an apportionment rule A (Section 4.1) that is anonymous, balanced, concordant, decent, and exact.*

The five organizing principles say nothing about the main issue we are aiming at, the preservation of proportionality. Perfect proportionality would require the existence of a proportionality constant $D > 0$ satisfying $x_j = v_j/D$ for all $j \leq \ell$. On the left-hand side the seat numbers x_j are markedly discrete items. On the right-hand side the quotients v_j/D are practically continuous quantities. Perfect proportionality is generally beyond reach, and we must be satisfied with some sort of approximate proportionality, $x_j \approx v_j/D$. This is where the multitude of apportionment methods comes into play. There are plenty of methods purporting to achieve approximate proportionality in a satisfactory manner.

A family of apportionment methods amenable to a fairly comprehensive analysis are divisor methods. They follow the motto "Divide and round". The essential ingredient to generate a divisor method is a rounding rule $[\![\cdot]\!]$ as introduced in Section 3.8.

4.4. DIVISOR METHODS

Definition. *The* divisor method A that is induced by the rounding rule $[\![\cdot]\!]$ *maps a house size* $h \in \mathbb{N}$ *and a vote vector* $(v_1, \dots v_\ell) \in (0;\infty)^\ell$ *into the set of seat vectors*

$$A(h;v) := \left\{ (x_1,\dots,x_\ell) \in \mathbb{N}^\ell(h) \;\middle|\; x_1 \in \left[\!\!\left[\frac{v_1}{D}\right]\!\!\right], \dots, x_\ell \in \left[\!\!\left[\frac{v_\ell}{D}\right]\!\!\right] \textit{ for some } D > 0 \right\}.$$

In other words the seat numbers x_j are obtained by applying the rounding rule $[\![\cdot]\!]$ to the quotients of the vote weight v_j and some common divisor $D > 0$. The divisor D is such that the seat numbers exhaust the given house size, $x_+ = h$. A divisor method is called *pervious* when the underlying rounding rule is pervious, $s(1) > 0$. It is called *impervious* when the rounding rule is impervious, $s(1) = 0$.

Often the input quantities are not vote counts v_j, but vote shares $w_j = v_j/v_+$. Since vote shares sum to unity, $w_+ = 1$, they must be scaled up to reach the house size h. It is then conducive to talk of *multiplier methods* rather than divisor methods. Divisors D for v_j, and multipliers μ for w_j are related through the evident identity

$$\frac{v_+}{D} w_j = \mu w_j, \qquad \text{that is,} \qquad \mu = \frac{v_+}{D}.$$

Both views could be subsumed under the neutral heading of *scaling methods*, but "divisor methods" is the term firmly established.

If seats were divisible items then the divisor v_+/h, the votes-per-seats ratio, would work and the *ideal shares* of seats $(v_j/v_+)h$ would provide perfectly proportional solutions. But seats are not divisible, and a final rounding step is unavoidable. For this reason divisor methods admit some leeway to adjust the divisor appropriately. This is incorporated into the definition by saying that "some" divisor $D > 0$ will do the job. Nevertheless it is safe to predict that a feasible divisor D lies in the vicinity of the votes-per-seats ratio v_+/h. That is, a feasible multiplier μ for the vote shares w_j will be close to the house size h.

The class of divisor methods comprises five traditional divisor methods of past and present prominence. They go along with the five traditional rounding rules (Section 3.12), and warrant a substitution of the generic symbol A. We propose six-letter identifiers that help memorizing the multitude of methods:

Identifier	Name of method
	Traditional divisor methods
DivDwn	Divisor method with downward rounding
DivStd	Divisor method with standard rounding
DivGeo	Divisor method with geometric rounding
DivHar	Divisor method with harmonic rounding
DivUpw	Divisor method with upward rounding
	Families of divisor methods
DivPwr_p	Divisor method with power-mean rounding, $p \in [-\infty; \infty]$
DivSta_r	Divisor method with stationary rounding, $r \in [0; 1]$

The divisor methods with downward and standard rounding are pervious, those with geometric, harmonic, and upward rounding are impervious.

The remainder of the section verifies that divisor methods are well-defined, and that they obey the five principles required by definition. Verification is straightforward, though a bit lengthy. Skipping these details now incurs no lasting loss later, since the arguments used resurface in various disguises again and again.

First and foremost we show that a divisor method A is well-defined, that is, that the sets $A(h; v)$ are nonempty. Fixing a positive vote vector $v = (v_1, \ldots, v_\ell)$ we use induction on the house size h. We distinguish two cases, whether A is pervious or impervious.

In the pervious case the first signpost is positive, $s(1) > 0$. We begin with house size $h = 0$. Let the parties i whose quotient $v_i/s(1)$ is maximum be assembled in the set

$$I := \left\{ i \le \ell \;\middle|\; \frac{v_i}{s(1)} = \max_{j \le \ell} \frac{v_j}{s(1)} \right\}.$$

We select the divisor to be $D = v_i/s(1)$. The quotients v_j/D of the parties $j \notin I$ are smaller than the signpost $s(1)$ and hence are rounded downwards, $x_j = 0$. The parties $i \in I$ have tied quotients, $v_i/D = s(1)$. The rounding rule offers two rounding options, downwards to zero or upwards to unity. The house size $h = 0$ enforces the first option and excludes the second. Thus $h = 0$ starts out with an apportionment set that is nonempty, $A(0; v) = \{(0, \ldots, 0)\}$. For the induction step we assume that the house size $h \in \mathbb{N}$ has a nonempty apportionment set, $(x_1, \ldots, x_\ell) \in A(h; v)$. The definition of divisor methods guarantees the existence of some divisor $D > 0$ such that for all $j \le \ell$ we have $x_j \in [\![v_j/D]\!]$. The fundamental relation says that the inclusion holds true if and only if $s(x_j) \le v_j/D \le s(x_j + 1)$, or equivalently,

$$\frac{v_j}{s(x_j + 1)} \le D \le \frac{v_j}{s(x_j)}.$$

The inequalities remain intact when D is pushed down to its minimum value, $d := \max_{j \le \ell} v_j / s(x_j + 1)$. Let i be a party attaining the maximum, $v_i/s(x_i + 1) = d$. Its quotient $v_i/d = s(x_i + 1)$ may be rounded downwards or upwards. Since we wish to progress from house size h to house size $h + 1$, we round upwards to obtain the seat numbers $y_i := x_i + 1$ and $y_j := x_j$ for $j \ne i$. The construction identifies the seat vector $(y_1, \ldots, y_\ell)$ as a member of the apportionment set $A(h + 1; v)$. Hence the set is nonempty. The induction proof for the pervious case is complete.

The impervious case has $s(1) = 0 < s(2)$. We begin with house size $h = \ell$. Let the parties i whose quotient $v_i/s(2)$ is maximum be assembled in the set

$$I := \left\{ i \le \ell \;\middle|\; \frac{v_i}{s(2)} = \max_{j \le \ell} \frac{v_j}{s(2)} \right\}.$$

We select the divisor to be $D = v_i/s(2)$. The quotients v_j/D of the parties $j \notin I$ are smaller than the signpost $s(2)$. Hence they are rounded to unity, $x_j = 1$. The parties $i \in I$ have tied quotients, $v_i/D = s(2)$. The rounding rule offers two options, to round downwards to unity or upwards to two. The preordained house size $h = \ell$ enforces the first option and excludes the second. Thus $h = \ell$ starts out with an apportionment set that is nonempty, $A(\ell; v) = \{(1, \ldots, 1)\}$. The inductive step is literally the same as for the pervious case. Hence the induction proof for the impervious case is complete, but covers house sizes $h \ge \ell$ only.

Admittedly the definition is deficient and needs to be amended. It fails to be applicable when the divisor method is impervious and the house size is smaller than the size of the party system, $h < \ell$. The amendment is obvious. The seats are allocated to the h strongest parties, one seat each. We trust that the omission of the deficient instances is a minor sin, in view of their practical irrelevance. From now on we leave these irrelevant instances unattended.

It remains to show that every divisor method fulfills the five organizing principles. (A) A divisor method is anonymous because a rearrangement of vote counts entails the same rearrangement of seat numbers. (B) It is balanced since equal vote counts, $v_j = v_k$, imply equal rounding sets, $[\![v_j/D]\!] = [\![v_k/D]\!]$, that contain at most two consecutive integers. (C) Concordance is a consequence of the fact that all rounding rules are set-monotonic, $v_j > v_k \Rightarrow [\![v_j/D]\!] \ge [\![v_k/D]\!]$ (Section 3.8). (D) Decency is immediate since a scaling of the vote weights v_j into αv_j is matched by scaling the divisors D into αD.

(E) Exactness is established as follows. Let $x = (x_1, \ldots, x_\ell)$ be a seat vector in $\mathbb{N}^\ell(h)$. We have $s(x_j) \le x_j \le s(x_j + 1)$, that is, $x_j \in [\![x_j]\!]$. With divisor $D(x) = 1$ the vector x is seen to be a member of its own apportionment set, $x \in A(h; x)$. Let us assume that the set contains a second seat vector $y \ne x$. If y has divisor $D(y) < 1 = D(x)$ then the components of y are no less than those of x, and the side condition $x_+ = h = y_+$ lets the vectors coincide contradicting the assumption $y \ne x$. A similar argument excludes $D(y) > 1$. Hence we get $D(y) = 1$, too. Because of $x \ne y$ and $x_+ = h = y_+$ there exist two parties $i \ne k$ with $x_i < y_i$ and $x_k > y_k$. But $x_j, y_j \in [\![x_j]\!]$ entails $x_i + 1 = y_i = s(x_i + 1)$ and $x_k - 1 = y_k = s(x_k)$. Hence in the signpost sequence one term hits its right limit and another its left limit, $s(x_i + 1) = x_i + 1$ and $s(x_k) = x_k - 1$. This contradicts the left-right disjunction 3.10.c, whence exactness obtains. Therefore, divisor methods are well-defined and obey the five organizing principles is complete.

The following result offers an easy check whether a candidate seat vector x belongs to the apportionment set of a divisor method under consideration or not. The check circumvents an explicit appeal to divisors. As usual, we set $v_j/0 = \infty$ for $v_j > 0$.

4.5. MAX-MIN INEQUALITY

Theorem. *Let the divisor method A be induced by the rounding rule with signpost sequence $s(0), s(1), s(2)$ etc. Then a seat vector $x \in \mathbb{N}^\ell(h)$ belongs to the apportionment set of a vote vector $v \in (0; \infty)^\ell$,*

$$x \in A(h; v),$$

if and only if

$$\max_{j \le \ell} \frac{v_j}{s(x_j + 1)} \le \min_{j \le \ell} \frac{v_j}{s(x_j)}.$$

Proof. We have $x \in A(h; v)$ if and only if there exists a divisor $D > 0$ satisfying $x_j \in [\![v_j/D]\!]$ for all $j \le \ell$. The fundamental relation $s(x_j) \le v_j/D \le s(x_j+1)$ yields $v_j/s(x_j+1) \le D \le v_j/s(x_j)$. Hence the existence of D entails the Max-Min Inequality. Conversely, every number D between the left maximum and the right minimum is a feasible divisor. □

As mentioned in the previous section it is often instructive to re-interpret divisor methods for vote counts v_j as multiplier methods for vote shares $w_j = v_j/v_+$. Since divisors and multipliers are inversely related, the multiplier version of the Max-Min Inequality reads

$$x \in A(h; w) \qquad \Longleftrightarrow \qquad \max_{j \le \ell} \frac{s(x_j)}{w_j} \le \min_{j \le \ell} \frac{s(x_j+1)}{w_j}.$$

In the remainder of this book we refer to either version as "the" Max-Min Inequality. Whether the reference is to the divisor version in the theorem, or to the multiplier version in the current paragraph, will be clear from the context.

The Max-Min Inequality is extremely fruitful, and at the heart of divisor methods. It designates three sets that are of pertinent importance:

$$D(v,x) := \left[\max_{j \le \ell} \frac{v_j}{s(x_j+1)};\ \min_{j \le \ell} \frac{v_j}{s(x_j)} \right],$$

$$I(v,x) := \left\{ i \le \ell \;\middle|\; \frac{v_i}{s(x_i+1)} = \max_{j \le \ell} \frac{v_j}{s(x_j+1)} \right\},$$

$$K(v,x) := \left\{ k \le \ell \;\middle|\; \frac{v_k}{s(x_k)} = \min_{j \le \ell} \frac{v_j}{s(x_j)} \right\}.$$

The *divisor interval* $D(v,x)$ comprises the divisors feasible for house size h and vote vector v. The *set of increment options* $I(v,x)$ assembles the parties i eligible to receive the $(h+1)$st seat as soon as the divisor falls below the smallest feasible value. The *set of decrement options* $K(v,x)$ consists of the candidate parties k that give up the hth seat when the divisor grows beyond the largest feasible value and only $h-1$ seats are available. Hence a seat vector $x \in A(h; v)$ for house size h allows an immediate passage to seat vectors for house sizes $h+1$ and $h-1$,

$$(x_1, \ldots, x_{i-1}, x_i + 1, x_{i+1}, \ldots, x_\ell) \in A(h+1; v) \quad \text{for all } i \in I(v,x),$$

$$(x_1, \ldots, x_{k-1}, x_k - 1, x_{k+1}, \ldots, x_\ell) \in A(h-1; v) \quad \text{for all } k \in K(v,x).$$

These relations entail a rather efficient way to carry out the calculations.

4.6. JUMP-AND-STEP CALCULATIONS

Given a house size h and a vote vector $v \in (0;\infty)^\ell$ we now tackle the task of calculating the apportionment set $A(h; v)$. Unfortunately no closed formula is available, yet a simple *jump-and-step algorithm* suffices to get the job done. As with all algorithms its initialization is crucial. With a good initialization the jump-and-step algorithm produces the apportionment in a few steps. With a bad initialization, the algorithm takes longer to produce the same answer. It is a misconception to believe that different initializations breed different methods. Nor is somebody who starts from a bad initialization and works diligently towards the result more serious about the problem than somebody else who uses a clever initialization and gets the job finished sooner.

5BT1965	Second Votes	Quotient	DivDwn	Quotient	DivStd
SPD	12 813 186	202.9	202	202.2	202
CDU	12 387 562	196.1	196	195.4997	195
CSU	3 136 506	49.7	49	49.5001	50
FDP	3 096 739	49.03	49	48.9	49
Sum (Divisor)	31 433 993	(63 160)	496	(63 363.6)	496
Divisor interval		[63 119.2; 63 198.7]		[63 363.5; 63 363.7]	

TABLE 4.1 *A narrow divisor interval.* The divisor method with downward rounding has a divisor interval large enough to contain the user-friendly divisor 63 160. The divisor interval of the divisor method with standard rounding is so narrow that all divisors are fractional, such as 63 363.6.

Feasible divisors are likely to be close to the votes-per-seats ratio (Section 4.4). Trusting that v_+/h is a promising divisor initialization we jump to an initial seat vector y with seat numbers $y_j \in [\![(v_j/v_+)h]\!]$. One of following cases a–c applies:

a. The component sum of y exhausts the house size, $y_+ = h$. In this case the initial seat vector is a solution, $x = y$. No further step is needed.

b. The component sum of y stays below the house size, $y_+ < h$. In this case a further seat is handed out to some increment option $i \in I(v, y)$. The increment step is repeated until the incremented seat vector x satisfies $x_+ = h$.

c. The component sum of y exceeds the house size, $y_+ > h$. In this case a seat is retracted from some decrement option $k \in K(v, y)$. The decrement step is repeated until the decremented seat vector x satisfies $x_+ = h$.

The jump-and-step algorithm terminates with a seat vector x in the apportionment set $A(h; v)$. Calculations conclude by selecting a user-friendly divisor D that encourages users to verify the relations $x_j \in [\![v_j/D]\!]$. If the divisor interval is degenerate, $D(v, x) = [D; D]$, then there is no other choice than D. If the divisor interval is non-degenerate, $D(v, x) = [a; b]$ with $a < b$, then every number in the interval would be feasible. The choice is to commercially round the interval's midpoint $(a + b)/2$ to as few significant digits as the interval's interior permits. This selection yields the user-friendly divisors D quoted in Chapters 1 and 2.

The solution may be succinctly paraphrased: *Every D votes justify roughly one seat*. The adverb "roughly" reminds us that the seat numbers $x_1, \ldots, x_\ell$ undergo a final rounding step. True, the user-friendly divisor D is also a rounded quantity. However, the two rounding operations have a distinct meaning. For seat numbers rounding is imperative, they cannot be but whole numbers. For divisors it is not essential and only a matter of aesthetics. The limits a and b of the divisor interval are rounded before they fit to print so that the rounded limits remain inside the interval. The left limit a is rounded upwards to six significant digits (or more if necessary for the sake of clarity), and the right limit b is rounded downwards.

The quoted divisor D is often a multiple of a power of ten. On rare occasions it is unavoidable to quote a fractional divisor. For example, in the 1965 Bundestag election the legal apportionment method was the divisor method with downward rounding. Its divisor interval is [63 119.2; 63 198.7]. The midpoint 63 158.9 may be rounded to become a multiple of ten, 63 160. In contrast, the divisor method with standard rounding yields the divisor interval [63 363.5; 63 363.7]. The interval is so narrow that the best we can do is to quote a divisor with one digit after the decimal point, 63 363.6. See Table 4.1.

Four questions suggest themselves. First, what does the set of *all* solutions look like? So far we have identified only one of its elements, a seat vector x. Section 4.7 scrutinizes the apportionment set $A(h;v)$ in its entirety. Second, how are ties to be handled whenever they occur? Section 4.8 comments on how to resolve ties. Third, how are the calculations carried out in practice? We discuss three initializations and illustrate them with the data of the Austrian EP election 2009: a good initialization (Section 4.9), a better initialization (Section 4.10), and a bad initialization (Section 4.11). Regrettably the bad initialization is the favorite of legislators and political scientists. Fourth, determination of a better initialization deserves a closer look. For stationary divisor methods the recommendation is quite constructive. Given a split $r \in [0;1]$, Section 6.1 argues in favor of the recommended divisor $v_+/\big(h + \ell(r-1/2)\big)$.

4.7. UNIQUENESS, MULTIPLICITIES, AND TIES

Uniqueness of an apportionment result x, and uniqueness of a feasible divisor D are complementary events: The apportionment is unique if and only if the divisor is not unique. That is, the apportionment is not unique if and only if the divisor is unique. Since divisors are characterized by the Max-Min Inequality, the statements are expressed through

$$\{x\} = A(h;v) \qquad \Longleftrightarrow \qquad \max_{j\le\ell} \frac{v_j}{s(x_j+1)} < \min_{j\le\ell} \frac{v_j}{s(x_j)},$$

$$\{x\} \subsetneqq A(h;v) \qquad \Longleftrightarrow \qquad \max_{j\le\ell} \frac{v_j}{s(x_j+1)} = \min_{j\le\ell} \frac{v_j}{s(x_j)}.$$

Evidently the two statements are equivalent; we prove the second.

For the direct implication assume that $x, y \in A(h;v)$ are two distinct solutions, $x \neq y$. Let $D(x)$ be a divisor for x, and $D(y)$ be a divisor for y. If $D(x) > D(y)$, monotonicity entails $x_j \le y_j$ for all $j \le \ell$. Because of equal component sums, $x_+ = h = y_+$, the two vectors coincide, $x = y$, contradicting the assumption that they are distinct. A similar argument excludes $D(x) < D(y)$. Thus both vectors share the same divisor, $D(x) = D(y) = D$, implying $x_j, y_j \in [\![v_j/D]\!]$ for all $j \le \ell$. Because of $x \neq y$ and $x_+ = h = y_+$ there are two components $i \neq k$ with $x_i < y_i$ and $x_k > y_k$. But $x_i, y_i \in [\![v_i/D]\!]$ implies $x_i + 1 = y_i$ and the tie $v_i/D = s(x_i+1)$. Similarly $x_k, y_k \in [\![v_k/D]\!]$ entails $x_k - 1 = y_k$ and the tie $v_k/D = s(x_k)$. Altogether we get

$$D = \frac{v_i}{s(x_i+1)} \le \max_{j\le\ell} \frac{v_j}{s(x_j+1)} \le \min_{j\le\ell} \frac{v_j}{s(x_j)} \le \frac{v_k}{s(x_k)} = D.$$

This establishes equality in the Max-Min Inequality. The direct implication is proved.

For the converse implication we assume that the two sides of the Max-Min Inequality share the same value, D. Since the increment options $i \in I(v,x)$ are tied, $v_i/D = s(x_i+1)$, they round to $[\![v_i/D]\!] = \{x_i, x_i+1\}$. The decrement options $k \in K(v,x)$ are also tied, $v_k/D = s(x_k)$, and round to $[\![v_k/D]\!] = \{x_k-1, x_k\}$. In the present case the option sets are disjoint, $I(v,x) \cap K(v,x) = \emptyset$, since otherwise some component j with $v_j/s(x_j+1) = D = v_j/s(x_j)$ would violate the fact that all signpost sequences are strictly increasing. Hence we may fix two distinct components, $i \in I(v,x)$ and $k \in K(v,x)$, and define the seat vector y through $y_i := x_i + 1$, $y_k := x_k - 1$, and $y_j := x_j$ for all $j \neq i, k$. Then y is distinct from x, but still a member of the apportionment set $A(h;v)$. This proves the converse implication.

(EP2009AT)	Votes	Quotient	Fifteen equally justified apportionments, DivDwn														
			#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
A	840 000	6	6−	6−	6−	6−	6−	5+	5+	5+	5+	5+	5+	5+	5+	5+	5+
B	700 000	5	5−	4+	4+	4+	4+	5−	5−	5−	5−	4+	4+	4+	4+	4+	4+
C	560 000	4	3+	4−	3+	3+	3+	4−	3+	3+	3+	4−	4−	4−	3+	3+	3+
D	420 000	3	2+	2+	3−	2+	2+	2+	3−	2+	2+	3−	2+	2+	3−	3−	2+
E	280 000	2	1+	1+	1+	2−	1+	1+	1+	2−	1+	1+	2−	1+	2−	1+	2−
F	140 000	1	0+	0+	0+	0+	1−	0+	0+	0+	1−	0+	0+	1−	0+	1−	1−
Sum(Div.)	2 940 000	(140 000)	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17

TABLE 4.2 *Manufactured ties.* All parties are tied since the quotients of votes and divisor 140 000 hit some downward-rounding signpost, $s(n) = n$. Trailing plus-signs indicate increment options (4), trailing minus-signs decrement options (2). There are $\binom{4+2}{4} = 15$ equally justified apportionments.

The proof of the converse implication also shows how an identified member x relates to the other seat vectors in the apportionment set $A(h; v)$. Let $a := \#I(v, x)$ be the number of increment options of x, and let $b := \#K(v, x)$ be the number of decrement options. The seat vector x has $a + b$ parties tied to signposts. When all parties are rounded to their higher level, the resulting component sum is $h + a$ and exceeds the target value by a seats. Hence a of the $a + b$ ties must be resolved at the lower level, and the remaining b ties at the upper level. Thus the cardinality of the apportionment set is given by the binomial coefficient,

$$\#A(h; x) = \binom{a+b}{a}.$$

Table 4.2 pushes the example to an extreme by having all vote counts tied to a signpost. To compactify the notation we append to a natural number n a trailing plus-sign to indicate an upward tie, $n+ := \{n, n+1\}$, or a trailing minus-sign to indicate a downward tie, $n- := \{n-1, n\}$. With four increment options and two decrement options, the apportionment set DivDwn$(17, v)$ consists of the $\binom{6}{2} = 15$ equally justified seat vectors that are enumerated in the table. Each seat vector with its trailing plus- or minus-signs is sufficiently informative to recover the other fourteen.

4.8. RESOLUTION OF TIES

In popular elections vote counts reach into the thousands or millions, and ties occur rather rarely. Yet many electoral laws include a provision how to resolve ties. The provision contains a message beyond dealing with a rare oddity. It documents the legislator's intention to design an electoral system such that the seat apportionment is definitive and determinate.

A common tie resolution procedure is to draw a lot. The example in Table 4.2 shows that the lot may be cast over the equally justified apportionments (in Table 4.2: fifteen), or by choosing the required number of parties from those in the tie group by simple random sampling without replacement (four out of six, in this example). There is a variety of other tie resolution procedures, particularly when a single seat is tied between two parties: the seat goes to the larger party, or to the smaller party, or to the party that registered earlier, or to the party with the oldest candidate, or with the candidate with the most personal votes, or with the candidate with the most children. Other provisions grant the election officer a deciding vote, or call a run-off election.

2006UsterZhCH	Party Votes	Quotient	DivStd	Voter Count	Quotient	DivStd
01 SP	71 847	10.7	11	1 996	10.7	11
02 SVP	64 728	9.6	10	1 798	9.7	10
03 FDP	36 613	5.4	5	1 017	5.47	5
04 EVP	13 600	2.0	2	378	2.0	2
05 CVP	9 756	1.45	1	271	1.46	1
06 SD	5 745	0.9	1	160	0.9	1
07 EDU	3 353	0.499	0	93	0.5	1–
08 SEDU	3 752	0.6	1	104	0.6	1
09 GP	13 369	2.0	2	371	2.0	2
10 GLP	16 476	2.45	2	458	2.46	2
11 JEDU	3 365	0.501	1	93	0.5	0+
Sum (Divisor)	242 604	(6720)	36	6 739	(186)	36

TABLE 4.3 *Uster, tied 2006 city council election.* With reference to voter counts a tie emerges between the seventh and eleventh list, EDU and JEDU. By taking recourse to party votes, the tie is resolved in favor of the JEDU that outperforms the EDU by a margin of 12 party votes.

A tie problem emerged during the 2006 election of the 36-seat city council of the City of Uster in the Swiss Canton of Zurich. The particulars of the incident entail a particular tie resolution decision. The pertinent electoral law governs all elections in the Canton of Zurich, including some with multiple districts. The number of seats to be filled in a district is known as the district magnitude. In each district, a voter may mark as many candidates on the ballot sheet as is given by the preordained district magnitude. Summing raw votes over multiple districts would aim at equality among ballot marks, not at equality among voters. In order to measure the support of a party with reference to human beings and not with reference to ballot marks, a party's vote total is converted into the party's *voter count* (Wählerzahl). It is defined to be the commercially rounded quotient of party votes and district magnitude,

$$\text{voter count} = \left\langle \frac{\text{party votes}}{\text{district magnitude}} \right\rangle .$$

For example, in Table 4.3 the SP voter count is $\langle 71\,847/36\rangle = \langle 1\,995.8\rangle = 1\,996$. The formula does not allow for voters not exhausting all the votes they have, or spreading their votes over several lists (panachage). Judging from past elections in Zurich, there is no evidence that cross voting entails a bias of any significance. The proposed voter count definition seems to serve all practical needs.

The law decrees the conversion of party votes into voter counts even when there is just a single district, as is the case in Uster. It so happens that a seat is tied between EDU and JEDU, in terms of voter counts. Both lists have voter count 93. The seats are apportioned using the divisor method with standard rounding. With divisor 186, gave both lists obtain the interim quotient 0.5. Hence one of the two quotients has to be rounded upwards, the other one downwards. See Table 4.3.

In the 2006 election, the election officer cast a lot and gave the seat to the EDU. An attentive voter complained by pointing out that, in terms of party votes, the JEDU ranked above the EDU by a margin of 12 votes. The electoral authorities overruled the lot decision, and awarded the tied seat to the JEDU.

4.9. GOOD INITIALIZATION OF THE JUMP-AND-STEP CALCULATIONS

In the next four sections we illustrate the calculatory scheme for divisor methods with the data from Table 1.3, the 2009 EP election in Austria. The example uses the divisor method with downward rounding; its signposts are $s_1(n) = n$, $n \geq 0$. We stick to the notation $s_1(n)$, thus indicating how to handle other signpost sequences. Three variants of the jump-and-step algorithm of Section 4.6 are presented. They differ in their divisor initializations. Variant 1 initializes the divisor with the votes-per-seats ratio, $D = v_+/h$. The Austrian data call for two increment steps to obtain the final seat vector. Variant 2 uses the recommended initialization, $D = v_+/(h + \ell/2)$, and instantly jumps to the final result (Section 4.10). Variant 3 starts out with divisor $D = \infty$. From start to finish it needs a maximum of increment steps, seventeen (Section 4.11). Variant 3 may be compressed into a table of highest comparative figures (Section 4.12). Use of comparative figures needs less ink for the print, but an advanced understanding of what the print means.

Variant 1 is displayed below. The divisor is initialized with the votes-per-seats ratio, $D = 2\,825\,027/17 = 166\,178.1$. When the counts in the "Votes" line are divided by D, the quotients in the next line are obtained. Rounding them downwards we jump to the "Initial Seats" vector $y = (5, 4, 3, 2, 1, 0)$. It stays below the house size by two seats, $y_+ = 15 < 17$, and thus calls for two increment steps.

According to Section 4.5 the increment options $I(v, y)$ are identified by computing the figures $v_j/s(y_j + 1)$. The computation is documented in the line labeled "Seat 16". Then these figures are compared. The highest comparative figure is marked with a dot (•). It tells us to allocate the sixteenth seat with the ÖVP.

The increment from the sixteenth to the seventeenth seat follows the same recipe, but is less laborious. Undotted comparative figures are copied, from the "Seat 16" line into the "Seat 17" line. Only the dotted ÖVP figure needs to be re-calculated. The highest comparative figure in the line "Seat 17", again dotted, allocates the seventeenth seat with the GRÜNE. The "Final Seats" vector $x = (6, 4, 3, 2, 2, 0)$ meets the house size, $x_+ = 17$. It is the solution sought. For those who trust that all calculations are correct the job is done.

Variant 1	ÖVP	SPÖ	Martin	FPÖ	GRÜNE	BZÖ	Sum (Divisor)
Votes	858 921	680 041	506 092	364 207	284 505	131 261	2 825 027
Quotient	5.2	4.1	3.05	2.2	1.7	0.8	(166 178.1)
Initial Seats	5	4	3	2	1	0	15
Seat 16	/(5+1)= 143 153.5•	/(4+1)= 136 008.2	/(3+1)= 126 523	/(2+1)= 121 402.3	/(1+1)= 142 252.5	/(0+1)= 131 261	*Increment:* ÖVP
Seat 17	/(6+1)= 122 703	136 008.2	126 523	121 402.3	142 252.5•	131 261	GRÜNE
Final Seats	6	4	3	2	2	0	17
Appendix	122 703	136 008.2•	126 523	121 402.3	/(2+1)= 94 835	131 261	

For those who wish to check the result a line "Appendix" is added. If there were an eighteenth seat, it would go to the SPÖ. The implied lower limit of the divisor interval is relevant, not the seat itself. The interval extends from $680\,041/5 = 136\,008.2$ to $284\,505/2 = 142\,252.5$. Its midpoint is $(136\,008.2 + 284\,505/2)/2 = 139\,130.4$. The midpoint is rounded to two significant digits to stay in the interval's interior, $D = 140\,000$. Hence the solution may be paraphrased by saying that every 140 000 votes justify roughly—up to downward rounding—one seat. This is the information conveyed in Table 1.3.

The increment steps compare the figures $v_j/s_1(x_j+1)$. Since the divisor method with downward rounding has signposts $s_1(n) = n$ a typical comparative figure happens to turn into $v_j/(x_j+1)$. This quotient has the form of an average of votes and seats, under the hypothesis that party j increases its current seat number x_j by being awarded the next seat. Actually, the next seat is awarded to the party whose votes-per-(seats+1) average is highest. For this reason the divisor method with downward rounding is also called the *highest average method*.

For a general divisor method the ratio $v_j/s(x_j+1)$ fails to be interpretable as a votes-per-seats average, however, and cannot be addressed as such. Instead we refer to the ratio $v_j/s(x_j+1)$ as a *comparative figure*, as is common practice in Sweden.

4.10. RECOMMENDED INITIALIZATION OF THE JUMP-AND-STEP CALCULATIONS

Variant 2 starts out with the initialization recommended in Section 6.1, $D = 2\,825\,027/(17+6/2) = 141\,251.4$. The interim quotients v_j/D are rounded downwards, and yield the initial seat vector $y = (6,4,3,2,2,0)$. Since it fits, $y_+ = 17$, the initial seat vector is the final solution, $x = y$. No increment or decrement steps are needed. For those who trust that the calculations are error-free the job is done.

Variant 2	ÖVP	SPÖ	Martin	FPÖ	GRÜNE	BZÖ	Sum (Divisor)
Votes v_j	858 921	680 041	506 092	364 207	284 505	131 261	2 825 027
Quotient	6.1	4.8	3.6	2.6	2.01	0.9	(141 251.4)
Seats x_j	6	4	3	2	2	0	17
$v_j/(x_j+1)$	122 703	136 008.2•	126 523	121 402.3	94 835	131 261	(• = max)
v_j/x_j	143 153.5	170 010.3	168 697.3	182 103.5	142 252.5•	∞	(• = min)

For those who wish to check the result two lines are appended. In the line labeled "$v_j/(x_j+1)$" the highest figure is marked. This is the critical divisor where the house size increases by one seat, from 17 to 18. In line "v_j/x_j" the lowest figure is marked. This is the critical divisor where the house size decreases by one seat, from 17 down to 16. The marked quantities provide the limits of the divisor interval, [136 008.2; 142 252.5]. Again the user-friendly divisor is determined by rounding the interval midpoint 139 130.35 to two significant digits, $D = 140\,000$. This is the information presented in Table 1.3.

4.11. BAD INITIALIZATION OF THE JUMP-AND-STEP CALCULATIONS

Variant 3 carries out an increment process of marathon length, from no seat to all seats. Effectively it starts with a huge initial divisor $D = \infty$, whence initially all seat numbers are rounded downwards to zero. For a start, nobody gets anything. From there on the calculations work their way via $1, 2, 3, \ldots$ seats up to the target house size 17. The determination of the divisor interval [136 009; 142 252] and the selection of the divisor $D = 140\,000$ proceed as in Section 4.9.

Variant 3 Votes	ÖVP 858 921	SPÖ 680 041	Martin 506 092	FPÖ 364 207	GRÜNE 284 505	BZÖ 131 261	Sum 2 825 027
Seat 1	/(0 + 1) = 858 921•	/(0 + 1) = 680 041	/(0 + 1) = 506 092	/(0 + 1) = 364 207	/(0 + 1) = 284 505	/(0 + 1) = 131 261	*Increment:* ÖVP
Seat 2	/(1 + 1) = 429 460.5	680 041•	506 092	364 207	284 505	131 261	SPÖ
Seat 3	429 460.5	/(1 + 1) = 340 020.5	506 092•	364 207	284 505	131 261	Martin
Seat 4	429 460.5•	340 020.5	/(1 + 1) = 253 046	364 207	284 505	131 261	ÖVP
Seat 5	/(2 + 1) = 286 307	340 020.5	253 046	364 207•	284 505	131 261	FPÖ
Seat 6	286 307	340 020.5•	253 046	/(1 + 1) = 182 103.5	284 505	131 261	SPÖ
Seat 7	286 307•	/(2 + 1) = 226 680.3	253 046	182 103.5	284 505	131 261	ÖVP
Seat 8	/(3 + 1) = 214 730.3	226 680.3	253 046	182 103.5	284 505•	131 261	GRÜNE
Seat 9	214 730.3	226 680.3	253 046•	182 103.5	/(1 + 1) = 142 252.5	131 261	Martin
Seat 10	214 730.3	226 680.3•	/(2 + 1) = 168 697.3	182 103.5	142 252.5	131 261	SPÖ
Seat 11	214 730.3•	/(3 + 1) = 170 010.3	168 697.3	182 103.5	142 252.5	131 261	ÖVP
Seat 12	/(4 + 1) = 171 784.2	170 010.3	168 697.3	182 103.5•	142 252.5	131 261	FPÖ
Seat 13	171 784.2•	170 010.3	168 697.3	/(2 + 1) = 121 402.3	142 252.5	131 261	ÖVP
Seat 14	/(5 + 1) = 143 153.5	170 010.3•	168 697.3	121 402.3	142 252.5	131 261	SPÖ
Seat 15	143 153.5	/(4 + 1) = 136 008.2	168 697.3•	121 402.3	142 252.5	131 261	Martin
Seat 16	143 153.5•	136 008.2	/(3 + 1) = 126 523	121 402.3	142 252.5	131 261	ÖVP
Seat 17	/(6 + 1) = 122 703	136 008.2	126 523	121 402.3	142 252.5•	131 261	GRÜNE
Final Seats	6	4	3	2	2	0	17
Appendix	122 703	136 008.2•	126 523	121 402.3	/(2 + 1) = 94 835	131 261	

With almost twenty seats this example requires a table of about a page's length. The seats of the German EP allocation (Table 1.8) would fill five pages, and the roughly six hundred seats in the German Bundestag (Table 2.1) thirty. The space required by Variant 3 grows linearly with the house size h. In contrast the complexities of Variants 1 and 2 do not depend on the house size h at all (Section 6.2). They are bounded by the number of participating parties or half of it, ℓ (Variant 1) or $\ell/2$ (Variant 2), see Sections 6.2 and 6.3.

4.12. HIGHEST COMPARATIVE FIGURES

The excessive growth of Variant 3 may be curbed by leaping over the repetitive appearances of non-dotted and hence non-active comparative figures. Only the dotted figures are relevant. Hence Variant 3 boils down to the sufficient extract shown below. Variant 3A is obtained by elevating the former table body (the comparative figures $v_j/s(x_j+1)$) to become table rows, and by moving the former rows (seats) into the current table body. In the table body the seventeen seats are allocated in the succession of highest comparative figures as indicated by the italicized numbers *1, ..., 17*:

Variant 3A	ÖVP	SPÖ	Martin	FPÖ	GRÜNE	BZÖ
Votes v_j	858 921	680 041	506 092	364 207	284 505	131 261
$v_j/1$	858 921-*1*	680 041-*2*	506 092-*3*	364 207-*5*	284 505-*8*	131 261
$v_j/2$	429 461-*4*	340 021-*6*	253 046-*9*	182 104-*12*	142 257-*17*	
$v_j/3$	286 307-*7*	226 680-*10*	168 697-*15*	121 402	94 835	
$v_j/4$	214 730-*11*	170 010-*14*	126 523			
$v_j/5$	171 784-*13*	136 008				
$v_j/6$	143 154-*16*					
$v_j/7$	122 703					
Seats	6	4	3	2	2	0

Variant 3A compactifies the information, but its growth is still linear in the house size. It needs about w_1h rows of comparative figures, where w_1 is the vote share of the strongest party and h is the house size. Variant 3A is a mechanical recipe conveying only little insight into the structure of apportionment methods. If you do not know what is going on, Variant 3A will not teach you.

4.13. AUTHORITIES

Experts often prefer a somewhat cryptic jargon by naming a seat apportionment method after an authority who fought for it. Unfortunately there exists no international agreement who deserves the honor most. For example the divisor method with downward rounding is the *Jefferson method* or the *D'Hondt method* or the *Hagenbach-Bischoff method*, you choose. Here is a list of celebrities associated with the five traditional divisor methods:

DivDwn *Thomas Jefferson* (1743–1826), principal author of the US Declaration of Independence, third US President 1801–1809
Victor D'Hondt (1841–1901), Professor of Law, Ghent University, co-founder of the Belgian L'Association réformiste pour l'adoption de la Représentation Proportionnelle 1881
Eduard Hagenbach-Bischoff (1833–1910), Professor of Physics, University of Basel, and cantonal politician

DivStd *Daniel Webster* (1782–1852), US statesman, Senator from Massachusetts, US Secretary of State
Jean-André Sainte-Laguë (1882–1950), Professor of Mathematics, Conservatoire national des arts et métiers, Paris
Hans Schepers (b. 1928), Physicist, Head of the Data Processing Unit, Scientific Services of the German Bundestag

DivGeo *Joseph Adna Hill* (1860–1938), Statistician, Assistant Director of the Census, US Bureau of the Census
Edward Vermilye Huntington (1874–1952), Professor of Mathematics, Harvard University, Cambridge, Massachusetts

DivHar *James Dean* (1776–1849), Professor of Astronomy and Mathematics, University of Vermont, Burlington, Vermont

DivUpw *John Quincy Adams* (1767–1848), US diplomat and statesman, sixth US President 1825–1829

Chapter 5 introduces another important family of apportionment methods, quota methods. In a way quota methods may be viewed as procedures complementary to divisor methods. Divisor methods fix the rounding rule and adjust the divisor. Quota methods fix the divisor and adjust the rounding rule.

CHAPTER 5

Quota Methods of Apportionment: Divide and Rank

Another important family of apportionment methods is quota methods. Relying on a fixed divisor of some intrinsic persuasiveness, called quota, they follow the motto "Divide and rank". The most prominent member of the family, the Hare-quota method with residual fit by greatest remainders, is discussed in the first part of the chapter. The second part addresses various variants of the quota, and various variants of the residual apportionment step. As a whole, the family of quota methods offers a more eclectic approach to apportionment problems than the family of divisor methods.

5.1. QUOTA METHODS

Quota methods solve the same seat apportionment problem to which divisor methods are applied: h seats are to be apportioned among parties $j \le \ell$ in proportion to their vote counts $v_j > 0$ (Section 4.1). Again a divisor is used to downscale the vote counts into the vicinity of the final seat numbers x_j. The point is that quota methods consider the divisor to be fixed, justified by its intrinsic persuasiveness. To emphasize this point the fixed divisor is called *quota*, and designated by the letter Q. A quota that appears compelling to some may appear debatable to others. Section 5.8 lists a variety of quotas that have found their way into electoral laws in former and present times.

Quota methods are two-step procedures. The first step is called the *main apportionment*. It calculates an interim quotient v_j/Q, and apportions its integral part $y_j = \lfloor v_j/Q \rfloor$ to party j. Let $m := y_+$ denote the number of seats allotted by the main apportionment. The quota Q is supposed to be such that it hands out no more seats than are available, $m \le h$, and that it misses no more than one seat per party, $h - m \le \ell$. The second step, the *residual fit*, allocates the $h - m \in \{0, \ldots, \ell\}$ residual seats left. The most popular procedure is the residual fit *by greatest remainders*. It ranks the fractional parts $(v_j/Q) - \lfloor v_j/Q \rfloor$ by decreasing size, and allocates one seat to each of the $h - m$ greatest remainders. For this reason quota methods are captured by the motto "Divide and rank".

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_5,

5.2. HARE-QUOTA METHOD WITH RESIDUAL FIT BY GREATEST REMAINDERS

The dominant quota method is the *Hare-quota method with residual fit by greatest remainders*. We abbreviate it by HaQgrR. The quota used is the votes-per-seats ratio, v_+/h, called *Hare-quota* in the context of quota methods. The induced quotients with the vote counts are the ideal shares of seats, $v_j/(v_+/h) = (v_j/v_+)h$.

The main apportionment allocates to party $j \leq \ell$ the integral part of its ideal share of seats, $y_j := \lfloor (v_j/v_+)h \rfloor$. Hence the number of seats accounted for in the main apportionment stays below the house size,

$$m = \sum_{j \leq \ell} \left\lfloor \frac{v_j}{v_+} h \right\rfloor \leq \sum_{j \leq \ell} \frac{v_j}{v_+} h = h.$$

The main apportionment exhausts the house size, $m = h$, if and only if all vote counts v_j are integer multiples of the Hare-quota v_+/h. As this is highly unlikely to happen, the method usually enters the second step, the fit of $h - m$ residual seats.

Since all fractional parts are strictly less than unity, $(v_j/v_+)h - \lfloor (v_j/v_+)h \rfloor < 1$, the number of residual seats is strictly smaller than the number of parties,

$$h - m = \sum_{j \leq \ell} \left(\frac{v_j}{v_+} h - \left\lfloor \frac{v_j}{v_+} h \right\rfloor \right) < \ell.$$

Since the numbers involved are integers the inequality tightens to $h - m \leq \ell - 1$. That is, the main apportionment takes care of at least $h + 1 - \ell$ seats. In most practical applications the house size is much larger than the number of parties whence the vast majority of seats is dealt with.

For example in the German Bundestag with $h = 598$ seats and $\ell = 6$ parties (Table 2.1), the main apportionment step would settle at least 593 seats, and hence would leave at most 5 seats for the residual fit. In case of the Bulgarian EP election where the method was applied with $h = 17$ and $\ell = 6$ (Table 1.5), the main apportionment step is guaranteed to account for at least 12 seats, and to leave at most 5 seats. With the 2009 data it actually allocates 13 seats in the main apportionment, and leaves 4 seats for the residual fit.

The second step, the residual fit by greatest remainders, ranks the ideal shares' fractional parts, and adds for the $h - m$ largest of them one seat to the preliminary seat numbers of the main apportionment. Thus the final seat numbers are $x_i = y_i + 1$ for the parties i with a larger remainder, and $x_k = y_k$ for the parties k with a smaller remainder. Ties emerge if and only if several parties share the same remainder and this particular remainder is the split to separate larger remainders from smaller remainders.

The Hare-quota method with residual fit by greatest remainders is anonymous, balanced, concordant, decent, and exact, as is easily verified. Hence it qualifies as an apportionment method in the sense of Section 4.3.

5.3. GREATEST REMAINDERS CALCULATIONS

We illustrate the pertinent calculations with the Bulgarian 2009 EP election. The main apportionment rests on the ideal shares of seats listed in the column “Quotient’ in Table 1.5. Their integral parts provide the preliminary seat numbers y_j:

$$y_{\text{GERB}} = 4,\ \ y_{\text{BSP}} = 3,\ \ y_{\text{DPS}} = 2,\ \ y_{\text{ATAKA}} = 2,\ \ y_{\text{NDSV}} = 1,\ \ y_{\text{SDS-DSB}} = 1.$$

Thus $m = 13$ of the $h = 17$ seats are dealt out in the main apportionment, and four are left to be allocated in the residual fit. After ranking the ideal shares’ fractional parts $f(j) := (v_j/v_+)h - \lfloor (v_j/v_+)h \rfloor$ in decreasing order, the four parties with the largest remainders have their preliminary seat numbers increased by one seat, while the two smallest parties stay as is:

$$f(\text{GERB}) = .880, f(\text{DPS}) = .832, f(\text{BSP}) = .706, f(\text{NDSV}) = .595 \Rightarrow x_i = y_i + 1,$$
$$f(\text{SDS-DSB}) = .592, f(\text{ATAKA}) = .395 \Rightarrow x_k = y_k.$$

This verifies the seat numbers in the last column of Table 1.5:

$$x_{\text{GERB}} = 5,\ \ x_{\text{BSP}} = 4,\ \ x_{\text{DPS}} = 3,\ \ x_{\text{ATAKA}} = 2,\ \ x_{\text{NDSV}} = 2,\ \ x_{\text{SDS-DSB}} = 1.$$

The same result is obtained using the stationary divisor method with split $r = .594$. In fact, every split from the interval

$$[f(\text{SDS-DSB}); f(\text{NDSV})] = [.592\,432; .594\,989]$$

separates the fractional parts of those that do not profit from the residual fit, from those that do. The strategy for selecting a user-friendly split r is the same as selecting a communicable divisor D (Section 4.6), with one extra rule. If the interval happens to contain the value .5, then we select $r = .5$; the extra rule emphasizes that standard rounding would do the job. Otherwise the midpoint of the interval is rounded to as few significant digits as the interval’s interior permits. In the Bulgarian example we thus publish the split value .594, as shown in the bottom line of Table 1.5.

The examples indicates that the results of the Hare-quota method with residual fit by greatest remainders can be replicated by means of some stationary divisor method. The replication approach extends to the large class of quotas called shift-quotas. They perturb the Hare-quota just a little bit so that the second apportionment step can still be carried out via a residual fit by greatest remainders.

5.4. SHIFT-QUOTA METHODS

The *shift-quota* with *shift* $s \in [-1; 1)$, denoted by $Q(s)$, is defined by

$$Q(s) := \frac{v_+}{h+s}.$$

The *shift-quota method* with residual fit by greatest remainders and with shift s is abbreviated as shQgrR_s. The shift $s = 0$ retrieves the Hare-quota method with residual fit by greatest remainder, $\text{shQgrR}_0 = \text{HaQgrR}$.

Although there is no practical interest in negative shifts, $s < 0$, all shift-quotas $Q(s)$, $s \in [-1;1)$, allow the main apportionment to be paired with a residual fit by greatest remainders. Indeed, the main allocation for party j is

$$y_j = \left\lfloor \frac{v_j}{Q(s)} \right\rfloor = \left\lfloor \frac{v_j}{v_+}\Big(h+s\Big) \right\rfloor .$$

Hence the main apportionment allocates at most $y_+ \le \sum_{j\le\ell}(v_j/v_+)(h+s) = h+s < h+1$ seats. Because of integrality the inequality tightens to $y_+ \le h$. The lower bound $y_+ > \sum_{j\le\ell}\big((v_j/v_+)(h+s)-1\big) = h+s-\ell \ge h-\ell-1$ tightens to $y_+ \ge h-\ell$. The range $h-\ell \le y_+ \le h$ confirms feasibility of a residual fit by greatest remainders.

The next theorem presents a Max-Min Inequality similar to that for divisor methods (Section 4.5). The theorem's charm lies in avoiding a direct appeal to remainders.

5.5. MAX-MIN INEQUALITY

Theorem. *Consider a shift-quota method $shQgrR_s$ with shift $s \in [-1;1)$. Then a seat vector $x \in \mathbb{N}^\ell(h)$ belongs to the apportionment set of a vote vector $v \in (0;\infty)^\ell$,*

$$x \in shQgrR_s(h;v),$$

if and only if

$$\max_{j\le\ell}\left(\frac{v_j}{v_+}\Big(h+s\Big) - x_j\right) \le \min_{j\le\ell}\left(\frac{v_j}{v_+}\Big(h+s\Big) + 1 - x_j\right).$$

Proof. It is convenient to switch to the vote shares $w_j = v_j/v_+$. The shift-quota $Q(s) = v_+/(h+s)$ then yields interim quotients $v_j/Q(s) = w_j(h+s)$.

For the proof of the direct implication, let x be a solution vector in $\mathrm{shQgrR}_s(h;v)$. We assemble the parties that receive no seat in the residual apportionment in the set I. The complement $K := I'$ comprises the parties that are awarded one of the residual seats,

$$I := \Big\{ i \le \ell \;\Big|\; x_i = \lfloor w_i(h+s)\rfloor \Big\}, \qquad K := \Big\{ k \le \ell \;\Big|\; x_k = \lfloor w_k(h+s)\rfloor + 1 \Big\}.$$

In the left maximum of the Max-Min Inequality parties $i \in I$ have a nonnegative difference, $w_i(h+s) - x_i = w_i(h+s) - \lfloor w_i(h+s)\rfloor =: f(i) \ge 0$, the fractional part of the interim quotient $w_i(h+s)$. Parties $k \in K$ have a negative difference and drop out. Hence we get

$$\max_{j\le\ell}\Big(w_j(h+s)-x_j\Big) = \max_{i\in I}\Big(w_i(h+s)-x_i\Big) = \max_{i\in I} f(i).$$

The right minimum equals $\min_{j\le\ell}\Big(w_j(h+s)+1-x_j\Big) = \min_{k\in K} f(k)$. The residual fit by greatest remainders entails that the maximum stays below the minimum, $\max_{i\in I} f(i) \le \min_{k\in K} f(k)$.

For the proof of the converse implication, the Max-Min Inequality is spelled out as

$$w_i(h+s) - x_i \le w_k(h+s) + 1 - x_k \qquad \text{for all } i,k \le \ell.$$

Summation over $k \ne i$ gives $(\ell-1)\Big(w_i(h+s)-x_i\Big) \le (1-w_i)(h+s) + (\ell-1) - (h-x_i)$, for all $i \le \ell$. This simplifies to $w_i(h+s) - x_i \le (\ell-1+s)/\ell < 1$. Summation over $i \ne k$ yields an inequality leading to $0 \le w_k(h+s)+1-x_k$, for all $k \le \ell$. Together we get

$$w_j(h+s) - 1 < x_j \le w_j(h+s) + 1 \qquad \text{for all } j \le \ell.$$

The inequality string leaves just two possibilities, $x_j = \lfloor w_j(h+s)\rfloor$ or $x_j = \lfloor w_j(h+s)\rfloor + 1$. Let the set I assemble the parties i with $x_i = \lfloor w_i(h+s)\rfloor$, and $K := I'$ the others. The first part of the proof shows that the Max-Min Inequality takes the form $\max_{i\in I} f(i) \le \min_{k\in K} f(k)$. Hence the parties $i \in I$ that are stuck with their seats from the main apportionment have an interim quotient with a fractional part less than or equal to that of the parties $k \in K$ that are awarded one of the residual seats. This establishes $x \in \mathrm{shQgrR}_s(h;v)$. □

The Max-Min Inequality gives rise to the *split interval*

$$R(v,x) := \left[\max_{j\le\ell}\left(\frac{v_j}{v_+}(h+s)-x_j\right);\ \min_{j\le\ell}\left(\frac{v_j}{v_+}(h+s)+1-x_j\right)\right].$$

The preceding proof reveals that $R(v,x)$ assembles the splits $r \in [0;1]$ that separate the fractional part $f(i)$ of the interim quotient of a party i that has to make do with the main apportionment, $x_i = \lfloor v_i/Q(s)\rfloor$, from the fractional part $f(k)$ of the interim quotient of a party k that gets one of the residual seats, $x_k = \lfloor v_k/Q(s)\rfloor + 1$. The attitude is complementary to divisor methods. Divisor methods fix the rounding rule and adjust the divisor. Quota methods fix the divisor and adjust the rounding rule.

5.6. SHIFT-QUOTA METHODS AND STATIONARY DIVISOR METHODS

Corollary. *Consider a shift-quota method* $shQgrR_s$ *with shift* $s \in [-1;1)$. *For all house sizes* $h \in \mathbb{N}$ *and for all vote vectors* $(v_1,\ldots,v_\ell) \in (0;\infty)^\ell$ *there is a split* $r^* \in [0;1]$, *generally depending on* s, h, *and* v, *such that the shift-quota method with shift* s *and the stationary divisor method with split* r^* *have the same solution sets,*

$$shQgrR_s(h;v) = DivSta_{r^*}(h;v).$$

Proof. The proof of the Max-Min Inequality 5.5 shows that, given an arbitrary seat vector $x \in \mathrm{shQgrR}_s(h;v)$, every split r^* in the split interval $R(v,x)$ establishes the assertion. □

The corollary mitigates the motto "Divide and rank". The ranking of parties by decreasing fractional parts of their interim quotients is no more than a transient step of the calculations. The persisting step is the partition of parties into a group with smaller fractional parts where quotients are rounded downwards, and a complementary group with larger fractional parts where quotients are rounded upwards. The partition is succinctly specified by publishing a split r^*. Once r^* is made known it is no longer needed to establish the ranking of the fractional parts. The motto "Divide and split" would subsume quota methods more pointedly.

5.7. AUTHORITIES

The Hare-quota method with residual fit by greatest remainders carries the name of *Thomas Hare* (1806–1891), an English barrister and proponent of proportional representation systems. In his writings *Hare* repeatedly referred to the votes-per-seats ratio; hence it may rightly be called the Hare-quota. However, the system that *Hare* fought for was a single transferable vote (STV) scheme. In terms of achieving proportionality the STV scheme is close to the apportionment methods discussed in this book. Nevertheless, the philosophy underlying STV schemes is rather different from quota methods, as are ballot structure and vote counting. Hence referring the "Hare-quota method with residual fit by greatest remainders" to *Hare* is a misnomer of sorts.

In the United States of America the Hare-quota method with residual fit by greatest remainders is called the Hamilton method, after *Alexander Hamilton* (1755–1804). *Hamilton* successfully proposed the method to the House of Representatives, only to then see it vetoed by President George Washington. In Germany the method is associated with the name of the mathematician *Horst Friedrich Niemeyer* (1931–2007).

The Hare-quota method with residual fit by greatest remainders often comes under the alternative name of *largest remainder method* (or LR method, for short). The name emphasizes the second step of the method, the residual fit by greatest remainders. It entirely neglects the main apportionment step although this is where most of the seats are apportioned. For this reason we maintain the acronyms of the type HaQgrR, even though they are somewhat bulky. Proponents of the short name LR method would presumably argue that there is no need to explicate the main apportionment step because it is natural and self-evident. History teaches otherwise.

5.8. QUOTA VARIANTS

A "quota", as the term is used by *Hare*, signifies a number of voters who justify the allocation of a seat to their representative. With this understanding a quota must be a whole number. Accordingly the original Hare-quota is not the votes-per-seats ratio itself, but its integral part. With a hopefully pardonable inversion of the historical roots we refer to the integral part of the votes-per-seats ratio as the *Hare-quota variant-1*, HQ1. On occasion legislators were in the mood for rounding the votes-per-seats ratio upwards, thus giving rise to the *Hare-quota variant-2*, HQ2. In summary the default Hare-quota HaQ is accompanied by the variants HQ1 and HQ2,

$$\mathrm{HaQ} = \frac{v_+}{h}, \quad \mathrm{HQ1} = \left\lfloor \frac{v_+}{h} \right\rfloor, \quad \mathrm{HQ2} = \left\lceil \frac{v_+}{h} \right\rceil.$$

Another set of quotas is associated with the name of *Henry Richmond Droop* (1831–1884). *Droop* (1881) proposed what came to be called the *Droop-quota*, DrQ $:= \lfloor v_+/(h+1) \rfloor + 1$. In all applications the Droop-quota is smaller than the Hare-quota. Hence it increases the number of seats allocated in the main apportionment, whence fewer seats are passed on to the residual fit. Since more seats are justified by a full quota of votes, the main apportionment reinforces its persuasive power.

Four Droop-quota variants may be met in practice and theory. Variant-1 omits the addition of unity, variant-2 revives it almost surely by rounding upwards instead of downwards, variant-3 uses standard rounding, and variant-4 remains unrounded,

$$\mathrm{DrQ} = \left\lfloor \frac{v_+}{h+1} \right\rfloor + 1, \quad \mathrm{DQ1} = \left\lfloor \frac{v_+}{h+1} \right\rfloor, \quad \mathrm{DQ2} = \left\lceil \frac{v_+}{h+1} \right\rceil,$$

$$\mathrm{DQ3} = \left\langle \frac{v_+}{h+1} \right\rangle, \quad \mathrm{DQ4} = \frac{v_+}{h+1}.$$

Variant 1 of the Hare-quota, and variants 1 and 3 of the Droop-quota may theoretically become to zero and invalidate their use as divisors. As a remedy they are then assigned the smallest possible positive integer, unity (Section 1.6).

The first seven quota variants are practically relevant. They are in current use, or have been used in the past:

HaQ	Bulgaria, EP 2009 election (Table 1.5)
HQ1	Italy, EP 2009 election (Table 1.17)
HQ2	Lithuania, EP 2009 election (Table 1.18)
DrQ	Ireland, EP 2009 election (Table 1.16)
DQ1	Solothurn, cantonal elections 1896–1977
DQ2	Solothurn, cantonal elections 1981–1993
DQ3	Slovakia, EP 2009 election (Table 1.28)

The Droop-quota variant-4 serves as a kind of fractional approximation to the integral Droop-quota DrQ and its variants 1–3. Moreover, it is a kind of a closure of the shift-quota family, $\lim_{s\to 1} Q(s) = \mathrm{DQ4}$.

5.9. RESIDUAL FIT VARIANTS

At least three alternatives are available to substitute for the residual fit by greatest remainders, as mentioned already in Section 1.7.

The variant gR1 relies on the greatest remainders of the parties' interim quotients, but includes only those parties that receive at least one seat in the main apportionment. We refer to gR1 as the *full-seat restricted residual fit by greatest remainders*. The variant WTA follows the imperative "winner take all" by awarding all residual seats to the strongest party. The variant -EL is peculiar to Greece (Section 1.4). All variants are or were used:

grR	Bulgaria, EP 2009 election (Table 1.5)
gR1	Lithuania, EP 2009 election (Table 1.18)
WTA	Solothurn, cantonal elections 1896–1917
-EL	Greece, EP 2009 election (Table 1.11)

While the Greek version is one of a kind, the other three variants reflect the transition from plurality voting systems to proportional representation systems. In 1896 the Swiss Canton of Solothurn moved from plurality to proportionality. It implemented DQ1WTA, the Droop-quota variant-1 method with residual fit by winner-take-all. While from today's viewpoint the winner-take-all imperative appears unacceptably biased, it is likely that at the time the effect was considered negligible. In the abolished plurality system all seats succumbed to the winner-take-all rule, in the novel proportional system the rule diverted a few residual seats only.

In 1917 Solothurn adopted the residual variant gR1. By eliminating parties that receive no seat in the main apportionment the variant implements a kind of polar imperative, loser-get-nil. Naturally it is a widespread game for stronger parties to devise stumbling stones for weaker parties to enter parliament. However, it is hard to defend such hurdles from a conceptual viewpoint. Quota methods presuppose that the quota represents voters, and so do quota remainders. It is human beings who are being treated by different standards, not just fractional parts of interim quotients.

For example Section 5.3 displays the fractional parts of the interim quotients for the Bulgarian 2009 EP election. These fractional numbers are indicative of human beings. Indeed the Hare-quota signifies batches of 128 619 voters. The four main apportionment seats of the GERB party thus represent $4 \times 128\,619 = 514\,476$ voters. This leaves $627\,693 - 514\,476 = 113\,217$ GERB voters unaccounted for in the main apportionment, to be looked after in the residual fit. In summary the voter numbers unaccounted for in the main apportionment, $u(j)$, are as follows:

$$u(\text{GERB}) = 113\,217,\ \ u(\text{DPS}) = 106\,959,\ \ u(\text{BSP}) = 90\,761,\ \ u(\text{NDSV}) = 76\,527,$$
$$u(\text{SDS-DSB}) = 76\,198,\ \ u(\text{ATAKA}) = 50\,814.$$

Fractional parts and unaccounted voters are related through the elementary formula $f(j) = u(j)/128\,619$ of course. The point is that the present display signifies human beings more readily than the fractional parts in Section 5.3. Since the residual fit commands four seats, the four strongest of the six support groups are able to achieve representation. Regrettably no seats are left to represent the two weakest voter groups. The full-seat restricted variant gR1 would look at these groups conditional on whether their fellow voters have secured representation or not.

5.10. QUOTA METHOD VARIANTS

The present book keeps the focus on the unabridged residual fit by greatest remainders, grR. Still we encounter difficulties. The reason is that most of the quota variants in Section 5.7 produce a whole number, and hence involve a rounding step. Therefore the ensuing quota methods with residual fit by greatest remainders, HQ1grR and so on, fail to be decent in the sense of Section 4.2. By abuse of terminology we continue to speak of “apportionment methods” and thus stretch the term’s meaning beyond its proper limits.

We claim that only the three Hare-quotas and the genuine Droop-quota—HaQ, HQ1, HQ2, and DrQ—are such that their main apportionment always can be paired with a residual fit by greatest remainders. Second we claim that the four Droop-quota variants DQ1–4 are so small that the main apportionment occasionally allocates more seats than are available. Hence the residual apportionment would have to retract seats, not to hand out yet more.

To verify the two claims we sort Hare- and Droop-quotas by decreasing magnitude,

$$\text{HQ2} = \left\lceil \frac{v_+}{h} \right\rceil \geq \text{HaQ} = \frac{v_+}{h} \geq \text{HQ1} = \left\lfloor \frac{v_+}{h} \right\rfloor,$$

$$\text{DrQ} = \left\lfloor \frac{v_+}{h+1} \right\rfloor + 1 \geq \text{DQ2} = \left\lceil \frac{v_+}{h+1} \right\rceil \geq \text{DQ3} = \left\langle \frac{v_+}{h+1} \right\rangle \geq \text{DQ1} = \left\lfloor \frac{v_+}{h+1} \right\rfloor.$$

The Droop-quota variant-4, $\text{DQ4} = v_+/(h+1)$, may be inserted in place of DQ3.

In all practical applications the vote total v_+ is so large that it obeys the condition $v_+ \geq h(h+1)$, that is,

$$\frac{v_+}{h} \geq \frac{v_+}{h+1} + 1. \qquad (*)$$

An application of the floor function shows that the Hare-quota variant-1 stays above the Droop-quota, HQ1 $\geq$ DrQ. Thus the first inequality string runs into the second,

$$\text{HQ2} \geq \text{HaQ} \geq \text{HQ1} \geq \text{DrQ} \geq \text{DQ2} \geq \text{DQ3–4} \geq \text{DQ1}.$$

We now invoke the shift-quotas of Section 5.4. Condition (*) implies $Q(-1) = v_+/(h-1) \geq v_+/h + 1 > \lceil v_+/h \rceil = \text{HQ2}$. On the other hand the Droop-quota satisfies $\text{DrQ} > v_+/(h+1) = Q(1)$. Therefore, there exists an admissible shift $s^* < 1$ with $\text{DrQ} = Q(s^*)$. Hence the first half of the inequality string is framed by shift-quotas,

$$Q(-1) > \text{HQ2} \geq \text{HaQ} \geq \text{HQ1} \geq \text{DrQ} = Q(s^*).$$

Let $y_j(Q) = \lfloor v_j/Q \rfloor$ denote the seats allocated to party j in the main apportionment that uses quota Q. Because of the framing the main apportionment totals $y_+(Q)$ inherit the feasibility bounds of the shift-quota totals (Section 5.4),

$$h - \ell \leq y_+(\text{HQ2}) \leq y_+(\text{HaQ}) \leq y_+(\text{HQ1}) \leq y_+(\text{DrQ}) \leq h.$$

This proves the first claim that the four larger quotas are such that their main apportionments always may be completed with a residual fit by greatest remainders. Moreover, all eight quotas respect the lower feasibility bound $h - \ell$. The quotas never overcharge the residual fit with more than ℓ seats.

The four smaller quotas can lead to seat totals beyond the upper bound h. To see this take a vote total of the form $v_+ = n(h+1)$, for some $n \in \mathbb{N}$. Then the Droop-quota variant-2 equals $\text{DQ2} = \lceil v_+/(h+1) \rceil = n$. If all vote counts v_j happen to be multiples of n, then the main apportionment hands out one seat too many,

$$y_+(\text{DQ2}) = \sum_{j \leq \ell} \left\lfloor \frac{v_j}{n} \right\rfloor = \sum_{j \leq \ell} \frac{v_j}{n} = h + 1.$$

Illustrative figures are easily contrived, $h = 9$ and $v = (40, 30, 20, 10)$. This proves the second claim that the four Droop-quota variants DQ1–4 are infeasible. Additional instructions are required whenever the main apportionment hands out too many seats.

Such instructions are incompatible with the philosophy of quota methods. A quota of votes is perceived as a sacrosanct measure telling how many voters must come together to be guaranteed a Member of Parliament to represent them. Suddenly a deficiency in the electoral system revokes the guarantee. Some seat is declared to be one too many, and is retracted. The only way-out is to soften the sacrosanct status of the quota, and to inject some dose of flexibility. However, if a sensible invocation of flexibility is what is asked for then the answer is divisor methods, not quota methods.

CHAPTER 6

Targeting the House Size: Discrepancy Distribution

Technical aspects are discussed that are common to divisor methods and to quota methods. The methods start with an initial seat apportionment possibly missing the target house size by some discrepancy. The range of variation of the discrepancy is analyzed. For stationary divisor methods, an efficient divisor initialization is recommended. The ensuing discrepancy distribution is determined in two complementary stochastic models. Either the vote shares are assumed to be uniformly distributed and the house size is allowed to be finite, or the vote shares follow an arbitrary absolutely continuous distribution and the house size grows to infinity. An invariance principle emerges showing that the limit discrepancy distribution is a convolution of uniformly distributed rounding residuals irrespective of the underlying vote share distribution.

6.1. SEAT TOTAL AND DISCREPANCY

Apportionment methods operate in two steps. The initial step jumps to some reasonable seat vector $y = (y_1, \ldots, y_\ell)$ without guaranteeing that its component sum y_+ is equal to the house size h. The finalizing step advances from y to a final seat vector $x = (x_1, \ldots, x_\ell) \in \mathbb{N}^\ell(h)$ by adjoining or removing individual seats until the house size is met. The present chapter investigates the properties of the initial seat assignment y.

Divisor methods choose some initial divisor D, divide it into the vote counts v_j, and obtain the interim quotients v_j/D. The underlying rounding rule $[\![\cdot]\!]$ then yields the initial seat numbers

$$y_j(D) \in \left[\!\!\left[\frac{v_j}{D}\right]\!\!\right] \quad \text{for all } j \leq \ell.$$

The resulting seat vector is denoted by $y(D) = \big(y_1(D), \ldots, y_\ell(D)\big)$. The question is how the *seat total* $y_+(D) := \sum_{j\leq\ell} y_j(D)$ compares to the target house size h.

Definition. *The difference* $y_+(D) - h$ *is called* discrepancy.

Quota methods constitute a special case. The divisor D is replaced by the quota Q, and the rounding rule applied is downward rounding, $y_j(Q) \in \lfloor\!\lfloor v_j/Q \rfloor\!\rfloor$. To maintain a maximum level of generality the development focuses on divisor methods.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*, DOI 10.1007/978-3-319-03856-8_6, © Springer International Publishing Switzerland 2014

In the case of a vanishing discrepancy, $y_+(D) - h = 0$, we are in the lucky situation that the initial seat vector represents a final solution, $x = y(D) \in \mathbb{N}^\ell(h)$. Unfortunately there is no divisor D that fits all vote vectors $v = (v_1, \ldots, v_\ell)$. For every divisor D there exists a vote vector v whose discrepancy is nonzero. Yet some initial divisors perform better than others.

A bad initialization is the choice $D = \infty$, with initial seat vector $y(\infty) = 0$. For a start nobody gets anything. The discrepancy $y_+(\infty) - h = -h$ means that there remains a deficiency of h seats, that is, all seats. The finalizing step then passes through a great many rounds to allocate one seat after the other until all h seats are dealt out. This marathon effort is the favorite approach of many legislators and political scientists. Section 4.11 exemplifies its inefficiency.

A good, universal initialization is the votes-per-seats ratio, $D = v_+/h$. For a start every party is allocated its ideal share of seats except for some rounding inaccuracy. The ensuing discrepancy $y_+(v_+/h) - h$ is bounded by $\pm\ell$ (Section 6.2). Hence at most ℓ seats remain to be handled. The gain in efficiency is spectacular, from h down to at most ℓ seats. Section 4.9 illustrates the pertinent calculations.

Stationary divisor methods permit an even more efficient divisor initialization. Given a split $r \in [0; 1]$ the *recommended divisor* is

$$D(r) := \frac{v_+}{h_r}, \qquad \text{where} \qquad h_r := h + \ell\left(r - \frac{1}{2}\right).$$

The recommended divisors $D(r)$ resemble the shift-quotas $Q(s) = v_+/(h + s)$ (Section 5.4). The term h_r is called the *adjusted multiplier*. The discrepancy $y_+\big(D(r)\big) - h$ is seen to be bounded by $\pm\ell/2$, whence its removal touches upon at most $\lfloor \ell/2 \rfloor$ seats (Section 6.3). This is the smallest possible range, the probability of a zero discrepancy is largest, and the discrepancy values $\pm z$ turn rapidly unlikely as z moves away from zero (Theorems 6.7 and 6.11). Section 4.10 presents one of the many instances where initial and final solutions coincide right away.

6.2. UNIVERSAL DIVISOR INITIALIZATION

Consider a general divisor method, with an arbitrary underlying rounding rule $[\![\cdot]\!]$. Let $D > 0$ be an arbitrary initial divisor. The initial seat numbers $y_j(D) \in [\![v_j/D]\!]$ and the interim quotients v_j/D satisfy $-1 \le y_j(D) - v_j/D \le 1$. Summation yields

$$-\ell \le y_+(D) - \frac{v_+}{D} \le \ell.$$

With universal divisor $D = v_+/h$ the difference $y_+(v_+/h) - h$ coincides with the discrepancy. In view of the inequalities it attains values between $\pm\ell$,

$$y_+\left(\frac{v_+}{h}\right) - h \in \big\{-\ell, \ldots, \ell\big\}.$$

The initial seat assignment to party j is its rounded ideal share, $y_j(v_+/h) \in [\![(v_j/v_+)h]\!]$.

Item	Proportion A	Percentage A	Proportion B	Percentage B
Item 1	28.39	28	28.59	29
Item 2	20.42	20	20.52	21
Item 3	14.47	14	13.67	14
Item 4	12.48	12	12.58	13
Item 5	11.38	11	11.58	12
Item 6	7.41	7	7.51	8
Item 7	5.45	5	5.55	6
Sum	100.00	97	100.00	103

TABLE 6.1 *Discrepancy range.* Proportions with a hundredth-of-a-percent accuracy are rounded to percentages using standard rounding and its recommended divisor, $100.00/100 = 1$. The discrepancies are extreme, $-\lfloor 7/2 \rfloor = -3$ percentage points for set A, and $\lfloor 7/2 \rfloor = 3$ percentage points for set B.

6.3. RECOMMENDED DIVISOR INITIALIZATION

A stationary rounding rule $[\![\cdot]\!]_r$ with split $r \in [0;1]$ bounds the difference of initial seat numbers $y_j(D)$ and interim quotients v_j/D in an asymmetric fashion, $-r \le y_j(D) - v_j/D \le 1-r$. Summation and symmetrization yield

$$-\ell r \le y_+(D) - \frac{v_+}{D} \le \ell(1-r), \qquad \text{that is,} \qquad -\frac{\ell}{2} \le y_+(D) - \frac{v_+}{D} + \ell r - \frac{\ell}{2} \le \frac{\ell}{2}.$$

The universal divisor v_+/h leads to the discrepancy range $[-\ell r; \ell - \ell r]$, according to the left display. The right symmetrization suggests the divisor $D(r) = v_+/h_r$ recommended in Section 6.1. Because of $v_+/D(r) - \ell(r - 1/2) = h$ the discrepancy $y_+\big(D(r)\big) - h$ then ranges through the interval $[-\ell/2; \ell/2]$ that is symmetric around zero. Due to integrality the interval can be tightened,

$$y_+\big(D(r)\big) - h \in \left\{ -\left\lfloor \frac{\ell}{2} \right\rfloor, \ldots, \left\lfloor \frac{\ell}{2} \right\rfloor \right\}.$$

The recommended divisor $D(r)$ is more successful than the universal divisor v_+/h. It involves the house size h, the number of parties ℓ, and the split r. The sensitivity pays off by halving the width of the discrepancy range of the universal divisor.

This is the narrowest discrepancy range generally possible. For example, Table 3.2 lists $\ell = 5$ continents and assigns $h = 100$ percentage points to them. The tabled percentages $y_j(v_+/h) = \langle\!\langle (v_j/v_+)h \rangle\!\rangle$ leave a discrepancy of $98 - 100 = -2$ percentage points. Thus the lower bound of the discrepancy range is attained, $-\lfloor 5/2 \rfloor = -2$.

Table 6.1 presents another example from *Happacher* (1996 [5]). Seven fractional percentages, with an accuracy of a hundredth of a percent, are commercially rounded to whole percentages. Presumably most people would spot little exciting differences when glancing over proportions A and B. All first and last digits are the same. The second digit differs only for item 3, but both numbers are properly rounded to the common value 14. Could anything go astray? In terms of whole percentages, set A forfeits three percentage points and drops to the minimum possible discrepancy, $-\lfloor 7/2 \rfloor = -3$. Set B fabricates three percentage points in excess and raises to the maximum possible discrepancy, $\lfloor 7/2 \rfloor = 3$.

For standard rounding ($r = 1/2$), universal and recommended divisors coincide and agree with the votes-per-seats ratio, v_+/h. The coincidence is indicative of the distinguished role played by standard rounding. For upward rounding ($r = 0$) the recommended divisor is

$$D_0 = \frac{v_+}{h - \ell/2},$$

a quantity somewhat larger than the votes-per-seats ratio. Hence the divisor is on the frugal side and balances the generous upward rounding of the finishing step. For downward rounding ($r = 1$) the recommended divisor

$$D_1 = \frac{v_+}{h + \ell/2}$$

is smaller than the votes-per-seats ratio. It hands out seats more generously and thereby counterbalances the final downward rounding. The divisor D_1 was recommended already more than a century ago by the Swiss proportional representation activist *Jules Gfeller* (1890):

> ... on obtiendra le diviseur au moyen de la division du total des suffrages par le nombre des candidats *plus la moitié du nombre des listes.*

> ... one obtains the divisor by means of the division of the total of the votes by the number of seats *plus half of the number of parties.*

The chapter's remainder is devoted to the determination of the discrepancy distribution. Without extra effort the results are derived for seat totals $y_+(D)$ with a general divisor $D > 0$. Section 6.7 returns to specific divisors and their discrepancy.

6.4. DISTRIBUTIONAL ASSUMPTIONS

How likely is it that the seat total attains a given value, $y_+(D) = m$? Of course, the seat total distribution depends on the distributional assumptions for the vote counts. The set of all vote counts is the orthant $(0;\infty)^\ell$, an unbounded set. It is advantageous to normalize the vote counts v_j into the vote shares $w_j = v_j/v_+$. The set of all weight vectors $w = (w_1, \ldots, w_\ell)$ constitutes the *probability simplex*

$$\Omega_\ell := \left\{ (w_1, \ldots, w_\ell) \in (0;1)^\ell \;\middle|\; w_+ = 1 \right\},$$

an open and bounded set in the affine subspace of vectors with component sum unity. Let $\lambda_{\ell-1}$ denote $(\ell-1)$-dimensional Lebesgue measure over the probability simplex Ω_ℓ. (The total mass is well-known to be $\lambda_{\ell-1}(\Omega_\ell) = \sqrt{\ell}/\big((\ell-1)!\big)$, but the exposition makes no use of this formula.) The subsequent results assume the distribution of the weight vectors to be (a) uniform or, more generally, (b) absolutely continuous.

Definition.

a. *The weight vectors w are said to follow the* uniform distribution *when their distribution is given by normalized Lebesgue measure,*

$$P(B) = \frac{\lambda_{\ell-1}(B)}{\lambda_{\ell-1}(\Omega_\ell)} \qquad \text{for all Borel subsets } B \subseteq \Omega_\ell.$$

b. *The weight vectors w are said to follow an* absolutely continuous distribution *when their distribution is given by a Lebesgue-integrable density function f,*

$$P(B) = \int_B f \, d\lambda_{\ell-1} \qquad \text{for all Borel subsets } B \subseteq \Omega_\ell.$$

The uniform distribution is the particular absolutely continuous distribution that has a constant density function, $f(w) = 1/\lambda_{\ell-1}(\Omega_\ell)$. An important consequence of absolutely continuous distributions is that the occurrence of ties disappears in null-sets. Given a multiplier $\mu > 0$, a scaled weight μw_j is a tie if and only if it hits a signpost, $\mu w_j = s(n)$. Hence the vectors $(w_1, \ldots, w_\ell) \in \Omega_\ell$ with a fixed component $w_j = s(n)/\mu$ lie in a lower-dimensional subspace and have $\lambda_{\ell-1}$-measure zero. It is safe to neglect ties, in the remainder of the present chapter.

In the absence of ties the set-valued rounding rule $[\![\cdot]\!]$ is substituted by a compatible rounding function $[\cdot] : [0; \infty) \mapsto \mathbb{N}$ (Section 3.8). The set-oriented notion $y_j(D) \in [\![v_j/D]\!]$ is replaced by the number-oriented notion $y_j(D) = [v_j/D]$. Recall that divisors for vote counts v_j, and multipliers for vote shares $w_j = v_j/v_+$ are related through

$$\frac{v_j}{D} = \frac{v_+}{D}\frac{v_j}{v_+} = \mu w_j, \qquad \mu = \frac{v_+}{D}.$$

The seat total attains the format $y_+(D) = \sum_{j\le\ell}[\mu w_j]$. It visibly displays the pertinent variables: ℓ, μ, and w_j. The event that the seat total attains a given value m is

$$B(m) := \left\{ (w_1, \ldots, w_\ell) \in \Omega_\ell \;\middle|\; \sum_{j\le\ell}[\mu w_j] = m \right\}.$$

The probabilities of the events $B(m)$ for $m \in \mathbb{N}$ constitute the seat total distribution.

From now on we restrict attention to stationary divisor methods with split $r \in [0; 1]$. The bounds $v_+/D - \ell r \le y_+(D) \le v_+/D + \ell(1-r)$ from Section 6.3 translate into $\mu - \ell r \le \sum_{j\le\ell}[\mu w_j]_r \le \mu + \ell(1-r)$, where $\mu > 0$. Under the specific assumption of a uniform distribution we determine the seat total distribution (Lemma 6.5), and the discrepancy distribution for a fixed house size h (Theorem 6.7).

The distributions prove dramatically practical. They apply to apportionment methods in electoral systems, to the transformation into percentages or the like, to the fitting of contingency tables, as well as to a wealth of other rounding problems. Their practicality comes as a surprise since a weight distribution that is uniform would appear to be a bold model to capture the peculiarities of this fan of applications. The reason for the wide applicability is that the distributions depend on their input only through the rounding residuals (Lemma 6.9). When the house size grows to infinity, the rounding residuals become exchangeable and uniformly distributed even under the rather general and very mild assumption that the weights follow an *arbitrary* absolutely continuous distribution (Invariance Principle 6.10).

In the subsequent formulas we write

$$t_{\text{pos}} := \frac{t + |t|}{2} = \max\{t, 0\}$$

for the *positive part* of the real number $t \in \mathbb{R}$, and shorten $(t_{\text{pos}})^n$ to t^n_{pos}. The following lemma calculates the seat total distribution for a given multiplier μ. We assume that the multipliers are not too close to zero, $\mu > \ell r$. The assumption guarantees that the seat total is positive, and leapfrogs the irrelevant event that the seat total is zero.

6.5. SEAT-TOTAL DISTRIBUTIONS

Lemma. *Consider the stationary divisor method with split* $r \in [0;1]$, *and a multiplier* $\mu > \ell r$. *If the weights are uniformly distributed then the seat total attains the values* $m = \lceil \mu - \ell r \rceil, \ldots, \lfloor \mu + \ell - \ell r \rfloor$ *with probabilities*

$$P\Big(\Big\{(w_1,\ldots,w_\ell) \in \Omega_\ell \,\Big|\, \sum_{j\le\ell}[\mu w_j]_r = m\Big\}\Big)$$
$$= \sum_{k=0}^{\ell} \frac{(-1)^k}{\mu^{\ell-1}} \binom{\ell}{k} \sum_{i=0}^{\ell-k} \binom{\ell-k}{i} \binom{m+k-1}{i+k-1} \Big(\mu - m + i(1-r) - kr\Big)_{\text{pos}}^{\ell-1}.$$

Proof. I. The seat total is bounded according to $\mu - \ell r \le \sum_{j\le\ell}[\mu w_j]_r \le \mu + \ell(1-r)$ (Section 6.3). Hence its values m range between $\lceil \mu - \ell r \rceil$ and $\lfloor \mu + \ell - \ell r \rfloor$. The assumption $\mu > \ell r$ secures $m \ge 1$, and bypasses the trivial singleton $\mathbb{N}^\ell(0) = \{0\}$.

The domain of attraction $A(y)$ of a seat vector $y \in \mathbb{N}^\ell(m)$ is defined to consist of the weight vectors w so that μw gets rounded to y,

$$A(y) := \Big\{(w_1,\ldots,w_\ell) \in \Omega_\ell \,\Big|\, [\mu w_1]_r = y_1, \ldots, [\mu w_\ell]_r = y_\ell\Big\}.$$

The intersection of two such domains consists of ties and hence is a null-set. Therefore, the probability sought is the sum of the probabilities of the domains of attraction,

$$P\Big(\Big\{(w_1,\ldots,w_\ell) \in \Omega_\ell \,\Big|\, \sum_{j\le\ell}[\mu w_j]_r = m\Big\}\Big) = \sum_{y\in\mathbb{N}^\ell(m)} P\big(A(y)\big) = \sum_{y\in\mathbb{N}^\ell(m)} \frac{\lambda_{\ell-1}\big(A(y)\big)}{\lambda_{\ell-1}(\Omega_\ell)}.$$

II. We fix a vector $y = (y_1,\ldots,y_\ell) \in \mathbb{N}^\ell(m)$. A component y_j attracts all values in the interval $[a_j; b_j]$. The limits are $a_j := y_j - 1 + r$ in case $y_j > 0$ and $a_j := 0$ in case $y_j = 0$, and $b_j := y_j + r$. The rectangle $[a;b] := [a_1;b_1] \times \cdots \times [a_\ell; b_\ell]$ with South-West corner $a = (a_1,\ldots,a_\ell)$ and North-East corner $b = (b_1,\ldots,b_\ell)$ assembles the vectors $v \in (0;\infty)^\ell$ that are rounded to y. Introducing for a set $B \subseteq \mathbb{R}^\ell$ the scaled set $\mu B := \{\mu v \mid v \in B\}$, we get $\mu A(y) = \big(\mu\Omega_\ell\big) \cap [a;b]$.

The set of corner vectors of the rectangle $[a;b]$ is partitioned into subsets $E^k(y)$, for $k = 0,\ldots,\ell$. The set $E^k(y)$ is defined to contain the corner vectors c such that k components come from b and $\ell - k$ from a. Every corner vector c induces an orthant $[c;\infty) := [c_1;\infty) \times \cdots \times [c_\ell;\infty)$. The inclusion-exclusion principle yields

$$\lambda_{\ell-1}\big(A(y)\big) = \frac{1}{\mu^{\ell-1}}\lambda_{\ell-1}\big(\mu A(y)\big) = \frac{1}{\mu^{\ell-1}} \sum_{k=0}^{\ell} \sum_{c\in E^k(y)} (-1)^k \, \lambda_{\ell-1}\big((\mu\Omega_\ell) \cap [c;\infty)\big).$$

In case $\mu \ge c_+$ the intersection is $(\mu\Omega_\ell) \cap [c;\infty) = c + (\mu - c_+)\Omega_\ell$. It originates from the probability simplex Ω_ℓ through a translation by the vector c and a scaling by the factor $\mu - c_+$. Lebesgue measure $\lambda_{\ell-1}$ is invariant under translation, and scales with power $\ell - 1$. In case $\mu < c_+$ the intersection is empty. Both cases combine into the volume formula $\lambda_{\ell-1}\big((\mu\Omega_\ell) \cap [c;\infty)\big) = (\mu - c_+)_{\text{pos}}^{\ell-1} \, \lambda_{\ell-1}(\Omega_\ell)$. Hence we obtain

$$\sum_{y\in\mathbb{N}^\ell(m)} \frac{\lambda_{\ell-1}\big(A(y)\big)}{\lambda_{\ell-1}(\Omega_\ell)} = \sum_{k=0}^{\ell} \frac{(-1)^k}{\mu^{\ell-1}} \sum_{y\in\mathbb{N}^\ell(m)} \sum_{c\in E^k(y)} \big(\mu - c_+\big)_{\text{pos}}^{\ell-1}.$$

III. The corner vector set $E^k(y)$ is further partitioned into the subsets $E_i^k(y)$, for $i = 0, \dots, \ell - k$. The set $E_i^k(y)$ is defined to contain the vectors $c \in E^k(y)$ wherein the $\ell - k$ components from a are such that i belongs to $y_j > 0$, and $\ell - k - i$ to $y_j = 0$. With a seat vector $y \in \mathbb{N}^\ell(m)$ fixed, the set $E_i^k(y)$ contains $\binom{\ell}{k}\binom{\ell-k}{i}$ many vectors that are permutations of the *generator* having

k	initial components equal to	$y_j + 1 - (1 - r)$,	where $y_j \geq 0$,
i	middle components equal to	$y_j - (1 - r)$,	where $y_j > 0$,
$\ell - k - i$	remaining components equal to	0,	where $y_j = 0$.

As the seat vector y varies through $\mathbb{N}^\ell(m)$, there are as many distinct generators as there are ways for $y_+ + k = m + k$ indistinguishable objects to occupy $i + k$ cells leaving none of the cells empty. This occupancy problem is well-known to have $\binom{m+k-1}{i+k-1}$ solutions. The vectors $c \in E_i^k(y)$ share the same component sum, $c_+ = m - i(1 - r) + kr$. □

The double sum comprises at most $(\ell + 1)(\ell + 2)/2$ terms and lends itself to rapid machine computation. For values m smaller or larger than the indicated range the double sum can be shown to be zero. In order to deduce the distribution of the discrepancy $y_+\big(D(r)\big) - h$ we pass from the recommended divisor $D(r) = v_+/h_r$ to the adjusted multiplier $v_+/D(r) = h + \ell(r - 1/2) = h_r$. Thus the discrepancy of interest is $\big(\sum_{j \leq \ell} [h_r w_j]_r\big) - h$, to be pursued in Theorem 6.7.

6.6. HAGENBACH-BISCHOFF INITIALIZATION

We briefly digress to appreciate the contributions of *Eduard Hagenbach-Bischoff*. As the leading proportional representation proponent in the Great Council of the Canton of Basel in Switzerland, he fought for the introduction of the divisor method with downward rounding, and succeeded. By profession a physics professor at Basel University, his numerous writings on the topic are still today a fruitful and reliable source. The 28-page pamphlet *Hagenbach-Bischoff* (1905 [15, 26]) is the first to calculate discrepancy probabilities.

Hagenbach-Bischoff remarks that the divisor method with downward rounding executes faster when the initial divisor is chosen somewhat smaller than the votes-per-seats ratio v_+/h. His choice is $\lfloor v_+/(h+1) \rfloor + 1$. This divisor coincides with the Droop-quota. (In view of this coincidence some authors erroneously file *Hagenbach-Bischoff*'s work under the heading of quota methods.) The feasibility bounds $h - \ell \leq y_+(\text{DrQ}) \leq h$ of Section 5.10 mean that the discrepancy is non-positive, $-\ell \leq y_+(\text{DrQ}) - h \leq 0$. How likely is it that the discrepancy equals $z = -1$ and one seat must be adjoined, or that it equals $z = -2$ and two seats are needed, and so on?

For the calculation of these probabilities *Hagenbach-Bischoff* makes two vital assumptions that are specified in the appendix “Mathematische Ergänzungen”. First, the initial divisor is stripped off the rounding step and is simplified to $v_+/(h + 1)$. With this simplification, the divisor induces the multiplier $\mu = h + s$ with shift $s = 1$. Second, he interprets the discrepancies via rounding residuals (as we do in Lemma 6.9). In our terms this means that the house size h grows beyond limits.

Hence *Hagenbach-Bischoff*'s results follow from Lemma 6.5 when $r = 1$, $m = h+z$, $\mu = h+s$, and $h \to \infty$. The limit of the inner sum is determined by the term $i = \ell - k$ and converges to $(\ell-1)!$, as we show in the proof of Theorem 6.7.a. Thus we obtain

$$\lim_{h\to\infty} P\Bigg(\bigg\{w \in \Omega_\ell \;\bigg|\; \Big(\sum_{j\le\ell} \lfloor (h+s)w_j \rfloor\Big) - h = z\bigg\}\Bigg) = \sum_{k=0}^{\ell} \frac{(-1)^k}{(\ell-1)!}\binom{\ell}{k}\Big(s-z-k\Big)_{\text{pos}}^{\ell-1}.$$

With $s = 1$, the formula yields the probabilities given by *Hagenbach-Bischoff*:

If **three** parties are present, then the probability is
1/2 that the discrepancy equals zero and the initial seat vector is final,
1/2 that the discrepancy equals -1 and is removed by adding one seat.
If **four** parties are present, then the probability is
1/6 that the discrepancy equals zero and the initial seat vector is final,
2/3 that the discrepancy equals -1 and is removed by adding one seat,
1/6 that the discrepancy equals -2 and is removed by adding two seats.
If **five** parties are present, then the probability is
1/24 that the discrepancy equals zero and the initial seat vector is final,
11/24 that the discrepancy equals -1 and is removed by adding one seat,
11/24 that the discrepancy equals -2 and is removed by adding two seats,
1/24 that the discrepancy equals -3 and is removed by adding three seats.

Section 6.1 recommends the divisor $v_+/(h+\ell/2)$. The shift $s = \ell/2$ responds to the size of the party system, in contrast to *Hagenbach-Bischoff*'s constant choice $s = 1$. The multiplier that is associated with the recommended divisor is the adjusted multiplier $h + \ell/2$. An application of the above formula with $s = \ell/2$, for $\ell = 3, 4, 5$, yields discrepancy distributions more favorable than those of *Hagenbach-Bischoff*:

If **three** parties are present, then the probability is
3/4 that the discrepancy equals zero and the initial seat vector is final,
1/8 that the discrepancy equals -1 and is removed by adding one seat,
1/8 that the discrepancy equals $+1$ and is removed by retracting one seat.
If **four** parties are present, then the probability is
2/3 that the discrepancy equals zero and the initial seat vector is final,
1/6 that the discrepancy equals -1 and is removed by adding on seat,
1/6 that the discrepancy equals $+1$ and is removed by retracting one seat.
If **five** parties are present, then the probability is
115/192 that the discrepancy equals zero and the initial seat vector is final,
19/96 that the discrepancy equals -1 and is removed by adding one seat,
19/96 that the discrepancy equals $+1$ and is removed by retracting one seat,
1/384 that the discrepancy equals -2 and is removed by adding two seats,
1/384 that the discrepancy equals $+2$ and is removed by retracting two seats.

The variation of the shift parameter, from 1 over 3/2 and 2 to 5/2, may appear negligible, but actually is significant. The event that the discrepancy vanishes acquires maximum likelihood, and the probability of a discrepancy $z \neq 0$ decreases rapidly as z moves away from zero. The general formulas that govern this behavior are as follows.

6.7. DISCREPANCY PROBABILITIES: FORMULAS

Theorem. *Consider the stationary divisor method with split* $r \in [0;1]$*, and a house size* $h > \ell/2$*. With adjusted multiplier* $h_r = h + \ell(r - 1/2)$ *let*

$$p_{\ell,r,h}(z) := P\left(\left\{(w_1,\dots,w_\ell) \in \Omega_\ell \;\middle|\; \left(\sum_{j\le\ell}[h_r w_j]_r\right) - h = z\right\}\right)$$

denote the probability that the discrepancy attains the values $z = -\lfloor \ell/2 \rfloor, \dots, \lfloor \ell/2 \rfloor$.

a. (Double-Sum Formula) *If the weight vectors* w *are uniformly distributed then*

$$p_{\ell,r,h}(z) = \sum_{k=0}^{\ell} \frac{(-1)^k}{h_r^{\ell-1}} \binom{\ell}{k} \sum_{i=0}^{\ell-k} \binom{\ell-k}{i} \binom{z+h+k-1}{i+k-1} \left(\frac{\ell}{2} - z - k - (\ell-k-i)(1-r)\right)_{\text{pos}}^{\ell-1}.$$

b. (Single-Sum Formula) *For large house sizes the probabilities* $p_{\ell,r,h}(z)$ *have a limit,*

$$\lim_{h\to\infty} p_{\ell,r,h}(z) = \sum_{k=0}^{\ell} \frac{(-1)^k}{(\ell-1)!} \binom{\ell}{k} \left(\frac{\ell}{2} - z - k\right)_{\text{pos}}^{\ell-1} =: g_\ell(z).$$

c. (Approximation Formula) *For large numbers of parties the limits* $g_\ell(z)$ *converge,*

$$\lim_{\ell\to\infty} \sqrt{\frac{\ell}{12}}\, g_\ell\left(\sqrt{\frac{\ell}{12}}\, z\right) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}, \qquad \textit{that is,} \qquad g_\ell(z) \approx \sqrt{\frac{6}{\ell\pi}} e^{-6z^2/\ell}.$$

Proof. a. The Double-Sum Formula is Lemma 6.5 with $\mu = h_r$ and $m = z + h$.

b. In the Double-Sum Formula, the ratio $\binom{z+h+k-1}{i+k-1} / h_r^{\ell-1}$ is of order $h^{i+k-\ell}$. Hence the inner sum has all ratios with $i < \ell - k$ tending to zero. (The terms originate from corner vectors with $\ell - k - i > 0$ many zeros.) The last ratio with $i = \ell - k$ converges to $1/(\ell-1)!$. (These are the corner vectors with no zeros.) Hence the Double-Sum Formula has limit $g_\ell(z)$. A reference to the split parameter is no longer needed because the limit is the same for all r.

c. Let the random variables $U_1, \dots, U_\ell$ be stochastically independent and uniformly distributed over the interval $[-1/2; 1/2]$. The sum U_+ has a density that for $u \in [-\ell/2; \ell/2]$ is given by

$$g_\ell(u) = \int g_{\ell-1}(t) g_1(u-t)\, dt = \int_{u-\frac{1}{2}}^{u+\frac{1}{2}} g_{\ell-1}(t)\, dt = \sum_{k=0}^{\ell} \frac{(-1)^k}{(\ell-1)!} \binom{\ell}{k} \left(\frac{\ell}{2} - u - k\right)_{\text{pos}}^{\ell-1}.$$

The last identity is proved by induction, see *Feller* (1971 [28]). The sum U_+ has expectation zero and variance $\ell/12$. Hence $\sqrt{\ell/12}\, g\big(\sqrt{\ell/12}\, z\big)$ is the density of the standardized sum, and the Central Limit Theorem applies. An Edgeworth expansion details the speed of convergence, see *Happacher* (1996 [90, 95]):

$$\left| g_\ell(z) - \left(1 - \frac{3}{20\ell} + \frac{18z^2}{5\ell^2} - \frac{36z^4}{5\ell^3}\right) \sqrt{\frac{6}{\ell\pi}}\, e^{-6z^2/\ell} \right| \le \ell^{-5/2}. \qquad \square$$

According to the Invariance Principle 6.10 the Single-Sum Formula continues to hold true even when the weight vector w follows an arbitrary absolutely continuous distribution, rather than being restricted to the uniform distribution. Yet we pause before accumulating more theory, and contemplate some practical aspects.

6.8. DISCREPANCY PROBABILITIES: PRACTICE

Discrepancy examples abound. In December 1999 a public German television station conducted a telephone poll to rank five pop music groups. The top group was fortunate to win by the narrow margin of one percentage point. However, the five percentages that flashed over the television screen implied a discrepancy of one percentage point,

$$\big(29 + 28 + 25 + 13 + 6\big) - 100 = 1.$$

This discrepancy point hit the press. The tabloid *Bild am Sonntag* made a mockery of the television people's inability to count to one hundred. In a subsequent press release the television company reaffirmed the public that the excess point had not been instrumental to determine the winner. What they wanted to say is that simple rounding, though deficient, still is concordant (Section 4.2). Simple rounding is all too likely to miss the target house size (Section 3.13). The probabilities that troubled the television people are given by the Double-Sum Formula ($\ell = 5, r = 1/2, h = 100$):

Discrepancy z	-2	-1	0	1	2	Sum
Double-Sum Formula [in %]	0.3	20.2	59.8	19.4	0.3	100.0

The television station risked a sizable chance of forty percent to encounter a nonzero discrepancy and to ridicule themselves.

The extreme discrepancies ± 3 in Table 6.1 are much less likely, occurring just twice per a hundred thousand instances. With reference to this setting ($\ell = 7, r = 1/2, h = 100$) Theorem 6.7 yields the following probabilities [in %]:

Discrepancy z	-3	-2	-1	0	1	2	3	Sum
a. Double-Sum	0.002	1.652	23.591	51.093	22.176	1.484	0.002	100.000
b. Single-Sum	0.002	1.567	22.880	51.102	22.880	1.567	0.002	100.000
c. Approximation	0.023	1.694	22.166	52.234	22.166	1.694	0.023	100.000

The three formulas send the same qualitative message. Since they are evaluated equally fast by modern computing equipment, we consider them interchangeable.

We note that the Double-Sum Formula yields a skewed distribution, with slightly more weight on negative discrepancies than on their positive counterparts. The skewness is caused by the scaled simplex in the proof of Lemma 6.5, $\mu\Omega_\ell$, that cuts the nonnegative orthant $(0;\infty)^\ell$ into two pieces. The piece towards the origin is a relatively compact pyramid, the piece away from the origin is a truncated pyramid receding to infinity. On lower-dimensional faces (in Lemma 6.5 for $i < \ell - k$, when corner vectors have components that are zero) the orientation matters; negative discrepancies (towards the origin) are more likely than positive discrepancies (away from the origin), see *Happacher* (1996 [55]). Within the relative interior ($i = \ell - k$) the two directions look locally alike, whence the limit Single-Sum distribution is symmetric.

The discrepancy distributions conform rather satisfactorily with empirical data, as we illustrate with the 1996 United States and Russian presidential elections. In both instances the vote counts are rounded to a tenth of a percent ($h = 1000$) using standard rounding ($r = 1/2$). For the 1996 US election the *International Herald Tribune* of 7 November 1996 reported the vote counts for the three leading candidates ($\ell = 3$) within the 50 states, the District of Columbia, and the whole country. This

Discrepancy z	0	± 1	± 2	± 3	$\pm\{4,\dots\}$	Sum
$\ell =$ 2	100	0				100
3	76	12				100
4	66	17	0			100
5	60	20	0			100
10	44	24	4	0	0	100
20	30	23	10	2	0	100
30	26	21	11	4	1	100
40	22	19	12	6	2	100
50	20	17	12	7	4	100
100	14	13	11	8	11	100

TABLE 6.2 *Discrepancy distributions for increasing numbers of party.* The Single-Sum probabilities 6.7.b are listed after being symmetrized around zero [in %]. The distributions spread out and flatten as ℓ grows. The probability of a vanishing discrepancy tends to zero of order $1.4/\sqrt{\ell}$.

provides a sample of size $N = 52$. The observed counts of the discrepancy values $-1, 0, 1$ corroborate the theoretically predicted counts:

Discrepancy z	-1	0	1	N
Observed counts	5	39	8	52
Predicted counts	7	39	6	52

The predicted counts—which are integers—are obtained from the expected values $Np_{3,1/2,1000}(z)$—which are reals—using the divisor method with standard rounding.

The 1996 Russian presidential election data were reported in the *Rossijskaja Gazeta* of 1 July 1996. The Russian people could vote for one of ten candidates, or against all of them ($\ell = 11$). The 89 Constitutional Subjects of the Russian Federation, the votes abroad, and the country-wide totals provide a sample of size $N = 91$. Discrepancy values as extreme as ± 5 and ± 4 did not materialize.

Discrepancy z	-3	-2	-1	0	1	2	3	N
Observed counts	0	9	18	37	20	6	1	91
Predicted counts	0	4	23	38	22	4	0	91

Again the counts observed agree with those predicted.

The most desired event is that the discrepancy is zero, so that the target house size is met. This probability decreases as the size of the party system increases. For the American data ($\ell = 3$) the percentages duly sum to one hundred in 39 out of 52 instances, that is, in seventy-five cases out of a hundred. With the Russian data ($\ell = 11$) this occurs in 37 out of 91 instances, that is, only in forty of a hundred cases. Theorem 6.7.c says that the probability of exhausting the house size converges to zero,

$$P\Big(\Big\{(w_1,\dots,w_\ell) \in \Omega_\ell \;\Big|\; \sum\nolimits_{j\le\ell} [h_r w_j]_r = h\Big\}\Big) \approx g_\ell(0) = \frac{1.4}{\sqrt{\ell}}.$$

As a rather simplistic rule of thumb we equate $1.4/\sqrt{\ell}$ to one half for all one-digit system sizes $\ell \le 9$. In other words the target house size is likely to be hit right away in about half of all applications. The other half misses the target by one seat too few or too many (or two or more seats, but very rarely so).

Whether nonzero discrepancies are judged irritating or not depends on the context. A nonzero discrepancy is entirely unacceptable when the apportionment concerns parliamentary seats. The public would not agree to some seats remaining vacant or others being brought to life on the ground of rounding effects. A nonzero discrepancy may still be a cause for ridicule when the allocation units are percentage points that eventually stay below or exceed one hundred, as experienced by the television company. But a nonzero discrepancy is increasingly tolerable when it accrues after the decimal point in the last digit where it is clearly attributable to rounding effects.

Yet some people take pains to secure impeccable totals. For example, vote share percentages in the *Augsburger Allgemeine* newspaper always sum to 100.0, not to 99.9 nor to 100.1. The newspaper prints electoral results by quoting vote counts along with vote shares. Vote shares are given in tenths of a percent, $n_j/1000$, in such a way that they persistently fit the correct total, $\sum_{j\le\ell} n_j = 1000$. Persistent fits cannot result from simple rounding. So what do the journalists do to balance the field? They fit the strongest party last. All other parties have their vote shares commercially rounded to a tenth of a percent, $n_j := \langle 1000\, v_j/v_+\rangle$ for all $j = 2, \dots, \ell$. The strongest party makes up for a possible imbalance, $n_1 := 1000 - \sum_{j=2}^{\ell} n_j$.

The *journalists' apportionment method* may be classified as a "Hare-quota method with standard rounding and with a residual fit by winner-take-all". Indeed, it relies on the quota $Q = v_+/1000$ to calculate interim quotients v_j/Q. Then standard rounding is applied, not downward rounding. The residual fit implements the winner-take-all rule, $n_1 := 1000 - \sum_{j=2}^{\ell} n_j$, meaning that a few tenths of a percent are added or subtracted according to whether the discrepancy is negative or positive. An obvious variant to ensure a fitting total is the loser-take-all rule, $n_\ell := 1000 - \sum_{j<\ell} n_j$.

The following lemma builds on a similar strategy, to temporarily deal with all but one party, and to eventually revive the omitted party to complete the field. The lemma expresses the discrepancy in terms of rounding residuals $u_j(h)$. Out of all ℓ parties, $\ell - 1$ are dealt with and one is omitted.

6.9. DISCREPANCY AND ROUNDING RESIDUALS

Lemma. *Consider the stationary divisor method with split $r \in [0;1]$. Let $h_r = h + \ell(r - 1/2)$ denote the adjusted multiplier. For weights $w = (w_1, \dots, w_\ell) \in \Omega_\ell$ such that $h_r w$ is tie-free, the rounding residual of party $j \le \ell$ is given by*

$$u_j(h) := \left\langle h_r w_j - r + \frac{1}{2} \right\rangle - \left(h_r w_j - r + \frac{1}{2} \right) \in \left[-\frac{1}{2}; \frac{1}{2} \right].$$

Then, omitting an arbitrary party j, the discrepancy satisfies

$$\left(\sum_{i\le\ell} [h_r w_i]_r \right) - h = u_+(h) = \left\langle \sum_{i\ne j} u_i(h) \right\rangle.$$

Proof. The fundamental relation for stationary rounding, $n - 1 + r \le t \le n + r$, translates into the fundamental relation for standard rounding, $n - 1/2 \le t - r + 1/2 \le n + 1/2$. Thus stationary rounding is always expressible through standard rounding, $[\![t]\!]_r = \langle\!\langle t - r + 1/2 \rangle\!\rangle$. In the presence of tie-freeness we revert to rounding functions, and equate $[h_r w_j]_r = \langle h_r w_j - r + 1/2 \rangle =: y_j$. Summation over the defining relation for the rounding residuals yields the first identity, $u_+(h) = y_+ - h_r + \ell(r - 1/2) = y_+ - h$.

As for the second identity we note that the rounding residuals (left) are continuous variables while the discrepancy (right) is a discrete variable. The breach of measurement levels is easily mended. Setting $z := y_+ - h \in \mathbb{Z}$ we separate the rounding residual of party j from the other rounding residuals,

$$\sum_{i \neq j} u_i(h) = z - u_j(h) \in \left[z - \frac{1}{2}; z + \frac{1}{2}\right], \qquad \text{that is,} \qquad \left\langle \sum_{i \neq j} u_i(h) \right\rangle = z.$$

The last statement employs the extended definition of standard rounding (Section 3.7). □

With the discrepancy exhibited as a function of the rounding residual and nothing else, the search for the discrepancy distribution raises an intermediate question: What is the distribution of the rounding residuals? If we assume that the house size grows large, then the answer is simple. The joint distribution of the rounding residuals is exchangeable, with uniform marginals. This is the main message of the following Invariance Principle. It interprets the quantities of interest to be random. The change of viewpoint is indicated by capital letters, such as W_j or $U_j(h)$. The Invariance Principle states that, almost invariably, large house sizes make the rounding residuals look like random variables $U_1, \ldots, U_\ell$ that are exchangeable. In addition, they are stochastically independent of the underlying weights $W_1, \ldots, W_\ell$. The common distribution of the limit variables U_j is the uniform distribution over their range $[-1/2; 1/2]$.

6.10. INVARIANCE PRINCIPLE FOR ROUNDING RESIDUALS

Theorem. *Consider the stationary divisor method with split $r \in [0; 1]$, and assume that the weights $W_1, \ldots, W_\ell$ follow an arbitrary absolutely continuous distribution over the probability simplex Ω_ℓ. Let $h_r = h + \ell(r - 1/2)$ denote the adjusted multiplier.*

Then the rounding residuals $U_j(h) := \langle h_r W_j - r + 1/2 \rangle - (h_r W_j - r + 1/2) \in [-1/2; 1/2]$ and the weights W_j, for $j \leq \ell$, jointly converge in distribution,

$$\big(U_1(h), \ldots, U_\ell(h), W_1, \ldots, W_\ell\big) \xrightarrow[h \to \infty]{\text{in distribution}} (U_1, \ldots, U_\ell, W_1, \ldots, W_\ell),$$

with limit variables $U_1, \ldots, U_\ell$ being uniformly distributed over the interval $[-1/2; 1/2]$, exchangeable, and stochastically independent of $W_1, \ldots, W_\ell$. Omitting any limit variable U_j, the remaining $\ell - 1$ variables U_i, $i \neq j$, are stochastically independent.

Proof. Initially the proof omits the last component, $j = \ell$. Setting $k := \ell - 1$ the components are assembled into the vectors $U(h) = (U_1(h), \ldots, U_k(h))$ and $W = (W_1, \ldots, W_k)$. By assumption the random vector W admits a Lebesgue density f on the k-dimensional simplex $\big\{(w_1, \ldots, w_k) \in (0; 1)^k \mid \sum_{j \leq k} w_j < 1\big\}$. Convergence in distribution is verified via the Lévy continuity theorem for characteristic functions, as in *Janson* (2012). A prime signifies the scalar product of two vectors, $s't := \sum_{j \leq k} s_j t_j$, and $\mathrm{i} := \sqrt{-1}$ is the imaginary unit.

The random variables $2\pi U_j(h)$ fall into the interval $[-\pi;\pi]$. Since the trigonometric system over the unit circle is complete, integer coefficients suffice. With integer vector $s = (s_1,\ldots,s_k) \in \mathbb{Z}^k$ and real vector $t = (t_1,\ldots,t_k) \in \mathbb{R}^k$ we introduce the characteristic functions

$$\varphi_h(s,t) := \mathrm{E}\Big(e^{2\pi \mathrm{i}\, s'U(h) + \mathrm{i}\, t'W}\Big).$$

The claim is that these characteristic functions converge to the characteristic function of the limit distribution,

$$\lim_{h\to\infty} \varphi_h(s,t) = \Big(\prod\nolimits_{j\le k} \mathrm{E}\left(e^{2\pi \mathrm{i} s_j U_j}\right)\Big) \cdot \mathrm{E}\Big(e^{\mathrm{i}\, t'W}\Big).$$

The claim is trivially true in case $s = 0$. Otherwise, when some component s_j is nonzero, the factor $\mathrm{E}\Big(e^{2\pi \mathrm{i} s_j U_j}\Big) = 0$ annihilates the right-hand side. Hence the claim simplifies,

$$s \neq 0 \qquad \Longrightarrow \qquad \lim_{h\to\infty} \varphi_h(s,t) = 0.$$

The periodicity $e^{2\pi \mathrm{i} z} = 1$ obliterates the integer $\langle h_r W_j - r + 1/2\rangle$, and leaves $\exp\Big(2\pi \mathrm{i}\, s'U(h)\Big) = \exp\Big(-2\pi \mathrm{i} \sum_{j\le k} s_j(h_r W_j - r + 1/2)\Big) = \exp\Big(-2\pi \mathrm{i} h_r\, s'W + 2\pi \mathrm{i}(r-1/2)s_+\Big)$. Thus $\varphi_h(s,t)$ is determined by the Fourier transform $\widehat{f}$ of the density f,

$$\varphi_h(s,t) = \mathrm{E}\Big(e^{\mathrm{i}\,(t - 2\pi h_r\, s)'W}\Big)\, e^{2\pi \mathrm{i}(r-1/2)s_+} = \widehat{f}(t - 2\pi h_r\, s)\, e^{2\pi \mathrm{i}(r-1/2)s_+}.$$

Now the Riemann / Lebesgue Lemma is invoked stating that Fourier transforms of Lebesgue densities vanish at infinity,

$$\lim_{h\to\infty} \|t - 2\pi h_r\, s\| = \infty \qquad \Longrightarrow \qquad \lim_{h\to\infty} \widehat{f}(t - 2\pi h_r\, s) = 0,$$

see, for example, *Bauer* (1991 [196]). This proves the convergence of $\big(U(h), W\big)$ to (U, W) where the components of $U := (U_1,\ldots,U_k)$ are stochastically independent of each other, and of W.

Independence evidently extends to the omitted weight $W_\ell = 1 - \sum_{j<\ell} W_j$. The omitted rounding residual is $U_\ell(h) = \big\langle \sum_{j<\ell} U_j(h)\big\rangle - \sum_{j<\ell} U_j(h)$, by Lemma 6.9. The transformation is almost surely continuous, whence the Continuous Mapping Theorem applies. As $h \to \infty$, the rounding residuals $U_\ell(h)$ now converge in distribution to $U_\ell := \langle \sum_{j<\ell} U_j\rangle - \sum_{j<\ell} U_j$. Clearly U_ℓ is stochastically independent of $W_1,\ldots,W_\ell$.

Finally the proof repeats the whole argument while omitting a component j other than the last, $j < \ell$. Then the convergence is to random variables $\widetilde{U}_k$, $k \neq j$, and the missing component to be filled in is $\widetilde{U}_j := \big\langle \sum_{k\neq j} \widetilde{U}_k\big\rangle - \sum_{k\neq j} \widetilde{U}_k$. Since the limit distribution is unique, we conclude that it is characterized by two properties. First, every $(\ell-1)$-element subset of $U_1,\ldots,U_\ell$ is such that its components are stochastically independent and identically distributed according to a uniform distribution over $[-1/2; 1/2]$. Second, the total component sum is a whole number, $U_+ \in \mathbb{Z}$. As both properties are invariant under permutations, the random variables $U_1,\ldots,U_\ell$ are exchangeable. □

The Invariance Principle exposes the essence of the problem. It invariably applies whatever density governs the weights W. It overcomes diverting technicalities owed to finite house sizes. Its result is the same for all divisor methods that are stationary.

As a first application we establish the universal validity of the discrepancy probabilities $g_\ell(z)$ of the Single-Sum Formula 6.7.b. More applications of the Invariance Principle are to be met in Theorems 7.3 and 7.14, in the next chapter.

6.11. DISCREPANCY LIMIT DISTRIBUTION

Theorem. *Consider the stationary divisor method with split $r \in [0;1]$, and assume that the weights $W_1, \dots, W_\ell$ follow an arbitrary absolutely continuous distribution over the probability simplex Ω_ℓ. Let $h_r = h + \ell(r - 1/2)$ denote the adjusted multiplier.*

Then, for all $z = -\lfloor \ell/2 \rfloor, \dots, \lfloor \ell/2 \rfloor$, the discrepancy probabilities are convergent,

$$\lim_{h\to\infty} P\left(\left\{\left(\sum_{j\le\ell} [h_r W_j]_r\right) - h = z\right\}\right) = g_\ell(z),$$

with limit probabilities $g_\ell(z)$ defined through the Single-Sum Formula 6.7.b.

Proof. Lemma 6.9 and the Invariance Principle 6.10 imply that the discrepancy converges in distribution to $\left\langle \sum_{i<\ell} U_i \right\rangle$, with random variables $U_1, \dots, U_{\ell-1}$ that are stochastically independent and identically distributed according to a uniform distribution over $[-1/2, 1/2]$. The limit probabilities $P\left(\left\{\left\langle \sum_{i<\ell} U_i \right\rangle = z\right\}\right) = P\left(\left\{\sum_{i<\ell} U_i \in [z - 1/2; z + 1/2]\right\}\right)$ are $g_\ell(z)$, as is verified in the proof of Theorem 6.7.c. □

This concludes the excursion into some more technical aspects of apportionment methodology. We are now ready to return to the main topic, what apportionment methods can achieve and how their merits respond to practical needs. The next chapter investigates whether a method produces seat numbers that are unbiased or biased, that is, whether on average parties receive their ideal shares of seats or not.

C H A P T E R 7

Favoring Some at the Expense of Others: Seat Biases

A party's seat bias is a quantitative measure assessing the deviation of the number of seats apportioned to the party from the party's ideal share of seats. Seat bias formulas for stationary divisor methods are calculated when parties are ordered by their vote strengths. The formulas are rather telling, and entail manifold consequences. The divisor method with standard rounding emerges as the unique stationary divisor method treating all parties in an unbiased fashion. In the presence of party alliances the seat bias formulas turn inscrutable. It becomes practically impossible to predict whether it is advantageous or disadvantageous for a party to join an alliance.

7.1. A PARTY'S SEAT EXCESS

Electoral laws are amended only occasionally, they commonly stay the same over quite some time. The question arises whether repeated applications of the same seat apportionment method entail systematic effects that are of interest to the electorate. For example, the divisor method with downward rounding is among the oldest procedures for proportional representation systems. Early on it was recognized that this particular method is biased, in that it favors stronger parties at the expense of weaker parties. This chapter offers detailed formulas for the calculation of seat biases.

Seat biases provide a measure whether, on average, the allocated seats deviate from the ideal share of seats. In contrast, the term "seat excess" captures the behavior in a particular realization. Before turning to the discussion of seat biases in Section 7.4, we begin the inquiry with an analysis of seat excesses. Again we denote by x_j and w_j the number of seats and the vote share of party j, and by h the house size.

Definition. *The* seat excess of party j *is the difference between the number of seats apportioned to the party and its ideal share of seats,* $x_j - w_j h$.

The intention is to average the seat excesses over all vote share vectors w in the probability simplex $\Omega_\ell = \big\{(w_1, \ldots, w_\ell) \in (0;1)^\ell \;\big|\; w_+ = 1\big\}$. The distributional assumptions allow us to concentrate on tie-free instances (Section 6.4). The approach is restricted to stationary divisor methods. The following lemma states that seat excesses comprise three terms, a systematic effect, a contribution to the discrepancy removal, and a rounding residual. Let the *sign function* $\operatorname{sgn}(t)$ attain the values -1, 0, or 1 according as t is negative, zero, or positive.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*, DOI 10.1007/978-3-319-03856-8_7, © Springer International Publishing Switzerland 2014

7.2. SEAT EXCESS TRISECTION

Lemma. *Consider the stationary divisor method with split $r \in [0;1]$. Then, for every party $j \le \ell$, a tie-free seat vector $x \in DivSta_r(h;w)$ has seat excess*

$$x_j - w_j h = \ell\left(r - \frac{1}{2}\right)\left(w_j - \frac{1}{\ell}\right) + \left(x_j - y_j\right) + u_j(h),$$

where $y_j := [h_r w_j]_r$ is the seat number belonging to the adjusted multiplier $h_r = h + \ell(r-1/2)$, and $u_j(h) := y_j - (h_r w_j - r + 1/2)$ is the associated rounding residual.

Furthermore, the modulus $|x_j - y_j|$ is equal to the count how often party j appears among the $|u_+(h)|$ smallest entries of the matrix

$$a_{ni} := \frac{n - 1/2 - \operatorname{sgn}\left(u_+(h)\right)u_i(h)}{w_i} \qquad \textit{for all } n \le \left\lfloor \frac{\ell}{2} \right\rfloor \textit{ and } i \le \ell.$$

Proof. The seat numbers $y_i(h)$ are obtainable via standard rounding, $y_i = \langle h_r w_i - r + 1/2 \rangle$ for all $i \le \ell$, and satisfy $y_+ - h = u_+(h)$ (Lemma 6.9). The trisection identity follows from the fact that the rounding residual $u_j(h)$ satisfies

$$y_j - w_j h = \ell\left(r - \frac{1}{2}\right)\left(w_j - \frac{1}{\ell}\right) + u_j(h).$$

Furthermore, if the discrepancy is zero, $u_+(h) = 0$, then the initial seat vector is final, $x = y$. If the discrepancy is positive, $u_+(h) = y_+ - h \in \{1, \ldots, \lfloor \ell/2 \rfloor\}$, then it is removed by decrementing every party j by $y_j - x_j$ many seats. This forces the multiplier h_r to shrink. As described in Section 4.6, party j contributes as many seats to the discrepancy removal as often as it appears among the $u_+(h)$ highest comparative figures of the array

$$\frac{s_r(y_i - n + 1)}{w_i}, \qquad \text{with } n \le \left\lfloor \frac{\ell}{2} \right\rfloor \text{ and } i \le \ell.$$

The common shift $-h_r = -(h_r w_i)/w_i$ does not change the order. The numerators turn into

$$y_i - n + r - h_r w_i = -\left(n - \frac{1}{2}\right) + y_i - \left(h_r w_i - r + \frac{1}{2}\right) = -\left(n - \frac{1}{2}\right) + u_i(h).$$

Reversing the sign and invoking $\operatorname{sgn}\left(u_+(h)\right) = 1$, we thereafter look for the $u_+(h)$ smallest among the values a_{ni}. This establishes the assertion, if the discrepancy is positive.

If the discrepancy is negative, $u_+(h) = y_+ - h < 0$, then we need to adjoin $|u_+(h)|$ seats; $x_j - y_j$ seats got to party j. With the multiplier growing beyond h_r, party j is awarded as many additional seats as it features among the $|u_+(h)|$ lowest comparative figures of

$$\frac{s_r\left(y_i + n\right)}{w_i}, \qquad \text{with } n \le \left\lfloor \frac{\ell}{2} \right\rfloor \text{ and } i \le \ell.$$

The same shift as before unifies the format of the numerators,

$$y_i + n - 1 + r - h_r w_i = \left(n - \frac{1}{2}\right) + y_i - \left(h_r w_i - r + \frac{1}{2}\right) = \left(n - \frac{1}{2}\right) + u_i(h).$$

Since $\operatorname{sgn}\left(u_+(h)\right) = -1$, the proof is complete. □

The seat excess trisection reflects the problem's complexity. The first term, measuring a method's response to the vote share w_j, is decisive. It is further analyzed in subsequent sections. The second term is party j's contribution to the discrepancy removal. These terms sum to the negative discrepancy, $\sum_{j\leq\ell}(x_j - y_j) = h - y_+$. The third term is the individual rounding effect of party j; their sum is the initial discrepancy, $u_+(h) = y_+ - h$ (Lemma 6.9). The two sums cancel each other.

Theorem 7.3 proves that the second and third terms vanish when the house size grows large and the rounding residuals are averaged out. We use capital letters, $X_j - hW_j$, to indicate consideration of all conceivable vote shares W_j and all ensuing seat numbers X_j, rather than reporting the lone realization of a single instance. As in the Invariance Principle 6.10 we treat the vote shares $W = (W_1, \dots, W_\ell)$ as a random vector with values in the probability simplex Ω_ℓ. Likewise $X = (X_1, \dots, X_\ell)$ is a random vector, with values in the finite solution range $\mathbb{N}^\ell(h)$.

7.3. SYSTEMATIC SEAT EXCESS OF A PARTY

Theorem. *Consider the stationary divisor method with split $r \in [0;1]$, and assume that the vote share vector $W = (W_1, \dots, W_\ell)$ follows an arbitrary absolutely continuous distribution over the probability simplex Ω_ℓ.*

Then, for every party $j \leq \ell$, the conditional expectations of the seat excess $X_j - W_j h$ given the vote shares $W = w$ converge in distribution,

$$\mathrm{E}\left(X_j - hW_j \,\middle|\, W_1 = w_1, \dots, W_\ell = w_\ell\right) \xrightarrow[h\to\infty]{\text{in distribution}} \ell\left(r - \frac{1}{2}\right)\left(w_j - \frac{1}{\ell}\right).$$

Proof. In view of the absolutely continuous vote share distribution all seat vectors are almost surely tie-free. Hence the Seat Excess Trisection 7.2 applies. The first term determines the limit,

$$\mathrm{E}\left(\ell\left(r - \frac{1}{2}\right)\left(W_j - \frac{1}{\ell}\right) \,\middle|\, W = w\right) = \ell\left(r - \frac{1}{2}\right)\left(w_j - \frac{1}{\ell}\right).$$

The second term is the seat adjustment $Z_j(h) := X_j - Y_j$. The Invariance Principle 6.10, with its variables $U = (U_1, \dots, U_\ell)$ and W, lets $Z_j(h)$ converge to $Z_j(U, W) := \mathrm{sgn}(U_+)N_j(U, W)$ where $N_j(U, W)$ counts how often party j appears among the $|U_+|$ smallest entries of the array

$$b_{ni} := \frac{n - 1/2 - \mathrm{sgn}(U_+)U_i}{W_i}, \quad \text{with } n \leq \left\lfloor \frac{\ell}{2} \right\rfloor \text{ and } i \leq \ell.$$

Due to the independence of U and W the limit of the conditional expectations is the unconditional expectation of $Z_j(U, w)$,

$$\mathrm{E}\left(Z_j(h) \,\middle|\, W = w\right) \xrightarrow[h\to\infty]{\text{in distribution}} \mathrm{E}\left(Z_j(U, W) \,\middle|\, W = w\right) = \mathrm{E}\big(Z_j(U, w)\big).$$

The variables $Z_1(U, w), \dots, Z_\ell(U, w)$ inherit exchangeability from $U_1, \dots, U_\ell$. Their common expectation is $\ell\,\mathrm{E}\big(Z_j(U, w)\big) = \mathrm{E}\big(Z_+(U, w)\big) = -E(U_+) = 0$, thus entailing $\mathrm{E}\big(Z_j(U, w)\big) = 0$.

The third term is the rounding residual $U_j(h)$. Its limit U_j is stochastically independent of W. Hence the conditional expectation vanishes in the limit,

$$\mathrm{E}\left(U_j(h) \,\middle|\, W = w\right) \xrightarrow[h\to\infty]{\text{in distribution}} \mathrm{E}\left(U_j \,\middle|\, W = w\right) = \mathrm{E}(U_j) = 0. \qquad \square$$

The systematic seat excess admits an appealing heuristic explanation. Within the integer intervals $[n-1;n]$, the rounding operation takes place at the split point r rather than at $1/2$, and party j misses its due share by $-(r-1/2)$ seat fractions. Every other party meets the same fate, and so a total of $\ell(r-1/2)$ seat fractions accrue. After regaining its proportional share w_j from the total, party j is left with the balance $w_j\ell(r-1/2)-(r-1/2)=\ell(r-1/2)(w_j-1/\ell)$, its systematic seat excess. Altogether the systematic seat excesses even out, because of $\sum_{j\le\ell}(w_j-1/\ell)=0$. If some parties are advantaged then others are disadvantaged. If some parties are disadvantaged then others are advantaged. One man's meat is another man's poison.

Only the divisor method with standard rounding ($r=1/2$) has systematic seat excesses that vanish for all parties under the provision that the house size is sufficiently large ($h\to\infty$). A single realization must endure a rounding effect simply because it is unavoidable that an interim quotient must be rounded to a whole number. But on average the method apportions to all parties as many seats as the ideal shares demand. No party gains a systematic profit, and no party falls victim to a systematic deficit.

Stationary divisor methods with a split larger than one-half, $r>1/2$, behave differently. Parties above average strength, with a vote share $w_j>1/\ell$, enjoy a positive seat excess. Barring rounding effects the methods promise them more seats than their ideal shares justify. For an extremely strong party, with $w_j\approx 1$, the excess comes close to $(\ell-1)(r-1/2)$ seat fractions. It increases with the number of participating parties, ℓ. These bonus seats are counterbalanced by seat losses of below-average parties. Below-average parties, with vote shares $w_j<1/\ell$, face deficits of up to half a seat. The deficit bound $-1/2$ is immediate from inserting an extremely weak party, $w_j\approx 0$:

$$\ell\left(r-\frac{1}{2}\right)\left(w_j-\frac{1}{\ell}\right)\ge-\left(r-\frac{1}{2}\right)\ge-\frac{1}{2}.$$

For example consider the divisor method with downward rounding ($r=1$). A party with one-third of the votes, $w_j=1/3$, may look forward to an excess of $(\ell-3)/6$ seat fractions. In a system with nine parties, $\ell=9$, the party may look forward to an excess of a full seat. The bonus seat is counterbalanced by two or more parties with less than one-ninth of the votes that miss their ideal shares by up to half a seat.

Stationary divisor methods with split point $r<1/2$ exhibit a complementary behavior. Above-average parties encounter a seat deficit that grows linearly with the number of participants, ℓ. Below-average parties are advantaged, but their predicted surpluses are bounded from above by half a seat.

The bipartition of the ensemble of all parties into two groups, namely the stronger group of parties with an above-average vote share and the complementary weaker group of parties with a below-average vote share, is a bit crude. A more sensitive approach is to rank parties according to their vote strengths, and then look at them separately. Ranking parties by vote strength is also more informative from a practical viewpoint.

7.4. RANK-ORDER OF PARTIES BY VOTE SHARES

It is of great practical interest whether an apportionment method possesses a systematic trend to handle stronger and weaker parties differently. The attributes "strong" and "weak" solely refer to voter support, in terms of vote counts or vote shares. Other conceivable indices are neglected, such as party budget, number of party members, campaign contributions, media presence, or the like.

We order parties by their vote shares from strongest to weakest, $w_1 \geq \cdots \geq w_\ell$. As a visual reminder we replace the otherwise favored party subscript j by the letter k when denoting rank-ordered vote shares w_k. Thus the subscript k turns into the party's rank-score: $k = 1$ is the strongest party, $k = 2$ is the second-strongest party, and the weakest party is last, $k = \ell$. Naturally the parties' rank-order can be ascertained only retrospectively, when all vote counts are available. Moreover, a party's rank-order generally varies from one election to the next. It may not be the same party that turns out strongest in one election or the other.

7.5. SEAT BIASES

Definition. *The* seat bias of the kth-strongest party *is the conditionally expected seat excess of party k given that the parties are rank-ordered,*

$$\mathrm{E}^{(0)}\big(X_k - hW_k\big) := \mathrm{E}\big(X_k - hW_k \;\big|\; W_1 \geq \cdots \geq W_\ell \geq 0\big),$$

for all $k \leq \ell$. The superscripted notation $\mathrm{E}^{(0)}$ reminds us of the conditioning event $\{W_1 \geq \cdots \geq W_\ell \geq 0\}$.

The seat bias of a party is the average of its seat excesses. The average is taken over the vote shares that preserve the party's rank-score. With a positive seat bias, a party expects to be alloted more seats than justified by its ideal share. A negative seat bias tells the party that on average its seat numbers fall short of the ideal share.

Seat biases are averages, and so they are measured in fractions, not in whole numbers. Theorem 7.7 shows that in a three-party system the divisor method with downward rounding entails seat biases of $5/12$ seat fractions for the strongest party, $-1/12$ for the middle party, and $-4/12$ for the weakest party. Thus the strongest party may expect five seats on top of its ideal share in the course of twelve elections. The five bonus seats are gathered together at the expense of the middle party that has to give up one seat, and of the weakest party that faces a loss of four seats.

The sum of the seat biases of all parties vanishes,

$$\sum_{k=1}^{\ell} \mathrm{E}^{(0)}\big(X_k - hW_k\big) = \mathrm{E}^{(0)}\big(X_+ - hW_+\big) = h - h = 0.$$

Hence either all seat biases are zero; in this case we call the apportionment method *unbiased*. Or else the method is *biased*, meaning that some parties enjoy a positive seat bias and other parties suffer from a negative seat bias. Favoritism of some spawns distress for others. Our wording always acknowledges the existence of both groups, those advantaged and those disadvantaged.

Seat bias formulas are laborious to obtain when the house size is finite. The labor is dispensable, and we skip the details. Empirical data and computational evidence are sufficiently reassuring that the limit formula for infinite house size is also applicable for finite house sizes. Theorem 7.3 yields the formula

$$\lim_{h\to\infty} \mathrm{E}^{(0)}\big(X_k - hW_k\big) = \ell\left(r - \frac{1}{2}\right)\left(\mathrm{E}^{(0)}\big(W_k\big) - \frac{1}{\ell}\right)$$

for all $k \leq \ell$. Evidently the seat bias formula inherits the structure of the systematic seat excess. A seat excess inserts the realized vote share w_k and can be evaluated *ex post* only. Quite differently, the seat bias relies on the average over all conceivable vote shares of a party with rank-order k, and thus constitutes an *ex ante* index.

7.6. PERCENTAGE HURDLES FOR VOTE EFFECTIVENESS

A last practical issue needs to be addressed. Some electoral systems impose a threshold t that vote shares must meet or exceed in order to participate in the apportionment process,

$$w_j \geq t \quad \text{for all} \quad j \leq \ell.$$

In the presence of ℓ parties the threshold $t = 1/\ell$ makes all vote shares equal, $w_j = 1/\ell$ for all $j \leq \ell$, and is of no practical interest. Values greater than $1/\ell$ are impossible. Therefore, thresholds are restricted to lie in the right-open interval $[0; 1/\ell)$.

Vote share thresholds are a model for the type of percentage hurdles encountered in practice. For example, Table 1.31 lists instances of no threshold ($t = 0$), and thresholds of 1.8 percent, or three, four, or five percent. The percentages refer to votes cast, or to valid votes. Since the modeling level always operates with effective votes, we no longer distinguish between votes cast, valid votes, and effective votes. We trust that the change of reference sets is negligible.

It should be emphasized that a vote share threshold excludes not just some parties from the apportionment calculations. Rather, it is a number of voters who are deprived from obtaining representation in parliament. Chapter 1 documents the calamity by recording the number of voters who fall victim to an imposed vote share threshold.

Theorem 7.7 calculates the seat bias of the kth-strongest party in the presence of a t-threshold, indicated by the superscripted notation $\mathrm{E}^{(t)}$,

$$\mathrm{E}^{(t)}\big(X_k - hW_k\big) := \mathrm{E}\big(X_k - hW_k \;\big|\; W_1 \geq \cdots \geq W_\ell \geq t\big).$$

The vote share vector W is assumed to be uniformly distributed. More precisely, the uniformity assumption is needed on the ordered and truncated subset $\{W_1 \geq \cdots \geq W_\ell \geq t\}$ only, not on the full probability simplex Ω_ℓ. The restricted uniformity assumption delimits the vote share of the kth-strongest party to lie between those of its two neighbors, $W_k \in [W_{k-1}; W_{k+1}]$. The smaller range of variation promises an expectation more stable with respect to a deviation from uniformity. This reasoning may explain the surprisingly vast domain of validity of the seat bias formula.

7.7. SEAT BIAS FORMULA

Theorem. *Consider the stationary divisor method with split* $r \in [0; 1]$ *and a threshold* $t \in [0; 1/\ell)$, *and assume the vote share vector* $W = (W_1, \dots, W_\ell)$ *to be uniformly distributed over the sub-simplex* $\{W_1 \geq \cdots \geq W_\ell \geq t\}$.

Then the seat bias of the kth-strongest party converges to

$$B_r^{(t)}(k) := \lim_{h \to \infty} \mathrm{E}^{(t)}\big(X_k - hW_k\big) = \left(r - \frac{1}{2}\right)\left(H_k^\ell - 1\right)(1 - \ell t)$$

for all $k \leq \ell$, *where* $H_k^\ell := \sum_{n=k}^{\ell}(1/n)$ *is a partial sum of the harmonic series.*

Proof. Section 7.5 provides the formula $B_r^{(t)}(k) = \ell(r - 1/2)\big(\mathrm{E}^{(t)}(W_k) - 1/\ell\big)$. Setting $V_k := (W_k - t)/(1 - \ell t)$ for all $k \leq \ell$, the vector $V := (V_1, \dots, V_\ell)$ has non-negative components summing to unity. Since W generates V by translation and scaling, V inherits the uniform distribution over its range $\{V_1 \geq \cdots \geq V_\ell \geq 0\}$. Substitution of W_k by V_k yields $\mathrm{E}^{(t)}(W_k) = t + (1 - \ell t)\,\mathrm{E}^{(0)}(V_k)$, that is, the threshold for V is zero. The expectation of V is the centroid of its range and hence emerges as the arithmetic mean of the range's vertices,

$$\mathrm{E}^{(0)}\big(V\big) = \frac{1}{\ell}\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} + \frac{1}{\ell}\begin{pmatrix}1/2\\1/2\\0\\\vdots\\0\end{pmatrix} + \frac{1}{\ell}\begin{pmatrix}1/3\\1/3\\1/3\\\vdots\\0\end{pmatrix} + \cdots + \frac{1}{\ell}\begin{pmatrix}1/\ell\\1/\ell\\1/\ell\\\vdots\\1/\ell\end{pmatrix} = \frac{1}{\ell}\begin{pmatrix}H_1^\ell\\H_2^\ell\\H_3^\ell\\\vdots\\H_\ell^\ell\end{pmatrix}.$$

This yields $E^{(0)}(V_k) = (1/\ell)H_k^\ell$, and $\mathrm{E}^{(t)}(W_k) - 1/\ell = t + (1/\ell - t)H_k^\ell - 1/\ell = (H_k^\ell - 1)(1/\ell - t)$. Insertion into the formula of Section 7.5 completes the proof. □

Since the factor that depends on the split point r satisfies $r - 1/2 = -\big((1-r) - 1/2\big)$, the mirror image $1 - r$ has a seat bias of opposite sign,

$$B_r^{(t)}(k) = -B_{1-r}^{(t)}(k).$$

Hence the stationary divisor method with split r is paired with the stationary divisor method with split $1 - r$. While the method with a split larger than one-half, $r > 1/2$, favors stronger parties at the expense of weaker parties, the method with split $r < 1/2$ favors weaker parties at the expense of stronger parties. The unique unbiased method is the divisor method with standard rounding, $B_{1/2}^{(t)}(k) = 0$ for all $k \leq \ell$. Every party may expect its due share of seats. No party is advantaged, nor is any party disadvantaged.

The last factor $1 - \ell t$ is linear in the threshold parameter t. It discounts the seat biases of all parties in a like manner. Practical thresholds are so small that the factor is insignificant numerically. In the sequel we neglect the vote share threshold and interpret the seat bias formula with $t = 0$,

$$B_r^{(0)}(k) = \left(r - \frac{1}{2}\right)\left(H_k^\ell - 1\right).$$

7.8. BIASEDNESS VERSUS UNBIASEDNESS

A prominent pair of biased apportionment methods are the divisor methods with downward rounding ($r = 1$), and the divisor method with upward rounding ($r = 0$). The noticeable seat bias of the divisor method with downward rounding has been appreciated ever since the inception of the method, in the eighteenth century in the United States of America, and in the nineteenth century in continental Europe. The monograph *Balinski / Young* (1982) tells the tale of the New World, how arguments swayed back and forth to promote a preferred apportionment method over its competitors. Among the plot's protagonists we meet illustrious figures such as *Thomas Jefferson, Alexander Hamilton, Daniel Webster*, to name but a few.

In the Old World the divisor method with downward rounding was resolutely advocated by the Belgian activist *Victor D'Hondt* and his Swiss contemporary *Eduard Hagenbach-Bischoff*. Once the method got firmly rooted in the minds of proportional representation protagonists, it was praised with all sorts of fanciful claims: that it is the only apportioning method existing, that it is the simplest method, that it is the method most widely used. Of these claims, the last is correct. The frequency in Table 1.31 is indicative of the method's predominance. Sixteen of the 27 Member States of the European Union use it for their EP elections.

George Pólya (1918) was first to investigate seat biases in a formal manner. For the divisor method with downward rounding ($r = 1$) in a three-party system he found the biases of the strongest, middle, and weakest parties to be

$$B_1^{(0)}(1) = \frac{5}{12}, \qquad B_1^{(0)}(2) = -\frac{1}{12}, \qquad B_1^{(0)}(3) = -\frac{4}{12}.$$

Pólya emphasized that he considered his calculations

> das wichtigste Resultat dieser Abhandlung.
>
> the most important result of this treatise.

He also pointed to the availability of unbiased procedures, such as the divisor method with standard rounding, or the Hare-quota method with residual fit by greatest remainders.

Unbiased apportionment methods are often preferred in the second stage of a two-stage systems, as evidenced in the Netherlands and in Poland (Section 1.8). In the Netherlands, sub-apportionments become necessary when parties join into an alliance. The promise of an unbiased sub-apportionment makes it easier for stronger parties to persuade weaker parties to join into a prospective alliance. Quantitatively, though, the weaker partners amplify the strength of the stronger partner. The irrationalities of alliances are discussed in Section 7.11.

In Poland the seats from the super-apportionment are passed on to the various district lists of the parties. Within a party it would seem strange to bias the sub-apportionment in favor of larger districts at the expense of smaller ones. As a consequence the sub-apportionments use an unbiased apportionment method. The within-party (sub-)apportionments strictly obey the motto "One person, one vote", while the between-party (super-)apportionment exerts noticeable seat biases. *Quod licet Iovi, non licet bovi.* Whether double standards are allowed or forbidden is ultimately decided by constitutional law rather than formal analyzes.

More variation is met when the seats of a house are apportioned among regional districts in proportion to census figures. The seats of the French Assemblée nationale are allocated to the 101 Départements using the divisor method with upward rounding, see *Balinski* (2004 [92]). The method is biased in favor of sparsely populated (rural) districts at the expense of densely populated (urban) districts. The bias has a straightforward explanation in terms of power politics. All that is needed is that the parliamentary majority finds their voter support in the countryside outnumbering that in the cities. Then over-representation of smaller districts and underrepresentation of larger districts appears to be the method of choice.

A glaring contraposition of smaller versus larger districts materialized in imperial Germany. Members of the Reichstag were elected in single-seat constituencies. While the constituencies were initially of almost equal size, industrialization and rural depopulation eventually led to blatant differences in size. However, the induced bias accorded with political cleavages. Typically the character of smaller constituencies was rural, of larger constituencies, metropolitan. Since the conservative Reichstag majority catered to the need of agrarian freeholders more than to the interests of industrial masses, there was no political incentive for them to redraw constituency boundaries.

Unbiasedness is not a virtue by itself. The central issue is its status in the relevant constitution, how much nonzero bias is constitutionally tolerable. While the tasks of apportioning seats among political parties, or of apportioning seats among geographical districts are the same formally, their constitutional appraisal may differ significantly. Many countries encompass some sparsely populated districts that, under standard rounding or downward rounding or any other pervious rounding rule, would end up with no parliamentary representation at all. For this reason many constitutions guarantee every district a minimum representation.

For example, Article I, Section 2 of the Constitution of the United States stipulates that *each state shall have at least one representative.* The clause may be understood to inject a mild dose of biasedness in favor of smaller states at the expense of larger states. Indeed, the 435 seats of the United States House of Representatives are apportioned to the 50 states of the Union using the divisor method with geometric rounding. Geometric rounding is impervious and guarantees every state at least one seat, as demanded by the Constitution. In the family power-mean divisor methods, geometric rounding ($p = 0$) and standard rounding ($p = 1$) are quite close to each other. Their differences are captured by the majorization order, to be discussed in Chapter 8.

For infinite house sizes the diversity of the power-mean family reduces to the three stationary divisor methods that are based on upward rounding, standard rounding, and downward rounding (Section 3.12). Thus the asymptotic results of Theorem 7.7 are inconclusive for distinguishing between the divisor methods with geometric rounding, and with standard rounding. Both methods are asymptotically unbiased. But the application is to the house size $h = 435$ that is finite, not asymptotic. This raises the question which house sizes h appear to be large enough for the asymptotic seat bias formulas $B_r^{(t)}(k)$ to acquire validity.

7.9. HOUSE SIZE RECOMMENDATION

In this section we argue in support of the following house size recommendation:

> *The seat bias formula $B_r^{(t)}(k)$ is applicable for all practical purposes provided the house size meets or exceeds twice the number of parties, $h \geq 2\ell$, even though Theorem 7.7 requires the assumption of large house sizes, $h \to \infty$.*

Seat biases promptly approach the limit value $B_r^{(t)}(k)$ for house sizes $h \geq 2\ell$. Hence each deployment of the method incurs a bias of constant size. Each time the method is used the kth-strongest party faces a bias of $B_r^{(t)}(k)$ seat fractions.

Empirical data reinforce the house size recommendation. *Schuster / Pukelsheim / Drton / Draper* (2003 [657]) present empirical three-party seat excess distributions calculated from $N = 49$ parliamentary elections in the German State of Bavaria. During 1966–1998 Bavaria conducted nine elections of its diet. In seven elections just three parties pass the five percent threshold and qualify for the seat apportionment. Since Bavaria is subdivided into seven electoral districts the data provide a sample of size $N = 7 \times 7 = 49$. House sizes varied between 19 and 65 seats. The voting attitudes of the electorate may be considered stable enough to allow the assumption that the seat excesses originate from the same distribution. The theoretical seat biases $B_r^{(t)}(k)$ are shown to conform rather satisfactorily with the empirical data. Another data set provided by those authors is from the Swiss Canton of Solothurn. During the period 1896–1997, the ten Solothurn districts provide a sample of $N = 143$ three-party elections, with house sizes ranging from 7 to 29 seats. Again theoretically calculated seat biases agree exceedingly well with the empirically observed seat biases. The remaining deviations are plausibly attributed to rounding effects and random variation.

The recommendation excludes house sizes below 2ℓ. With too small a house size the seat excesses $x_k - hw_k$ vary erratically. Even their averages inhibit the identification of a systematic trend. The lamento of *Riker* (1982 [260]) of what happens when $h = 10$ seats are apportioned among $\ell = 12$ parties is void.

The exclusion of house sizes $h < 2\ell$ is an example of the many boundary effects that are present in apportionment methodology. Section 4.4 mentions special provisions when an impervious divisor method meets too small a house sizes, $h < \ell$. Section 5.10 compares various quota definitions by restricting attention to sufficiently large vote totals, $v_+ > h(h+1)$. Lemma 6.5 calculates seat total distributions assuming multipliers $\mu > \ell r$ large enough in order not to dilute an already longish proof.

7.10. TOTAL POSITIVE BIAS: THE STRONGER THIRD, THE WEAKER TWO-THIRDS

The discriminating power for the seat biases $B_r^{(t)}(k)$ emanates from the factor $H_k^\ell - 1$, where $H_k^\ell = \sum_{n=k}^{\ell}(1/n)$. If the factor is positive, the kth-strongest party is advantaged for splits $r > 1/2$, and disadvantaged for splits $r < 1/2$. If it is negative, the advantages switch sign. The sign switches when $H_k^\ell = 1$. The logarithmic approximation

$$H_k^\ell = \sum_{n=k}^{\ell} \frac{1}{n} \approx \int_k^\ell \frac{1}{x}\, dx = \log \ell - \log k$$

determines the approximate *no-bias rank-score* $k_\ell := \langle \ell/e \rangle = \langle 0.37\ell \rangle \in \{1, \ldots, \ell-1\}$. Roughly the stronger third of the parties has a positive rank-score factor. The weaker two-thirds have a negative factor. Hence the stronger third enjoys positive seat biases, while the weaker two-thirds is exposed to negative seat biases. The trend is familiar from other areas of life: a minority of well-to-do enjoy a surplus, at the expense of a majority of not-so-well-to-do who must come to grips with less than their fair share.

The approximation $H_k^\ell \approx \log(\ell + 1/2) - \log(k - 1/2) \approx \log \ell - \log k$ implies that the seat biases decrease logarithmically as a function of the rank-scores $k \le \ell$. Conversely we may fix the rank-score, k, and consider a growing number of competing parties, ℓ. For example we may be interested in the front runner, $k = 1$, when more and more weaker competitors join the campaign. Then seat biases increase logarithmically, according to $\log \ell$.

An index that captures the aggregate seat biases of a system of ℓ parties is the *total positive bias, TPB*. It is the sum of all positive seat biases accumulating among ℓ parties under a stationary divisor method with split $r \in [0; 1]$,

$$TPB(\ell, r) := \left(r - \frac{1}{2}\right) \sum_{k=1}^{k_\ell} \left(H_k^\ell - 1\right) \approx \left(r - \frac{1}{2}\right) \left(\frac{\ell}{e} - 1\right).$$

The approximate equality is established under the assumption $\ell \ge 5$, whence $k_\ell \ge 2$. For $k \le k_\ell$ we have $H_k^\ell = H_k^{k_\ell - 1} + H_{k_\ell}^\ell$, and get

$$\sum_{k=1}^{k_\ell} \left(H_k^{k_\ell - 1} + H_{k_\ell}^\ell - 1\right) \approx \left(\sum_{n=1}^{k_\ell - 1} \sum_{k=1}^{n} \frac{1}{n}\right) + k_\ell \left(\log \ell - \log \frac{\ell}{e}\right) - k_\ell = k_\ell - 1.$$

Approximating the no-bias rank-score $k_\ell = \langle \ell/e \rangle$ by ℓ/e yields the term $\ell/e - 1$. In summary, the trisection by party strength maximizes the system's total positive bias. It is employed for discriminatory purposes already by *Balinski / Young* (1982 [75, 126]).

7.11. ALLIANCES OF LISTS

The advocates of the divisor method with downward rounding recognized early on that the method is compromised by being biased in favor of stronger parties at the expense of weaker parties. To console weaker participants, *Hagenbach-Bischoff* (1896) devised the construct of an *alliance of lists* (German: Listenverbindung, French: apparentement des listes). His intention was to allow weaker parties to gather strength by joining their lists into an alliance, and thereby to evade or at least to lessen any threatening seat biases. In electoral systems with multiple districts it may well happen that in one district parties A and B join into a first alliance and parties C and D into a second alliance, while in another district parties A and C register an alliance and parties B and D stand alone, and in a third district yet another partition is realized. It is more to the point to speak of alliances of lists rather than of alliances of parties.

BY2008Friedberg	Votes	Quotient	DivDwn
Alliance $\{2, 3, 5\}$	194 141	16.04	16
List 1	150 615	12.4	12
List 4	28 428	2.4	2
List 6	12 010	0.99	0
Sum (Divisor)	385 194	(12 100)	30

Partners	Votes	Quotient	DivDwn
Alliance $\{2, 3, 5\}$*: Sub-apportionment*			
List 2	145 292	13.2	13
List 3	30 558	2.8	2
List 5	18 291	1.7	1
Sum (Divisor)	194 141	(11 000)	16

TABLE 7.1 *Discordance victory via an alliance, divisor method with downward rounding.* The alliance gets 16 seats; 13 of them go to party 2. The end result is discordant: List 2 has fewer votes but more seats than list 1. Without alliance the divisor is 12 000, and party 2 is allotted 12 seats.

Many electoral laws permit parties to register alliances, in particular when the law stipulates the notoriously biased divisor method with downward rounding. The stipulations persist even when an unbiased apportionment method is used. It is a safe bet that such laws have precursors that formerly used the divisor method with downward rounding, and that the unbiased method was adopted only later on. The legislators who prepared the amendment lacked insight and forgot to ban alliances. Alliances radiate a seductive charm for party officials, as a smart lever to divert the electoral outcome into a more desirable direction. From the voters' viewpoint alliances interfere with the principle of a direct election since calculations get obscured by a sizable dose of indirectness. The functioning of alliances is non-transparent indeed.

An alliance triggers a two-stage apportionment calculation. In a first stage, called super-apportionment in Chapter 1, the seats are apportioned among the alliances that are registered with the electoral bureau, and among the lists that stand alone. Every alliance then calls for a second stage. This is a separate calculation, called sub-apportionment, in order to distribute the alliance's joint seats among its partners. The superposition of the biases accounts for the lack of transparency.

The 2008 local election in the Bavarian City of Friedberg may serve as an example, see Table 7.1. Six parties campaigned for the city council. As usual we rank-order the parties from 1 (strongest) to 6 (weakest). The second-, third-, and fifth-strongest lists registered an alliance, $\{2, 3, 5\}$. The other parties stood for themselves. In technical language the other parties formed a *singleton alliance* each: $\{1\}$, $\{4\}$, and $\{6\}$. The vote count aggregate for the alliance $\{2, 3, 5\}$ exceeded the vote count for party 1, parties 4 and 6 trailed behind. Hence the four alliances finished in the rank-order $\{2, 3, 5\}$, $\{1\}$, $\{4\}$, $\{6\}$. The final seat apportionment admits three views.

First, party 2 emerged as the ultimate winner. In the super-apportionment between alliances, the alliance $\{2, 3, 5\}$ finished with 16 seat compared to the cumulative 15 seats they would have won separately. In the within-partner sub-apportionment the bonus seat went to the strongest of the three partners, party 2. Thus party 2 won a seat it would not have won without the alliance. The end result is a discordant seat assignment. Party 2 ended up having fewer votes but more seats than party 1. Second, *Hagenbach-Bischoff* nurtured the consoling view that the purpose of the alliance was to move parties 3 and 5 from back to front. It did, but it made no difference. They received as many seats with the alliance as they would have received without. Third, there is the desolate experience of the trailing party 6. Without an alliance, party 6 would have won a seat to obtain representation in the council. Alas, the alliance pocketed its seat. Party 6 did nothing but look on, yet had to pay the bill.

In the general set-up an alliance L is a subset of parties, $L \subseteq \{1,\dots,\ell\}$. Each stand-alone party j constitutes a singleton alliance, $\{j\}$. The ensemble of all alliances yields a partition of the party system $\{1,\dots,\ell\}$. Let a be the number of alliances, that is, the number of elements in the partition. In the example, the six-party system $\{1,2,3,4,5,6\}$ is partitioned into $a = 4$ alliances, namely $L_1 = \{2,3,5\}$, $L_2 = \{1\}$, $L_3 = \{4\}$, and $L_4 = \{6\}$. The following seat bias formula disregards the vote share threshold by setting $t = 0$.

7.12. SEAT BIAS FORMULA AND ALLIANCES OF LISTS

Corollary to Theorem 7.7. *Suppose the party system $\{1,\dots,\ell\}$ is partitioned into a alliances of lists $L_1,\dots,L_a$. If the kth-strongest of the ℓ parties is among the p partners of the alliance L, that is $k \in L$ and $\#L = p$, then the seat bias of the kth-strongest party is practically approximated by*

$$B_r\left(k \mid L, a\right) := \left(r - \frac{1}{2}\right)\left(\left(\frac{a}{\ell} + \frac{p-1}{\sum_{i\in L} H_i^\ell}\right)H_k^\ell \;-\; 1\right).$$

Proof. Theorem 7.7 provides the formula $\lim_{h\to\infty} \mathrm{E}^{(0)}(X_k - hW_k) = (r-1/2)\big(\ell\,\mathrm{E}^{(0)}(W_k)-1\big)$. Let $W_L := \sum_{i\in L} W_i$ denote the aggregated vote shares of alliance L. The super-apportionment entails for alliance L the seat bias

$$\mathrm{E}^{(0)}\big(X_L - hW_L\big) \approx \left(r - \frac{1}{2}\right)\left(a\,\mathrm{E}^{(0)}\big(W_L\big) - 1\right).$$

Of this bias, party k carries its proportionate share $\mathrm{E}^{(0)}(W_k/W_L)$. Within alliance L, the sub-apportionment has $X_L := \sum_{i\in L} X_i$ seats to allocate. With a view towards its p partners, party k is exposed to the seat bias

$$\mathrm{E}^{(0)}\left(X_k - X_L\frac{W_k}{W_L}\right) \approx \left(r - \frac{1}{2}\right)\left(p\,\mathrm{E}^{(0)}\left(\frac{W_k}{W_L}\right) - 1\right).$$

The two stages yield the aggregate bias

$$\begin{aligned}
&\mathrm{E}^{(0)}\left(\frac{W_k}{W_L}\right)\left(r - \frac{1}{2}\right)\left(a\,\mathrm{E}^{(0)}\big(W_L\big) - 1\right) + \left(r - \frac{1}{2}\right)\left(p\,\mathrm{E}^{(0)}\left(\frac{W_k}{W_L}\right) - 1\right)\\
&= \left(r - \frac{1}{2}\right)\left(a\,\mathrm{E}^{(0)}\left(\frac{W_k}{W_L}\right)\mathrm{E}^{(0)}\big(W_L\big) + (p-1)\,\mathrm{E}^{(0)}\left(\frac{W_k}{W_L}\right) - 1\right).
\end{aligned}$$

Under the uniformity assumption of Theorem 7.7 the random variables W_k/W_L and W_L are stochastically dependent to such a low degree that the dependence may be neglected, for all practical purposes. Approximate independence entails

$$\mathrm{E}^{(0)}\left(\frac{W_k}{W_L}\right)\mathrm{E}^{(0)}\big(W_L\big) \approx \mathrm{E}^{(0)}\big(W_k\big) = \frac{1}{\ell}\,H_k^\ell, \quad \text{and} \quad \mathrm{E}^{(0)}\left(\frac{W_k}{W_L}\right) \approx \frac{\mathrm{E}^{(0)}\big(W_k\big)}{\mathrm{E}^{(0)}\big(W_L\big)} = \frac{H_k^\ell}{\sum_{i\in L} H_i^\ell}.$$

Insertion of these approximations yields $B_r(k \mid L, a)$. □

The cases $a = 1$ and $a = \ell$ are instructive though irrelevant. When there is a single alliance, $a = 1$, there are $p = \ell$ partners and the super-apportionment becomes redundant. Accordingly the seat biases stay the same, $B_r(k \mid \{1, \dots, \ell\}, 1) = B_r^{(0)}(k)$, since $\sum_{i \le \ell} H_i^\ell = \ell$. When all alliances are singletons, $a = \ell$, then each party stands on its own, $p = 1$, and it is the sub-apportionment that is redundant. Again the seat biases stay put, $B_r(k \mid \{k\}, \ell) = B_r^{(0)}(k)$.

The relevant cases have $1 < a < \ell$. There are several alliances, and at least one of them has two or more partners. The number of feasible partitions grows rapidly. In a system with six parties there are 201 partitions, with seven parties it is 875. The mere information which alliances are registered, often tucked away in a small-print footnote on the ballot sheet, does not tell voters whether the party of their choice is going to gain or lose. Party officials would not know either (unless they evaluate the above formula), but they are destined to play the game.

Here are the rules of the game, tailored to the divisor method with downward rounding ($r = 1$).

Rule 1: If you and your partners are the only alliance, you win. With just one non-singleton alliance L, all of its $p \ge 2$ partners $k \in L$ increase their seat biases,

$$B_1(k \mid L, 1 + \ell - p) = B_1^{(0)}(k) + \frac{p-1}{2} \left(\frac{1}{\sum_{i \in L} H_i^k} - \frac{1}{\ell} \right) H_k^\ell > B_1^{(0)}(k).$$

A positive seat bias grows more positive, a negative seat bias becomes less negative. For example, in the 2008 communal elections in Bavaria alliances got registered in 668 communities. Two-thirds of the communities (456) saw just one non-singleton alliance. Rule 1 says that the alliances' partners could look forward to bonus seats simply because their competitors were napping.

Rule 2: If you refrain from joining an alliance, you lose. A party that maintains its independence and does not join into an alliance, $k \in \{k\}$, decreases its seat bias,

$$B_1\big(k \mid \{k\}, a\big) = B_1^{(0)}(k) - \frac{1}{2} \left(1 - \frac{a}{\ell}\right) H_k^\ell < B_1^{(0)}(k).$$

A stronger party loses some of its positive seat bias, a weaker party exacerbates its negative seat bias. The seats thus released benefit competing parties.

Rule 3: If there are competing alliances, you gamble. In the presence of several non-singleton alliances party officials need to evaluate the seat bias formula $B_1(k \mid L, a)$ to get a feeling of what to expect. The bias to which the party is exposed to may be beneficial, disadvantageous, or neutral. It is a lottery.

7.13. SEAT BIASES OF SHIFT-QUOTA METHODS

Pólya (1918) already pointed out that the Hare-quota method with residual fit by greatest remainders is unbiased, as is the divisor method with standard rounding (Section 7.8). The following proof runs parallel to the analysis of stationary divisor methods. The arguments gain in lucidity by considering the larger family of shift-quota methods, shQgrR_s, with shift $s \in [-1; 1)$ (Section 5.4). Recall that the shift-quota is $Q(s) = v_+/(h + s)$. Specifically, the Hare-quota method with residual fit by greatest remainders emerges for a vanishing shift, $\text{shQgrR}_0 = \text{HaQgrR}$.

Theorem. *Consider the shift-quota method with shift* $s \in [-1; 1)$.

a. *A tie-free solution* $x \in shQgrR_s(h; w)$ *has seat excess trisection*

$$x_j - w_j h = s\left(w_j - \frac{1}{\ell}\right) + \left(x_j - y_j - \left(\frac{1}{2} - \frac{s}{\ell}\right)\right) + u_j(h),$$

where $y_j := \lfloor w_j(h+s) \rfloor = \langle w_j(h+s) - 1/2 \rangle$ *is the number of seats allocated to party* j *by the main apportionment, and* $u_j(h) := y_j - \big((h+s)w_j - 1/2\big) \in [-1/2; 1/2]$ *is the associated rounding residual.*

b. *If the vote share vector* W *follows an absolutely continuous distribution, then the systematic seat bias of party* $j \le \ell$ *is given by*

$$\mathrm{E}\left(X_j - hW_j \,\middle|\, W_1 = w_1, \ldots, W_\ell = w_\ell\right) \xrightarrow[h\to\infty]{\text{in distribution}} s\left(w_j - \frac{1}{\ell}\right).$$

c. *If the vote share vector* W *follows uniform distribution, then the seat bias of the* k*-strongest party has limit*

$$\lim_{h\to\infty} \mathrm{E}\left(X_k - hW_k \,\middle|\, W_1 \ge \cdots \ge W_\ell \ge 0\right) = \frac{s}{\ell}\left(H_k^\ell - 1\right).$$

Proof. a. The trisection follows from the identity $y_j - w_j h = s\big(w_j - 1/\ell\big) - (1/2 - s/\ell) + u_j(h)$.

b. Again X_j and W_j are taken to be random variables with ranges $\mathbb{N}^\ell(h)$ and Ω_ℓ. The first term of the trisection determines the systematic part of the seat bias. As for the second term, the difference $Z_j(h) := X_j - Y_j \in \{0, 1\}$ is an indicator function signaling whether party j benefits from the residual apportionment. The negative discrepancy tells the number of seats to be handled by the residual fit, $h - Y_+ = \ell/2 - s - U_+(h)$. Party j gets one of them if and only if its fractional part $W_j(h+s) - Y_j = 1/2 - U_j(h)$ is among the $h - Y_+$ greatest fractional parts $1/2 - U_1(h), \ldots, 1/2 - U_\ell(h)$. The Invariance Principle 6.10 allows us to replace the rounding residuals $U_j(h)$ by their limits U_j. Thus the conditional expectations converge to random variables $Z_j(U, w)$, for all $j \le \ell$. They inherit exchangeability from $U_1, \ldots, U_\ell$. The sum of their expectations is $\ell\,\mathrm{E}\big(Z_j(U,w)\big) = \mathrm{E}\big(Z_+(U,w)\big) = \mathrm{E}(\ell/2 - s - U_+) = \ell/2 - s$. This entails $\mathrm{E}\big(Z_j(U,w) - (1/2 - s/\ell)\big) = 0$ whence the limit expectation of the second term vanishes. The third term is seen to have conditional expectation zero as in the proof of Theorem 7.3.

c. The seat biases are obtained as in the proof of Theorem 7.7. □

In summary the family of shift-quota methods contains a unique member that is unbiased, $s = 0$, the Hare-quota method with residual fit by greatest remainders. The limiting extreme $s = 1$ attests to the bias of the Droop-quota method with residual fit by greatest remainders. It favors stronger parties at the expense of weaker parties. Since its coefficient $1/\ell$ is smaller than the coefficient $1/2$ of the seat biases of the divisor method with downward rounding, the Droop seat biases are less pronounced.

So far we have investigated how a fixed apportionment method behaves under repeated applications. Repetitions motivate the study of a method's average behavior, the average being taken over all possible vote shares. In a nutshell: one method, many vote share vectors. The next chapter reverts the emphasis: many methods, one vote share vector. The comparison invokes the majorization order, a relation with fruitful applications in many fields of the natural and behavioral sciences.

CHAPTER 8

Preferring Stronger Parties to Weaker Parties: Majorization

Majorization provides a partial order of apportionment methods. When passing from one method to another, every group of stronger parties gets more seats and the complementary group of weaker parties fewer seats, or they keep what they have. Specifically, one divisor method majorizes another if and only if their signpost ratios are strictly increasing. The family of stationary divisor methods is shown to be monotonically parameterized, as is the family of power-mean divisor methods. Thus the five traditional divisor methods are ordered by majorization, from upward rounding, via harmonic, geometric and standard rounding, to downward rounding.

8.1. BIPARTITIONS BY VOTE STRENGTHS

Majorization is a well-established order to compare two seat vectors $x = (x_1, \ldots, x_\ell)$ and $y = (y_1, \ldots, y_\ell)$ when both of them apportion the house size h among ℓ parties. Party i's isolated interest is whether the passage from x to y does any good, $x_i < y_i$. Since the component sums of x and y are equal to h, some other party j must be off worse, $x_j > y_j$. Hence some party is better off, another party worse. Moreover, consideration of just two parties disregards the other $\ell - 2$ parties. Pairwise party comparisons are insufficient to order two seat vectors x and y.

Majorization takes a broader view. It divides the party system into complementary subsets, $I \subseteq \{1, \ldots, \ell\}$ and $I' := \{1, \ldots, \ell\} \setminus I$. A bipartition is exhaustive of the entire system, every party features either in I or in I'. Interest is in bipartitions that confront stronger parties with weaker parties. As always in this book, the strength of party j is reflected by its vote count v_j, or equivalently, by its vote share $w_j = v_j / v_+$. A *group of stronger parties* is a subset of parties, I, such that all of its members are at least as strong as the parties in the complementary set I',

$$v_i \geq v_j \quad \text{for all } i \in I \text{ and for all } j \in I'.$$

This leaves just $\ell - 1$ bipartitions to look at: the strongest party versus the $\ell - 1$ weaker parties, the two strongest parties versus the $\ell - 2$ weaker parties, and so on until the $\ell - 1$ strongest parties versus the weakest party. Bookkeeping turns easy when we rank-order the parties from strongest to weakest, $v_1 \geq \cdots \geq v_\ell$, as in Section 7.4.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_8, © Springer International Publishing Switzerland 2014

The group of the k strongest parties then constitutes the top section $1, \ldots, k$, and its complement is the tail section of the $\ell - k$ weakest parties $k+1, \ldots, \ell$.

Let A be a general apportionment method (Section 4.3). Since the method is anonymous the re-arrangement by rank-order does no harm. Moreover the method is concordant, whence all seat vectors $x \in A(h; v)$ have components ordered by decreasing magnitude provided the parties' rank-order is strict, $v_1 > \cdots > v_\ell \;\Rightarrow\; x_1 \geq \cdots \geq x_\ell$. For notational convenience we introduce the set of the seat vectors that have their components arranged in decreasing order,

$$N^\ell_\geq(h) := \left\{ x \in \mathbb{N}^\ell(h) \;\middle|\; x_1 \geq \cdots \geq x_\ell \right\}.$$

The group of the k strongest parties thus accumulates $x_1 + \cdots + x_k$ seats.

8.2. MAJORIZATION OF TWO SEAT VECTORS

Definition. *A vector $x \in \mathbb{N}^\ell_\geq(h)$ is said to be* majorized *by a vector $y \in \mathbb{N}^\ell_\geq(h)$, denoted by $x \preceq y$, when the sum of the k largest components of x is less than or equal to the sum of the k largest components of y,*

$$x_1 + \cdots + x_k \leq y_1 + \cdots + y_k, \qquad \textit{for all } k < \ell.$$

Hence when passing from x to y, every group of stronger parties gets more seats or stays put. Consequently, every group of weaker parties is allocated fewer seats or stays put. In other words the passage from x to y is beneficial for stronger parties and disadvantageous for weaker parties. The definition forgoes the sum of all components, $k = \ell$, since they are equal anyway, $x_+ = y_+ = h$. Vector majorization equips the set $\mathbb{N}^\ell_\geq(h)$ with a partial order. That is, the relation is reflexive ($x \preceq x$), transitive ($x \preceq y$ and $y \preceq z \Rightarrow x \preceq z$), and antisymmetric ($x \preceq y$ and $y \preceq x \Rightarrow x = y$).

A transitivity example for house size $h = 100$ is provided by the seat vectors

$$x = (40, 30, 20, 10) \preceq y = (41, 29, 20, 10) \preceq z = (41, 29, 21, 9).$$

The example illustrates the superiority of the majorization order over pairwise party comparisons. The passage from x to y involves the two strongest parties only, $(40, 30) \preceq (41, 29)$. The second-strongest party must give up a seat to the strongest party. The passage from y to z makes the fourth-strongest party give up a seat to the third-strongest party, $(20, 10) \preceq (21, 9)$. So far, so good. However, the passage from x to z affects all four parties and invites $\binom{4}{2} = 6$ pairwise comparisons. Sometimes a weaker party has to give up a seat to a stronger party, but not always. Take a look at the second- and third-strongest parties, with $(30, 20)$ seats in x, and $(29, 21)$ seats in z. The stronger party has to give up a seat that is transferred to the weaker party. The *pairwise give-up relation from weaker to stronger parties* fails to be transitive. However, the relation provides a sufficient condition to verify majorization.

8.3. A SUFFICIENT CONDITION VIA PAIRWISE COMPARISONS

Lemma. *Let the vectors $x, y \in \mathbb{N}^\ell_\geq(h)$ be given. If in all pairwise comparisons the stronger party grows or the weaker party shrinks then x is majorized by y,*

$$\Big(x_i \le y_i \text{ or } x_k \ge y_k \quad \text{for all } \ i < k\Big) \qquad \Longrightarrow \qquad x \preceq y.$$

Proof. The proof is indirect. Assume x not to be majorized by y. Then some party i is first to violate the defining inequalities, $\sum_{j\le k} x_j \le \sum_{j\le k} y_j$ for all $k < i$, but $\sum_{j\le i} x_j > \sum_{j\le i} y_j$. The ith terms must satisfy $x_i > y_i$. Due to equal component sums some later party $k > i$ must have $x_k < y_k$. Thus the two parties $i < k$ fulfill $x_i > y_i$ and $x_k < y_k$. □

Next we extend the majorization relation to the set of all apportionment methods. Majorization demands that all seat vectors of the first method are majorized by all seat vectors of the second method whenever the rank-order of the parties is strict.

8.4. MAJORIZATION OF TWO APPORTIONMENT METHODS

Definition. *An apportionment method A is said to be* majorized *by an apportionment method B, denoted by $A \prec B$, when the two methods are distinct and when, for all house sizes h, for any number of parties ℓ, and for every weight vector $(v_1, \ldots, v_\ell) \in (0;\infty)^\ell$ the methods satisfy*

$$v_1 > \cdots > v_\ell \qquad \Longrightarrow \qquad A(h;v) \preceq B(h;v),$$

where the set-notation $A(h;v) \preceq B(h;v)$ means $x \preceq y$ for all seat vectors $x \in A(h;v)$ and $y \in B(h;v)$.

The following theorem provides a check whether of two divisor methods $A \neq B$ one majorizes the other. Recall that a divisor method is characterized by a signpost sequence $s(0), s(1), s(2), \ldots$ (Section 3.10). Since every signpost sequence starts with $s(0) = 0$, the only significant terms are $s(n)$ with $n \ge 1$. The theorem uses the conventions $0/0 = 0$ and $\epsilon/0 = \infty$ for $\epsilon > 0$. They may become relevant when $n = 1$.

8.5. MAJORIZATION OF DIVISOR METHODS

Theorem. *Suppose that A is a divisor method with signpost sequence $s(n)$, $n \in \mathbb{N}$, and B is a divisor method with signpost sequence $t(n)$, $n \in \mathbb{N}$. Then method A is majorized by method B if and only if the signpost ratios $s(n)/t(n)$, $n \ge 1$, are strictly increasing,*

$$A \prec B \qquad \Longleftrightarrow \qquad \frac{s(n)}{t(n)} < \frac{s(n+1)}{t(n+1)} \quad \text{for all } n \ge 1.$$

Proof. The proof of the direct implication is by contraposition, assuming that some $N \ge 1$ satisfies

$$\frac{s(N)}{t(N)} \ge \frac{s(N+1)}{t(N+1)}, \qquad \text{that is,} \qquad \frac{s(N+1)}{s(N)} \le \frac{t(N+1)}{t(N)}.$$

This implies $s(N) > 0$. Setting $a := s(N+1)/s(N)$ we get $1 < a \le t(N+1)/t(N)$. For $\ell = 2$ parties with weights $v_1 := a/(a+1) > v_2 := 1/(a+1)$ we claim that $x := (N+1, N-1)$ is a solution vector in $A(2N; v_1, v_2)$. Indeed, the applicable Max-Min Inequality

$$\max\left\{\frac{s(N+1)}{v_1}, \frac{s(N-1)}{v_2}\right\} \le \min\left\{\frac{s(N+2)}{v_1}, \frac{s(N)}{v_2}\right\}$$

is verified simply by comparing each term on the left with each term on the right. Similarly the Max-Min Inequality

$$\max\left\{\frac{t(N)}{v_1}, \frac{t(N)}{v_2}\right\} \le \min\left\{\frac{t(N+1)}{v_1}, \frac{t(N+1)}{v_2}\right\}$$

shows that the vector $y := (N, N)$ lies in $B(2N; v_1, v_2)$. Since x is not majorized by y the proof of the direct implication is complete.

For the proof of the converse implication consider any house size h, some ordered weights $v_1 > \cdots > v_\ell$, and arbitrary seat vectors $x \in A(h; v)$ and $y \in B(h; v)$. Strict monotonicity of the signpost ratios enables us to establish the desired majorization relation, $A(h; v) \preceq B(h; v)$, by referring to Lemma 8.3. (The argument implies that divisor methods are immune against the lack of transitivity of the pairwise give-up relation mentioned in Section 8.2.) Verification of the premise of Lemma 8.3 is by contraposition, assuming that for two parties $i > k$ the stronger party i and the weaker party k satisfy $x_i > y_i$ and $x_k < y_k$. The integer inequalities imply $x_i - 1 \ge y_i$ and $x_k + 1 \le y_k$. Due to concordance $v_i > v_k$ entails $y_i \ge y_k$. We get $x_k + 1 \le x_i - 1 < x_i$. Hence strict monotonicity of the signpost ratios yields $s(x_k+1)/t(x_k+1) < s(x_i)/t(x_i)$. Now we select multipliers μ and ν satisfying $x_j \in [\![\mu v_j]\!]$ and $y_j \in [\![\nu v_j]\!]$ for all $j \le \ell$. The signpost estimates $\nu v_i \le t(y_i + 1)$ and $t(y_k) \le \nu v_k$, and $s(x_i) \le \mu v_i$ and $\mu v_k \le s(x_k + 1)$ combine into

$$\frac{v_i}{v_k} = \frac{\nu v_i}{\nu v_k} \le \frac{t(y_i+1)}{t(y_k)} \le \frac{t(x_i)}{t(x_k+1)} < \frac{s(x_i)}{s(x_k+1)} \le \frac{\mu v_i}{\mu v_k} = \frac{v_i}{v_k},$$

a contradiction. □

Signpost ratio sequences always converge to unity, $\lim_{n\to\infty} s(n)/t(n) = 1$, since the localization property $s(n) \in [n-1; n]$ entails $(n-1)/n \le s(n)/t(n) \le n/(n-1)$. Monotone convergence as in the Theorem implies $s(n) < t(n)$ for all $n \ge 1$. Thus, if a method A is majorized by another method B, all signposts of A are shifted to larger values before reaching the signposts of B.

A signpost shift to larger values increases the likelihood that interim quotients are rounded downwards, and decreases the likelihood of rounding them upwards. Nevertheless, the effects of a signpost shift retain a dose of randomness as far as just a single party is concerned, whether strong or weak. However, for groups of stronger parties any randomness evaporates. Majorization delivers deterministic facts, not stochastic predictions. Every group of stronger parties does at least as well when an apportionment method A is replaced by a method B that majorizes A. No group of stronger parties will ever do worse, no group of weaker parties will ever do better.

Majorization is easily recognized among the stationary divisor methods DivSta_r by means of their split parameter $r \in [0; 1]$. An analogous result holds for the power-mean divisor methods DivPwr_p in terms of their power parameter $p \in [-\infty; \infty]$. An increase of the parameter reflects an increase in the majorization order.

8.6. MAJORIZATION-INCREASING PARAMETERIZATIONS

Theorem.

a. (Stationary methods) *A stationary divisor method* $DivSta_r$ *is majorized by another stationary divisor method* $DivSta_R$ *if and only if* $r < R$.

b. (Power-mean methods) *A power-mean divisor method* $DivPwr_p$ *is majorized by another power-mean divisor method* $DivPwr_P$ *if and only if* $p < P$.

Proof. a. With stationary signposts $s_r(n) = n-1+r$, signpost ratio monotonicity $(n-1+r)/(n-1+R) < (n+r)/(n+R)$ instantly reduces to $r < R$.

b. The power-mean signposts $\widetilde{s}_p(n)$ are the mean of order p of the interval endpoints $n-1$ and n (Section 3.12). Verification of signpost ratio monotonicity reduces to proving that the function

$$g(p) := \left(\frac{(n-1)^p + n^p}{n^p + (n+1)^p}\right)^{1/p}$$

is strictly increasing, from $g(-\infty) = (n-1)/n$ to $g(\infty) = n/(n+1)$. The proof follows from Proposition 5.B.3 in *Marshall / Olkin* (1979 [130]). □

Applying the result to the power-mean divisor methods with powers $-\infty$, -1, 0, 1, and ∞, the five traditional divisor methods are seen to be ordered by majorization,

$$\text{DivUpw} \prec \text{DivHar} \prec \text{DivGeo} \prec \text{DivStd} \prec \text{DivDwn}.$$

From left to right the methods are more supportive of groups of stronger parties, and detrimental to groups of weaker parties.

8.7. MAJORIZATION PATHS

The family of stationary divisor methods generates a sequence of seat vectors, and so does the family of power-mean divisor methods. In these sequences every seat vector is majorized by its successor. Although the parameter increases continuously, the evolving sequence of seat vectors is finite because of the inherent discretization step. Both paths start with DivUpw (= DivSta_0 = $\text{DivPwr}_{-\infty}$), pass through DivStd (= $\text{DivSta}_{.5}$ = DivPwr_1), and finish with DivDwn (= DivSta_1 = DivPwr_∞).

Table 8.1 presents an example. The top portion displays the path in the stationary family. The bottom part shows the path in the power-mean family. Both paths happen to feature eight seat vectors each. They have six seat vectors in common, while the third and the seventh are peculiar to just one path. Every seat vector is majorized by its successor. From left to right, groups of stronger parties accumulate more seats or maintain what they have. Groups of weaker parties do worse or stay put. The fate of a single party is less deterministic. The strongest party never loses a seat since it is a singleton group. The weakest party never wins a seat. Parties in-between may oscillate, as does the fifth strongest party.

Majorization path of stationary divisor methods, DivSta_r

Rank	Votes	$r = 0$		.3		.44		.47		.5		.8		.97		1
1	42 919	41	+1	42		42	+1	43		43	+1	44		44	+1	45
2	13 048	13		13		13		13		13		13		13		13
3	10 879	11		11		11		11		11		11		11		11
4	10 581	10		10	+1	11		11		11		11		11		11
5	9 547	10	−1	9		9		9	+1	10	−1	9	+1	10		10
6	5 708	6		6		6		6		6		6	−1	5		5
7	2 502	3		3		3		3	−1	2		2		2		2
8	1 898	2		2		2		2		2		2		2	−1	1
9	1 461	2		2		2	−1	1		1		1		1		1
10	1 457	2		2	−1	1		1		1		1		1		1
Sum	100 000	100		100		100		100		100		100		100		100

Majorization path of power-mean divisor methods, DivPwr_p

Rank	Votes	$p = -\infty$		0		.4		.6		1		18		50		∞
1	42 919	41	+1	42	+1	43		43		43	+1	44		44	+1	45
2	13 048	13		13		13		13		13		13		13		13
3	10 879	11		11		11		11		11		11		11		11
4	10 581	10		10		10	+1	11		11		11		11		11
5	9 547	10	−1	9		9		9	+1	10	−1	9	+1	10		10
6	5 708	6		6		6		6		6		6		6	−1	5
7	2 502	3		3		3		3	−1	2		2		2		2
8	1 898	2		2		2		2		2		2	−1	1		1
9	1 461	2		2		2	−1	1		1		1		1		1
10	1 457	2		2	−1	1		1		1		1		1		1
Sum	100 000	100		100		100		100		100		100		100		100

TABLE 8.1 *Majorization paths.* Top: Stationary divisor methods. Bottom: Power-mean divisor methods. Both paths start with DivUpw, pass through DivStd, and finish with DivDwn. In contiguous seat columns, a seat is given up from a weaker party (−1) to a stronger party (+1).

A seat transfer to a party i from a party k can occur only when the two are tied,

$$\frac{s(x_i + 1)}{v_i} = \frac{s(x_k)}{v_k}.$$

The tie equation is instrumental to verify that the paths in Table 8.1 are complete. For the stationary family, the tie equation and its solution are

$$\frac{x_i + r}{v_i} = \frac{x_k - 1 + r}{v_k}, \qquad \text{that is,} \qquad r_1(i,k) := \frac{x_i v_k - (x_k - 1)v_i}{v_i - v_k}.$$

In the presence of majorization we know that the receiving party i must be stronger than the donor k, that is, $i < k$ and $v_i > v_k$. As the split point r increases from $r_0 = 0$ onwards, the smallest of the solutions $r_1(i,k)$ is encountered first,

$$r_1 := \min_{i<k} r_1(i,k) = r_1(1,5) = \frac{1289}{8343} \approx .155.$$

At the change-point r_1, a seat is transferred from the fifth strongest party to the strongest party, thus giving rise to the second seat vector. Beyond r_1 calculations start afresh. The second seat vector persists until the next change-point r_2 is reached,

$$r_2 := \min_{i<k} r_2(i,k) = r_2(4,10) = \frac{3989}{9124} \approx .437.$$

The interval of constancy is $[.155; .437]$; Table 8.1 quotes the convenient value .3. (A traditional parameter value is preferred whenever possible.) The change-points $r_3, \ldots, r_7$ are obtained in the same manner.

For the family of power-mean divisor methods the tie equation takes the form

$$\left(\frac{x_i^p + (x_i+1)^p}{(x_k-1)^p + x_k^p}\right)^{1/p} = \frac{v_i}{v_k}.$$

No explicit solution is available, but a numerical solution p is easily obtained by machine calculation. The change-points $p_1, \ldots, p_7$ yield the path shown in Table 8.1.

Finally we show that the parameterization of the shift-quota methods (Section 5.4) is majorization-increasing, too, and that they are also framed by the divisor method with upward rounding and by the divisor method with downward rounding.

8.8. MAJORIZATION OF SHIFT-QUOTA METHODS

Theorem. *Two shift-quota methods with shifts $s < S$ satisfy*

$$DivUpw \prec shQgrR_s \prec shQgrR_S \prec DivDwn.$$

Proof. For house size h and vote vector $v \in (0;\infty)^\ell$ fixed, Corollary 5.6 provides splits r^* and R^* that reproduce the shift-quota solution sets, $\mathrm{shQgrR}_s(h;v) = \mathrm{DivSta}_{r^*}(h;v)$ and $\mathrm{shQgrR}_S(h;v) = \mathrm{DivSta}_{R^*}(h;v)$. Let x be a seat vector in the first set, y, in the second. Unless the two vectors are equal, there exists a party j with seat numbers $x_j > y_j$. From $x_j - 1 + r^* \leq v_j/Q(s)$ and $v_j/Q(S) \leq y_j + R^*$ we get

$$\frac{v_j}{Q(S)} + 1 - R^* \leq y_j + 1 \leq x_j \leq \frac{v_j}{Q(s)} + 1 - r^*.$$

Insertion of the shift-quotas $Q(s) = v_+/(h+s)$ yields $0 < (S-s)v_j/v_+ = v_j/Q(S) - v_j/Q(s) \leq R^* - r^*$. Hence the splits are ordered, $r^* < R^*$. Theorem 8.6.a yields $\mathrm{DivUpw}(h;v) \preceq \mathrm{DivSta}_{r^*}(h;v) \preceq \mathrm{DivSta}_{R^*}(h;v) \preceq \mathrm{DivDwn}(h;v)$. □

Seat biases in favor of some parties at the expense of others, and the preferential treatment of groups of stronger parties as compared to groups of weaker parties are important characteristics of an apportionment method, but do not exhaustively tell how sensibly it behaves in practice. The next chapter judges the resulting seat vectors from a wholistic viewpoint. The whole vector and its parts must fit together in a coherent way that is devoid of irritating paradoxes.

CHAPTER 9

Securing System Consistency: Coherence and Paradoxes

Apportionment methods are assessed from the collective viewpoint whether all variables act together in a consistent manner. A decisive requirement is coherence, demanding that the solution for an apportionment problem as a whole agrees with the solutions of all embedded subproblems. Coherence is achieved only by divisor methods. Furthermore, divisor methods respond sensibly to variations of particular variables such as house size, vote weights, or number of parties. In contrast, quota methods may exhibit a counterintuitive behavior of a seemingly paradoxical nature.

9.1. THE WHOLE AND ITS PARTS

The more parties participate in the seat apportionment, the more daunting becomes the apportionment problem. An established strategy to attack a large problem with many variables is to perceive it as a union of partial problems with fewer variables. A viable solution for the whole problem should comprise viable solutions for its partial problems. *Balinski / Young* (1982 [141]) put it this way: An inherent principle of any fair division is that *every part of a fair division should be fair.* Conversely the whole solution should be retrievable by concatenating solutions of partial problems. The whole and its parts must fit together in a coherent way.

A proportional representation system is taken to be coherent when all of its partial systems remain proportional. If the seat vector $(x_j)_{j=1,\ldots,\ell}$ is a solution for the apportionment of x_+ seats in a large system of ℓ parties, then a party subsystem $I \subset \{1,\ldots,\ell\}$ ought to admit the sub-vector $(x_i)_{i\in I}$ as a solution for the apportionment of $\sum_{i\in I} x_i$ seats among the parties in I. The definition includes a kind of converse to the effect that concatenated partial solutions retrieve a grand solution. We denote partial sums 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 vote vectors $v \in (0;\infty)^\ell$, every seat vector $x \in A(x_+;v)$ fulfills the following two properties* a *and* b *for all party subsets $I \subset \{1,\ldots,\ell\}$:*

a. (Coherence of partial solutions) $(x_i)_{i\in I} \in A\big(x_I;(v_i)_{i\in I}\big)$.

b. (Concatenation coherence) *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* $\big((y_i)_{i\in I},(z_k)_{k\in I'}\big) \in A\big(x_+;\big((v_i)_{i\in I},(v_k)_{k\in I'}\big)\big)$.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*, DOI 10.1007/978-3-319-03856-8_9, © Springer International Publishing Switzerland 2014

Coherence of partial solutions (a) means that all partial vectors that can be extracted from a grand seat vector are solutions of the associated partial problems. Concatenation coherence (b) says that two disjoint partial solutions with fitting component sums, x_I and $x_{I'}$, concatenate to form a grand seat vector. If the partial solutions are singletons then they arise from x, $A\big(x_I; (v_i)_{i\in I}\big) = \{(x_i)_{i\in I}\}$ and $A\big(x_{I'};$ $(v_k)_{k\in I'}\big) = \{(x_k)_{k\in I'}\}$, and retrieve the given seat vector x uniquely. If at least one of the partial problems admits multiple solutions then the grand problem does so, too.

Apportionment methods are stipulated to be anonymous, balanced, concordant, decent, and exact (Section 4.3). The five organizing principles imply that every coherent apportionment method must be a divisor method.

9.2. COHERENCE THEOREM

Theorem. *An apportionment method A is coherent if and only if A is a divisor method.*

Proof. The theorem is due to *Balinski / Young* (1982 [141–148]) where the term *uniformity* is used in place of *coherence*. The proof of the direct implication requires the construction of a signpost sequence that discloses A as a divisor method. The authors accomplish the task via a detailed analysis of the larger class of rank-index methods. Since the argument is fairly complicated and lengthy we abstain from reproducing it here.

The proof of the converse implication is immediate. Let A be a divisor method. Coherence of partial solutions holds since any divisor D that works for a grand seat vector x also works for all partial solutions $(x_i)_{i\in I}$. Concatenation coherence follows since if the partial solution set $A\big(x_I; (v_i)_{i\in I}\big)$ comprises two or more solution vectors then the divisor is unique (Section 4.7). Hence the divisor for $(x_i)_{i\in I}$ also applies for all other vectors in the partial solution set. Thus every divisor method is seen to be coherent. □

The theorem confirms the superiority of divisor methods. Other methods fail to be coherent. In particular, the Hare-quota method with residual fit by greatest remainders is incoherent. Its results often coincide with the results of the divisor method with standard rounding. Hence incoherence can surface only when the two methods yield distinct solutions. This hint makes it easy to locate instances of incoherence.

To this end Table 9.1 reconsiders the 2009 Bundestag election data of Table 2.1. The solution of the divisor method with standard rounding embraces the two-party subsystem of SPD and CSU coherently. The parties split their aggregate 188 seats 146 : 42 no matter whether the six parties are evaluated altogether, or whether SPD and CSU are treated as a separate two-party system. In contrast, the Hare-quota method with residual fit by greatest remainders provides two solutions, one in the whole ensemble (147 : 41) and one that considers the two parties separately (146 : 42). The two solutions differ, thus illustrating the incoherence the Hare-quota method with residual fit by greatest remainders.

(17BT2009)	Votes	Quotient	DivStd	Quotient	HaQgrR
CDU	11 828 277	173.4	173	173.517	173
SPD	9 990 488	146.497	146•	146.558	147•
FDP	6 316 080	92.6	93	92.655	93
LINKE	5 155 933	75.6	76	75.636	76
GRÜNE	4 643 272	68.1	68	68.115	68
CSU	2 830 238	41.502	42•	41.519	41•
Sum (Divisor, Split)	40 764 288	(68 196)	598	(.54)	598

TABLE 9.1 *Coherence of DivStd, and incoherence of HaQgrR.* Both methods give the same answer to the partial problem of allocating 188 seats to SPD (146) and CSU (42). This partial solution is part of the grand DivStd apportionment, but distinct from the grand HaQgrR apportionment (147 : 41).

Quite generally, two-party problems always admit seat vectors $x = (x_1, x_2)$ that most people would agree to accept as *the natural two-party apportionment* of H seats,

$$x_1 \in \left\langle\!\!\left\langle \frac{v_1}{v_1 + v_2}\, H \right\rangle\!\!\right\rangle, \qquad x_2 \in \left\langle\!\!\left\langle \frac{v_2}{v_1 + v_2}\, H \right\rangle\!\!\right\rangle.$$

This solution is delivered by the divisor method with standard rounding as well as by the Hare-quota method with residual fit by greatest remainders. In fact, for two-party systems the two methods always agree,

$$\text{DivStd}(H; v_1, v_2) = \text{HaQgrR}(H; v_1, v_2),$$

for all house sizes $H \in \mathbb{N}$ and for all vote pairs $(v_1, v_2) \in (0; \infty)^2$.

An obvious question is whether the natural two-party apportionments allow a coherent extension to arbitrarily large party systems. It turns out that the unique answer is the divisor method with standard rounding.

9.3. COHERENCE EXTENSION FROM TWO PARTIES TO MANY

Theorem. *The unique coherent extension of the natural two-party apportionments to party systems of arbitrary size $\ell \geq 2$ is the divisor method with standard rounding.*

Proof. Evidently DivStd is a coherent extension of the natural two-party apportionments. We need to prove uniqueness: If A is a coherent extension then $A = \text{DivStd}$. That is, for all house sizes $h \in \mathbb{N}$, all system sizes $\ell \geq 2$, and all vote vectors $v \in (0, \infty)^\ell$ we claim

$$A(h; v) = \text{DivStd}(h; v).$$

As for the direct inclusion, $A(h; v) \subseteq \text{DivStd}(h; v)$, we assume indirectly that some seat vector $z \in A(h; v)$ is not contained in $\text{DivStd}(h; v)$. For arbitrary seat vectors $y \in \text{DivStd}(h; v)$ let $c(y)$ be the count of components $j \leq \ell$ with $y_j \neq z_j$. We select a vector $x \in \text{DivStd}(h; v)$ such that $c(x)$ is minimum. By assumption we have $c(x) \geq 2$. Let the parties i and k be such that $x_i > z_i$ and $x_k < z_k$. Of the two sums $x_i + x_k$ and $z_i + z_k$ one is less than or equal to the other. Without loss of generality we treat the case $H := x_i + x_k \leq z_i + z_k =: \widetilde{H}$. Since all rounding rules are set-monotonic (Section 3.8) we get

$$\left\langle\!\!\left\langle \frac{v_i}{v_i + v_k}\, H \right\rangle\!\!\right\rangle \leq \left\langle\!\!\left\langle \frac{v_i}{v_i + v_k}\, \widetilde{H} \right\rangle\!\!\right\rangle.$$

Because of coherence of DivStd the left set contains x_i; due to coherence of A the right set includes z_i. In view of $x_i > z_i$ this is possible only when the rounding sets are equal to $\{x_i, z_i\}$. It follows that $H = \widetilde{H}$, whence $\langle\!\langle v_k H/(v_i + v_k) \rangle\!\rangle = \{x_k, z_k\}$. Thus (x_i, x_k) and (z_i, z_k) are tied solutions for apportioning H seats proportionally to the two-party vote vector (v_i, v_k). Concatenation coherence implies that the vector

$$y := (x_1, \ldots, x_{i-1}, z_i, x_{i+1}, \ldots, x_{k-1}, z_k, x_{k+1}, \ldots, x_\ell)$$

lies in DivStd$(h; v)$. Evidently y differs from z in two fewer components than x. This contradicts the minimality of x, thereby verifying the direct inclusion $A(h; v) \subseteq$ DivStd$(h; v)$. A similar reasoning establishes the converse inclusion. Thus the proof of uniqueness is complete. □

In summary coherence captures an intrinsic consistency aspect of apportionment methods. The notion appears innocuous, but entails remarkably strong consequences. The vast class of all apportionment methods is constricted to the smaller class of divisor methods. Conformity with the natural two-party apportionments leaves just a single procedure, the divisor method with standard rounding.

Next we turn to other consistency aspects of proportional representation systems. How does an apportionment method respond to an increase of the house size? To the growth of one party relative to another party? To an enlargement of the party system? Whichever question is asked, the answer points in the same direction. Divisor methods always behave reasonably. Quota methods mostly do too, but occasionally deliver solutions that seem paradoxical. We consider each question in greater detail.

9.4. HOUSE SIZE MONOTONICITY

An apportionment method is called *house size monotone* when an increase of the house size never leads to a decrease of some party's seat number. If x is a seat vector for house size h and an extra seat becomes available then some party i gets $x_i + 1$ seats and all other parties $j \neq i$ retain the x_j seats they already have. Formalities to properly track ties are more tedious. For two seat vectors x and y we write $x \leq y$ when all components are non-decreasing, $x_j \leq y_j$ for all $j \leq \ell$. Formally, house size monotonicity of an apportionment method A demands that for all seat vectors $x \in A(h; v)$ there exists a seat vector $y \in A(h + 1; v)$ satisfying $x \leq y$.

It would be tempting to want the inequality $x \leq y$ to hold true for *all* vectors $y \in A(h+1; v)$. But this is asking too much, and it is not hard to see why. The example in Table 4.2 displays fifteen tied seat vectors x for house size 17. Every vector y that componentwise dominates all fifteen vectors x satisfies $y \geq (6, 5, 4, 3, 2, 1)$ and has a component sum of $y_+ \geq 21$. This would inhibit any comparisons with house sizes 18, 19, and 20. The definition demands that there exists a vector $y \in A(h + 1; v)$ that improves upon a given seat vector $x \in A(h; v)$.

Every divisor method is house size monotone. Indeed, the monotone progression from house size h with seat vector x, to house size $h + 1$ with some seat vector $y \geq x$ is part of the argument that divisor methods are well-defined (Section 4.4).

A non-monotone apportionment method B is prone to a disconcerting behavior. When two seat vectors $x \in B(h;v)$ and $y \in B(h+1;v)$ violate the componentwise inequality $x \le y$, some party i gets fewer seats than before, $x_i > y_i$, even though there are more seats to go around, $x_+ = h < h+1 = y_+$. *Janson / Linusson* (2012) construct a striking example. In a five-party system with vote vector $v = (280, 275, 270, 90, 85)$, the Hare-quota method with residual fit by greatest remainders produces seat vector $x = (1,1,1,1,1)$ for house size five, but $y = (2,2,2,0,0)$ for house size six. The extra seat works wonders. Three parties double their representation, and two parties disappear from the scene. The example shows that the Hare-quota method with residual fit by greatest remainders fails to be house size monotone. Violation of monotonicity was observed during the 1880 United States census when it was dubbed the *Alabama paradox*, see Section 9.7. Lack of house size monotonicity raises suspicions whether the apportionment method harbors other awkward peculiarities.

9.5. VOTE RATIO MONOTONICITY

Vital input quantities of an apportionment problem are the vote counts. From one election to the next they generally vary, of course. We consider a profile where the house size and the number of parties stay the same. Let the two consecutive elections be recorded in the vote vectors $u = (u_1, \dots, u_\ell) \in (0;\infty)^\ell$ and $v = (v_1, \dots, v_\ell) \in (0;\infty)^\ell$, with ensuing seat vectors $x \in A(h;u)$ and $y \in A(h;v)$.

How would party i assess its performance relative to a competing party k? A natural approach is to measure the votes of party i in multiples of the votes of party k, that is, to compare the ratio u_i/u_k in the first election with the ratio v_i/v_k in the second. If the vote ratio is increasing, party i improves its standing relative to party k. A sensible method should react to the improvement by a potential increase of the seats of party i or a potential decrease of the seats of party k,

$$\frac{u_i}{u_k} < \frac{v_i}{v_k} \qquad \Longrightarrow \qquad x_i \le y_i \quad \text{or} \quad x_k \ge y_k.$$

An apportionment method A is called *vote ratio monotone* when the implication holds for all house sizes h, vote vectors $u, v \in (0;\infty)^\ell$, seat vectors $x \in A(h;u)$ and $y \in A(h;v)$, and parties $i, k \le \ell$. The implication inhibits the case $x_i > y_i$ and $x_k < y_k$, entailing $x_i/x_k > y_i/y_k$. Lack of vote ratio monotonicity means that, while party i does better than party k in terms of vote ratios, it does worse in terms of seat ratios.

All divisor methods are vote ratio monotone. To prove the claim we consider two cases. In case $x_i \le y_i$ the conclusion is evidently true. In case $x_i > y_i$ the scaled vote counts satisfy $\mu u_i \ge \nu v_i$, where μ and ν are appropriate multipliers for x and y. This gives $\mu/\nu \ge v_i/u_i > v_k/u_k$. Since $\mu u_k > \nu v_k$ implies $x_k \ge y_k$, the claim is proved.

Quota methods fail to be vote ratio monotone. This is shown by the Hare-quota example in Table 9.2. The table lists all concordant seat vectors obtained by rounding the ideal shares upwards or downwards. In the first election party A receives at least five seats, in the second at most four seats. Party D fares better. In the first election it gets no seat, in the second one seat. In any case party A loses a seat, and party D wins a seat. On the other hand party A gathers strength relative to party D. It is $717/93 = 7.7$ times stronger in the first election, but $570/73 = 7.8$ times stronger in the second. This behavior contradicts vote ratio monotonicity.

Party	Votes	Quotient	#1	#2	Votes	Quotient	#3	#4	#5
A	717	5.019	6	5	570	3.990	4	4	3
B	96	0.672	1	1	285	1.995	2	1	2
C	94	0.658	0	1	72	0.504	0	1	1
D	93	0.651	0	0	73	0.511	1	1	1
Sum (HaQ)	1000	(1000/7)	7	7	1000	(1000/7)	7	7	7

TABLE 9.2 *Lack of vote ratio monotonicity of Hare-quota methods with a concordant residual fit.* Party A's vote ratio relative to D improves from 717/93 = 7.7 (left) to 570/73 = 7.8 (right). Yet A loses a seat and D gains a seat, in every pairing of the seat vectors #1–#2 (left) and #3–#5 (right).

9.6. SYSTEM SIZE CONFORMITY

Does it matter whether in a system with ℓ parties those that are too weak to obtain representation are carried along until the end, or put aside at the beginning? The EP election in Spain is an instance to exemplify the issue. Table 1.12 includes a line of 28 "Others". The line comprises an aggregation of 28 weak parties each drawing fewer than 90 000 votes, and each failing to win representation. In Chapter 1 the figures are displayed solely for descriptive purposes, so as to record all effective votes and not just those that justified one or more seats. However, descriptive purposes are one thing, computational steps another. Hence we should specify: Does it matter *computationally* whether have-nots are carried along or not?

For divisor methods the answer is *No, it does not matter*. In Table 1.12 the divisor 280 000 stays the same irrespective of whether the 28 *Others* are included in the calculations or not.

For quota methods the answer is elusive, *It depends*. For the Spanish data in Table 1.12, the answer is *Yes, it matters*. Table 9.3 presents the details. Including the 28 have-nots, the Hare-quota is 312 305.9. Without them, the Hare-quota drops down to 304 671.6 and all interim quotients grow larger. The interim quotient of the strongest party leaps forwards substantially by half a seat, from 21.358 to 21.894. The progression of the fifth-strongest party is minute, from 1.447 to 1.483. While its former remainder earns a seat ($.447 > .4$), the latter remainder does not ($.483 < .5$). The seat is transferred from the fifth-strongest party to the strongest party. The transfer comes as a surprise since vote relations among the top seven parties remain the same as before, and have nothing to do with the 28 have-nots at the end of the field.

A vexing seat transfer of this sort troubled the 1993 Council of the Free and Hanseatic City of Hamburg. The council comprised four party groups of sizes 58 : 36 : 20 : 5 and two indeps. The task was to allocate the 15 seats of the Committee of Advisory Deputies. Surprisingly, it mattered whether the two indeps were included in the calculations, or not. The apportionment method used was the Hare-quota method with residual fit by greatest remainder. The seat allocation was 7 : 4 : 3 : 1 : 0 : 0 when the indeps were included in the calculation, but 7 : 5 : 2 : 1 when they were omitted. Thus the procedural standing of indeps proved decisive for the second- and third strongest parties to divide their seven seats as 4 : 3, or 5 : 2.

The problem is not bound to parties receiving no seats. We chose have-nots for a start only because they facilitate the exposition. The phenomenon also arises when new parties that join are accompanied by new seats to proportionally enlarge the prevailing

(EP2009ES)	Votes	Quotient Including	HaQgrR 28 Others	Quotient Without	HaQgrR 28 Others
PP	6 670 377	21.358	21	21.894	22•
PSOE	6 141 784	19.666	20	20.159	20
CpE	808 246	2.588	3	2.653	3
IU-ICV-EU/IA-BA	588 248	1.884	2	1.931	2
UPyD	451 866	1.447	2	1.483	1•
EdP-V	394 938	1.265	1	1.296	1
II	178 121	0.570	1	0.585	1
28 Others	381 716	—	0	—	
Sum (Split)	15 615 296	(.4)	50	(.5)	50
Hare-Quota		312 305.9		304 671.6	

TABLE 9.3 *System size conformity.* The Hare-quota method with residual fit by greatest remainders swaps a seat between the strongest and the fifth-strongest parties contingent on the inclusion or exclusion of the 28 *Others*, each having fewer than 90 000 votes and remainders below .3.

house size. Such instances do occur. For example, the German Democratic Republic acceded to the Federal Republic of Germany in 1990 while the Bundestag's legislative period was in full swing. The apportionment method then in use was the Hare-quota method with residual fit by greatest remainders. The Bundestag decided to enlarge its house size and the sizes of its committees by proportionally adding new seats. Had the apportionments been recalculated from scratch, the old parties would have had to swap some seats for no other reason than the purportedly proportional enlargement.

All examples in this section fit under the heading of coherence. While coherence requires the apportionment method to treat *all* party subsets in a consistent manner, the troublesome examples highlight instances where *some* particular subset falls victim to incidences of incoherence. The remedy is offered by Theorem 9.1, stating that breaches of coherence are avoided when the procedure used is a divisor method.

9.7. QUOTA METHOD PARADOXES

A breach of house size monotonicity of the Hare-quota method with residual fit by greatest remainders first came to light in the United States of America. The Constitution decrees to conduct a census every ten years, and to apportion the seats in the House of Representatives among the States of the Union accordingly. However, it does not specify the apportionment method to be used. In the aftermath of every census fierce debates erupted in the House of Representatives which method to use. The discussions produced a wealth of political arguments and statistical ramifications advertising proposed methods, and criticizing counter-proposals. *Balinski / Young* (1982) tell the tale, in a vivid and enlightening manner.

The size of the House of Representatives has been fixed at 435 members by statute since Arizona and New Mexico became states in 1912. Until then the determination of the house size was part of the agenda of the decennial apportionment legislation. The Bureau of the Census would prepare tables for different house sizes for the purpose of aiding the decision-making process. After the 1880 census *Charles Wesley Seaton*, chief clerk of the Census Office, computed apportionments for all house sizes between 275 and 350 using the Hare-quota method with residual fit by greatest remainders (overseas called Hamilton method). As reported by *Balinski / Young* (1982 [38]), *Seaton* noted a startling effect that he described in a letter to Congress dated 25 October 1881:

> While making these calculations I met with the so-called "Alabama" paradox where Alabama was alloted 8 Representatives out of a total of 299, receiving but 7 when the total became 300.

The label *Alabama paradox* conquered a firm place in the field, as a synonym for a breach of house size monotonicity. Use of the term "paradox" is but an invitation to pause and ask what is going on. On the way from house size 299 to 300 all interim quotients grow larger, of course. Since growth is proportional to size, the larger states Illinois and Texas leap forwards substantially and pass the split point .6715. The progression of small Alabama is much smaller and lets it fall below the split point. Thus Alabama loses the residual seat won earlier, while Illinois and Texas pick up an extra seat each. The data of the 1880 census are shown in Table 9.4.

In the present context vote ratio monotonicity turns into *population monotonicity*,

$$\frac{w_i}{v_i} > \frac{w_k}{v_k} \qquad \Longrightarrow \qquad x_i \leq y_i \quad \text{or} \quad x_k \geq y_k.$$

The inequality in the premise is logically equivalent to $v_i/v_k < w_i/w_k$, but allows a more persuasive interpretation in terms of population growth. The ratio w_i/v_i expresses the current census figure w_i of state i as a multiple of its previous figure v_i. If the ratio is larger than the ratio of another state k, as assumed in the premise, then the population of state i grows at a rate faster than that of state k. The conclusion demands that state i meets or exceeds its former representation and thereby does potentially better, or that state k does worse or stays put. The notion of population monotonicity plays a pivotal role in the exposition of *Balinski / Young* (1982 [108, 117]). Lack of monotonicity is what the authors term the *population paradox*. That is, state i grows faster than state k and yet state i loses one or more seats to state k.

Vote ratio monotonicity and population monotonicity pose problems that are identical as far as numbers are concerned. However, political perceptions and practical consequences of the concepts differ. While the translation of votes into seats remains the prerogative of the political legislator, the representation of geographical districts involves considerable administrative expertise. All modern states entertain statistical offices updating census figures and monitoring population mobility meticulously and laboriously. It would seem counterproductive to belittle these efforts by eventually employing an apportionment method flawed by paradoxes. This is the reason why the German Bundestag disposed of the Hare-quota method with residual fit by greatest remainders. The lack of population monotonicity was considered irritating when apportioning the 299 constituencies to the sixteen German States. In 2008 the Bundestag adopted the divisor method with standard rounding for the allocation of the 299 constituencies as well as for the apportionment of the notional 598 Bundestag seats.

The expansion of the United States from formerly 15 to currently 50 states gives rise to profiles with a growing number of participants. As discussed in Section 9.6 quota methods may trigger strange seat transfers when a new state accedes the Union. *Balinski / Young* (1982 [44]) term the lack of conformity the *new states paradox*. They illustrate the problem with the accession of Oklahoma in 1907. Giving Oklahoma its ideal share of five seats, New York and Maine would have to swap a seat if the Hamilton apportionment were calculated from scratch.

US1880Census	Population	Quotient	HaQgrR	Quotient	HaQgrR
New York	5 082 871	30.783	31	30.886	31
Pennsylvania	4 282 891	25.938	26	26.025	26
Ohio	3 198 062	19.368	19	19.433	19
Illinois	3 077 871	18.640•	18	18.702•	19
Missouri	2 168 380	13.132	13	13.176	13
Indiana	1 978 301	11.981	12	12.021	12
Massachusetts	1 783 085	10.799	11	10.835	11
Kentucky	1 648 690	9.985	10	10.018	10
Michigan	1 636 937	9.914	10	9.947	10
Iowa	1 624 615	9.839	10	9.872	10
Texas	1 591 749	9.640•	9	9.672•	10
Tennessee	1 542 359	9.341	9	9.372	9
Georgia	1 542 180	9.340	9	9.371	9
Virginia	1 512 565	9.160	9	9.191	9
North Carolina	1 399 750	8.477	8	8.505	8
Wisconsin	1 315 497	7.967	8	7.993	8
Alabama	1 262 505	7.646•	8	7.671•	7
Mississippi	1 131 597	6.853	7	6.876	7
New Jersey	1 131 116	6.850	7	6.873	7
Kansas	996 096	6.033	6	6.053	6
South Carolina	995 577	6.029	6	6.050	6
Louisiana	939 946	5.692	6	5.711	6
Maryland	934 943	5.662	6	5.681	6
California	864 694	5.237	5	5.254	5
Arkansas	802 525	4.860	5	4.876	5
Minnesota	780 773	4.728	5	4.744	5
Maine	648 936	3.930	4	3.943	4
Connecticut	622 700	3.771	4	3.784	4
West Virginia	618 457	3.745	4	3.758	4
Nebraska	452 402	2.740	3	2.749	3
New Hampshire	346 991	2.101	2	2.108	2
Vermont	332 286	2.012	2	2.019	2
Rhode Island	276 531	1.675	2	1.680	2
Florida	269 493	1.632	1	1.638	1
Colorado	194 327	1.177	1	1.181	1
Oregon	174 768	1.058	1	1.062	1
Delaware	146 608	0.888	1	0.891	1
Nevada	62 266	0.377	1	0.378	1
Sum (Split)	49 371 340	(.643)	299	(.6715)	300
Hare-Quota		165 121.5		164 571.1	

TABLE 9.4 *Alabama paradox, US census 1880.* An increase of the house size from 299 to 300 increases all interim quotients, for the larger states Illinois and Texas more so than for Alabama (•). With its remainder falling below the split point .6715, Alabama loses its residual seat from before.

Another detail of Table 9.4 is worth mentioning. The bottom line features Nevada, a state so small receiving no seat in the main apportionment nor in the residual fit. Yet the table lists a seat for Nevada. The seat is owed to the constitutional warranty that *each state shall have at least one representative*, mentioned in Section 7.8. The constitutional order is executed as follows. The lacking seats are filled-in from the residual seats before entering the residual apportionment stage. For house sizes 299 and 300 the main apportionment provides all states with at least one representative except Delaware and Nevada. Hence from the residual seats Delaware and Nevada receive one seat each, and only the remaining residual seats are fed into the residual fit by greatest remainders. Chapter 12 has more to say how to handle seat restrictions.

CHAPTER 10

Appraising Electoral Equality: Goodness-of-Fit Criteria

Perfect electoral equality is a conceptual ideal defying reality, a certain degree of inequality must be practically tolerated. Ways and means are discussed how to numerically evaluate the residue of disproportionality that apportionment methods inevitably carry along. Three approaches are explored. A first approach is based on all-embracing goodness-of-fit criteria, that is, functions that map a system's deviations from ideal equality into a real number. Different criteria are seen to justify different methods. A second approach utilizes stability criteria based on pairwise comparisons. The aim is to reduce a pending imbalance by way of a seat transfer from some party that is advantaged to another party that is disadvantaged. A third approach examines whether realized and ideal shares of seats of all parties come to lie as near as may be.

10.1. OPTIMIZATION OF GOODNESS-OF-FIT CRITERIA

It is an established approach of all sciences to explore a complex system by means of real functions. Naturally a single number cannot possibly mirror the full complexity of a multi-dimensional system. Yet valuable information may be retrieved when the criterion function is geared to the essentials of the system.

Apportionment methods aim at achieving proportionality between the input vote counts $(v_1, \ldots, v_\ell)$, and the output seat numbers $(x_1, \ldots, x_\ell)$. Usually the vote total v_+ is much larger than the sum of all seats, $x_+ = h$. Therefore, the vectors are standardized so that their component sum is the same, unity. Goodness-of-fit criteria measure how the output, the seat share vector $x/h = (x_1/h, \ldots, x_\ell/h)$, conforms to the input, the vote share vector $v/v_+ = (v_1/v_+, \ldots, v_\ell/v_+)$.

Perfect proportionality would require the two vectors to be equal. Because of the discreteness of the seat numbers, however, equality is almost always beyond reach and some degree of inequality must be tolerated. The question is, which inequality? Inequality among whom? An election features at least three groups of protagonists who may claim a constitutional right to equality among their group members. The largest group is the v_+ voters. A middle size group is the h Members of Parliament who are elected by the voters. The smallest, institutional group is the ℓ parties that mediate between the first group, those voting, and the second group, those elected.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_10,

For each of these groups there exists an adequate criterion function to measure a lack of group inherent equity. Minimization of different criteria points to different apportionment methods, though. The voter-oriented criterion justifies the divisor method with standard rounding (Section 10.2). The criterion that minimizes parliamentary inequity yields the divisor method with geometric rounding (Section 10.3). Minimization of inequality among parties leads to the Hare-quota method with residual fit by greatest remainders (Section 10.4).

People are not generally agreed about the relative merits of one criterion function versus the other. There is no obligation to dispose of other apportionment methods in favor of the one that optimizes some specific criterion. Every criterion mirrors but a partial view of the whole system. On the other hand, if an apportionment method is optimal with respect to some criterion, the method evidently harmonizes well with the aspects that are captured by the given criterion.

10.2. VOTER ORIENTATION: DivStd

The group of voters comes first among all actors in an election. We consider an arbitrary seat vector $x = (x_1, \ldots, x_\ell) \in \mathbb{N}^\ell(h)$, and relate it to a given vote vector $v = (v_1, \ldots, v_\ell) \in (0; \infty)^\ell$. How do we appraise the disproportionality that is present in the seat vector x, from the voters' viewpoint? A single voter who casts a vote for party j is just one individual among v_j supporters of party j. Together all v_j voters secure a success of x_j seats. Hence a single voter contributes the success share x_j/v_j. Standardization leads to the success value attributable to an individual voter,

$$\frac{x_j/h}{v_j/v_+},$$

see Section 2.7. If perfect proportionality were possible, seat shares and vote shares would coincide. The ideal success values would equal unity. However, realized success values differ from the ideal value unity almost always. If the ratio is larger than unity, voters enjoy a bit of good luck and are more successful than promised by pure proportionality. If smaller, voters have to endure a bit of bad luck. In order to neutralize the direction of the deviation, the difference between the realized success value of a voter of party j and the ideal success value unity is squared,

$$\left(\frac{x_j/h}{v_j/v_+} - 1\right)^2.$$

This per-voter index is counted once for each of the v_j voters of party j,

$$v_j \left(\frac{x_j/h}{v_j/v_+} - 1\right)^2.$$

An aggregation of these terms over all groups of voters, $j \leq \ell$, gives rise to a criterion accounting for each voter in the entire electorate,

$$f_{h,v}(x) := \sum_{j \leq \ell} v_j \left(\frac{x_j/h}{v_j/v_+} - 1\right)^2.$$

A seat vector that minimizes this goodness-of-fit criterion may be claimed to minimize electoral inequality, from the voters' viewpoint.

Theorem. *A seat vector $x \in \mathbb{N}^\ell(h)$ minimizes the above goodness-of-fit criterion if and only if x belongs to the divisor method with standard rounding,*

$$f_{h,v}(x) \le f_{h,v}(y) \quad \text{for all } y \in \mathbb{N}^\ell(h) \qquad \Longleftrightarrow \qquad x \in \mathit{DivStd}(h; v).$$

Proof. With $w_j = v_j/v_+$, plain algebra yields $f_{h,v}(x) = (v_+/h^2)\left(\sum_{j\le\ell} x_j^2/w_j\right) - v_+$. Hence the function to be minimized simplifies to $F(x) := \sum_{j\le\ell} x_j^2/w_j$.

For the proof of the direct implication let x be minimum. A seat transfer from a party i with $x_i \ge 1$ to a party $k \ne i$ yields a new seat vector y with components $y_i := x_i - 1$, $y_k := x_k + 1$, and $y_j := x_j$ for all $j \ne i, k$. Minimality secures $F(x) \le F(y)$. This inequality reduces to $x_i^2/w_i + x_k^2/w_k \le (x_i - 1)^2/w_i + (x_k + 1)^2/w_k$, that is, $(x_i - 1/2)/w_i \le (x_k + 1/2)/w_k$. A passage to the maximum (left) and the minimum (right) verifies the Max-Min Inequality 4.5 of the divisor method with standard rounding, establishing $x \in \mathrm{DivStd}(h; w) = \mathrm{DivStd}(h; v)$.

The proof of the converse implication starts with a vector $x \in \mathrm{DivStd}(h; w)$. For a seat vector $y \in \mathbb{N}^\ell(h)$ other than x, $y \ne x$, we assemble the parties with differing seat numbers in the sets $I := \{i \le \ell \mid y_i < x_i\} = \{i \le \ell \mid y_i \le x_i - 1\}$, and $K := \{k \le \ell \mid y_k > x_k\} = \{k \le \ell \mid x_k + 1 \le y_k\}$. For all $i \in I$ and $k \in K$ the definitions and the Max-Min Inequality yield

$$\frac{(x_i + y_i)/2}{w_i} \le \frac{x_i - 1/2}{w_i} \le \frac{x_k + 1/2}{w_k} \le \frac{(x_k + y_k)/2}{w_k}.$$

It follows that $\delta := \max_{i\in I}(x_i+y_i)/w_i - \min_{k\in K}(x_k+y_k)/w_k \le 0$. As surpluses balance deficits we get $\sum_{i\in I}(x_i - y_i) = \sum_{k\in K}(y_k - x_k) =: S > 0$. We obtain

$$\left(\sum_{j\le\ell} \frac{x_j^2}{w_j}\right) - \left(\sum_{j\le\ell} \frac{y_j^2}{w_j}\right) = \left(\sum_{i\in I} \frac{x_i + y_i}{w_i}\,(x_i - y_i)\right) - \left(\sum_{k\in K} \frac{x_k + y_k}{w_k}\,(y_k - x_k)\right) \le \delta S \le 0.$$

This establishes $F(x) \le F(y)$. □

With vote shares $w_j = v_j/v_+$, the criterion function admits the alternative form

$$f_{h,v}(x) = \frac{v_+}{h} \sum_{j\le\ell} \frac{(x_j - w_j h)^2}{w_j h}.$$

The sum is the familiar chi-square statistic. However, an allusion to the chi-square distribution is inappropriate. The prevalent assumption is that the vote share vector $(W_1, \ldots, W_\ell)$ is uniformly distributed (Section 6.4). Then the sums $\sum_{j\le\ell}(X_j - hW_j)^2/(hW_j)$ converge in distribution (as $h \to \infty$, then $\ell \to \infty$, with a proper standardization). But the limit is a Lévy-stable distribution, not a chi-square distribution.

Yet the rationale for the coexistence of numerator and denominator remains the same as with the chi-square statistic. It is illustrated best with a toy example. Suppose the house size is $h = 100$ and the electorate divides into three parties with vote counts $v = (45, 35, 20)$. Because of exactness every apportionment method comes up with the same answer, the seat vector $(45, 35, 20)$. Being exact the solution is devoid of any disproportionality, $f_{h,v}(45, 35, 20) = 0$.

How about seat vectors that suffer from a malapportionment of two seats for the strongest party? The two seats may be allocated (1) to the middle party, or (2) to the middle and the weakest parties, or (3) to the weakest party. The three alternatives incur the following disproportionality indices:

$$\begin{aligned}
&(1) \qquad f_{h,v}(45-2, 35+2, 20+0) = 0.089 + 0.114 + 0 = 0.203,\\
&(2) \qquad f_{h,v}(45-2, 35+1, 20+1) = 0.089 + 0.029 + 0.05 = 0.168,\\
&(3) \qquad f_{h,v}(45-2, 35+0, 20+2) = 0.089 + 0 + 0.2 = 0.289.
\end{aligned}$$

The index is 0.203 when the two diverted seats benefit the middle party (1). It is aggravated to 0.289 when they boost the weakest party (3). The increase appears quite reasonable since a two-seat surplus weighs heaviest with the weakest party. Disproportionality is least, 0.168, when the two seats are shared by the middle and weakest party (2). This is appealing since a kind of local equity is realized in a state missing exactness. As *Gauss* (1821) argued when advertising the method of least squares: The squared-error function weighs deviations in a quite natural manner

> indem man sich gewiss lieber den einfachen Fehler zweimal als den doppelten einmal gefallen läßt.
>
> as one certainly bears the simple error twice more willingly than the double error once.

Thus the divisor method with standard rounding, minimizing the sum of the squared deviations of the voters' realized success values from the ideal success value unity, harmonizes exceedingly well with equality among the individuals in the electorate.

10.3. PARLIAMENTARY ORIENTATION: DivGeo

The next group of electoral protagonists to consider is those elected. How would Members of Parliament assess any disproportionality in a given seat vector x? The representative weight of each of the x_j Members of Parliament of party j is the average number of voters he or she represents,

$$\frac{v_j}{x_j},$$

see Section 2.8. When the denominator is zero the usual convention applies, $v_j/0 = \infty$. Ideally all Members of Parliament enjoy the same representative weight, $v_j/x_j = v_+/h$ for all $j \leq \ell$. Practically, deviations from the votes-per-seats ratio are unavoidable. Again the sign of a deviation is neutralized by squaring,

$$\left(\frac{v_j}{x_j} - \frac{v_+}{h}\right)^2.$$

This per-seat index is counted once for each of the x_j representatives of party j,

$$x_j\left(\frac{v_j}{x_j} - \frac{v_+}{h}\right)^2.$$

The sum of these terms over all groups of representatives, $j \le \ell$, results in a criterion incorporating every Member of Parliament,

$$f_{h,v}(x) := \sum_{j\le\ell} x_j \left(\frac{v_j}{x_j} - \frac{v_+}{h}\right)^2.$$

A seat vector that minimizes this goodness-of-fit criterion may be claimed to minimize electoral inequality, from the viewpoint of the Members of Parliament.

Theorem. *A seat vector $x \in \mathbb{N}^\ell(h)$ minimizes the above goodness-of-fit criterion if and only if x belongs to the divisor method with geometric rounding,*

$$f_{h,v}(x) \le f_{h,v}(y) \quad \text{for all } y \in \mathbb{N}^\ell(h) \qquad \Longleftrightarrow \qquad x \in \mathit{DivGeo}(h;v).$$

Proof. Simple algebra yields $f_{h,v}(x) = \left(\sum_{j\le\ell} v_j^2/x_j\right) - v_+^2/h$. Hence the function that needs to be minimized is $F(x) := \sum_{j\le\ell} v_j^2/x_j$. Recall the convention $\epsilon/0 = \infty$ for $\epsilon > 0$.

For the proof of the direct implication let x be minimal. When the seat vector y is obtained by a seat transfer from party i with $x_i \ge 1$ to some other party $k \ne i$, minimality $F(x) \le F(y)$ implies $v_i^2/x_i + v_k^2/x_k \le v_i^2/(x_i-1) + v_k^2/(x_k+1)$, that is, $v_k^2/\big(x_k(x_k+1)\big) \le v_i^2/\big((x_i-1)x_i\big)$. Now the Max-Min Inequality 4.5 that goes along with the divisor method with geometric rounding establishes $x \in \mathrm{DivGeo}(h;v)$.

For the proof of the converse implication we assume $x \in \mathrm{DivGeo}(h;v)$. Given a vector $y \in \mathbb{N}^\ell(h)$, $y \ne x$, we again use the sets $I := \{i \le \ell \mid y_i \le x_i - 1\}$ and $K := \{k \le \ell \mid x_k + 1 \le y_k\}$. For all $k \in K$ and $i \in I$ the definitions and the Max-Min Inequality yield

$$\frac{v_k}{\sqrt{x_k y_k}} \le \frac{v_k}{\sqrt{x_i(x_k+1)}} \le \frac{v_i}{\sqrt{(x_i-1)x_i}} \le \frac{v_i}{\sqrt{x_i y_i}}.$$

It follows that $\delta := \max_{k\in K} v_k^2/(x_k y_k) - \min_{i\in I} v_i^2/(x_i y_i) \le 0$. With $\sum_{i\in I}(y_i - x_i) = \sum_{k\in K}(x_k - y_k) =: S > 0$ we get

$$\left(\sum_{j\le\ell} \frac{v_j^2}{x_j}\right) - \left(\sum_{j\le\ell} \frac{v_j^2}{y_j}\right) = \left(\sum_{k\in K} \frac{v_k^2}{x_k y_k}(y_k - x_k)\right) - \left(\sum_{i\in I} \frac{v_i^2}{x_i y_i}(x_i - y_i)\right) \le \delta S \le 0.$$

This establishes $F(x) \le F(y)$. □

With the vote shares $w_j = v_j/v_+$ the criterion takes the form

$$f_{h,v}(x) = \left(\frac{v_+}{h}\right)^2 \sum_{j\le\ell} \frac{(x_j - w_j h)^2}{x_j}.$$

The sum is a modification of the chi-square statistic where the denominator is the realized seat number x_j, not the ideal share of seats $w_j h$. As an illustration we reconsider the toy example from the previous section. The transfer of two seats from the strongest party to the others leads to the following criterion values:

(1) $f_{h,v}(45-2, 35+2, 20+0) = 0.093 + 0.108 + 0 = 0.201,$

(2) $f_{h,v}(45-2, 35+1, 20+1) = 0.093 + 0.028 + 0.048 = 0.169,$

(3) $f_{h,v}(45-2, 35+0, 20+2) = 0.093 + 0 + 0.182 = 0.275.$

The new denominator does not hamper the rationale for the criterion's usefulness. Qualitatively the numbers send the same message as in the previous section.

The divisor method with geometric rounding secures electoral equality among the Members of Parliament. It does so by minimizing the sum of the squared deviations of the deputies' realized representative weights from the votes-per-seats ratio.

10.4. PARTY ORIENTATION: HaQgrR

The third relevant group in an election system is the parties. They constitute the political institutions mediating between the people voting and the representatives elected. The success of party j is expressed through the number of seats apportioned to it,

$$x_j.$$

With w_j denoting party j's vote share, perfect proportionality would promise the party the ideal share of seats $w_j h$. Practically some deviation between the realized seat number x_j and the ideal share $w_j h$ is unavoidable, and a nonzero seat excess $x_j - w_j h$ must be tolerated. The seat excess can be positive, or negative. The direction of the deviation may again be neutralized by squaring,

$$\left(x_j - w_j h\right)^2.$$

Aggregation of the squared seat excesses of all parties gives rise to the criterion function

$$\sum_{j \le \ell} \left(x_j - w_j h\right)^2.$$

Minimization of the squared-error criterion generalizes considerably. The structural form of the goodness-of-fit criteria covered by the following theorem is

$$f_{h,v,\varphi(t)}(x) := \sum_{j \le \ell} \varphi\left(x_j - \frac{v_j}{v_+} h\right),$$

where the *score function* $\varphi(t)$, $t \in \mathbb{R}$, is assumed to be such that its *slope function*

$$\psi(t) := \frac{\varphi(t) - \varphi(t-1)}{t - (t-1)} = \varphi(t) - \varphi(t-1)$$

is non-decreasing on $\mathbb{R}$ and strictly increasing on $[0; 1]$. Feasible score functions are all functions that are convex and have a unique minimum at zero, such as the square t^2 and the modulus $|t|$. Also every strictly convex function on $\mathbb{R}$ is feasible, such as t^2, e^t and e^{-t}. The following theorem is due to *Pólya* (1919d).

Theorem. *A seat vector $x \in \mathbb{N}^\ell(h)$ minimizes a goodness-of-fit criterion $f_{h,v,\varphi(t)}$ that is specified by a score function $\varphi(t)$ as defined above if and only if x belongs to the Hare-quota method with residual fit by greatest remainders,*

$$f_{h,v,\varphi(t)}(x) \le f_{h,v,\varphi(t)}(y) \quad \text{for all } y \in \mathbb{N}^\ell(h) \qquad \Longleftrightarrow \qquad x \in HaQgrR(h; v).$$

Proof. In the objective function the terms $\varphi(x_j - w_j h)$ telescope via the slope function into $\varphi(-w_j h) + \sum_{n=1}^{x_j} \psi(n - w_j h)$, whence $f_{h,v,\varphi(t)}(x) = \sum_{j \le \ell} \big(\varphi(-w_j h) + \sum_{n=1}^{x_j} \psi(n - w_j h)\big)$. The function to be minimized is seen to be

$$F(x) := \sum_{j \le \ell} \sum_{n=1}^{x_j} \psi(n - w_j h).$$

The proof of the direct implication is in two steps. The first step delivers bounds for the entries of a seat vector $x \in \mathbb{N}^\ell(h)$ that is assumed to be minimum,

$$x_j \ge \lfloor w_j h \rfloor \quad \text{for all } j \le \ell, \tag{1}$$

$$x_j \le \lfloor w_j h \rfloor + 1 \quad \text{for all } j \le \ell. \tag{2}$$

We demonstrate (1), by contraposition. Assuming that there is a component k with $x_k < \lfloor w_k h \rfloor$ we show that x is non-minimum. It is impossible that all other components $i \ne k$ fulfill $x_i \le \lfloor w_i h \rfloor$ as this would lead to the contradiction $h = x_+ < \lfloor w_k h \rfloor + \sum_{i \ne k} \lfloor w_i h \rfloor \le h$. Hence some component $i \ne k$ satisfies $x_i > \lfloor w_i h \rfloor$. Transferring a seat from i to k we generate a rival seat vector y with entries $y_i := x_i - 1$, $y_k := x_k + 1$, and $y_j := x_j$ for all $j \ne i, k$. We get

$$F(x) - F(y) = \psi(x_i - w_i h) - \psi(x_k + 1 - w_k h). \tag{3}$$

But $x_k < \lfloor w_k h \rfloor$ entails $x_k + 1 - w_k h \le 0$, while $x_i > \lfloor w_i h \rfloor$ implies $x_i - w_i h > 0$. The monotonicity behavior of ψ yields $\psi(x_k + 1 - w_k h) \le \psi(0) < \psi(x_i - w_i h)$. It follows that $F(x) > F(y)$, whence x is non-minimum. A similar reasoning demonstrates (2).

The second step verifies the Max-Min Inequality 5.5 that belongs to HaQgrR. Let x continue to be minimum. We fix an arbitrary component i with $x_i \ge 1$, and again generate a rival vector y by transferring a seat from i to an arbitrary component $k \ne i$. Minimality implies $F(x) \le F(y)$. We claim that the following implication holds true:

$$F(x) \le F(y) \qquad \Longrightarrow \qquad x_i - w_i h \le x_k + 1 - w_k h. \tag{4}$$

The implication is shown by contraposition. Assuming that $x_i - w_i h > x_k + 1 - w_k h$, the bounds (1) and (2) imply $0 \le \lfloor w_k h \rfloor + 1 - w_k h \le x_k + 1 - w_k h < x_i - w_i h \le \lfloor w_i h \rfloor + 1 - w_i h \le 1$. Thus the arguments lie in the interval $[0; 1]$ where the slope function is strictly increasing, $\psi(x_k + 1 - w_k h) < \psi(x_i - w_i h)$. From (3) we get $F(x) > F(y)$, thus proving (4). Now (4) provides the inequality $x_i - w_i h \le x_k + 1 - w_k h$, to begin with for all i with $x_i \ge 1$ and for all $k \ne i$. Clearly the inequality holds true also for $k = i$, and for those components i that have $x_i = 0$. This leads to the Max-Min Inequality 5.5, and establishes $x \in \text{HaQgrR}(h; v)$.

The proof of the converse implication runs along the pattern of the previous proofs. Consider a seat vector $x \in \text{HaQgrR}(h; v)$, a competitor $y \in \mathbb{N}^\ell(h)$, $y \ne x$, and the subscripts where the two vectors differ, $I := \{i \le \ell \mid y_i < x_i\}$ and $K := \{k \le \ell \mid y_k > x_k\}$. We get $\sum_{i \in I}(x_i - y_i) = \sum_{k \in K}(y_k - x_k) =: S > 0$. Monotonicity of the slope function ψ and the Max-Min Inequality imply $\delta := \max_{i \in I} \psi(x_i - w_i h) - \min_{k \in K} \psi(x_k + 1 - w_k h) \le 0$. We obtain

$$\begin{aligned} F(x) - F(y) &= \left(\sum_{i \in I} \sum_{n = y_i + 1}^{x_i} \psi(n - w_i h) \right) - \left(\sum_{k \in K} \sum_{n = x_k + 1}^{y_k} \psi(n - w_k h) \right) \\ &\le \left(\sum_{i \in I} (x_i - y_i) \psi(x_i - w_i h) \right) - \left(\sum_{k \in K} (y_k - x_k) \psi(x_k + 1 - w_k h) \right) \le \delta S \le 0. \end{aligned}$$

Thus minimality of x is established, $F(x) \le F(y)$. □

Despite the theorem's ostensible generality of admitting a wide range of score functions $\varphi(t)$, the results are impaired by a lack of sensitivity. For the toy example of the previous sections we get

$$(1) \qquad f_{h,v,\varphi(t)}(45-2, 35+2, 20+0) = \varphi(-2) + \varphi(2) + \varphi(0),$$

$$(2) \qquad f_{h,v,\varphi(t)}(45-2, 35+1, 20+1) = \varphi(-2) + \varphi(1) + \varphi(1),$$

$$(3) \qquad f_{h,v,\varphi(t)}(45-2, 35+0, 20+2) = \varphi(-2) + \varphi(0) + \varphi(2).$$

Whatever the score function $\varphi(t)$, the criterion function $f_{h,v,\varphi(t)}$ fails to sense whether the two extra seats benefit the middle party or the weakest party. A conventional score function is the modulus function, $\varphi(t) = |t|$. It induces the goodness-of-fit criterion

$$f_{|t|,h,v}(x) = \sum_{j \le \ell} \left| x_j - \frac{v_j}{v_+} h \right|.$$

This criterion assigns the common weight 4 to all three cases of the toy example and thus fails to detect any distinctions whatsoever. Moreover, the lack of discriminating power jumps to the eye when in Sections 10.2–10.4 the square is replaced by the modulus,

$$\sum_{j \le \ell} v_j \left| \frac{x_j/h}{v_j/v_+} - 1 \right| = \sum_{j \le \ell} x_j \left| \frac{v_j}{x_j} - \frac{v_+}{h} \right| = \frac{v_+}{h} \sum_{j \le \ell} \left| x_j - \frac{v_j}{v_+} h \right|.$$

No matter whether voters, or Members of Parliament, or parties constitute the reference set where to measure inequality: the score is the same (up to the votes-per-seats constant). *Gauss* had good reasons why he discarded the modulus as less informative, and praised the high-level sensitivity of squared-error criteria.

10.5. CURTAILMENT OF OVERREPRESENTATION: DivDwn

Voters with a success value larger than unity point to some kind of overrepresentation of their party. The extreme overshoot amounts to

$$f_{h,v}(x) := \max_{j \le \ell} \frac{x_j/h}{v_j/v_+}.$$

Minimization of this goodness-of-fit criterion curtails worst-case overrepresentation.

Theorem. *If a seat vector $x \in \mathbb{N}^\ell(h)$ belongs to the divisor method with downward rounding then x minimizes the above goodness-of-fit criterion,*

$$x \in \mathit{DivDwn}(h;v) \qquad \Longrightarrow \qquad f_{h,v}(x) \le f_{h,v}(y) \quad \text{for all } y \in \mathbb{N}^\ell(h).$$

Proof. Given a seat vector $x \in \text{DivDwn}(h;v)$ and a competing vector $y \in \mathbb{N}^\ell(h)$, $y \ne x$, we set $K := \{k \le \ell \mid x_k + 1 \le y_k\}$. The following inequality string establishes the assertion:

$$f_{h,v}(x) = \max_{j \le \ell} \frac{x_j}{v_j} \overset{(1)}{\le} \min_{j \le \ell} \frac{x_j+1}{v_j} \overset{(2)}{\le} \min_{k \in K} \frac{x_k+1}{v_k} \overset{(3)}{\le} \min_{k \in K} \frac{y_k}{v_k} \overset{(4)}{\le} \max_{k \in K} \frac{y_k}{v_k} \overset{(5)}{\le} \max_{j \le \ell} \frac{y_j}{v_j} = f_{h,v}(y).$$

Inequality (1) is the Max-Min Inequality for DivDwn, (2) restricts the minimum to the subset K, (3) invokes the definition of K, (4) switches from a minimum to the maximum, and (5) extends the maximum to the superset of all ℓ parties. □

Generally the converse implication may fail to hold true. To see this consider the apportionment of four seats among three parties that are equally strong. The divisor method with downward rounding yields the tied solution $(2-, 1+, 1+)$, that is, it offers the three equally justified seat vectors $(2, 1, 1)$, $(1, 2, 1)$, and $(1, 1, 2)$. All of them attain the objective criterion's minimum, $3/2$. However, the seat vector $z = (2, 2, 0) \in \mathbb{N}^3(4)$ does so, too, even though it does not belong to DivDwn. In fact, z is unbalanced and belongs to no apportionment method at all.

To appreciate the workings of the criterion we evaluate the criterion function for the toy example of the previous section:

$$(1) \qquad f_{h,v}(45-2, 35+2, 20+0) = \max\{0.956, 1.057, 1\} = 1.057,$$

$$(2) \qquad f_{h,v}(45-2, 35+1, 20+1) = \max\{0.956, 1.029, 1.05\} = 1.05,$$

$$(3) \qquad f_{h,v}(45-2, 35+0, 20+2) = \max\{0.956, 1, 1.1\} = 1.1.$$

The criterion tends to spotlight weaker parties more than stronger parties. The reason is that quotients of weaker parties have smaller denominators v_j/v_+ and hence respond to changes in the numerator with bigger leaps than do quotients of stronger parties. In the example the stepsize of the strongest party is $1/45 = 0.02$, of the middle party, $1/35 = 0.03$, of the weakest party, $1/20 = 0.05$. For this reason the minimization of the criterion is more likely to push the quotients of weaker parties down to unity or even below. The consequence is that the divisor method with downward rounding is biased in favor of stronger parties at the expense of weaker parties (Chapter 7). For the same reason the method is more preferential towards stronger party groups than towards weaker party groups in terms of majorization (Chapter 8).

10.6. ALLEVIATION OF UNDERREPRESENTATION: DivUpw

Voters with a success value smaller than unity are subject to underrepresentation. The lowest success value is

$$f_{h,v}(x) := \min_{j \le \ell} \frac{x_j/h}{v_j/v_+}.$$

Maximization of this goodness-of-fit criterion alleviates worst-case underrepresentation. The motivation is akin to the strain of welfare economics that aims at maximizing the income of the poorest person in order to reduce the unequal distribution of wealth.

Theorem. *If a seat vector $x \in \mathbb{N}^\ell(h)$ belongs to the divisor method with upward rounding then x maximizes the above goodness-of-fit criterion,*

$$x \in \mathit{DivUpw}(h; v) \qquad \Longrightarrow \qquad f_{h,v}(x) \ge f_{h,v}(y) \quad \text{for all } y \in \mathbb{N}^\ell(h).$$

Proof. The proof runs parallel to the proof of Theorem 10.5. □

The converse implication does not hold true in general; an example is contained in *Pukelsheim* (1993 [312]).

The above criterion, of maximizing the smallest success value, is the counterpart of the goal in Section 10.5, of minimizing the largest success value. The success values for weaker parties are now moved upwards towards unity or even above. The current criterion advantages weaker parties more than stronger parties. Therefore, the divisor method with upward rounding is biased in favor of weaker parties at the expense of stronger parties (Chapter 7). In terms of majorization the method is more preferential towards groups of weaker parties than towards groups of stronger parties (Chapter 8).

The preceding sections exhibit, and delimit, the scope of an optimization approach to apportionment problems. The plethora of goodness-of-fit criteria divulge an impression of arbitrariness. All are attractive, none is mandatory. Moreover , the set of optimality candidates $\mathbb{N}^{\ell}(h)$ is intimidatingly large, perhaps too large. It contains proposals markedly impractical for the apportionment problem, such as allotting all seats to one party and none to the others, or such "optimal" solutions that upon second glance turn out to be unacceptable for lack of balancedness (Section 10.5). Occasionally optimality fully characterizes a method (Theorems 10.2–10.4). On other occasions a method implies optimality but not vice versa (Theorems 10.5–10.6).

Despite these reservations the optimization approach involves arguments inviting further scrutiny. In the apportionment problem the value in dispute is one seat at the least, whether it stays where it is or whether it is transferred to another party. It takes two to quarrel. The idea suggests a path of improvement by way of pairwise comparisons: Does the transfer of a seat promise an improvement? If so, carry it out, and investigate further transfers.

10.7. OPTIMIZATION OF STABILITY CRITERIA

The concept of pairwise comparisons is exemplified with the 2009 Bundestag election data (Table 2.1). Can the seat apportionment be improved by swapping a seat from one party to another? Or does the apportionment prove to be stable? The answer depends on the precise meaning of the notion of "stability". From the voters' viewpoint (and from the viewpoint of the German Federal Constitutional Court, see Section 2.7) the stability criterion ought to be based on the voters' success values. The success value of a voter of the weakest party, CSU, is found to be

$$\frac{42/598}{2\,830\,238/40\,764\,288} = \frac{0.070234}{0.069429} = 101.2 \text{ percent.}$$

The success values of the voters of the five stronger parties are, 99.7 percent for CDU voters, 99.6 for SPD voters, 100.4 for FDP voters, 100.5 for LINKE voters, and 99.8 for GRÜNE voters.

In this election the CSU voters are most successful. Their 42 CSU-seats secure a 101.2 percent success. The SPD voters are least successful, their 146 seats mean a success of only 99.6 percent. The success-value stability disparity between a CSU voter and an SPD voter amounts to

$$g_{h,v;\text{CSU},\text{SPD}}(42, 146) = |101.16 - 99.62| = 1.54 \text{ percentage points.}$$

The disparity is the difference in absolute terms, of the individual success values of a CSU voter and an SPD voter.

It is tempting to try and decrease the disparity by swapping a seat from the overrepresented CSU to the underrepresented SPD. However, the test is negative,

$$g_{h,v;\mathrm{CSU,SPD}}(41,147) = |98.75 - 100.30| = 1.55 \text{ percentage points.}$$

The seat transfer makes the success-value disparity grow, not shrink. In fact, there are $\binom{6}{2} = 15$ pairings that can be drawn from the six parties. Whichever seat transfer is examined the criterion function is seen to increase and deteriorate, rather than to decrease and improve. In this sense the apportionment in Table 2.1 is seen to be stable. The general definition of stability is as follows.

Definition. *A seat vector* $x \in \mathbb{N}^\ell(h)$ stabilizes *the criterion functions* $g_{h,v,i,k}(x_i, x_k)$ *when for all parties* $i, k \leq \ell$ *with* $x_i \geq 1$ *and* $i \neq k$ *it fulfills*

$$g_{h,v,i,k}(x_i, x_k) \leq g_{h,v,i,k}(x_i - 1, x_k + 1).$$

The remainder of the chapter shows that success-value stability is a concept distinguishing the divisor method with standard rounding (Section 10.8). Of course, other stability criteria are conceivable. Representative-weight stability characterizes the divisor method with harmonic rounding (Section 10.9). Some stability criteria turn out to be unworkable, and all *relative* stability criteria go along with the divisor method with geometric rounding (Section 10.10).

10.8. SUCCESS-VALUE STABILITY: DivStd

For a general seat vector $x \in \mathbb{N}^\ell(h)$, the *success-value criterion* of a voter of party i and a voter of party k is given by

$$g_{h,v,i,k}(x_i, x_k) := \left| \frac{x_i/h}{v_i/v_+} - \frac{x_k/h}{v_k/v_+} \right|.$$

A stabilizing seat vector x inhibits any desire to allocate one of its seat differently because any such transfer impairs the criterion function.

Theorem. *A seat vector* $x \in \mathbb{N}^\ell(h)$ *stabilizes the success-value criterion if and only if* x *belongs to the divisor method with standard rounding,*

$$g_{h,v,i,k}(x_i, x_k) \leq g_{h,v,i,k}(x_i - 1, x_k + 1) \quad \textit{for all } i, k \leq \ell \quad \Longleftrightarrow \quad x \in \mathit{DivStd}(h; v).$$

Proof. The stability definition refers to parties $i, k \leq \ell$ with $x_i \geq 1$ and $i \neq k$, only. The theorem also admits the cases $x_i = 0$ and $i = k$. No harm is done though. The case $x_i = 0$ implies $g_{h,v,i,k}(0, x_k) = (x_k/h)/(v_k/v_+) < (1/h)/(v_i/v_+) + \big((x_k+1)/h\big)/(v_k/v_+) = g_{h,v,i,k}(0-1, x_k + 1)$. The case $i = k$ evidently yields $g_{h,v,i,i}(x_i, x_i) = 0 < (2/h)/(v_i/v_+) = g_{h,v,i,i}(x_i - 1, x_i + 1)$.

It is notationally convenient to work with the vote shares $w_j = v_j/v_+$. For the direct implication assume x to be stable. Any two parties i and k fulfill either $x_i/w_i \leq x_k/w_k$; in this case the inequality $(x_i - 1/2)/w_i \leq (x_k + 1/2)/w_k$ is obvious. Or they have $x_i/w_i > x_k/w_k$. In this case the assumption forces $g_{h,v,i,k}(x_i, x_k) = x_i/(w_i h) - x_k/(w_k h) \leq g_{h,v,i,k}(x_i - 1, x_k + 1) = (x_k+1)/(w_i h) - (x_i - 1)/(w_k h)$. Thus the inequality $(x_i - 1/2)/w_i \leq (x_k + 1/2)/w_k$ holds true, in any case. Now the appropriate Max-Min Inequality establishes $x \in \mathrm{DivStd}(h; v)$.

For the converse implication we assume $x \in \mathrm{DivStd}(h; v)$. Any two parties i and k with $x_i/w_i \leq x_k/w_k$ fulfill $g_{h,v,i,k}(x_i, x_k) = x_k/(w_k h) - x_i/(w_i h) \leq (x_k + 1)/(w_k h) - (x_i - 1)/(w_i h) = g_{h,v,i,k}(x_i - 1, x_k + 1)$. The complementary case $x_i/w_i > x_k/w_k$ refers to the Max-Min Inequality to extract $(x_i - 1/2)/w_i \leq (x_k + 1/2)/w_k$, that is, $x_i/w_i - x_k/w_k \leq (x_k+1)/w_k - (x_i - 1)/w_i$. We get $g_{h,v,i,k}(x_i, x_k) = x_i/(w_i h) - x_k/(w_k h) \leq g_{h,v,i,k}(x_i - 1, x_k + 1)$. This establishes stability. □

Thus the divisor method with standard rounding is the unique method that is success-value stable. This provides yet another justification of the method, beyond the optimization of the success-value oriented goodness-of-fit criterion in Section 10.2.

10.9. REPRESENTATIVE-WEIGHT STABILITY: DivHar

Stability from the parliamentary viewpoint justifies the divisor method with harmonic rounding, not the divisor method with geometric rounding (Section 10.3). For a general seat vector $x \in \mathbb{N}^\ell(h)$ the *representative-weight criterion* for a representative of party i and a representative of party k is given by

$$g_{h,v,i,k}(x_i, x_k) := \left| \frac{v_i}{x_i} - \frac{v_k}{x_k} \right| .$$

Representative-weight stability conforms to electoral equality from the viewpoint of the Members of Parliament.

Theorem. *A seat vector $x \in \mathbb{N}^\ell(h)$ stabilizes the representative-weight criterion if and only if x belongs to the divisor method with harmonic rounding,*

$$g_{h,v,i,k}(x_i, x_k) \leq g_{h,v,i,k}(x_i - 1, x_k + 1) \quad \text{for all } i, k \leq \ell \quad \iff \quad x \in \mathit{DivHar}(h; v).$$

Proof. The harmonic signposts are $\widetilde{s}_{-1}(0) = \widetilde{s}_{-1}(1) = 0$, and $\widetilde{s}_{-1}(n) = 2\Big((n-1)^{-1} + n^{-1}\Big)^{-1}$ for $n \geq 2$, see Section 3.12.

For the proof of the direct implication we assume x to be stable. Any two parties i and k satisfy either $x_i/v_i \leq x_k/v_k$; this case yields $\widetilde{s}_{-1}(x_i)/v_i \leq \widetilde{s}_{-1}(x_k + 1)/v_k$. Or we have $x_i/v_i > x_k/v_k$; in this case we obtain $g_{h,v,i,k}(x_i, x_k) = v_k/x_k - v_i/x_i \leq g_{h,v,i,k}(x_i - 1, x_k + 1) = v_i/(x_i - 1) - v_k/(x_k + 1)$. In any case we get $v_k/\widetilde{s}_{-1}(x_k + 1) \leq v_i/\widetilde{s}_{-1}(x_i)$. Now the appropriate Max-Min Inequality establishes $x \in \mathrm{DivHar}(h; v)$.

For the proof of the converse implication we assume $x \in \mathrm{DivStd}(h; v)$. Any two parties i and k with $x_i/v_i \leq x_k/v_k$ fulfill $g_{h,v,i,k}(x_i, x_k) = v_i/x_i - v_k/x_k \leq v_i/(x_i - 1) - v_k/(x_k + 1) = g_{h,v,i,k}(x_i - 1, x_k + 1)$. The complementary case $x_i/v_i > x_k/v_k$ uses the Max-Min Inequality to establish stability. □

In summary, parliament-oriented equality is such that distinct specifications justify distinct methods. Optimization of the goodness-of-fit criterion in Section 10.3 advances the divisor method with geometric rounding. Representative-weight stability fosters the divisor method with harmonic rounding. It seems only natural that different questions have different answers. The rare occurrence of identical answers is what is truly remarkable. Voter-oriented equality points to one and the same procedure, the divisor method with standard rounding (Sections 10.2 and 10.8).

10.10. UNWORKABLE STABILITY CRITERIA

Other stability criteria for pairwise comparisons are readily invented. However, some turn out to be unworkable. They lead into a never ending circle of improvements. For instance the constitutional courts of the German states of Bavaria (1961) and Lower Saxony (1978) propose in passing that electoral equality requires the seat numbers of two parties to be in the same ratio as their vote counts. The courts' proposal sounds persuasive, but fails practically. Their wording points to the disparity criterion

$$g_{h,v,i,k}(x_i, x_k) := \left| \frac{x_i}{x_k} - \frac{v_i}{v_k} \right|.$$

Here is an example of a vicious circle, the apportionment of 16 seats among three parties proportionate to the vote counts $v = (729, 534, 337)$. With votes-per-seats ratio $1600/16 = 100$, the ideal shares of seats are $v/100 = (7.29, 5.34, 3.37)$. The ideal shares point to three seat vectors of interest, $x = (8, 5, 3)$, $y = (7, 6, 3)$, and $z = (7, 5, 4)$. The search for a stable vector loops endlessly from x to y to z, to x to y to z, *ad infinitum*:

$$g_{h,v;1,2}(x_1, x_2) = \left| \frac{8}{5} - \frac{729}{534} \right| = 0.24 > 0.20 = \left| \frac{7}{6} - \frac{729}{534} \right| = g_{h,v;1,2}(y_1, y_2),$$

$$g_{h,v;2,3}(y_2, y_3) = \left| \frac{6}{3} - \frac{534}{337} \right| = 0.42 > 0.34 = \left| \frac{5}{4} - \frac{534}{337} \right| = g_{h,v;2,3}(z_2, z_3),$$

$$g_{h,v;3,1}(z_3, z_1) = \left| \frac{4}{7} - \frac{337}{729} \right| = 0.11 > 0.09 = \left| \frac{3}{8} - \frac{337}{729} \right| = g_{h,v;3,1}(x_3, x_1).$$

None of the solutions is stable, but each has its merits. The vector x belongs to the divisor methods with standard rounding, y to the divisor method with geometric rounding, and z to the divisor method with harmonic rounding.

The example is taken from *Huntington* (1928). The paper champions comparison tests, that is, pairwise comparisons by means of a stability criterion. The goal is to improve a given seat vector, or to identify it as a stable solution. A plethora of thirty-two criteria is presented. Twelve of them turn out to be *unworkable*, in the sense that they may entail endless circles of improvement. The other twenty stability criteria are shown to lead to five procedures, namely the five traditional divisor methods listed in Section 4.4. *Huntington* classifies these findings as *a confusion of miscellaneous results*, and dismisses the twenty workable stability criteria as *undesirable*.

Huntington disposes of the thirty-two stability criteria for the good reason that he has a better criterion in the offering. His remedy is to divide a stability criterion by the smaller of its terms, and thus to create a *relative* stability criterion. In this way every (absolute) stability criterion gets matched with a relative companion variant. For instance, the *relative* success-value stability criterion and the *relative* representative-weight stability criterion are given by

$$\frac{\left| \dfrac{x_i/h}{v_i/v_+} - \dfrac{x_k/h}{v_k/v_+} \right|}{\min\left\{ \dfrac{x_i/h}{v_i/v_+}, \dfrac{x_k/h}{v_k/v_+} \right\}}, \qquad \frac{\left| \dfrac{v_i}{x_i} - \dfrac{v_k}{x_k} \right|}{\min\left\{ \dfrac{v_i}{x_i}, \dfrac{v_k}{x_k} \right\}}.$$

Sparing his readers the sight of bulky formulas throughout his exposition, *Huntington* assures them of his belief that *in the present problem it is clearly the relative or percentage difference, rather than the mere absolute difference, which is significant.* The author's persuasive technical result says that the relative variants of the thirty-two stability measures justify one and the same procedure, the divisor method with geometric rounding. This is *Huntington*'s desirable apportionment method, successfully promoted under the winning label *method of equal proportions*.

Unfortunately we fear that any rescaling blemishes the relative stability criteria to become almost surely unconstitutional. No constitution includes provisions that computational unambiguity is sufficient to equip a vote for party i with two (or more) distinct weights, some weight when it is compared to a vote for a second party j, and another weight when it is compared to a vote for a third party k. Exaggerating *Huntington*'s classification, twelve stability criteria are unworkable, twenty are undesirable, and thirty-two are unconstitutional. The crux is that any claim to exclusiveness is futile. Electoral systems are complex systems. They have plenty of facets, and admit a great many criteria highlighting one aspect or another.

10.11. IDEAL-SHARE STABILITY: DivStd

The stability concept extends to the parties' ideal shares of seats. Consider a situation where party i's seat number x_i exceeds the ideal share of seats by more than half a seat, $x_i > w_i h + 1/2$, and party k's seat number x_k falls short by more than half a seat, $x_k < w_k h - 1/2$. The transfer of a seat from party i to party k generates a vector y, via $y_i = x_i - 1$, $y_k = x_k + 1$, and $y_j = x_j$ for all $j \neq i, k$, such that parties i and k both move closer to their ideal shares of seats while the other parties maintain their status. For a seat vector to be enduring it ought to be immune against this kind of betterment.

Generally, given a house size h and a vote share vector $(w_1, \ldots, w_\ell)$, a seat vector $x \in \mathbb{N}^\ell(h)$ is said to be *ideal-share stable* when

$$\left(x_j \leq w_j h + \frac{1}{2} \quad \text{for all } j \leq \ell \right) \quad \text{or} \quad \left(x_j \geq w_j h - \frac{1}{2} \quad \text{for all } j \leq \ell \right).$$

A seat vector x fails to be ideal-share stable if and only if some parties i and k satisfy $x_i > w_i h + 1/2$ and $x_k < w_k h - 1/2$, as discussed in the previous paragraph. Ideal-share stability is a trait of the divisor method with standard rounding.

Theorem. *The divisor method with standard rounding is the only divisor method for which all solution vectors are ideal-share stable.*

Proof. Let $x \in \text{DivStd}(h; w)$ be a solution vector of the divisor method with standard rounding. With an appropriate multiplier μ the Max-Min Inequality yields $(x_j - 1/2)/w_j \leq \mu \leq (x_j + 1/2)/w_j$, that is, $x_j \leq w_j \mu + 1/2$ and $x_j \geq w_j \mu - 1/2$, for all $j \leq \ell$. The case $\mu \leq h$ entails the first stability inequality, the case $h \leq \mu$ the second. This proves ideal-share stability of DivStd.

Let $A \neq \text{DivStd}$ be another divisor method that is ideal-share stable. Since divisor methods are coherent (Theorem 9.2) it suffices to consider two-party systems. For some house size h and some vote share vector $w = (w_1, w_2)$ we get $A(h; w) \neq \text{DivStd}(h; w)$. Hence we may select a seat vector $x \in A(h; w)$ not belonging to DivStd, $x \notin \text{DivStd}(h; w)$. The negation of the Max-Min Inequality for DivStd is $(x_1+1/2)/w_1 < (x_2-1/2)/w_2$. This yields $x_1 = w_1(x_1+1/2)+w_2(x_1+1/2)-1/2 < w_1(x_1+1/2)+w_1(x_2-1/2)-1/2 = w_1h-1/2$ and, similarly, $x_2 > w_2h+1/2$. Hence $x \in A(h; w)$ is not ideal-share stable. □

This book calls the seat fractions w_jh—to which a party with vote share w_j were entitled if seats were divisible and fractional seats could be realized—the "ideal share of seats" of party j. Other authors speak of the "exact quota of seats" of party j. We close the chapter by contemplating the merits of nomenclature.

10.12. IDEAL SHARE OF SEATS VERSUS EXACT QUOTA OF SEATS

In a parliament of size h a party with vote share w_j has associated with it the ideal share of seats w_jh. This section philosophizes on the meaning of the terminology chosen. The ideal share of seats is the fractional number of seats a party could claim if fractional seats were available. But seats are indivisible. The seat fractions w_jh become practical only when rounded to integral values. The integers framing w_jh are its floor, $\lfloor w_jh \rfloor$, and its ceiling, $\lceil w_jh \rceil$.

The party's seat number x_j is said to *stay within the ideal frame* when it coincides with either one of the neighboring whole numbers of the ideal share,

$$x_j \in \Big\{ \lfloor w_jh \rfloor,\ \lceil w_jh \rceil \Big\}.$$

This restricts x_j to be one of two consecutive integers. The integers collapse to a singleton in the rare instances when the ideal share happens to be a whole number. When x_j fails to stay within the ideal frame, it is said to *violate the ideal frame*.

The ideal share of seats is often termed the *exact quota of seats*. Our substitutions of "ideal" for "exact", and of "share" for "quota" are intentional. The attribute "exact" subsists as one of the organizing principles of apportionment methods (Section 4.2). If for all parties the ideal shares of seats are whole numbers, $x_j = w_jh \in \mathbb{N}$ for all $j \leq \ell$, then they are the only acceptable solution to the apportionment problem and the attribute "exact" is to the point. Otherwise the seat fractions w_jh are idealized quantities that are continuous and in no way exact.

The term "quota", too, has its place in electoral parlance. There is the Hare-quota and the Droop-quota and a whole lot of quota variants (Section 5.8). They signify a *quota of votes*, a bunch of votes or voters needed to justify a seat. A typical phrase would read like *The candidate was just 99 votes short of the quota*, indicating a decreed, non-negotiable number of 99 votes lacking to meet the quota. The same understanding underlies the usage of a fishing quota, sales quota, production quota, women's quota, and the like. Once decreed, somebody is held responsible when the quota is missed.

Applying the term "quota" to seats imposes a non-negotiable entitlement. Non-adherence to a "quota of seats" would create a tenuous position. In contrast, the term *share of seats* offers more leeway. Reference to a "share" indicates the size of the seat allotment only broadly. An apportionment method is called for to turn the broad meaning of seat shares into a precise whole number. Anybody who owns some shares in a company is in a similar situation. The shares disclose their precise value in euros and cents only when sold, traded, or subjected to some other financial transaction.

Party	Votes	Ideal share of $h = 95$	Divisor method with standard rounding for house sizes from 95 to 104										Ideal share of $h = 104$
A	5 023	47.7185	50	50	50	50	50	50	50	50	50	50	52.2392
B	557	5.2915	5	6•	6	6	6	6	6	6	6	6	5.7928
C	556	5.2820	5	5	6•	6	6	6	6	6	6	6	5.7824
D	555	5.2725	5	5	5	6•	6	6	6	6	6	6	5.7720
E	554	5.2630	5	5	5	5	6•	6	6	6	6	6	5.7616
F	553	5.2535	5	5	5	5	5	6•	6	6	6	6	5.7512
G	552	5.2440	5	5	5	5	5	5	6•	6	6	6	5.7408
H	551	5.2345	5	5	5	5	5	5	5	6•	6	6	5.7304
I	550	5.2250	5	5	5	5	5	5	5	5	6•	6	5.7200
K	549	5.2155	5	5	5	5	5	5	5	5	5	6•	5.7096
Sum	10 000	95.0000	95	96	97	98	99	100	101	102	103	104	104.0000

TABLE 10.1 *Violations of the ideal frame.* Party A violates its ideal frame of seats, by +2 seats for house size 95 and by −2 seats for house size 104. The intermediate seats go to the equally weak nine parties B–K to promptly raise their level by one seat.

Violation of the ideal shares of seats loses its offensiveness when looking at concrete numbers. Typical examples juxtapose a single strong party with many weak parties. The contrived data in Table 10.1 use the divisor method with standard rounding. For house size 95 the strongest party A may claim an ideal share of 47.7 seat fractions but is awarded an excess of two seats, 50. For house size 104 the party gets 50 seats, two seats below its ideal share of 52.2 seat fractions. Violation of the ideal frame of the strongest party is counterbalanced by fairly looking after the nine mini-parties B–K to raise them promptly from five seats to six, one after the other. The example in Table 10.1 is an artifact, as are all other examples that can be found in the literature. The divisor method with standard rounding violates the ideal frames of seats only very rarely. None of the empirical data sets in this book lends itself as an illustration that the divisor method with standard rounding violates the ideal frame of seats.

In conclusion the term “exact quota” is too narrow, and promises a rigorousness that fails to materialize in the light of real data. The notions of “ideal share of seats” and “ideal frame of seats” are wider, and more appropriate to serve practical needs. Staying within the ideal frame of seats does not guarantee flawless apportionments. The Hare-quota method with residual fit by greatest remainders always stays within the ideal frames. Yet it suffers from limitations, as outlined in Section 9.7.

Optimization of goodness-of-fit or stability criteria originate from a view towards the whole system. A complementary view investigates some of the particular variables separately and thus highlights isolated system properties. This is the topic of the next chapter.

CHAPTER 11

Tracing Peculiarities: Vote Thresholds and Majority Clauses

Various sorts of vote thresholds are studied in detail. The minimum vote share and the maximum vote share that are compatible with a given number of seats are determined. Particular cases are the threshold of representation, and the threshold of exclusion. Amendments are presented ensuring that a party with a straight majority of votes, however narrow, is guaranteed a straight majority of seats. A specific amendment, the majority-minority partition clause, applies to a majority of votes of a single party, as well as to a majority of aggregated votes of a coalition of parties. The issue is illustrated with an example from the Bundestag-Bundesrat Conference Committee in Germany.

11.1. VOTE SHARE VARIATION FOR A GIVEN SEAT NUMBER

The apportionment task determines a seat vector $x = (x_1, \ldots, x_\ell)$ suitable for the vote counts v_j or vote shares $w_j = v_j/v_+$ of the parties $j \leq \ell$. A converse task is to find all vote shares w_j that possibly lead to a given seat number x_j. If the vote shares are too small, they entail fewer than x_j seats. If too large, they produce too many seats. Hence the feasible vote shares form an interval $[a(x_j); b(x_j)]$, extending from the *minimum vote share given x_j seats*, $a(x_j)$, to the *maximum vote share given x_j seats*, $b(x_j)$.

Two indices are of particular interest. The minimum vote share given one seat is the lowest vote share such that a party may win a seat and obtain parliamentary representation. Therefore, $a(1)$ is called the *threshold of representation.* The maximum vote share given no seat is the highest vote share such that a party is excluded from parliament because of too few votes. Hence $b(0)$ is termed *threshold of exclusion.* Clearly the threshold of representation lies below the threshold of exclusion, $a(1) \leq b(0)$.

Generally the interval $[a(x_j); b(x_j)]$ captures the vote shares w_j that potentially lead to precisely x_j seats. It is called the *support interval given x_j seats.* (In the stochastic jargon of Theorem 7.3 it is the support interval of the conditional distribution of the random vote shares W_j given the event $\{X_j = x_j\}$.) First we derive upper and lower bounds for the seat excesses $x_j - w_j h$ that later, in Theorem 11.5, are converted into formulas for $a(x_j)$ and $b(x_j)$.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*, DOI 10.1007/978-3-319-03856-8_11,

11.2. SEAT EXCESS BOUNDS: GENERAL DIVISOR METHODS

Lemma. *Let A be a divisor method with associated signpost sequence $s(0), s(1)$ etc. For every seat vector $x \in A(h;w)$ the seat excess of a party $j \leq \ell$ satisfies*

$$-(1-w_j)\Big(s(x_j+1)-x_j\Big) - w_j \sum_{i\neq j}\Big(x_i - s(x_i)\Big) \leq x_j - w_j h, \tag{1}$$

$$x_j - w_j h \leq (1-w_j)\Big(x_j - s(x_j)\Big) + w_j \sum_{i\neq j}\Big(s(x_i+1)-x_i\Big). \tag{2}$$

The lower bound (1) holds with equality if and only if $w = (v_1,\dots,v_\ell)/v_+$ where $v_j := s(x_j+1)$, and $v_i := s(x_i)$ for $i \neq j$. The upper bound (2) holds with equality if and only if $w = (v_1,\dots,v_\ell)/v_+$ where $v_j := s(x_j)$, and $v_i := s(x_i+1)$ for $i \neq j$.

Proof. The quantification "for every seat vector $x \in A(h;w)$" is a lazy version meaning, in full length, that the statement holds true for every house size h, for every number ℓ of parties, for every vote share vector $w \in (0;1)^\ell$, $w_+ = 1$, and, finally, for every seat vector $x \in A(h;w)$.

The divisor method A results in seat numbers x_i satisfying $s(x_i) \leq \mu w_i \leq s(x_i+1)$, where μ is a multiplier such that the house size is met, $x_+ = h$. With residuals $u_i := \mu w_i - x_i \in [s(x_i)-x_i; s(x_i+1)-x_i]$, the seat numbers turn into $x_i = \mu w_i - u_i$, for all $i \leq \ell$. Summation gives $h = \mu - u_+$. Subtracting from $x_j = \mu w_j - u_j$ the identity $w_j h = w_j \mu - w_j u_+$ we get

$$x_j - w_j h = -u_j + w_j u_+ = -(1-w_j)u_j + w_j \sum_{i\neq j} u_i.$$

Because of the opposing signs the residuals u_j and u_i are estimated in opposite directions to obtain the bounds (1) and (2).

Equality holds in (1) if and only if $\mu w_j - x_j = u_j = s(x_j+1)-x_j$ and $\mu w_i - x_i = u_i = s(x_i)-x_i$ for all $i \neq j$. This means $w_j = s(x_j+1)/\mu$, and $w_i = s(x_i)/\mu$ for $i \neq j$, as asserted. Equality in (2) follows similarly. Note that the equality characterization admits vanishing weights, $w_i = 0$, beyond the convention of assuming all weights to be positive. Vanishing weights emerge if some party gets no seats, $x_i = 0$, or if the signpost sequence is impervious, $s(1) = 0$. □

The reference to a general signpost sequence makes the result look rather abstract. Its message becomes concrete when specialized. For example in a two-party system with parties i and j the inequalities $x_j - s(x_j) \leq 1$ and $s(x_i+1) - x_i \leq 1$ are such that one of them is strict, because the left-right disjunction rules out that both hold with equality (Section 3.10.c). Thus the above bounds (1) and (2) entail $-1 < x_j - w_j h < 1$. The seat numbers must be of the form

$$x_j = \lfloor w_j h \rfloor \qquad \text{or} \qquad x_j = \lceil w_j h \rceil.$$

This proves that, in two-party systems, all divisor methods stay within the ideal frames of seats (Section 10.12). More can be said when the divisor method is stationary.

11.3. SEAT EXCESS BOUNDS: STATIONARY DIVISOR METHODS

Theorem. *Let DivSta$_r$ be a stationary divisor method with split $r \in [0;1]$. For every seat vector $x \in$ DivSta$_r(h;w)$ the seat excess of a party $j \le \ell$ satisfies*

$$-(1-w_j)r - w_j(1-r)M \le x_j - w_j h \le (1-w_j)(1-r)I + w_j r(\ell-1),$$

where $I := 1$ when $x_j \ge 1$ and $I := 0$ when $x_j = 0$, and where $M := \min\{\ell-1, h-x_j\}$.

Proof. The differences $s(x_i+1) - x_i = r$ are all constant. The differences $x_i - s(x_i) = 1-r$ are also constant as long as $x_i \ge 1$. The initial value $x_i = 0$ has $0 - s(0) = 0$. The distinction is captured by the indicator function I, thus establishing the upper bound.

The lower bound involves the sum $S := \sum_{i \ne j}\big(x_i - s(x_i)\big)$ that remains after attending to party j. If sufficiently many seats are left, $h - x_j \ge \ell - 1$, then all competing parties can obtain representation, $S = (1-r)(\ell-1)$. Otherwise at most $h - x_j$ parties may be allocated a seat each, $S = (1-r)(h-x_j)$. The distinction combines into $S = (1-r)M$. □

From a practical viewpoint the case $h - x_j < \ell - 1$ may be neglected. It requires an unrealistically small house size h (Section 7.9). Or it awards party j close to all seats; this is equally unrealistic. The realistic case $h - x_j \ge \ell - 1$ has $M = \ell - 1$.

For the divisor method with standard rounding ($r = 1/2$) the theorem implies

$$-\frac{1}{2} - \frac{\ell-2}{2}\, w_j \le x_j - w_j h \le \frac{1}{2} + \frac{\ell-2}{2}\, w_j.$$

In three-party systems we have $\ell - 2 = 1$. In view of $w_j < 1$ we get $-1 < x_j - w_j h < 1$. Thus the divisor method with standard rounding stays within the ideal frames of seats not just for two-party systems, but also for three-party systems. However, no divisor method stays within the ideal frames for all system sizes ℓ.

11.4. DIVISOR METHODS AND IDEAL FRAMES

Theorem.

a. *Every divisor method A is such that every seat vector $x \in A(h;w)$ has either all its components staying above the lower ideal frame, $x_j \ge \lfloor w_j h \rfloor$ for all $j \le \ell$, or all its components staying below the upper ideal frame, $x_j \le \lceil w_j h \rceil$ for all $j \le \ell$.*

b. *The divisor method with downward rounding is the only divisor method that always stays above the lower ideal frame, that is, every seat vector $x \in$ DivDwn$(h;w)$ and all parties $j \le \ell$ satisfy $x_j \ge \lfloor w_j h \rfloor$.*

c. *The divisor method with upward rounding is the only divisor method that always stays below the upper ideal frame, that is, every seat vector $x \in$ DivUpw$(h;w)$ and all parties $j \le \ell$ satisfy $x_j \le \lceil w_j h \rceil$.*

d. *No divisor method A stays within the ideal frame at all times, that is, there exist a house size h, a system size ℓ, a vote share vector $w \in (0;1)^\ell$, $w_+ = 1$, a seat vector $x \in A(h;w)$, and a party $j \le \ell$ that satisfy $x_j < \lfloor w_j h \rfloor$ or $x_j > \lceil w_j h \rceil$.*

Proof. a. Let the seat vector $x \in A(h;w)$ belong to a divisor method A with signpost sequence $s(n)$. We assume that some parties i and k satisfy $x_i < \lfloor w_i h \rfloor$ and $x_k > \lceil w_k h \rceil$. Let μ be a multiplier for x. From $x_i \in [\![\mu w_i]\!]$ we get $\mu w_i \le s(x_i+1) \le x_i + 1 \le \lfloor w_i h \rfloor \le w_i h$, and $\mu \le h$. Similarly $x_k \in [\![\mu w_k]\!]$ gives $\mu w_k \ge s(x_k) \ge x_k - 1 \ge \lceil w_k h \rceil \ge w_k h$, and $\mu \ge h$. Since the multiplier $\mu = h$ is unique, $s(x_i + 1) = x_i + 1$ and $s(x_k) = x_k - 1$ are tied. The ties contradict the left-right disjunction (Section 3.10.c), whence the assumption must be discarded.

b. The divisor method with downward rounding has $r = 1$. The first inequality of Theorem 11.3 yields the middle step in the string $-1 < -(1 - w_j) \leq x_j - w_j h \leq x_j - \lfloor w_j h \rfloor$. Integrality tightens the inequality to $0 \leq x_j - \lfloor w_j h \rfloor$. Hence the divisor method with downward rounding always stays above the lower ideal frame.

It remains to establish uniqueness. Let A be a divisor method with signpost sequence $s(n)$. Assuming $A \neq$ DivDwn, there exist a house size h and a vote share vector w such that some seat vector $x \in A(h; w)$ does not belong to the divisor method with downward rounding, $x \notin$ DivDwn$(h; w)$. Violation of the Max-Min Inequality means that two parties i and k have $(x_i + 1)/w_i < x_k/w_k$, that is, $w_i x_k - w_k(x_i + 1) > 0$. Now we construct a new problem, with

$$L := 1 + \left\lceil \frac{w_i}{w_i x_k - w_k(x_i + 1)} \right\rceil \geq 2$$

parties, weights $v_1 := w_i$ and $v_2 = \cdots = v_L := w_k$, and house size $H := x_i + (L - 1)x_k$. Method A yields the seat vector $y \in A(h; v)$ with components $y_1 := x_i$ and $y_2 = \cdots = y_L := x_k$ since the Max-Min Inequality for $x \in A(h; w)$, implies the Max-Min Inequality for $y \in A(h; v)$,

$$\max_{J \leq L} \frac{s(y_J)}{v_J} \leq \max_{j \leq \ell} \frac{s(x_j)}{w_j} \leq \min_{j \leq \ell} \frac{s(x_j + 1)}{w_j} \leq \min_{J \leq L} \frac{s(y_J + 1)}{v_J}.$$

The first party's ideal share of seats fulfills $(v_1/v_+)H = w_i(x_i + (L - 1)x_k)/(w_i + (L - 1)w_k) \geq x_i + 1$, that is, $L - 1 \geq w_i/(w_i x_k - w_k(x_i + 1))$. Integrality tightens the inequality to $L - 1 \geq \lceil w_i/(w_i x_k - w_k(x_i + 1)) \rceil$ by choice of L. The first party stays strictly below its lower ideal frame, $y_1 = x_i < x_i + 1 \leq \lfloor (v_1/v_+)H \rfloor$. Hence the apportionment method A does not stay above the lower ideal frames at all times.

c. Part c is established by an analogous construction as in part b.

d. Part d follows since parts b and c determines two distinct methods. □

This ends the diversion in how far divisor methods meet or violate the ideal frames. We now revert to vote shares w_j proper, and exhibit their support interval when the seat number x_j is given.

11.5. VOTE SHARES FOR GIVEN SEAT NUMBERS: STATIONARY DIVISOR METHODS

Theorem. *Let DivSta$_r$ be a stationary divisor method with split $r \in [0; 1]$. For every seat vector $x \in$ DivSta$_r(h; w)$ a party with x_j seats has vote share w_j obeying*

$$a(x_j) := \frac{x_j - (1 - r)I}{h - (1 - r)I + r(\ell - 1)} \leq w_j \leq \frac{x_j + r}{h + r - (1 - r)M} =: b(x_j),$$

where $I := 1$ when $x_j \geq 1$ and $I := 0$ when $x_j = 0$, and where $M := \min\{\ell - 1, h - x_j\}$.

Proof. The seat excess bounds of Theorem 11.3 are easily rearranged into the present bounds. □

We specialize the formulas when $M = \ell - 1 \leq h - x_j$. These are the only cases of practical interest. Then the minimum vote share given one seat, $a(1)$, and the maximum vote share given zero seats, $b(0)$, are

$$a(1) = \frac{r}{h - 1 + r\ell}, \qquad b(0) = \frac{r}{h + 1 - (1 - r)\ell}.$$

Occasionally the legislator concocts modifications making it more difficult to obtain a first seat. Such modifications have repercussions on the support intervals.

11.6. VOTE SHARES FOR GIVEN SEAT NUMBERS: MODIFIED DIVISOR METHODS

The divisor method with standard rounding is sometimes modified by raising the first signpost above one-half, $s(1) \geq 1/2$, while maintaining the other signposts, $s(n) = n - 1/2$ for all $n \geq 2$. The goal is to keep "very weak" parties out of parliament. The signpost $s(1) = 0.7$ is in use in Sweden, whence we term this variant the *Swedish modification* of the divisor method with standard rounding. The modification $s(1) = 1$ has also been tried. Then the first seat must be "fully earned", in the sense that it cannot arise from upward rounding of quotients below unity. The motto "Below one is none" sounds persuasive to those not affected by it. We call the procedure the *full-seat modification* of the divisor method with standard rounding. The following lines illustrate the derivation of the minimum and maximum vote shares given x_j seats.

It is instructive to examine a more general setting. We start from a stationary divisor method with split $r \in [0; 1]$, but modify it by raising the first signpost above its regular level, $t := s(1) \in [r; 1]$. The minimum vote share given one seat is tackled first. With $x_j = 1$ the upper bound of Lemma 11.2 reads $1 - w_j h \leq (1 - w_j)(1 - t) + w_j S$ where the sum $S := \sum_{i \neq j} \big(s(x_i + 1) - x_i\big)$ is the decisive term. The furthest spread of the sum materializes when some party $i \neq j$ is allotted the rest of the seats, $x_i = h - x_j = h - 1$, and the other $\ell - 2$ parties $k \neq i, j$ get nothing, giving $S = r + (\ell - 2)t$. It remains to solve for w_j. Thus the minimum vote share given one seat, for the stationary divisor method with first signpost modified into $t \geq r$, is

$$a(1) = \begin{cases} \dfrac{t}{h - (1 - r) + t(\ell - 1)} & \text{generally when } t = s(1) \in [r; 1], \\[2ex] \dfrac{0.7}{h - 1.2 + 0.7\ell} & \text{specifically for } t = 0.7 \text{ and } r = \dfrac{1}{2}, \\[2ex] \dfrac{1}{h - 1.5 + \ell} & \text{specifically for } t = 1 \text{ and } r = \dfrac{1}{2}. \end{cases}$$

The maximum vote share given zero seats is dealt with similarly. With $x_j = 0$ the lower bound in Lemma 11.2 yields $w_j h \leq t - w_j t + w_j S$, now with sum $S := \sum_{i \neq j} \big(x_i - s(x_i)\big)$. The sum involves vanishing terms $x_i - s(x_i) = 0$ for $x_i = 0$, small terms $x_i - s(x_i) = 1 - t$ for $x_i = 1$, and large terms $x_i - s(x_i) = 1 - r$ for $x_i \geq 2$. The sum is maximum provided all parties $i \neq j$ can get two or more seats, $S = (\ell - 1)(1 - r)$. Hence assuming $h \geq 2(\ell - 1)$ the maximum vote share given no seat, for the stationary divisor method with first signpost modified into $t \geq r$, is

$$b(0) = \begin{cases} \dfrac{t}{h + t - (1 - r)(\ell - 1)} & \text{generally when } t = s(1) \in [r; 1], \\[2ex] \dfrac{0.7}{h + 1.2 - 0.5\ell} & \text{specifically for } t = 0.7 \text{ and } r = \dfrac{1}{2}, \\[2ex] \dfrac{1}{h + 1.5 - 0.5\ell} & \text{specifically for } t = 1 \text{ and } r = \dfrac{1}{2}. \end{cases}$$

The assumption $h \geq 2(\ell - 1)$ conforms with the house size recommendation $h \geq 2\ell$ from Section 7.9, and so we omit the pathological cases $h < 2(\ell - 1)$.

Before discussing practical implications in Section 11.8, we derive the corresponding results for the shift-quota methods. We recall from the definition in Section 5.4 that all shift-quota methods are combined with a residual fit by greatest remainders.

11.7. VOTE SHARES FOR GIVEN SEAT NUMBERS: SHIFT-QUOTA METHODS

Theorem. *Let shQgrR$_s$ be a shift-quota method, with shift $s \in [-1;1)$. For every seat vector $x \in$ shQgrR$_s(h;w)$ the seat excess of party $j \le \ell$ satisfies*

$$s\left(w_j - \frac{1}{\ell}\right) - \left(1 - \frac{1}{\ell}\right) \le x_j - w_j h \le s\left(w_j - \frac{1}{\ell}\right) + \left(1 - \frac{1}{\ell}\right).$$

Proof. A party with vote share w_i has interim quotient $v_i/Q(s) = w_i(h+s)$. These quotients are rounded downwards or upwards according as their remainders are small or large. Let $r^* \in [0;1]$ be a split that decides about smallness or largeness (Corollary 5.6). The remaining arguments are adapted from the proof of Lemma 11.2. The residuals $u_i := (h+s)w_i - x_i \in [-(1-r^*);r^*]$ turn the seat excesses into $x_i - w_i h = s w_i - u_i$. Summation gives $0 = s - u_+$. Subtracting from $x_j - w_j h = s w_j - u_j$ the identity $0 = s/\ell - u_+/\ell$ we get

$$x_j - w_j h = s\left(w_j - \frac{1}{\ell}\right) - u_j + \frac{1}{\ell}u_+ = s\left(w_j - \frac{1}{\ell}\right) - \left(1 - \frac{1}{\ell}\right)u_j + \frac{1}{\ell}\sum_{i \ne j} u_i.$$

Because of the opposing signs the residuals u_j and the $\ell - 1$ terms u_i are estimated in opposite directions to obtain the lower and upper bounds. □

The seat excess inequalities are easily rearranged to exhibit the minimum and maximum vote shares given x_j seats:

$$a(x_j) := \frac{x_j - 1 + (1+s)/\ell}{h+s} \le w_j \le \frac{x_j + 1 - (1-s)/\ell}{h+s} =: b(x_j),$$

Hence the minimum vote share given one seat, $a(1)$, and the maximum vote share given no seat, $b(0)$, of the shift-quota method with shift s are given by

$$a(1) = \begin{cases} \dfrac{(1+s)/\ell}{h+s}, \\ \dfrac{1/\ell}{h}, \\ \dfrac{2/\ell}{h+1}; \end{cases} \quad \text{and} \quad b(0) = \begin{cases} \dfrac{1-(1-s)/\ell}{h+s} & \text{for general } s \in [-1;1) \quad (\text{shQgrR}_s), \\ \dfrac{1-1/\ell}{h} & \text{specifically for } s = 0 \quad (\text{HaQgrR}), \\ \dfrac{1}{h+1} & \text{specifically for } s = 1 \quad (\text{DQ4grR}). \end{cases}$$

Specifically, for the Hare-quota method with residual fit by greatest remainders ($s = 0$) the seat excess bounds simplify,

$$\left| x_j - w_j h \right| \le 1 - \frac{1}{\ell}.$$

A first consequence is that the method's ideal shares of seats are rounded downwards when too close to their floor, and upwards when too far away,

$$w_j h < \lfloor w_j h \rfloor + \frac{1}{\ell} \qquad \Longrightarrow \qquad x_j = \lfloor w_j h \rfloor,$$
$$w_j h > \lfloor w_j h \rfloor + 1 - \frac{1}{\ell} \qquad \Longrightarrow \qquad x_j = \lfloor w_j h \rfloor + 1.$$

A second consequence is that the identity HaQgrR$(h;v) =$ DivSta$_{r^*}(h;v)$ in Corollary 5.6 holds true with a split r^* from the sub-interval $[1/\ell; 1 - 1/\ell]$. The edge regions of the full generic interval $[0;1]$ are superfluous.

A third consequence is that within an ℓ-party system the method majorizes the stationary divisor method with split $1/\ell$, and is majorized by the stationary divisor method with split $1 - 1/\ell$,

$$\mathrm{DivSta}_{1/\ell} \prec \mathrm{HaQgrR} \prec \mathrm{DivSta}_{1-1/\ell}.$$

Since apportionment rules are required to admit party systems of arbitrary size $\ell \geq 2$ (Section 4.1), the passage to the limit $\ell \to \infty$ recovers the corner points DivUpw and DivDwn from Section 8.8. The vote threshold discussion concludes with an overview.

11.8. OVERVIEW OF VOTE THRESHOLDS

Authors use the term "threshold" in various meanings. Generally, thresholds serve to exclude groups of voters from representation in parliament. One such provision is the percentage threshold discussed in Section 7.6. Percentage thresholds are taken to form a first category of thresholds; we call them *explicit thresholds*. They are most visible to the public, and they are most explicitly appreciated by party whips.

A second category consists of *implicit thresholds*. They are stashed somewhere in the electoral law, but are neither easily recognized nor widely publicized. An example is the Romanian threshold for indeps, see Section 1.9.

A third category are *natural thresholds*. They present representation hurdles peculiar to the apportionment method used. Two indices suggest themselves to map the idea into a quantitative measure. First, should the definition be based on the smallest vote share enabling representation, $a(1)$? Second, should it build on the smallest vote share guaranteeing representation, that is, the largest vote share when representation is possibly denied, $b(0)$?

The Swiss Federal Court deliberated on the issue in a 2003 case. The court arrived at the clear decision that the constitutionally mandated natural threshold is $b(0)$, the smallest vote share above which representation is guaranteed. The other option, $a(1)$, only indicates when representation is possible but not certain. This vagueness is insufficient to substantiate a constitutional right. Thus it is the rightmost column in Table 11.1 that shows the constitutionally binding natural threshold, $b(0)$.

Every method has its own formula for the natural threshold. Yet there are some common qualitative dependencies. The decisive variable in the denominators is the house size, h. The more seats become available, the smaller is the natural threshold. This antitonic trend is certainly what common sense would expect to happen. Technically some of the formulas require the house size to be not too small, $h \geq \ell - 1$ or $h \geq 2(\ell - 1)$, as explicated in the previous sections. All these requirements are met by the general recommendation that the house size should meet or exceed twice the number of parties, $h \geq 2\ell$ (Section 7.11). The other variable in the denominator is the size of the party system, ℓ. The dependence on ℓ is only weak, and sometimes not at all present (DivDwn, DQ4grR). If present, a growing party system is accompanied with a growing natural threshold. This behavior sounds plausible, too.

Apportionment method	$a(1)$	$b(0)$ Natural threshold
DivStd: Divisor method with standard rounding	$\frac{0.5}{h-1+0.5\ell}$	$\frac{0.5}{h+1-0.5\ell}$
• Swedish modification, with $s(1) = 0.7$	$\frac{0.7}{h-1.2+0.7\ell}$	$\frac{0.7}{h+1.2-0.5\ell}$
• Full-seat modification, with $s(1) = 1$	$\frac{1}{h-1.5+\ell}$	$\frac{1}{h+1.5-0.5\ell}$
DivDwn: Divisor method with downward rounding	$\frac{1}{h-1+\ell}$	$\frac{1}{h+1}$
HaQgrR: Hare-quota method with grR	$\frac{1/\ell}{h}$	$\frac{1-1/\ell}{h}$
DQ4grR: Droop-quota variant-4 method with grR	$\frac{2/\ell}{h+1}$	$\frac{1}{h+1}$

TABLE 11.1 *Threshold formulas.* The minimum vote share given one seat, $a(1)$, is the threshold enabling representation. The maximum vote share given no seat, $b(0)$, is the threshold guaranteeing representation. The formulas depend on the method, the house size, and the size of the party system.

The natural thresholds are illustrated with two examples for the full-seat modification of the divisor method with standard rounding. The modification was written into the electoral law of the German State of North Rhine-Westphalia, only to be retracted almost instantaneously. According to a 2009 decision of the state's constitutional court the full-seat modification violates the principle of electoral equality to an extent that it is unconstitutional. The following examples are taken from the 2004 local elections that preceded the court's decision.

Four communes had councils with $h = 20$ seats and with $\ell = 4$ parties campaigning. The natural threshold of the full-seat restricted variant of the divisor method with standard rounding is $1/19.5 = 5.1$ percent. It exceeds the five percent threshold. Once recognized it is easy to illustrate the effect with numbers close to the North Rhine-Westphalian 2004 local elections. The vote vector $v = (2\,501, 701, 501, 199)$ results in the seat vector $x = (13, 4, 3, 0)$. A feasible divisor is $D = 200$, the divisor interval is $D(v, x) = [199; 200.8]$. The fourth party does not obtain representation, their 199 votes of a total of 3 902 votes constitute a share of 5.1 percent.

Three communes had a twenty-seat council and $\ell = 5$ parties standing in the election. The natural threshold increases to $1/19 = 5.3$ percent. Again the predicted threshold is easily demonstrated by appropriate data. Indeed, the vote vector $v = (1\,381, 301, 301, 181, 119)$ leads to the seat vector $x = (12, 3, 3, 2, 0)$. A feasible divisor is 120, the divisor interval is $D(v, x) = [119; 120.086]$. The fifth party is left out although their 119 votes out of a total of 2 283 votes are a share of 5.2 percent. The Constitutional Court for the State of North Rhine-Westphalia, having previously barred the five percent threshold from communal elections, decided that the full-seat modification is also unconstitutional.

11.9. PRESERVATION OF A STRAIGHT MAJORITY

Thresholds deal with weak parties whether they achieve parliamentary representation or not. At the other end of the scale one may worry whether strong parties are appor-

Party	Votes	HaQgrR=DivHar=DivGeo=DivStd=DivDwn
A	18 594 670	248
B	12 950 200	173
C	3 664 459	49
D	1 980 006	26
Sum	37 189 335	496

TABLE 11.2 *Majority preservation failure.* Party A wins a straight majority of 18 594 670 votes versus their opponents' 18 594 665 votes. All common methods apportion to party A only 248 seats, and fail to award it a straight seat majority in parliament.

tioned their due share of seats. An issue of political interest is majority preservation. Does an apportionment method always ensure that a party with a straight majority of votes—that is, a party that wins more votes than all its competitors together—is apportioned a straight majority of seats? People might want to insist that sensible methods are majority preserving. Alas, the opposite is true. No reasonable apportionment method is majority preserving.

Table 11.2 quotes an intriguing example from the 1982 Bundestag records (*Bundestagsdrucksache 9/1913* of 12 August 1982). The example was manufactured to demonstrate that the Hare-quota method with residual fit by greatest remainders may fail to preserve the majority. Party A wins a straight majority of votes, 18 594 670. Its opponents get five votes less, 18 594 665. Yet HaQgrR awards party A only half of the seats, 248 of 496, and not a straight majority. The example is reproduced in textbooks such as *Nohlen* (2009 [123]) as if the deficiency were peculiar to this particular apportionment method. However, other commonly used methods apportion the 496 seats in exactly the same fashion. Hence if the example were to invalidate one method, it would invalidate all of them. The sobering message is that no reasonable apportionment method is always majority preserving.

The divisor method with downward rounding, DivDwn, is a prime candidate for a majority preservation procedure. It is biased in favor of stronger parties at the expense of weaker parties (Section 7.8), and it prefers groups of stronger parties to groups of weaker parties (Theorem 8.6). Yet it may fail to preserve a straight majority, as evidenced in Table 11.2. The reason is that the house size in the table is even, 496. If the house size h is odd, then DivDwn is indeed majority preserving.

More precisely, the divisor method with downward rounding is the only stationary divisor method that is majority preserving for odd house sizes, $h = 2n + 1$. To see this we evaluate $b(n)$, the maximum vote share given n seats. With house size $2n + 1$ an allotment of n seats misses a straight majority by just one seat. A method is majority preserving if and only if no vote share that is leading to n seats exceeds one-half, $b(n) \leq 1/2$. By Theorem 11.5 a stationary divisor method with split r has

$$b(n) = \frac{n + r}{2n + 1 + r - (\ell - 1)(1 - r)} = \frac{n + r}{2(n + r) - (\ell - 2)(1 - r)}.$$

For two-party systems we get $b(n) = 1/2$; in this case all stationary divisor methods are majority preserving. For larger systems, $\ell \geq 3$, the inequality $b(n) \leq 1/2$ holds true if and only if $r = 1$. This excludes all stationary divisor methods except DivDwn. For odd house sizes the divisor method with downward rounding is the unique stationary method that is majority preserving.

NW2009Coesfeld	Vote Counts	Quotient	DivStd	Quotient	HaQgrR
CDU	54 233	27.1	27	27.043	27
SPD	23 648	11.8•	12	11.792	12
GRÜNE	11 798	5.9•	6	5.883	6
FDP	10 329	5.2	5	5.150	5
VWG	5 303	2.7•	3	2.644	3
LINKE	2 983	1.49	1	1.487	1
Sum (Divisor|Split)	108 294	(2000)	54	(.5)	54

TABLE 11.3 *Council election, Coesfeld county, 2009.* Despite a straight majority of votes the CDU received but half the seats, 27 of 54. Of the quotients that are rounded upwards (SPD, GRÜNE, VWG), the one with the least claim (VWG, remainder .7) loses a seat that is redirected to the CDU.

11.10. RESIDUAL SEAT REDIRECTION CLAUSE

A majority clause expresses the legislator's definitive intention that the electoral law delivers a parliamentary majority for all vote margins no matter how small. Conversely, it is conceivable that a party with fewer than half of the votes is awarded more than half of the seats by nothing else than pure computational luck. A legislator who feels obliged to implement a majority denial clause has yet to be found.

An acute majority preservation failure occurred during the 2009 communal elections in the German State of North Rhine-Westphalia. In the election of the council of the County of Coesfeld, 54 233 citizens voted for the CDU, versus 54 061 who cast their votes for other lists. Although the CDU beat their competitors by a margin of 172 voters, the divisor method with standard rounding would have apportioned them just half of the seats, 27 out of 54, rather than a straight majority. The same seat apportionment would have resulted from the Hare-Quota method with residual fit by greatest remainders. The North Rhine-Westphalian law for communal elections rectifies the failure by granting the CDU an additional seat that is redirected from one of the competing lists whose interim quotient is rounded upwards (SPD, GRÜNE, VWG). The seat is taken away from the list whose quotient has a fractional part that is smallest (VWG: .7). The smallest remainder signals the feeblest claim to be rounded upwards. Thus the seat apportionment that was put into effect was $(28, 12, 6, 5, 2, 1)$, not the unmodified seat vector $(27, 12, 6, 5, 3, 1)$. See Table 11.3.

This *residual seat redirection clause* serves its majority preservation purpose well only for the Hare-quota method with residual fit by greatest remainders. Indeed, a straight majority of votes for party j forces the ideal share of seats to exceed $h/2$, $v_j/v_+ > 1/2 \Rightarrow (v_j/v_+)h > h/2$, although ever so little. With a tiny remainder the main apportionment does not suffice, $y_j = \lfloor (v_j/v_+)h \rfloor < h/2$. The clause augments it by a residual seat, $x_j = y_j + 1 > h/2$. In Table 11.3 the CDU has quotient 27.043. Since the fractional part .043 is too small the clause redirects a residual seat from VWG to CDU in order to provide the CDU with a straight majority of 28 seats.

This majority clause was put forward already by *Gfeller* (1890). It got reinvented by *Horst Friedrich Niemeyer*, as reported by *Niemeyer / Niemeyer* (2008). The term *Hare/Niemeyer procedure* that is popular in Germany embraces not only the Hare-quota method with residual fit by greatest remainders, but also its modification by the residual seat redirection clause.

SH2013Boostedt	Second Votes	Quotient	DivStd
CDU	2 815	8.49	8
SPD	2 155	6.503	7
FWG	549	1.7	2
Sum (Divisor)	5 519	(331.4)	17

TABLE 11.4 *Council election, Boostedt community, 2013.* A straight majority of the electorate voted CDU, yet the party fell short of a straight majority of seats. The law fails to provide for a majority clause. The case generated some sardonic press comments.

When an electoral law is amended to introduce a new apportionment method an existing majority clause must be reassessed. The reassessment occasionally evades the attention of the legislator. In North Rhine-Westphalia the former Hare-quota method with residual fit by greatest remainders was replaced by the current divisor method with standard rounding. The residual seat redirection clause was left untouched. It happened to work out fine in the case of Coesfeld (Table 11.3). We illustrate its potential deficiency with the contrived data from Table 10.1, for house size 104. The divisor method with standard rounding awards the majority party A only 50 seats. The residual seat redirection clause swaps just one residual seat from party K to party A. But 51 seats in a house of 104 seats still stay below a straight majority. The seat redirection clause cannot be recommended. It may serve its purpose, or not.

11.11. HOUSE SIZE AUGMENTATION CLAUSE

In Schleswig-Holstein the former divisor method with downward rounding was replaced by the current divisor method with standard rounding. Trusting that house sizes would be mostly odd the old law had no need for a majority clause. No majority clause was written into the new law either. During the 2013 communal elections the community of Boostedt experienced a majority preservation failure. A straight majority of the electorate voted CDU, yet the party was apportioned only eight of seventeen seats, see Table 11.4. The case received considerable press coverage. Evidently majority clauses do have some relevance. We present two clauses that both perform flawlessly whatever apportionment method is used.

A viable majority clause is to augment the house size if necessary. Additional seats are created on behalf of the majority party until it reaches a straight majority of seats. This *house size augmentation clause* has the advantage of not retracting a seat from a party that was looking forward to receiving it. The clause responds rather mechanically to the objective of translating a straight majority of votes into a straight majority of seats. It presupposes that there is some leeway for the house size to be augmented if called for by extraordinary circumstances. In Boostedt (Table 11.4) the CDU would have received two extra seats to establish a straight majority, with ten of nineteen seats.

In Table 10.1 the last column with house size 104 depicts a majority preservation failure. The majority party's 50 seats lag behind the aggregated allotment of its nine opponents (54 seats) by a margin of four seats. The house size augmentation clause would create five additional seats so that the majority party is granted a straight majority of 55 seats. The clause would raise the house size from the notional level of

104 seats to a terminal level of 109 seats. If the house size had been 109 from the very beginning then the resulting seat allocation would have been just the same: 55 seats for party A, and six seats for each party B–K. Hence even though the clause looks mechanical, it often yields a proportional end result. It does so in the present example and in many others, but not always (try $h = 100$ in Table 10.1).

11.12. MAJORITY-MINORITY PARTITION CLAUSE

Another viable majority clause maintains partial proportionality in a more principled fashion. If the underlying apportionment method produces a seat vector that grants the majority party a straight majority of seats, then there is no need to interfere. Otherwise the clause starts afresh and separates the seat apportionment for the majority party from the seat apportionment for the minority parties. The majority party gets as many seats as needed to establish the smallest possible straight majority of seats. The remaining seats are apportioned among the remaining parties using the underlying apportionment method. We call this modification the *majority-minority partition clause*. In Boostedt (Table 11.4) the CDU would have received nine seats and the remaining eight seats would have been shared proportionally among the two minority parties (SPD 6, FWG 2).

The majority-minority partition clause extends from a single majority party to a coalition of parties with an aggregate vote majority. Suppose that there is a coalition of parties whose partners altogether win a straight majority of votes so narrow that the apportionment method denies them a straight majority of seats. Then the generalized variant of the majority-minority partition clause proceeds in two steps. The first step is to apportion the smallest possible straight majority of seats among the coalition partners. The second step apportions the remaining seats among the remaining parties. Both steps are calculated using the pertinent apportionment method.

If the underlying apportionment method is a divisor method then the clause results in seat vectors that are house size monotone (Section 9.6). To see this we note that the clause merges two apportionment strings. The first string is the sequential apportionment of seats among the majority parties, the second, the sequential apportionment of seats among the minority parties. Neither of the two sequences violates house size monotonicity. Since the two strings are joined together in a coherent fashion (Theorem 9.2) monotonicity is maintained. Table 11.5 provides an example.

The majority-minority partition clause has a historical precursor known under its Latin name *itio in partes*, separation into parts. The clause was an inventive novelty of parliamentary decision-making first codified in the Peace of Westphalia 1648. It stipulated that a resolution would be carried by the plenum only if carried separately by either part of the plenum. Thus it secured procedural parity between two unequal parts that were anxious to safeguard their constitutional identities. In those days the two parts were the opposing blocs of the contracting states, the Catholic bloc (*Corpus catholicorum*) and the Protestant bloc (*Corpus evangelicorum*). Nowadays, in the context of proportional representation systems in contemporary democracies, the two parts are the government majority and the opposition minority.

Political Group	Size	DivStd	MMP	DivStd	MMP	DivStd	MMP	DivStd
Government majority parties								
SPD	249	5	6•	6	7•	7	7	7
B90/GRÜNE	55	1	1	1	1	1	2•	2
Opposition minority parties								
CDU/CSU	247	4	4•	5	5•	6	6•	7
FDP	47	1	1	1	1	1	1	1
Sum	598	11	12	13	14	15	16	17
Divisor		55	45|55	45	38.2|45	38.2	35|38.2	35

TABLE 11.5 *Majority-minority partition clause for small committees, 15th German Bundestag 2002.* For committee sizes 12, 14, 16, the majority-minority partition (MMP) clause carries out two apportionments. The smallest possible seat majority is apportioned among the government majority parties. The remaining committee seats are apportioned among the remaining parties.

11.13. THE 2002 GERMAN CONFERENCE COMMITTEE DILEMMA

The 2002 German Bundestag faced a majority preservation problem that we use to illustrate the majority-minority partition clause. The Bundestag had to apportion the 16 seats of the Conference Committee among its four Political Groups. The Conference Committee mediates between the Bundestag (Federal Diet, First Chamber) and the Bundesrat (Federal Council, Second Chamber) in cases of disagreement. Existence and functions of the Conference Committee are set forth in the Basic Law, and are not at the discretion of the sitting parliament. The Bundesrat nominates one committee member per each of the 16 Federal states. They are referred to as the *Bundesrat bench*. The Basic Law stipulates that the *Bundestag bench* has the same size, 16. The number cannot be changed without amending the Basic Law.

In the 2002 Bundestag the government majority was formed by the Political Groups of SPD (249 seats, as of 1 February 2005), and of BÜNDNIS 90/Die GRÜNEN (55). The opposition minority consisted of CDU/CSU (247), and of FDP (47). The divisor method with standard rounding allocates the 16 committee seats in the division 7 : 1 to both, SPD and GRÜNE, and to CDU/CSU and FDP. Hence the government majority and the opposition minority were tied, with eight seats each. Facing a troubled apportionment decision the Bundestag habitually tables the alternative apportionments of the divisor method with downward rounding, and of the Hare-quota method with residual fit by greatest remainders. For the 2002 data both methods reproduced the troublesome tie and offered no help. The government majority, displeased with the failure to preserve its majority, decided to apportion 15 seats proportionately, and to allocate the 16th seat with the strongest Political Group, SPD. Hence the implemented seat allocation was 8 and 1 for the government majority, versus 6 and 1 for the opposition minority.

The opposition complained to the Federal Constitutional Court since they perceived the allotment an act of caprice. In an opaque decision the court ruled that the minutes of the Bundestag's Ways and Means Committee gave insufficient evidence of the reasonings underlying the apportionment of the 16 seats. The court ordered the Bundestag to reconsider the issue, and to document its decision-making process more fully in order to enable the court to assess its constitutional merits.

Table 11.5 illustrates the majority-minority partition clause with the 2005 Bundestag data. For committee size $h = 11$, the regular method gives six seats to the government majority, and five seats to the remaining parties. The majority is preserved, and the majority-minority partition clause remains dormant. It gets activated for committee size $h = 12$, since otherwise the two parts were tied with six seats each. Instead, seven seats are apportioned among the majority parties and, separately, five seats among the remaining parties. The differentials relative to the regular method are marked by a dot (•). For committee size $h = 13$, the regular method again preserves the majority, thereby re-balancing the over-all imbalance of the previous step.

The case spurring the discussion was the apportionment of the 16 seats of the Conference Committee's Bundestag bench. The implemented allocation was 8 and 1 seats for the coalition government, versus the division of 6 and 1 for the opposition parties. The majority-minority partition clause differs, with 7 and 2 seats for the coalition parties, versus the same 6 and 1 division for the remaining parties. Hence the opposition's litigation might have resulted in the junior coalition partner doubling their representation, from one to two committee seats. However, before further action was taken the Bundestag adjourned early. The issue was closed unresolved, as were all other pending issues, and filed away in the archives.

Another type of system restrictions is upper and lower limits for the eventual seat, and to these we turn next.

CHAPTER 12

Truncating Seat Ranges: Minimum-Maximum Restrictions

Some electoral systems restrict the eventual seat numbers to respect a guaranteed minimum or, occasionally, to obey a preordained maximum. Minimum restrictions or maximum restrictions form an obstacle for quota methods, but are accommodated easily by divisor methods. Restrictions commonly apply when representing geographical districts. They also arise when proportional representation of political parties is combined with the election of persons in single-seat constituencies. The varied incorporation of restrictions is exemplified with election data from the United Kingdom and Germany, and with the composition of the EP.

12.1. MINIMUM REPRESENTATION FOR ELECTORAL DISTRICTS

The Constitution of the United States of America guarantees every state of the Union at least one representative when the seats of the House of Representatives are apportioned among the various states (Section 7.8). Other countries follow suit. France allocates a minimum parliamentary representation to their Départements, the Swiss Confederation to their Cantons, the European Union to their Member States. Clearly it is reasonable to ensure that the territorial units that are subject to the legislation of a parliament are represented in this parliament.

Generally it is not automatic that an apportionment method secures a minimum representation. There is a sole exception. If the minimum restriction is one seat, then it is satisfied by every divisor method that is impervious (Section 4.4). Even tiny interim quotients guarantee at least one seat because of being rounded upwards to unity.

The exceptional setting applies to the United States House of Representatives. The legally decreed apportionment method is the divisor method with geometric rounding, advertised by *Huntington* (1928) under the winning label "method of equal proportions" (Section 10.10). Being impervious the divisor method with geometric rounding is compatible with the United States Constitution. No further action is needed to make sure that each state has at least one representative.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_12,

In other settings the minimum restriction asks for two seats or more. For instance, the composition of the EP guarantees every Member State at least six seats. Moreover, a state's seat allocation must obey a maximum restriction, 96 seats. In electoral systems that combine proportional representation with the election of persons in single-seat constituencies, the seats a party's candidates win in the constituencies are called direct seats. They impose a minimum restriction on this party's seat allocation in the proportionality calculations. As the direct-seat restrictions are known only when the election is over, they come with a more dynamic flavor than static restrictions that are decreed beforehand. Whether static or dynamic, observance of restrictions can only be guaranteed provided the apportionment methods are appropriately modified.

The modifications called for by minimum or maximum restrictions are difficult to achieve with quota methods (Section 12.2), but easy to implement with divisor methods (Section 12.3). For this reason divisor methods become the uncontested methods of choice in the presence of seat restrictions. The direct-seat restricted variant of the divisor method with downward rounding is exemplified with the 2012 election of the London Assembly (Section 12.4). The direct-seat restricted divisor method with standard rounding is illustrated with a re-evaluation of the 2009 Bundestag election (Section 12.5). The simultaneous observance of minimum and maximum restrictions is explicated with the allocation of the seats of the EP between the Member States of the European Union (Sections 12.6–12.8).

12.2. QUOTA METHOD AMBIGUITIES

As for the Hare-quota method with residual fit by greatest remainders, a *pragmatic modification* to handle minimum restrictions has been met already in Table 9.4. With the 1880 census data, the main apportionment of 300 seats allocates no seat to Delaware nor to Nevada. It is a gamble whether the two states would profit from the 18 seats that remain for the residual fit, Delaware would, Nevada would not. Hence the modification interrupts the transition from the main apportionment to the residual fit. First the endangered states receive one seat each out of the allocation of the 18 remaining seats, thus satisfying the one-seat minimum restriction for all 38 states. Thereafter the residual fit continues with the reduced number of remaining seats, 16 instead of 18, and with only 36 states after setting aside Delaware and Nevada.

In general the pragmatic modification utilizes the seats remaining after the main apportionment as a reservoir to tap for the correction of any deficiencies still persisting. In the presence of maximum restrictions, the pragmatic modification would fill the reservoir with the seat overflows emerging after the main apportionment. Clearly the strategy lacks severity. The reservoir may contain too few seats to even out all minimum restrictions, or it may be flooded with too many overflow seats so that a residual fit by greatest remainders is no longer practicable. Such difficulties are unlikely to materialize in settings as in Table 9.4, with many participants and with the lowest non-trivial restriction of one seat each. But the pragmatic modification would usually be unable to handle a six-seat minimum restriction as in the EP, or with a dynamic direct-seat restrictions as in Section 12.4.

A *principled modification* with a sound numerical base is proposed by *Balinski/ Young* (1982 [133]). The idea is to modify the Hare-quota in such a way that it honors the presence of minimum or maximum restrictions. More precisely, let party $j \le \ell$ be restricted to be allocated at least a_j seats and at most b_j seats. It goes without saying that the restrictions are taken to be compliant with a view towards lower and upper levels, $a_j \le b_j$, and with a view towards the house size, $a_+ \le h \le b_+$, lest the problem becomes void. By convention the *middle* of three numbers $a_j, v_j/Q, b_j$ is the quantity middle in value,

$$\operatorname{med}\left(a_j, \frac{v_j}{Q}, b_j\right) := \begin{cases} b_j & \text{in case } v_j/Q > b_j, \\ v_j/Q & \text{in case } v_j/Q \in (a_j; b_j], \\ a_j & \text{in case } v_j/Q \le a_j. \end{cases}$$

Evidently the median comes to lie in the interval $[a_j; b_j]$. The *modified quota* Q is defined to be the particular value that solves the equation

$$\sum_{j \le \ell} \operatorname{med}\left(a_j, \frac{v_j}{Q}, b_j\right) = h.$$

The term $\operatorname{med}(a_j, v_j/Q, b_j)$ is interpreted as the *ideal share of seats of party j subject to all restrictions being met.* Its floor constitutes the main apportionment for party j. The remaining seats are handed out via a residual fit by greatest remainders among the parties that strictly stay below the maximum restriction b_j. The only drawback is the determination of the solution Q which is laborious.

Unfortunately the first, pragmatic modification, and the second, principled modification, may produce two solutions that are distinct. As a consequence the question of how to incorporate restrictions into a quota method admits no clear and unique answer. No harm is done, as far as we know none of the contemporary electoral systems employs either of the two modifications.

The two modifications retrieve the Hare-quota method with residual fit by greatest remainders in case no restrictions prevail, $a_j = 0$ and $b_j = h$ for all $j \le \ell$. In all other cases it is a gamble whether the pragmatic modification is applicable, while the principled modification treats the quota as a flexible quantity to solve an appropriate equation. But if a sensible invocation of flexibility is what is asked for, then the answer is divisor methods, not quota methods (Section 5.10).

12.3. MINIMUM-MAXIMUM RESTRICTED VARIANTS OF DIVISOR METHODS

Divisor methods adapt to minimum-maximum restrictions more easily. They simply truncate the underlying rounding rule as need be. That is, the rounding rule never reaches below the lower level a_j nor beyond the upper level b_j. It's all very well to say that, yet a precise description requires a few notational prerequisites.

Suppose the divisor method A relies on the rounding rule $[\![\cdot]\!]$ with signpost sequence $s(n)$, $n \in \mathbb{N}$. We consider the apportionment of h seats to ℓ parties. The vector $a = (a_1, \ldots, a_\ell) \in \mathbb{N}^\ell$ is taken to embody the minimum restrictions, $b = (b_1, \ldots, b_\ell) \in \mathbb{N}^\ell$ the maximum restrictions. The restrictions are assumed to be compliant, $a_j \le b_j$ for all $j \le \ell$ and $a_+ \le h \le b_+$. In this setting the *truncated rounding rule* $[\![\cdot]\!]_{a_j}^{b_j}$ adapts the notion of a median to the set-oriented language of rounding rules. It maps a positive number $t > 0$ into the set $[\![t]\!]_{a_j}^{b_j}$ defined as

$$[\![t]\!]_{a_j}^{b_j} := \begin{cases} \{b_j\} & \text{in case } t > s(b_j), \\ [\![t]\!] & \text{in case } t \in \big(s(a_j); s(b_j)\big], \\ \{a_j\} & \text{in case } t \le s(a_j). \end{cases}$$

For $t = 0$ the no input–no output law continues to hold true, $[\![0]\!]_{a_j}^{b_j} := \{0\}$.

Now the divisor method A is modified into the *minimum-maximum restricted variant* A_a^b by mapping a vote vector $(v_1, \ldots v_\ell) \in (0; \infty)^\ell$ into the set of seat vectors

$$A_a^b(h; v) := \left\{ (x_1, \ldots, x_\ell) \in \mathbb{N}^\ell(h) \;\middle|\; x_1 \in \left[\!\!\left[\frac{v_1}{D}\right]\!\!\right]_{a_1}^{b_1}, \ldots, x_\ell \in \left[\!\!\left[\frac{v_\ell}{D}\right]\!\!\right]_{a_\ell}^{b_\ell} \text{ for some } D > 0 \right\}.$$

This means that an interim quotient v_j/D is rounded to the seat number $x_j \in [\![v_j/D]\!]$, except when the minimum restriction warrants more seats, $x_j = a_j$ in case $v_j/D \le s(a_j)$, or when the maximum restriction imposes fewer seats, $x_j = b_j$ in case $v_j/D > s(b_j)$. The divisor D ensures that all available seats are handed out, $x_+ = h$.

The phrase that describes the solution recalls that the rounding of an interim quotient is possibly overruled by a restriction: *Every D votes justify roughly one seat, except when a minimum restriction warrants more seats or a maximum restriction imposes fewer seats.* The examples that follow furnish evidence that the minimum-maximum restricted variant of a given divisor method is carried out with ease.

12.4. DIRECT-SEAT RESTRICTED VARIANT OF DivDwn

The 2012 election of the London Assembly makes use of the direct-seat restricted variant of the divisor method with downward rounding. The assembly size is 25 seats; fourteen of them are filled from single-seat constituencies by plurality vote. In a plurality vote system candidates of stronger parties are likely to come in first. In fact, eight direct seats are gained by candidates of the Labour Party, the other six direct seats by candidates of the Conservative Party. These direct seat wins persist and restrict the final proportional apportionment of the overall 25 seats.

Proportionality calculations are based on London-wide list votes. A five percent threshold applies to the 2 215 008 valid votes, amounting to 110 751. As a consequence 255 614 valid votes are rejected and nine parties and indeps are eliminated, leaving 1 959 394 votes and four parties for consideration in the apportionment process. The divisor method with downward rounding is used, subject to safeguarding the direct seat gains. In 2012, the restrictions remain inactive. Every 73 000 votes justify roughly one seat. The Labour Party wins twelve seats, thus carrying their eight direct seats plus

LondonAssembly2012	Direct Seats	List Votes	Quotient	DivDwn
Labour Party	8	911 204	12.5	12
Conservative Party	6	708 528	9.7	9
Green Party	0	189 215	2.6	2
Liberal Democrats	0	150 447	2.1	2
Sum (Divisor)	14	1 959 394	(73 000)	25

TABLE 12.1 *London Assembly, election 2012.* In order to safeguard the 14 direct seats that are elected by plurality vote in single-seat constituencies, the direct-seat restricted variant of the divisor method with downward rounding is employed. With the 2012 data, the restrictions remain inactive.

filling four seats from their list of nominees. Similarly the nine final seats for the Conservative Party support their six direct seats. See Table 12.1.

Since all direct seats are considered seeded, the seat vector $y = (8, 6, 0, 0)$ provides an initialization of the jump-and-step algorithm shortcutting the laborious tabulation that people usually go through (Section 4.11). Thereafter the comparative figures $v_j/s(y_j + 1)$ are evaluated as usual:

London 2012	Labour	Cons	Green	LibDem	Sum
Direct Seats	8	6	0	0	14
List Votes	911 204	708 528	189 215	150 447	1 959 394
Seat 15	$/(8+1) =$ 101 244.9	$/(6+1) =$ 101 218.3	$/(0+1) =$ 189 215•	$/(0+1) =$ 150 447	*Increment:* Green
Seat 16	101 244.9	101 218.3	$/(1+1) =$ 94 607.5	150 447•	LibDem
Seat 17	101 244.9•	101 218.3	94 607.5	$/(1+1) =$ 75 223.5	Labour
Seat 18	$/(9+1) =$ 91 120.4	101 218.3•	94 607.5	75 223.5	Cons
Seat 19	91 120.4	$/(7+1) =$ 88 566	94 607.5•	75 223.5	Green
Seat 20	91 120.4•	88 566	$/(2+1) =$ 63 071.7	75 223.5	Labour
Seat 21	$/(10+1) =$ 82 836.7	88 566•	63 071.7	75 223.5	Cons
Seat 22	82 836.7•	$/(8+1) =$ 78 725.3	63 071.7	75 223.5	Labour
Seat 23	$/(11+1) =$ 75 933.7	78 725.3•	63 071.7	75 223.5	Cons
Seat 24	75 933.7•	$/(9+1) =$ 70 852.8	63 071.7	75 223.5	Labour
Seat 25	$/(12+1) =$ 70 092.6	70 852.8	63 071.7	75 223.5•	LibDem
Final Seats	12	9	2	2	25
Appendix	70 092.6	70 852.8•	63 071.7	$/(2+1) =$ 50 149	

In every line the party with the highest comparative figure is flagged (•) to receive the next seat. An appendix for a twenty-sixth seat reveals the entire divisor interval [70 852.8; 75 223.5]; Table 12.1 quotes the user-friendly divisor 73 000.

SP2011Lothian	Direct Seats	Votes	Quotient	DivDwn•
Scottish National Party	8	110 953	6.9•	8
Labour	1	70 544	4.4	4
Conservatives	0	33 019	2.1	2
Green	0	21 505	1.3	1
Indep Margo MacDonald	0	18 732	1.2	1
Liberal Democrats	0	15 588	0.97	0
Sum (Divisor)	9	270 341	(16 000)	16

TABLE 12.2 *Lothian district, Scottish Parliament election 2011.* The direct-seat restricted variant of the divisor method with downward rounding allocates to parties as many seats as indicated by the quotient's integral part or, if larger, by the direct seat gains. The proviso applies to the SNP.

The variant is publicized under the name of an *additional member system.* In view of the English history of plurality elections, the fourteen direct seats are well understood and need not be further explained to the public. The other nine "additional" seats are installed to achieve or at least to move towards proportionality. The attribute "additional member" may be interpreted mischievously as if the direct seat deputies were the true representatives of the people, while the additional members are seen as decorative add-ons. This view is untenable. All parliamentary members are guaranteed their right to statutory equality. The alternate label *mixed member proportional system* is a bit more remote from such misgivings.

Generally the direct-seat restricted variant may produce a seat apportionment distinct from that of the unrestricted parent method. An example occurred in the 2011 general election for the Scottish Parliament. For the purpose of the election the Scottish territory is divided into eight electoral districts. In Scotland they are referred to as regions. The districts comprise between eight to ten single-seat constituencies. Voters cast two votes, a *constituency vote* and a *regional vote.* Every constituency returns a Member of the Scottish Parliament by plurality of constituency votes. In our parlance they figure as direct seats. The district magnitude is obtained by adding seven seats to the number of the district's constituencies. Every district evaluates its regional votes separately, using the direct-seat restricted variant of the divisor method with downward rounding. In seven districts the restrictions remain inactive. The seats allocated are the same as when employing the usual, unrestricted divisor method with downward rounding.

The restriction is activated in just one district, Lothian, see Table 12.2. With nine constituencies, it has a district magnitude of 16 seats. The Scottish National Party gains eight direct seats, Labour one. Every 16 000 votes justify roughly one seat, except when the direct seats warrant more. Hence the eight direct seats of the Scottish National Party prevail, even though the proportionality calculations yield but six seats. To highlight the intervention of the direct-seat restriction, the quotient is marked by a dot, 6.9•, and so is the header of the final seat column, "DivDwn•".

The seat apportionment may be determined just as in the previous 2012 London Assembly example. That is, the distribution of the direct seats provides the initialization $(8, 1, 0, 0, 0, 0)$. Thereafter seats 10, 11, ..., 16 are incremented one after the other as determined by the highest value of the comparative figures $v_j/(y_j + 1)$, $j \leq \ell$.

Alternatively, a faster calculation starts with the recommended divisor $D = v_+/(h + \ell/2) = 270\,341/19 = 14\,228.5$ of Section 4.10. Since the Scottish National Party's eight direct seats dominate their proportionate due, $\lfloor 7.8 \rfloor = 7$, the initial seat distribution comes out to be $y = (8, 4, 2, 1, 1, 1)$. Its sum is 17, one seat too many. The relevant comparative figures are v_j/y_j. The lowest comparative figure points to the party to decrement, Liberal Democrats. Now the seat vector $x = (8, 4, 2, 1, 1, 0)$ meets the house size $h = 16$, whence it is final. The line "Appendix" shows that a divisor above 16 509.5 decrements the next seat, this time of the Conservatives. Thus the divisor interval is [15 588; 16 509.5]; Table 12.2 quotes the user-friendly value 16 000.

Lothian 2011	SNP	Labour	Cons	Green	Indep	LibDem	Sum (Divisor)
Direct Seats	8	1	0	0	0	0	9
List Votes v_j	110 953	70 544	33 019	21 505	18 732	15 588	270 341
Quotient	7.8•	4.96	2.3	1.5	1.3	1.1	(14 228.5)
Initial Seats	8	4	2	1	1	1	17
	/8=	/4=	/2=	/1=	/1=	/1=	
Seat 16	13 869.1	17 636	16 509.5	21 505	18 732	15 588•	(Decrement •)
	8	4	2	1	1	0	16
Appendix v_j/x_j	—	17 636	16 509.5•	21 505	18 732	∞	

The seat vector of the direct-seat restricted variant is $x = (8, 4, 2, 1, 1, 0)$. In contrast, the seat vector of the unrestricted parent method, the divisor method with downward rounding, is $z = (7, 4, 2, 1, 1, 1)$. The difference between the two vectors exposes the impact of the restrictions, $x - z = (1, 0, 0, 0, 0, -1)$. The strongest party exceeds its proportionate due by one seat, the weakest party falls short by one seat. Let $|x - z|$ signify the L_1-norm of the difference vector,

$$|x - z| := \sum_{j \leq \ell} |x_j - z_j|.$$

Generally the count of transfers needed to pass back and forth between two vectors $x, z \in \mathbb{N}^\ell(h)$ equals $|x - z|/2$. We interpret the transfer count $|x - z|/2$ as the *unproportionality index* of x because it assesses the deviation from the proportionality solution z. The index is measured in terms of concrete parliamentary seats, not by means of some abstract goodness-of-fit criterion as in Chapter 10.

12.5. DIRECT-SEAT RESTRICTED VARIANT OF DivStd

The unproportionality index proves particularly enlightening in the electoral system for the German Bundestag. Direct seat gains based on first votes have to be matched with the proportionate success derived from second votes. Table 2.2 shows the 2009 evaluation according to the then valid Federal Election Law. The sub-apportionments for SPD, FDP, LINKE, and GRÜNE support their direct seats (if any) without further ado. Solely the seat apportionment for the CDU is special. The CDU receives 173 seats by the (unrestricted) the divisor method with standard rounding. In seven states the CDU wins more direct seats than allotted to its state list, thus creating 21 overhang seats. The lasting number of CDU seats is $173 + 21 = 194$.

(17BT2009)	Direct Seats	Second Votes	Quotient	DivStd•	DivStd	Difference
Sub-apportionment to districts: CDU						
Schleswig-Holstein	9	518 457	7.51•	9	9	0
Mecklenburg-Vorpommern	6	287 481	4.2•	6	5	1
Hamburg	3	246 667	3.6	4	4	0
Niedersachsen	16	1 471 530	21.3	21	24	−3
Bremen	0	80 964	1.2	1	1	0
Brandenburg	1	327 454	4.7	5	5	0
Sachsen-Anhalt	4	362 311	5.3	5	6	−1
Berlin	5	393 180	5.7	6	6	0
Nordrhein-Westfalen	37	3 111 478	45.1	45	51	−6
Sachsen	16	800 898	11.6•	16	13	3
Hessen	15	1 022 822	14.8	15	17	−2
Thüringen	7	383 778	5.6•	7	6	1
Rheinland-Pfalz	13	767 487	11.1•	13	13	0
Baden-Württemberg	37	1 874 481	27.2•	37	31	6
Saarland	4	179 289	2.6•	4	3	1
Sum (Divisor)	173	11 828 277	(69 000)	194	194	12 − 12

TABLE 12.3 *Direct-seat restricted sub-apportionment of CDU seats to districts.* In the restricted variant the better performance is decisive, direct seats or proportionate seats. Marked quotients (•) are overruled by direct seats. The unproportionality index amounts to 12 seats.

As an alternative Table 12.3 uses the direct-seat restricted variant to apportion the 194 CDU seats among the fifteen CDU district lists, DivStd• $= x$. These seat numbers are identical to the final results in Table 2.2, DivStd+Overhang, simply because of how minimum restrictions take effect. In order to determine the unproportionality index of x, Table 12.3 adjoins a penultimate column with the unrestricted apportionment of 194 seats, DivStd $= z$ (divisor 60 700). The ultimate column contains the difference between the actual apportionment and proportional solution, $x - z$. The unproportionality index amounts to $|x - z|/2 = 12$ seats. That is, the realized apportionment x is twelve seats away from the proportionate optimum z.

The merger of direct seats and proportionate seats is an asset of the Bundestag electoral system. If the two categories of votes—first votes and second votes—and the two types of seats—direct seats and list seats—are to be reconciled, the creation of additional seats must be brought from back to front. This is what the 2013 amendment of the Federal Election Law achieves. It includes a *house size adjustment step* that is likely to raise the Bundestag size beyond the notional level of 598 seats. The adjusted multiplier secures a flawless applicability of the direct-seat restricted variant.

The Federal Election Law 2013 is the topic of the next chapter whence full details are postponed. Here we only illustrate how the unproportionality levels depend on the strategy that is adopted for the house size adjustment step. The following overview lists the unproportionality indices of the CDU sub-apportionment for the adjustment strategies (a)–(f) that are detailed below:

Strategy for the House Size Adjustment Step	Bundestag Size 2009	CDU Seats	Unproportionality Index, CDU Sub-apportionment
(a) Notional size	598	173	21
(b) Direct seats +10%	653	190	15
(c) Quota-based estimates	658	191	15
(d) *Status quo* CDU seats	666	194	12
(e) Federal Election Law 2013	671	195	12
(f) Zero unproportionality	801	233	0

Different adjustment strategies yield different Bundestag sizes and different CDU seat numbers. The unproportionality index compares two apportionments, the direct-seat restricted variant and its parent method, the divisor method with standard rounding. The index counts by how many seats the two CDU sub-apportionments differ.

Strategy (a) apportions the 598 notional seats proportional to second votes, resulting in 173 CDU seats (Table 2.1). As there are 173 direct seats, the direct-seat restricted variant simply confirms the direct seat wins and otherwise cannot move at all. The allocation's unproportionality index is found to be 21 seats.

Strategy (b) adds a ten percent buffer to a party's direct seat wins. In order to produce $173 + 17 = 190$ CDU seats, the Bundestag size is raised to 653. The buffer seats enter the proportionality part. The unproportionality index decreases to 15 seats.

Strategy (c) determines the Bundestag size from first and second votes. When in district i party j wins d_{ij} direct seats and v_{ij} second votes, the seat estimate is taken to be $\max\{d_{ij}, \lfloor (v_{ij}/v_{++})598 \rfloor\}$. The overall party estimates are met when the Bundestag size grows to 658 seats. The CDU unproportionality index stays at 15 seats.

Strategy (d) reproduces the same CDU seat total they actually received in 2009, $173 + 21 = 194$. The pertinent calculations are presented in Table 12.3. The unproportionality index of the direct-seat restricted CDU sub-apportionment is 12 seats.

Strategy (e) results from using the 2013 Federal Election Law, to be detailed in Chapter 13. The unproportionality index remains unchanged, 12 seats. This option exceeds the *status quo* seats of the CDU, providing 195 seats rather than 194.

Strategy (f) shows the ultimate Bundestag size of 801 seats, when the direct seat gains of the CDU fit into the sub-apportionment calculation and all direct-seat restrictions remain inactive. The unproportionality index would be zero.

The political and constitutional impact of active restrictions can only be appreciated within the entire electoral system. In Germany an active restriction transfers seats between district lists *within* the same party. The electorate's division along party lines remains untouched. Active restrictions do not change the political composition of the Bundestag. They affect only the personal composition of the Bundestag. More direct seats of a party in a district entail more representatives from this district, whence the party sends from other districts fewer representatives than perfect proportionality would indicate. Since every Member of the Bundestag is a representative of the whole people (Section 2.5), it would appear to be of secondary importance whether a party's deputy originates from one district or the other.

In Scotland active restrictions cause seat transfers *between* parties (Section 12.4). If one party sends more Members to the Scottish Parliament than proportionality indicates, another party's allotment comprises fewer members. The political composition of the Scottish Parliament differs from what is warranted by perfect proportionality. On the other hand active restrictions do not interfere with the agreed balance between districts. The prespecified district magnitudes are observed meticulously. Every district sends seven representatives in addition to its direct seats. Regional representation is often considered a more sensitive issue than party representation, not just in Scotland. A similar sensitivity emerges when considering the composition of the EP.

12.6. COMPOSITION OF THE EP: CONSTRAINTS

The term *composition of the EP* refers to the allocation of the EP seats between the Member States of the European Union. The house size of the EP comprises 751 seats. With the July 2013 accession of Croatia the Union has 28 Member States. So far the task sounds identical to the apportionment of the 435 seats of the United States House of Representatives among the 50 states of the Union. The two problems differ in the restrictions they have to obey. The United States Constitution solely demands that each state shall have at least one representative.

The restrictions on the composition of the EP are more delicate. Article 14 Section 2 of the 2010 Treaty of Lisbon, part of the Union's primary law, imposes three constitutional constraints. The first two are numerical, the third is structural:

There is a minimum threshold of six members per Member State.

No Member State shall be allocated more than ninety-six seats.

Representation of citizens shall be degressively proportional.

The remainder of this section digresses and explicates the rationale underlying the triple constraints. We also comment on the population figures to be used in the calculations. Then two principled methods are described that provably comply with the legal stipulations, the Cambridge Compromise (Section 12.7) and its downgraded-population variant (Section 12.8). The two allocations arising from the methodological approach are compared with the actually enacted allocation for the 2014–2019 legislative period that emerged from political negotiations in the EP (Section 12.9). Finally we sketch the Jagiellonian Compromise, a qualified majority voting system for the European Council honoring the constitutional promise that all Union citizens are equal (Section 12.10).

Minimum threshold. The minimum threshold of six members per Member State enables a Member State's seat allocation to reflect the political division of its citizenry. This could barely be achieved with a minimum of one or two seats when only the strongest party would obtain representation and, possibly, the second strongest party. In 2009 Cyprus spreads its six representatives among four domestic parties and three EP Political Groups (Table 1.6), Estonia among four parties plus one indep and four Political Groups (Table 1.10), Luxembourg among four parties as well as four Political Groups (Table 1.19), and Malta among two parties and two Political Groups (Table 1.21). The relevance of a minimum threshold may fade away once the European Political Parties start performing, but at present this is wishful thinking.

Maximum allocation. The limiting maximum prevents large Member States from getting so many seats that would dominate the others. A similar restriction was part of the 1919 Weimar Constitution in Germany. For the representation of the German states in the *Reichsrat* (Second Chamber) Article 61 decreed that no state was to be allotted more than two fifth of all seats. Throughout the functioning of the Reichsrat 1919–1934 the restriction applied to the largest German state, Prussia.

The EP cap of ninety-six seats out of 751 is in absolute numbers, not relative terms. Originally the limit was ninety-nine seats, expressing a common agreement that no seat allocation would reach into the three-digit range. In view of the growth of the Union the composition of the EP became a central topic of the 2000 Intergovernmental Conference in Nice. Members of the negotiating teams classify the final outcome on the

EP composition to be one of the most illogical and arbitrary decisions of the meeting, see *Gray / Stubb* (2001). They report that *the Presidency handed out seats like loose change.* No written explanation nor oral statement was issued why the maximum cap was reduced from ninety-nine to ninety-six seats.

With 28 Member States and a house size of 751, every seat allocation procedure that is reasonably proportionate to census figures will activate the ninety-six seat limit only for Germany. The accession of further Member States, for instance of a large country like Turkey, might render the maximum restriction inactive. If some Member States were to leave the Union, the limit may affect other large states besides Germany. In any case it is worth remembering that demographic mobility may interfere and re-order the Member States' populations in a manner different from now. With 28 members the ranking of census figures is inevitably dynamic and changing.

Degressive proportionality. The term "degressive proportionality" is a neologism of unknown origin. It emerged during the political debates in the times of the European Union's precursors. Politicians welcomed degressive proportionality as a rhetorical phrase of some persuasion without defining the term in any more precise way. It expressed a vague and non-technical understanding that larger Member States have lesser representative privileges than smaller Member States. Despite its opaque ambiguity degressive proportionality was included into the text of a *Treaty establishing a Constitution for Europe* that was drafted by the 2002–2003 *Convention on the Future of Europe* in Rome. While the treaty never entered into force, degressive proportionality flourished and struck roots in the 2010 Treaty of Lisbon. According to the Union's legal proceedings the obligation to finally equip the term with a precise operational meaning lies with the EP.

The EP's first attempt was a failure, generally speaking. A 2007 parliament resolution defines degressive proportionality to mean the following:

> [Parliament] considers that the principle of degressive proportionality means that the ratio between the population and the number of seats of each Member State must vary in relation to their respective populations in such a way that each Member from a more populous Member State represents more citizens than each Member from a less populous Member State and conversely, but also that no less populous Member State has more seats than a more populous Member State.

The definition has two parts. The first part, up to the comma, addresses the representative weight of the x_i seats of a Member State i with population p_i. The representative weight is the population-per-seats average, p_i/x_i (Section 2.8). The definition requires the representative weights to be monotone. That is, two Member States i and j with decreasing population figures must exhibit non-increasing representative weights,

$$p_i > p_j \qquad \Longrightarrow \qquad \frac{p_i}{x_i} \geq \frac{p_j}{x_j}.$$

The second part, after the comma, demands that decreasing population figures entail non-increasing seat numbers,

$$p_i > p_j \qquad \Longrightarrow \qquad x_i \geq x_j.$$

In our terminology this says that degressive proportionality implies concordance.

The concordance requirement is a condition essential to democratic representation. Its denial would admit allocations awarding to a larger state fewer seats than to a smaller state, an unacceptable absurdity. Concordance is one of the organizing principles of apportionment methods (Section 4.2). In contrast the first requirement, representative-weight monotonicity, does not stand up to scrutiny and cannot be sustained even though it sounds plausible and innocuous. Concordance and representative-weight monotonicity may exclude each other, typically when several Member States get the same number of seats. The 2007 seat allocation, negotiated by the resolution's rapporteurs and used to substantiate their proposal, verified representative-weight monotonicity by pure coincidence.

The EP's second attempt will be a success. In 2013 the EP decreed its composition for the upcoming elections in 2014. The resolution amends the flawed 2007 definition of degressive proportionality by injecting the words *before rounding to whole numbers*:

> [...] the ratio between the population and the number of seats of each Member State before rounding to whole numbers must vary in relation to their respective populations [...]

The 2013 allocation, still emerging from negotiations, includes no evidence that any rounding to whole numbers is going on whence it cannot be checked for compliance with the 2013 definition of degressive proportionality. However, other sections of the 2013 resolution express the commitment that, next time, the EP will establish a seat allocation system that is objective, fair, durable and transparent.

The attributes "objective, fair, durable and transparent" meticulously capture the spirit of the Cambridge Compromise and its variants. These methods have in common that their seat numbers are of the form

$$x_i \in b + \left[\!\left[\frac{p_i^t}{D} \right]\!\right].$$

Here $b \in \mathbb{N}$ is a preordained number of base seats, $t \leq 1$ is an exponent to downgrade the population figures, and the divisor $D > 0$ is determined so that the house size is met, $x_+ = h$. The 2013 degressive proportionality definition neglects the rounding step $[\![\cdot]\!]$, and simply requires the population-per-(seats before rounding) average to be monotone. The *representative-weight-before-rounding function* is

$$f(p) = \frac{p}{b + p^t/D} = \frac{1}{bp^{-1} + (1/D)p^{t-1}}.$$

It is increasing in $p > 0$, whence monotonicity is immediate, $p_i > p_j \Rightarrow f(p_i) \geq f(p_j)$. Monotonicity is strict, or the function is constant. It is constant if and only if there is no base seat nor downgrading, $b = 0$ and $t = 1$.

All citizens are treated equally if and only if the underlying divisor method is unmodified. Otherwise, with modifications $b \geq 1$ or $t < 1$, the allocation obeys strict degressive proportionality. In plain words every allocation method is degressively proportional, in the sense of the 2013 definition, provided it first hands out some base seats or it evaluates exponentially downgraded population figures by means of a divisor method. Degressivity is no longer a concern to worry about, but is promoted from the odd constraint to a natural matter of fact.

Population figures. Population figures are the key input for the composition of the EP. It sounds a triviality that Member States announce their population figures accurately and consistently. However, the triviality transpires to be a daunting task garnished with immense obstacles and unexpected traps. Whom to count? When to count? When to estimate? How to estimate? The determination of population figures is a challenge that needs to be resolved by the Union's 28 Member States in unison.

Instead we import the figures from elsewhere. The three constitutional organs of the Union are the EP, the European Council, and the European Commission. The Treaty of Lisbon decrees a QMV rule for council decisions. If a member of the council so requests, it must be verified that the Member States constituting the qualified majority represent at least 62 percent of the total population of the Union. The *QMV-population* that is relevant for council's QMV decisions during the calender year 2013 was published in the *Official Journal of the European Union* L16, 16–17 EN (19.1.2013). Croatia's population figure was adjoined in a Note de Transmission 9855/13 JUR 260 INST 246 (23.5.2013).

Irritatingly the QMV-population figures in the *Official Journal* suffer from a preposterous format. Populations are quoted as fractional numbers, in batches of a tenth of a thousand. For example Germany's 81 843 743 citizens are disfigured into a QMV-population of 81 843.7 cohorts of a thousand. The fault does not lie with EuroStat, the Statistical Office of the European Union. Being aware that they count human beings all statistical offices exhibit population figures in whole numbers. The least we can do is to revert council's cohorts to whole numbers, such as 81 843 700 for Germany. These are the numbers displayed in Tables 12.4–6.

12.7. CAMBRIDGE COMPROMISE

The Cambridge Compromise derives its name from a workshop at the University of Cambridge as reported by *Grimmett et al.* (2011). The label to be adopted for it is "5+Upw". The label is indicative of the base stage and the proportionality stage that together establish the composition of the EP:

Each Member State receives five base seats.

The remaining seats are apportioned in proportion to population figures using the divisor method with upward rounding with a maximum cap of 91 seats.

Technically this means that a Member State i with population p_i is allocated

$$x_i \in 5 + \left\lceil\!\left\lceil \frac{p_i}{D} \right\rceil\!\right\rceil_1^{91}$$

seats; the divisor $D > 0$ is determined so that the house size is met, $x_+ = 751$.

This dual-stage procedure obeys both, the minimum threshold and the maximum allocation. Every state gets at least one proportionality seat since the divisor method with upward rounding is impervious. This seat plus the five base seats verify the minimum threshold of six seats per Member State. The maximum allocation of 96 seats is met since every Member State gets five base seats plus at most 91 proportionality seats. Degressive proportionality is realized (Section 12.6). The Cambridge Compromise is a legitimate allocation method for the composition of the EP.

Furthermore, the dual-stage construction of the Cambridge Compromise adequately fits the spirit of the Treaty of Lisbon. The treaty's implicit message is that the Union comprises two kinds of constitutional subjects, the Member States and the Union's citizens. There is a potential ambiguity in the term "Member State" over whether it refers to a state's citizenry or to a state's government. The states' governments are represented in the European Council. The meaning applicable to the EP's composition is "citizenry". Thus the dual-stage Cambridge Compromise realizes a kind of *dual electoral equality.* The base stage ascertains equality among the Member States' citizenries, and the proportionality stage ensures equality among the Union's citizens.

The proportionality stage of the Cambridge Compromise employs the divisor method with upward rounding. It does so because the method is notorious for being biased in favor of smaller Member States at the expense of larger Member States (Section 7.8). Thus biasedness resurfaces as a manifestation of a broadly understood notion of degressive proportionality. In terms of majorization, the method treats groups of smaller Member States at least as well as other methods (Sections 8.6 and 8.8). If degressive proportionality admonishes larger states to leave an advantage to smaller states, then the divisor method with upward rounding is superior to other methods.

As for the implementation of the minimum threshold of six seats two options suggest themselves. The first option is incorporated into the Cambridge Compromise by handling the maintenance of the threshold as a separate stage. Since an impervious method is to be used it is sufficient to provide every Member State with five base seats. With 28 Member States a total of 140 base seats are taken care of. This leaves 611 seats for the proportionality stage. The second option would use the minimum-maximum restricted variant of the divisor method with upward rounding (Section 12.3). It is not hard to see that the result is less degressively proportional, in whatever broad sense.

We exemplify the differentials of the two options by comparing the top and bottom thirds of the Member States (Section 7.10). Omitting Germany since it hits the maximum cap, the top third of the Member States ranges from France to Belgium, the bottom third from Ireland to Malta. The Cambridge Compromise (with allocations $x_i \in 5 + \llbracket p_i/D \rrbracket_1^{91}$) supplies the top and bottom thirds with 446 versus 73 seats. The minimum-maximum restricted variant (with allocations $x_i \in \llbracket p_i/D \rrbracket_6^{96}$) assigns to the top and bottom thirds 489 versus 56 seats. Evidently the the Cambridge Compromise is more favorable to smaller states than the minimum-maximum restricted variant. Moreover, the Cambridge Compromise deviates from the decreed 2013 composition less than the minimum-maximum restricted variant.

12.8. DOWNGRADED-POPULATION VARIANT

This variant evades an activation of the maximum restriction. In cases when the Cambridge Compromise is such that the maximum cap is inactive the variant remains silent. In cases of an active capping the variant re-distributes the ensuing overshoot in a non-linear fashion that benefits the smaller states more than the larger states. The effect is easy to appreciate with the above example. The downgraded-population variant awards the top and bottom thirds of the Member States 435 versus 76 seats. The result favors the smaller states yet more than the Cambridge Compromise, and also comes yet closer to the 2013 composition.

The *downgraded-population variant* involves an additional parameter, an exponent t, whence we label it 5+Upw^t. A Member State i with population p_i is allocated

$$x_i(t) \in 5 + \left[\!\left[\frac{p_i^t}{D} \right]\!\right]$$

seats. When the pure Cambridge Compromise is such that the maximum cap is inactive, the exponent is set to unity, $t = 1$, and the Cambridge Compromise allocation persists. Otherwise an exponent $t < 1$ is calculated to downgrade all population figures to p_i^t until the largest state $i = 1$ fits the maximum cap, $x_1(t) = 96$. The divisor $D > 0$ is determined so that the house size is met, $x_+(t) = 751$. Use of the exponent makes the maximum restriction in the seat formula for $x_i(t)$ dispensable.

Usually several exponents $t_1 < \cdots < t_n$ satisfy the definition. The corresponding seat vectors are ordered by majorization, $x(t_1) \preceq \cdots \preceq x(t_n)$, as is shown in the following theorem. Majorization means that the lowest seat vector is less preferential to groups of larger states and more preferential to groups of smaller states. Abiding to the philosophy of degressive proportionality the downgraded-population variant selects the seat vector $x(t_1)$ lowest in the majorization order. For instance, the 2013 QMV-populations admit the seat vectors $x(.927) \preceq x(.928) \preceq x(.93) \preceq x(.9345)$. The downgraded-population variant selects the vector $x(.927)$ that is shown in Table 12.4.

Theorem. *Let the divisor method A be based on an impervious signpost sequence $s(n)$, $n \geq 0$. Then, for all base seats $b \in \mathbb{N}$ and house sizes $h \in \mathbb{N}$ that are jointly feasible, the weighted population vectors $(p_1^t, \dots, p_\ell^t) \in (0; \infty)^\ell$ with exponents $t > 0$ satisfy the following two statements:*

a. *The seat vectors $x(t) \in b + A(h; p_1^t, \dots, p_\ell^t)$, $t > 0$, are increasing in the majorization order, that is, $0 < t < T \Rightarrow x(t) \preceq x(T)$.*

b. *The downgraded-population variant is well-defined, that is, there exists an exponent $t \leq 1$ such that the largest Member State is allocated 96 seats.*

Proof. a. It suffices to consider a seat vector $y(t) \in A(h - \ell b; p_1^t, \dots, p_\ell^t)$. It is accompanied by the Max-Min Inequality $\max_{j \leq \ell}\big(s\big(y_j(t)\big)/p_j^t\big) \leq \min_{i \leq \ell}\big(s\big(y_i(t)+1\big)/p_i^t\big)$ (Theorem 4.5). For all $i \leq \ell$ and $j \leq \ell$ we get $t \log(p_i/p_j) \leq \log\big(s\big(y_i(t)+1\big)\big)/s\big(y_j(t)\big)$. Imperviousness guarantees $s\big(y_i(t)+1\big) \geq 1$; in case $s\big(y_j(t)\big) = 0$ we set the right-hand side equal to infinity. With ordered populations $p_1 > \cdots > p_\ell$ the sign of $\log(p_i/p_j)$ determines the *exponent interval*

$$\left[\max_{i>j} \left(\log \frac{s\big(y_i(t)+1\big)}{s\big(y_j(t)\big)} \Big/ \log \frac{p_i}{p_j} \right);\ \min_{i<j} \left(\log \frac{s\big(y_i(t)+1\big)}{s\big(y_j(t)\big)} \Big/ \log \frac{p_i}{p_j} \right) \right].$$

All values T in this interval reproduce the seat vector $y(t)$. Let T be equal to the upper limit, and let $i < j$ be two states with $T = \log\big(s\big(y_i(t)+1\big)/s\big(y_j(t)\big)\big/\log(p_i/p_j)$. The equation implies a tie, $s\big(y_i(t)+1\big)/p_i^T = s\big(y_j(t)\big)/p_j^T$. Hence the transfer of a seat from the smaller state j to the larger state i gives rise to the next seat vector $y(T)$. Lemma 8.3 shows that the sequence of vectors thus obtained increasing in the majorization order, $y(t) \preceq y(T)$.

b. Majorization implies that the largest state $i = 1$ never loses a seat, $0 < t < T \Rightarrow y_1(t) \leq y_1(T)$ (Section 8.7). Since advancement is by transfers of single seats, the seat numbers $y_1(t)$, $t > 0$, form a discrete interval. The interval starts with $\lim_{t\to 0} y_1(t) = \lceil h/\ell \rceil - b$. Indeed, as t tends to zero all weighted population become almost uniform. Hence larger states are allocated $\lceil h/\ell \rceil - b$ seats and smaller states $\lfloor h/\ell \rfloor - b$ seats, where the division into larger and smaller states is such that the house size $h - \ell b$ is met. The interval finishes with $\lim_{t\to\infty} y_1(t) = h - \ell b - (\ell - 1)$. Indeed, as t tends to infinity the largest state turns dominant. Hence the finishing seat vector allocates just one seat to the $\ell - 1$ smaller states, and gives all other seats to the largest state.

Let $M \in \mathbb{N}$ be a preordained maximum cap that satisfies $\lim_{t\to 0} y_1(t) \leq M \leq \lim_{t\to\infty} y_1(t)$. (Table 12.4 has $22 \leq M = 91 \leq 584$.) Then the equation $y_1(t) = M$ is solvable. In case the ordinary Cambridge Compromise solution suffices, $y_1(1) \leq M$, no action is taken. In case the ordinary Cambridge Compromise exceeds the maximum cap, $y_1(1) > M$, all solutions of the equation $y_1(t) = M$ fulfill $t < 1$. (The exponent interval in part a suggests an algorithm how to step down from 1 to t.) Hence the downgraded-population variant is well-defined. □

The elaborate weighting of populations causes citizens from different Member States to be treated unequally. Although long ago there existed electoral systems granting priests and professors multiple votes, under the courteous premise that they were wiser creatures than the rest of the electorate, contemporary democracies strive for equality. The proportionality stage of the downgraded-population variant is not in line with the motto “One person, one vote” because it converts the true population figures into population indices of dubious constitutional legitimacy. On the other hand it offers a formulaic passage from the negotiated 2013 composition to the more principled Cambridge Compromise.

12.9. COMPOSITION OF THE EP: ALLOCATIONS

Table 12.4 compares the Cambridge Compromise and its downgraded-population variant with the 2013 composition enacted for the 2014–2019 legislative period, with QMV-population as explained in Section 12.6. The data provide simple examples that in the 2007 definition of degressive proportionality concordance and representative-weight monotonicity may become incompatible. For instance both Ireland (QMV-population 4 582 800) and Croatia (4 398 150) have eleven seats. If the two states become eligible for another seat, concordance forces it to go to Ireland. But a twelfth seat lowers Ireland’s representative weight (381 900) below the representative weight of Croatia (399 832). This is forbidden by representative-weight monotonicity. The 2007 definition lacks general validity. Even the 2013 composition features six discordant pairs of states wherein the smaller carries the larger representative weight (FR-DE, UK-DE, ES-IT, ES-DE, IE-FI, IE-SK). Every pair violates degressive proportionality as defined in 2007. In contrast, the allocations of the Cambridge Compromise and of its downgraded-population variant conform to the 2013 definition of degressive proportionality (Section 12.6).

EP2013 Composition	Cambridge Compromise QMV-Pop.	5+Quotient	5+Upw•	Downgraded-Pop. Variant (t = .927) QMV-Pop.^{t}	5+Quotient	5+Upw^t	2013 Seats
Germany	81 843 700	5+97.4•	96	21 643 990.3	5+90.3	96	96
France	65 397 900	5+77.9	83	17 580 355.2	5+73.3	79	74
United Kingdom	62 989 600	5+74.99	80	16 979 395.5	5+70.8	76	73
Italy	60 820 800	5+72.4	78	16 436 764.1	5+68.5	74	73
Spain	46 196 300	5+54.996	60	12 737 695.7	5+53.1	59	54
Poland	38 538 400	5+45.9	51	10 767 712.3	5+44.9	50	51
Romania	21 355 800	5+25.4	31	6 229 614.7	5+25.98	31	32
Netherlands	16 730 300	5+19.9	25	4 968 072.9	5+20.7	26	26
Greece	11 290 900	5+13.4	19	3 450 478.9	5+14.4	20	21
Belgium	11 041 300	5+13.1	19	3 379 712.3	5+14.1	20	21
Portugal	10 541 800	5+12.5	18	3 237 740.2	5+13.5	19	21
Czech Republic	10 505 400	5+12.5	18	3 227 375.3	5+13.5	19	21
Hungary	9 957 700	5+11.9	17	3 071 096.3	5+12.8	18	21
Sweden	9 482 900	5+11.3	17	2 935 110.6	5+12.2	18	20
Austria	8 443 000	5+10.1	16	2 635 497.0	5+10.99	16	18
Bulgaria	7 327 200	5+8.7	14	2 310 987.4	5+9.6	15	17
Denmark	5 580 500	5+6.6	12	1 795 419.7	5+7.5	13	13
Slovakia	5 404 300	5+6.4	12	1 742 807.7	5+7.3	13	13
Finland	5 401 300	5+6.4	12	1 741 910.8	5+7.3	13	13
Ireland	4 582 800	5+5.5	11	1 495 782.2	5+6.2	12	11
Croatia	4 398 150	5+5.2	11	1 439 830.4	5+6.004	12	11
Lithuania	3 007 800	5+3.6	9	1 012 364.2	5+4.2	10	11
Slovenia	2 055 500	5+2.4	8	711 335.5	5+2.97	8	8
Latvia	2 041 800	5+2.4	8	706 939.5	5+2.9	8	8
Estonia	1 339 700	5+1.6	7	478 339.2	5+1.99	7	6
Cyprus	862 000	5+1.03	7	317 844.9	5+1.3	7	6
Luxembourg	524 900	5+0.6	6	200 683.2	5+0.8	6	6
Malta	416 100	5+0.5	6	161 806.6	5+0.7	6	6
Sum (Divisor)	508 077 850	(840 000)	751	—	(239 800)	751	751

TABLE 12.4 *Cambridge Compromise, and its downgraded-population variant.* The two methods allocate five base seats plus proportionality seats. Using the divisor method with upward rounding proportionality is referred to the QMV-populations, or to the downgraded QMV-populations. The QMV-populations are the official 2013 figures, including Croatia, for qualified majority decisions in the council. The last column lists the seat allocations actually enacted in June 2013.

Table 12.4 invokes the 2013 QMV-population (Section 12.6). The Cambridge Compromise and its sibling set out by awarding five base seats to each Member State. The remaining 611 seats are handed out proportionally but in different ways. The Cambridge Compromise allocates one non-base seat for every 840 000 citizens or part thereof unless the maximum cap imposes 96 seats. The downgraded-population variant builds on the calculatory indices $p_i^{.927}$. It allots one non-base seat for every 239 800 index units. The exponent .927 downgrades all population figures so that the largest state ends up getting five base seats plus 91 proportionality seats. Hence it is no longer necessary to mention the maximum cap explicitly, though implicitly its presence determines the exponent .927.

A quantitative distance measure is the number of seat transfers needed to pass back and forth between the two allocations (Section 12.4). The Cambridge Compromise allocation and the 2013 composition are 29 seat transfers apart. The downgraded-population allocation and the 2013 composition are closer, calling for only 18 transfers of the 751 seats. Moreover, the downgraded-population allocation entails spacings between the five largest states that appear to be more natural than the spacings of the Cambridge Compromise (96, 79, 76, 74, 59 versus 96, 83, 80, 78, 60).

EP2013 Composition	Cambridge Compromise Limited-Loss Variant QMV-Pop.	5+Min..Max	5+Quotient	5+Upw•	2013 Seats
Germany	81 843 700	5+89..91	5+95.6•	96	96
France	65 397 900	5+67..91	5+76.4	82	74
United Kingdom	62 989 600	5+66..91	5+73.5	79	73
Italy	60 820 800	5+66..91	5+71.01	77	73
Spain	46 196 300	5+47..91	5+53.9	59	54
Poland	38 538 400	5+44..91	5+44.995	50	51
Romania	21 355 800	5+25..91	5+24.9	30	32
Netherlands	16 730 300	5+19..91	5+19.5	25	26
Greece	11 290 900	5+14..91	5+13.2	19	21
Belgium	11 041 300	5+14..91	5+12.9•	19	21
Portugal	10 541 800	5+14..91	5+12.3•	19	21
Czech Republic	10 505 400	5+14..91	5+12.3•	19	21
Hungary	9 957 700	5+14..91	5+11.6•	19	21
Sweden	9 482 900	5+13..91	5+11.1•	18	20
Austria	8 443 000	5+11..91	5+9.9•	16	18
Bulgaria	7 327 200	5+10..91	5+8.6•	15	17
Denmark	5 580 500	5+6..91	5+6.5	12	13
Slovakia	5 404 300	5+6..91	5+6.3	12	13
Finland	5 401 300	5+6..91	5+6.3	12	13
Ireland	4 582 800	5+4..91	5+5.4	11	11
Croatia	4 398 150	5+4..91	5+5.1	11	11
Lithuania	3 007 800	5+4..91	5+3.5	9	11
Slovenia	2 055 500	5+1..91	5+2.4	8	8
Latvia	2 041 800	5+1..91	5+2.4	8	8
Estonia	1 339 700	5+1..91	5+1.6	7	6
Cyprus	862 000	5+1..91	5+1.01	7	6
Luxembourg	524 900	5+1..91	5+0.6	6	6
Malta	416 100	5+1..91	5+0.5	6	6
Sum (Divisor)	508 077 850	—	(856 500)	751	751

TABLE 12.5 *Limited-loss variant of the Cambridge Compromise.* Every Member State is allocated five base seats. The remaining seats are apportioned proportionately to QMV-populations, with a minimum restriction of the 2013 composition minus two seats, and a maximum cap of 91 seats.

The trouble with the downgraded-population variant is its dubious constitutional status. Does the invocation of the calculatory indices p_i^{927} violate the principle of a direct election? What do the indices mean? The 416 100 Maltese citizens are transformed into index 161 806.6. Is the ratio 161 806.6/416 100 ≈ .4 an indication that a full citizen in Malta corresponds to a "forty percent citizen in the Union?

If the purpose of the downgraded-population variant is to dampen the losses of the currently overrepresented middle-size Member States, then it might be worthwhile to explicitly impose a limit to such losses and employ a minimum-maximum restricted variant of the Cambridge Compromise. As an example we may insist that no Member State loses more than two seats of its 2013 composition. The *limited-loss variant* adapts the proportionality stage of the Cambridge Compromise as follows:

> The remaining seats are apportioned in proportion to population figures using the divisor method with upward rounding with a maximum cap of 91 seats and with a minimum restriction of the *status quo* allocation minus two seats.

Table 12.5 shows the ensuing allocation. For instance the minimum restriction for Hungary is 14 seats. Indeed, the addition of five base seats gives at least 19 seats, two seats below the *status quo* allocation (21). Hungary would be one of seven states that turn out to be shielded by the minimum restriction. The limited-loss variant evades dubious calculatory diversions and maintains a direct reference to population figures. It allocates one non-base seat for every 856 500 citizens or part thereof, except when the minimum restrictions warrant more seats or the maximum cap imposes fewer seats. In Table 12.5 all exceptions are marked with a dot(•). The limitation of losses may find its constitutional justification in the pragmatic principle of electoral continuity that exempts the legislator from amending the electoral system too abruptly when a gentle transition is also feasible.

12.10. JAGIELLONIAN COMPROMISE

The *Jagiellonian Compromise* proposes a qualified majority voting system for the European Council. The council is the decision-making body of the Member States' governments. Hence the Jagiellonian Compromise does not belong to the systems for representing people that are in the focus of this book. Yet the procedure pays due attention to the status of the Union's citizens. In this sense the Jagiellonian Compromise is in agreement with the Cambridge Compromise, its later namesake. In view of the philosophical affinity we take the space to briefly outline its essence.

In a QMV system, every Member State is equipped with a voting weight. A group of Member States qualifies as a majority when their cumulative voting weights meet or exceed a preordained quota.

The Jagiellonian Compromise derives its voting weights and quota directly from the Member States' population figures, and it does so in a remarkably transparent way. The *voting weight* of Member State i simply is the square root of its QMV-population, $\sqrt{p_i}$. The *quota* q is calculated from the QMV-populations by averaging the sum of the square roots and the square root of the sum,

$$q := \frac{1}{2}\left(\sum_{i\le 28}\sqrt{p_i} + \sqrt{\sum_{i\le 28} p_i}\right).$$

Table 12.6 exhibits weights and quota for the year 2013. For the sake of simplicity all quantities are commercially rounded. The quota 60 723 amounts to 61.4 percent of the sum of the voting weights.

In the past the council and its precursors agreed on weights and quota by negotiation. The negotiated systems cast doubt on the expertise of European diplomats and their advisers. In the European Economic Community 1958–1972, Germany, France and Italy each had weight 4, the Netherlands and Belgium 2 each, and Luxembourg one. The quota was set to be 12. This system gave Luxembourg no decision power whatsoever. The other states commanded twelve votes or more, or ten votes or less. Luxembourg was never decisive to reach the quota, or to fail it. In 1973 Luxembourg was assigned two votes, and Ireland and Denmark three each. For the ten Member States 1981–1985 the quota was set at 45. Among the 1024 possible voting profiles Luxembourg was decisive as often as were Ireland and Denmark. The three states had the same decision power despite of the ostensible distinctness of their voting weights. Diplomatic good-will cannot substitute for factual expertise.

EuropeanCouncil2013	QMV-Population	Voting Weight	Power Share
Germany	81 843 700	9 047	9.15
France	65 397 900	8 087	8.18
United Kingdom	62 989 600	7 937	8.02
Italy	60 820 800	7 799	7.89
Spain	46 196 300	6 797	6.87
Poland	38 538 400	6 208	6.28
Romania	21 355 800	4 621	4.67
Netherlands	16 730 300	4 090	4.14
Greece	11 290 900	3 360	3.40
Belgium	11 041 300	3 323	3.36
Portugal	10 541 800	3 247	3.28
Czech Republic	10 505 400	3 241	3.28
Hungary	9 957 700	3 156	3.19
Sweden	9 482 900	3 079	3.11
Austria	8 443 000	2 906	2.94
Bulgaria	7 327 200	2 707	2.74
Denmark	5 580 500	2 362	2.39
Slovakia	5 404 300	2 325	2.35
Finland	5 401 300	2 324	2.35
Ireland	4 582 800	2 141	2.16
Croatia	4 398 150	2 097	2.12
Lithuania	3 007 800	1 734	1.75
Slovenia	2 055 500	1 434	1.45
Latvia	2 041 800	1 429	1.44
Estonia	1 339 700	1 157	1.17
Cyprus	862 000	928	0.94
Luxembourg	524 900	724	0.73
Malta	416 100	645	0.65
Sum	508 077 850	98 905	100.00
Quota		60 723	61.40

TABLE 12.6 *Jagiellonian Compromise proposal for the European Council.* A Member State's voting weight is the square root of its QMV-population. A group of Member States qualifies as a majority when their cumulative weights meet or exceed the quota, 60 723. The quota is the average of the sum of the voting weights (98 905) and the square-root of the population total ($\sqrt{508\,077\,850}$). The system gives all Union citizens the same power to contribute indirectly via their governments.

The number of voting profiles in which Member State i turns decisive by passing from Yea to Nay or from Nay to Yea provides a meaningful measure for the state's decision power. The 28 Member States of the current Union allow 268 million profiles of dividing into Yea and Nay camps. Diplomatic insight no longer suffices to count how often a state is tipping the scales. Elaborate procedures are needed to carry out the counting, in general. Specifically, the Jagiellonian Compromise comes with a surprise: A state's share of power is proportional to its voting weight! The proof of the statement is difficult, but the statement itself makes life easy. All that is needed to determine the power distribution among the 28 Member States is to normalize the voting weights. The resulting *power share* $\beta_i := \sqrt{p_i}/\sum_{j\leq 28}\sqrt{p_j}$ of Member State i is shown in the last column of Table 12.6 (in percent). The power share β_i is also called its *normalized Banzhaf index* whence the use of the letter β.

The crux of the Jagiellonian Compromise is the quota formula. The formula is due to *Słomczyński / Życzkowski* (2006), two members of the Jagiellonian University Kraków, thus explaining the attribute "Jagiellonian". The system is a veritable "compromise" because it gently mediates between the decision power biases and other characteristics of several QMV systems currently in use or under discussion. The seminal monograph of *Felsenthal / Machover* (1998) teaches how to comprehensively analyze QMV systems. In particular it shows how to appraise a QMV system from the citizens' viewpoint.

The basic assumption is that decision-making is a repetitive business. Therefore, system indices must be evaluated by their likely values, not by a single realization. The citizens' influence whether a proposal is carried is modeled by a thought experiment in an idealized democracy. First a popular vote is taken, and then the government executes the majority's will. From an *a priori* viewpoint it appears constitutionally compelling to assume that citizens cast their votes independently of each other. The critical proposals are those that are supported or dismissed with the same likelihood, one-half. Of course the influence of a single citizen tends to zero, particularly in a state i where the population p_i is large. On the other hand a large population of voters gives rise to an almost infinite set of nip-and-tuck races where the last vote becomes decisive. Since a limit of the form $0 \times \infty$ has no meaning, a more sensitive analysis is called for. As the population grows, $p_i \to \infty$, an individual citizen's decision power decreases as $1/\sqrt{p_i}$. This is a consequence of the Central Limit Theorem. Let $X_n = 1$ indicate a Yea of citizen n, and $X_n = 0$ a Nay. The outcome is determined by the Yea total, $\sum_{n \le p_i} X_n$, a divergent sum. It needs to be scaled by $1/\sqrt{p_i}$ to be stabilized; this is the point where the square root makes its appearance. It can then be shown that if Member State i has a decision power β_i on the government level, the system conveys indirect power $\beta_i/\sqrt{p_i}$ to every citizen.

The discussion of the Jagiellonian Compromise now is quickly concluded. As mentioned above the government of Member State i has direct decision power $\beta_i = \sqrt{p_i}/\sum_{j \le 28} \sqrt{p_j}$. Since $\beta_i/\sqrt{p_i} = 1/\sum_{j \le 28} \sqrt{p_j}$ is the same constant for all Member States i, the citizens of all Member States share the same indirect decision power. This is the second surprise that comes with the Jagiellonian Compromise: All Union citizens have the same indirect decision power!

The final conclusion is exceedingly gratifying. The Jagiellonian Compromise and the Cambridge Compromise both boost the democratic foundations of the Union. The Jagiellonian Compromise guarantees that all Union citizens participate with provable equality in the decision-making processes of the European Council, even though participation is indirect through their governments. The Cambridge Compromise implements a dual concept of equal representation in the EP by achieving equality of the Member States' citizenries as well as equality of the Union's citizens. The next chapter turns to a multi-goal problem of a different type, the reconciliation of proportional representation and the election of persons.

CHAPTER 13

Proportionality and Personalization: BWG 2013

The 2013 amendment of the German Federal Election Law (Bundeswahlgesetz 2013, BWG 2013) is described by example. The Bundestag's seat apportionment obeys strict proportionality by political parties. All second votes in the country acquire practically equal success values. Thereafter a party's country-wide seats are allocated among its various state-lists of nominees in a way granting precedence to the direct seats won via first votes in single-seat constituencies. In this way the system combines proportional representation with the election of persons. The per-party sub-apportionment uses the direct-seat restricted variant of the divisor method with standard rounding. Its feasibility is ensured by an initial house size adjustment that usually raises the Bundestag size above the nominal level of 598 seats. The amended electoral system is exemplified with the September 2013 election of the 18th Bundestag.

13.1. THE 2013 AMENDMENT OF THE FEDERAL ELECTION LAW

The success of Germany's post-war Federal Election Law rests on its multi-purpose character. The law serves three goals, to achieve proportionality between political parties across the whole country, to proportionally divide a party's seats among the party's various state-lists of nominees, and to combine the system's proportional nature with plurality elections of individual candidates in single-seat constituencies.

Until 1980 the combination of proportional representation with the election of persons worked out alright because only two or three parties passed the five percent threshold and participated in the apportionment calculations. The diversification of the party system exacerbated the weaknesses of the old law, as reviewed in Sections 2.4 and 7.4. Eventually the Federal Election Law was amended in 2013. The amendment was carried by the broad consensus of four of the five political groups in the German Bundestag, with only the LINKE party abstaining.

The present chapter elucidates the new provisions with the election to the 18th German Bundestag on 22 September 2013. The new law introduces two novel elements. The first novelty is an initial calculation to adjust the size of the Bundestag. The adjustment ensures that the two components, proportional representation and the election of persons, can be combined successfully. In 2013 the Bundestag size is raised to 631 seats. Thereafter the super-apportionment apportions the $h = 631$ seats to the parties in proportion to their country-wide second votes (Section 13.2).

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_13,

18BT2013	Second Votes	Quotient	DivStd
Super-apportionment			
CDU	14 921 877	255.4	255
SPD	11 252 215	192.6	193
LINKE	3 755 699	64.3	64
GRÜNE	3 694 057	63.2	63
CSU	3 243 569	55.52	56
Sum (Divisor)	36 867 417	(58 420)	631

TABLE 13.1 *Super-apportionment of 631 seats by second votes, election to the 18th Bundestag 2013.* Every 58 420 votes justify roughly one seat. The seat apportionment is carried out using the divisor method with standard rounding, DivStd. All second votes acquire practically equal success values.

The second novelty concerns the per-party sub-apportionments. They do not use the unabridged divisor method with standard rounding, but its direct-seat restricted variant (Section 13.3). This innovation classifies as a novelty within Germany only. When adopting the German two-votes system other countries instantly rectified the notorious deficiencies of the system by imposing minimum restrictions, see the examples of Scotland and London in Section 12.4. The law's house size adjustment strategy is cumbersome and generous (Section 13.4). It is tempting to envision strategies that are less complicated and more parsimonious (Section 13.5).

13.2. APPORTIONMENT OF SEATS AMONG PARTIES

The amended Federal Election Law's very first action is to determine the definitive size of the Bundestag. The September 2013 results entail an adjustment to a level of 631 seats. This is the size how the Bundestag officiates during the 2013–2017 legislative period. The super-apportionment, that is, the apportionment of the 631 seats among the political parties, is explained in this section.

The apportionment is proportionate to the parties' country-wide totals of second votes. A party's vote total is the sum of second-votes over all sixteen German states, except that the CDU campaigns in only fifteen states (all but Bavaria) and that the CSU stands in just one state (Bavaria). The seat apportionment is carried out by means of the divisor method with standard rounding. Every 58 420 second votes justify roughly one seat, see Table 13.1.

The divisor method with standard rounding is distinguished by excellent properties guaranteeing that the second votes are dealt with as fairly as is practical. First, all voters' second votes enjoy a practically equal success value, in the sense of minimizing the squared-error deviations from the ideal one-hundred percent success (Section 10.2). Second, the seat apportionment is success-value stable since an attempted seat transfer never entails an improvement of the success-value disparities but worsens them or leaves them as is (Section 10.8). Third, all political parties receive their ideal share of seats as far as is practically possible, because the apportionment is ideal-share stable (Section 10.11). Fourth, the seat biases of all parties are zero. That is, on average the unavoidable rounding operations entail seat profits and seat deficits that are entirely random and not subject to a systematic trend (Theorem 7.7). Fifth, the verbalization "Scale all vote counts by a common divisor and then round interim quotients to whole numbers as in day-to-day life" arguably is the simplest instruction possible.

18BT2013 (*continued*)	Dir.	Second Votes	Quotient	DivStd•
Sub-apportionment to state-lists: CDU				
SH Schleswig-Holstein	9	638 756	10.7	11
MV Mecklenburg-Vorpommern	6	369 048	6.2	6
HH Hamburg	1	285 927	4.8	5
NI Niedersachsen	17	1 825 592	30.6	31
HB Bremen	0	96 459	1.6	2
BB Brandenburg	9	482 601	8.1•	9
SA Sachsen-Anhalt	9	485 781	8.1•	9
BE Berlin	5	508 643	8.52	9
NW Nordrhein-Westfalen	37	3 776 563	63.3	63
SN Sachsen	16	994 601	16.7	17
HE Hessen	17	1 232 994	20.7	21
TH Thüringen	9	477 283	8.0•	9
RP Rheinland-Pfalz	14	958 655	16.1	16
BY Bayern	—	—	—	—
BW Baden-Württemberg	38	2 576 606	43.2	43
SL Saarland	4	212 368	3.6	4
Sum (Divisor)	191	14 921 877	(59 700)	255

District	Dir.	Second Votes	Quotient	DivStd	Dir.	Second Votes	Quotient	DivStd
Sub-apportionment to state-lists: SPD					*LINKE*			
SH	2	513 725	8.8	9	0	84 177	1.4	1
MV	0	154 431	2.6	3	0	186 871	3.1	3
HH	5	288 902	4.9	5	0	78 296	1.3	1
NI	13	1 470 005	25.1	25	0	223 935	3.7	4
HB	2	117 204	2.0	2	0	33 284	0.6	1
BB	1	321 174	5.49	5	0	311 312	5.2	5
SA	0	214 731	3.7	4	0	282 319	4.7	5
BE	2	439 387	7.51	8	4	330 507	5.51	6
NW	27	3 028 282	51.8	52	0	582 925	9.7	10
SN	0	340 819	5.8	6	0	467 045	7.8	8
HE	5	906 906	15.503	16	0	188 654	3.1	3
TH	0	198 714	3.4	3	0	288 615	4.8	5
RP	1	608 910	10.4	10	0	120 338	2.0	2
BY	0	1 314 009	22.46	22	0	248 920	4.1	4
BW	0	1 160 424	19.8	20	0	272 456	4.54	5
SL	0	174 592	3.0	3	0	56 045	0.9	1
Sum	58	11 252 215	(58 500)	193	4	3 755 699	(60 000)	64

District	Dir.	Second Votes	Quotient	DivStd
Sub-apportionment to state-lists: GRÜNE				
SH	0	153 137	2.53	3
MV	0	37 716	0.6	1
HH	0	112 826	1.9	2
NI	0	391 901	6.47	6
HB	0	40 014	0.7	1
BB	0	65 182	1.1	1
SA	0	46 858	0.8	1
BE	1	220 737	3.6	4
NW	0	760 642	12.6	13
SN	0	113 916	1.9	2
HE	0	313 135	5.2	5
TH	0	60 511	1.0	1
RP	0	169 372	2.8	3
BY	0	552 818	9.1	9
BW	0	623 294	10.3	10
SL	0	31 998	0.53	1
Sum	1	3 694 057	(60 600)	63

TABLE 13.2 *Sub-apportionments of party-seats to state-lists, election to 18th Bundestag 2013.* The CDU direct seats (Dir.) overrule the interim quotient in three states (marked •), due to the direct-seat restricted variant of the divisor method with standard rounding. For the other parties the direct seats are supported by the proportionality seats whence the restrictions remain dormant.

13.3. ASSIGNMENT OF CANDIDATES TO SEATS

Since the CSU stands only in one state, Bavaria, its seat apportionment is completed instantly. The super-apportionment awards the party 56 seats (Table 13.1). Moreover the party wins 45 direct seats on the basis of first votes. Therefore, the CSU sends 45 constituency winners into the Bundestag, plus the eleven nominees on its Bavarian state-list that rank top apart from any constituency winners.

The other four parties stand in more states than just one. Again every direct-seat winner is declared elected due to the first-vote plurality victory in her or his constituency. These direct-seat wins impose a minimum restriction on the finalizing sub-apportionment of a party's country-wide seats. Therefore, the apportionment method used is not the simple divisor method with standard rounding (DivStd), but its direct-seat restricted variant (DivStd•). In every state the direct-seat winners get their seats, and the remaining seats are assigned to the top-ranked nominees on the state-list after removing the direct-seat winners. See Table 13.2.

The size adjustment procedure ensures that every per-party sub-apportionment gets sufficiently many seats to support all direct seats. In the 2013 election the direct-seat restrictions remain dormant in the three sub-apportionments for SPD, LINKE, and GRÜNE whose seat allocations coincide with those from the (unabridged) divisor method with standard rounding. Hence these seat allocations enjoy the method's excellent properties reviewed in the previous section. Note, however, that success-value optimality has two distinct meanings, previously in the super-apportionment, and now in the sub-apportionment. Previously, the super-apportionment evaluates second votes by party affiliation for which the principle of electoral equality is paramount. Now, the sub-apportionments re-evaluate second votes by federal provenance, whether they are cast for the (same) party in one state, or in another state.

The CDU sub-apportionment is distinguished by having to cope with direct-seat restrictions that are active. Every 59 700 second votes justify roughly one seat, except when a direct-seat restriction warrants more seats. In Table 13.2 the quotients that are superseded by direct-seat wins are marked with a dot (•). The unproportionality index (Section 12.4) between the direct-seat restricted apportionment, $x = \text{DivStd}\bullet$, and the unrestricted apportionment, $z = \text{DivStd}$ (not shown in the table), is $|x - z|/2 = 3$. In plain words, the direct-seat restrictions have the effect that three CDU seats are assigned to state-lists other than indicated by proportionality. The deviation from proportionality serves two goals. First it is instrumental to evade overhang seats and negative voting weights that formerly troubled the law. Second it allows to successfully combine the two characteristic components of the Bundestag electoral system, proportional representation and the election of persons.

13.4. INITIAL ADJUSTMENT OF THE BUNDESTAG SIZE

A central novelty is the initial adjustment of the Bundestag size to ensure that the system can accommodate the actual election results. Of course there are many ways to achieve this goal. The way chosen is cumbersome. It carries some of the burden that is inevitable when reaching a broad political consensus. The transient 2011 amendment introduced separate apportionments in each of the states. These separate per-state

18BT2013-Step 1	German Pop. 31.12.2012	Quotient	DivStd
Schleswig-Holstein	2 686 085	21.7	22
Mecklenburg-Vorpommern	1 585 032	12.8	13
Hamburg	1 559 655	12.6	13
Niedersachsen	7 354 892	59.3	59
Bremen	575 805	4.6	5
Brandenburg	2 418 267	19.49	19
Sachsen-Anhalt	2 247 673	18.1	18
Berlin	3 025 288	24.4	24
Nordrhein-Westfalen	15 895 182	128.1	128
Sachsen	4 005 278	32.3	32
Hessen	5 388 350	43.4	43
Thüringen	2 154 202	17.4	17
Rheinland-Pfalz	3 672 888	29.6	30
Bayern	11 353 264	91.52	92
Baden-Württemberg	9 482 902	76.4	76
Saarland	919 402	7.4	7
Sum (Divisor)	74 324 165	(124 050)	598

TABLE 13.3 *Step 1 of the house size adjustment calculations, election to the 18th Bundestag 2013.* The 598 nominal seats are allocated to the states by population figures as of 31 December 2012. Every 124 050 Germans justify roughly one seat.

apportionments resurface in the adjustment calculation not out of technical necessity, but more as a face-saving device for futile former proposals.

The adjustment of the Bundestag size proceeds in three steps. Step 1 allocates the 598 nominal Bundestag seats to the sixteen states on the basis of population figures. Step 2 evaluates each state separately by apportioning a state's seats among parties in proportion to second votes. The maximum of direct-seat wins and proportionality seats is earmarked as the target seat number for this party in this state. Step 3 raises the Bundestag size until every party gets at least as many seats as are called for by the party's total target seats.

Step 1: Allocation of nominal seats to states. The first step allocates the 598 nominal seats to the sixteen states on the basis of the German population of the end of the previous year, 31 December 2012. The population figures are published by the German Statistical Office usually during the summer. The divisor method with standard rounding is applied, see Table 13.3.

It may happen that a direct-seat winner is not affiliated with a party or that the affiliated party fails the five percent threshold. Such a seat is allotted to the winning candidate, and simultaneously deducted from the allocation of the state where the constituency is located. For example in 2002 two PDS candidates won their Berlin constituencies, but the PDS failed the five percent threshold. The allocation for Berlin would then have to be reduced from 24 to 22 seats, for use in the subsequent Step 2.

Step 2: Target seats. The second step consists of sixteen apportionment calculations, one for each state. The seats from Step 1 are apportioned among the state's parties by second votes using the divisor method with standard rounding. In Table 13.4 the applicable state divisors are collected in the CDU box; they are also used in the other boxes. For instance in Schleswig-Holstein, the 638 756 CDU votes are divided by 61 000. The interim quotient 10.47 is rounded to 10 seats. At this point the direct seat wins come into play. The better result of direct seats (9) and of proportionality seats (10) is earmarked as the target seat number, 10. For the CDU, the per-state target

18BT2013-Step 2			CDU			
	State divisor	Dir.	Second Votes	Quotient	DivStd	Target
SH	61 000	9	638 756	10.47	10	10
MV	60 000	6	369 048	6.2	6	6
HH	60 000	1	285 927	4.8	5	5
NI	66 000	17	1 825 592	27.7	28	28
HB	65 000	0	96 459	1.48	1	1
BB	60 000	9	482 601	8.0	8	9
SA	60 000	9	485 781	8.1	8	9
BE	62 000	5	508 643	8.2	8	8
NW	63 600	37	3 776 563	59.4	59	59
SN	61 000	16	994 601	16.3	16	16
HE	61 000	17	1 232 994	20.2	20	20
TH	60 000	9	477 283	8.0	8	9
RP	63 000	14	958 655	15.2	15	15
BY	58 300	—	—	—	—	—
BW	60 600	38	2 576 606	42.52	43	43
SL	67 000	4	212 368	3.2	3	4
Total target seats						242

	SPD					LINKE				
	Dir.	Sec. Votes	Quotient	DivStd	Target	Dir.	Sec. Votes	Quotient	DivStd	Target
SH	2	513 725	8.4	8	8	0	84 177	1.4	1	1
MV	0	154 431	2.6	3	3	0	186 871	3.1	3	3
HH	5	288 902	4.8	5	5	0	78 296	1.3	1	1
NI	13	1 470 005	22.3	22	22	0	223 935	3.4	3	3
HB	2	117 204	1.8	2	2	0	33 284	0.51	1	1
BB	1	321 174	5.4	5	5	0	311 312	5.2	5	5
SA	0	214 731	3.6	4	4	0	282 319	4.7	5	5
BE	2	439 387	7.1	7	7	4	330 507	5.3	5	5
NW	27	3 028 282	47.6	48	48	0	582 925	9.2	9	9
SN	0	340 819	5.6	6	6	0	467 045	7.7	8	8
HE	5	906 906	14.9	15	15	0	188 654	3.1	3	3
TH	0	198 714	3.3	3	3	0	288 615	4.8	5	5
RP	1	608 910	9.7	10	10	0	120 338	1.9	2	2
BY	0	1 314 009	22.54	23	23	0	248 920	4.3	4	4
BW	0	1 160 424	19.1	19	19	0	272 456	4.496	4	4
SL	0	174 592	2.6	3	3	0	56 045	0.8	1	1
Total target seats					183					60

	GRÜNE				
	Dir.	Sec. Votes	Quotient	DivStd	Target
SH	0	153 137	2.51	3	3
MV	0	37 716	0.6	1	1
HH	0	112 826	1.9	2	2
NI	0	391 901	5.9	6	6
HB	0	40 014	0.6	1	1
BB	0	65 182	1.1	1	1
SA	0	46 858	0.8	1	1
BE	1	220 737	3.6	4	4
NW	0	760 642	12.0	12	12
SN	0	113 916	1.9	2	2
HE	0	313 135	5.1	5	5
TH	0	60 511	1.0	1	1
RP	0	169 372	2.7	3	3
BY	0	552 818	9.48	9	9
BW	0	623 294	10.3	10	10
SL	0	31 998	0.48	0	0
Total target seats					61

TABLE 13.4 *Step 2 of the house size adjustment calculations, election to the 18th Bundestag 2013.* The evaluations operate per state, that is, rowwise. The state divisor that is shown in the top box is applied to this state throughout the other three boxes. The maximum of the direct seats (Dir.) and the proportionality seats (DivStd) yields the target seat numbers.

18BT2013-Step 3	Min.	Second Votes	Quotient	DivStd	Quotient	DivStd	Quotient	DivStd
CDU	242	14 921 877	241.8	242	255.3	255	255.4	255
SPD	183	11 252 215	182.4	182•	192.51	193	192.6	193
LINKE	60	3 755 699	60.9	61	64.3	64	64.3	64
GRÜNE	61	3 694 057	59.9	60•	63.2	63	63.2	63
CSU	56	3 243 569	52.6	53•	55.49	55•	55.52	56
Sum	602	36 867 417	(61 700)	598	(58 450)	630	(58 420)	631

TABLE 13.5 *Step 3 of the house size adjustment calculations, election to the 18th Bundestag 2013.* A party's total target seats constitute a minimum restriction (Min.). Up to house size 630 some minima are missed (•), thereafter they are met. Hence the definitive Bundestag size is 631 seats.

seats sum to 242 target seats altogether. For the SPD, FDP, LINKE and GRÜNE the total target seats are 183, 60, and 61 seats, respectively. The CSU is not shown in the table. Its interim quotient $3\,243\,569/58\,300 = 55.6$ leads to 56 proportionality seats. In view of 45 direct CSU seats, the earmarked target seat number is 56 seats.

Step 3: Adjustment of the Bundestag size. The final step interprets a party's country-wide target seats as the minimum number of seats to which the party is entitled. If these minima are reached or exceeded with the 598 nominal seats, then the nominal Bundestag size is also the final size. If not, the Bundestag size is raised. The first house size supporting all minima constitutes the definitive Bundestag size. The underlying apportionment method is the divisor method with standard rounding. Table 13.5 exhibits the last house size that fails, 630, and the first house size that complies, 631. The qualifiers "last" and "first" are to the point because of the house size monotonicity of divisor methods (Section 9.4). Now the definitive Bundestag size of 631 seats triggers the super-apportionment and the ensuing sub-apportionments as explicated in Sections 13.2 and 13.3.

13.5. ALTERNATIVE HOUSE SIZE ADJUSTMENT STRATEGIES

Step 2 reacts too sensitively to the non-uniform voter turnouts in the sixteen states. The differences are exacerbated by the heterogeneous distribution of ineffective votes. The two effects manifest themselves through the state divisors' variability in Table 13.4. The divisors extend from 58 300 in Bavaria to 67 000 in Saarland. That is, in Bavaria every 58 300 votes justify roughly one seat, in the Saarland every 67 000 votes. In Table 13.5 the initial federal divisor 61 700 is lowered to the final value 58 420 close to the Bavarian state divisor 58 300. The diminishment of the divisor forces an enlargement of the house size. However, the diminishment is due to the variability of the voter turnout and the uneven spread of ineffective votes. This is beside the point. The point is to reconcile proportional representation and the election of persons.

Section 12.5 compares various strategies (a)–(f) for the adjustment of the Bundestag size. With the 2013 data they produce the following results. In each instance the particular strategy accounts for the size of the Bundestag. These seats are apportioned among the parties, using the divisor method with standard rounding, and lead to the seat number for the CDU as listed. In order to accommodate the 191 direct seats of the CDU, the CDU sub-apportionment applies the direct-seat restricted variant of the divisor method with standard rounding. The unproportionality index measures the effect of the restrictions when active. It is the number of seats that are apportioned differently, due to active direct-seat restrictions, than what perfect proportionality would indicate (that is, use of the unabridged divisor method with standard rounding).

Strategy for the House Size Adjustment Step	Bundestag Size 2013	CDU Seats	Unproportionality Index, CDU Sub-apportionment
(a) Notional size	598	242	4
(b) Direct seats +10%	598	242	4
(c) Quota-based estimates	598	242	4
(d) BWG 2008 CDU seats	607	246	3
(e) BWG 2013 apportionment	631	255	3
(f) Zero unproportionality	661	268	0

It is evident that the complicated and laborious strategy (e) of the current law performs poorly. The unproportionality index is lowered from 4 to 3 seats only insignificantly. Here is a review how the strategies proceed.

Strategy (a) apportions the 598 nominal seats using the divisor method with standard rounding. The CDU receives 242 seats. The sub-apportionment of these seats must cope with the 191 direct seat wins of the CDU. Therefore the direct-seat restricted variant is applied. It allocates four seats differently than compared to perfect proportionality, whence its unproportionality index is 4 seats.

Strategy (b) makes sure that the CDU proportional allotment exceeds the number of direct seats by ten percent. That is, the house size must guarantee the CDU at least $191 + 19 = 210$ seats. Since this is achieved by the nominal house size of 598 seats, no further action is needed. The end result is the same as with strategy (a).

Strategy (c) replaces the separate per-state evaluations in Step 2 by applying a common divisor to all eighty state-lists of nominees. The divisor used is the votes-per-seats ratio, $36\,867\,417/598 = 61\,651.2$. Interim quotients are truncated to their integral part in order to obtain parsimonious seat estimates. With target seats taken to be the maximum of direct seats and estimated seats, the target seat total for the CDU turns out to be 241 seats. Again no further action is needed.

Strategy (d) argues that the CDU allotment with the old 2008 law would have consisted of 242 proportionality seats plus 4 overhang seats. Hence the Bundestag size is raised to 607 seats. Then the super-apportionment allocates 246 seats to the CDU.

Strategy (e) is the current 2013 law described above. Strategy (f) raises the Bundestag size to 661 seats, which would have enabled the defunct 2008 law to accommodate all direct seats into the proportionality calculations and to circumvent the generation of overhang seats.

Adjustment strategies affect the balance of constituency seats and list seats. With a nominal house size of 598 seats, there are 299 constituency seats and 299 list seats. If the house size is enlarged then more list seats are brought to life. Therefore, another option is to reduce the 299 constituencies to 275 say, and to adopt a nominal Bundestag size of 550 seats. Under present conditions the adjustment strategy in the current 2013 law promises a Bundestag size of about 600 seats. In 2002 the Bundestag appointed a reform committee to reduce the nominal house size from then 656 seats to below 600. It would be easy to amend the current law so that it realizes the 2002 reduction goal.

CHAPTER 14

Representing Districts and Parties: Double Proportionality

Double-proportional methods achieve fairness in two directions, the geographical division of the country and the political division of the electorate. Initially all seats are apportioned among districts proportionately to population figures and, independently, among parties proportionately to vote counts. The core of double proportionality is the finalizing sub-apportionment, the allotment of seats to districts and parties in a way that each district meets its district magnitude and each party exhausts its overall seats. Hence there are two sets of electoral keys, district divisors and party divisors. Once the divisors are published the sub-apportionment is verified quite easily.

14.1. THE 2012 PARLIAMENT ELECTION IN THE CANTON OF SCHAFFHAUSEN

When an electoral area is subdivided into regional districts the electoral system often is expected to honor the subdivision of the electoral region by geographical districts in a similar way as it honors the division of the electorate by political parties. Historically, the idea that a Member of Parliament represents a local district precedes the view that parliamentary representation provides a mirror image of the division along party lines, see Sections 1.10 and 2.1. Procedures that successfully meet the double challenge are double-proportional divisor methods. Because of the dual objective, the notational requirements and abstract analysis of double-proportional apportionment methods are more elaborate than those of simple-proportional apportionment methods.

It is instructive to begin with a concrete example, the September 2012 election of the parliament of the Swiss Canton of Schaffhausen. The Cantonal Parliament (Kantonsrat) comprises sixty seats. For its election the whole canton is subdivided into six districts. The seat apportionment is carried out using the double-proportional variant of the divisor method with standard rounding. We show how the sixty seats are allocated to the six districts, how the sixty seats are apportioned among the twelve parties that stood in the 2012 election, and how the double-proportional sub-apportionment works that finalizes the seat allotment.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8_14,

Prior allocation of seats to districts: District magnitudes. The number of seats allocated to a district is called the *district magnitude*. The district magnitudes for the 2012 election are obtained from the census figures as of 31 December 2010, using the Hare-quota method with residual fit by greatest remainders and guaranteeing all districts at least one seat. The seat guarantee holds without further ado, whence the quota method ambiguities from Section 12.2 play no role. See Table 14.1.

The minimum restricted variant of the divisor method with standard rounding yields the same allocation. Every 1 250 citizens justify roughly one seat. The district magnitudes range from twenty-eight seats in the largest district, the City of Schaffhausen, to a single seat in the smallest district, the exclave Buchberg-Rüdlingen. The smallest district is a single-seat constituency where formerly the representative was elected by plurality vote. Votes cast for candidates other than the constituency winner were wasted. On the other hand the magnitude of the largest district exceeds twice the number of parties, $28 > 2 \times 12 = 24$, whence the largest district obeys the house size recommendation of Section 7.9. However, the volatility of district magnitudes loses its importance when a double-proportional system is used.

In Schaffhausen, and in similar electoral systems, the district magnitude also fixes the number of candidates that may be marked on the ballot sheet. Since parties, candidates, and election officials need time to make appropriate preparations, the district magnitudes were publicized in January 2012 well ahead of the September election.

Super-apportionment of seats to parties: Overall party-seats. In every district the votes of all candidates of a party are aggregated into the *party votes* (Parteistimmen). Since a voter in the City of Schaffhausen may mark up to 28 candidates while a voter in Buchberg-Rüdlingen can mark only one, different districts yield party votes on different scales. However, electoral equality pertains to human beings, not to marks on the ballot sheets. Therefore party votes are converted into voter counts in the same way as in the Canton of Zurich (Section 4.8). First they are divided by the district magnitude, and then the resulting quotient is commercially rounded:

$$\text{voter count} = \left\langle \frac{\text{party votes}}{\text{district magnitude}} \right\rangle.$$

Party votes and district magnitudes are conveniently documented in Table 14.3. In the Schaffhausen district, 55 905 SVP party votes yield voter count 1 997 (since $55\,905/28 = 1\,996.6$). In Buchberg-Rüdlingen, the 309 SVP votes stay put (since $309/1 = 309$). The sum of the SVP voter counts over the six districts is 6 740 (Table 14.2).

The reference to voter counts ensures that the super-apportionment honors all voters equally, irrespective of the district where they reside. Since the seat apportionment is carried out using the divisor method with standard rounding, the voters' contributions to the political division of the cantonal parliament obey the principle of electoral equality in a best possible fashion. That is, the success values of the voters are as equal as is practically feasible (Section 10.2), the seat apportionment is success-value stable (Section 10.8) as well as ideal-share stable (Section 10.11), and the seat biases of all parties vanish identically (Theorem 7.7). The resulting seat number of a party is referred to as their *overall party-seats*. For example, the strongest party SVP is apportioned 16 overall party-seats, the weakest party JUSO one. See Table 14.2.

Sh2012DistrictMagn.	Population	Min.	Quotient	HaQgrR
Schaffhausen	34 943	1	27.458	28
Klettgau	15 453	1	12.143	12
Neuhausen	10 185	1	8.003	8
Reiat	8 986	1	7.061	7
Stein	5 222	1	4.103	4
Buchberg-Rüdlingen	1 567	1	1.231	1
Sum (Split)	76 356	6	(.3)	60

TABLE 14.1 *District magnitudes, Schaffhausen 2012.* The 60 seats are allocated to districts proportionately to the census figures as 31 December 2010 using the Hare-quota method with residual fit by greatest remainders. The minimum restriction of at least one seat per district remains dormant. The minimum restricted variant of the divisor method with standard rounding yields the same result.

Sh2012Super-app.	Voter count	Quotient	DivStd
SVP	6 740	16.1	16
SP	5 314	12.7	13
FDP	3 778	9.0	9
AL	1 886	4.51	5
ÖBS	1 878	4.49	4
CVP	1 232	2.9	3
JSVP	1 117	2.7	3
EDU	889	2.1	2
JFSH	827	2.0	2
SVP Sen.	618	1.48	1
EVP	551	1.3	1
JUSO	384	0.9	1
Sum (Divisor)	25 214	(418)	60

TABLE 14.2 *Super-apportionment, Schaffhausen 2012.* The determination of the overall party-seats is based on the parties' canton-wide totals of the per-district voter counts. Every 418 voters justify roughly one seat. Since the divisor method with standard rounding is used the resulting overall party-seats realize practically equal success values for all voters in the whole canton.

Sh2012Sub-app.		SVP	SP	FDP	AL	ÖBS	CVP	District divisor
		16	13	9	5	4	3	
Schaffhausen	28	55 905-5	70 837-6	46 656-4	34 800-4	27 243-2	12 596-1	10 700
Klettgau	12	23 901-4	11 871-2	11 980-2	2 802-1	3 431-1	2 350-0	5 400
Neuhausen	8	4 493-2	5 252-3	3 309-2	781-0	1 003-0	2 054-1	2 000
Reiat	7	8 749-2	4 380-1	3 493-1	968-0	2 087-1	443-0	3 100
Stein	4	2 519-2	1 681-1	464-0	301-0	782-0	1 064-1	1 400
Buchberg-Rüdlingen	1	309-1	92-0	85-0	98-0			400
Party divisor		1.16	1.05	1	0.9	1.2	1	

(*continued*)		JSVP	EDU	JFSH	SVP Sen.	EVP	JUSO	District divisor
		3	2	2	1	1	1	
Schaffhausen	28	8 214-1	9 204-1	11 126-1	5 031-1	7 178-1	5 617-1	10 700
Klettgau	12	5 650-1	3 952-1	1 336-0	1 348-0	3 006-0	917-0	5 400
Neuhausen	8	644-0	457-0	377-0	820-0	348-0	292-0	2 000
Reiat	7	1 241-1	936-0	1 106-1	1 033-0		318-0	3 100
Stein	4	201-0	159-0	202-0	149-0		100-0	1 400
Buchberg-Rüdlingen	1	45-0		63-0	38-0			400
Party divisor		0.8	1	0.7	0.9	1.2	1	

TABLE 14.3 *Sub-apportionment, Schaffhausen 2012.* The Schaffhausen SVP party votes (55 905) are divided by the Schaffhausen divisor (10 700) and SVP divisor (1.16). The resulting quotient 4.504 justifies 5 seats. The other seat numbers are obtained similarly. The published divisors guarantee that each district meets its district magnitude and that each party exhausts its overall party-seats.

Sub-apportionment: Joint allotment of seats to districts and parties. The final sub-apportionment consists of the allocation of all 60 seats to the lists of nominees presented to the electorate in the 6 districts by the 12 parties. The maximum number of potential lists would be $6 \times 12 = 72$. But some parties do not stand in some districts, and so only 65 lists materialize. The sub-apportionment delivers a joint allotment of seats to districts and parties aiming at three goals:

(1) Each district meets its district magnitude.

(2) Each party exhausts its overall party-seats.

(3) Proportionality is observed among parties within a given district as well as among districts within a given party.

The goals are achieved by the double-proportional variant of the divisor method with standard rounding. Two sets of divisors are needed. The first consists of a *district divisor* $C_i > 0$ for every district i. The second set contains a *party divisor* $D_j > 0$ for every party j. The divisors ensure that goals (1) and (2) are satisfied meticulously. Goal (3) is realized in that district divisors scale the party votes within a given district, while party divisors scale the party votes within a given party. The way how these divisors are determined is the theme of the subsequent sections.

Once the divisors are obtained and published it is rather easy to determine the seat numbers. The party votes v_{ij} that in district i are cast for party j are divided by the two associated divisors to obtain the interim quotient $v_{ij}/(C_i D_j)$. Standard rounding yields the *double-proportional seat number* x_{ij}. In the absence of ties we get

$$x_{ij} = \left\langle \frac{v_{ij}}{C_i D_j} \right\rangle. \qquad (\dagger)$$

Table 14.3 summarizes the double-proportional solution. The inner box shows party votes and seat numbers separated by a hyphen "-", for all districts i and for all parties j. Appreciation and verification of the solution is aided by the information arranged on the outside: district magnitudes on the left, overall party-seats along the top, district divisors on the right, and party divisors along the bottom.

For instance, in the first district the Schaffhausen SVP's 55 905 party votes are divided by the Schaffhausen divisor (10 700) and the SVP divisor (1.16). The resulting interim quotient, 4.504, is rounded to 5. The Schaffhausen list of the SVP is allocated five seats. Similarly the JUSO's 5 617 party votes in Schaffhausen are divided by the Schaffhausen divisor and the JUSO divisor (1). The interim quotient is 0.52 and justifies one seat. Across the whole table the seat numbers x_{ij} sum rowwise to the district magnitude and columnwise to the overall party-seats. The seat apportionment of the 2012 Cantonal Parliament election in the Canton of Schaffhausen is complete.

The merits of double proportionality vividly surface in the single-seat district Buchberg-Rüdlingen. Formerly the election was by plurality vote, for the last time in 2004. Voter turnout in 2004 amounted to $580/1\,068 = 54$ percent. The 2012 turnout of $730/1\,136 = 64$ percent is a significant increase of ten percentage points. The increased turnout are people who voted for somebody else than the prospective winner. Presumably these people have become aware that, though their candidates can hardly overturn the traditional winner, their votes nevertheless contribute to the canton-wide apportionment.

In fact the 98 AL votes in Buchberg-Rüdlingen are instrumental to secure a fifth seat. AL is canton-wide ahead of ÖBS by a narrow margin of eight votes only (Table 14.2). With 8 votes fewer in Buchberg-Rüdlingen AL would be tied with ÖBS, each with 1 878 votes, other things being equal. Hence 91 of the 98 AL voters are decisive. With the old plurality system they would have had no say whatsoever. All 421 non-winning votes would have been wasted and useless. Now the voters in the Buchberg-Rüdlingen district contribute to the canton-wide electoral success as much as do the voters in all other districts.

14.2. FROM THE EXAMPLE TO THE GENERAL SET-UP

The example of the 2012 Schaffhausen Cantonal Parliament election features many aspects of greater generality. We comment on some of them.

Organizing principles. The Schaffhausen example aptly illustrates which organizing principles (Section 4.2) extend to double-proportional divisor methods. Three do, two do not. Anonymity applies since districts may be permuted at will, and so may be parties. Decency holds true since in (†) a scaling of the numerator is instantly matched by a scaling in the denominator. Exactness is established as in Section 4.4.

Balancedness may fail,

$$v_{ij} = v_{mn} \qquad \not\Rightarrow \qquad \left| x_{ij} - x_{mn} \right| \leq 1.$$

A violation of balancedness is easily constructed from the Schaffhausen data by raising the Klettgau CVP party votes, 2 350, to the level of the Stein SVP party votes, 2 519. The super-apportionment in Table 14.2 and the divisors in Table 14.3 stay the same. The seat numbers, zero and two, violate balancedness.

Discordant seat assignments. Concordance may also fail,

$$v_{ij} > v_{mn} \qquad \not\Rightarrow \qquad x_{ij} \geq x_{mn}.$$

As an example we raise the CVP party votes in Klettgau by yet another vote to 2 520, or even to 4 668. Still they receive no seats, while the SVP in Stein gets two.

Discordant seat assignments also occur within a party. For a within-party comparison we switch to voter counts. The Klettgau SP voter count 989 gets two seats, the Neuhausen SP voter count 657 receives three. This discordance pair may be explained by relating voter counts to district totals. The Klettgau SP voter count 989 out of a Klettgau total of 6 045 is a share of 16 percent. The Neuhausen SP voter count 657 out of a Neuhausen total of 2 482 is 26 percent. If five seats were to be apportioned according to weights 16 and 26, the first participant would get two seats, the second three. From this viewpoint this discordance pair looks somewhat reasonable.

Discordance pairs within the same district are more conspicuous. In Klettgau, the 2 802 AL party votes produce one seat, the 3 006 EVP party votes yield none. Fortunately, discordant seat assignments are rare due to the excellent mirror image of votes and seats that results from double-proportionality. The few discordant seat assignments that possibly emerge in a practical instance are unstructured and unpredictable. They are caused by the externally prescribed district magnitudes and overall party-seats that restrict the feasibility range of double-proportional seat matrices in a queer way. Occasional discordances are unavoidable in general.

Winner-take-one modification. Specifically, in single-seat districts, concordance can be rescued. We recommend to do so. In fact, in a single-seat district discordance is truly perplexing. It means that the only seat is allocated not to the strongest party, but to the runner-up or another weaker party. Luckily this does not happen in the 2012 Schaffhausen election. Nevertheless a discordance accident is easily manufactured by raising the AL votes in Buchberg-Rüdlingen from 98 to 262. In the super-apportionment of Table 14.2 the overall party-seats remain unaffected. But double-proportionality would award the sole Buchberg-Rüdlingen seat to the 262 AL voters, not to the 309 SVP voters. The assignment would invert the former plurality vote and thus irritate the public, press, and parties.

In single-seat districts concordance may be forced by way of the following *winner-take-one* modification: *In every district the strongest party is allocated at least one seat.* In recognition of the principle of electoral equality the wording addresses every district, whether large or small. The ensuing seat guarantees in the sub-apportionment have repercussions on the super-apportionment. Every party must receive at least as many seats as this party accumulates by means of the seat guarantees. The induced minimum restrictions in the super-apportionment are of no real worry though. Since they concern stronger parties, in practice they are fulfilled automatically.

In the 2012 Schaffhausen election the strongest party in Klettgau, Reiat, Stein, and Buchberg-Rüdlingen is the SVP, in Schaffhausen and Neuhausen the SP. Hence the super-apportionment must impose the minimum restrictions of at least four seats for the SVP, and at least two seats for the SP. They are visibly fulfilled since the SVP actually is apportioned 16 seats and the SP 13 (Table 14.2). In the sub-apportionment, larger districts automatically assign a seat to the strongest party and hence are unaffected by the winner-take-one modification. However, in Buchberg-Rüdlingen the modification matters. It makes sure that the single Buchberg-Rüdlingen seat goes to the 309 SVP voters even when the number of AL voters increases to 262 as in the previous paragraph, or even to 308, other things being equal.

The Schaffhausen electoral law omits the winner-take-one modification. We believe that the modification is always recommendable when some of the district magnitudes are as low as one or two seats. The extensive simulation studies of *Maier* (2009 [109]) strongly support the recommendation. In single-seat districts the winner-take-one modification perpetuates the merits of the former plurality vote, and overcomes its demerits of wasting all minority votes.

General set-up. Applicability of the winner-take-one modification is not bound to the divisor method with standard rounding that Schaffhausen uses. The underlying divisor method may be quite arbitrary. It may be distinct from the methods employed in the prior allocation of seats to districts, or in the super-apportionment of seats to parties. The two precursory steps precede the double-proportional final solely to fix the marginals, the district magnitudes and the overall party-seats. We designate the district magnitude of district i by r_i, and the overall party-seats of party j by s_j. Generally assuming that there are k districts and ℓ parties the seat numbers x_{ij} are assembled into a $k \times \ell$ seat matrix $x = (\!(x_{ij})\!)$. As usual the rounding rule that belongs to the underlying divisor method is denoted by $[\![\cdot]\!]$ (Section 4.4).

Definition. *A seat matrix x is called a* double-proportional seat apportionment *when there exist positive row divisors $C_1, \dots, C_k > 0$ and positive column divisors $D_1, \dots, D_\ell > 0$ such that for all rows $i \le k$ and columns $j \le \ell$ we have*

$$x_{ij} \in \left[\!\!\left[\frac{v_{ij}}{C_i D_j} \right]\!\!\right], \qquad x_{i+} := \sum\nolimits_{j \le \ell} x_{ij} = r_i, \qquad x_{+j} := \sum\nolimits_{i \le k} x_{ij} = s_j. \qquad (\ddagger)$$

District magnitudes and overall party-seats provide the *row marginals* r_i and *column marginals* s_j (German: Spaltenmarginalien) that must be attained by the row and column sums of a seat matrix x. The seat matrix x remains the same whether the original weights v_{ij} are scaled, or the row-normalized weights v_{ij}/v_{i+}, or the column-normalized weights v_{ij}/v_{+j}, or the overall-normalized weights v_{ij}/v_{++}, or the unrounded voter counts v_{ij}/r_i. This is easily seen by adjusting row or column divisors appropriately. The definition also tolerates a truncated rounding rule $[\![\cdot]\!]_{a_{ij}}^{b_{ij}}$ with minimum restrictions a_{ij} and maximum restrictions b_{ij} (Section 12.3). For instance the winner-take-one modification is accommodated by setting the minimum restrictions equal to unity when in district i party j is strongest, $a_{ij} = 1$, and otherwise $a_{ij} = 0$.

Since it is irrelevant how the prespecified marginals are calculated, we no longer distinguish between voter counts and party votes. From now on we address the quantities v_{ij} as vote weights, or simply as *weights.*

Definition. *A* weight matrix *v is defined to be a nonnegative $k \times \ell$ matrix with no row nor column vanishing,*

$$v = ((v_{ij})) \in [0; \infty)^{k \times \ell}, \qquad v_{i+} > 0 \textit{ for all } i \le k, \qquad v_{+j} > 0 \textit{ for all } j \le \ell.$$

The definition admits vanishing weights, $v_{ij} = 0$, but excludes the nonsensical constellation that a whole row or column consists of zeros. The admission of vanishing weights is a new feature of double-proportional problems that is structurally innocuous but technically demanding. By contrast, simple-proportional apportionment problems assume all weights to be positive (Section 4.1). The anonymity principle allows to move vanishing weights to the end of the vote vector and forget about them. This simplification is no longer feasible in double-proportional problems. It often happens that some party j does not stand in some district i, entailing a vanishing vote weight $v_{ij} = 0$. In the Schaffhausen example the weight matrix features seven structural zeros; in Table 14.3 they show up as empty cells.

Uniqueness. The most important technical question is whether a double-proportional seat apportionment is unique. The answer of the upcoming Uniqueness Theorem 14.3 is *Yes, except for ties.* The exception of ties is harmless. Ties are theoretically challenging but practically negligible. The Schaffhausen 2012 election is tie-free, as is every other empirical data set we know of. Hence the double-proportional apportionment in Table 14.3 is unique and final, and not up for discussion. There is no reason to worry when somebody publishes other district and party divisors, or when programming experts concoct lines of code they fancy will benefit the parties of their choice, or when the computer code is flawed and erroneous. The only test to pass is that the seat numbers and the published divisors satisfy the defining relations (‡).

Uniqueness does not hold for the district divisors C_i and the party divisors D_j as we know from Section 4.7. The presence of two sets of divisors introduces yet another degree of freedom. Evidently the scaled divisors $\widetilde{C}_i = \alpha C_i$ and $\widetilde{D}_j = D_j/\alpha$ would do the job, too, whatever the scaling constant $\alpha > 0$.

The new degree of freedom serves to embellish the final divisors C_i and D_j from (‡) into user-friendly divisors $\widetilde{C}_i$ and $\widetilde{D}_j$ apt for publication. First, the scaling constant is set equal to the median of the final party divisors, $D_0 := \text{med}(D_1, \ldots, D_\ell)$. If ℓ is even, we choose the lower median. Second, with weights $\widehat{v}_{ij} := v_{ij}/(C_i D_0)$ every party is treated separately. Its s_j seats are allocated to the districts $1, \ldots, k$ in proportion to $(\widehat{v}_{1j}, \ldots, \widehat{v}_{kj})$. The user-friendly divisor from Section 4.6 is selected for publication, $\widetilde{D}_j$. Third, with weights $\widetilde{v}_{ij} = v_{ij}/\widetilde{D}_j$ every district is treated separately. Its r_i seats are apportioned to parties $1, \ldots, \ell$ according to $(\widetilde{v}_{i1}, \ldots, \widetilde{v}_{i\ell})$. The user-friendly divisor of Section 4.6 is selected for publication, $\widetilde{C}_i$.

As a result, one or more of the published party divisors equal unity and the others vary above and below. In Table 14.3 four party divisors are unity, four lie above and four below. The district divisors are of a size close to what would emerge if every district were apportioned separately on its own.

Existence. The other technical question is that of existence. Can we be sure that a set of double-proportional seat numbers satisfying (‡) exists? The answer is in the affirmative. Existence is implied by the Optimality Theorem 14.5 below, which builds on the Critical Inequalities 14.4. The inequalities characterize the solutions of a double-proportional divisor method in the same way as the Max-Min Inequality 4.5 characterizes the solutions of a simple-proportional divisor method. While reassuring, the abstract existence statement is dispensable in concrete instances. It plainly suffices to publish district and party divisors that do the job. Table 14.3 is an example.

Algorithms. Even when we know that there exists a double-proportional seat apportionment, there still remains the task of finding it. The task boils down to calculating district divisors $C_1, \ldots, C_k$ and party divisors $D_1, \ldots, D_\ell$ that give rise to a seat matrix x satisfying the defining relations (‡). The task is resolved by the Alternating Scaling algorithm, or by the Tie-and-Transfer algorithm. For all practical purposes the Alternating Scaling algorithm is generally fast, but it may stall in singular instances when the scaling leads to a weight matrix exhibiting many ties in a weird pattern. The Tie-and-Transfer algorithm is generally slow but always safe. The algorithms and their ramifications are to be discussed in Sections 14.7 and 14.8.

The functioning of the Alternating Scaling algorithm is easy to understand. The basic ingredient is a simple-proportional divisor method. The idea is to repeat the method many times until a solution is reached. In the first step all districts are handled separately. Such a step is called a *scaling of rows*. Then row marginals are met, column marginals possibly not. In the second step the original weights are re-scaled with the previous district divisors, and then all parties are handled separately. This step is called a *scaling of columns*. Column marginals are met, row marginals possibly not. In the third step the previous weights are re-scaled using the previous party divisors, and again all districts are handled on their own. This constitutes another scaling of rows, to be followed by another scaling of columns, and so on. The alternation terminates as soon as row and column marginals are met simultaneously. Upon termination the then current seat matrix is the desired double-proportional seat apportionment.

The Schaffhausen example in Table 14.3 requires five scalings of rows and columns. The three odd scalings 1, 3, 5 concern rows and call for $3 \times 6 = 18$ applications of the simple-proportional divisor method with standard rounding. Scalings 2 and 4 treat columns and need $2 \times 12 = 24$ simple-proportional apportionments. Altogether the

Alternating Scaling algorithm applies the simple-proportional divisor method $18+24 = 42$ times. It terminates with the double-proportional seat matrix that is displayed in Table 14.3. The final divisors are embellished for publication as outlined above.

Proposed first by *Balinski / Demange* (1989a, 1989b) double-proportional divisor methods are a rather recent addition to proportional representation methodology. The quest that parliamentary representation simultaneously merges the geographical and political divisions of the electorate is much older of course. The two most popular approaches to preserving the regional and political dimensions grant one dimension precedence over the other. These approaches are rudimentary applications of the Alternating Scaling algorithm by instantly terminating after the first step.

The Alternating Scaling algorithm as introduced above starts with a scaling of rows. Termination after the first step then means that the district magnitudes are apportioned in each district separately. No attention is paid to overall party-seats. Party divisors remain on their initialization level, $D_j = 1$, and disappear from the defining relations (‡). The electoral area is perceived as an assembly of its districts, the geographical subdivision predominates. This is the system used in the Canton of Schaffhausen before the 2008 adoption of the double-proportional divisor method. The system is still in use in other Swiss cantons. In the EP elections separate district apportionments are applied in Belgium, France and Ireland, see Table 1.31.

Because of the symmetry between rows and columns, the Alternating Scaling algorithm could also be specified by starting with a scaling of columns. Termination after the first step then means that the overall party-seats are apportioned among the districts, separately for each party. No attention is paid to the per-district seat totals. District divisors remain on their initialization level, $C_i = 1$, and disappear from the defining relations (‡). The predominant goal is to mirror the political division of the country. The per-party sub-apportionments are familiar from the German Bundestag elections. They are also employed in the EP election in Poland, see Table 1.23.

It is a political decision whether precedence is given to districts and their district magnitudes, or to parties and their overall party-seats, or whether a third option is preferred that strikes a balance between districts and parties. Double-proportional divisor methods offer an attractive way to realize the third option. The improved balance is paid for by a few more calculations than are usually called for. Since the additional labor is executed by computer, the actual price is nil.

In the remainder of the chapter we embark on a more detailed investigation of the technical questions raised above: uniqueness, existence, and algorithms.

14.3. UNIQUENESS OF A DOUBLE-PROPORTIONAL SEAT APPORTIONMENT

For the general analysis the house size continues to be h. The electoral area is supposed to be subdivided into k districts. The number of parties participating in the apportionment calculations again is ℓ. Double-proportionality makes sense only when there are at least two districts and two parties, $k \geq 2$ and $\ell \geq 2$. The row marginals $r = (r_1, \ldots, r_k)$ assemble the district magnitudes. The column marginals $s = (s_1, \ldots, s_\ell)$ signify the overall party-seats. Both vectors are assumed to consist of positive integers that sum to the preordained house size, $r_+ = h = s_+$.

Helpful further notions are the *set of feasible seat matrices* $\mathbb{N}^{k\times\ell}(r,s)$, and the *solution set* of double-proportional seat apportionments $A(r,s;v)$, defined through

$$\mathbb{N}^{k\times\ell}(r,s) := \left\{x \in \mathbb{N}^{k\times\ell} \,\middle|\, x_{i+} = r_i,\ i \le k, \text{ and } x_{+j} = s_j,\ j \le \ell\right\},$$
$$A(r,s;v) := \left\{x \in \mathbb{N}^{k\times\ell}(r,s) \,\middle|\, x_{ij} \in \left[\!\!\left[\frac{v_{ij}}{C_iD_j}\right]\!\!\right],\ i \le k,\ j \le \ell, \right.$$
$$\left. \text{for some } C_1,\ldots,C_k,D_1,\ldots,D_\ell > 0\right\}.$$

The set of feasible seat matrices comprises the $k \times \ell$ integer matrices whose rows are summing to the row marginals r and whose columns are summing to the column marginals s. The solution set $A(r,s;v)$ additionally depends on a given weight matrix $v \in [0;\infty)^{k\times\ell}$. It contains the feasible seat matrices x that result from a scaling of the weights v_{ij} with row divisors C_i and column divisors D_j, and then rounding the interim quotients $v_{ij}/(C_iD_j)$ to the seat numbers x_{ij}. The underlying rounding rule $[\![\cdot]\!]$ is assumed to rely on the signpost sequence $s(n)$, $n \in \mathbb{N}$ (Section 3.10).

The next theorem states a sufficient condition for a solution matrix x to be unique, $A(r,s;v) = \{x\}$. The proof makes use of the notion of a cycle. In a general $k \times \ell$ array, a *cycle via* $(i_1,\ldots,i_q)$ *and* $(j_1,\ldots,j_q)$ is defined to be a succession of cells where a move in columns $j_1,\ldots,j_q$ alternates with a move in rows $i_1,\ldots,i_q$,

$$(i_1,j_1),\ (i_2,j_1),\ (i_2,j_2),\ (i_3,j_2),\ \ldots,\ (i_{q-1},j_{q-1}),\ (i_q,j_{q-1}),\ (i_q,j_q),\ (i_1,j_q).$$

The $q \ge 2$ rows are taken to be distinct, and so are the columns. Hence every row or column is visited at most once.

Uniqueness Theorem. *Consider a double-proportional seat apportionment x in the solution set $A(r,s;v)$. If at most three seat numbers x_{ij} have interim quotients $v_{ij}/(C_iD_j)$ hitting positive signposts, then the solution x is unique,*

$$A(r,s;v) = \{x\}.$$

Proof. The proof is by contraposition. Let us assume that the set $A(r,s;v)$ contains two distinct matrices $x \ne y$. The difference $z := y - x$ is non-vanishing, but its rows and columns sum to zero. We select $q \ge 2$ distinct rows $i_1,\ldots,i_q$ and columns $j_1,\ldots,j_q$ so that along the cell cycle

$$(i_1,j_1),\ (i_2,j_1),\ (i_2,j_2),\ (i_3,j_2),\ \ldots,\ (i_{q-1},j_{q-1}),\ (i_q,j_{q-1}),\ (i_q,j_q),\ (i_1,j_q)$$

the entries of z alternate in sign, $-,+,\ldots$, as follows. We start with a cell (i_1,j_1) where $z_{i_1j_1} < 0$. In column j_1 we pick a cell (i_2,j_1) with $z_{i_2j_1} > 0$. Next we search in row i_2 a cell (i_2,j_2) where $z_{i_2j_2} < 0$. Then we look for a row i_3 such that $z_{i_3j_2} > 0$, etc. We finish when meeting a row or column already visited. The succession of cells then consists of a cycle that is possibly preceded by an initial section. Discarding the initial section the remaining cycle is preserved and relabeled. We select this cycle provided its first cell has $z_{i_1j_1} < 0$. Otherwise its last cell satisfies $z_{i_1j_q} < 0$; then we select the cycle obtained from reversing the sequences of row and column indices. Along the selected cycle, z has sign pattern $-,+,\ldots$ as desired. All vote weights along the cycle are positive since otherwise $v_{ij} = 0$ implies $x_{ij} = y_{ij} = 0$ and $z_{ij} = 0$.

Denote the divisors of x by C_i and D_j, and those of y by E_i and F_j. The fundamental relation (Section 3.8) turns the roundings $x_{ij} \in [\![v_{ij}/(C_iD_j)]\!]$ and $y_{ij} \in [\![v_{ij}/(E_iF_j)]\!]$ into the inequalities

$$\frac{s(x_{ij})}{v_{ij}} \le \frac{1}{C_iD_j} \le \frac{s(x_{ij}+1)}{v_{ij}} \qquad \text{and} \qquad \frac{s(y_{ij})}{v_{ij}} \le \frac{1}{E_iF_j} \le \frac{s(y_{ij}+1)}{v_{ij}}.$$

Along the selected cycle we form two strings of product inequalities, setting $i_{q+1} := i_1$,

$$\prod_{p\le q} \frac{s(x_{i_pj_p})}{v_{i_pj_p}} \le \prod_{p\le q} \frac{1}{C_{i_p}D_{j_p}} = \prod_{p\le q} \frac{1}{C_{i_{p+1}}D_{j_p}} \le \prod_{p\le q} \frac{s(x_{i_{p+1}j_p}+1)}{v_{i_{p+1}j_p}}, \tag{1}$$

$$\prod_{p\le q} \frac{s(y_{i_{p+1}j_p})}{v_{i_{p+1}j_p}} \le \prod_{p\le q} \frac{1}{E_{i_{p+1}}F_{j_p}} = \prod_{p\le q} \frac{1}{E_{i_p}F_{j_p}} \le \prod_{p\le q} \frac{s(y_{i_pj_p}+1)}{v_{i_pj_p}}. \tag{2}$$

The sign pattern of z yields $y_{i_pj_p} < x_{i_pj_p}$ and $y_{i_{p+1}j_p} > x_{i_{p+1}j_p}$, that is, $y_{i_pj_p} + 1 \le x_{i_pj_p}$ and $x_{i_{p+1}j_p} + 1 \le y_{i_{p+1}j_p}$. Due to signpost monotonicity the inequalities (1) and (2) intertwine,

$$\prod_{p\le q} \frac{s(x_{i_pj_p})}{v_{i_pj_p}} \le \prod_{p\le q} \frac{s(x_{i_{p+1}j_p}+1)}{v_{i_{p+1}j_p}} \le \prod_{p\le q} \frac{s(y_{i_{p+1}j_p})}{v_{i_{p+1}j_p}} \le \prod_{p\le q} \frac{s(y_{i_pj_p}+1)}{v_{i_pj_p}} \le \prod_{p\le q} \frac{s(x_{i_pj_p})}{v_{i_pj_p}}.$$

Hence equality holds throughout inequality (1). Because of $\prod_{p\le q} s(x_{i_{p+1}j_p}+1)/v_{i_{p+1}j_p} > 0$ the common value is positive. This entails equality in all factor inequalities and gives

$$s(x_{i_pj_p}) = \frac{v_{i_pj_p}}{C_{i_p}D_{j_p}} > 0 \qquad \text{and} \qquad s(x_{i_{p+1}j_p}+1) = \frac{v_{i_{p+1}j_p}}{C_{i_{p+1}}D_{j_p}} > 0.$$

Since every cycle visits at least two rows and two columns, $q \ge 2$, four or more interim quotients that are associated with the seat matrix x are tied to positive signposts. □

An equivalent formulation says that non-uniqueness forces at least four interim quotients to be tied to positive signposts. This is why the statement is often paraphrased by saying that the solution is “unique up to ties”.

A mere abundance of ties is not enough though. The ties must form a pattern that includes a cycle alternating between decrement options and increment options,

$$[\![s(x_{i_pj_p})]\!] = \{x_{i_pj_p} - 1, x_{i_pj_p}\} = x_{i_pj_p}-,$$
$$[\![s(x_{i_{p+1}j_p}+1)]\!] = \{x_{i_{p+1}j_p}, x_{i_{p+1}j_p}+1\} = x_{i_{p+1}j_p}+ .$$

The trailing plus- and minus-signs indicate the rounding options (Section 4.7), they are congruent with the sign pattern of the matrix z in the proof. It is extremely unlikely that many ties arise and, in addition, exhibit such a particular pattern. Hence double-proportional seat apportionments are unique, for all practical purposes.

14.4. CRITICAL INEQUALITIES

Generally we are interested only in cycles shunning zero entries of the weight matrix v. These cycles deserve a distinctive name.

Definition. *Given a weight matrix* $v \in [0;\infty)^{k\times\ell}$, *a* v-cycle *is defined to be a cycle via* $(i_1,\dots,i_q)$ *and* $(j_1,\dots,j_q)$ *such that in every cell the weight is positive,* $v_{i_pj_p} > 0$ *and* $v_{i_{p+1}j_p} > 0$ *for all* $p \le q$*, where* $i_{q+1} := i_1$.

As an illustration we consider a 3×3 matrix with diagonal weights zero and off-diagonal weights positive,

$$v = \begin{pmatrix} 0 & v_{12} & v_{13} \\ v_{21} & 0 & v_{23} \\ v_{31} & v_{32} & 0 \end{pmatrix}.$$

No cycle via two rows and columns is a v-cycle since it involves a zero of the diagonal. Yet the cycle via $(2,3,1)$ and $(1,2,3)$ is a v-cycle calling on the positive weights $v_{21}, v_{31}, v_{32}, v_{12}, v_{13}, v_{23}$.

The next theorem is the double-proportional analogue of the Max-Min Inequality 4.5. It substitutes the Max-Min Inequality by the ensemble of *critical inequalities*

$$\prod_{p \leq q} \frac{s(x_{i_p j_p})}{v_{i_p j_p}} \leq \prod_{p \leq q} \frac{s(x_{i_{p+1} j_p} + 1)}{v_{i_{p+1} j_p}}$$

that are induced by v-cycles. The restriction to v-cycles ascertains that all denominators are positive, $v_{i_p j_p} > 0$ and $v_{i_{p+1} j_p} > 0$. However, double proportional settings permit scenarios with vanishing weights, $v_{ij} = 0$. The consequences are contingent on whether the underlying divisor method is pervious ($s(1) > 0$) or impervious ($s(1) = 0$). In case of perviousness the no input–no output law is imposed, $v_{ij} = 0 \Rightarrow x_{ij} = 0$. In case of imperviousness we adjoin the *no output–no input law*, $x_{ij} = 0 \Rightarrow v_{ij} = 0$. The two laws collapse into an equivalence, $v_{ij} = 0 \Leftrightarrow x_{ij} = 0$, that is, $v_{ij} > 0 \Leftrightarrow x_{ij} > 0$. We merge the two scenarios into the single notion of "preserving the zeros of v".

Definition. *Given a weight matrix $v \in [0;\infty)^{k\times\ell}$, a seat matrix $x \in \mathbb{N}^{k\times\ell}$ is said to* preserve the zeros of v *when x obeys the no input–no output law and, in the presence of an impervious rounding rule, when x also obeys the no output–no input law.*

Critical Inequalities Theorem. *Let the double-proportional divisor method be induced by the rounding rule with signpost sequence $s(n)$, $n \in \mathbb{N}$. Then a seat matrix $x \in \mathbb{N}^{k\times\ell}(r,s)$ belongs to the solution set for a weight matrix $v \in [0;\infty)^{k\times\ell}$,*

$$x \in A(r,s;v),$$

if and only if x preserves the zeros of v and for every v-cycle via $(i_1,\ldots,i_q)$ and $(j_1,\ldots,j_q)$ the matrix x fulfills the inequality

$$\prod_{p \leq q} \frac{s(x_{i_p j_p})}{v_{i_p j_p}} \leq \prod_{p \leq q} \frac{s(x_{i_{p+1} j_p} + 1)}{v_{i_{p+1} j_p}}.$$

Proof. The direct implication follows from (1) in the proof of the Uniqueness Theorem 14.3.

The converse implication is more demanding. We need to find divisors C_i and D_j that satisfy $x_{ij} \in [\![v_{ij}/(C_i D_j)]\!]$ or, equivalently, $s(x_{ij}) \leq v_{ij}/(C_i D_j) \leq s(x_{ij}+1)$. The case $v_{ij} = 0$ instantly verifies $s(0) \leq 0 \leq s(1)$. In case $v_{ij} > 0$ we divide by v_{ij} and take logarithms,

$$\log \frac{s(x_{ij})}{v_{ij}} \leq \alpha_i + \beta_j \leq \log \frac{s(x_{ij}+1)}{v_{ij}}.$$

On the left-hand side we set $\log(0/v_{ij}) = -\infty$ when $s(x_{ij}) = 0$. The existence of $\alpha_i := -\log C_i$ and $\beta_j := -\log D_j$ is at issue.

For a concise formulation we pass to the linear space $\mathbb{R}^{k\times\ell}$ of real $k\times\ell$ matrices. Let the linear subspace $L\subset\mathbb{R}^{k\times\ell}$ consist of the matrices a having entries $a_{ij}=\alpha_i+\beta_j$ for some $\alpha_i,\beta_j\in\mathbb{R}$. Let the generalized rectangle $R\subset\mathbb{R}^{k\times\ell}$ be the Cartesian product of the intervals

$$I_{ij}:=\begin{cases}\left[\log\dfrac{s(x_{ij})}{v_{ij}};\log\dfrac{s(x_{ij}+1)}{v_{ij}}\right] & \text{in case } v_{ij}>0 \text{ and } s(x_{ij})>0, \qquad \text{(I-1)}\\[2ex] \left(-\infty;\log\dfrac{s\,(x_{ij}+1)}{v_{ij}}\right] & \text{in case } v_{ij}>0 \text{ and } s(x_{ij})=0, \qquad \text{(I-2)}\\[2ex] (-\infty;\infty) & \text{in case } v_{ij}=0. \qquad \text{(I-3)}\end{cases}$$

The intervals are nonempty. This is evident in (I-1) and (I-3). In (I-2) it follows from $s(x_{ij}+1)/v_{ij}>0$ since x preserves the zeros of v. Hence the rectangle R is nonempty, too.

There are two possibilities. Either the sets L and R intersect. If so, there are numbers α_i and β_j fulfilling $\alpha_i+\beta_j\in I_{ij}$. The divisors $C_i:=e^{-\alpha_i}$ and $D_j:=e^{-\beta_j}$ yield the assertion.

Or the sets L and R are disjoint. But disjointness contradicts the validity of the critical inequalities. To see this, we investigate the orthogonal complement $L^\perp$. Orthogonality refers to the Euclidean inner product $\langle a,b\rangle=\operatorname{trace}a'b=\sum_{i\le k}\sum_{j\le\ell}a_{ij}b_{ij}$ in the space $\mathbb{R}^{k\times\ell}$. Denoting by $\mathbb{1}_\ell:=(1,\ldots,1)'\in\mathbb{R}^\ell$ the *unity vector* in $\mathbb{R}^\ell$, the original space is $L=\{\alpha\mathbb{1}_\ell'+\mathbb{1}_k\beta' \mid \alpha\in R^k,\beta\in\mathbb{R}^\ell\}$. The inner product with a matrix $b\in\mathbb{R}^{k\times\ell}$ turns into two inner products, in $\mathbb{R}^k$ and $\mathbb{R}^\ell$, namely $\langle\alpha\mathbb{1}_\ell'+\mathbb{1}_k\beta',b\rangle=\operatorname{trace}\mathbb{1}_\ell\alpha'b+\operatorname{trace}\beta\mathbb{1}_k'b=\langle\alpha,b\mathbb{1}_\ell\rangle+\langle b'\mathbb{1}_k,\beta\rangle$. These vanish for all $\alpha\in\mathbb{R}^k$ and $\beta\in\mathbb{R}^\ell$ if and only if $b\mathbb{1}_\ell=0$ and $b'\mathbb{1}_k=0$. Thus the subspace $L^\perp$ consists of the matrices b whose rows and columns sum to zero, $L^\perp=\{b\in\mathbb{R}^{k\times\ell}\mid b\mathbb{1}_\ell=0,b'\mathbb{1}_k=0\}$.

Such matrices loom behind the Uniqueness Theorem 14.3. For a cycle via $i_{(q)}:=(i_1,\ldots,i_q)$ and $j_{(q)}:=(j_1,\ldots,j_q)$, the *cycle matrix* $c(i_{(q)},j_{(q)})$ is defined to be the $k\times\ell$ matrix with entry -1 in cell (i_p,j_p) and $+1$ in cell (i_{p+1},j_p) for all $p\le q$, and zeros elsewhere. The negative of a cycle matrix is the cycle matrix belonging to the *reverse cycle*, that is, the cycle via $(i_1,i_q,i_{q-1},\ldots,i_4,i_3,i_2)$ and $(j_q,\ldots,j_1)$. All cycle matrices lie in $L^\perp$. Moreover they are elementary matrices, in the following sense.

The *support* of a matrix $b\in\mathbb{R}^{k\times\ell}$ is the subset of cells where the entries of b are nonzero. An *elementary matrix of the subspace* $L^\perp$ is defined to be a nonzero matrix in $L^\perp$ whose support does not properly include the support of any other nonzero matrix in $L^\perp$. If b and $\widetilde{b}$ are elementary matrices in $L^\perp$ with the same support then they are scalar multiples of each other, $\widetilde{b}=\gamma b$ for some $\gamma\ne0$. Indeed, for some cell (i,j) in their common support consider the scalar $\gamma:=\widetilde{b}_{ij}/b_{ij}$. Then the matrix $d:=\widetilde{b}-\gamma b$ belongs to $L^\perp$ and has its support properly included in the support of b. This forces $d=0$, and $\widetilde{b}=\gamma b$. There are only finitely many distinct supports, and so the subspace $L^\perp$ has only finitely many elementary matrices, up to scalar multiples. Clearly the elementary matrices of the subspace $L^\perp$ are the cycle matrices, up to scalar multiples.

Now Theorem 22.6 in *Rockafellar* (1970 [203]) is applied. If L and R are disjoint, then there exists an elementary matrix in $L^\perp$, hence a cycle matrix $c(i_{(q)},j_{(q)})$, such that all $a\in R$ fulfill

$$\langle a,-c(i_{(q)},j_{(q)})\rangle>0, \qquad \text{that is,} \qquad \sum_{p\le q}a_{i_pj_p}>\sum_{p\le q}a_{i_{p+1}j_p}.$$

Since the left-hand sum is bounded from below by the right-hand sum, the terms $a_{i_pj_p}\in I_{i_pj_p}$ stay finite and the intervals $I_{i_pj_p}$ stem from definition (I-1). The intervals $I_{i_{p+1}j_p}$, being bounded from above, may stem from (I-1) or (I-2). All weights occurring in (I-1) and (I-2) are positive whence the current cycle is a v-cycle. Legitimate values in (3) are the limits of the intervals, $a_{i_pj_p}=\log\big(s(x_{i_pj_p})/v_{i_pj_p}\big)$ and $a_{i_{p+1}j_p}=\log\big(s(x_{i_{p+1}j_p})/v_{i_{p+1}j_p}\big)$. Exponentiation yields

$$\prod_{p\le q}\frac{s(x_{i_pj_p})}{v_{i_pj_p}}>\prod_{p\le q}\frac{s(x_{i_{p+1}j_p}+1)}{v_{i_{p+1}j_p}}.$$

This contradicts the validity of the critical inequalities. Hence L and R cannot be disjoint. □

The theorem provides a divisor-free check whether a feasible seat matrix is a double-proportional solution. Verification of the critical inequalities may be laborious, but at least there are only finitely many of them. The next section presents a fruitful application. The double-proportional seat apportionments are characterized by attaining the minimum of an objective function f_v. Since the objective function is minimized over the finite set of feasible seat matrices, $\mathbb{N}^{k\times\ell}(r,s)$, the existence of the minimum is guaranteed and so is the existence of double-proportional seat apportionments.

14.5. EXISTENCE OF DOUBLE-PROPORTIONAL SEAT APPORTIONMENTS

Chapter 10 appraises divisor methods through optimum properties relative to a variety of objective functions that are responsive to basic concepts of proportional representation systems. In a similar vein the present section characterizes double-proportional seat apportionments as optimum solutions of an appropriate objective function f_v. However, the merits of the objective function are less conceptual, but more technical.

Suppose a weight matrix $v \in [0;\infty)^{k\times\ell}$ is given. The objective function $f_v : \mathbb{N}^{k\times\ell} \to (0;\infty]$ maps a $k \times \ell$ integer matrix x to a positive value or to infinity. In case x preserves the zeros of v the definition is

$$f_v(x) := \prod_{i\le k, j\le \ell:\ v_{ij}>0} \left(\prod_{n\le x_{ij}\ :\ s(n)>0} \frac{s(n)}{v_{ij}} \right).$$

The product over n has range $1,\ldots,x_{ij}$ in case of perviousness, and $2,\ldots,x_{ij}$ in case of imperviousness. Empty products equal unity. If the products are nonempty then the terms $s(n)/v_{ij}$ are positive and finite, and so is the value of the objective function, $0 < f_v(x) < \infty$. In case x does not preserve the zeros of v we define $f_v(x) := \infty$. These seat matrices are far from being of any real interest.

The objective function is going to be minimized over the set of feasible seat matrices, $\mathbb{N}^{k\times\ell}(r,s)$, where rows are summing to the prespecified row marginals r and columns to the prespecified column marginals s. This is a finite set, whence the existence of the minimum is immediate. The minimum attains a finite value provided there exists at least some feasible seat matrix z that preserves the zeros of v, because of $\min_{y\in\mathbb{N}^{k\times\ell}(r,s)} f_v(y) \le f_v(z) < \infty$. The next result states that a feasible seat matrix attains the minimum if and only if it is a double-proportional seat apportionment.

Optimality Theorem. *Assume that for the given weight matrix v there exists some seat matrix $z \in \mathbb{N}^{k\times\ell}(r,s)$ that preserves the zeros of v. Then a seat matrix $x \in \mathbb{N}^{k\times\ell}(r,s)$ minimizes the objective function f_v if and only if x is a double-proportional seat apportionment,*

$$f_v(x) \le f_v(y) \quad \text{for all } y \in \mathbb{N}^{k\times\ell}(r,s) \qquad \Longleftrightarrow \qquad x \in A(r,s;v).$$

Proof. For the proof of the direct implication let x be a feasible seat matrix that attains the minimum. Then x preserves the zeros of v since $f_v(x) \le f_v(z) < \infty$. Let an arbitrary v-cycle via $i_{(q)} := (i_1, \ldots, i_q)$ and $j_{(q)} := (j_1, \ldots, j_q)$ be given. We need to verify the critical inequality

$$\prod_{p \le q} \frac{s(x_{i_p j_p})}{v_{i_p j_p}} \le \prod_{p \le q} \frac{s(x_{i_{p+1} j_p} + 1)}{v_{i_{p+1} j_p}}.$$

The inequality is trivial if the left-hand side is zero. This is the case for a pervious rounding rule if at least one of the seat numbers $x_{i_p j_p}$ is zero, and for an impervious rounding rule if some seat number $x_{i_p j_p}$ is unity. Otherwise all seat numbers $x_{i_p j_p}$ are large enough to allow a one-seat decrement. We introduce the seat matrix $y := x + c(i_{(q)}, j_{(q)})$ where $c(i_{(q)}, j_{(q)})$ is the cycle matrix of the given v-cycle. Minimality of x secures $f_v(x) \le f_v(y)$. The second products in the objective function f_v differ because of $y_{i_p j_p} = x_{i_p j_p} - 1$ and $y_{i_{p+1} j_p} = x_{i_{p+1} j_p} + 1$. Either the last term is too much and needs to be divided out, or it is missing and needs to be factored in,

$$f_v(x) \le f_v(y) = \frac{\prod_{p \le q} \dfrac{s(x_{i_{p+1} j_p} + 1)}{v_{i_{p+1} j_p}}}{\prod_{p \le q} \dfrac{s(x_{i_p j_p})}{v_{i_p j_p}}} \; f_v(x).$$

After cancellation of $f_v(x) > 0$ a rearrangement of terms verifies the critical inequality. Theorem 14.4 states that x is a double-proportional seat apportionment, $x \in A(r, s; v)$.

The converse implication claims that every double-proportional seat apportionment $x \in A(r, s; v)$ satisfies $f_v(x) \le f_v(y)$ for all seat matrices $y \in \mathbb{N}^{k \times \ell}(r, s)$. We start a bit offside and substitute for x a matrix z with no more than a scaling structure. No assumption is made on the row and column sums of z. Given some divisors E_i and F_j we choose numbers $z_{ij} \in [\![v_{ij}/(E_i F_j)]\!]$ to construct the matrix z. By construction z preserves the zeros of v. Let $y \in \mathbb{N}^{k \times \ell}(r, s)$ be a competing seat matrix preserving the zeros of v. In every cell (i, j) with $z_{ij} > y_{ij}$ or $z_{ij} < y_{ij}$ the weight is positive, $v_{ij} > 0$. In the definition of f_v the second products are estimated using signpost monotonicity, and the fundamental relation, $s(z_{ij})/v_{ij} \le E_i^{-1} F_j^{-1} \le s(z_{ij} + 1)/v_{ij}$, (omitting the condition $s(n) > 0$ for ease of reading)

$$z_{ij} > y_{ij} \;\Rightarrow\; \prod_{n \le z_{ij}} \frac{s(n)}{v_{ij}} = \left(\prod_{n \le y_{ij}} \frac{s(n)}{v_{ij}} \right) \prod_{m = y_{ij}+1}^{z_{ij}} \frac{s(m)}{v_{ij}} \le \left(\prod_{n \le y_{ij}} \frac{s(n)}{v_{ij}} \right) \left(E_i^{-1} F_j^{-1} \right)^{z_{ij} - y_{ij}},$$

$$z_{ij} < y_{ij} \;\Rightarrow\; \prod_{n \le z_{ij}} \frac{s(n)}{v_{ij}} = \left(\prod_{n \le y_{ij}} \frac{s(n)}{v_{ij}} \right) \prod_{m = z_{ij}+1}^{y_{ij}} \frac{v_{ij}}{s(m)} \le \left(\prod_{n \le y_{ij}} \frac{s(n)}{v_{ij}} \right) (E_i F_j)^{y_{ij} - z_{ij}}.$$

If $z_{ij} = y_{ij}$ or if $v_{ij} = 0$, then the exponent $y_{ij} - z_{ij}$ is zero and the divisor terms may be included for free. Upon application of the first product, the divisor terms combine into

$$\prod_{i \le k, j \le \ell :\, v_{ij} > 0} (E_i F_j)^{y_{ij} - z_{ij}} = \prod_{i \le k} \prod_{j \le \ell} (E_i F_j)^{y_{ij} - z_{ij}} = \left(\prod_{i \le k} E_i^{r_i - z_{i+}} \right) \left(\prod_{j \le \ell} F_j^{s_j - z_{+j}} \right).$$

The final comparison of $f_v(z)$ and $f_v(y)$ employs the inverse of the last expression,

$$\left(\prod_{i \le k} E_i^{z_{i+} - r_i} \right) \left(\prod_{j \le \ell} F_j^{z_{+j} - s_j} \right) f_v(z) \le f_v(y).$$

Only now is z reverted to x. Since row and column sums of x are fitted we get $f_v(x) \le f_v(y)$. □

The main message of the Optimality Theorem is that it answers the existence question in the affirmative. Since the set $\mathbb{N}^{k\times\ell}(r,s)$ is finite, it always contains a matrix x that minimizes f_v. Under the mild assumption that some feasible seat matrix preserves the zeros of v, every minimizer x is certified to belong to the double-proportional solution set,

$$x \in A(r,s;v),$$

which hence is nonempty. Practically, it suffices to produce a double-proportional seat apportionment x. It verifies both, the input assumption and the output conclusion. From a retrospective viewpoint no further action is needed.

From a prospective viewpoint it is of course reassuring to know that a solution exists before we set out to find it. Reassurance may be gained by applying a *greedy construction* to verify the assumption. The construction is illustrated with the Schaffhausen data in Table 14.3. Since the underlying divisor method with standard rounding is pervious, only the no input–no output law must be heeded. The seven empty cells force $z_{ij} = 0$. The construction consists of a prelude, and a finale. The prelude plunders the largest district and saturates as many weak parties as possible. The 28 seats of the first district easily furnish the 22 seats wanted by AL (5), ÖBS (4), CVP (3), JSVP (3), EDU (2), JFSH (2), SVP Sen. (1), EVP (1), and JUSO (1). The finale attends to the $60 - 22 = 38$ seats remaining. Luckily the three parties left do not feature any empty cells. This leaves a problem of six districts, with row marginals $(6, 12, 8, 7, 4, 1)$, and three parties, with column marginals $(16, 13, 9)$:

$$\begin{array}{r} \\ 6 \\ 12 \\ 8 \\ 7 \\ 4 \\ 1 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 16 & 13 & 9 \end{array} \\ \begin{pmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{pmatrix} \end{array} \quad \to \quad \begin{array}{r} \\ 6 \\ 12 \\ 8 \\ 7 \\ 4 \\ 1 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 16 & 13 & 9 \end{array} \\ \begin{pmatrix} 6 & & \\ 10 & 2 & \\ & 8 & \\ & 3 & 4 \\ & & 4 \\ & & 1 \end{pmatrix} \end{array}$$

The greedy finale places as many seats in the first cell as is permitted by its marginals, 6. This closes the first row. Next the first cell in the second row gets its largest permissible number of seats, 10. This finishes the first column. Continuing, the displayed matrix is obtained. Prelude and finale produce a matrix that preserves the zeros of v. The Optimality Theorem 14.5 tells us that the Schaffhausen problem is solvable.

The finale of the greedy construction works fine because there are no zero weights left to worry about. Therefore, if right from the beginning the weight matrix features no zeros then the greedy construction guarantees existence:

> *If the weight matrix v has no zeros then a double-proportional seat apportionment exists, $A(r,s;v) \neq \emptyset$.*

Or the other way round: The only troublemakers are the zeros in the weight matrix v. Empirical election data are trouble-free for the reason that their vote matrices contain few zeros or none. In no event are there enough zeros to cause upheaval. The Schaffhausen example is typical.

Yet it is a diverting intellectual exercise to conceive sparse matrices with many zeros in intriguing patterns. The greedy construction runs into problems when not ending with a finale where all zeros of v have been disposed of during the prelude (or iterated preludes). Fortunately the theory of graphs and networks provides a check that is completely versatile, under the heading of *flow inequalities*.

For pervious rounding rules the no input–no output law applies. For any row subset $I \subseteq \{1, \ldots, k\}$ we denote the partial sum of the associated row marginals by $r_I := \sum_{i \in I} r_i$. Similarly $s_J := \sum_{j \in J} s_j$ designates the partial sum of column marginals for $J \subseteq \{1, \ldots, \ell\}$. The flow scheme pretends that the seats enter "the system" via the row marginals, possibly access nonzero cells of the weight matrix v, and leave the system via the column marginals. For the part of the flow that enters through a row subset I, the accessible columns form the set $J_v(I) := \{j \leq \ell \mid v_{ij} > 0 \text{ for some } i \in I\}$, called *the set of columns connected in v with I*. The columns in the complement $J_v(I)'$ are inaccessible due to weights zero, $v_{ij} = 0$ for all $i \in I$ and $j \in J_v(I)'$. In this flow scheme the sum of the seats that enter through the rows in I cannot exceed the sum of the column marginals in $J_v(I)$. This relation is captured by the flow inequalities:

$$r_I \leq s_{J_v(I)} \qquad \text{for all } I \subseteq \{1, \ldots, k\}. \tag{1}$$

The flow inequalities are not only necessary, but also sufficient:

> *There exists a feasible seat matrix $x \in \mathbb{N}^{k \times \ell}(r, s)$ that obeys the no input–no output law for the weight matrix v if and only if r, s, and v satisfy the flow inequalities* (1).

For a proof and a review of the pertinent literature see *Oelbermann* (2013 [35]). The flow inequalities confirm the *no zero–no problem* conclusion of the greedy construction. If the weight matrix v has no zeros, then a nonempty row subset I is connected in v with all columns, and the inequality $r_I \leq s_+ = h$ turns into a triviality.

For impervious rounding rules the no output–no input law must be obeyed, too. The law justifies the minimum restrictions $a_{ij} = 1$ when $v_{ij} > 0$, and $a_{ij} = 0$ otherwise. Minimum restrictions may surface for other good reasons, such as the winner-take-one modification. Similarly, the no input–no output law translates into the maximum restrictions $b_{ij} = 0$ when $v_{ij} = 0$, and $b_{ij} = h$ otherwise. In order not to deal with a void problem we assume the restrictions to be compliant, $a_{ij} \leq b_{ij}$, $a_{i+} \leq r_i \leq b_{i+}$, and $a_{+j} \leq s_j \leq b_{+j}$, for all rows $i \leq k$ and for all columns $j \leq \ell$. The flow inequalities acquire a form symmetric in the marginals r and s, and in the restrictions a and b,

$$r_I + a_{I' \times J} \leq s_J + b_{I \times J'} \qquad \text{for all } I \subseteq \{1, \ldots, k\} \text{ and } J \subseteq \{1, \ldots, \ell\}. \tag{2}$$

It goes without saying that $a_{I' \times J} := \sum_{(i,j) \in I' \times J} a_{ij}$ denotes the sum of the entries of a in the block $I' \times J$. Again the flow inequalities are necessary and sufficient:

> *There exists a feasible seat matrix $x \in \mathbb{N}^{k \times \ell}(r, s)$ with $a_{ij} \leq x_{ij} \leq b_{ij}$ for all $i \leq k$ and $j \leq \ell$ if and only if r, s, a, and b satisfy the flow inequalities* (2).

This version of the flow inequalities is given by *Maier* (2009 [35]). Either form, (1) or (2), allows to check the flow inequalities by machine computation using the max-flow min-cut theorem of graph theory. We may conclude that there are various ways to check whether there exists a feasible seat matrix that preserves the zeros of v.

Finally we remark that the Optimality Theorem suggests a redraft of the Uniqueness Theorem 14.3. Suppose that x is a double-proportional seat apportionment in the solution set $A(r,s;v)$. If there exists another solution $y \neq x$, then the proof of the Uniqueness Theorem exhibits a v-cycle such that x fulfills the associated critical inequality with equality. Conversely, if there is a v-cycle via $i_{(q)}$ and $j_{(q)}$ such that x fulfills the critical inequality with equality, then the direct part of the proof of the Optimality Theorem shows that $y := x + c(i_{(q)}, j_{(q)})$ is another optimal solution besides x. Hence non-uniqueness relates to the critical inequalities in the following fashion:

> *The solution set $A(r,s;v)$ contains another solution besides x if and only if there exists a v-cycle such that x fulfills the critical inequality actually with equality.*

The negation of the two statements characterizes uniqueness:

> *A solution $x \in A(r,s;v)$ is unique if and only if x fulfills all critical inequalities with strict inequality.*

This characterization of double-proportional uniqueness is pleasing theoretically. It runs parallel to what Section 4.7 establishes for simple-proportional divisor methods. Nevertheless we prefer to word the Uniqueness Theorem 14.3 without a reference to the critical inequalities, so as to stay closer to practical needs.

14.6. A DUAL VIEW

In this section we briefly digress and present some statements that are not necessary, but enlightening. The Optimality Theorem breeds a dual approach aiding the understanding of the Alternating Scaling and Tie-and-Transfer algorithms. The Optimality Theorem restricts the domain of action to the set of feasible seat matrices $\mathbb{N}^{k\times\ell}(r,s)$. The task is to search this domain for a member x that has a scaling structure. The tool to solve this *primal problem* is the *primal objective function* f_v.

The dual view exchanges the roles of the protagonists. Now the domain of action is the set of matrices that have a scaling structure. The task is to identify a member x whose row sums match the row marginal r and whose column sums meet the column marginals s. Not surprisingly, the tool to solve the *dual problem* is another objective function, the *dual objective function* g_v.

The set of matrices that have a scaling structure is parameterized by the pertinent row and column divisors. Hence the domain of definition for the dual problem is the positive orthant $(0;\infty)^{k+\ell}$. Suppose we are given a vector of row divisors $E := (E_1,\dots,E_k) \in (0;\infty)^k$, and a vector of column divisors $F := (F_1,\dots,F_\ell) \in (0;\infty)^\ell$. We choose integers $z_{ij} \in [\![v_{ij}/(E_iF_j)]\!]$ to construct the matrix z. The dual objective function is defined through

$$g_v(E,F) := \left(\prod_{i\le k} E_i^{z_{i+}-r_i}\right)\left(\prod_{j\le\ell} F_j^{z_{+j}-s_j}\right) f_v(z).$$

The function g_v is well-defined. If z is tied, then some cell has $v_{ij}/(E_iF_j) = s(z_{ij})$. The factor $E_iF_js(z_{ij})/v_{ij} = 1$ emerges by absorbing the terms $E_i^{z_{i+}}$ and $F_j^{z_{+j}}$ into $f_v(z)$. When spelling out $g_v(E,F)$ to full length, the products over n may hence extend up to z_{ij}, or stop short at $z_{ij}-1$. Either way yields the same value $g_v(E,F)$.

The function g_v appears inconspicuously in the course of the Optimality Theorem. The second half of the proof shows that all feasible seat matrices $y \in \mathbb{N}^{k\times\ell}(r,s)$ satisfy

$$g_v(E,F) \le f_v(y).$$

Thus the primal problem and the dual problem are intimately related. The dual objective function provides lower bounds for the primal objective function, and the primal objective function provides upper bounds for the dual objective function. Moreover, equality holds when inserting a double-proportional seat apportionment $x \in A(r,s;v)$ into the primal objective function, and any associated divisor vectors C and D into the dual objective function. Hence we get a *strong duality theorem*,

$$g_v(C,D) = \max_{\substack{E\in(0;\infty)^k\\ F\in(0;\infty)^\ell}} g_v(E,F) = \min_{y\in\mathbb{N}^{k\times\ell}(r,s)} f_v(y) = f_v(x).$$

Accordingly, there are two broad types of algorithms, primal algorithms and dual algorithms. Primal algorithms maintain row and column marginals while approaching the desired scaling structure. Dual algorithms maintain the scaling structure while approaching the desired row and column marginals. Both algorithms to be discussed are of dual type, the Alternating Scaling algorithm and the Tie-and-Transfer algorithm.

14.7. ALTERNATING SCALING ALGORITHM

The algorithms are discussed assuming that the solution set is nonempty, $A(r,s;v) \neq \emptyset$ (Section 14.5). Again A denotes a fixed divisor method, with underlying rounding rule $[\![\cdot]\!]$, r and s are prespecified row and column marginals, and v is a given weight matrix.

The Alternating Scaling algorithm constructs seat matrices $x(t)$ and weight matrices $v(t)$. The seat matrices $x(t)$ enjoy the desired scaling structure by means of row and column divisors, but they alternate in the verification of the marginals. The weight matrices $v(t)$ record the interim quotients used. The starting values are $x(0) := 0$ and $v(0) := v$. The algorithm advances in steps of two, $t = 0, 2, 4$ etc.

- Odd steps $t+1$ compose the seat matrix $x(t+1)$ rowwise. If row i of the predecessor matrix $x(t)$ fits then it is copied into $x(t+1)$ and its row divisor is set to unity, $\rho_i(t+1) = 1$. Otherwise row i of $x(t+1)$ is an apportionment of r_i seats proportionate to row i of the previous weight matrix $v(t)$. With a selected divisor $\rho_i(t+1)$, the associated interim quotients are stored in the weight matrix $v(t+1)$:

$$\text{row } i \text{ of } x(t+1) \in A\big(r_i; \text{row } i \text{ of } v(t)\big), \tag{AS-1}$$

$$v_{ij}(t+1) := \frac{v_{ij}(t)}{\rho_i(t+1)}, \tag{AS-2}$$

  for all rows $i \le k$ and for all columns $j \le \ell$.

- Even steps $t+2$ compose the seat matrix $x(t+2)$ columnwise. If column j of $x(t+1)$ fits then it is copied into $x(t+2)$ and its column divisor is set to unity, $\sigma_j(t+2) = 1$. Otherwise column j of $x(t+2)$ is an apportionment of s_j seats proportionate to column j of the weight matrix $v(t+1)$. With a selected divisor $\sigma_j(t+2)$, the ensuing interim quotients are stored in the weight matrix $v(t+2)$:

$$\text{column } j \text{ of } x(t+2) \in A\big(s_j; \text{column } j \text{ of } v(t+1)\big), \tag{AS-3}$$

$$v_{ij}(t+2) := \frac{v_{ij}(t+1)}{\sigma_j(t+2)}, \tag{AS-4}$$

  for all columns $j \le \ell$ and for all rows $i \le k$.

The products of *incremental row divisors* $\rho_i(t+1)$ generate *cumulative row divisors*, and those of *incremental column divisors* $\sigma_j(t+2)$ give *cumulative column divisors*,

$$\rho_i(1)\rho_i(3)\cdots\rho_i(t+1) =: C_i(t+1) =: C_i(t+2),$$
$$\sigma_j(2)\sigma_j(4)\cdots\sigma_j(t+2) =: D_j(t+2) =: D_j(t+3).$$

The cumulative divisors equip all seat matrices $x(t)$ with a scaling structure relative to the original weight matrix v,

$$x_{ij}(t) \in \left[\!\!\left[\frac{v_{ij}}{C_i(t)D_j(t)}\right]\!\!\right] \qquad \text{for all } i \le k, j \le \ell, \text{ and for all } t \ge 1.$$

As soon as a seat matrix $x(T)$ fits both marginals, rows and columns, it is a double-proportional seat apportionment, $x(T) \in A(r, s; v)$. The terminal divisors $C_i(T)$ and $D_j(T)$ are embellished as explained in the subsection on uniqueness in Section 14.2.

Generally, for any matrix $z \in \mathbb{N}^{k\times\ell}$, failure to meet the marginals r and s is measured by the L_1-error function, or *flaw count*,

$$f(z) := \sum_{i\le k} |z_{i+} - r_i| + \sum_{j\le\ell} |z_{+j} - s_j|.$$

The flaw count says how many seats are malapportioned, from the point of view of the target marginals. Every row of z that is *overfitted*, $z_{i+} > r_i$, increases the flaw count by its surplus seats, $z_{i+} - r_i$, and every row that is *underfitted*, $z_{i+} < r_i$, increases it by its deficit seats $r_i - z_{i+}$, as do the columns. Flaw counts are always even because every malapportioned seat is counted twice, once as a surplus seat and once as a deficit seat. If the flaw count is zero, then both marginals are fitted. Hence the flaw counts of the Alternating Scaling sequence $x(1), x(2), \ldots$ ought to decrease monotonically until terminating with zero, $f\big(x(T)\big) = 0$.

Unfortunately ties may interfere and create problems. We demonstrate the obstacle by example. Suppose there are two districts with 10 seats each, and four parties with overall party-seats $6, 6, 4, 4$. Suppose further that in an odd step when rows are fitted part (AS-1) yields the tied seat vector $3+, 3-, 2+, 2-$ in both districts. Then the 2×4 seat matrix x with both rows equal to $3, 3, 2, 2$ meets the marginals with flaw count zero, $f(x) = 0$, and hits the target. However, in view of the ties the seat matrix y

with both rows equal to $4, 2, 3, 1$ is equally justified. It has a despicable flaw count, $f(y) = 8$. In fact, there are $\binom{4}{2}\binom{4}{2} = 36$ equally justified choices for the seat matrix (Section 4.7). It would seem erratic to allow the algorithm to propose a seat matrix $y(t)$ with many flaws, when the tied situation suggests to choose an equally justified matrix $x(t)$ with fewer flaws. The sound handling of ties calls for an amendment.

To this end a seat matrix $z \in \mathbb{N}^{k\times\ell}$ is said to be a *tied variant* of a seat matrix $x(t)$ in the Alternating Scaling algorithm, denoted by $z \equiv x(t)$, when step t is odd and all rows of both matrices solve the apportionment tasks (AS-1), or step t is even and all columns of both matrices solve the apportionment tasks (AS-3). The amendment to the Alternating Scaling algorithm decrees that the algorithm selects only such seat matrices that minimize the flaw count among their tied variants:

- Every seat matrix $x(t)$ that is produced in the course of the Alternating Scaling algorithm has a flaw count that is minimum among its tied variants,

$$f\big(x(t)\big) = \min_{z\equiv x(t)} f(z). \tag{AS-5}$$

The minimization problem that is defined by (AS-5) may be solved by direct enumeration since the multiplicities of simple-proportional apportionment problems are well explored (Section 4.7). Minimization can also be achieved by the construction of paths to transfer seats from overfitted rows to underfitted rows (or columns), in the spirit of the Tie-and-Transfer algorithm in the next section. Practically, empirical data are tie-free. If no competing variant z exists that is distinct from $x(t)$ then amendment (AS-5) is trivially fulfilled. No action is needed, the condition may be disbanded. Theoretically, amendment (AS-5) is reasonable. Its impact is visible in the contrived example above, and in the proof of the Monotonicity Lemma below.

Monotonicity Lemma. *The sequence of seat matrices $x(1), x(2), \ldots$ that is produced by the Alternating Scaling algorithm has non-increasing flaw counts, $f\big(x(t)\big) \geq f\big(x(t+1)\big)$ for all $t \geq 1$.*

Proof. For an even step t the seat matrix $x(t)$ has parties fitted. Its flaws come from districts, $f\big(x(t)\big) = \sum_{i\leq k} |x_{i+}(t) - r_i|$. The next step $t+1$ is odd and fits districts. Consider a fixed district $i \leq k$. We get $r_i = z_{i+}$ for all tied variants z of $x(t+1)$ including $x(t+1)$ itself. We claim that there exists a tied variant z such that in every district i the ℓ differences $x_{i1}(t) - z_{i1}$, $\ldots, x_{i\ell}(t) - z_{i\ell}$ have the same sign. In case district i is fitted its row in $x(t+1)$ is a copy of the row in $x(t)$ and all the differences are zero anyway.

In case district i is overfitted at time t, $x_{i+}(t) > r_i$, its row divisor $\rho_i(t+1)$ cannot possibly be strictly smaller than unity lest the district becomes even more overfitted than before. This leaves two possibilities, $\rho_i(t+1) > 1$ or $\rho_i(t+1) = 1$. If $\rho_i(t+1) > 1$ then all interim quotients decrease, $v_{ij}(t) > v_{ij}(t)/\rho_i(t+1) = v_{ij}(t+1)$. Monotonicity of rounding rules implies $x_{ij}(t) \geq z_{ij}$; all desired differences are nonnegative. If $\rho_i(t+1) = 1$ then row i must be tied. We select a tied variant z such that in row i as many tied seat numbers, that is $x_{ij}(t) \in x_{ij}(t)- = \{x_{ij}(t)-1, x_{ij}(t)\}$, are lowered to $z_{ij} = x_{ij}(t) - 1$ as are needed to reach $z_{i+} = r_i$. Such a choice of z has all differences $x_{ij}(t) - z_{ij}$ nonnegative.

In case district i is underfitted, $x_{i+}(t) < r_i$, similar arguments allow to adjust the selected tied variant z further so that all differences $x_{ij}(t) - z_{ij}$ are nonpositive. With tied variant z so selected, every district $i \le k$ has the property that ℓ the differences $x_{i1}(t) - z_{i1}, \ldots, x_{i\ell}(t) - z_{i\ell}$ have the same sign. Now the triangle inequality justifies the first inequality,

$$\begin{aligned} f\big(x(t)\big) &= \sum_{i\le k}\Big|\sum_{j\le\ell}\big(x_{ij}(t) - z_{ij}\big)\Big| = \sum_{i\le k}\sum_{j\le\ell}\Big|x_{ij}(t) - z_{ij}\Big| \\ &\ge \sum_{j\le\ell}\Big|\sum_{i\le k}\big(x_{ij}(t) - z_{ij}\big)\Big| = \sum_{j\le\ell}\Big|z_{+j} - s_j\Big| = f(z) \\ &\ge f\big(x(t+1)\big). \end{aligned}$$

The second inequality is a consequence of the amendment (AS-5). For an odd step $t+1$ monotonicity is established in the same way. ☐

The Schaffhausen example terminates in five steps, $T = 5$. On its way to the solution shown in Table 14.3 the approximating seat matrices pass through flaw counts 10, 4, 2, 2, 0. Simulation studies show that, quite generally, the Alternating Scaling algorithm produces seat matrices $x(t)$ whose flaw counts decrease rapidly in the beginning. It may need some time to work the few remaining flaws down to zero though. Nevertheless the Alternating Scaling algorithm works fine in all empirical examples.

The Alternating Scaling algorithm becomes possibly inefficient only in the presence of many ties. The following 3×3 example is taken from *Oelbermann* (2013 [48]):

$$\begin{array}{c} \\ 1 \\ 1 \\ 1 \\ \\ \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 1 & 1 \end{array} \\ \begin{pmatrix} 20 & 50 & 50 \\ 50 & 20 & 20 \\ 50 & 20 & 20 \end{pmatrix} \\ \begin{array}{ccc} 2.5 & 1 & 1 \end{array} \end{array} \begin{array}{c} \\ 100 \\ 40 \\ 40 \\ \\ \end{array}, \qquad \begin{pmatrix} 0.08 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1- & 0+ \\ 1- & 0+ & 0+ \\ 0+ & 0+ & 1- \end{pmatrix}.$$

The first matrix lists the vote weights. It is framed by unity marginals on the left and the top, and by the row and column divisors for the divisor method with standard rounding on the right and the bottom. The second matrix shows the interim quotients. The third matrix exhibits the final seat numbers, with trailing plus- and minus signs indicating the tie options. This particular tie pattern allows the completion of another three cycles, whence the solution set $A(r, s; v)$ contains four tied variants.

In this example it matters which incremental divisors $\rho_i(t+1)$ and $\sigma_j(t+2)$ are used. If the divisors are as described in Section 4.4 then the algorithm terminates in step $T = 33$, with flaw counts $f\big(x(1)\big) = \cdots = f\big(x(32)\big) = 2$ and $f\big(x(33)\big) = 0$. If for overfitted rows and columns the largest feasible divisors are applied and for underfitted rows and columns the smallest, then the algorithm reaches the solution faster (in this example, but not in others), namely in step $T = 12$. However, if always the upper limits of the divisor intervals are used then step 2 returns to step 1. The algorithm stalls and oscillates endlessly without ever reaching the solution. We conclude that the divisor choice makes a significant difference to the performance of the Alternating Scaling algorithm, in the presence of ties. The remedy against long or endless spells of constant flaw counts is the Tie-and-Transfer algorithm.

14.8. TIE-AND-TRANSFER ALGORITHM

The Tie-and-Transfer algorithm operates on single seats that appear to be malapportioned. It removes a seat from a bad location and transfers it to a better place. The algorithm may be set up quite generally to move all seats around. With a large house size and many seats this takes a while. It is recommendable to start with an initialization that promises to allocate many seats correctly.

The Tie-and-Transfer algorithm is initialized with a scaling of columns. We denote the resulting seat matrix by $x(0)$, the column divisors used by $\sigma_j(0)$, and the accompanying interim quotients by $v_{ij}(0) := v_{ij}/\sigma_j(0)$. Thus the initial seat matrix fulfills $x_{+j}(0) = s_j$ for all $j \leq \ell$.

The algorithm maintains fitted columns for all subsequent seat matrices $x(t)$, $t = 1, \ldots, T$. Hence the flaw count originates from rows only,

$$f\big(x(t)\big) = \sum_{i \leq k} |x_{i+}(t) - r_i| = \sum_{i \in I^+(t)} \big(x_{i+}(t) - r_i\big) + \sum_{i \in I^-(t)} \big(r_i - x_{i+}(t)\big),$$

where the sums range over the overfitted rows and over the underfitted rows,

$$I^+(t) := \{i \leq k \mid x_{i+}(t) > r_i\}, \qquad I^-(t) := \{i \leq k \mid x_{i+}(t) < r_i\}.$$

The remaining rows are those fitted, $I^=(t) := \{i \leq k \mid x_{i+}(t) = r_i\}$.

The aim is to transfer seats from overfitted rows $I^+(t)$ to underfitted rows $I^-(t)$. A *path via* $(i_1, \ldots, i_q)$ *and* $(j_1, \ldots, j_{q-1})$ is defined to be a list of cells of the type

$$(i_1, j_1),\ (i_2, j_1),\ (i_2, j_2),\ (i_3, j_2),\ \ldots,\ (i_{q-1}, j_{q-1}),\ (i_q, j_{q-1}).$$

The q rows i_p are taken to be distinct, as are the $q-1$ columns j_p. That is, a path is "a cycle without the last two cells". It starts in row i_1 and, after alternating through rows and columns, it finishes in row i_q.

The Tie-and-Transfer algorithm consists of a tie update to create a path for a possible seat transfer, and the seat transfer proper. Consider step $t = 0, 1, 2$ etc.

- The *tie update routine* calculates row divisors $\rho_i(t+1)$ and column divisors $\sigma_j(t+1)$ such that the new interim quotients

  $$v_{ij}(t+1) := \frac{v_{ij}(t)}{\rho_i(t+1)\sigma_j(t+1)}$$

  not only justify the old seat numbers, $x_{ij}(t) \in [\![v_{ij}(t+1)]\!]$, but also give rise to a path via $(i_1, \ldots, i_q)$ and $(j_1, \ldots, j_{q-1})$ that starts in an overfitted row $i_1 \in I^+(t)$, keeps alternating from a decrement option to an increment option,

  $$v_{i_p j_p}(t+1) = s\big(x_{i_p j_p}(t)\big) \quad \text{and} \quad v_{i_{p+1} j_p}(t+1) = s\big(x_{i_{p+1} j_p}(t) + 1\big)$$

  for all $p < q$, and finishes in an underfitted row $i_q \in I^-(t)$.

- The *seat transfer routine* updates the seat matrix from $x(t)$ to $x(t+1)$ by setting

  $$x_{i_p j_p}(t+1) := x_{i_p j_p}(t) - 1 \quad \text{and} \quad x_{i_{p+1} j_p}(t+1) := x_{i_{p+1} j_p}(t) + 1$$

  for all $p < q$, and $x_{ij}(t+1) := x_{ij}(t)$ otherwise.

The properties of the path constructed ensure that an overfitted row i_1 has to give up a seat and an underfitted row i_q gains one, and that the net effect on all other rows and on all columns is zero. Thus the flaw count is strictly decreasing, $f\big(x(t+1)\big) = f\big(x(t)\big) - 2$. The algorithm terminates when no flaws are left, $f\big(x(T)\big) = 0$, of course.

In the sequel the two routines are discussed in greater detail, and then illustrated by example. There is not much to say about the seat transfer routine. It simply executes the seat transfer that is indicated by the path found in the tie update routine.

The core of the Tie-and-Transfer algorithm is the tie update routine. It follows a philosophy somewhat complementary to what we have advertised so far. Up to now we have taken pains to keep the user-friendly divisor in the interior of the divisor interval if possible (Section 4.6). Hence the interim quotients v_j/D of simple-proportional divisor methods stay away from their framing signposts $s(x_j)$ and $s(x_j+1)$. In contrast, the tie update perturbs the divisors ever so little until some interim quotient does hit a signpost. Eventually sufficiently many quotients agree with signposts to complete a path. When a lower signpost is hit, $v_{ij}(t) = s(x_{ij})$, a decrement option emerges, $x_{ij}-$. When an upper signpost is hit, $v_{ij}(t) = s(x_{ij}+1)$, an increment option $x_{ij}+$ is met.

The tie update routine consists of two subroutines, the *path finding subroutine* and the *tie creating subroutine*. Usually they need to be iterated a few times. The language suited best to describe the path finding subroutine is provided by graph theory. There, the subroutine is called a breadth-first search. This search strategy is fast and finds paths in which rows and columns appear at most once, as required by the notion of a "path via $(i_1, \ldots, i_q)$ and $(j_1, \ldots, j_{q-1})$". We skip the details because the underlying idea is easily conveyed without turning to graph theory.

The path finding subroutine assembles two sets, the *set of labeled rows* I_L that can be reached by a path starting in some overfitted row in $I^+(t)$, and the *set of labeled columns* J_L that are visited by these paths. Initially all overfitted rows are labeled, $I_L = I^+(t)$, but no column, $J_L = \emptyset$. Then, for every labeled row $i \in I_L$, the unlabeled columns $j \in J'_L$ are scanned whether the cell (i,j) contains a decrement option $x_{ij}-$; if so, column j is adjoined to the set of labeled columns J_L. Next, for every labeled column $j \in J_L$, the unlabeled rows $i \in I_L$ are checked whether the cell (i,j) features an increment option $x_{ij}+$; if so, row i is adjoined to the set of labeled rows I_L. Thereafter the scanning process turns back to scan columns, then rows, etc. It pauses when the sets of labeled rows and columns stall. There are two possibilities. Either the set of labeled rows contains an underfitted row, $I_L \cap I^-(t) \neq \emptyset$. Then the labeling procedure identifies a path from an overfitted row to an underfitted row. The job is done.

Or there is no such path, $I_L \cap I^-(t) = \emptyset$. Then the tie creating subroutine is called for help. Upon its invocation the state of affairs may be depicted as follows:

$$\begin{array}{cc} & \begin{array}{cc} \qquad J_L \qquad\qquad & \qquad J'_L \qquad \end{array} \\ \begin{array}{c} I_L \\ \\ I'_L \end{array} & \left(\begin{array}{c|c} & s\big(x_{ij}(t)\big) < v_{ij}(t) \\ \hline \\ v_{ij}(t) < s\big(x_{ij}(t)+1\big) & \end{array}\right) \end{array}$$

The block $I_L \times J_L$ comprises all rows and columns labeled so far. The presence of labels indicates that here the path finding strategy works successfully. The block $I'_L \times J'_L$ conveys no particular information. The informative blocks are the two off-diagonal blocks. In the block $I_L \times J'_L$ all lower signposts are smaller than the interim quotients, $s(x_{ij}(t)) < v_{ij}(t)$. If this were not so, some column $j \in J'_L$ would contain a decrement option $x_{ij}(t)-$ and the column would have been labeled before the pause, $j \in J_L$; this is a contradiction. With a similar argument the nonzero interim quotients in the block $I'_L \times J_L$ are seen to be smaller than the upper signposts, $v_{ij}(t) < s(x_{ij}(t)+1)$.

An off-diagonal block provides no information if a pertinent index set is empty or if $s(x_{ij}(t))$ or $v_{ij}(t)$ vanish throughout. For a pervious rounding rule a signpost is zero if and only if its argument is zero, $s(x_{ij}(t)) = 0 \Leftrightarrow x_{ij}(t) = 0$; we omit the reasoning for impervious rounding rules. An interim quotient is zero if and only if the weight itself is zero, $v_{ij}(t) = 0 \Leftrightarrow v_{ij} = 0$. Hence a non-informative block has entries summing to zero, $x_{I_L \times J'_L}(t) = 0$ or $v_{I'_L \times J_L} = 0$. If both sums are zero then they clash with the flow inequalities, as is not hard to see. But the flow inequalities hold true since the problem is assumed to be solvable, $A(r, s; v) \neq \emptyset$ (Section 14.5). Therefore, at least one of the two off-diagonal blocks is informative, and at least one of the following decisive factors ρ_0 and σ_0 is positive and finite,

$$\rho_0 := \max\left\{0, \frac{s(x_{ij}(t))}{v_{ij}(t)} \;\middle|\; i \in I_L, j \in J'_L, s(x_{ij}(t)) > 0\right\} \in [0; 1),$$

$$\sigma_0 := \min\left\{\infty, \frac{s(x_{ij}(t)+1)}{v_{ij}(t)} \;\middle|\; i \in I'_L, j \in J_L, v_{ij}(t) > 0\right\} \in (1; \infty].$$

The scene is now set for creating new ties. The case $\rho_0 \geq 1/\sigma_0$ uses update factor ρ_0. The row divisors of labeled rows are divided by ρ_0 and the column divisors of labeled columns are multiplied by ρ_0. The other divisors stay put. The effect on a weight is block dependent. In blocks $I_L \times J_L$ and $I'_L \times J'_L$ the interim quotients stay as is because the update factor cancels out or equals unity. In block $I_L \times J'_L$ we get $s(x_{ij}(t)) \leq \rho_0 v_{ij}(t) \leq v_{ij}(t) \leq s(x_{ij}(t)+1)$, that is, $x_{ij}(t) \in [\![\rho_0 v_{ij}(t)]\!]$. The first inequality follows from the definition of ρ_0; there is at least one equality $s(x_{i_0 j_0}(t)) = \rho_0 v_{i_0 j_0}(t)$. Therefore, a new decrement option $x_{i_0 j_0}-$ surfaces in block $I_L \times J'_L$. In block $I'_L \times J_L$ we get $s(x_{ij}(t)) \leq v_{ij}(t) \leq (1/\rho_0) v_{ij}(t) \leq \sigma_0 v_{ij}(t) \leq s(x_{ij}(t)+1)$, that is, $x_{ij}(t) \in [\![v_{ij}(t)/\rho_0]\!]$. Altogether the new weights still justify the old seats.

The case $1/\sigma_0 \geq \rho_0$ uses update factor $1/\sigma_0$. The updated quotients create a new increment option $x_{i_0 j_0}+$ in block $I'_L \times J_L$. As soon as the tie creating subroutine is finished control is returned to the path finding subroutine. In this way unlabeled rows and columns are processed until a path is found allowing another flaw-reducing seat transfer. This completes the general discussion of the Tie-and-Transfer algorithm.

As an illustration we apply the Tie-and-Transfer algorithm to the 3×3 example from Section 14.7. In the following display the first matrix recalls the input weights and marginals, but now with the initializing column divisors at the bottom,

$$\begin{array}{cc} & \begin{array}{ccc} 1 & 1 & 1 \end{array} \\ \begin{array}{c} 1 \\ 1 \\ 1 \end{array} & \begin{pmatrix} 20 & 50 & 50 \\ 50 & 20 & 20 \\ 50 & 20 & 20 \end{pmatrix} \\ & \begin{array}{ccc} 100 & 70 & 70 \end{array} \end{array}, \qquad \begin{pmatrix} 1/5 & 5/7 & 5/7 \\ 1/2 & 2/7 & 2/7 \\ 1/2 & 2/7 & 2/7 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & 1 \\ 1- & 0 & 0 \\ 0+ & 0 & 0 \end{pmatrix}.$$

The second matrix shows the interim quotients, $v(0)$. The third matrix is the seat matrix, $x(0)$. Its columns are fitted, its rows are not. The first row is overfitted, the third is underfitted. They yield two flaws, $f(x(0)) = 2$. The second row is alright.

The tie update routine begins with its path finding subroutine. It labels the first row, $I_L = I^+(0) = \{1\}$, but no column, $J_L = \emptyset$. In the first row none of the unlabeled columns $j = 1, 2, 3$ contains a decrement option. The sets I_L and J_L stall right away. The path finding subroutine pauses and waits for new ties to be created.

The tie creating subroutine starts its first pass. The block $I_L \times J'_L = \{1\} \times \{1, 2, 3\}$ consists of the first row. The decisive factors are found to be $\rho_0 = 7/10$ and $\sigma_0 = \infty$. The divisor update in the first row yields the updated quotients and seats,

$$\begin{pmatrix} 7/50 & 1/2 & 1/2 \\ 1/2 & 2/7 & 2/7 \\ 1/2 & 2/7 & 2/7 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1- & 1- \\ 1- & 0 & 0 \\ 0+ & 0 & 0 \end{pmatrix}.$$

The two decrement options $1-$ in the first row are new. The path finding subroutine resumes its work, labels the second and third column, $J_L = \{2, 3\}$, and pauses again.

The tie creating subroutine takes over for its second pass. The decisive factors turn out to be $\rho_0 = 0$ and $\sigma_0 = 7/4$. The updated quotients and seats are

$$\begin{pmatrix} 2/25 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1- & 1- \\ 1- & 0+ & 0+ \\ 0+ & 0+ & 0+ \end{pmatrix}.$$

Four increment options $0+$ are new. Now the path finding subroutine labels the second and third row, and the first column. The labeling of the underfitted third row stops the tie update routine. A (shortest) path to be handed over to the seat transfer routine runs via rows $(1, 3)$ and column (3). The final seat matrix is the same as in Section 14.7.

Small examples like the preceding 3×3 problem are transparent and instructive. Calculations can be carried out by paper and pencil, and need not be tucked away in a machine. However, such examples cannot provide more than a didactic crutch. They are academic artifacts conveying only a faint feeling of the practical worth of double-proportional methods. These methods prove their true value when the vote counts are truly big, and when the political challenge is truly immense. As an instance of larger numbers we turn to a 27×8 problem drawn from the EP elections 2009. Although larger, the problem must also be classified as an academic toy example since the institutional provisions of the EP do not (yet) support double proportionality.

14.9. DOUBLE PROPORTIONALITY FOR EP ELECTIONS

Electoral reform has been a permanent item on the EP's agenda. The EP's electoral system must comply with two objectives, to represent the Union's citizens, and to reflect the diversity of Europe's peoples. *Andrew Duff* (2011), the EP's Committee on Constitutional Affairs' rapporteur on electoral procedure, proposes a supplement of an additional 25 seats to be elected from a single, pan-EU constituency. A pan-European perspective may also be realized by a double-proportional approach, as we illustrate with the data of the 2009 EP elections.

The sample evaluation of the 2009 EP elections is purely hypothetical and highly speculative. Since votes were cast for domestic parties and not for European parties, vote aggregation for the EP's 2009 Political Groups is not authoritative. Neither are the output seat numbers. Nevertheless we use the numbers to indicate the procedural steps that need to be carried out in a double proportional electoral system.

Table 14.4 starts with the first step, the prior allocation of the house size of 751 seats to the 27 Member States. Croatia is left out because it was not present in the EP elections 2009. The reference base is the QMV-population 2013, as in Table 12.4. The non-synchronous data sources do not impair the sample calculations. Of course, in a real election the population figures and the election results would be synchronized. The Cambridge Compromise solution (Section 12.7) allocates to each Member State five base seats, leaving $751 - 27 \times 5 = 616$ seats. These are allocated proportionately to the QMV-populations, using the divisor method with upward rounding. The last column "5+Upw•" shows the resulting seat numbers. In double-proportional terminology, the last column contains the row marginals that are carried over into Table 14.6.

Table 14.5 continues with the second step, the super-apportionment of the 751 seats to the EP's Political Groups. We include the non-attached Members of Parliament as the pseudo-group NA, since in the absence of proper parties we do not know any better (Section 1.11). The column "Votes" copies the respective column from Table 1.32 (where 736 seats are at issue). Vote aggregation by Political Groups is the weak point of the illustration because at present voters must cast their votes for domestic parties, and are not given a pan-European choice. Therefore, neither the union-wide vote counts nor the resulting seat numbers are authoritative. Nevertheless we use these numbers to proceed. The divisor method with standard rounding is employed because of its excellent properties as reviewed in Section 13.2. In double-proportional terminology, the last column "DivStd" contains the column marginals that reappear in the heading of Table 14.6.

Table 14.6 finishes with the third step, the double-proportional apportionment. The crux of the method are the row and column divisors (last column "State divisor", bottom line "Party divisor") that allow an immediate verification of the apportionment. Generally, a vote count is divided by the two corresponding divisors, and then standard rounding of the quotient yields the seat number. Specifically, the German EPP vote count (9 968 153) is divided by the divisor for Germany (251 000), and by the EPP divisor (0.9575). The resulting quotient 41.48 justifies 41 seats. Double proportionality guarantees that each Member State retains its seats allocated in the prior apportionment (Table 14.4), and that each party (in this example: each Political Group) exhausts its union-wide party seats from the super-apportionment (Table 14.5).

Some discordances are inevitable. In Ireland, the 254 669 S & D voters are represented by two deputies while the 256 123 GUE/NGL voters are awarded but one. Such local inhomogeneities are the price for the global consistency that double proportionality achieves. Since the union-wide S & D allotment of 191 seats is almost six times larger than the 33 GUE/NGL seats, the domestic S & D lists do relatively better than the domestic GUE/NGL lists. Determination of the divisors takes the Alternating Scaling algorithm eight row scalings and eight column scalings. That is, the final seat apportionment is reached after $8 \times 27 + 8 \times 8 = 280$ applications of the simple-proportional divisor method with standard rounding.

EP2009DoubleProp DistrictMagnitudes	Cambridge Compromise QMV-Pop.	5+Quot.	5+Upw•
DE Germany	81 843 700	5+99.6•	96
FR France	65 397 900	5+79.6	85
UK United Kingdom	62 989 600	5+76.7	82
IT Italy	60 820 800	5+74.03	80
ES Spain	46 196 300	5+56.2	62
PL Poland	38 538 400	5+46.9	52
RO Romania	21 355 800	5+25.99	31
NL Netherlands	16 730 300	5+20.4	26
EL Greece	11 290 900	5+13.7	19
BE Belgium	11 041 300	5+13.4	19
PT Portugal	10 541 800	5+12.8	18
CZ Czech Republic	10 505 400	5+12.8	18
HU Hungary	9 957 700	5+12.1	18
SE Sweden	9 482 900	5+11.5	17
AT Austria	8 443 000	5+10.3	16
BG Bulgaria	7 327 200	5+8.9	14
DK Denmark	5 580 500	5+6.8	12
SK Slovakia	5 404 300	5+6.6	12
FI Finland	5 401 300	5+6.6	12
IE Ireland	4 582 800	5+5.6	11
LT Lithuania	3 007 800	5+3.7	9
SI Slovenia	2 055 500	5+2.5	8
LV Latvia	2 041 800	5+2.5	8
EE Estonia	1 339 700	5+1.6	7
CY Cyprus	862 000	5+1.05	7
LU Luxembourg	524 900	5+0.6	6
MT Malta	416 100	5+0.5	6
Sum (Divisor)	503 679 700	(821 600)	751

TABLE 14.4 *Prior allocation of 751 seats to 27 Member States.* The Cambridge Compromise apportionment shown in the table is based on the 2013 QMV-populations. Croatia is omitted since it was not present in the 2009 elections; the omission explains the differences with Table 12.4.

EP2009DoubleProp Super-apportionment	Votes	Quotient	DivStd
EPP	52 324 413	272.2	272
S & D	36 776 044	191.3	191
ALDE	16 058 094	83.55	84
GREENS/EFA	12 070 029	62.8	63
ECR	7 610 712	39.6	40
EFD	7 153 584	37.2	37
GUE/NGL	6 280 876	32.7	33
NA	5 970 692	31.1	31
Sum (Divisor)	144 244 444	(192 200)	751

TABLE 14.5 *Super-apportionment of 751 seats among the eight 2009 Political Groups.* The union-wide votes of the Political Groups substitute for the non-existing pan-European ballots of the electorate. Normally the group of non-attached seats (NA) would require a separate handling.

The pan-European dimension lies in the fact that, in the first place, votes contribute to a party's success on the European level. Voters in a Member State may well be instrumental to secure a further seat for their European party, but the double-proportional sub-apportionment may allocate this seat elsewhere. For instance, some additional eighty thousand S & D voters in Italy would raise the Italian S & D votes from 7 997 770 to 8 080 000. In the super-apportionment (Table 14.5) the S & D interim

EP2009DP	EPP	S & D	ALDE	GRE/EFA	ECR	EFD	GUE/NGL	NA	State
Sub-apportm.	272	191	84	63	40	37	33	31	divisor
DE 96	9968153-41	5472566-23	2888084-12	3194509-13			1969239-7		251000
FR 85	4799908-30	2838160-18	1455841-9	2803759-16		257437-2	915634-5	891847-5	169000
UK 82		2460249-16	2080613-13	1767218-11	4131386-18	2498226-17	126184-1	1181845-6	162000
IT 80	12966334-39	7997770-24	2476695-7			3125418-10			350000
ES 62	6670377-28	6141784-25	808246-3	689062-3			294124-1	451866-2	253000
PL 52	3787998-33	908765-8			2017607-11				121000
RO 31	2074019-14	1504218-10	702974-5					419094-2	150000
NL 26	913233-6	548691-4	1034065-6	412537-2	155270-1	169882-1	323269-2	772746-4	163500
EL 19	1655722-7	1878982-8		178987-1		366637-1	669212-2		261800
BE 19	1288422-4	1259998-4	1485854-4	1319341-4	296699-1			647170-2	350000
PT 18	1427300-8	946475-6					761718-4		178000
CZ 18	180451-2	528132-6			741946-6		334577-4		87000
HU 18	1632309-11	503140-3			153660-1			427773-3	151000
SE 17	744851-5	773513-5	603799-3	575029-3			179182-1		172700
AT 16	858921-5	680041-4		284505-2				870299-5	170000
BG 14	832510-5	476618-3	569343-4					308052-2	160000
DK 12	297199-2	503439-3	474041-2	371603-2		357942-2	168555-1		200000
SK 12	324081-6	264722-4	74241-1			45960-1			61530
FI 12	455874-3	292051-2	418251-3	206439-2		162930-1	98690-1		136050
IE 11	532889-4	254669-2	525375-3	34585-0		99709-1	256123-1		158000
LT 9	147756-3	102347-2	88870-2		46293-1	67237-1			50000
SI 8	200429-4	85407-2	98450-2						50000
LV 8	245288-3	77447-1	59326-1	76436-1	58991-1		77447-1		80000
EE 7	48492-1	34508-1	164383-3	116830-2	8860-0	2206-0			60000
CY 7	109209-3	67794-2	12630-0				106922-2		40000
LU 6	62202-2	38641-2	37013-1	33387-1					26000
MT 6	100486-3	135917-3		5802-0					41000
Party div.	0.9575	0.9563	1	1.0114	1.45	0.934	1.085	1.13	

TABLE 14.6 *Hypothetical sub-apportionment, EP elections 2009.* The double-proportional divisor method with standard rounding is applied to the Political Groups' vote counts (Chapter 1). The divisors guarantee each Member State its seat allocation, and each Political Group its union-wide party-seats. Sample calculation: The German EPP votes (9 968 153) are divided by the divisor for Germany (251 000), and by the EPP divisor (0.9575). The resulting quotient 41.48 justifies 41 seats.

quotient is 191.3, while the divisor is 192 200. Hence a growth by eighty thousand votes promises to raise the interim quotient beyond the next signpost, 191.5, and to secure another seat for S & D. And so it happens. The additional S & D voters in Italy would give rise to 192 seats in the pan-European super-apportionment. But the seat does not go to the Italian S & D. This is not at all surprising, since roughly half of the large Italian divisor 350 000 must be reached before a seat is transferred from one Italian party to another. The additional seat would benefit Slovakia and equip the Slovakian S & D with a fifth seat. When vote counts vary the induced seat transfer paths are unpredictable and, by themselves, non-informative.

The merits of double proportionality lie in its behavior across the whole electoral region. A double-proportional apportionment (Table 14.6) ensures all Union citizens that domestic affiliations are respected as preordained by the prior seat allocation to all Member States (Table 14.4), and that the pan-European representation of the electorate's political division honors the motto of "One person, one vote" (Table 14.5).

Comments and References

The pertinent literature is reviewed, and further perspectives are mentioned.

Chapter 1. Exposing Methods: The 2009 European Parliament Elections

The 2009 EP election data are taken from the paper *Oelbermann / Palomares / Pukelsheim* (2010). We do not know of any official source including the data of all 27 Member States to the depth that the calculations for the translation of votes into seats could be repeated. Available data in the Internet need to be treated with caution whether they are still preliminary, still semi-official, or indeed official and final. Nor do the 27 Member States use a uniform terminology to describe their electoral systems. The European Union speaks with 23 official languages and writes with three alphabets, Latin, Greek, and Cyrillic. Some Member States exhibit their electoral provisions in their mother tongues only. Others provide unofficial translations into English, occasionally with an irritating lack of proficiency concerning the conversion of votes into seats. Much of the European diversity reflects in the electoral provisions even though it is one and the same political body that is being elected, the EP. Therefore, Chapter 1 not only reports the data, but also attempts to identify a common terminological ground for the many apportionment methods used.

Electoral systems are also a prime topic of the field of political science of course. Many political science books and monographs are devoted to a systematic study of proportional representation systems, such as *Behnke* (2007), *Gallagher / Mitchell* (2008), *Grofman / Lijphart* (1986), *Nohlen* (2009), *Rose* (1974), *Taagepera / Shugart* (1989), and many others. Naturally those authors aim at a unification of electoral terms as we do. But the goals are different, and so is the outcome. The political science viewpoint emphasizes the political consequences of electoral procedures. Central topics are political power and its distribution, and democratic government and its establishment. The theme of the present book is more restrictive, methodological analysis. The aim is to study apportionment methods in proportional representation systems, how they function and which quantitative consequences they entail.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*, DOI 10.1007/978-3-319-03856-8, © Springer International Publishing Switzerland 2014

Chapter 2. Imposing Constitutionality: The 2009 Bundestag Election

The definitive history of electoral experiences in imperial Germany 1871–1918 is told by historian *Margaret Anderson* (2000). Recounting innumerable election incidents from the time, the author develops her thesis that these decades prepared the ground for German democracy to grow and strike roots after 1918. However, as soon as the revolution of 1918 disposed of the leadership figure of the emperor, the newly introduced proportional representation system was contested not to produce proper leaders, voiced in articles with titles such as "Proportional Representation and the Selection of Leaders" (Verhältniswahl und Führerauslese) of renowned jurist *Walter Jellinek* (1926) and proportional representation activist *Richard Schmidt* (1929). Another activist, *Hans Gustav Erdmannsdörffer* (1932), dreamed of a leader such as *Thomas Mann*.

The discussions of electoral matters during the Weimar Republic bore fruit when, after the victory over the leader regime of Nazi Germany, the electoral system was to be newly designed. The body to draft a new Basic Law was the Parliamentary Council. It delegated the design of the electoral system to its Committee on Electoral Procedure, *Rosenbach* (1994) meticulously reviews its records. Unfortunately they do not fully explain how the current electoral system for the Bundestag came into being. The German two-votes system has acquired a high reputation internationally. *Shugart / Wattenberg* (2001) provocatively pose the question whether the system implements the "best of both worlds", of the world of proportional representation and of the world of the election of persons.

The amended two-votes system was used for the first time in 1957, and right away caused some irritation because of the occurrence of overhang seats. The 1957 incident was convincingly explained by a malapportionment of constituencies among states. Much of the ensuing discussion conveys the impression that, having been the cause on one occasion, malapportionment was the only explanation conceivable. Other than that, overhang seats were declared to be a "necessary consequence" (notwendige Folge) of a proportional representation system that is combined with the election of persons. No mentioning was made that already *Geyerhahn* (1902) had proposed options to avoid overhang seats, nor that the export of the German system was accompanied by amendments that did not produce any overhang seats, see the examples of Scotland and London in Section 12.4. In Germany the issue of overhang seats grew particularly virulent when the party system started to diversify and continued to do so after the 1990 re-unification.

The fascinating story of the electoral heritage over the last 2000 years is told by *Szpiro* (2010). Elections in medieval Augsburg are studied by *Rogge* (1994) who tells the story of the rejected election of carpenter *Marx Neumüller* (Section 2.5). The material on *Llull* and *Cusanus* is taken from *Hägele / Pukelsheim* (2001, 2008). An Internet edition of *Llull's* three electoral tracts is provided by *Drton / Hägele / Haneberg / Pukelsheim / Reif* (2004), see www.uni-augsburg.de/llull.

The role of electoral equality in the decisions of German constitutional courts is studied by *Pukelsheim* (2000a, 2000b, 2000c). Emphasis is on equality of the success values of the voters' votes, equality of the representative weights of the Members of Parliament, and equality of the ideal share of seats of the parties (Sections 2.7–2.9).

Chapter 3. From Reals to Integers: Rounding Functions, Rounding Rules

The task of rounding real numbers to integers is centuries-old. In a witty letter to the editor *Seal* (1950) conjectures that for every allegedly novel result on rounding effects one can find a prior reference that precedes it.

The floor brackets for downward rounding and the ceiling brackets for upward rounding, $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$, are attributed to *Iverson* (1962 [12]). The angle brackets for standard rounding, $\langle\cdot\rangle$, are employed by *Abramowitz / Stegun* (1970 [223]).

Commercial rounding is the accepted rounding procedure for business and commerce; the term is a literal translation from German "kaufmännische Rundung". We find it convenient to use the term "commercial rounding" for the real-valued rounding function; *Kopfermann* (1991 [108]) calls it "Standardrundung". We reserve the term "standard rounding" for the associated set-valued rounding rule. The even-number rounding function is recommended by *Wallis / Roberts* (1956 [175]) specifically for statistical purposes, and by *Bronstein / Semendjajew* (1991 [98]) quite generally.

The inevitableness of ties necessitates the handling of set-valued rounding procedures. For the sake of clarity we prefer to explicitly distinguish rounding functions from rounding rules. Their distinct characters are visualized through simple and double delimiting brackets. Double ceiling brackets for the rule of upward rounding proved useful already in *Pukelsheim* (1993 [307]).

The name "signpost" stems from *Balinski / Young* (1982 [62]). The attribute "stationary" for the signpost family in Section 3.11 is due to *Balinski / Rachev* (1993, 1997). *Kopfermann* (1991 [202]) and *Janson* (2012 [29]) speak of "linear" divisor methods. Other authors use the term "parametric" divisor methods, such as *Oyama* (1991), *Oyama / Ichimori* (1995), *Palomares / Ramírez* (2003). *Balinski / Ramírez* (2012) establish a property that characterizes stationary divisor methods within the class of all divisor method: If voters swing from party A to party B while all other vote counts remain the same, then at some point a seat is transferred from A to B without affecting other parties. While this book restricts attention to the two families of stationary and power-mean signposts, other such sequences can be constructed, see *Dorfleitner / Klein* (1999) or *Marshall / Olkin / Pukelsheim* (2002). *Janson* (2012 [257, 263]) describes the electoral systems of a large set of countries, including Estonia and Macau (Section 3.9).

Chapter 4. Divisor Methods of Apportionment: Divide and Round

Chapters 4 and 5 heavily rely on the seminal monograph *Balinski / Young* (1982). The definition of a divisor method in Section 4.4 emphasizes the flexible nature of the divisor. By contrast many authors derive the name "divisor" method from the signpost sequence $s(n)$, $n \in \mathbb{N}$. The reason is that the signposts serve as divisors in the calculation of the comparative figures $v_j/s(n)$, see Section 4.12. We remark that the dividing points $d(a)$ of *Balinski / Young* (1982 [99]) are related to the signpost sequence by way of a shift of argument, $d(a) = s(a+1)$.

The formulation of the jump-and-step algorithm in Section 4.6 is motivated by work of *Happacher / Pukelsheim* (1996). However, the algorithm was known and discussed ever since the inception of the divisor method with downward rounding, see *Gfeller* (1890), *Hagenbach-Bischoff* (1892 [20]).

Chapter 5. Quota Methods of Apportionment: Divide and Rank

The multitude of quotas used is bewildering (Section 5.8). They received their three-letter acronyms in the sequence how they came across in empirical data. For this reason the labeling is somewhat at odds with the quotas' order in Section 5.10. The Max-Min Inequality 5.5 appears in *Kopfermann* (1991 [196]). The family of quotas that we call shift-quotas, $Q(s) = v_+/(h + s)$, is studied also by *Kopfermann* (1991 [195]) and *Janson* (2012 [70]). Practically, only the shift $s = 0$ (Hare-quota) and the limiting value $s = 1$ (Droop-quota variant-4) are of interest. Theoretically, the embedding of the Hare- and Droop-quotas into a single one-parameter family is pleasing.

The original connotation of the term "quota" signifies a definitive number of voters to be represented by a Member of Parliament, see *Hare* (1857, 1860, 1865), *Droop* (1868, 1869, 1881), *Hart* (1992). The then novel proportional representation system was pitted against the established plurality vote. Since a plurality winner is determined by assessing whole numbers of votes, a "quota" originally had to be a whole number, too. Fractions were deemed unacceptable:

> ... *rejecting the fractional numbers of the dividend*, in *Hare* (1857 [17]).
> *Les fractions ne comptent pas*, in *Morin* (1862 [26]).
> *Brüche werden nicht gerechnet*, in *Getz* (1864 [50]).

Nowadays the term "quota" signifies any quantity, possibly fractional, provided only that it is computationally convenient. Section 10.12 has more to say on the ambiguous usage of the term "quota".

Chapter 6. Targeting the House Size: Discrepancy Distribution

The seat-total distribution in Section 6.5 and the discrepancy distribution in Theorem 6.7 are due to *Happacher* (1996, 2001). For three-party systems and the divisor method with downward rounding, *Hagenbach-Bischoff* (1905) carried out similar calculations. Other related results may be found in *Mosteller / Youtz / Zahn* (1967), *Diaconis / Freedman* (1979), *Kopfermann* (1991 [185]). The examples in Section 6.8 are taken from *Happacher* (2001) and *Happacher / Pukelsheim* (1998, 2000).

The statement that rounding residuals tend to a uniform distribution irrespective of the underlying weight distribution is part of the folklore of the analysis of rounding effects, see *Seal* (1950). Nevertheless we are unaware of a formal proof anywhere in the literature. The Invariance Principle 6.10 is proved for Riemann-integrable densities by *Heinrich / Pukelsheim / Schwingenschlögl* (2004, 2005). The present version for Lebesgue-integrable densities is due to *Janson* (2013).

Chapter 7. Favoring Some at the Expense of Others: Seat Biases

Pólya (1918, 1919a, 1919b, 1919c, 1919d) is first to derive seat bias formulas in three-party systems. *Pólya* emphasizes that the seat biases vanish uniformly only for the Hare-quota method with residual fit by greatest remainders and for the divisor method with standard rounding.

The interpretation of the systematic seat excess in Section 7.3 is given by *Janson* (2013). The Seat Bias Formula 7.7 with no threshold, $t = 0$, is due to *Schuster / Pukelsheim / Drton / Draper* (2003). *Schwingenschlögl / Pukelsheim* (2006) adjoin the threshold factor $1-\ell t$. A stringent proof for the whole formula is obtained by *Heinrich / Pukelsheim / Schwingenschlögl* (2005). The seat bias results for list alliances are taken from *Pukelsheim / Leutgäb* (2009). In these papers the notion of seat bias is interpreted to be the average seat excess assuming that all vote shares are equally likely.

Other bias concepts are investigated by *Balinski / Young* (1982 [118–128]). *Schwingenschlögl* (2008) shows that those results conform with the Seat Bias Formula 7.7. In essence, conformance is a consequence of the Invariance Principle 6.10. *Janson* (2013) establishes a complementary asymptotic analysis. It includes the results in Sections 7.2 and 7.3, and the handling of the shift-quota methods in Section 7.13. His approach is first to average over equally likely house sizes h in a finite range $\{0, \ldots, H\}$, and then to expand the range beyond limits, $H \to \infty$.

During the first half of the twentieth century the United States of America experienced a fierce and at times vicious dispute over seat biases. By evaluating sample data from previous apportionment instances, statistician *Walter Francis Willcox* from Cornell University amassed overwhelming evidence that the divisor method with standard rounding is unbiased and that the divisor method with geometric rounding is biased. See *Willcox* (1911, 1916a, 1916b, 1950, 1951, 1952), *Durand* (1947), *Leonhard* (1961). The empirical evidence meant nothing to mathematician *Edward Vermilye Huntington* from the Massachusetts Institute of Technology who continued to champion his favorite choice, the divisor method with geometric rounding. See *Huntington* (1921, 1928, 1931, 1941), *Balinski / Young* (1977), *Bartlow* (2006).

Chapter 8. Preferring Stronger Parties to Weaker Parties: Majorization

The exposition follows *Marshall / Olkin / Pukelsheim* (2002). Table 8.1 elaborates an example of *Balinski / Rachev* (1997). The majorization result for the shift-quota methods is inspired by *Lauwers / Puyenbroeck* (2006a, 2006b), our presentation follows *Janson* (2012 [112]).

A majorization comparison always encompasses all parties. It splits the whole party system into a group of stronger parties and into the complementary group of weaker parties. The relation of one method favoring small states relative to another method of *Balinski / Young* (1982 [118]), called the relation that one method gives-up to another method in *Balinski / Rachev* (1997), compares two parties, a stronger party versus a weaker party, but neglects the others. Either way monotone signpost ratios constitute the crucial criterion to be checked. We find the majorization relation more satisfying conceptually. It boasts an impressive range of applications, as recounted by *Marshall / Olkin* (1979) and *Marshall / Olkin / Arnold* (2011).

Chapter 9. Securing System Consistency: Coherence and Paradoxes

The term "coherence" is coined by *Balinski* (2003). *Balinski / Young* (1982 [141]) use "uniformity" instead, *Young* (1994 [171]) "consistency". Since "uniformity" and "consistency" are employed often with other meanings, we abide by "coherence". It nicely alludes to the interaction of the whole and its parts, and to the logical compatibility of abstract premises with concrete procedures. The extension by coherence from two parties to many is presented in *Young* (1994 [50, 190]).

Incompatibility of abstract expectations with concrete outcomes is what is colloquially captured by the term "paradox". Hence the chapter also reviews the paradoxes generally attributed to quota methods, although paradoxes that can be explained as easily as these hardly qualify to be paradoxes of any depth in the logical sense. Table 9.2 shows that the quota methods and the concordance requirement outrule general vote ratio monotonicity; the example is taken from *Young* (1994 [60]). Another example is given by *Brams / Straffin* (1982). Vote ratio monotonicity plays a pivotal role in the exposition of *Balinski / Young* (1982 [108, 117]).

Chapter 10. Appraising Electoral Equality: Goodness-of-Fit Criteria

The search for optimality properties of apportionment methods is as old as the proportional representation movement. The optimality results for the divisor methods with standard rounding (Section 10.2) and with geometric rounding (Section 10.3) are due to *Sainte-Laguë* (1910a, 1910b). The author's interests then moved on to other areas, see *Fouilhé* (1950), *Chastenet de Géry* (1994), *Gropp* (1998).

Pólya (1919d) proved the optimality theorem for the Hare-quota method with residual fit by greatest remainders (Section 10.4). *Hagenbach-Bischoff* (1882) promoted the divisor method with downward rounding on the ground of its optimality property in Section 10.5. Both authors published extensively on the subject, with an expressed intention to reach out to the public, see the works of *Pólya* and *Hagenbach-Bischoff* (1882, 1888, 1890, 1892, 1896, 1905, 1908). Yet the writings of these authors seem to have fallen into oblivion. *Eduard Hagenbach-Bischoff* even has to endure the misfortune to be quoted as a plural by *Peter Felix Müller* (1959 [79, 80]).

In 1905 *Hagenbach-Bischoff* succeeded to have the divisor method with downward rounding installed for the election of the Great Council of the Canton of Basel. In Switzerland the method is known as the Hagenbach-Bischoff procedure, or simply as the Swiss procedure. *Hagenbach-Bischoff* (1891 [3]) did not approve of having the system named after him, but pointed out that it is due to *Victor D'Hondt* (1878, 1882, 1883, 1885). *D'Hondt* spelled his name with a capital letter D, the University of Ghent files his nachlass under letter H.

The pairwise comparison approach that substantiates success-value stability of the divisor method with standard rounding (Section 10.8) is mentioned first by the German statistician *Ladislaus von Bortkiewicz* (1919), see also *Bortkiewicz* (1920). *Bortkiewicz* (1910) discusses paradoxes in the Prussian three-class franchise. The pairwise comparison approach was perfected by *Huntington* (1921). Apparently *Huntington* was inspired by the report of statistician *Joseph Adna Hill* (1911). *Huntington*'s research led him to favor the divisor method with geometric rounding (Section 10.10). In 1941 the method became the legal procedure for the apportionment of the House of Representatives seats among the States of the Union, see the narrative of *Balinski / Young* (1982 [58]). The method is also known as the method of equal proportions, or as the *Huntington* method, or as the *Hill / Huntington* method. *Hill* himself never claimed any particular involvement in the apportionment discussion, see *Hill* (1910, 1929, 1935), *Goldenweiser* (1939).

The result on ideal-share stability of the divisor method with standard rounding is due to *Balinski / Young* (1982 [132]). Further optimization approaches are developed by *Grilli di Cortona / Manzi / Pennisi / Ricca / Simeone* (1999) and many others. A criterion motivated by statistical efficiency considerations and leading to the divisor method with upward rounding may be found in *Pukelsheim* (1993 [304]).

Chapter 11. Tracing Peculiarities: Vote Thresholds and Majority Clauses

The results in Sections 11.1–11.8 are well-known and belong to the usual repertoire of apportionment methodology. The arrangement of the material is a mixture of *Balinski / Young* (1982), *Kopfermann* (1991), *Balinski / Rachev* (1997), *Palomares / Ramírez* (2003), *Janson* (2012). The North Rhine-Westphalian examples in Section 11.8 are from *Pukelsheim / Maier / Leutgäb* (2009). The discussion of the majority clauses in Sections 11.9–11.13 follows *Pukelsheim / Maier* (2006, 2008). The residual seat redirection clause was proposed long ago by *Gfeller* (1890) as mentioned in Section 11.10.

Chapter 12. Truncating Seat Ranges: Minimum-Maximum Restrictions

Balinski (2004 [192–193]) points out that it is ambiguous in which way quota methods incorporate side restrictions, and emphasizes that the resulting seat apportionments depend on the way chosen (Section 12.2).

The Cambridge Compromise (Section 12.7) emerged from a January 2011 workshop at the University of Cambridge, see the report *Grimmett / Laslier / Pukelsheim / Ramírez González / Rose / Słomczyński / Zachariasen / Życzkowski* (2011). The downgraded-population variant is proposed by *Grimmett / Oelbermann / Pukelsheim* (2012); *Arndt* (2008) advocates a similar approach (Section 12.8). Other ways to adjust to the population figures are discussed by *Martínez Aroza / Ramírez González* (2008), *Słomczyński / Życzkowski* (2012). As for the Jagiellonian Compromise see the references in Section 12.10. *Birkmeier* (2011) extends the approach by admitting abstentions.

Chapter 13. Proportionality and Personalization: BWG 2013

For a review of the German electoral system including an analysis from the quantitative viewpoint see *Behnke* (2007). Already *Brams / Fishburn* (1984) discuss proportional representation in variable-size legislatures. For a compendium of the full history and all details of the German Federal Election Law see *Schreiber* (2009). Constitutional jurist *Hans Meyer* (2010) reviews the long and agonizing struggle to deal with overhang seats in a satisfactory manner. English translations of major decisions of the German Federal Constitutional Court are compiled by *Kommers / Miller* (2012 [238–268]).

Our description of the 2013 amendment of the Federal Election Law follows *Pukelsheim / Rossi* (2013). Section 13.5 proposes to set the minimum number of seats for a party to be equal to the number of direct seats plus a ten percent overhead. This proposal is due to *Peifer / Lübbert / Oelbermann / Pukelsheim* (2012) who call it a *direct-seat oriented proportionality adjustment* (direktmandatsbedingte Proporzanpassung). *Bochsler* (2012) points out that the two-votes system with a variable-size legislature bears a potential of misuse which he exemplifies with data from the election of the Albanian parliament.

Chapter 14. Representing Districts and Parties: Double Proportionality

Pukelsheim / Schuhmacher (2004, 2011) review the introduction of the double-proportional divisor method with standard rounding in the Swiss Cantons of Zurich, Aargau, and Schaffhausen. In addition, double proportionality received an overwhelming support in a September 2013 referendum in the Canton of Zug, and in the Canton of Nidwalden. For a general exposition of the critical interplay between territorial and political representation see *Bochsler* (2010). The analysis of the critical inequalities in Section 14.4 and the Optimality Theorem 14.5 is due to *Gaffke / Pukelsheim* (2008a). The optimality theorem is based on an objective function that is akin to a criterion investigated by *Carnal* (1993).

Quite a few papers illustrate the hypothetical usefulness of double proportionality, see *Gassner* (2000) for Belgium, *Balinski / Ramírez* (1999) and *Balinski* (2002) for Mexico, *Zachariassen / Zachariassen* (2006) for the Farœ Islands, *Ramírez / Pukelsheim / Palomares / Martínez* (2006) for Spain, *Pennisi / Ricca / Simeone* (2006, 2009) for Italy. The application of double proportionality to the EP elections in Section 14.9 is one of the ten steps towards a unified electoral system proposed by *Oelbermann / Pukelsheim* (2011). Double-proportional divisor method are not the only way to give the idea of biproportionality a concrete form. Alternative approaches are proposed by *Cox / Ernst* (1982), *Serafini* (2012).

Double-proportional divisor methods of apportionment originate with the papers *Balinski / Demange* (1989a, 1989b), and are further elaborated by *Balinski / Rachev* (1993, 1997) and *Balinski* (2006). *Balinski / Pukelsheim* (2007) show that double proportional apportionment methods are coherent. A generic structure of appropriate algorithms is developed by *Gaffke / Pukelsheim* (2008b). Network flow algorithms are studied by *Rote / Zachariasen* (2007), see also the overview *Pukelsheim / Ricca / Scozarri / Serafini / Simeone* (2012).

The Tie-and-Transfer algorithm is succinctly described by *Zachariasen* (2006) and *Maier* (2009). *Maier* (2009) and *Maier / Zachariassen / Zachariasen* (2010) demonstrate the effectiveness of the Alternating Scaling algorithm and the Tie-and-Transfer algorithm by means of an extensive real-life benchmark study. These two algorithms, hybrid combinations of the two, and other algorithms are implemented in the free software BAZI that is available from the Internet site `www.uni-augsburg.de/bazi`.

Bibliography

Numbers in square brackets refer to pages on which the reference is quoted.

Abramowitz, M. / Stegun, I.A. (1970): *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York. [216]

Anderson, M.L. (2000): *Practicing Democracy. Elections and Political Culture in Imperial Germany.* Princeton University Press, Princeton NJ. [215]

Arndt, F. (2008): Ausrechnen statt aushandeln: Rationalitätsgewinne durch ein formalisiertes Modell für die Bestimmung der Zusammensetzung des Europäischen Parlaments. *Zeitschrift für ausländisches öffentliches Recht und Völkerrecht—Heidelberg Journal of International Law 68,* 247–279. [220]

Balinski, M.L. (2002): Une "dose" de proportionnelle: Le systéme électoral mexicain. *Pour la Science, Avril 2002,* 58–59. [221]

Balinski, M.L. (2003): Quelle équité? *Pour la Science, Septembre 2003,* 82–87. [219]

Balinski, M.L. (2004): *Le suffrage universel inachevé.* Belin, Paris. [103, 220]

Balinski, M.L. (2006): Apportionment: Uni- and bi-dimensional. Pages 43–53 in: *Mathematics and Democracy. Recent Advances in Voting Systems and Collective Choice.* Springer, Berlin. [221]

Balinski, M.L. / Demange, G. (1989a): An axiomatic approach to proportionality between matrices. *Mathematics of Operations Research 14,* 700–719. [193, 221]

Balinski, M.L. / Demange, G. (1989b): Algorithms for proportional matrices in reals and integers. *Mathematical Programming 45,* 193–210. [193, 221]

Balinski, M.L. / Pukelsheim, F. (2007): Die Mathematik der doppelten Gerechtigkeit. *Spektrum der Wissenschaft, April 2007,* 76–80. [221]

Balinski, M.L. / Rachev, S.T. (1993): Rounding proportions: Rules of rounding. *Numerical Functional Analysis and Optimization 14,* 475–501. [216, 221]

Balinski, M.L. / Rachev, S.T. (1997): Rounding proportions: Methods of rounding. *Mathematical Scientist 22,* 1–26. [216, 218, 220, 221]

Balinski, M.L. / Ramírez, V. (1999): Parametric methods of apportionment, rounding and production. *Mathematical Social Sciences 37,* 107–122. [221]

Balinski, M.L. / Ramírez, V. (2012): Parametric vs. divisor methods of apportionment. *Annals of Operations Research.* DOI 10.1007/s10479-012-1120-7. [216]

Balinski, M.L. / Young, H.P. (1977): On Huntington methods of apportionment. *SIAM Journal on Applied Mathematics 33,* 607–618. [218]

Balinski, M.L. / Young, H.P. (1982): *Fair Representation. Meeting the Ideal of One Man, One Vote.* Yale University Press, New Haven CT (*Second Edition* [with identical pagination], Brookings Institution Press, Washington DC, 2001). [52, 102, 105, 117, 118, 123, 124, 158, 216, 218, 219, 220]

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8,

Bartlow, T.L. (2006): Mathematics and politics. Edward V. Huntington and apportionment of the United States Congress. *Proceedings of the Canadian Society for History and Philosophy of Mathematics 19,* 29–54. [218]

Bauer, H. (1991): *Wahrscheinlichkeitstheorie. 4. Auflage.* De Gruyter, Berlin. [93]

Behnke, J. (2007): *Das Wahlsystem der Bundesrepublik Deutschland. Logik, Technik und Praxis der Verhältniswahl.* Nomos, Baden-Baden. [214, 220]

Birkmeier, O. (2011): *Machtindizes und Fairness-Kriterien in gewichteten Abstimmungssystemen mit Enthaltungen.* Logos Verlag, Berlin. [220]

Bochsler, D. (2010): *Territory and Electoral Rules in Post-Communist Democracies.* Palgrave Macmillan, Basingstoke. [221]

Bochsler, D. (2012): A quasi-proportional electoral system "only for honest men"? The hidden potential for manipulating mixed compensatory electoral systems. *International Political Science Review 33,* 401–420. [221]

von Bortkiewicz, L. (1910): Mathematisch-statistisches zur preussischen Wahlrechtsreform. *Jahrbücher für Nationalökonomie und Statistik, Dritte Folge 39,* 692–699. [219]

von Bortkiewicz, L. (1919): Ergebnisse verschiedener Verteilungssysteme bei der Verhältniswahl. *Annalen für soziale Politik und Gesetzgebung 6,* 592–613. [219]

von Bortkiewicz, L. (1920): Zur Arithmetik der Verhältniswahl. *Sitzungsberichte der Berliner Mathematischen Gesellschaft 18,* 17–24. [219]

Brams, S.J. / Fishburn, P.C. (1984): Proportional representation in variable-size legislatures. *Social Choice and Welfare 1,* 211–229. [221]

Brams, S.J. / Straffin, P.D. (1982): The apportionment problem. *Science 217,* 437–438. [219]

Bronstein, I.N. / Semendjajew, K.A. (1991): *Taschenbuch der Mathematik. 25. Auflage.* Teubner, Leipzig. [216]

Carnal, H. (1993): Mathématiques et politique. *Elemente der Mathematik 48,* 27–32. [221]

Chastenet de Géry, J. (1994): Sainte-Laguë, André (1882–1950), Professeur de Mathématiques générales en vue des applications. Pages 522–525 in: *Les Professeurs du Conservatoire national des arts et métiers, Dictionnaire biographique 1794–1955, Tome II.* Paris. [219]

Cox, L.H. / Ernst, L.R. (1982): Controlled rounding. *INFOR—Information Systems and Operational Research 20,* 423–432. [221]

Diaconis, P. / Freedman, D. (1979): On rounding percentages. *Journal of the American Statistical Association 74,* 359–364. [217]

Dorfleitner, G. / Klein, T. (1999): Rounding with multiplier methods: An efficient algorithm and applications in statistics. *Statistical Papers 40,* 143–157. [216]

Droop, H.R. (1868): *On Methods of Electing Representatives.* Macmillan, London. [217]

Droop, H.R. (1869): *On the Political and Social Effects of Different Methods of Electing Representatives.* Maxwell, London. [217]

Droop, H.R. (1881): On methods of electing representatives. *Journal of the Statistical Society of London 44,* 141–196 (Reprinted in: *Voting Matters 24* (2007) 7–46). [76, 217]

Drton, M. / Hägele, G. / Haneberg, D. / Pukelsheim, F. / Reif, W. (2004): Ramon Llulls Traktate zu Wahlverfahren: Ziele und Realisierung einer Internet-Edition. Pages 131–140 in: *Mediaevistik und Neue Medien,* Thorbecke, Stuttgart. [215]

Duff, A. (2011): The electoral reform of the European Parliament. Pages 32-51 in: *Report on a Proposal for a Modification of the Act Concerning the Election of the Members of the European Parliament by Direct Universal Suffrage of 20 September 1976 (2009/2134(INI)).* European Parliament, Committee on Constitutional Affairs, A7-0176/2011 (28.4.2011), RR\865675EN.doc, PE440.210v04-00. [210]

Durand, E.D. (1947): Tribute to Walter F. Willcox. *Journal of the American Statistical Association 42,* 5–10. [218]

Erdmannsdörffer, H.G. (1932): *Sehr geehrter Herr Thomas Mann! Brief vom 24. August 1932.* Thomas-Mann-Archiv, ETH Zürich (www.uni-augsburg.de/bazi/Erdmannsdoerffer1932.html). [215]

Feller, W. (1971): *An Introduction to Probability Theory and Its Applications, Volume I, Third Edition.* Wiley, New York. [88]

Felsenthal, D.S. / Machover, M. (1998): *The Measurement of Voting Power. Theory and Practice, Problems and Paradoxes.* Elgar, Cheltenham. [176]

Fouilhé, A. (1950): Nécrologie: André Sainte-Laguë. *Agrégation—Bulletin Officiel de la Société des Agrégés de l'Université, Nouvelle série 13,* 147–148. [219]

Gaffke, N. / Pukelsheim, F. (2008a): Divisor methods for proportional representation systems: An optimization approach to vector and matrix apportionment problems. *Mathematical Social Sciences 56,* 166–184. [221]

Gaffke, N. / Pukelsheim, F. (2008b): Vector and matrix apportionment problems and separable convex integer optimization. *Mathematical Methods of Operations Research 67,* 133–159. [221]

Gallagher, M. / Mitchell, T. (2008): *The Politics of Electoral Systems.* Oxford University Press, Oxford. [214]

Gassner, M.B. (2000): *Représentations parlementaires. Méthodes mathématiques biproportionnelles de répartition des sièges.* Édition de l'Université de Bruxelles, Bruxelles. [221]

Gauss, C.F. (1808): Theorematis Arithmetici, Demonstratio Nova. *Carl Friedrich Gauss Werke 2,* 3–8. [45]

Gauss, C.F. (1821): Anzeige, Theoria Combinationis observationum erroribus minimis obnoxiae, pars prior. *Carl Friedrich Gauss Werke 4,* 95–100. [129]

Getz, G. (1864): Zum Wahlgesetz. *Frankfurter Reform,* Zeitung vom 29. Januar 1864, 50–51. [217]

Geyerhahn, S. (1902): Das Problem der verhältnismässigen Vertretung. Ein Versuch seiner Lösung. *Wiener Staatswissenschaftliche Studien, Band 3, Heft 4.* Mohr, Tübingen. [31, 215]

Gfeller, J. (1890): Du transfer des suffrages et de la répartition des sièges complémentaires. *Représentation proportionnelle—Revue mensuelle 9,* 120–131. [83, 151, 216, 220]

Goldenweiser, E.A. (1939): Joseph A. Hill. *Journal of the American Statistical Association 34,* 407–411. [220]

Gray, M. / Stubb, A. (2001): The Treaty of Nice: Negotiating a poisoned chalice? *Journal of Common Market Studies 39,* 12–20. [166]

Grilli di Cortona, P. / Manzi, C. / Pennisi, A. / Ricca, F. / Simeone, B. (1999): *Evaluation and Optimization of Electoral Systems.* SIAM, Philadelphia PA. [220]

Grimmett, G.R. / Laslier, J.-F. / Pukelsheim, F. / Ramírez González, V. / Rose, R. / Słomczyński, W. / Zachariasen, M. / Życzkowski, K. (2011): *The Allocation between the EU Member States of the Seats in the European Parliament.* European Parliament, Directorate-General for Internal Policies, Policy Department C: Citizen's Rights and Constitutional Affairs, Note 23.03.2011 (PE 432.760). [168, 220]

Grimmett, G.R. / Oelbermann, K.-F. / Pukelsheim, F. (2012): A power-weighted variant of the EU27 Cambridge Compromise. *Mathematical Social Sciences 63,* 136–140. [220]

Grofman, B. / Lijphart, A. (1986): *Electoral Laws and Their Political Consequences.* Agathon Press, New York. [214]

Gropp, H. (1998): On configurations and the book of Sainte-Laguë. *Discrete Mathematics 191,* 91–99. [219]

Hägele, G. / Pukelsheim, F. (2001): Llull's writings on electoral systems. *Studia Lulliana 41,* 30–38. [215]

Hägele, G. / Pukelsheim, F. (2008): The electoral systems of Nicholas of Cusa in the Catholic Concordance and beyond. Pages 229–249 in: *The Church, the Councils, & Reform. The Legacy of the Fifteenth Century.* Catholic University of America Press, Washington. [215]

Hagenbach-Bischoff, E. (1882): *Die proportionale Vertretung bei den Basler Grossrathswahlen.* Schweighauserische Buchdruckerei, Basel. [219]

Hagenbach-Bischoff, E. (1888): *Die Frage der Einführung einer Proportionalvertretung statt des absoluten Mehres.* H. Georg's Verlag, Basel. [219]

Hagenbach-Bischoff, E. (1890): La solution du probléme de la répartition proportionnelle. *Représentation proportionnelle—Revue mensuelle 9,* 159–171. [219]

Hagenbach-Bischoff, E. (1891): Zur Wahlreform. *Züricher Post,* Separatdruck aus der Nr. 53 und 54 vom 4. und 5. März 1891. [219]

Hagenbach-Bischoff, E. (1892): *Die Anwendung der Proportionalvertretung bei den Schweizerischen Nationalratswahlen. Gutachten aus Auftrag des leitenden Comités des schweizerischen Wahlreformvereins.* Vereinsbuchdruckerei, Basel. [216, 219]

Hagenbach-Bischoff, E. (1896): Emploi de listes associées. Anwendung gekoppelter Listen. *Bulletin de la Société suisse pour la Représentation Proportionnelle—Bulletin des Schweizerischen Wahlreform-Vereins für Proportionale Volksvertretung 12 Nos. 10 & 11,* 78–85. [105, 219]

Hagenbach-Bischoff, E. (1905): *Die Verteilungsrechnung beim Basler Gesetz nach dem Grundsatz der Verhältniswahl.* Buchdruckerei zum Basler Berichthaus, Basel. [86, 217, 219]

Hagenbach-Bischoff, E. (1908): Verteilung der Sitze bei der Proportionalvertretung. *Zeitschrift für die gesamte Staatswissenschaft 64,* 344–346. [219]

Happacher, M. (1996): *Die Verteilung der Diskrepanz bei stationären Multiplikatorverfahren zur Rundung von Wahrscheinlichkeiten.* Logos Verlag, Berlin. [82, 88, 89, 217]

Happacher, M. (2001): The discrepancy distribution of stationary multiplier rules for rounding probabilities. *Metrika 53,* 171–181. [217]

Happacher, M. / Pukelsheim, F. (1996): Rounding probabilities: Unbiased multipliers. *Statistics & Decisions 14,* 373–382. [216]

Happacher, M. / Pukelsheim, F. (1998): And round the world away. Pages 93–108 in: *Proceedings of the Conference in Honor of Shayle R. Searle.* Cornell University, Ithaca NY. [217]

Happacher, M. / Pukelsheim, F. (2000): Rounding probabilities: Maximum probability and minimum complexity multipliers. *Journal of Statistical Planning and Inference 85,* 145-158. [217]

Hare, T. (1857): *The Machinery of Representation. Second Edition.* Maxwell, London. [217]

Hare, T. (1860): On the application of a new statistical method to the ascertainment of the votes of majorities in a more exhaustive manner. *Journal of the Statistical Society of London 23,* 337–356 (Reprinted in: *Voting matters 25* (2008) 1–12). [217]

Hare, T. (1865): *The Election of Representatives, Parliamentary and Municipal. A Treatise, Third Edition, with a Preface, Appendix, and other Additions.* Longman, Green, Longman, Roberts & Green, London. [217]

Hart, J. (1992): *Proportional Representation. Critics of the British Electoral System 1820–1945.* Oxford University Press, Oxford. [217]

Heinrich, L. / Pukelsheim, F. / Schwingenschlögl, U. (2004): Sainte-Laguë's chi-square divergence for the rounding of probabilities and its convergence to a stable law. *Statistics & Decisions 22,* 43–60. [217]

Heinrich, L. / Pukelsheim, F. / Schwingenschlögl, U. (2005): On stationary multiplier methods for the rounding of probabilities and the limiting law of the Sainte-Laguë divergence. *Statistics & Decisions 23,* 117–129. [217, 218]

Hill, J.A. (1910): Class report. Pages 81–82 in: *Class of 1885, Harvard College, Twenty-Fifth Anniversary.* Wheeler, Cambridge MA. [220]

Hill, J.A. (1911): Method of apportioning representatives. *Sixty-Second Congress, First Session, House Reports, Volume 1, Report 12,* 43–108. [219]

Hill, J.A. (1929): Apportionment methods described. *Congressional Digest, February 1929,* 42–44. [220]

Hill, J.A. (1935): Class report. Pages 121–123 in: *Class of 1885, Harvard College, Fiftieth Anniversary.* Wheeler, Cambridge MA. [220]

D'Hondt, V. (1878): *Question électorale. La représentation proportionnelle des partis.* Bruylant-Christophe, Bruxelles. [219]

D'Hondt, V. (1882): *Système pratique et raisonné de représentation proportionnelle.* Muquardt, Bruxelles. [219]

D'Hondt, V. (1883): Formule du minimum dans la représentation proportionnelle. *Représentation proportionnelle—Revue mensuelle 2,* 117–130. [219]

D'Hondt, V. (1885): *Exposé du systéme pratique de représentation proportionnelle adopté par le Comité de l'Association Réformiste Belge.* Vanderhaeghen, Gand. [219]

Huntington, E.V. (1921): A new method of apportionment of representatives. *Journal of the American Statistical Association 17,* 859–870 (Abstract "The mathematical theory of the apportionment of representatives" in: *Proceedings of the National Academy of Sciences of the United States of America 7,* 123–127). [218, 219]

Huntington, E.V. (1928): The apportionment of representatives in Congress. *Transactions of the American Mathematical Society 30,* 85–110. [138, 156, 218]

Huntington, E.V. (1931): Methods of apportionment in Congress. *American Political Science Review 25,* 961–965. [218]

Huntington, E.V. (1941): The rôle of mathematics in congressional apportionment. *Sociometry—A Journal of Inter-Personal Relations 4,* 278–282. (Reply by W.F. Willcox, 283-298. Rejoinder by E.V. Huntington, 299-301.) [218]

Iverson, K.E. (1966): *A Programming Language.* Wiley, New York. [216]

Janson, S. (2012): *Proportionella valmetoder.* Typescript (20.08.2012), Uppsala. [92, 216, 217, 218, 220]

Janson, S. (2013): Asymptotic bias of some election methods. *Annals of Operations Research,* DOI 10.1007/s10479-012-1141-2. [217, 218]

Janson, S. / Linusson, S. (2012): The probability of the Alabama paradox. *Journal of Applied Probability 49,* 773–794. [121]

Jellinek, W. (1926): Verhältniswahl und Führerauslese. *Archiv des öffentlichen Rechts 50,* 71–99. [215]

Kommers, D.P. / Miller, R.A. (2012): *The Constitutional Jurisprudence of the Federal Republic of Germany. Third Edition, Revised and Expanded.* Duke University Press, Durham. [220]

Kopfermann, K. (1991): *Mathematische Aspekte der Wahlverfahren. Mandatsverteilung bei Abstimmungen.* Bibliographisches Institut, Mannheim. [54, 216, 217, 220]

Lauwers, L.L. / Van Puyenbroeck, T. (2006a): The Hamilton apportionment method is between the Adams method and the Jefferson method. *Mathematics of Operations Research 31,* 390–397. [218]

Lauwers, L.L. / Van Puyenbroeck, T. (2006b): The Balinski-Young comparison of divisor methods is transitive. *Social Choice and Welfare 26,* 603–606. [218]

Leonard, W.R. (1961): Walter F. Willcox: Statist. *American Statistician 15,* 16–19. [218]

Maier, S. (2009): *Biproportional Apportionment Methods: Constraints, Algorithms, and Simulation.* Dr. Hut, München. [190, 201, 221]

Maier, S. / Zachariassen, P. / Zachariasen, M. (2010): Divisor-based biproportional apportionment in electoral systems: A real-life benchmark study. *Management Science 56,* 373–387. [221]

Marshall, A.W. / Olkin, I. (1979): *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York. [114, 218]

Marshall, A.W. / Olkin, I. / Arnold, B.C. (2011): *Inequalities: Theory of Majorization and Its Applications.* Springer, New York. [218]

Marshall, A.W. / Olkin, I. / Pukelsheim, F. (2002): A majorization comparison of apportionment methods in proportional representation. *Social Choice and Welfare 19,* 885–900. [216, 218]

Martínez-Aroza, J. / Ramírez-González, V. (2008): Several methods for degressively proportional allotments. A case study. *Mathematical and Computer Modelling 48,* 1439–1445. [220]

Meyer, H. (2010): *Die Zukunft des Bundestagswahlrechts. Zwischen Unverstand, obiter dicta, Interessenkalkül und Verfassungsverstoß.* Nomos, Baden-Baden. [220]

Morin, A. (1862): *De la représentation des minorités.* Fick, Genève. [217]

Mosteller, F. / Youtz, C. / Zahn, D. (1967): The distribution of sums of rounded percentages. *Demography 4,* 850–858. [217]

Müller, P.F. (1959): *Das Wahlsystem. Neue Wege der Grundlegung und Gestaltung. Mit einer ergänzenden Abhandlung zum mathematischen Teil von Professor Dr. Fridolin Bäbler.* Polygraphischer Verlag, Zürich. [219]

Niemeyer, H.F. / Niemeyer, A.C. (2008): Apportionment methods. *Mathematical Social Sciences 56,* 240–253. [151]

Nohlen, D. (2009): *Wahlrecht und Parteiensystem. Zur Theorie und Empirie der Wahlsysteme. 6. Auflage.* Budrich, Opladen. [150, 214]

Oelbermann, K.-F. (2013): *Biproportionale Divisormethoden und der Algorithmus der alternierenden Skalierung.* Logos Verlag, Berlin. [201, 206]

Oelbermann, K.-F. / Palomares, A. / Pukelsheim, F. (2010): The 2009 European Parliament elections: From votes to seats in 27 ways. *Evropská volební studia—European Electoral Studies 5,* 148–182 (Erratum, *ibidem 6* (2011) 85). [214]

Oelbermann, K.-F. / Pukelsheim, F. (2011): Future European Parliament elections: Ten steps towards uniform procedures. *Zeitschrift für Staats- und Europawissenschaften—Journal for comparative Government and European policy 9,* 9–28. [221]

Oyama, T. (1991): On a parametric divisor method for the apportionment problem. *Journal of the Operations Research Society of Japan 34,* 187–221. [216]

Oyama, T. / Ichimori T. (1995): On the unbiasedness of the parametric divisor method for the apportionment problem. *Journal of the Operations Research Society of Japan 38,* 301–321. [216]

Palomares, A. / Ramírez, V. (2003): Thresholds of the divisor methods. *Numerical Algorithms 34,* 405–415. [216, 220]

Peifer, R. / Lübbert, D. / Oelbermann, K.-F. / Pukelsheim, F. (2012): Direktmandatsorientierte Proporzanpassung: Eine mit der Personenwahl verbundene Verhältniswahl ohne negative Stimmgewichte. *Deutsches Verwaltungsblatt 127,* 725–730. [221]

Pennisi, A. / Ricca, F. / Simeone B. (2006): Bachi e buchi della legge elettorale italiana nell'allocazione biproporzionale di seggi. *Sociologia e ricerca sociale 79,* 55–76. [221]

Pennisi, A. / Ricca, F. / Simeone B. (2009): Una legge elettorale sistematicamente erronea. *Polena: Political and Electoral Navigations—Rivista Italiana di Analisi Elettorale 6,* 65–72. [221]

Pólya, G. (1918): Über die Verteilungssysteme der Proportionalwahl. *Zeitschrift für schweizerische Statistik und Volkswirtschaft 54,* 363–387. [102, 108, 217]

Pólya, G. (1919a): Über Sitzverteilung bei Proportionalwahlverfahren. *Schweizerisches Zentralblatt für Staats- und Gemeinde-Verwaltung 20,* 1–5. [217]

Pólya, G. (1919b): Über die Systeme der Sitzverteilung bei Proportionalwahl. *Wissen und Leben—Schweizerische Halbmonatsschrift 12,* 259–268, 307–312. [217]

Pólya, G. (1919c): Proportionalwahl und Wahrscheinlichkeitsrechnung. *Zeitschrift für die gesamte Staatswissenschaft 74,* 297–322. [217]

Pólya, G. (1919d): Sur la représentation proportionnelle en matière électorale. *Enseignement Mathématiques 20,* 355–379. [131, 217, 219]

Pukelsheim, F. (1993): *Optimal Design of Experiments.* Wiley, New York (Reprinted as *Classics In Applied Mathematics, 50,* SIAM, Philadelphia PA, 2006). [134, 216, 220]

Pukelsheim, F. (2000a): Mandatszuteilungen bei Verhältniswahlen: Erfolgswertgleichheit der Wählerstimmen. *Allgemeines Statistisches Archiv—Journal of the German Statistical Society 84,* 447–459. [215]

Pukelsheim, F. (2000b): Mandatszuteilungen bei Verhältniswahlen: Vertretungsgewichte der Mandate. *Kritische Vierteljahresschrift für Gesetzgebung und Rechtswissenschaft 83,* 76–103. [215]

Pukelsheim, F. (2000c): Mandatszuteilungen bei Verhältniswahlen: Idealansprüche der Parteien. *Zeitschrift für Politik—Organ der Hochschule für Politik München 47,* 239–273. [215]

Pukelsheim, F. / Leutgäb, P. (2009): List apparentements in local elections: A lottery. *Homo Oeconomicus 26,* 489–500). [218]

Pukelsheim, F. / Maier, S. (2006): A gentle majority clause for the apportionment of committee seats. Pages 177–188 in: *Mathematics and Democracy. Recent Advances in Voting Systems and Collective Choice.* Springer, Berlin. [220]

Pukelsheim, F. / Maier, S. (2008): Parlamentsvergrößerung als Problemlösung für Überhangmandate, Pattsituationen und Mehrheitsklauseln. *Zeitschrift für Parlamentsfragen 39,* 312–322. [220]

Pukelsheim, F. / Maier, S. / Leutgäb, P. (2009): Zur Vollmandat-Sperrklausel im Kommunalwahlgesetz. *Nordrhein-Westfälische Verwaltungsblätter 22,* 85–90. [220]

Pukelsheim, F. / Ricca, F. / Scozarri, A. / Serafini, P. / Simeone, B. (2012): Network flow methods for electoral systems. *Networks 59,* 73–88. [221]

Pukelsheim, F. / Rossi, M. (2013): Imperfektes Wahlrecht. *Zeitschrift für Gesetzgebung 28,* 209–226. [221]

Pukelsheim, F. / Schuhmacher, C. (2004): Das neue Zürcher Zuteilungsverfahren für Parlamentswahlen. *Aktuelle Juristische Praxis—Pratique Juridique Actuelle 13,* 505–522. [221]

Pukelsheim, F. / Schuhmacher, C. (2011): Doppelproporz bei Parlamentswahlen: Ein Rück- und Ausblick. *Aktuelle Juristische Praxis—Pratique Juridique Actuelle 20,* 1581–1599. [221]

Ramírez, V. / Palomares, A. / Martínez, M.L. (2006): Degressively proportional methods for the allotment of the European Parliament seats amongst the EU Member States. Pages 205–220 in: *Mathematics and Democracy. Recent Advances in Voting Systems and Collective Choice.* Springer, Berlin. [221]

Riker, W.H. (1982): *Liberalism Against Populism. A Confrontation Between the Theory of Democracy and the Theory of Social Choice.* Freemann, San Francisco. [104]

Rockafellar, R.T. (1970): *Convex Analysis.* Princeton University Press, Princeton NJ. [197]

Rogge, J. (1994): Ir freye wale zu haben. Möglichkeiten, Probleme und Grenzen der politischen Partizipation in Augsburg zur Zeit der Zunftverfassung (1368–1548). *Bürgertum—Beiträge zur europäischen Gesellschaftsgeschichte 7,* 243–277. [215]

Rose, R. (1974): *Electoral Behavior: A Comparative Handbook.* Free Press, New York. [214]

Rosenbach, H. (1994): *Der Parlamentarische Rat 1948-1949. Akten und Protokolle. Band 6: Ausschuß für Wahlrechtsfragen.* Boldt, Boppard am Rhein. [215]

Rote, G. / Zachariasen, M. (2007): Matrix scaling by network flow. Pages 848–854 in: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Proceedings in Applied Mathematics, Volume 125.* SIAM, Philadelphia PA. [221]

Sainte-Laguë, A. (1910a): La représentation proportionnelle et la méthode des moindres carrés. *Annales scientifiques de l'École normale supérieure, Troisiéme série 27,* 529–542 (Résumé: *Comptes rendus hebdomadaires des Séances de l'Académie des Sciences 151,* 377–378). [40, 219]

Sainte-Laguë, A. (1910b): La représentation proportionnelle et les mathématiques. *Revue générale des Sciences pures et appliquées 21,* 846–852. [219]

Schmidt, R. (1929): *Verhältniswahl und Führerauslese.* Schulze, Gleiwitz. [215]

Schreiber, W. (2009): *Kommentar zum Bundeswahlgesetz. 8. Auflage.* Heymann, Köln. [220]

Schuster, K. / Pukelsheim, F. / Drton, M. / Draper, N.R. (2003): Seat biases of apportionment methods for proportional representation. *Electoral Studies 22,* 651–676. [104, 218]

Schwingenschlögl, U. (2008): Asymptotic equivalence of seat bias models. *Statistical Papers 49,* 191–200. [218]

Schwingenschlögl, U. / Pukelsheim, F. (2006): Seat biases in proportional representation systems with thresholds. *Social Choice and Welfare 27,* 189–193. [218]

Seal, H.L. (1950): Spot the prior reference. *Journal of the Institute of Actuaries Students' Society 10,* 255–258. [216, 217]

Serafini, P. (2012): Allocation of the EU Parliament seats via integer linear programming and revised quotas. *Mathematical Social Sciences 63,* 107–113. [221]

Shugart, M.S. / Wattenberg, M.P. (2001): *Mixed-Member Electoral Systems: The Best of Both Worlds?* Oxford University Press, Oxford. [215]

Słomczyński, W. / Życzkowski, K. (2006): Penrose voting system and optimal quota. *Acta Physica Polonica B 37,* 3133–3143. [175]

Słomczyński, W. / Życzkowski, K. (2012): Mathematical aspects of degressive proportionality. *Mathematical Social Sciences 63,* 94–101. [220]

Szpiro, G.G. (2010): *Numbers Rule. The Vexing Mathematics of Democracy, from Plato to the Present.* Princeton University Press, Princeton NJ. [215]

Taagepera, R. / Shugart, M.S. (1989): *Seats and Votes. The Effects and Determinants of Electoral Systems.* Yale University Press, New Haven CT. [214]

Wallis, W.A. / Roberts, H.V. (1956): *Statistics. A new approach.* Methuen, London. [216]

Willcox, W.F. (1911): Apportionment of representatives. *Sixty-Second Congress, First Session, House Reports, Volume 1, Report 12,* 7–42. [218]

Willcox, W.F. (1916a): The apportionment of representatives. *American Economic Review, Supplement 6,* 3–16. [218]

Willcox, W.F. (1916b): The recent apportionment in New York State. *Cornell Law Quarterly 2,* 1–19. [218]

Willcox, W.F. (1950): The apportionment problem and the size of the House: A return to Webster. *Cornell Law Quartely 35,* 367–389. [218]

Willcox, W.F. (1951): Methods of apportioning seats among the states in the American House of Representations. *International Statistical Institute 1947 Conference Proceedings 3,* 858–867. [218]

Willcox, W.F. (1952): Last words on the apportionment problem. *Law and Contemporary Problems 17,* 290–301. [218]

Young, H.P. (1994): *Equity. In Theory and Practice.* Princeton University Press, Princeton NJ. [219]

Zachariasen, M. (2006): *Algorithmic Aspects of Divisor-Based Biproportional Rounding.* Technical Report no. 06/05, Department of Computer Science, University of Copenhagen. [221]

Zachariassen, P. / Zachariassen, M. (2006): A comparison of electoral formulae for the Farœse Parliament. Pages 235–251 in: *Mathematics and Democracy. Recent Advances in Voting Systems and Collective Choice.* Springer, Berlin. [221]

Index

A number in boldface type refers to the page where the item is defined or introduced. A range of three or more pages may span a page not showing the item.

F. Pukelsheim, *Proportional Representation: Apportionment Methods and Their Applications*,
DOI 10.1007/978-3-319-03856-8, © Springer International Publishing Switzerland 2014